

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 530 057**

51 Int. Cl.:

C12P 19/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **23.01.2006 E 06719427 (4)**

97 Fecha y número de publicación de la concesión europea: **26.11.2014 EP 1851335**

54 Título: **Composiciones y procedimientos de uso de mezclas estandarizadas**

30 Prioridad:

21.01.2005 US 646157 P

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

26.02.2015

73 Titular/es:

**GENE EXPRESS, INC. (50.0%)
975 Research Drive
Toledo, OH 43614, US y
MEDICAL UNIVERSITY OF OHIO AT TOLEDO
(50.0%)**

72 Inventor/es:

**WILLEY, JAMES C.;
AUSTERMILLER, BRAD;
CRAWFORD, ERIN L.;
KNIGHT, CHARLES;
OSBORN, TERRY y
ZAHORCHAK, ROBERT**

74 Agente/Representante:

ARIAS SANZ, Juan

ES 2 530 057 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

Composiciones y procedimientos de uso de mezclas estandarizadas

5 Ciertas realizaciones de la presente invención se hicieron bajo las becas de investigación números ES02679 y ES01247 de los Institutos Nacionales de la Salud, la beca nº RR00044 de la División de Recursos de Investigación, el contrato 91-2 de los Institutos de la Salud y el contrato CH61 de la Organización Internacional del Plomo y Cinc, que pueden tener ciertos derechos sobre ellas. Ciertas realizaciones de la invención se hicieron bajo las becas de investigación nº NIH CA85147, CA 95806 y CA 103594, que pueden tener ciertos derechos sobre ellas. Ciertas realizaciones de la presente invención se hicieron bajo la beca de investigación número EO 1640 de los Institutos Nacionales de la Salud, que pueden tener ciertos derechos sobre ellas.

10 ANTECEDENTES DE LA INVENCION

15 Con la secuenciación del genoma humano, aparece la esperanza de acelerar el desarrollo de fármacos y descubrir mejores ensayos de diagnóstico. Esta esperanza ha creado la necesidad de desarrollar procedimientos mejorados para la medida de la expresión multigénica. Son particularmente necesarios procedimientos susceptibles de un control de calidad apropiado, por ejemplo, para satisfacer las directrices regulatorias. La presente invención se refiere a composiciones y procedimientos dirigidos a abordar estas esperanzas y necesidades.

SUMARIO DE LA INVENCION

20 Es un primer aspecto descrito en la presente memoria un procedimiento que comprende proporcionar una primera muestra que comprende un primer ácido nucleico; amplificar dicho primer ácido nucleico y obtener una relación en la que dicha relación puede dar cuenta de menos de aproximadamente 1.000 moléculas de dicho primer ácido nucleico en dicha primera muestra. Se describe además que dicha relación puede dar cuenta de menos de aproximadamente 100 moléculas, menos de aproximadamente 10 moléculas o menos de aproximadamente 1 molécula de dicho primer ácido nucleico en dicha primera muestra. Se describe también en la presente memoria que dicha relación compara una primera relación de producto amplificado de dicho primer ácido nucleico a producto coamplificado de un molde competitivo de dicho primer ácido nucleico con una segunda relación de producto amplificado de un segundo ácido nucleico en dicha primera muestra a producto coamplificado de un molde competitivo de dicho segundo ácido nucleico. Típicamente, dicho primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico se coamplifican en un primer recipiente y dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico se coamplifican en un segundo recipiente. El molde competitivo del primer o segundo ácido nucleico puede comprender una secuencia referenciada en la Tabla 4. Se describe además que el segundo ácido nucleico sirve como primer ácido nucleico de referencia, por ejemplo, como control de carga. El primer ácido nucleico de referencia puede corresponder a al menos un gen seleccionado de GADP, ACTB y β -actina. La relación puede comparar además el producto amplificado de una serie de otro(s) ácido(s) nucleico(s) con el producto coamplificado de molde(s) competitivo(s) de dicha serie de otro(s) ácido(s) nucleico(s). Al menos uno de dichos otros ácidos nucleicos puede servir como segundo ácido nucleico de referencia. El segundo ácido nucleico de referencia puede corresponder a al menos un gen seleccionado de GADP, ACTB y β -actina. Se describe además que la relación comprende el uso de electroforesis capilar de microfluidos, matriz de oligonucleótidos, espectrometría de masas o cromatografía. Se describe también en la presente memoria que la relación no implica tomar medidas a tiempo real ni la generación de una curva patrón. La relación puede controlar las fuentes de variación seleccionadas de carga de ADNc, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y eficacia de amplificación intramuestra. La relación es capaz de detectar una diferencia de menos de aproximadamente 2 veces o una diferencia de menos de aproximadamente 1 vez. La relación es capaz de detectar una diferencia de menos de aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 %, una diferencia de menos de aproximadamente un 30 % o una diferencia de menos de aproximadamente un 20 %. Se describe además en la presente memoria que la relación es capaz de detectar una diferencia de menos de aproximadamente 2 veces o una diferencia de menos de aproximadamente 1 vez en aproximadamente 100 moléculas o menos o en aproximadamente 10 moléculas o menos de dicho primer ácido nucleico en dicha primera muestra. La diferencia detectada puede ser una diferencia de menos de aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 %, una diferencia de menos de aproximadamente un 30 % o una diferencia de menos de aproximadamente un 20 %. Se describe también en la presente memoria que la relación proporciona un coeficiente de variación de menos de aproximadamente un 25 %, menos de aproximadamente un 20 %, menos de aproximadamente un 15 %, menos de aproximadamente un 10 % o menos de aproximadamente un 5 % entre dicha primera muestra y una segunda muestra de dicho primer ácido nucleico. Se describe además en la presente memoria que la primera y dicha segunda muestras se amplifican en momentos diferentes, la primera y dicha segunda muestras se amplifican en laboratorios diferentes o la primera y dicha segunda muestras se proporcionan de sujetos diferentes. El primer ácido nucleico puede comprender una secuencia referenciada en la Tabla 1 o 2. Los procedimientos descritos en la presente memoria pueden reducir o eliminar los falsos negativos, preferiblemente los falsos positivos se reducen a un número estadísticamente no significativo. El ácido nucleico puede ser una molécula de ARN o una molécula de ADN. Típicamente, la relación es sustancialmente constante más allá de la fase exponencial de dicha amplificación de dicho primer ácido nucleico.

Se describe también en la presente memoria un procedimiento de valoración de un primer ácido nucleico proporcionado en una primera muestra, que comprende coamplificar dicho primer ácido nucleico, una serie de otro(s) ácido(s) nucleico(s), un molde competitivo de dicho primer ácido nucleico y molde(s) competitivo(s) de dicho(s) otro(s) ácido(s) nucleico(s), en el que dichos moldes competitivos están a concentraciones conocidas uno respecto a otro, produciendo un primer producto amplificado de los mismos; diluir dicho primer producto amplificado y coamplificar además dicho primer producto amplificado diluido de dicho primer ácido nucleico y de dicho molde competitivo de dicho primer ácido nucleico, produciendo un segundo producto amplificado del mismo. Típicamente, el número es de al menos aproximadamente 1 otro ácido nucleico, al menos aproximadamente 100 otros ácidos nucleicos, o el número es de al menos aproximadamente 1.000 otros ácidos nucleicos. Típicamente, la dilución produce una dilución de al menos aproximadamente 100 veces, al menos aproximadamente 1.000 veces o al menos aproximadamente 10.000 veces. Preferiblemente, el procedimiento da cuenta de menos de aproximadamente 1.000 moléculas de dicho primer ácido nucleico en dicha muestra, de menos de aproximadamente 100 moléculas de dicho primer ácido nucleico en dicha primera muestra, de menos de aproximadamente 10 moléculas de dicho primer ácido nucleico en dicha primera muestra o de aproximadamente 1 molécula de dicho primer ácido nucleico en dicha primera muestra. Preferiblemente, al menos uno de dichos moldes competitivos comprende una secuencia referenciada en la Tabla 4. El primer ácido nucleico puede comprender una secuencia referenciada en la Tabla 1 o 2. Uno de los otros ácidos nucleicos puede servir como primer ácido nucleico de referencia, tal como control de carga. El primer ácido nucleico de referencia puede corresponder a al menos un gen seleccionado de GADP, ACTB y β -actina. Se describe además en la presente memoria que el procedimiento de valoración comprende obtener una primera relación, comparando dicha primera relación dicho segundo producto amplificado de dicho primer ácido nucleico con dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico; obtener una segunda relación, comparando dicha segunda relación dicho primer producto amplificado de dicho primer ácido nucleico de referencia con dicho primer producto amplificado de dicho molde competitivo de dicho primer ácido nucleico de referencia y comparar dichas primera y segunda relaciones. Se describe también en la presente memoria que otro de dichos otros ácidos nucleicos sirve como segundo ácido nucleico de referencia. El segundo ácido nucleico de referencia puede corresponder a al menos un gen seleccionado de GADP, ACTB y β -actina. Se describe además en la presente memoria que el procedimiento de valoración comprende obtener una primera relación, comparando dicha primera relación dicho segundo producto amplificado de dicho primer ácido nucleico con dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico; obtener una tercera relación, comparando dicha tercera relación dicho primer producto amplificado de dicho segundo ácido nucleico de referencia con dicho primer producto amplificado de dicho molde competitivo de dicho segundo ácido nucleico de referencia y comparar dicha primera y dicha tercera relaciones. Se describe además en la presente memoria que el procedimiento comprende además diluir y coamplificar además dicho primer producto amplificado diluido de dicho primer ácido nucleico de referencia y de dicho molde competitivo de dicho primer ácido nucleico de referencia, produciendo un segundo producto amplificado del mismo. Se describe además en la presente memoria que la valoración comprende obtener una primera relación, comparando dicha primera relación dicho segundo producto amplificado de dicho primer ácido nucleico con dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico; obtener una cuarta relación, comparando dicha cuarta relación dicho segundo producto amplificado de dicho primer ácido nucleico de referencia con dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico de referencia y comparar dicha primera y dicha cuarta relaciones. Se describe además en la presente memoria que la relación comprende el uso de electroforesis capilar de microfluidos, matriz de oligonucleótidos, espectrometría de masas o cromatografía. Se describe también en la presente memoria que la relación no implica tomar medidas a tiempo real ni la generación de una curva patrón. La relación puede controlar las fuentes de variación seleccionadas de carga de ADNc, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y eficacia de amplificación intramuestra. La valoración puede detectar una diferencia de menos de aproximadamente 2 veces, una diferencia de menos de aproximadamente 1 vez, una diferencia de menos de aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 %, una diferencia de menos de aproximadamente un 30 % o una diferencia de menos de aproximadamente un 20 %. Se describe además en la presente memoria que la relación proporciona un coeficiente de variación de menos de aproximadamente un 25 %, menos de aproximadamente un 20 %, menos de aproximadamente un 15 %, menos de aproximadamente un 10 % o menos de aproximadamente un 5 % entre dicha primera muestra y una segunda muestra de dicho primer ácido nucleico. Se describe también en la presente memoria que la primera y dicha segunda muestras se amplifican en momentos diferentes, la primera y dicha segunda muestras se amplifican en laboratorios diferentes o la primera y dicha segunda muestras se proporcionan de sujetos diferentes. El primer ácido nucleico puede comprender una secuencia referenciada en la Tabla 1 o 2. Los procedimientos descritos en la presente memoria pueden reducir o eliminar los falsos negativos, preferiblemente los falsos positivos se reducen a un número estadísticamente no significativo. El ácido nucleico puede ser una molécula de ARN o una molécula de ADN. Típicamente, la relación es sustancialmente constante más allá de la fase exponencial de dicha amplificación de dicho primer ácido nucleico. Se describe además en la presente memoria que las muestras se diluyen antes de la amplificación.

Es otro aspecto de la invención un procedimiento de valoración de un primer ácido nucleico en una primera muestra como se caracteriza en las reivindicaciones adjuntas, que comprende proporcionar una mezcla estandarizada que comprende un molde competitivo de dicho primer ácido nucleico y un molde competitivo de un segundo ácido nucleico en dicha primera muestra, en el que dichos moldes competitivos están a concentraciones conocidas uno

respecto a otro; combinar dicha primera muestra con dicha mezcla estandarizada; coamplificar dicho primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico, produciendo un primer producto amplificado del mismo; diluir dicho primer producto amplificado; coamplificar además dicho primer producto amplificado de dicho primer ácido nucleico y de dicho molde competitivo de dicho primer ácido nucleico, produciendo un segundo producto amplificado del mismo y coamplificar dicho segundo ácido nucleico amplificado y dicho molde competitivo de dicho segundo ácido nucleico, produciendo el primer producto amplificado del mismo. El primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico pueden coamplificarse en un primer recipiente y dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico pueden coamplificarse en un segundo recipiente. Típicamente, la dilución produce una dilución de al menos aproximadamente 100 veces, una dilución de al menos aproximadamente 1.000 veces o una dilución de al menos aproximadamente 10.000 veces. El procedimiento puede dar cuenta menos de aproximadamente 1.000 moléculas de dicho primer ácido nucleico en dicha primera muestra, de menos de aproximadamente 100 moléculas de dicho primer ácido nucleico en dicha primera muestra, de menos de aproximadamente 10 moléculas de dicho primer ácido nucleico en dicha primera muestra o de aproximadamente 1 molécula de dicho primer ácido nucleico en dicha primera muestra. La mezcla estandarizada puede comprender además cantidades suficientes de dichos moldes competitivos para valorar dicho primer ácido nucleico en más de aproximadamente 10^6 otras muestras, más de aproximadamente 10^8 otras muestras, más de aproximadamente 10^{10} otras muestras, más de aproximadamente 10^{11} otras muestras o más de aproximadamente 10^{12} otras muestras. La mezcla estandarizada puede comprender además una serie de otro(s) molde(s) competitivo(s) de otro(s) ácido(s) nucleico(s), en la que dicho(s) molde(s) competitivo(s) están a concentraciones conocidas uno respecto a otro; permitiendo así la valoración de dichos otros ácidos nucleicos en dicha primera muestra. El número de otros moldes competitivos puede ser de al menos aproximadamente 100 o al menos aproximadamente 1.000. En algunas realizaciones, el segundo ácido nucleico sirve como primer ácido nucleico de referencia, tal como para control de carga. El primer ácido nucleico de referencia puede corresponder a al menos un gen seleccionado de GADP, ACTB y β -actina. El procedimiento de valoración comprende obtener una primera relación, comparando dicha primera relación dicho segundo producto amplificado de dicho primer ácido nucleico con dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico; obtener una segunda relación, comparando dicha segunda relación dicho primer producto amplificado de dicho primer ácido nucleico de referencia con dicho primer producto amplificado de dicho molde competitivo de dicho primer ácido nucleico de referencia y comparar dichas primera y segunda relaciones. Al menos uno de dichos otros ácidos nucleicos sirve como segundo ácido nucleico de referencia y dicho segundo ácido nucleico de referencia puede corresponder a al menos un gen seleccionado de GADP, ACTB y β -actina. El procedimiento puede comprender además coamplificar dicho segundo ácido nucleico de referencia y dicho molde competitivo de dicho segundo ácido nucleico de referencia, produciendo un primer producto amplificado del mismo. La valoración puede comprender obtener una primera relación, comparando dicha primera relación dicho segundo producto amplificado de dicho primer ácido nucleico con dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico; obtener una tercera relación, comparando dicha tercera relación dicho primer producto amplificado de dicho segundo ácido nucleico de referencia con dicho primer producto amplificado de dicho molde competitivo de dicho segundo ácido nucleico de referencia y comparar dichas primera y tercera relaciones. La mezcla estandarizada puede comprender además cantidades suficientes de dichos moldes competitivos para valorar dicho primer ácido nucleico en más de aproximadamente 10^6 otras muestras, más de aproximadamente 10^8 otras muestras, más de aproximadamente 10^{10} otras muestras, más de aproximadamente 10^{11} otras muestras o más de aproximadamente 10^{12} otras muestras. El procedimiento puede comprender además diluir y coamplificar además dicho primer producto amplificado diluido de dicho primer ácido nucleico de referencia y dicho molde competitivo de dicho primer ácido nucleico de referencia, produciendo el segundo producto amplificado del mismo. En una realización, el primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico se coamplifican además en un primer recipiente y dicho primer ácido nucleico de referencia y dicho molde competitivo de dicho primer ácido nucleico de referencia se coamplifican además en un segundo recipiente. En otra realización, la valoración comprende obtener una primera relación, comparando dicha primera relación dicho segundo producto amplificado de dicho primer ácido nucleico con dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico; obtener una cuarta relación, comparando dicha cuarta relación dicho segundo producto amplificado de dicho primer ácido nucleico de referencia con dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico de referencia y comparar dicha primera y cuarta relaciones. En diversas realizaciones, la relación comprende el uso de electroforesis capilar de microfluidos, matriz de oligonucleótidos, espectrometría de masas o cromatografía. En algunas realizaciones, la relación no implica tomar medidas a tiempo real ni la generación de una curva patrón. La relación puede controlar las fuentes de variación seleccionadas de carga de ADNc, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y eficacia de amplificación intramuestra. La valoración puede detectar una diferencia de menos de aproximadamente 2 veces, una diferencia de menos de aproximadamente 1 vez, una diferencia de menos de aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 %, una diferencia de menos de aproximadamente un 30 % o una diferencia de menos de aproximadamente un 20 %. En algunas realizaciones, la relación proporciona un coeficiente de variación de menos de aproximadamente un 25 %, menos de aproximadamente un 20 %, menos de aproximadamente un 15 %, menos de aproximadamente un 10 % o menos de aproximadamente un 5 % entre dicha primera muestra y una segunda muestra de dicho primer ácido nucleico. En algunas realizaciones, la primera y dicha segunda muestras se amplifican en momentos diferentes, la primera y dicha segunda muestras se amplifican en laboratorios diferentes o la primera y dicha segunda muestras se proporcionan de sujetos diferentes. El primer ácido nucleico puede

comprender una secuencia referenciada en la Tabla 1 o 2. Los procedimientos descritos en la presente memoria pueden reducir o eliminar los falsos negativos, preferiblemente los falsos positivos se reducen a un número estadísticamente no significativo. El ácido nucleico puede ser una molécula de ARN o una molécula de ADN.

5 Se describe también en la presente memoria un procedimiento para valorar un primer ácido nucleico que comprende proporcionar una serie de mezclas estandarizadas diluidas en serie que comprenden un molde competitivo de dicho primer ácido nucleico y un molde competitivo de un segundo ácido nucleico presente en una serie de muestras que comprenden dicho primer ácido nucleico, en el que dichos moldes competitivos están a concentraciones conocidas uno respecto a otro; combinar una de dichas muestras que comprenden dicho primer ácido nucleico con la primera de dichas mezclas estandarizadas diluidas en serie; coamplificar dicho primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico, produciendo el producto amplificado del mismo; obtener una primera relación, comparando dicha primera relación dicho producto amplificado de dicho primer ácido nucleico con dicho producto amplificado de dicho molde competitivo de dicho primer ácido nucleico; determinar si dicha primera relación está dentro de aproximadamente 1:10 a aproximadamente 10:1, si no, repetir dichas etapas de combinación, coamplificación, obtención y determinación con la segunda de dichas mezclas estandarizadas diluidas en serie; coamplificar dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico, produciendo el producto amplificado de mismo; obtener una segunda relación, comparando dicha segunda relación dicho producto amplificado de dicho segundo ácido nucleico con dicho producto amplificado de dicho molde competitivo de dicho segundo ácido nucleico y comparar dichas primera y segunda relaciones. El procedimiento puede comprender además diluir dicho producto amplificado de dicho primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico y coamplificar además dicho producto amplificado diluido, produciendo un producto amplificado adicionalmente del mismo. Además, el procedimiento puede comprender además diluir dicho producto amplificado de dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico y coamplificar además dicho producto amplificado diluido, produciendo producto amplificado adicionalmente del mismo. El número de muestras puede comprender una serie de muestras diluidas en serie de dicho segundo ácido nucleico. Se describe además en la presente memoria que una de dichas muestras se selecciona para proporcionar dicho segundo ácido nucleico aproximadamente calibrado con dicho molde competitivo de dicho segundo ácido nucleico en dicha primera de las mezclas estandarizadas diluidas en serie. Se describe además en la presente memoria que dicho primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico se coamplifican en un primer recipiente y que dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico se coamplifican en un segundo recipiente. El segundo ácido nucleico puede servir como primer ácido nucleico de referencia, tal como control de carga. El primer ácido nucleico de referencia puede ser de GADP, ACTB o β -actina. El primer ácido nucleico de referencia puede estar presente a dos concentraciones diferentes en dos mezclas estandarizadas diluidas en serie. La serie de mezclas estandarizadas diluidas en serie puede comprender además suficientes cantidades de dichos moldes competitivos para valorar dicho primer ácido nucleico en más de aproximadamente 10^6 muestras, más de aproximadamente 10^8 muestras, más de aproximadamente 10^{10} muestras o más de aproximadamente 10^{12} muestras. La serie de mezclas estandarizadas diluidas en serie puede comprender además una serie de otro(s) molde(s) competitivo(s) de otro(s) ácido(s) nucleico(s), en los que dicho(s) molde(s) competitivo(s) están a concentraciones conocidas uno respecto a otro, permitiendo así la valoración de dicho(s) otro(s) ácido(s) nucleico(s). Este número puede ser de al menos aproximadamente 100 otros moldes competitivos o de al menos aproximadamente 1.000 otros moldes competitivos. El al menos uno de dichos otros ácidos nucleicos puede servir como segundo ácido nucleico de referencia. El segundo ácido nucleico de referencia puede corresponder a al menos un gen seleccionado de GADP, ACTB y β -actina. El procedimiento de valoración puede comprender adicionalmente coamplificar dicho segundo ácido nucleico de referencia y dicho molde competitivo de dicho segundo ácido nucleico de referencia, produciendo producto amplificado del mismo; obtener una tercera relación, comparando dicha tercera relación dicho producto amplificado de dicho segundo ácido nucleico de referencia con dicho producto amplificado de dicho molde competitivo de dicho segundo ácido nucleico de referencia y comparar dicha primera y dicha tercera relaciones. La serie de mezclas estandarizadas diluidas en serie puede comprender además suficientes cantidades de dicho número de otro(s) molde(s) competitivo(s) para valorar dicho(s) otro(s) ácido(s) nucleico(s) en más de aproximadamente 10^6 muestras, en más de aproximadamente 10^8 muestras o en más de aproximadamente 10^{10} muestras. La serie de mezclas estandarizadas diluidas en serie puede comprender además suficientes cantidades de dicho número de otro(s) molde(s) competitivo(s) para valorar dicho(s) otro(s) ácido(s) nucleico(s). La serie de mezclas estandarizadas diluidas en serie puede comprender además suficientes cantidades de dicho número de otro(s) molde(s) competitivo(s) para valorar dicho(s) otros ácido(s) nucleico(s) en más de aproximadamente 10^{11} muestras o en más de aproximadamente 10^{12} muestras. El procedimiento puede efectuarse de tal forma que dicho primer ácido nucleico y dicho(s) otro(s) ácido(s) nucleico(s) varíen en una cantidad dentro de un intervalo de más de aproximadamente 2 órdenes de magnitud. El procedimiento puede detectar una diferencia de menos de aproximadamente 2 veces dentro de dicho intervalo, una diferencia de menos de un 50 % dentro de dicho intervalo o una diferencia de menos de un 20 % dentro de dicho intervalo. Se describe además en la presente memoria que dicho primer ácido nucleico y dicho(s) otro(s) ácido(s) nucleico(s) varían en una cantidad dentro de un intervalo de aproximadamente 3 o más órdenes de magnitud y que dicha valoración detecta una diferencia de menos de aproximadamente 2 veces dentro de dicho intervalo, una diferencia de menos de aproximadamente un 50 % dentro de dicho o intervalo o una diferencia de menos de aproximadamente un 20 % dentro de dicho intervalo. Se describe además en la presente memoria que el primer ácido nucleico y dicho(s) otro(s) ácido(s) nucleico(s) varían en cantidad dentro de un intervalo de aproximadamente 4 o más órdenes de magnitud y que la valoración detecta una diferencia de menos de aproximadamente 2 veces dentro de dicho

intervalo, una diferencia de menos de aproximadamente un 50 % dentro de dicho o intervalo o una diferencia de menos de aproximadamente un 20 % dentro de dicho intervalo. Se describe también en la presente memoria que el primer ácido nucleico y dicho(s) otro(s) ácido(s) nucleico(s) varían en cantidad dentro de un intervalo de aproximadamente 6 o más órdenes de magnitud y que la valoración detecta una diferencia de menos de aproximadamente 2 veces dentro de dicho intervalo, una diferencia de menos de aproximadamente 1 vez dentro de dicho intervalo, una diferencia de menos de un 80 % dentro de dicho intervalo, una diferencia de menos de un 50 % dentro de dicho intervalo, una diferencia de menos de un 30 % dentro de dicho intervalo o una diferencia de menos de un 20 % dentro de dicho intervalo. Se describe además en la presente memoria que el primer ácido nucleico y dicho(s) otro(s) ácido(s) nucleico(s) varían en cantidad dentro de un intervalo de aproximadamente 7 o más órdenes de magnitud y que dicho recuento detecta una diferencia de menos de aproximadamente 2 veces dentro de dicho intervalo, detecta una diferencia de menos de aproximadamente 1 vez dentro de dicho intervalo, una diferencia de menos de aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 % o una diferencia de menos de aproximadamente un 20 %. El intervalo puede ser también de aproximadamente 8, aproximadamente 9, aproximadamente 10 o aproximadamente 15. En algunos procedimientos, los aproximadamente 7 órdenes de magnitud se extienden por un intervalo de aproximadamente 7 unidades logarítmicas de expresión génica y dichos aproximadamente 7 órdenes de magnitud pueden incluir aproximadamente 10^{-3} , aproximadamente 10^{-2} , aproximadamente 0,1, aproximadamente 1, aproximadamente 10, aproximadamente 10^2 , aproximadamente 10^3 y aproximadamente 10^4 copias/célula. El primer ácido nucleico puede comprender una secuencia referenciada en la Tabla 1 o 2. El molde competitivo de dicho primer o dicho segundo ácido nucleico puede comprender una secuencia referenciada en la Tabla 4. El molde competitivo de dicho primer ácido nucleico puede estar a una serie de concentraciones respecto a dicho molde competitivo de dicho segundo ácido nucleico. La serie de concentraciones puede proporcionar diluciones en serie de 10 veces de dicho molde competitivo de dicho primer ácido nucleico respecto a dicho molde competitivo de dicho segundo ácido nucleico. Al menos dos de dichas series de concentraciones pueden extenderse por aproximadamente 1 orden de magnitud, al menos dos de dichas series de concentraciones se extienden por aproximadamente 3 órdenes de magnitud, o al menos dos de dichas series de concentraciones se extienden por aproximadamente 6 órdenes de magnitud. Las series de concentraciones pueden incluir al menos dos concentraciones seleccionadas de aproximadamente 10^{-11} M, aproximadamente 10^{-12} M, aproximadamente 10^{-13} M, aproximadamente 10^{-14} M, aproximadamente 10^{-15} M y aproximadamente 10^{-16} M. Las series de concentraciones pueden incluir al menos 3 concentraciones seleccionadas de aproximadamente 10^{-11} M, aproximadamente 10^{-12} M, aproximadamente 10^{-13} M, aproximadamente 10^{-14} M, aproximadamente 10^{-15} M y aproximadamente 10^{-16} M. Las series de concentraciones pueden incluir al menos 6 concentraciones de aproximadamente 10^{-11} M, aproximadamente 10^{-12} M, aproximadamente 10^{-13} M, aproximadamente 10^{-14} M, aproximadamente 10^{-15} M y aproximadamente 10^{-16} M. La primera o dicha segunda relación pueden obtenerse con el uso de electroforesis capilar de microfluidos, matriz de oligonucleótidos, espectrometría de masas o cromatografía. Se describe adicionalmente en la presente memoria que la primera o dicha segunda relación no implican tomar medidas a tiempo real ni la generación de una curva patrón. Las mezclas estandarizadas pueden controlar las fuentes de variación seleccionadas de carga de ADNc, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y eficacia de amplificación intramuestra. Las mezclas estandarizadas de dicha serie pueden dar cuenta menos de aproximadamente 1.000 moléculas de dicho primer ácido nucleico en una de dichas muestras, menos de aproximadamente 100 moléculas de dicho primer ácido nucleico en una de dichas muestras, menos de aproximadamente 10 moléculas de dicho primer ácido nucleico en una de dichas muestras o aproximadamente 1 molécula de dicho primer ácido nucleico en una de dichas muestras. Las mezclas estandarizadas de dicha serie pueden proporcionar un coeficiente de variación de menos de aproximadamente un 25 %, menos de aproximadamente un 15 % o menos de aproximadamente un 10 % entre dos de dichas muestras que comprenden dicho primer ácido nucleico. El procedimiento puede efectuarse en muestras obtenidas de diferentes sujetos, diferentes laboratorios o en diferentes momentos. El procedimiento puede reducir o eliminar falsos negativos a un número no significativo, tal como un número estadísticamente no significativo. El procedimiento puede estar informatizado, comprendiendo la informatización instruir a un manipulador robótico para seleccionar dicha primera de las mezclas estandarizadas diluidas en serie para combinación. La informatización puede comprender obtener dicha primera relación, tal como determinar un área bajo la curva. La informatización puede comprender instruir a dicho manipulador robótico para seleccionar dicha segunda de las mezclas estandarizadas diluidas en serie basándose en dicha primera relación. El ácido nucleico valorado puede ser una molécula de ARN o una molécula de ADN.

Se describe además en la presente memoria un procedimiento para preparar una mezcla estandarizada de reactivos, comprendiendo dicho reactivos suficiente molde competitivo para valorar las cantidades de una serie de ácidos nucleicos en más de aproximadamente 10^6 muestras, en el que dichas mezclas estandarizadas permiten la comparación directa de dichas cantidades entre 2 de dichas muestras. El número puede ser de 2 ácidos nucleicos, es de al menos aproximadamente 96 ácidos nucleicos, al menos aproximadamente 100 ácidos nucleicos o al menos aproximadamente 1.000 ácidos nucleicos. El procedimiento puede implicar el uso de reactivos que sean suficientes para valorar dichas cantidades en más de aproximadamente 10^8 muestras, más de aproximadamente 10^{10} muestras, más de aproximadamente 10^{11} muestras o más de aproximadamente 10^{12} muestras. El procedimiento puede emplear reactivos que comprenden además un cebador de codificación y/o un cebador inverso para cebar la amplificación de dicho molde competitivo de dicho número de ácido(s) nucleico(s). El molde competitivo, dicho cebador de codificación y/o dicho cebador inverso pueden comprender una secuencia referenciada en la Tabla 4. El cebador de codificación y/o cebador inverso puede tener sustancialmente la misma temperatura de asociación que

otro cebador de codificación y/o cebador inverso en dicha mezcla estandarizada. El cebador de codificación y/o cebador inverso puede permitir la detección de aproximadamente 600 moléculas o menos de dicho(s) ácido(s) nucleico(s), de aproximadamente 60 moléculas o menos de dicho(s) ácido(s) nucleico(s) o de aproximadamente 6 moléculas o menos de dicho(s) ácido(s) nucleico(s). Al menos uno de dichos ácidos nucleicos puede comprender una secuencia referenciada en la Tabla 1 o 2. Uno de dicho número de ácidos nucleicos puede servir como primer ácido nucleico de referencia. El primer ácido nucleico de referencia puede ser un control de carga y puede ser de GADP, ACTB o β -actina. También, otro de dicho número de ácidos nucleicos sirve como segundo ácido nucleico de referencia. El segundo ácido nucleico de referencia puede ser un gen seleccionado de GADP, ACTB y β -actina. La valoración puede efectuarse con el uso de electroforesis capilar de microfluidos, matriz de oligonucleótidos, espectrometría de masas o cromatografía. Se describe además en la presente memoria que el procedimiento no implica la toma de medidas a tiempo real ni la generación de una curva patrón. Las mezclas estandarizadas pueden controlar las fuentes de variación tales como carga de ADNc, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y eficacia de amplificación intramuestra. Las mezclas estandarizadas de dicha serie pueden dar cuenta menos de aproximadamente 1.000 moléculas de dicho primer ácido nucleico en una de dichas muestras, de menos de aproximadamente 100 moléculas de dicho primer ácido nucleico en una de dichas muestras, de menos de aproximadamente 10 moléculas de ácido nucleico en una de dichas muestras o de aproximadamente 1 molécula de ácido nucleico en una de dichas muestras. Las mezclas estandarizadas de dicha serie pueden proporcionar un coeficiente de variación de menos de aproximadamente un 25 %, menos de aproximadamente un 15 % o menos de aproximadamente un 10 % entre 2 de dichas muestras que comprenden dicho primer ácido nucleico. El procedimiento puede efectuarse en muestras obtenidas de diferentes sujetos, diferentes laboratorios o en diferentes momentos. El procedimiento puede reducir o eliminar los falsos negativos a un número no significativo, tal como un número estadísticamente no significativo. El ácido nucleico valorado puede ser una molécula de ARN o una molécula de ADN.

Se describe adicionalmente en la presente memoria un procedimiento que comprende preparar una serie de mezclas estandarizadas de reactivos, comprendiendo dicho reactivo suficiente molde competitivo para valorar las cantidades de un número de ácidos nucleicos en más de aproximadamente 10^6 muestras, en el que dichas mezclas estandarizadas permiten la comparación directa de dichas cantidades entre 2 de dichas muestras. El número puede ser de 2 ácidos nucleicos, es de al menos aproximadamente 96 ácidos nucleicos, al menos aproximadamente 100 ácidos nucleicos o al menos aproximadamente 1.000 ácidos nucleicos. El procedimiento puede implicar el uso de reactivos que sean suficientes para valorar dichas cantidades en más de aproximadamente 10^8 muestras, más de aproximadamente 10^{10} muestras, más de aproximadamente 10^{11} muestras o más de aproximadamente 10^{12} muestras. En algunos procedimientos, las cantidades pueden variar dentro de un intervalo de más de aproximadamente 2 órdenes de magnitud, dentro de un intervalo de aproximadamente 3 o más órdenes de magnitud, dentro de un intervalo de 4 o más órdenes de magnitud, dentro de un intervalo de aproximadamente 6 o más órdenes de magnitud o dentro de un intervalo de aproximadamente 7 o más órdenes de magnitud. La serie puede permitir la detección de una diferencia de menos de aproximadamente 2 veces dentro de dicho intervalo, de una diferencia de menos de aproximadamente un 50 % dentro de dicho intervalo, de una diferencia de menos de aproximadamente un 20 % dentro de dicho intervalo. También se engloban en la presente memoria aproximadamente 8 o más órdenes de magnitud, aproximadamente 9 o más órdenes de magnitud, aproximadamente 10 o más órdenes de magnitud o aproximadamente 15 o más órdenes de magnitud. Los aproximadamente 7 órdenes de magnitud pueden extenderse por un intervalo de aproximadamente 7 unidades logarítmicas de expresión génica y pueden incluir aproximadamente 10^{-3} , aproximadamente 10^{-2} , aproximadamente 0,1, aproximadamente 1, aproximadamente 10, aproximadamente 10^2 , aproximadamente 10^3 y aproximadamente 10^4 copias/célula. El procedimiento puede emplear reactivos que comprenden además un cebador de codificación y/o un cebador inverso para cebar la amplificación de dicho molde competitivo de dicho número de ácido(s) nucleico(s). El molde competitivo, dicho cebador de codificación y/o dicho cebador inverso pueden comprender una secuencia referenciada en la Tabla 4. El cebador de codificación y/o dicho cebador inverso pueden tener sustancialmente la misma temperatura de asociación que otro cebador de codificación y/o cebador inverso en dicha mezcla estandarizada. El cebador de codificación y/o dicho cebador inverso puede permitir la detección de aproximadamente 600 moléculas o menos de dicho(s) ácido(s) nucleico(s), de aproximadamente 60 moléculas o menos de dicho(s) ácido(s) nucleico(s) o de aproximadamente 6 moléculas o menos de dicho(s) ácido(s) nucleico(s). Al menos uno de dichos ácidos nucleicos puede comprender una secuencia referenciada en la Tabla 1 o 2. Los moldes competitivos pueden comprender un primer molde competitivo del primero de dichos ácidos nucleicos y un segundo molde competitivo del segundo de dichos ácidos nucleicos, en los que dicho primer molde competitivo está a una serie de concentraciones respecto a dicho segundo molde competitivo. El segundo ácido nucleico puede servir como primer ácido nucleico de referencia, tal como control para carga, y ser de GADP, ACTB, o β -actina. En el procedimiento, la serie de concentraciones puede proporcionar diluciones en serie de 10 veces de dicho primer molde competitivo respecto a dicho segundo molde competitivo. Al menos dos de dichas series de concentraciones se extienden por aproximadamente 1 orden de magnitud, aproximadamente 3 órdenes de magnitud o aproximadamente 6 órdenes de magnitud. Las series de concentraciones pueden incluir concentraciones seleccionadas de aproximadamente 10^{-11} M, aproximadamente 10^{-12} M, aproximadamente 10^{-13} M, aproximadamente 10^{-14} M, aproximadamente 10^{-15} M y aproximadamente 10^{-16} M. La valoración puede efectuarse con el uso de electroforesis capilar de microfluidos, matriz de oligonucleótidos, espectrometría de masas o cromatografía. Se describe además en la presente memoria que el procedimiento no implica la toma de medidas a

tiempo real ni la generación de una curva patrón. Las mezclas estandarizadas pueden controlar las fuentes de variación tales como carga de ADNc, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y eficacia de amplificación intramuestra. Las mezclas estandarizadas de dicha serie pueden dar cuenta de menos de aproximadamente 1.000 moléculas de dicho primer ácido nucleico en una de dichas muestras, de menos de aproximadamente 100 moléculas de dicho primer ácido nucleico en una de dichas muestras, de menos de aproximadamente 10 moléculas de ácido nucleico en una de dichas muestras o de aproximadamente 1 molécula de ácido nucleico de dicho primer ácido nucleico en una de dichas muestras. Las mezclas estandarizadas de dicha serie pueden proporcionar un coeficiente de variación de menos de aproximadamente un 25 %, menos de aproximadamente un 15 % o menos de aproximadamente un 10 % entre 2 de dichas muestras que comprenden dicho primer ácido nucleico. El procedimiento puede efectuarse en muestras obtenidas de diferentes sujetos, diferentes laboratorios o en diferentes momentos. El procedimiento puede reducir o eliminar los falsos negativos a un número no significativo, tal como un número estadísticamente no significativo. El ácido nucleico puede ser una molécula de ARN o una molécula de ADN.

Se describen también en la presente memoria composiciones para uso en los procedimientos descritos en la presente memoria. Se describe por tanto una composición que comprende una mezcla estandarizada de reactivos, comprendiendo dichos reactivos suficiente molde competitivo para valorar cantidades de un número de ácidos nucleicos en más de aproximadamente 10^6 muestras, en la que dicha mezcla estandarizada permite la comparación directa de dichas cantidades entre 2 de dichas muestras. Se describe también una composición que comprende una serie de mezclas estandarizadas diluidas en serie de reactivos, comprendiendo dichos reactivos suficiente molde competitivo para valorar las cantidades de un número de ácidos nucleicos en más de aproximadamente 10^6 muestras, en la que dichas mezclas estandarizadas permiten la comparación directa de dichas cantidades entre 2 de dichas muestras.

Otro aspecto descrito en la presente memoria es una base de datos que comprende valores numéricos correspondientes a las cantidades de un primer ácido nucleico en un número de muestras, en el que dichos valores numéricos son directamente comparables entre aproximadamente 5 de dichas muestras. En la base de datos, el número puede ser de al menos aproximadamente 10 muestras, al menos aproximadamente 100 muestras, al menos aproximadamente 1.000 muestras, al menos aproximadamente 5.000 muestras o al menos aproximadamente 10.000 muestras. Las muestras pueden obtenerse de diferentes sujetos, de diferentes especies, de diferentes laboratorios o en diferentes momentos. En la base de datos, las cantidades pueden mostrar un coeficiente de variación de menos de aproximadamente un 25 %, menos de aproximadamente un 20 %, menos de aproximadamente un 15 %, menos de aproximadamente un 10 % o menos de aproximadamente un 5 % entre dichas 2 muestras. La base de datos puede comprender además valores numéricos correspondientes a las cantidades de un número de otro(s) ácido(s) nucleico(s) en dicho número de muestras. El número puede ser de al menos aproximadamente 100 otros ácidos nucleicos, al menos aproximadamente 1.000 otros ácidos nucleicos o al menos aproximadamente 10.000 otros ácidos nucleicos. Las cantidades pueden obtenerse usando electroforesis capilar de microfluidos, matriz de oligonucleótidos, espectrometría de masas o cromatografía. Se describe además que las cantidades no se obtienen usando medidas a tiempo real ni por la generación de una curva patrón. Los valores numéricos pueden corregirse para fuentes de variación tales como carga de ADNc, eficacia de la amplificación intraácido nucleico, eficacia de la amplificación interácido nucleico, eficacia de la amplificación interespecimen, eficacia de la amplificación intermuestra y eficacia de la amplificación intramuestra. Los valores numéricos corresponden típicamente a números de moléculas de dicho primer ácido nucleico en dicho número de muestras. En la base de datos, al menos uno de dichos valores numéricos puede corresponder a menos de aproximadamente 1.000 moléculas, menos de aproximadamente 100 moléculas, menos de aproximadamente 10 moléculas o aproximadamente 1 molécula de dicho primer ácido nucleico en al menos una de dichas muestras. Los valores numéricos pueden corresponder a una diferencia de menos de aproximadamente 2 veces o una diferencia de menos de aproximadamente 1 vez en dicho primer ácido nucleico entre 2 de dichas muestras. Los valores numéricos pueden corresponder a una diferencia de menos de aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 %, una diferencia de menos de aproximadamente un 30 % o una diferencia de menos de aproximadamente un 20 % en dicho primer ácido nucleico entre 2 de dichas muestras. Los valores numéricos pueden variar dentro de un intervalo de más de aproximadamente 2 órdenes de magnitud, dentro de un intervalo de aproximadamente 4 o más órdenes de magnitud, dentro de un intervalo de aproximadamente 6 o más órdenes de magnitud o dentro de un intervalo de aproximadamente 7 o más órdenes de magnitud. Los aproximadamente 7 órdenes de magnitud pueden extenderse por un intervalo de aproximadamente 7 unidades logarítmicas de expresión génica a magnitudes que incluyen aproximadamente 10^{-3} , aproximadamente 10^{-2} , aproximadamente 0,1, aproximadamente 1, aproximadamente 10, aproximadamente 10^2 , aproximadamente 10^3 y aproximadamente 10^4 copias/célula. Los valores numéricos típicamente no comprenden un número estadísticamente significativo de falsos positivos. Los valores numéricos pueden usarse en al menos una etapa del desarrollo de fármacos seleccionada de cribado de fármaco diana, identificación del compuesto principal, validación preclínica, ensayo clínico y tratamiento de pacientes. La validación preclínica puede ser un bioensayo y/o un estudio animal. Se describe además en la presente memoria que la comparación directa en la base de datos no usa recursos bioinformáticos. El ácido nucleico comprende una molécula de ARN o una molécula de ADN. La base de datos puede indicar el nivel de expresión génica correspondiente a un estado biológico, tal como un estado normal o estado patológico.

Se describe también en la presente memoria que la base de datos comprende índices numéricos, estando obtenidos dichos índices numéricos mediante cómputo matemático de 2 valores numéricos, siendo dichos 2 valores numéricos correspondientes a cantidades de 2 ácidos nucleicos en un número de muestras, en la que dichos índices numéricos son directamente comparables entre 5 de dichas muestras. Los índices numéricos pueden indicar un estado biológico. Se describe además en la presente memoria que al menos un índice numérico es un índice numérico equilibrado. El índice numérico puede calcularse dividiendo un numerador entre un denominador, correspondiendo dicho numerador a dicha cantidad de uno de dichos 2 ácidos nucleicos y correspondiendo dicho denominador a dicha cantidad del otro de dichos 2 ácidos nucleicos. El numerador puede corresponder a un gen asociado positivamente a dicho estado biológico y dicho denominador corresponde a un gen asociado negativamente a dicho estado biológico. En la base de datos, dicho estado biológico puede ser un estado patológico, una predisposición a un estado patológico, una respuesta a fármaco terapéutico, una predisposición a respuesta a fármaco terapéutico, una respuesta adversa a fármaco, una predisposición a una respuesta adversa a fármaco, una toxicidad de fármaco o una predisposición a toxicidad de fármaco.

Se describe también en la presente memoria un procedimiento para obtener un índice numérico que indica un estado biológico, que comprende proporcionar 2 muestras correspondientes a cada uno de un primer estado biológico y un segundo estado biológico; valorar una cantidad de cada uno de 2 ácidos nucleicos en cada una de dichas 2 muestras, en el que dicha valoración puede dar cuenta menos de aproximadamente 1.000 moléculas de cada uno de dichos 2 ácidos nucleicos; proporcionar dichas cantidades como valores numéricos, en los que dichos valores numéricos son directamente comparables entre un número de muestras; computar matemáticamente dichos valores numéricos correspondientes a cada uno de dichos primer y dicho segundo estados biológicos y determinar un cómputo matemático que discrimine dichos primer y dicho segundo estados biológicos, obteniendo así dicho índice numérico. El procedimiento de determinación de dicho cómputo matemático puede implicar el uso de software. Los 2 ácidos nucleicos pueden estar asociados con dicho primer estado biológico y no con dicho segundo estado biológico. Los 2 ácidos nucleicos pueden estar asociados positivamente con dicho primer estado biológico y el otro de dichos 2 ácidos nucleicos está asociado negativamente con dicho primer estado biológico. El cómputo matemático puede comprender dividir un numerador entre un denominador, correspondiendo dicho numerador a dicho ácido nucleico asociado positivamente con dicho primer estado biológico y correspondiendo dicho denominador a dicho ácido nucleico asociado negativamente con dicho primer estado biológico. El primer estado biológico puede ser un estado patológico y dicho segundo estado biológico es un estado no patológico. Puede ser una afección relacionada con la angiogénesis, una afección relacionada con antioxidantes, una afección relacionada con la apoptosis, una afección relacionada con el sistema cardiovascular, una afección relacionada con el ciclo celular, una afección relacionada con la estructura celular, una afección relacionada con citocinas, una afección relacionada con la respuesta de defensa, una afección relacionada con el desarrollo, una afección relacionada con la diabetes, una afección relacionada con la diferenciación, una afección relacionada con la replicación y/o reparación de ADN, una afección relacionada con células endoteliales, una afección relacionada con el receptor de folato, una afección relacionada con receptores hormonales, una afección relacionada con la inflamación, una afección relacionada con el metabolismo intermedio, una afección relacionada con el transporte de membrana, una afección relacionada con el metabolismo oxidativo, una afección relacionada con la neurotransmisión, una afección relacionada con el cáncer, una afección relacionada con la maduración de proteínas, una afección relacionada con la transducción de señal, una afección relacionada con la respuesta al estrés, una afección relacionada con la estructura de tejido, una afección relacionada con los factores de transcripción, una afección relacionada con el transporte o una afección relacionada con el metabolismo xenobiótico. Se describe adicionalmente que la comparación directa no usa recursos bioinformáticos.

Se describe además en la presente memoria un procedimiento que comprende estado y no con un segundo estado biológico; proporcionar 2 muestras correspondientes a cada uno de dicho primer estado biológico y dicho segundo estado biológico; valorar una cantidad de cada uno de dichos 2 ácidos nucleicos en dichas 2 muestras, en el que dicha valoración puede dar cuenta menos de aproximadamente 1.000 moléculas de cada uno de dichos 2 ácidos nucleicos y computar matemáticamente dichas cantidades correspondientes a cada uno de dichos primer y segundo estados biológicos para determinar un índice numérico, discriminando dicho índice numérico dichos primer y segundo estados biológicos.

Se describe también en la presente memoria un procedimiento de identificación de un estado biológico que comprende valorar una cantidad de cada uno de los 2 ácidos nucleicos en una primera muestra, en el que dicha valoración puede dar cuenta de menos de aproximadamente 1.000 moléculas de cada uno de dichos 2 ácidos nucleicos en dicha primera muestra; proporcionar dichas cantidades como valores numéricos, en los que dichos valores numéricos son directamente comparables entre un número de muestras, y usar dichos valores numéricos para proporcionar un índice numérico, indicando dicho índice numérico dicho estado biológico.

Se describe también en la presente memoria un procedimiento de identificación de un estado biológico que comprende valorar una cantidad de un ácido nucleico en una primera muestra, en el que dicha valoración puede dar cuenta de menos de aproximadamente 1.000 moléculas de dicho ácido nucleico en dicha primera muestra, y proporcionar dicha cantidad como un valor numérico en el que dicho valor numérico es directamente comparable entre un número de otras muestras.

Se describen además en la presente memoria procedimientos comerciales. Se describe por tanto un procedimiento

comercial que comprende recoger un primer espécimen que comprende un primer ácido nucleico; medir una cantidad de dicho primer ácido nucleico en una primera muestra de dicho primer espécimen, en el que dicha medida puede dar cuenta de menos de aproximadamente 1.000 moléculas de dicho primer ácido nucleico en dicha primera muestra, y proporcionar dicha cantidad como un valor numérico en el que dicho valor numérico permite la comparación directa con una cantidad de dicho primer ácido nucleico en una segunda muestra. La primera y segunda muestras pueden medirse en momentos diferentes o en laboratorios diferentes. La segunda muestra puede obtenerse de dicho primer espécimen o de un segundo espécimen. El primer y segundo especímenes pueden recogerse de diferentes sujetos o de diferentes especies. La etapa de medida puede efectuarse al menos aproximadamente 100 veces al día, al menos aproximadamente 1.000 veces al día o al menos aproximadamente 4.000 veces al día. El primer espécimen puede comprender al menos aproximadamente 1.0000 células. El primer espécimen puede comprender un espécimen humano, que puede recogerse sin información identificativa. La información de recogida puede comprender la acreditación de cumplir el protocolo de investigación. La información de identificación puede recogerse en un momento posterior a dicha recogida de dicho primer espécimen. La información puede recogerse a través de un sitio web. El procedimiento puede comprender además identificar cuáles de dichos ácidos nucleicos seleccionados se someten a electroforesis conjuntamente. Las cantidades de dichos ácidos nucleicos identificados pueden someterse a electroforesis simultáneamente. El valor numérico puede proporcionarse por correo electrónico. La valoración puede comprender proporcionar una mezcla estandarizada que comprende un molde competitivo de dicho primer ácido nucleico y un molde competitivo de un segundo ácido nucleico en dicho primer espécimen, en la que dichos moldes competitivos están a concentraciones conocidas uno respecto a otro; combinar dicha mezcla estandarizada con una primera muestra de dicho espécimen; coamplificar dicho primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico, produciendo un primer producto amplificado del mismo; diluir dicho primer producto amplificado; coamplificar además dicho primer producto amplificado diluido de dicho primer ácido nucleico y de dicho molde competitivo de dicho primer ácido nucleico, produciendo un segundo producto amplificado del mismo y coamplificar dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico, produciendo un primer producto amplificado del mismo. El primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico pueden coamplificarse en un primer recipiente y dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico se coamplifican en un segundo recipiente. El procedimiento puede comprender además obtener una primera relación, comparando dicha primera relación dicho segundo producto amplificado de dicho primer ácido nucleico y dicho segundo producto amplificado de dicho molde competitivo de dicho primer ácido nucleico; obtener una segunda relación, comparando dicha segunda relación dicho primer producto amplificado de dicho segundo ácido nucleico y dicho primer producto amplificado de dicho molde competitivo de dicho segundo ácido nucleico, y comparar dichas primera y segunda relaciones. El segundo ácido nucleico puede servir como ácido nucleico de referencia. La mezcla estandarizada puede comprender adicionalmente cantidades suficientes de dichos moldes competitivos para valorar dicho primer ácido nucleico en más de aproximadamente 10^6 muestras.

Se describen también en la presente memoria procedimientos comerciales para el desarrollo de fármacos. Se describe por tanto un procedimiento comercial de mejora del desarrollo de fármacos, que comprende recoger un primer espécimen que comprende un ácido nucleico de una primera entidad biológica con administración de un fármaco candidato en la primera etapa del desarrollo de fármacos; recoger un segundo espécimen que comprende dicho ácido nucleico de una segunda entidad biológica en una segunda etapa del desarrollo de fármacos; valorar la cantidad de dicho ácido nucleico en cada uno de dichos primer y segundo especímenes; comparar directamente dichas cantidades y alterar una etapa de dicho desarrollo de fármacos basándose en dicha comparación. Se describe además un procedimiento comercial de mejora del desarrollo de fármacos, que comprende proporcionar una base de datos que comprende valores numéricos correspondientes a las cantidades del primer ácido nucleico en un número de muestras, en el que dichos valores numéricos son directamente comparables entre 5 de dichas muestras; recoger un primer espécimen que comprende dicho primer ácido nucleico de una entidad biológica con administración de un fármaco candidato en una etapa del desarrollo de fármacos; valorar la cantidad de dicho primer ácido nucleico en una primera muestra de dicho primer espécimen; comparar directamente dicha cantidad con al menos uno de dichos valores numéricos de dicha base de datos y alterar una etapa de dicho desarrollo de fármacos basándose en dicha comparación. La primera o dicha segunda entidad biológica es típicamente al menos una entidad seleccionada de un virus, una célula, un tejido, un cultivo *in vitro*, una planta, un animal y un sujeto que participa en un ensayo clínico. La primera o dicha segunda etapa del desarrollo de fármacos puede ser cribado de fármaco diana, identificación del compuesto principal, validación preclínica, ensayo clínico y/o tratamiento de pacientes. La validación preclínica puede ser un bioensayo y/o un estudio animal. La alteración puede comprender una estratificación de un ensayo clínico. La estratificación puede implicar identificar sujetos que tienen efectos secundarios reducidos. La alteración puede reducir el tiempo para dicho desarrollo de fármacos.

Se describe además también un procedimiento comercial de mejora del desarrollo de fármacos, que comprende proporcionar una base de datos que comprende índices numéricos, estando obtenidos dichos índices numéricos mediante cómputo matemático de 2 valores numéricos correspondientes a las cantidades de 2 ácidos nucleicos en un número de muestras, en el que dichos índices numéricos son directamente comparables entre 5 de dichas muestras; recoger un primer espécimen que comprende dichos 2 ácidos nucleicos de una entidad biológica con administración de un fármaco candidato en una etapa del desarrollo de fármacos; valorar la cantidad de cada uno de dichos 2 ácidos nucleicos en una primera muestra de dicho primer espécimen; usar dichas 2 cantidades para computar matemáticamente un primer índice numérico; comparar directamente dicho primer índice numérico con al

menos uno de dichos índices numéricos en dicha base de datos y alterar una etapa de dicho desarrollo de fármacos basándose en dicha comparación.

BREVE DESCRIPCIÓN DE LAS FIGURAS

- 5 Los rasgos novedosos de la invención se exponen particularmente en las reivindicaciones adjuntas. Se obtendrá una mejor comprensión de los objetos, rasgos y ventajas de la presente invención mediante referencia a la siguiente descripción detallada, que expone realizaciones ilustrativas en que se utilizan los principios de la invención y los dibujos adjuntos, los cuales:
- 10 La Figura 1 ilustra una tabla que muestra valores numéricos para un número de ácidos nucleicos correspondientes a las medidas de expresión para algunos candidatos a genes quimiorresistentes a carboplatino en carcinoma pulmonar no microcítico (NSCLC) primario.
- La Figura 2 ilustra una tabla que muestra los valores numéricos para un número de ácidos nucleicos correspondientes a las medidas de expresión de una serie de genes en células epiteliales de vías aéreas de donante de pulmón.
- La Figura 3 ilustra un proceso global “en dos etapas” para evaluar ácidos nucleicos en algunas realizaciones.
- 15 La Figura 4 ilustra una tabla que proporciona valores numéricos para un número de ácidos nucleicos correspondientes a las medidas de expresión de un número de genes derivados de ARN de referencia humano de Stratagene, medidos usando realizaciones de enfoques tanto en dos etapas como no en dos etapas. Se proporcionan también las correspondientes secuencias usadas como cebador de codificación (F), cebador inverso (R) y molde competitivo (CT) para cada uno de los genes (SEC ID N° 1-282).
- 20 La Figura 5 ilustra la relación entre la cantidad de ácido nucleico usada en una reacción PCR y el número de copias de transcritos de ARNm/célula que puede medirse para un número dado de células/reacción PCR.
- La Figura 6 ilustra una mezcla estandarizada usada en algunas realizaciones de la presente invención.
- La Figura 7 ilustra el recálculo de los valores numéricos basados en un primer ácido nucleico de referencia (β -actina) a valores numéricos basados en un segundo ácido nucleico de referencia (ciclofilina).
- 25 La Figura 8 ilustra el uso de una serie de mezclas estandarizadas según algunas realizaciones de la presente invención.
- La Figura 9 ilustra el uso de un ácido nucleico que sirve como referencia para equilibrar una muestra con una mezcla estandarizada de una serie de mezclas estandarizadas diluidas en serie.
- 30 La Figura 10 ilustra la dilución de ADNc que proporciona un ácido nucleico de referencia (β -actina) en equilibrio con 600.000 moléculas del ácido nucleico de referencia de un molde competitivo en una mezcla estandarizada.
- La Figura 11 ilustra una serie de mezclas estandarizadas diluidas en serie A-F que comprenden una serie de concentraciones de moldes competitivos de ácidos nucleicos diana (6.000.000, 600.000, 60.000, 6.000, 600 y 60 moléculas/ μ l, respectivamente) respecto a una concentración dada de molde competitivo de β -actina (600.000 moléculas/ μ l).
- 35 La Figura 12 ilustra el uso de la mezcla E inicialmente, basándose en los niveles de expresión de la mayoría de genes.
- La Figura 13 ilustra una situación en la que la mezcla inicial usada no proporciona un molde competitivo del ácido nucleico diana (c-myc) suficientemente en equilibrio con la cantidad de ácido nucleico diana a la dilución de ADNc usada.
- 40 La Figura 14 ilustra la selección de una mezcla posterior, la mezcla C, basándose en los resultados obtenidos usando la primera mezcla.
- La Figura 15 ilustra la situación en que la mezcla posterior seleccionada, la mezcla C, proporciona un molde competitivo del ácido nucleico diana (c-myc) suficientemente en equilibrio con la cantidad de ácido nucleico diana a la dilución de ADNc usada.
- 45 La Figura 16 ilustra el cálculo del “cociente de cocientes” basado en los datos obtenidos usando una mezcla apropiada.
- La Figura 17 ilustra una serie de electroferogramas de diversos genes.
- La Figura 18 ilustra un sistema global para la valoración de ácidos nucleicos, una o más etapas del cual pueden informatizarse en diversas realizaciones.

- 5 La Figura 19 ilustra la relación no lineal entre la cantidad de producto amplificado de glutatión peroxidasa GSH-Px (cuadrados blancos) o de gliceraldehído-3-fosfato dehidrogenasa GAPDH (cuadrados negros) y la cantidad total de partida de ARN para cantidades crecientes de ARN, por ejemplo, más allá de la fase exponencial de amplificación. Las líneas rectas representan las cantidades teóricas de producto de PCR (GSH-Px o GAPDH) que se obtendrían si la amplificación permaneciera exponencial a lo largo del proceso de amplificación.
- La Figura 20 ilustra la relación lineal entre el cociente de (producto amplificado de ácido nucleico/producto coamplificado de su molde competitivo) y la cantidad total de partida de ARN para el primer y segundo ácidos nucleicos correspondientes a GSH-Px (cuadrados blancos) y GAPDH (cuadrados negros), respectivamente.
- 10 La Figura 21 ilustra que la relación de (producto amplificado del primer ácido nucleico/producto coamplificado de su molde competitivo)/(producto amplificado del segundo ácido nucleico/producto coamplificado de su molde competitivo) a la cantidad total de partida de ARN permanece constante, o sustancialmente constante, para dos ácidos nucleicos diferentes cuando se amplifican de acuerdo con diversas realizaciones de la presente invención.
- La Figura 22 tabula el número de fuentes de variación y los procedimientos de control.
- 15 La Figura 23 ilustra el control de una o más fuentes de error en algunas realizaciones en comparación con RT-PCR a tiempo real en dos especímenes diferentes en cuatro experimentos diferentes.
- La Figura 24 ilustra el desarrollo y uso de una base de datos de valores numéricos de algunas realizaciones descritas en la presente memoria.
- La Figura 25 ilustra el uso de índices numéricos en la identificación de un estado biológico.
- 20 La Figura 26 ilustra el proceso global respecto al uso de cribados de micromatrices con realizaciones de la presente invención.
- La Figura 27 ilustra el proceso global de algunas realizaciones de un procedimiento comercial para evaluar ácidos nucleicos.
- La Figura 28 ilustra el proceso global de algunas realizaciones de un procedimiento comercial para mejorar el desarrollo de fármacos.
- 25 La Figura 29 ilustra experimentos que comparan un enfoque no en dos pasos con uno en dos pasos según algunas realizaciones de la presente invención.
- La Figura 30 ilustra los resultados de experimentos que comparan un enfoque no en dos etapas con uno en dos etapas según algunas realizaciones de la presente invención.
- 30 La Figura 31 es una gráfica que muestra la correlación de los valores de expresión génica obtenidos por enfoques de 96 genes en dos etapas o no en dos etapas.
- La Figura 32 ilustra un procedimiento para diseñar moldes competitivos para uso en algunas realizaciones de la presente invención.
- La Figura 33 ilustra el cálculo de la expresión génica basándose en valores densitométricos para el producto separado electroforéticamente de NT y CT de GST.
- 35 La Figura 34 ilustra fotografías negativas de los gels analizados por densitometría.
- La Figura 35 ilustra un cálculo basado en los valores de densitometría.
- La Figura 36 ilustra que se observó un aumento similar en la expresión del gen CYP1A1 tanto en análisis Northern como en algunas realizaciones de los procedimientos dados a conocer en la presente memoria.
- 40 La Figura 37 ilustra los resultados de experimentos que comparan algunas realizaciones de la presente invención con análisis de micromatrices de oligonucleótidos.
- La Figura 38 ilustra el mayor intervalo dinámico lineal obtenido usando algunas realizaciones en dos etapas frente al análisis de micromatrices.
- 45 Cada una de estas figuras proporciona solo ilustración, y no pretende en modo alguno ser limitante con respecto a la presente invención. Por ejemplo, los especialistas en la materia apreciarán fácilmente las variaciones y modificaciones de los esquemas ilustrados basándose en las enseñanzas proporcionadas en la presente memoria.

DESCRIPCIÓN DETALLADA DE LA INVENCION

La presente invención está dirigida a procedimientos para evaluar ácidos nucleicos como se caracterizan en las reivindicaciones adjuntas. Algunos aspectos de la presente invención se refieren a mejoras frente a Willey y Willey *et*

al. patentes de EE.UU. nº 5.043.390, 5.639.606 y 5.876.978.

I. Procedimientos para valorar un ácido nucleico

Un aspecto de la presente invención se refiere a procedimientos para valorar las cantidades de un ácido nucleico en una muestra. En algunas realizaciones, la invención permite la medida de pequeñas cantidades de ácido nucleico, por ejemplo, cuando el ácido nucleico se expresa en bajas cantidades en un espécimen, cuando permanecen intactas pequeñas cantidades del ácido nucleico y/o cuando se proporcionan pequeñas cantidades de un espécimen. Por ejemplo, en algunas realizaciones la práctica de la invención valora la expresión génica en muestras pequeñas de especímenes biológicos.

“Especímen”, como se usa en la presente memoria, puede hacer referencia al material recogido para análisis, por ejemplo, un frotis de cultivo, un pellizco de tejido, una extracción de biopsia, un vial de fluido corporal, por ejemplo, saliva, sangre y/u orina, etc., que se toma para investigación, diagnóstico u otros fines de cualquier entidad biológica. “Entidad biológica”, como se usa en la presente memoria, puede hacer referencia a cualquier entidad capaz de albergar un ácido nucleico incluyendo cualquier especie, por ejemplo un virus, una célula, un tejido, un cultivo *in vitro*, una planta, un animal y/o un sujeto que participa en un ensayo clínico. “Muestra”, como se usa en la presente memoria, hace referencia a material de espécimen usado para un ensayo, reacción, tanda, prueba y/o experimento dado. Por ejemplo, una muestra puede comprender una alícuota del material de espécimen recogido, hasta e incluyendo todo el espécimen. Como se usan en la presente memoria, los términos ensayo, reacción, tanda, prueba y/o experimento pueden usarse intercambiamente. Algunas realizaciones de la presente invención pueden practicarse usando pequeñas cantidades de partida de ácido nucleico que dan cantidades cuantificables.

En algunas realizaciones, el espécimen recogido puede comprender menos de aproximadamente 100.000 células, menos de aproximadamente 10.000 células, menos de aproximadamente 5.000 células, menos de aproximadamente 1.000 células, menos de aproximadamente 500 células, menos de aproximadamente 100 células, menos de aproximadamente 50 células o menos de aproximadamente 10 células. En algunas realizaciones, los procedimientos de la presente invención son capaces de valorar la cantidad de ácido nucleico presente en una muestra que comprende menos de aproximadamente 100.000 células. Por ejemplo, una muestra de biopsia puede comprender menos de aproximadamente 100.000 células. En algunas realizaciones, el procedimiento es capaz de valorar la cantidad de ácido nucleico en menos de aproximadamente 10.000 células, menos de aproximadamente 5.000 células, menos de aproximadamente 1.000 células, menos de aproximadamente 500 células, menos de aproximadamente 100 células, menos de aproximadamente 50 células o menos de aproximadamente 10 células. Especímen biológico pequeño puede hacer referencia también a las cantidades recogidas típicamente en biopsias, por ejemplo, biopsias endoscópicas (usando cepillo y/ pinzas), biopsias de aspiración con aguja (incluyendo biopsias de aspiración con aguja fina), así como cantidades proporcionadas en poblaciones celulares clasificadas (por ejemplo, poblaciones celulares clasificadas por flujo) y/o materiales microdisccionados (por ejemplo, tejidos microdisccionados capturados con láser). Por ejemplo, la biopsias por sospecha de lesiones cancerosas en pulmón, mama, próstata, tiroides y páncreas se realizan normalmente por biopsia de aspiración con aguja fina (FNA), la médula ósea se obtiene también por biopsia, y los tejidos de cerebro, embrión en desarrollo y modelos animales pueden obtenerse mediante muestras microdisccionadas capturadas con láser.

En algunas realizaciones, valoración, evaluación y/o medida de un ácido nucleico pueden hacer referencia a proporcionar una medida de la cantidad de ácido nucleico en un espécimen y/o muestra, por ejemplo, para determinar el nivel de expresión de un gen. En algunas realizaciones, proporcionar una medida de una cantidad hace referencia a detectar la presencia o ausencia del ácido nucleico de interés. En algunas realizaciones, proporcionar una medida de una cantidad puede hacer referencia a cuantificar la cantidad de ácido nucleico, por ejemplo, proporcionando una medida de la concentración o grado de la cantidad de ácido nucleico presente. En algunas realizaciones, proporcionar una medida de la cantidad de ácido nucleico hace referencia a dar cuenta de la cantidad de ácido nucleico, por ejemplo, indicando el número de moléculas de ácido nucleico presentes en una muestra. Puede hacerse referencia al ácido nucleico de interés como ácido nucleico diana, y puede hacerse referencia al gen de interés, por ejemplo el gen que se está evaluando, como gen diana.

En algunas realizaciones, los procedimientos de la presente invención son capaces de dar cuenta de menos de aproximadamente 1.000 moléculas de ácido nucleico en una muestra, por ejemplo, de aproximadamente 800, aproximadamente 600 o aproximadamente 400 moléculas de ácido nucleico. En algunas realizaciones, se puede dar cuenta en una muestra de menos de aproximadamente 100 moléculas, por ejemplo aproximadamente 60 moléculas, preferiblemente menos de aproximadamente 10 moléculas, por ejemplo aproximadamente 6 moléculas, o más preferiblemente menos de aproximadamente 1 molécula de un ácido nucleico. Por ejemplo, en realizaciones preferidas una sola molécula de molde de ácido nucleico puede dar lugar a un producto amplificado detectable. En algunas realizaciones, los procedimientos de la presente invención pueden medir menos de aproximadamente 10.000.000, menos de aproximadamente 5.000.000, menos de aproximadamente 1.000.000, menos de aproximadamente 500.000, menos de aproximadamente 100.000, menos de aproximadamente 50.000, menos de aproximadamente 10.000, menos de aproximadamente 8.000, menos de aproximadamente 6.000, menos de aproximadamente 5.000 o menos de aproximadamente 4.000 moléculas de ácido nucleico en una muestra. Puede hacerse referencia también al número de moléculas de un ácido nucleico como el número de copias del ácido nucleico encontradas en una muestra y/o espécimen.

La práctica de algunas realizaciones de la presente invención permite medir transcritos escasos con significancia estadística. Por ejemplo, en algunas realizaciones, puede determinarse el número de copias de un ácido nucleico correspondiente a un transcrito génico, por ejemplo, el número de copias/célula, cuando el gen se expresa con bajo número de copias. Dar cuenta de menos de aproximadamente 1.000 moléculas puede permitir la medida de menos de aproximadamente 10 copias/célula de al menos 100 transcritos génicos diferentes en un espécimen biológico pequeño, por ejemplo, a partir de la cantidad de material usada típicamente para obtener una medida génica, por ejemplo, para medir las pocas copias de ácido nucleico correspondientes a un gen. En algunas realizaciones, los procedimientos de la presente invención son capaces de medir y/o dar cuenta de menos de aproximadamente 10 copias/célula de al menos 100 transcritos génicos diferentes en un espécimen biológico pequeño a partir de la cantidad de material usada típicamente para obtener una medida génica. En algunas realizaciones, dar cuenta de menos de aproximadamente 10.000 moléculas puede permitir la medida de menos de aproximadamente 10 copias/célula de al menos 100 transcritos génicos diferentes en un espécimen biológico pequeño, por ejemplo, a partir de la cantidad de material usada típicamente para obtener una medida génica, por ejemplo, para medir las pocas copias de un ácido nucleico correspondientes a un gen.

En aún otras realizaciones, pueden obtenerse más medidas a partir de un espécimen y/o muestra dado, por ejemplo, del tamaño usado típicamente para medir las pocas copias de un ácido nucleico correspondientes a un gen. Por ejemplo, la práctica de algunas realizaciones de la invención dadas a conocer en la presente memoria puede medir y/o dar cuenta de menos de aproximadamente 100, menos de aproximadamente 50, menos de aproximadamente 20, menos de aproximadamente 10, menos de aproximadamente 8 o menos de aproximadamente 5 copias/célula de al menos aproximadamente 20, al menos aproximadamente 50, al menos aproximadamente 80, al menos aproximadamente 100, al menos aproximadamente 120, al menos aproximadamente 150 o al menos aproximadamente 200 ácidos nucleicos diferentes en una muestra, por ejemplo, correspondientes a diferentes transcritos génicos.

El material expresado puede ser endógeno para la entidad biológica, por ejemplo, transcritos de un gen expresado naturalmente en un tipo celular dado, o el material expresado para medir puede ser de naturaleza exógena. Por ejemplo, en algunas realizaciones pueden usarse los procedimientos de la presente invención para cuantificar los genes transfectados después de terapia génica y/o un gen informador en ensayos de transfección transitoria, por ejemplo, para determinar la eficacia de transfección (Morales, M. J., y Gottlieb, D. I., "A polymerase chain reaction-based method for detection and quantification of reporter gene expression in transient transfection assays", Analytical Biochemistry, 210, 188-194 (1993)).

Como se usa en la presente memoria, "ácido nucleico" puede hacer referencia a una forma polimérica de nucleótidos y/o moléculas de tipo nucleótido de cualquier longitud. En realizaciones preferidas, el ácido nucleico sirve como molde para la síntesis de un ácido nucleico complementario, por ejemplo, mediante la incorporación complementaria de bases de unidades nucleotídicas. Por ejemplo, un ácido nucleico puede comprender ADN de origen natural, por ejemplo ADN genómico, ARN, por ejemplo ARNm, y/o puede comprender una molécula sintética incluyendo, pero sin limitación, ADNc y moléculas recombinantes generadas de cualquier manera. Por ejemplo, el ácido nucleico puede generarse a partir de síntesis química, transcripción inversa, replicación de ADN o una combinación de estos procedimientos de generación. El ligamiento entre las subunidades puede proporcionarse por fosfatos, fosfonatos, fosforamidatos, fosfotioatos o similares, o por grupos no fosfato como son conocidos en la materia, tales como ligamientos de tipo peptídico utilizados en los ácidos nucleicos peptídicos (PNA). Los grupos de ligamiento pueden ser quirales o aquirales. Los polinucleótidos pueden tener cualquier estructura tridimensional, englobando moléculas monocatenarias, bicatenarias y de triple hélice que pueden ser, por ejemplo, moléculas de ADN, ARN o ADN/ARN híbridas. Una molécula de tipo nucleótido puede hacer referencia a un resto estructural que puede actuar sustancialmente como un nucleótido, por ejemplo, exhibiendo complementariedad de bases con una o más de las bases que aparecen en ADN o ARN y/o ser capaz de incorporación complementaria de bases. Los términos "polinucleótido", "molécula de polinucleótido", "molécula de ácido nucleico", "secuencia polinucleotídica" y "secuencia de ácido nucleico" pueden usarse intercambiamente con "ácido nucleico" en la presente memoria. En algunas realizaciones específicas, los ácidos nucleicos para medir pueden comprender una secuencia correspondiente a un gen referenciado en la Tabla 1 o 2 en la Figura 1 o 2, respectivamente.

En algunas realizaciones, el espécimen recogido comprende ARN para medir, por ejemplo, ARNm expresado en un cultivo de tejido. En algunas realizaciones, el espécimen recogido comprende ADN para medir, por ejemplo, ADNc transcrito de forma inversa a partir de transcritos. En algunas realizaciones, el ácido nucleico para medir se proporciona en una mezcla heterogénea de otras moléculas de ácido nucleico.

A. *Enfoque en dos etapas*

La presente invención proporciona un procedimiento de valoración del ácido nucleico proporcionado en una muestra, que comprende coamplificar el ácido nucleico, un número de otros(s) ácido(s) nucleico(s), un molde competitivo del ácido nucleico y un molde(s) del (de los) otro(s) ácido (s) nucleico(s), por ejemplo, produciendo un primer producto amplificado del mismo. El primer producto amplificado puede diluirse y coamplificarse además entonces, produciendo por ejemplo un segundo producto amplificado del mismo. Amplificar y amplificar entonces además el ácido nucleico y el molde competitivo del ácido nucleico puede considerarse como dos rondas de amplificación y puede hacerse referencia a un proceso que emplea dos rondas de amplificación como un proceso "en dos etapas" o

enfoque “en dos etapas”.

La Figura 3 ilustra esquemáticamente algunas realizaciones del proceso global “en dos etapas” descrito en la presente memoria, por ejemplo, cuando el ácido nucleico amplificado es ADNc. Los detalles experimentales que comparan un enfoque en dos etapas con un enfoque no en dos etapas pueden encontrarse en el ejemplo I siguiente.

5 En la etapa 301 de la Figura 3, por ejemplo, puede extraerse ARN de las células o tejidos del espécimen.

En la etapa 302 de la Figura 3, el ARN extraído puede transcribirse de forma inversa, proporcionando ADNc. En algunas realizaciones, el ácido nucleico amplificado es un ácido nucleico distinto de ADNc, como se describe anteriormente. En algunas realizaciones, aunque la eficacia de transcripción inversa puede ser variable, la representación de un ácido nucleico en comparación con otro en el producto de ADNc resultante puede no estar afectada. Es decir, en algunas realizaciones, la cantidad de ADNc del ácido nucleico diana en comparación con la cantidad de ADNc de un segundo ácido nucleico (por ejemplo, un segundo ácido nucleico que sirve como ácido nucleico de referencia) puede permanecer equivalente o sustancialmente equivalente a la cantidad de ARNm de ácido nucleico diana en comparación con la cantidad de ARNm del segundo ácido nucleico.

10 En el paso 303 de la Figura 3, se coamplifican ADNc nativo y su molde competitivo en una primera ronda de amplificación. El ADNc nativo puede comprender tanto el ácido nucleico diana como uno o más de los otros ácidos nucleicos, que pueden coamplificarse con un molde competitivo del ácido nucleico diana y un molde competitivo de uno o más de los otros ácidos nucleicos. Por ejemplo, el ADNc puede diluirse en serie y amplificarse entonces una o más diluciones en serie.

Los moldes competitivos de al menos dos ácidos nucleicos están a concentraciones conocidas uno respecto a otro. “Molde competitivo”, como se usa en la presente memoria, puede hacer referencia a un ácido nucleico que compite con un ácido nucleico diana durante una reacción de amplificación. Es decir, cuando está presente en una mezcla de reacción para amplificar el ácido nucleico diana, el molde competitivo compite por servir como molde para dicha amplificación. En algunas realizaciones, por ejemplo, el molde competitivo de un ácido nucleico dado tiene una estructura que permite su amplificación en la misma o sustancialmente la misma medida que el ácido nucleico dado. En realizaciones preferidas, puede amplificarse un molde competitivo de un ácido nucleico dado usando uno o más de los mismos cebadores que del ácido nucleico dado y/o se amplifica con la misma o sustancialmente la misma eficacia que el ácido nucleico dado. En realizaciones preferidas, se amplifica un molde competitivo de un ácido nucleico dado usando los mismos cebadores, comparte homología de secuencia y/o se amplifica con la misma o sustancialmente la misma eficacia que el ácido nucleico dado. En algunas realizaciones, se hace referencia a los moldes competitivos como patrones internos o como patrón interno molde competitivo.

El término “molde nativo”, como se usa en la presente memoria, puede hacer referencia a ácido nucleico obtenido directa o indirectamente de un espécimen que puede servir como molde para amplificación. Por ejemplo, puede hacer referencia a moléculas de ADNc correspondientes a un gen cuya expresión se va a medir, en que el ADNc se amplifica y cuantifica. En algunas realizaciones específicas, al menos un molde competitivo usado comprende una secuencia referenciada en la Tabla 4 de la Figura 4.

El término “cebador” hace referencia en general a un ácido nucleico capaz de actuar como punto de iniciación de la síntesis junto con una hebra complementaria cuando las condiciones son adecuadas para la síntesis de un producto de extensión de cebador. En algunas realizaciones específicas, al menos un cebador usado comprende una secuencia referenciada en la Tabla 4 de la Figura 4. También, la tabla de la Figura 4 muestra la secuencia de cebador y posición de varios genes cuya expresión puede medirse.

Preferiblemente, el molde competitivo tiene un rasgo distintivo del ácido nucleico diana, que permite distinguir a su producto amplificado del producto amplificado del ácido nucleico diana. Por ejemplo, el molde competitivo puede comprender mutantes de ácido nucleico para evaluar. Las mutaciones pueden ser mutaciones puntuales, inserciones, inversiones, deleciones o similares. Por ejemplo, en algunas realizaciones, un molde competitivo comprende al menos un nucleótido que es diferente del correspondiente nucleótido en el ácido nucleico para evaluar. En algunas realizaciones, el molde competitivo comprende al menos aproximadamente 2, al menos aproximadamente 5, al menos aproximadamente 10, al menos aproximadamente 15 o al menos aproximadamente 20 nucleótidos que son diferentes. Se proporcionan en algunas realizaciones deleciones, inserciones, inversiones, sustituciones y/u otras alteraciones más largas. Por ejemplo, pueden generarse moldes competitivos acortados artificialmente según el procedimiento descrito por Celi *et al.*, Nucleic Acids Res. 21: 1047 (1993).

En algunas realizaciones preferidas, el molde competitivo comprende una alteración que causa una pérdida y/o ganancia de uno o más sitios de escisión en el molde competitivo, en comparación con su correspondiente ácido nucleico. Por ejemplo, puede sustituirse una base en una secuencia de molde competitivo dando como resultado la ganancia y/o pérdida de un sitio de reconocimiento de endonucleasa de restricción, un sitio de escisión química u otro sitio de escisión específica. Pueden usarse diversos programas para identificar y ajustar 1 o 2 o más secuencias de desapareamiento de base para sitios de reconocimiento conocidos. Por ejemplo, puede usarse el programa Map del paquete de software Genetics Computer Group (Devereux *et al.*, *supra*, 1984). En este programa, se obtienen secuencias de ADNc para un ácido nucleico dado, y se evalúa entonces en la secuencia la presencia de uno o dos

desapareamientos de pares de bases para endonucleasas de restricción conocidas.

En algunas realizaciones, el molde competitivo comprende una alteración que causa una pérdida y/o ganancia de uno o más sitios de reconocimiento específicos en el molde competitivo en comparación con su correspondiente ácido nucleico. Por ejemplo, puede sustituirse una base en una secuencia de molde competitivo, dando como resultado la ganancia y/o pérdida de un sitio de unión de proteína tal como un sitio de unión de factor de transcripción. Resultarán evidentes para los especialistas en la materia otros cambios estructurales para distinguir el producto amplificado de un molde competitivo del producto amplificado de su correspondiente ácido nucleico.

La amplificación puede conseguirse mediante cualquier procedimiento conocido en la materia y/o dado a conocer en la presente memoria para amplificar moléculas de ácido nucleico. Cuando se usa amplificación por reacción en cadena de la polimerasa (PCR), las condiciones pueden incluir la presencia de ribonucleótidos y/o desoxirribonucleótidos, di-, tri- tetra- y penta-fosfatos y/o de orden superior; cebadores para amplificación por PCR para al menos un ácido nucleico y su correspondiente molde competitivo y al menos un agente inductor de la polimerización, tal como transcriptasa inversa, ARN polimerasa y/o ADN polimerasa. Los ejemplos de ADN polimerasas incluyen, pero sin limitación, ADN polimerasa de *E. coli*, Sequenase 2.0®, ADN polimerasa T4 o el fragmento Klenow de ADN polimerasa 1, T3, ARN polimerasa SP6, polimerasa de AMV, M-MLV y/o Vent, así como ThermoSequenase™ (Amersham) o Taquenase™ (ScienTech, St Louis, Mo.). Ejemplos adicionales incluyen polimerasas termostables aisladas de *Thermus aquaticus*, *Thermus thermophilus*, *Pyrococcus woesei*, *Pyrococcus furiosus*, *Thermococcus litoralis* y *Thermotoga maritima*. El agente inductor de la polimerización y los nucleótidos pueden estar presentes en un tampón adecuado, que puede incluir constituyentes que son cofactores o que afectan a condiciones tales como pH y similares a diversas temperaturas adecuadas. Los cebadores de PCR usados son preferiblemente moléculas nucleotídicas monocatenarias, pero pueden practicarse bicatenarias, tricatenarias y/o de orden superior con la presente invención. Como se usa en la presente memoria, "producto amplificado" puede hacer referencia a cualquier ácido nucleico sintetizado al menos en parte por incorporación complementaria de bases usando otro ácido nucleico como molde. Puede hacerse referencia también a un producto amplificado como un amplicón y/o amplicero en la presente memoria. La amplificación puede llevarse a cabo durante un número de ciclos de PCR, por ejemplo al menos aproximadamente 10, al menos aproximadamente 20, al menos aproximadamente 30, al menos aproximadamente 40 o al menos aproximadamente 50 ciclos en algunas realizaciones.

De acuerdo con la invención, se coamplifican más de un ácido nucleico (y su correspondiente molde competitivo). En realizaciones preferidas, el número de otros ácidos nucleicos es de al menos uno. En algunas realizaciones, el número es de al menos aproximadamente 50 otros ácidos nucleicos, al menos 100 otros ácidos nucleicos, al menos aproximadamente 200, al menos aproximadamente 300 otros ácidos nucleicos, al menos aproximadamente 500 otros ácidos nucleicos, al menos aproximadamente 800 otros ácidos nucleicos, al menos aproximadamente 1.000 otros ácidos nucleicos, al menos aproximadamente 5.000 otros ácidos nucleicos, al menos aproximadamente 10.000 otros ácidos nucleicos, al menos aproximadamente 50.000 otros ácidos nucleicos o al menos aproximadamente 100.000 otros ácidos nucleicos. Puede usarse un molde competitivo para cada ácido nucleico adicional para evaluar y, en realizaciones preferidas, pueden medirse simultáneamente una pluralidad de ácidos nucleicos en una muestra.

Al menos uno de los otros ácidos nucleicos puede servir como ácido nucleico de referencia. "Ácido nucleico de referencia", como se usa en la presente memoria, puede hacer referencia a un ácido nucleico que se amplifica así como al ácido nucleico para evaluar. El ácido nucleico puede "normalizarse" a un ácido nucleico de referencia. En algunas realizaciones, el ácido nucleico de referencia sirve como control de carga, por ejemplo para controlar el ADNc cargado en la reacción. Por ejemplo, en algunas realizaciones preferidas, el ácido nucleico de referencia comprende un ácido nucleico que no se espera que varíe (o que varíe significativamente) entre el espécimen biológico dado y/o en respuesta a ciertos estímulos. Por ejemplo, ARNm de un gen expresado constitutivamente puede proporcionar el ácido nucleico de referencia. En algunas realizaciones, pueden proporcionar el ácido nucleico de referencia genes constitutivos conocidos o potenciales incluyendo, pero sin limitación, gliceraldehído-3-fosfato deshidrogenasa (GAPD o GAPDH), β -actina, ARN de 28S, ARN de 18S y/u otros genes de proteína ribonuclear de ser humano, ratón y/o rata. Pueden usarse también otros genes constitutivos que se han usado como patrones internos en análisis Northern de expresión génica. Véanse, por ejemplo, Devereux *et al.*, Nucleic Acids Res. 12: 387 (1984); Barbu *et al.*, Nucleic Acids Res. 17: 7115 (1989). En algunas realizaciones, un molde competitivo de un ácido nucleico de referencia puede comprender un ácido nucleico que tiene una secuencia similar a cualquier hebra de ADNc de un gen constitutivo, pero que tiene un rasgo distintivo como se describe anteriormente.

Muchos genes diferentes pueden proporcionar ácidos nucleicos de referencia. La elección del ácido nucleico de referencia puede depender de los tejidos para ensayar y/o de los estados biológicos que se están estudiando. Por ejemplo, la β -actina varía poco entre muestras de células epiteliales bronquiales normales (véase, por ejemplo, Crawford, E. L., Khuder, S. A., Durham, S. J., *et al.* (2000). "Normal bronchial epithelial cell expression of glutathione transferase P1, glutathione transferase M3, and glutathione peroxidase is low in subjects with bronchogenic carcinoma" (Cancer Res. 60, 1609-1618), pero puede variar aproximadamente más de 100 veces en muestras de diferentes tejidos, tales como células epiteliales bronquiales en comparación con linfocitos.

En la etapa 304 de la Figura 3, se diluyen el producto amplificado de ADNc nativo y molde competitivo (obtenido en la ronda 1) antes de la amplificación adicional en la ronda 2. En algunas realizaciones, pueden diluirse el producto

amplificado del ácido nucleico diana y su correspondiente molde competitivo. En algunas realizaciones, pueden diluirse el producto amplificado de un ácido nucleico de referencia y su correspondiente molde competitivo. La dilución del producto amplificado puede conseguirse mediante cualquier técnica conocida en la materia y/o descrita en la presente memoria. Por ejemplo, la dilución puede implicar la retirada de una alícuota de una mezcla que comprende un primer producto amplificado, y la transferencia a un recipiente que contiene tampón adicional. En algunas realizaciones, la dilución produce una dilución de al menos aproximadamente 1.000.000 veces, una dilución de al menos aproximadamente 500.000 veces, una dilución de al menos aproximadamente 100.000 veces, una dilución de al menos aproximadamente 50.000 veces, una dilución de al menos aproximadamente 10.000 veces, una dilución de al menos aproximadamente 5.000 veces, una dilución de al menos aproximadamente 1.000 veces, una dilución de al menos aproximadamente 500 veces o una dilución de al menos aproximadamente 100 veces.

En la etapa 305 de la Figura 3, se amplifican además en la ronda 2 el producto amplificado diluido de ADNc nativo y molde competitivo (obtenido en la ronda 1). De acuerdo con la invención, se coamplifican adicionalmente el producto amplificado diluido de un ácido nucleico diana y su correspondiente molde competitivo en una segunda ronda de amplificación. En algunas realizaciones, pueden coamplificarse además el producto amplificado diluido de un ácido nucleico de referencia y su correspondiente molde competitivo en una segunda ronda de amplificación. Como se menciona anteriormente, puede hacerse referencia al uso de dos rondas como un enfoque “en dos etapas”. En algunas realizaciones, el ácido nucleico diana y/o el ácido nucleico de referencia pueden someterse a más de dos rondas de amplificación. Por ejemplo, pueden diluirse de nuevo y amplificarse además el segundo producto amplificado del ácido nucleico diana y su correspondiente molde competitivo y/o pueden diluirse de nuevo y amplificarse además el segundo producto amplificado del ácido nucleico de referencia y su correspondiente molde competitivo.

Pueden amplificarse en un recipiente dado diversos ácidos nucleicos y los correspondientes moldes competitivos durante la ronda 1 y/o ronda 2 de un proceso en dos etapas. Por ejemplo, en algunas realizaciones, se coamplifican en un recipiente dado más de un ácido nucleico (cada uno con su correspondiente molde competitivo). En algunas realizaciones, se llevan a cabo amplificaciones repetidas en un recipiente dado con menos ácidos nucleicos diferentes (cada uno con su correspondiente molde competitivo). Por ejemplo, en algunas realizaciones preferidas, los productos amplificados se amplifican además con cebadores de un ácido nucleico correspondiente a un gen. Por ejemplo, puede conseguirse la coamplificación del primer producto amplificado diluido de un ácido nucleico y del molde competitivo del ácido nucleico usando un par de cebadores para coamplificar el ácido nucleico particular y su correspondiente molde competitivo secado en el recipiente usado en la ronda 2. Por ejemplo, pueden tomarse alícuotas de cebadores de genes individuales en recipientes de reacción individuales y secarse, por ejemplo en placas de 384 pocillos. Pueden prepararse por adelantado múltiples placas cargadas con cebadores (por ejemplo, aproximadamente 10, aproximadamente 100, aproximadamente 500 placas). Por ejemplo, en algunas realizaciones, los cebadores preparados de este modo son estables a 4 °C durante meses.

En la etapa 306 de la Figura 3, pueden compararse cantidades de productos amplificados. De acuerdo con la invención, la cantidad de producto amplificado de un ácido nucleico diana se compara con la cantidad de producto amplificado de su molde competitivo, por ejemplo, la comparación implica obtener una relación, por ejemplo una primera relación que refleja las cantidades amplificadas de ácido nucleico diana en comparación con las cantidades amplificadas de su molde competitivo. En realizaciones preferidas, esta relación se proporciona como un cociente, por ejemplo un primer cociente de cantidad de producto amplificado de un ácido nucleico a cantidad de producto amplificado de su molde competitivo, por ejemplo, cuando se coamplifican el ácido nucleico y su molde competitivo.

La cantidad de producto amplificado de un ácido nucleico diana se compara entonces con un ácido nucleico de referencia. Se compara el ácido nucleico de referencia mismo con un molde competitivo del ácido nucleico de referencia. Concretamente, se compara la cantidad de producto amplificado del ácido nucleico de referencia con la cantidad de producto amplificado de su molde competitivo. Esta comparación implica obtener una relación, por ejemplo una segunda relación, que refleja la cantidad amplificada de ácido nucleico de referencia en comparación con la cantidad amplificada de su molde competitivo. En realizaciones preferidas, esta relación se proporciona como un cociente, por ejemplo un segundo cociente de cantidad de producto amplificado del ácido nucleico de referencia a cantidad de producto amplificado de su molde competitivo, por ejemplo, cuando se coamplifican el ácido nucleico de referencia y su molde competitivo.

De acuerdo con la invención, la comparación del ácido nucleico diana con un ácido nucleico de referencia implica comparar la primera y segunda relaciones descritas anteriormente. Por ejemplo, puede obtenerse una relación que refleja cómo la primera relación se compara con la segunda relación. En algunas realizaciones, esta relación compara la primer cociente con el segundo cociente, por ejemplo como cociente del primer y segundo cocientes.

Los adjetivos “primero”, “segundo”, “tercero” y demás, como se usan en la presente memoria, no indican necesariamente ningún orden de preferencia, importancia, cronología o grado de calidad, concentración y/o cantidad. En lugar de ello, los términos se usan para diferenciar los nombres calificados por los adjetivos, por ejemplo, un primer y segundo cocientes puede significar dos cocientes diferentes; un segundo ácido nucleico puede significar un ácido nucleico diferente al que se hace referencia como primer ácido nucleico.

En un proceso en dos etapas según la invención, se obtiene una primera relación comparando el segundo producto

amplificado del ácido nucleico diana con un segundo producto amplificado del molde competitivo del ácido nucleico diana; se obtiene una segunda relación comparando el primer producto amplificado del ácido nucleico de referencia con el primer producto amplificado del molde competitivo del ácido nucleico de referencia y se comparan la primera y segunda relaciones. En realizaciones más preferidas, la relación obtenida comparando la primera y segunda relaciones permanece sustancialmente constante más allá de la fase exponencial de amplificación del ácido nucleico. Sustancialmente constante puede hacer referencia a variaciones de \pm aproximadamente 1 %, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 15 % o aproximadamente 20 % de un número constante absoluto.

En algunas realizaciones, otro de los ácidos nucleicos amplificados puede servir como segundo ácido nucleico de referencia. En dichas realizaciones, medir la cantidad de ácido nucleico de referencia puede comprender obtener una tercera relación que compara el primer producto amplificado de este segundo ácido nucleico de referencia con el primer producto amplificado del molde competitivo del segundo ácido nucleico de referencia, y comparar la primera y tercera relaciones. También, en algunas realizaciones, los datos calculados usando un primer ácido nucleico de referencia pueden recalcularse respecto a los de otro ácido nucleico de referencia.

En algunas realizaciones, usar dos o más ácidos nucleicos de referencia puede proporcionar una comprensión de la variación interespecimen y/o internuestra entre los ácidos nucleicos de referencia. En algunas realizaciones, por ejemplo, pueden usarse β -actina y GAPD como primer y segundo ácido nucleicos de referencia. Por ejemplo, existe una correlación significativa entre el cociente de expresión de β -actina/GAPD y el tamaño celular (Willey, J. C., Crawford, E. L. y Jackson, C. M. (1998) "Expression measurement of many genes simultaneously by quantitative RT-PCR using standardized mixtures of competitive templates". *Am. J. Respir. Cell Mol. Biol.* 19, 6-17), pudiendo hacerse uso de estos 2 ácidos nucleicos de referencia preferidos en algunas realizaciones. En algunas realizaciones, pueden usarse como referencia cualquier ácido nucleico o combinación de ácidos nucleicos medidos, incluyendo todos los ácidos nucleicos medidos. El número de genes que debe cuantificarse para normalización de cualquiera de los ácidos nucleicos medidos para dar como resultado una normalización adecuada puede variar dependiendo de las muestras que se estén estudiando.

Como se menciona anteriormente, en algunas realizaciones un procedimiento en dos etapas puede comprender una amplificación en dos etapas del ácido nucleico que sirve como ácido nucleico de referencia. En algunas de dichas realizaciones, puede obtenerse una cuarta relación comparando el segundo producto amplificado del ácido nucleico de referencia con el segundo producto amplificado de su molde competitivo. En algunas realizaciones, se comparan la primera y cuarta relaciones, obteniendo por ejemplo un cociente del primer y cuarto cocientes. En aún otras realizaciones, cuando se amplifica en dos rondas el ácido nucleico que sirve como ácido nucleico de referencia, pueden usarse el primer producto amplificado del ácido nucleico diana y el primer producto amplificado de su molde competitivo para obtener la primera relación, por ejemplo el primer cociente.

Cuando el enfoque "en dos etapas" se extiende a más de dos rondas de amplificación, puede diluirse y amplificarse aún más más el segundo producto amplificado de un ácido nucleico y de un molde competitivo del ácido nucleico, por ejemplo, produciendo un tercer producto amplificado del mismo. Las etapas de dilución y amplificación adicional pueden repetirse al menos aproximadamente 1 vez, al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 5 veces, al menos aproximadamente 10 veces, al menos aproximadamente 20 veces, al menos aproximadamente 50 veces o al menos aproximadamente 100 veces o más.

En algunas realizaciones, comparar la primera y segunda y segunda y/o primera y tercera y/o primer y cuarta relaciones puede proporcionar un "cociente de cocientes" correspondiente a un valor numérico. En algunas realizaciones, los valores numéricos para diversos ácidos nucleicos medidos, por ejemplo para diversas medidas de expresión génica, se proporcionan como una base de datos, como se describe con más detalle a continuación. Por ejemplo, dicha base de datos puede usarse con datos de expresión génica en ensayos de diagnóstico clínico.

En algunas realizaciones, obtener las comparaciones, por ejemplo el primer, segundo, tercera y/o cuarto cocientes, implica medir las cantidades de producto amplificado de cada uno de ácido nucleico, molde competitivo de ácido nucleico, ácido(s) nucleico(s) de referencia y molde(s) competitivo(s) del ácido nucleico de referencia. Puede usarse cualquier procedimiento capaz de cuantificar ácidos nucleicos que tengan un rasgo distintivo (por ejemplo, que tengan diferentes tamaños y/o secuencias). Los procedimientos de cuantificación pueden implicar separar y/o aislar el producto amplificado, por ejemplo, mediante el uso de electroforesis, hibridación en fase sólida tal como matrices, espectrometría, cromatografía, HPLC y/u otros procedimientos conocidos en la materia para separar diferentes moléculas de ácido nucleico.

La electroforesis usada puede ser una o más de electroforesis en gel (por ejemplo, electroforesis en gel de agarosa y/o poliacrilamida), electroforesis capilar (por ejemplo, usando un dispositivo de electroforesis capilar como PE 310 o un dispositivo de CE de microfluidos CE como Agilent 2100 o el sistema de alto rendimiento Calipertech AMS 90) y/u otros tipos de dispositivos de electroforesis conocidos en la materia. Véanse, por ejemplo, (G. Gilliland, S. Perrin, K. Blanchard y H.F. Bunn, *Proc. Natl. Acad. Sci. USA* 87, 2725-2729 (1990); M.J. Apostolakos, W.H. Schuermann, M. W. Frampton *et al.*, *Analytical Biochemistry* 213, 277-284 (1993)). Además, la electroforesis capilar (CE), en particular la tecnología de CE de microfluidos, puede permitir la medida de ácido nucleico en volúmenes muy pequeños. Véanse, por ejemplo, T. S. Kanigan *et al.*, en "Advances in Nucleic Acid and Protein Analyses,

Manipulation, and Sequencing”, P. A. Limbach, J. C. Owicki, R. Raghavachari, W. Tan, Eds. Proc. SPIE 3926: 172, (2000). Otros dispositivos de electroforesis que pueden usarse incluyen, por ejemplo, Agilent o ABI 310. En algunas realizaciones, se mostró que para la separación del producto amplificado en gel de agarosa, PerkinElmer 310 CE (ABI Prism 310 Genetic Analyzer) y 2100 Bioanalyzer microfluidic CE (Agilent, Santa Clara, California, EE.UU.) proporcionan resultados estadísticamente similares y reproducibles. E.L. Crawford, L.A. Warner, D.A. Weaver y J.C. Willey, “Quantitative end-point RT-PCR expression measurement using the Agilent 2100 Bioanalyzer and standardized RT-PCR”, solicitud Agilent, 1-8 de sept. de 2001.

Cuando se van a separar productos amplificados por electroforesis, el tamaño de los moldes competitivos y/o ácido(s) nucleico(s) de referencia puede seleccionarse para diferir de los del ácido nucleico diana. Por ejemplo, en algunas realizaciones, el producto amplificado generado a partir del ácido nucleico de referencia y el ácido nucleico diana son de tamaños suficientemente diferentes para separarse por electroforesis. Además, en algunas realizaciones, el producto amplificado generado a partir del molde competitivo para un ácido nucleico dado y el ácido nucleico dado son de tamaños suficientemente diferentes para separarse por electroforesis.

En algunas realizaciones, se consigue la diferencia de tamaño usando un molde competitivo para un ácido nucleico dado que sea más largo o más corto que el ácido nucleico dado. En algunas realizaciones, este diferencial de tamaño puede conseguirse mediante digestión con endonucleasa de restricción del producto amplificado, en que el molde competitivo difiere de su correspondiente ácido nucleico por la adición o falta de un sitio de endonucleasa de restricción. Por ejemplo, en una realización específica, se prepararon moldes competitivos de GAPD que se separaban de GAPD nativo basándose en digestión con EcoRI o BamHI. Puede usarse también la separación basada en digestión con otra endonucleasa de restricción. Además, en algunas realizaciones, puede usarse el mismo sitio de reconocimiento tanto para el ácido nucleico de referencia como para el ácido nucleico para medir.

Además, en algunas realizaciones, la longitud del producto amplificado después de digestión con endonucleasa de restricción es un factor a considerar. Por ejemplo, en ciertas realizaciones, se prefieren diferencias de tamaño de ácido nucleico mayores para una separación adecuada en geles de agarosa, por ejemplo, diferencias preferiblemente de aproximadamente 40, aproximadamente 50, aproximadamente 80, aproximadamente 100 o aproximadamente 120 pares de bases.

Los productos separados pueden cuantificarse mediante cualquier procedimiento conocido en la materia y/o descrito en la presente memoria incluyendo, por ejemplo, el uso de sondas radiomarcadas, autorradiografía y, preferiblemente, por espectrofotometría y/o densitometría, por ejemplo densitometría de geles teñidos con bromuro de etidio. Otros procedimientos que pueden usarse para cuantificar el producto amplificado incluyen la cromatografía, por ejemplo, cromatografía líquida de alta resolución (HPLC), cromatografía de gases y/o espectrometría de masas, por ejemplo, desorción/ionización mediante láser asistida por matriz-tiempo de vuelo-espectrometría de masas (MALDI-TOF-MS) (“Una previsión económica del mercado de expresión génica” <http://www.researchandmarkets.com/reports/5545>).

En algunas realizaciones, se miden los productos amplificados usando hibridaciones en fase sólida. Algunas realizaciones, por ejemplo, comprenden el uso de una matriz, incluyendo microperlas y/o micromatrices. Las matrices incluyen, por ejemplo, matrices de oligonucleótidos, incluyendo matrices de ADNc, ADN y/o ARN. Dichas matrices pueden comprender una macromatriz, una micromatriz (por ejemplo, una matriz de microfluidos) y/o una nanomatriz. En algunas realizaciones, pueden marcarse el producto amplificado y/o el oligonucleótido que hibrida con el mismo, por ejemplo, con un resto detectable. Por ejemplo, pueden marcarse uno o más de los nucleótidos de la reacción de amplificación con un resto detectable. Los restos detectables que pueden usarse incluyen restos fluorescentes, restos radiactivos, puntos cuánticos y/o sistemas luminiscentes.

En algunas realizaciones, las matrices para uso en la práctica de la presente invención comprenden oligonucleótidos inmovilizados sobre un soporte sólido, en que puede unirse un primer conjunto de oligonucleótidos inmovilizados a una secuencia del producto amplificado del ácido nucleico que no es común al producto amplificado del molde competitivo del ácido nucleico, y en que puede unirse un segundo conjunto de oligonucleótidos inmovilizados a una secuencia del producto amplificado del molde competitivo del ácido nucleico que no es común al producto amplificado del ácido nucleico, por ejemplo, secuencias que extienden la conexión entre el extremo 5' del molde competitivo y el extremo 3' desalineado truncado del molde competitivo (por ejemplo, que puede prepararse según el procedimiento de Celi). El producto amplificado del ácido nucleico y del molde competitivo del ácido nucleico puede permitirse unir a la matriz y obtenerse un cociente a partir de los dos conjuntos. En aún otras realizaciones, el enfoque en dos etapas puede practicarse sin el uso de hibridaciones en fase sólida, por ejemplo, sin el uso de matrices.

El uso de dos rondas en realizaciones preferidas de un proceso en dos etapas puede reducir la cantidad umbral de ácido nucleico que puede medirse en una muestra. El umbral inferior de detección puede definirse como la cantidad mínima de analito que puede detectarse fiablemente por encima del fondo. El límite de detección puede definirse como la concentración o cantidad de analito más baja que puede detectarse con una razonable certidumbre. Sin limitarse a una hipótesis y/o teoría particular, puede haber una cantidad mínima de ADNc que puede usarse para conseguir una medida estadísticamente significativa. El umbral inferior de detección en medidas de expresión génica puede considerarse en términos del número mínimo de moléculas de ADNc en una reacción para amplificación o el

número mínimo de células.

La Figura 5 ilustra esquemáticamente cómo la cantidad de ADNc usada en una reacción PCR tiene una relación directa con el número de copias de transcritos de ARN/célula que pueden medirse para un número dado de células usadas. El número mínimo de células depende entonces de las copias de ARNm/célula en una muestra, así como de la eficacia de la extracción de ARN y/o transcripción inversa. Por ejemplo, se considera el número de células para proporcionar ARN suficiente para dar como resultado al menos 10 moléculas de ADNc para un gen particular. Se supone en general que la extracción de ARN es cercana a aproximadamente 100 %, mientras que la transcripción inversa es aproximadamente un 10 % eficaz. Por tanto, si se estudia una población homogénea de células y cada célula contiene 10 copias de ARNm para un gen, permanecerá 1 copia por célula después de la transcripción inversa. Debido a consideraciones estequiométricas, las muestras de ADNc incluidas en una reacción de PCR que contiene menos de aproximadamente 10 moléculas de transcrito son cuestionables, en algunos tipos de PCR. En dichas realizaciones, está preferiblemente presente ADNc que representa aproximadamente 10 células en la reacción PCR, como se ilustra en la Figura 5. Si se estudia una población celular heterogénea en que 1 célula de cada 10 expresa un transcrito particular, está preferiblemente presente ADNc que representa aproximadamente 1.000 células en la reacción PCR.

En ciertas realizaciones, el uso de dos rondas puede superar algunas de las limitaciones ilustradas en la Figura 5. Se considera típica una muestra de ADNc de aproximadamente 10 µl que representa aproximadamente 1.000 células y comprende aproximadamente 6×10^5 moléculas de ácido nucleico de β-actina. Los genes expresados a nivel medio (100 veces menos que β-actina) se representan por aproximadamente 6.000 moléculas en la muestra. Un número de genes que pueden ser funcionalmente importantes se expresan 10.000 veces menos que la β-actina y, para dichos genes, habría aproximadamente 60 moléculas representadas en la muestra. En una muestra 100 veces menor de aproximadamente 100 nl, los genes expresados 10.000 veces menos que la β-actina estarían representados por aproximadamente 0,6 copias o menos.

En ciertas realizaciones de la presente invención, pueden usarse aproximadamente 10 nl de un producto amplificado en la ronda 1 de aproximadamente 10 µl en aproximadamente dos volúmenes de reacción de aproximadamente 100 nl. Debido a que se consigue rutinariamente una amplificación de más de aproximadamente 1.000.000 veces en la reacción de la ronda 1, aproximadamente 10 nl de aproximadamente 10 µl de reacción de la ronda 1 contendrán un amplio producto amplificado de ácido nucleico y molde competitivo para medir con confianza estadística después de la ronda 2.

Además, en algunas realizaciones preferidas, el uso de dos rondas puede aumentar el número de medidas obtenibles a partir de una muestra pequeña de ácido nucleico. Por ejemplo, en algunas realizaciones, pueden obtenerse al menos aproximadamente 10.000, al menos aproximadamente 50.000, al menos aproximadamente 80.000, al menos aproximadamente 100.000, al menos aproximadamente 150.000 medidas de ácido nucleico a partir de la misma cantidad de ácido nucleico de partida usada típicamente para obtener una medida usando los procesos proporcionados en Willey y Willey *et al.* '390, '606 y '978. En algunas realizaciones, pueden obtenerse al menos aproximadamente 200.000, al menos aproximadamente 500.000, al menos aproximadamente 800.000, al menos aproximadamente 1.000.000 o al menos aproximadamente 1.500.000 medidas de ácido nucleico a partir de la misma cantidad de ácido nucleico de partida usada típicamente para obtener una medida usando los procesos proporcionados en Willey y Willey *et al.* '390, '606 y '978, preferiblemente sin pérdida de sensibilidad para detectar transcritos escasos. Por ejemplo, en algunas realizaciones, puede generarse suficiente producto amplificado para medir los ácidos nucleicos correspondientes a varios genes en aproximadamente 100 a aproximadamente 1.000 muestras celulares. Usando los procesos proporcionados en Willey y Willey *et al.* '390, '606, y '978, se usa típicamente ADNc que representa de aproximadamente 100 a aproximadamente 1.000 células para medir un ácido nucleico en una reacción PCR. Haciendo referencia de nuevo a la Figura 5, usar esta cantidad permite la detección de transcritos que se expresan a aproximadamente 0,1 a aproximadamente 1 copia por célula (o aproximadamente 1 a aproximadamente 10 copias por 10 células) con significación estadística. Puede usarse la misma cantidad de ADNc en una primera ronda de amplificación en ciertas realizaciones de la presente invención. Puesto que este ADNc se coamplifica con un molde competitivo del ácido nucleico para medir, y puesto que la relación de ADNc endógeno a su molde competitivo permanece constante o sustancialmente constante, el producto amplificado de la ronda 1 puede diluirse y amplificarse además en una segunda ronda con cebadores específicos de un ácido nucleico dado sin cambiar significativamente las cantidades relativas de producto amplificado.

Además, en algunas realizaciones, el uso de dos rondas puede aumentar el número de ácidos nucleicos que pueden medirse en una muestra dada. Algunas realizaciones, por ejemplo, permiten duplicar la medida de muchos genes en pequeñas cantidades de material de espécimen.

En algunas realizaciones, los procedimientos de la presente invención reducen los falsos positivos a un número estadísticamente no significativo, y en algunas realizaciones, se eliminan. Por ejemplo, cuando se usa un molde competitivo en un número de medidas de ácido nucleico, puede no haber falsos negativos y un número estadísticamente no significativo de falsos positivos.

B. *Uso de una mezcla estandarizada*

- De acuerdo con la invención, el enfoque en dos etapas de valoración de un ácido nucleico en una muestra comprende el uso de una mezcla estandarizada. "Mezcla estandarizada", como se usa en la presente memoria, puede hacer referencia a una mezcla que comprende un número de patrones internos, por ejemplo, un número de moldes competitivos, a concentraciones conocidas. La mezcla estandarizada comprende un molde competitivo de al menos un ácido nucleico diana y un molde competitivo de al menos un ácido nucleico de referencia, en que los moldes competitivos están a concentraciones conocidas uno respecto a otro. En realizaciones más preferidas, los moldes competitivos están a concentraciones fijas respecto a otros, hasta incluir todos los otros moldes competitivos en la mezcla.
- La Figura 6 ilustra una mezcla estandarizada usada en algunas realizaciones de la presente invención. El rasgo 601 ilustra una muestra, la muestra A, que comprende un número de ácidos nucleicos para medir, correspondiente a los genes 1-6-n, así como un ácido nucleico que sirve como referencia, correspondiente a β -actina en esta ilustración.
- El rasgo 602 ilustra una mezcla estandarizada de patrones internos que comprende moldes competitivos del ácido nucleico de referencia (patrón de β -actina) así como moldes competitivos de ácidos nucleicos diana (patrones de genes 1 a 6-n). En algunas realizaciones, el número de molde(s) competitivo(s) puede ser de al menos 1 otro molde competitivo además de un ácido nucleico diana, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 500, al menos aproximadamente 1.000, al menos aproximadamente 5.000, al menos aproximadamente 10.000, al menos aproximadamente 50.000 o al menos aproximadamente 100.000 otros moldes competitivos. Por ejemplo, los moldes competitivos de varios genes para medir pueden incluirse en una mezcla estandarizada dada, como se ilustra en el rasgo 602.
- El rasgo 603 (flechas verticales bidireccionales) ilustra la relación entre los patrones internos en una mezcla estandarizada. Un molde competitivo de cada uno de un número de genes puede estar a una concentración fija respecto a otros moldes competitivos en una mezcla estandarizada. En consecuencia, en algunas realizaciones, cuando se combina una muestra de ADNc con una mezcla estandarizada, se fija la concentración de cada molde competitivo respecto al ADNc que representa su correspondiente gen.
- El rasgo 604 (flechas horizontales bidireccionales) ilustra la relación entre un patrón interno y su correspondiente ADNc de una muestra y cómo cada ácido nucleico se mide respecto a su molde competitivo respectivo en la mezcla estandarizada. Debido a que el molde competitivo de cada uno de estos ácidos nucleicos está presente a una concentración fija respecto a otros moldes competitivos, las mezclas estandarizadas pueden permitir valorar un ácido nucleico diana respecto a otros ácidos nucleicos que se están midiendo con la mezcla estandarizada 602. Por ejemplo, la muestra A 601 puede combinarse con la mezcla estandarizada 602, formando por ejemplo una mezcla maestra usada para coamplificaciones adicionales. Por ejemplo, la mezcla maestra puede usarse para coamplificar ácido nucleico correspondiente al gen 1 y su molde competitivo (patrón de gen 1), así como para coamplificar ácido nucleico correspondiente al gen 2 y su molde competitivo (patrón de gen 2).
- En un enfoque en dos etapas que usa mezcla estandarizada 602, pueden coamplificarse un ácido nucleico diana y su molde competitivo respectivo, produciendo el primer producto amplificado del mismo. Los productos amplificados pueden diluirse y coamplificarse además una o más veces, como se describe con más detalle anteriormente. En algunas realizaciones, el primer producto amplificado del ácido nucleico de referencia puede diluirse y amplificarse además una o más veces, también como se describe anteriormente.
- El rasgo 606 ilustra un número de otras muestras, muestras B_{1-n} 605, que comprenden también ácidos nucleicos, correspondientes a los genes 1 a 6-n, y un ácido nucleico de referencia correspondiente a β -actina. En algunas realizaciones, el número de moléculas de ARNm de β -actina obtenido a partir de una célula puede variar de aproximadamente 100 a aproximadamente 1.000, dependiendo por ejemplo de la eficacia en la extracción de ARN, el tamaño y/o las características de la célula.
- En algunas realizaciones, otro ácido nucleico puede servir como segundo ácido nucleico de referencia. Por ejemplo, en algunas realizaciones puede recalcularse la expresión génica medida con referencia a ARNm de β -actina respecto a la de otro ácido nucleico de referencia, si se desea. Por ejemplo, si otro ácido nucleico, por ejemplo GAPDH o cualquier otro de los genes 1 a 6-n 602, parece variar menos que la β -actina en las muestras B_{1-n} 605, los datos pueden recalcularse ("normalizarse") a esa referencia sin alterar la medida de expresión relativa, por ejemplo, la medida de expresión relativa en una muestra. Cuando se recalculan los datos de medida de ácido nucleico, las cantidades relativas medidas entre ácidos nucleicos pueden permanecer iguales o sustancialmente iguales.
- La Figura 7 ilustra un recálculo usando ciclofilina como segundo gen de referencia, en que la expresión génica se proporciona como un cociente de (moléculas de NT de gen diana)/(10⁶ moléculas de NT de β -actina). En la Figura 7, NT hace referencia a molde nativo, y el gen diana es c-myc.
- El cociente 701 ilustra el valor de expresión génica para el gen diana como cociente de (moléculas NT de c-myc)/(10⁶ moléculas NT de β -actina). El cociente 702 ilustra el valor de expresión génica de un primer gen de referencia como el cociente de (moléculas de NT de ciclofilina)/(10⁶ moléculas de NT de β -actina). El cociente 703 ilustra el factor de conversión para recalcularse respecto a ciclofilina. El cociente 703 proporciona la inversa del

cociente 702, a saber $(10^6 \text{ moléculas de NT de } \beta\text{-actina})/(\text{moléculas de NT de ciclofilina})$. La conversión puede conseguirse multiplicando el cociente 701 por el cociente 703, proporcionando el cociente 704. El cociente 704 ilustra el cociente $(\text{moléculas de NT de c-myc})/(\text{moléculas de NT de ciclofilina})$, un valor de expresión génica del gen diana respecto al nuevo gen de referencia.

- 5 En otras realizaciones, puede conseguirse la conversión de $(\text{moléculas de ácido nucleico diana})/(\text{moléculas de un primer ácido nucleico de referencia})$ a $(\text{moléculas de ácido nucleico diana})/(\text{moléculas de un segundo ácido nucleico de referencia})$, por ejemplo, invirtiendo el valor de expresión génica de la segunda referencia, por ejemplo, a $(\text{moléculas del primer ácido nucleico de referencia})/(\text{moléculas del segundo gen de referencia})$ y multiplicando este factor por los datos. El valor de moléculas del primer ácido nucleico de referencia puede cancelarse, dejando el
10 segundo gen de referencia en el denominador.

El recálculo puede lograrse usando una hoja de cálculo en algunas realizaciones. En algunos casos, el recálculo respecto a una nueva referencia puede alterar el valor numérico de la cantidad medida de ácido nucleico dado sin alterar los valores numéricos de los ácidos nucleicos entre sí. Sin limitarse a una hipótesis y/o teoría particular, esto puede explicarse porque puede decirse que se ligan cantidades medidas de ácido nucleico a través del uso de una
15 mezcla común estandarizada de moldes competitivos 602. Por tanto, el cociente entre dos ácidos nucleicos en una muestra sería el mismo o sustancialmente el mismo usando β -actina, ciclofilina o una combinación de ácidos nucleicos como ácido nucleico de referencia.

El rasgo 605 (flechas bidireccionales) ilustra cómo puede medirse cada uno de estos ácidos nucleicos en muestras adicionales respecto a su molde competitivo respectivo en la mezcla estandarizada 602. Como con la muestra A 601, cada uno de estos ácidos nucleicos puede valorarse respecto a otros ácidos nucleicos medidos con la mezcla estandarizada 602. Además, es posible comparar los datos de análisis de la muestra A 601 con los datos de análisis de las muestras B_{1-n} 604. Por ejemplo, debido a que el número de moléculas para cada molde competitivo es conocido dentro de la mezcla estandarizada, es posible calcular todos los datos en forma de moléculas/moléculas de ácido nucleico de referencia. En algunas realizaciones, la mezcla estandarizada 602 comprende suficientes
20 cantidades de moldes competitivos para valorar uno o más de los ácidos nucleicos diana en un gran número de muestras B_{1-n} 604, por ejemplo, en más de aproximadamente 10^4 muestras, en más de aproximadamente 10^5 muestras, en más de aproximadamente 10^6 muestras, en más de aproximadamente 10^7 muestras, en más de aproximadamente 10^8 muestras, en más de aproximadamente 10^9 muestras, en más de aproximadamente 10^{10} muestras, en más de aproximadamente 10^{11} muestras, en más de aproximadamente 10^{12} muestras, en más de aproximadamente 10^{13} muestras, en más de aproximadamente 10^{14} muestras o en más de aproximadamente 10^{15} muestras. En algunas realizaciones preferidas, el uso de una mezcla estandarizada común para múltiples muestras puede reducir el tiempo para obtener medidas de ácido nucleico. Por ejemplo, la re Preparación de reactivos para reacciones PCR puede consumir tiempo y puede conducir también a fuentes de error.

Pueden coamplificarse un ácido nucleico y su molde competitivo (y/o coamplificarse además) en el mismo o diferentes recipientes que uno o más otros ácidos nucleicos y su correspondiente molde competitivo. Véanse, por ejemplo, Apostolakos, M. J., Schuermann, W. H., Frampton, M. W., Utell, M. J. y Willey, J. C. (1993) "Measurement of gene expression by multiplex competitive polymerase chain reaction". Anal. Biochem. 213, 277-284; Willey, J. C., Crawford, E. L. y Jackson, C. M. (1998) "Expression measurement of many genes simultaneously by quantitative RT-PCR using standardized mixtures of competitive templates". Am. J. Respir. Cell Mol. Biol. 19, 6-17. El recipiente usado puede ser cualquier objeto capaz de permitir que exista una mezcla de reacción dentro de él y/o sobre él. Por ejemplo, el recipiente puede comprender un pocillo, tubo, depósito y/o canal de nano- y/o microfluidos, capilar, surco, superficie y/u otro envase.

En algunas realizaciones preferidas, el uso de una mezcla estandarizada 602 permite comparar directamente diferentes ácidos nucleicos amplificados en recipientes separados. En algunas realizaciones, por ejemplo, se coamplifican en un recipiente un ácido nucleico y su molde competitivo, mientras que se coamplifican en un recipiente diferente otro ácido nucleico y su molde competitivo. En cualquier caso, como ilustra el rasgo 603, el ácido nucleico puede medirse respecto a su molde competitivo patrón interno respectivo en la mezcla estandarizada y el otro ácido nucleico puede servir como ácido nucleico de referencia. Es decir, en realizaciones preferidas, el uso de una mezcla estandarizada permite la concentración de patrón interno de un ácido nucleico respecto a otras que permanecen fijas a lo largo de diferentes medidas.

Como ilustra el rasgo 603, el uso de una mezcla estandarizada común permite hacer comparaciones directas entre muestras B_{1-n} 604. Las diferentes muestras pueden amplificarse en diferentes momentos, por ejemplo, en diferentes días, en el mismo o diferentes experimentos en el mismo laboratorio y/o en diferentes experimentos en diferentes laboratorios. Crawford, E. L., Peters, G. J., Noordhuis, P., et al. (2001) "Reproducible gene expression measurement among multiple laboratories obtained in a blinded study using standardized RT (StaRT)-PCR". Mol. Diagn. 6, 217-225; Crawford, E. L., Warner, K. A., Khuder, S. A., et al. (2002) "Multiplex standardized Ki-PCR for expression analysis of many genes in small samples". Biochem. Biophys. Res. Commun. 293, 509-516; Crawford, E. L., Khuder, S. A., Durham, S. J., et al. (2000) "Normal bronchial epithelial cell expression of glutathione transferase P1, glutathione transferase M3, and glutathione peroxidase is low in subjects with bronchogenic carcinoma". Cancer Res. 60, 1609-1618; DeMuth, J. P., Jackson, C. M., Weaver, D. A., et al. (1998) "The gene expression index c-myc x E2F1/p21 is highly predictive of malignant phenotype in human bronchial epithelial cells". Am. J. Respir. Cell. Mol.

Biol. 19, 18-24; Mollerup, S., Ryberg, D., Hewer, A., Phillips, D. H. y Haugen, A. (1999) "Sex differences in lung CYP1A1 expression and DNA adduct levels among lung cancer patients". Cancer Res. 59, 3317-3320; Rots, M. G., Willey, J. C, Jansen, G., *et al.* (2000) "mRNA expression levels of methotrexate resistance-related proteins in childhood leukemia as determined by a standardized competitive template-based RT-PCR method". Leukemia 14, 2166-2175; Rots, M. G., Pieters, R., Peters, G. J., *et al.* (1999) "Circumvention of methotrexate resistance in childhood leukemia subtypes by rationally designed antifolates". Blood 94, 3121-3128; Allen, J. T., Knight, R. A., Bloor, C. A. y Spiteri, M. A. (1999) "Enhanced insulin-like growth factor binding protein-related protein 2 (connective tissue growth factor) expression in patients with idiopathic pulmonary fibrosis and pulmonary sarcoidosis". Am. J. Respir. Cell. Mol. Biol. 21, 693-700; Loitsch, S. M., Kippenberger, S., Dauletbaev, N., Wagner, T. O. y Bargon, J. (1999) "Reverse transcription-competitive multiplex PCR improves quantification of mRNA in clinical samples-application to the low abundance CFTR mRNA". Clin. Chem. 45, 619-624; Vondracek, M. T., Weaver, D. A., Sarang, Z., *et al.* (2002) "Transcript profiling of enzymes involved in detoxification of xenobiotics and reactive oxygen in human normal and Simian virus 40 T antigen-immortalized oral keratinocytes". In. J. Cancer 99, 776-782. En realizaciones preferidas, se hacen medidas usando la misma mezcla estandarizada y dilución del patrón interno de moldes competitivos.

Además, en algunas realizaciones, las medidas obtenidas usando diversos enfoques de cuantificación son directamente comparables cuando se usa una mezcla estandarizada común. Por ejemplo, se obtuvieron resultados estadísticamente similares usando una mezcla estandarizada común y cuantificando el producto amplificado mediante diversos tipos de electroforesis, o mediante separación electroforética con un Caliper AMS 90 SE30 o hibridando con micromatrices de acuerdo con algunas realizaciones de la presente invención. En otro ejemplo, se obtuvieron medidas de expresión génica reproducibles cuando se cuantificó el producto amplificado usando MALDI-TOF MS en lugar de usar electroforesis. Ding C. y Cantor, C. R. (2003) "A high-throughput gene expression analysis technique using competitive PCR and matrix-assisted laser desorption ionization time-of-flight MS". Proc. Natl. Acad. Sci. USA 100, 3059-3064.

El uso de mezclas estandarizadas puede aplicarse también a otros procedimientos para medir ácidos nucleicos, por ejemplo, en RT-PCR a tiempo real. Por ejemplo, en algunas realizaciones, obtener un cociente de producto amplificado de un ácido nucleico a producto amplificado de un molde competitivo del ácido nucleico puede comprender el uso de análisis de RT-PCR a tiempo real. Como otro ejemplo, puede usarse una mezcla estandarizada de acuerdo con algunas realizaciones de la presente invención en combinación con técnicas de molde competitivo descritas, por ejemplo, en Siebert, P. D., *et al.*, Nature 359: 557-558 (1992); Siebert, P. D., *et al.*, BioTechniques 14: 244-249 (1993) y Clontech Brochure, 1993, "Reverse Transcriptase-PCR (RT-PCR)". Por ejemplo, pueden desarrollarse sondas fluorescentes para uso de una mezcla estandarizada con RT-PCR a tiempo real.

C. *Uso de mezclas estandarizadas diluidas en serie*

En algunas realizaciones, se usa una serie de mezclas estandarizadas diluidas en serie para valorar las cantidades de ácido nucleico. "Mezclas estandarizadas diluidas en serie" puede hacer referencia a dos o más mezclas estandarizadas en que uno o más de los reactivos de las mezclas estandarizadas está diluido en serie. En algunas realizaciones, uno o más reactivos en las mezclas estandarizadas está diluido en serie respecto a uno diferente o más de los reactivos en las mezclas. Por ejemplo, en realizaciones preferidas, se diluye en serie un molde competitivo de un primer ácido nucleico respecto a un molde competitivo de un segundo ácido nucleico cuando el segundo ácido nucleico puede actuar como ácido nucleico de referencia. En algunas realizaciones, el ácido nucleico de referencia puede estar presente a dos concentraciones diferentes en dos de las mezclas estandarizadas diluidas en serie. Se hace referencia también en la presente memoria a una de las series de mezclas diluidas en serie como "mezcla".

La Figura 8 ilustra el uso de una serie de mezclas estandarizadas según algunas realizaciones de la presente invención. En la figura, "SMIS" hace referencia a una mezcla estandarizada de patrones estándar, preparada de acuerdo con realizaciones de la presente invención.

El rasgo 801 ilustra una muestra, la muestra A, que comprende un número de ácidos nucleicos para medir, correspondientes a los genes 1-12, así como un ácido nucleico que sirve como referencia, correspondiente a β -actina en esta ilustración.

El rasgo 802 ilustra una serie de 6 mezclas estandarizadas, las mezclas A-F, que comprenden diluciones de 10 veces de moldes competitivos de diferentes genes respecto a los moldes competitivos de un gen de referencia, β -actina en esta ilustración.

El rasgo 803 ilustra la relación entre moldes competitivos del ácido nucleico de referencia (patrón de β -actina) en comparación con moldes competitivos de ácidos nucleicos diana (patrones de genes 1 a 12) en las diferentes mezclas diluidas en serie. El uso de las series puede permitir la medida de los ácidos nucleicos correspondientes a diferentes genes expresados dentro de un intervalo, por ejemplo un intervalo de más de 6 órdenes de magnitud.

El rasgo 804 (flechas bidireccionales) ilustra cómo estos diferentes ácidos nucleicos de la muestra 801 están en

equilibrio (concretamente calibrados) con diferentes concentraciones de sus correspondientes moldes competitivos en las diferentes mezclas. “Equilibrar” o estar en equilibrio con, como se usa en la presente memoria, puede hacer referencia a calibrar las cantidades de dos ácidos nucleicos. Por ejemplo, los genes 9 y 10 de la muestra A 801, expresados a bajo nivel, están en equilibrio con la mezcla E que comprende 600 moléculas/μl de molde competitivo del gen 9 y el gen 10. Los genes 9 y 10 se miden preferiblemente usando la mezcla E. Los genes 6 y 7 se expresan a mayor nivel en la muestra A 801 y están en equilibrio con la mezcla C y la mezcla D, respectivamente. El gen 6 se mide preferiblemente usando la mezcla C y el gen 7 se mide preferiblemente con la mezcla D.

En algunas realizaciones, el uso de una serie permite la medida de ácidos nucleicos dentro de un intervalo de concentraciones. Cuando la práctica de la invención valora la expresión génica, como en la Figura 8, algunas realizaciones permiten la medida de la expresión génica dentro de uno o más órdenes de magnitud. Por ejemplo, en algunas realizaciones, las cantidades de dos ácidos nucleicos para medir varían dentro de un intervalo de menos de aproximadamente un orden de magnitud, más de aproximadamente un orden de magnitud o más de aproximadamente 2 órdenes de magnitud. En algunas realizaciones, las cantidades de dos ácidos nucleicos diferentes para medir, por ejemplo, los niveles de ARNm expresados a partir de dos o más genes diferentes, varían dentro de un intervalo de aproximadamente 3 o más órdenes de magnitud, de aproximadamente 4 o más órdenes de magnitud, de aproximadamente 5 o más órdenes de magnitud, de aproximadamente 6 o más órdenes de magnitud o de aproximadamente 7 o más órdenes de magnitud, por ejemplo, extendiéndose un intervalo de aproximadamente 7 unidades logarítmicas de expresión génica que incluyen aproximadamente 10^{-3} , aproximadamente 10^{-2} , aproximadamente 0,1, aproximadamente 1, aproximadamente 10, aproximadamente 10^2 , aproximadamente 10^3 y aproximadamente 10^4 copias/célula. En algunas realizaciones, la cantidades de dos ácidos nucleicos diferentes para medir varían dentro de un intervalo de aproximadamente 8 o más, aproximadamente 9 o más o aproximadamente 10 o más órdenes de magnitud, por ejemplo, extendiéndose un intervalo de aproximadamente 10 unidades logarítmicas de expresión génica de aproximadamente 10^{-3} , aproximadamente 10^{-2} , aproximadamente 0,1, aproximadamente 1, aproximadamente 10, aproximadamente 10^2 , aproximadamente 10^3 , aproximadamente 10^4 , aproximadamente 10^5 o aproximadamente 10^6 copias/célula. Dichos intervalos de expresión génica pueden ser importantes en la detección de agentes biológicos de guerra biológica, por ejemplo.

El rasgo 805 ilustra una muestra diferente, la muestra B, que comprende también ácidos nucleicos correspondientes a los genes 1-2 y a β-actina.

El rasgo 806 (flechas bidimensionales) ilustra que los diferentes ácidos nucleicos en la muestra B 805 están también en equilibrio con diferentes concentraciones de sus correspondientes moldes competitivos en las diferentes mezclas. Un gen dado en una muestra diferente puede estar en equilibrio con la misma mezcla, permitiendo que la experiencia pasada con la medida de un gen dado informe de la selección de la mezcla apropiada.

En algunas realizaciones, la serie puede comprender una dilución en serie de 10 veces de una mezcla estandarizada que comprende molde competitivo de más o menos de los 12 genes de la Figura 8. Por ejemplo, puede prepararse una serie para una mezcla estandarizada de 96 ácidos nucleicos o una mezcla estandarizada que comprende diversos números de ácidos nucleicos como se detalla anteriormente.

En algunas realizaciones, el procedimiento para valorar una cantidad de ácido nucleico implica proporcionar una serie de mezclas estandarizadas diluidas en serie que comprenden un molde competitivo del ácido nucleico y un molde competitivo de otro ácido nucleico presente en un número de muestras que comprenden el ácido nucleico, en que los moldes competitivos están a concentraciones conocidas uno respecto a otro; combinar una de las muestras que comprenden el ácido nucleico con una de las mezclas estandarizadas diluidas en serie; coamplificar el ácido nucleico y su molde competitivo, produciendo producto amplificado del mismo; obtener una primera relación que compara el producto amplificado del ácido nucleico con el producto amplificado de su molde competitivo; determinar si la relación corresponde a un cociente dentro de aproximadamente 1:10 a aproximadamente 10:1 y, si no, repetir las etapas de combinación, coamplificación, obtención y determinación usando una segunda de las mezclas estandarizadas diluidas en serie. Además, en algunas realizaciones, pueden coamplificarse el otro ácido nucleico y su molde competitivo produciendo producto amplificado del mismo; obtenerse una segunda relación que compara el producto amplificado del otro ácido nucleico con su molde competitivo y comparar la primera y segunda relaciones.

En algunas realizaciones, puede usarse un enfoque “en dos etapas”. Por ejemplo, en algunas realizaciones el procedimiento comprende además diluir el producto amplificado de ácido nucleico y su correspondiente molde competitivo y coamplificar además el producto amplificado, produciendo producto amplificado adicional del mismo.

En algunas realizaciones, pueden usarse diferentes concentraciones de moldes competitivos de ácido nucleico de referencia. Por ejemplo, cuando varía la expresión de un primer ácido nucleico de referencia en comparación con un segundo ácido nucleico de referencia, el uso de más de una concentración puede ser provechoso para determinar la variación intermuestra y/o interespecímen en la expresión de los correspondientes genes de referencia. Por ejemplo, algunas realizaciones usan dos concentraciones diferentes de moldes competitivos de GAPD, ya que la expresión de GAPD respecto a β-actina puede variar tanto como aproximadamente 100 veces de un tipo de tejido a otro. Tener dos concentraciones diferentes de molde competitivo de GAPD respecto a la de β-actina puede posibilitar una mejor comparación de GAPD con β-actina en diversas muestras.

La Figura 9 ilustra cómo, en algunas realizaciones, puede usarse ácido nucleico que sirve como referencia para equilibrar una muestra con una mezcla estandarizada de la serie de mezclas estandarizadas diluidas en serie.

La etapa 901 ilustra el equilibrado cuantitativo de una muestra de ácido nucleico. Puede hacerse referencia también a equilibrado cuantitativo, como se usa en la presente memoria, como calibración cualitativa. La muestra de ácido nucleico puede diluirse proporcionando una serie de muestras diluidas en serie y seleccionarse una de las series para combinar con la mezcla estandarizada dependiendo de la concentración del ácido nucleico de referencia en la dilución. Por ejemplo, en la etapa 901, se diluye en serie material de ADNc, proporcionando una serie de muestras que tienen diluciones en serie de ácido nucleico de β -actina.

La etapa 902 ilustra que se selecciona una dilución que proporcione moléculas de molde nativo (NT) de β -actina aproximadamente equivalentes a las moléculas de molde competitivo (CT) de β -actina en una mezcla SMIS. En algunas realizaciones, puede diluirse un espécimen hasta que uno cualquiera (o más) de los ácidos nucleicos está aproximadamente equilibrado con, concretamente aproximadamente calibrado con, la cantidad de molde competitivo de ese ácido nucleico en la mezcla estandarizada. Por tanto, en realizaciones preferidas, se selecciona la primera de una serie de muestras para combinar con la mezcla estandarizada, proporcionando un ácido nucleico de referencia calibrado o aproximadamente calibrado con su molde competitivo en la mezcla estandarizada. Puede aparecer la calibración aproximada cuando el ácido nucleico está dentro de un intervalo de aproximadamente 10 veces, un intervalo de 9 veces, un intervalo de 8 veces, un intervalo de 7 veces, un intervalo de 6 veces, un intervalo de 5 veces, un intervalo de 3 veces, un intervalo de 2 veces o un intervalo de 1 vez o menos, de la concentración de molde competitivo de ese ácido nucleico particular en la mezcla estandarizada. En realizaciones preferidas, el cociente de NT/CT para el ácido nucleico de referencia está entre aproximadamente 1:10 y aproximadamente 10:1 (por ejemplo, para que la medida esté dentro del intervalo dinámico lineal).

La Figura 10 ilustra la selección de una dilución de ADNc que proporcione un ácido nucleico de referencia (β -actina en esta ilustración) en equilibrio con 600.000 moléculas del molde competitivo del ácido nucleico de referencia en la mezcla estandarizada, por ejemplo, para que el ácido nucleico pueda competir igualmente (o de forma sustancialmente igual) con las 600.000 moléculas de molde competitivo. En realizaciones preferidas, todas las mezclas estandarizadas de una serie dada contienen un número dado de moléculas de un ácido nucleico particular, permitiendo usar cualquiera de las mezclas estandarizadas en el equilibrado. Por ejemplo, A-F pueden contener cada una molde competitivo de β -actina aproximadamente 10^{-12} M, de modo que puede usarse cualquiera de las mezclas A-F para equilibrar con una muestra de ADNc. Típicamente, se usa la mezcla F para equilibrar el ADNc de β -actina en una muestra.

La Figura 11 ilustra una serie de mezclas estandarizadas diluidas en serie que comprenden una o más mezclas en que 1 μ l contiene 600.000 moléculas de molde competitivo de β -actina, correspondiente a 1 μ l de una mezcla estandarizada que contiene molde competitivo de β -actina 10^{-12} M. En ese caso, por ejemplo, puede diluirse material de ADNc hasta calibrar 1 μ l con 600.000 moléculas de molde competitivo de β -actina. Típicamente, esta es la cantidad de ADNc derivado de 100 a 1.000 células en el caso de β -actina. Aunque el número de copias de ARNm de β -actina/célula varía de una célula a otra, usando una estimación conservadora de 600 copias de ARNm de β -actina/célula y suponiendo una eficacia de la transcripción inversa del 10 %, una muestra de ADNc que contiene 600.000 moléculas de ADNc de β -actina puede derivar de 1.000 células.

Esta cantidad puede usarse para proporcionar suficiente ADNc para cuantificar genes expresados a bajos niveles, por ejemplo, genes expresados con bajo número de copias, por ejemplo, a aproximadamente 0,1 copias/célula, 0,05 copias/célula y/o 0,01 copias/célula. Con el ADNc de referencia en equilibrio con β -actina aproximadamente 10^{-12} M en la reacción PCR, algunas realizaciones pueden cuantificar el ácido nucleico de muestra que está en equilibrio con su CT aproximadamente 10^{-16} M o menos. En algunas realizaciones específicas, cuando el ADNc de referencia está en equilibrio con β -actina aproximadamente 10^{-12} M en un volumen de reacción PCR de 10 μ l, puede haber aproximadamente 600.000 moléculas de NT de β -actina y aproximadamente 600.000 moléculas de CT de β -actina en la reacción, y el número de moléculas de ácido nucleico de muestra en equilibrio con su CT aproximadamente 10^{-16} M o aproximadamente 10^{-17} M puede ser de aproximadamente 60 o aproximadamente 6, respectivamente. Aproximadamente 60 o aproximadamente 6 moléculas de ácido nucleico pueden traducirse en aproximadamente 0,1 a aproximadamente 0,01 moléculas/célula.

Este equilibrado puede proporcionar al menos aproximadamente 10 copias presentes el inicio de la amplificación evitando, por ejemplo, problemas estequiométricos. En algunas realizaciones, cuando se busca menos sensibilidad, puede usarse menos ADNc. Por ejemplo, en algunas realizaciones, puede usarse una cantidad de ADNc aproximadamente en equilibrio con 60.000 moléculas de CT de β -actina, permitiendo un consumo reducido de ADNc, por ejemplo por aproximadamente 10 veces.

Puede seleccionarse la primera de las mezclas estandarizadas diluidas en serie para combinar con la muestra de ácido nucleico. La Figura 12 ilustra que puede usarse la mezcla E inicialmente, basándose en los niveles de expresión de la mayoría de genes. Parece haber una distribución estequiométrica y/o estocástica de la expresión entre genes (véase, por ejemplo, Kuznetsova, *et al.*, "General Statistics of Stochastic Process of Gene Expression in Eukaryotic Cells", *Genetics*, vol. 161, 1321-1332, julio de 2002), con una media aproximadamente 2 órdenes de

magnitud menor que la expresión de β -actina, por ejemplo, en células epiteliales bronquiales humanas. Sin limitarse a una teoría y/o hipótesis dada, la distribución de los niveles de expresión génica en células indica que los transcritos de ARNm de muchos genes estarán equilibrados con la mezcla E en algunas realizaciones.

5 La Figura 12 ilustra además que el uso de una serie de mezclas estandarizadas diluidas en serie de algunas realizaciones puede permitir la medida de la expresión génica dentro del espectro completo observado. Como la Figura 12 ilustra mediante una codificación por color, pueden usarse diferentes mezclas para medir los genes expresados a diferentes niveles con buena reproducibilidad. Debido a que hay de aproximadamente 100 a aproximadamente 1.000 copias/célula de β -actina para la mayoría de tipos celulares, este nivel de sensibilidad permite la medida de una molécula por aproximadamente 100 a aproximadamente 1.000 células. En el otro extremo del espectro de expresión, una mezcla estandarizada que comprende concentraciones mayores de moldes competitivos puede permitir la medida de genes más altamente expresados. Por ejemplo, la mezcla A en algunas realizaciones puede permitir la medida de más de 10^7 moléculas/ 10^6 moléculas de β -actina (de aproximadamente 1.000 a aproximadamente 10.000 copias/célula). Los ejemplos de genes expresados a estos niveles incluyen UGB (Genbank nº U01101) y vimentina (X56134).

15 En otras realizaciones, puede usarse inicialmente una mezcla diferente basándose en la experiencia pasada y/o la predicción de las cantidades de ácido nucleico esperadas. Por ejemplo, pueden usarse inicialmente mezcla A, mezcla B, mezcla C, mezcla E o mezcla F. En realizaciones preferidas, la mezcla seleccionada es aquella que contiene una concentración de molde competitivo que es probable que se calibre aproximadamente (por ejemplo, dentro de un intervalo de aproximadamente 10 veces) con el gen o genes que se están valorando. En realizaciones preferidas, puede seleccionarse una mezcla estandarizada apropiada basándose en los datos de algunas realizaciones de bases de datos de expresión estandarizadas descritas en la presente memoria.

25 Después de combinar una muestra que comprende un ácido nucleico para medir con una de las series de mezclas estandarizadas diluidas en serie, pueden coamplificarse entonces el ácido nucleico y su molde competitivo, por ejemplo, como se describe con detalle anteriormente. También como se describe anteriormente, puede obtenerse un cociente que compara la cantidad de producto amplificado del ácido nucleico con la cantidad de producto amplificado de su correspondiente molde competitivo. Aunque el ácido nucleico de referencia en la muestra se equilibró con su molde competitivo en la mezcla, el ácido nucleico diana no puede equilibrarse. Cuando las cantidades de producto amplificado de un ácido nucleico diana y de su molde competitivo difieren en gran medida, puede repetirse la coamplificación usando una mezcla diferente de la serie de mezclas diluidas en serie. Es decir, puede seleccionarse una segunda y/o posterior mezcla estandarizada diluida en serie para combinar con la muestra de ácido nucleico.

35 La Figura 13 ilustra una situación en que la mezcla inicial no proporcionaba molde competitivo del ácido nucleico diana suficientemente en equilibrio con la cantidad de ácido nucleico diana en la dilución de ADNc. El ácido nucleico diana en esta ilustración corresponde a c-myc; IS hace referencia a un patrón interno de molde competitivo. Como ilustra la Figura 12, el producto amplificado de NT de c-myc no está dentro de aproximadamente 10 veces la cantidad de producto amplificado de CT de c-myc. En algunas realizaciones, el software determina el área bajo la curva para NT y CT y calcula el cociente de NT/CT para el ácido nucleico diana.

40 En realizaciones preferidas, la siguiente mezcla seleccionada de la serie está basada en el cociente obtenido cuando se compara el producto amplificado del ácido nucleico diana con el producto amplificado de su molde competitivo. Por ejemplo, cuando el cociente es de menos de aproximadamente 1/10, se usará posteriormente una mezcla más diluida de la serie; cuando la proporción NT/CT es de más de aproximadamente 10/1, se usará una mezcla más concentrada de la serie. La Figura 12 ilustra la situación en que se obtiene un cociente alto, indicando que debería usarse a continuación una mezcla más concentrada, por ejemplo la mezcla C. En algunas realizaciones, puede usarse software para determinar automáticamente cuál mezcla debería seleccionarse a continuación.

45 La Figura 14 ilustra además la selección de la mezcla C. El cociente de NT/CT obtenido para el ácido nucleico diana (c-myc en esta ilustración) se representa en una gráfica. La posición en la gráfica puede indicar cuál mezcla debería usarse para el ácido nucleico expresado a ese nivel. En algunas realizaciones, descritas con más detalle a continuación, el software comunica automáticamente la mezcla correcta para usar a un robot.

50 Puede combinarse entonces otra muestra del ácido nucleico, por ejemplo a la misma dilución de ADNc, con la mezcla estandarizada diluida en serie seleccionada posteriormente. Después de combinar, pueden coamplificarse el ácido nucleico y su molde competitivo, por ejemplo como se describe con detalle anteriormente. También como se describe anteriormente, puede obtenerse un cociente que compara la cantidad de producto amplificado del ácido nucleico con la cantidad de producto amplificado de su correspondiente molde competitivo.

55 La Figura 15 ilustra la situación en que la siguiente mezcla seleccionada proporciona molde competitivo del ácido nucleico diana suficientemente en equilibrio con la cantidad de ácido nucleico diana a la dilución de ADNc. Como ilustra la Figura 15, el producto amplificado de NT de c-myc está dentro de aproximadamente una cantidad de 10 veces el producto amplificado de CT de c-myc. En algunas realizaciones, el software determina el área bajo la curva para NT y CT y calcula el cociente de NT/CT para el ácido nucleico diana. En algunas realizaciones, el software puede comparar también este cociente con el cociente de NT/CT del ácido nucleico que sirve como referencia.

- En realizaciones preferidas, la cantidad de ADNc de muestra puede mantenerse constante mientras se usa una mezcla estandarizada diferente. Como otro ejemplo, si se usaba mezcla D y la cantidad de producto amplificado de NT era más de 10 veces mayor que la del correspondiente CT, el experimento puede repetirse con la misma cantidad de partida de ADNc, pero usando la mezcla C, que tiene una concentración aproximadamente 10 veces mayor del molde competitivo, o mezcla A o mezcla B. Cuando la cantidad de producto amplificado es menos de 10 veces menor que la del correspondiente CT, el experimento puede repetirse con la misma cantidad de partida de ADNc, pero usando mezcla E o mezcla F. La mezcla más diluida y/o la mezcla más concentrada seleccionadas pueden ser la siguiente mezcla más diluida y/o más concentrada en la serie o una mezcla diluida en serie diferente en la serie dependiendo, por ejemplo, de la magnitud del cociente obtenido.
- Una realización altamente preferida, en términos de consumo de ADNc y coste reducido, implica el uso de 1 µl de ADNc equilibrado en la ronda 1 de un proceso en dos etapas con cada una de las seis mezclas de molde competitivo (A-F); el uso de 10 nl del producto amplificado en la ronda 1 en paralelo con un volumen de 100 nl de amplificaciones de la ronda 2 para medir las cantidades de todos los 96 ácidos nucleicos usando la mezcla E (que contiene moldes competitivos a una concentración que estará en equilibrio con la mayoría de genes) y la repetición de las etapas anteriores para los ácidos nucleicos que no están en equilibrio con la mezcla E usando la mezcla apropiada.
- Cuando se usa una mezcla apropiada, puede valorarse la cantidad de ácido nucleico diana de acuerdo con los procedimientos descritos en la presente memoria. La Figura 16 ilustra el cálculo de el “cociente de cocientes” basado en los datos obtenidos usando una mezcla apropiada.
- La Figura 17 ilustra una serie de electroferogramas, por ejemplo, como pueden obtenerse en realizaciones preferidas, en que se valoran conjuntamente múltiples ácidos nucleicos. Se proporcionan en el ejemplo II siguiente detalles adicionales respecto a la práctica de diversos esquemas de etapas anteriores.
- Como se indica anteriormente, en algunas realizaciones el procedimiento para valorar ácidos nucleicos usando una serie de mezclas estandarizadas diluidas en serie está informatizado. La Figura 18 ilustra esquemáticamente un sistema global para valorar ácidos nucleicos, una o más etapas del cual pueden estar informatizadas en diversas realizaciones.
- En la etapa (a), un programa de software puede determinar la concentración deseada de reactivos de molde competitivo para usar. Esta etapa puede comprender seleccionar una dilución de muestra y/o seleccionar una mezcla de la serie de mezclas diluidas en serie para combinar. Por ejemplo, la informatización puede comprender instruir a un manipulador robótico para seleccionar la primera de las mezclas estandarizadas diluidas en serie para combinar, por ejemplo la mezcla E como se detalla anteriormente.
- En la etapa (b), un programa de software puede causar que al menos un reactivo se dispense en uno o más recipientes, en los que se van a realizar las reacciones de amplificación, y el producto amplificado pueda dirigirse a un dispositivo adecuado para separar, identificar y/o marcar, por ejemplo, mediante flujo a una máquina de electroforesis capilar de microfluidos (CE). En algunas realizaciones, esta etapa puede comprender instruir a un manipulador robótico para dispensar una mezcla seleccionada y/o dilución de muestra a un recipiente, coamplificar los ácidos nucleicos y sus correspondientes moldes competitivos y separar el producto amplificado.
- En la etapa (c), puede analizarse la información respecto a los productos amplificados separados. Por ejemplo, la etapa (c) puede comprender obtener una relación que compara el producto amplificado de un ácido nucleico con el producto amplificado de su molde competitivo. Por ejemplo, después de suficiente electroforesis en gel, puede digitalizarse automáticamente la imagen de los geles y analizarse automáticamente la imagen para valorar las cantidades de producto amplificado, por ejemplo, determinando automáticamente el área bajo las curvas. Por ejemplo, el software puede determinar el área bajo la curva para NT y CT de un ácido nucleico dado y calcular el cociente de NT/CT.
- En algunas realizaciones, se incorporan las etapas de cálculo a una hoja de cálculo. Por ejemplo, en algunas realizaciones un usuario puede introducir los valores brutos (por ejemplo, de alturas máximas o área bajo la curva) para los productos de PCR NT, CT y heterodiméricos de un gen dado para medir en una hoja de cálculo, y puede calcularse automáticamente el valor de expresión del gen. En algunas realizaciones, puede usarse software para introducir automáticamente los valores del producto amplificado por NT y CT para cada uno o más ácidos nucleicos para medir en una hoja de cálculo, calculando automáticamente un valor numérico, por ejemplo, un valor numérico correspondiente a la expresión génica.
- La información de la etapa (c) puede proporcionarse en un “informe” enviado a una “base de datos” y/o enviarse a la etapa (d), que puede repetir el proceso para análisis adicional de los datos recibidos. Por ejemplo, si el cociente calculado no está dentro del intervalo deseado (por ejemplo, dentro de un cociente de aproximadamente 1:10 a aproximadamente 10:1), como se describe anteriormente, puede elegirse una nueva concentración deseada de reactivos de molde competitivo (concretamente, diferente de las concentraciones originales seleccionadas en la etapa (a)) y repetirse la etapas (b)-(c). En algunas realizaciones, puede usarse software para determinar automáticamente cuál mezcla debería seleccionarse a continuación, basándose en las consideraciones descritas

anteriormente. En algunas realizaciones, un programa de software puede instruir a un manipulador robótico para combinar una muestra con la nueva mezcla.

Se describe también en la presente memoria un programa informático para ejecutar ciertas realizaciones de procedimientos de la presente invención. El programa informático incluye un medio legible informáticamente e instrucciones almacenadas en el medio legible informáticamente. Se describe además que las instrucciones incluyen una o más de las etapas indicadas anteriormente. El programa informático puede incluir además instrucciones para dispensar el producto amplificado a matrices para medida, así como instrucciones para marcar fluorescentemente producto amplificado y/o ácido nucleico con el que hibridan. El producto amplificado puede marcarse, por ejemplo, marcando uno más nucleótidos en la reacción de amplificación con un resto detectable, por ejemplo un resto fluorescente. El programa informático puede incluir adicionalmente instrucciones para medir cantidades de ácido nucleico, por ejemplo, comparando las intensidades fluorescentes de las matrices del producto amplificado de un ácido nucleico dado y su molde competitivo.

D. *Relación sustancialmente constante*

Algunas realizaciones de la presente invención descritas anteriormente proporcionan una relación para valorar el ácido nucleico en que la relación permanece constante o sustancialmente constante más allá de la fase exponencial de amplificación. En las amplificaciones de ácido nucleico, por ejemplo PCR, la cantidad de producto amplificado puede dejar de aumentar exponencialmente después de un número indefinido de ciclos. Por ejemplo, en cierto punto y por razones desconocidas, puede limitarse la reacción de amplificación y puede aumentar la cantidad de producto amplificado a una velocidad desconocida y/o no exponencial. Por ejemplo, la velocidad de amplificación por PCR puede ser baja en los ciclos tempranos cuando la concentración de los moldes es baja. Después de un número impredecible de ciclos, la reacción puede entrar en una fase de amplificación lineal logarítmica. En los ciclos tardíos, la velocidad de amplificación puede retardarse a medida que la concentración de productos de PCR se vuelve mayor, por ejemplo, suficientemente alta para competir con los cebadores por la unión a moldes. Se ha reseñado que el rendimiento del producto amplificado en reacciones PCR, por ejemplo, varía tanto como 6 veces entre muestras idénticas procesadas simultáneamente. Gilliland, G., *et al.*, Proc. Natl. Acad. Sci. 87: 2725-2729, 1990. Las técnicas de PCR se describen en general en las patentes de EE.UU. n° 4.683.195, 4.683.202 y 4.965.188. Otros investigadores han analizado muestras amplificadas durante un número de ciclos conocido por proporcionar amplificación exponencial (Horikoshi, T., *et al.*, Cancer Res. 52: 108-116 (1992); Noonan, K. E., *et al.*, Proc. Natl. Acad. Sci. 87: 7160-7164 (1990); Murphy, L. D., *et al.*, Biochemistry 29: 10351-10356 (1990); Carre, P. C., *et al.*, J. Clin. Invest. 88: 1802-1810 (1991); Chelly, J., *et al.*, Eur. J. Biochem. 187: 691-698 (1990); Abbs, S., *et al.*, J. Med. Genet. 29: 191-196 (1992); Feldman, A. M. *et al.*, Circulation 83: 1866-1872 (1991). Algunas realizaciones de la presente invención permiten la cuantificación de la amplificación por PCR en cualquier fase del proceso de PCR, incluyendo la fase de meseta.

Algunas realizaciones de la presente invención se refieren a obtener una relación constante (o sustancialmente constante) más allá de la fase exponencial de amplificación de ácido nucleico, permitiendo así determinar la cantidad inicial de ácido nucleico por extrapolación a partir de las cantidades de punto final de producto amplificado. En algunas realizaciones, la fase exponencial para amplificar el ácido nucleico no tiene que definirse para cada conjunto de condiciones experimentales, ahorrando tiempo y materiales. Por ejemplo, algunas realizaciones no implican medidas a tiempo real. Algunas realizaciones no implican la generación de una curva patrón y/o la generación de múltiples curvas patrón, por ejemplo, cuando la curva patrón se usa para determinar el intervalo exponencial de amplificación para un ácido nucleico dado para medir y/o cuando las curvas patrón comparan las cantidades medidas de un ácido nucleico con otro.

Las Figuras 19 y 20 ilustran cómo ciertas realizaciones de la presente invención pueden proporcionar una relación constante o sustancialmente constante usando medidas de punto final más allá de la fase exponencial de amplificación. Los datos mostrados en las gráficas se obtuvieron como se detalla en el Ejemplo III siguiente. La Figura 19 ilustra que la cantidad de producto amplificado frente a cantidad de partida total de ARN no permanece lineal al aumentar las cantidades de ARN, por ejemplo, más allá de la fase exponencial de amplificación. Es decir, existe una relación no lineal entre la cantidad de producto amplificado (cuadrados blancos: glutation peroxidasa GSH-Px; cuadrados negros: gliceraldehído-3-fosfato deshidrogenasa GAPDH) y la cantidad de partida total de ARN para cantidades crecientes de ARN, por ejemplo, más allá de la fase exponencial de amplificación. Las líneas rectas representan las cantidades teóricas de producto de PCR (GSH-Px o GAPDH) que se obtendrían si la amplificación permaneciera exponencial a lo largo del proceso de amplificación.

La Figura 20 ilustra que puede obtenerse una relación lineal cuando se representa el cociente de (producto amplificado de ácido nucleico/producto coamplificado de su molde competitivo) frente a la cantidad de partida total de ARN para el primer y segundo ácidos nucleicos correspondientes a GSH-Px (cuadrados blancos) y GAPDH (cuadrados negros), respectivamente. La Figura 20 ilustra dos relaciones lineales (o sustancialmente lineales) para dos ácidos nucleicos diferentes, coamplificada cada una con su molde competitivo respectivo, con $r_2 = 0,982$ para GSH-Px y $r_2 = 0,973$ para GAPDH para el intervalo de ARN total estudiado.

La Figura 21 ilustra que la relación de (producto amplificado del primer ácido nucleico/producto coamplificado de su molde competitivo)/(producto amplificado del segundo ácido nucleico/producto coamplificado de su molde

competitivo) a cantidad de partida total de ARN permanece constante, o sustancialmente constante, para dos ácidos nucleicos diferentes cuando se amplifican de acuerdo con diversas realizaciones de la presente invención. En consecuencia, algunas realizaciones de la presente invención usan una relación que compara al menos dos relaciones de al menos dos ácidos nucleicos en una muestra, a saber, una primera relación que compara el producto amplificado de un primer ácido nucleico con el producto coamplificado de un molde competitivo del primer ácido nucleico, y una segunda relación que compara el producto amplificado de un segundo ácido nucleico con el producto coamplificado de un molde competitivo del segundo ácido nucleico. Se proporcionan anteriormente detalles adicionales de algunas de estas coamplificaciones. En algunas realizaciones, la relación buscada compara además el producto amplificado de un número de otro(s) ácido(s) nucleico(s) con el producto coamplificado del (de los) molde(s) competitivo(s) de dicho número de otro(s) ácido(s) nucleico(s).

E. Sensibilidad

Algunas realizaciones de la presente invención descritas anteriormente proporcionan una relación para valorar ácido nucleico en que la relación con sensibilidad. La sensibilidad puede definirse como la capacidad de un procedimiento de producir un cambio en la señal por un cambio definido en la cantidad de analito, concretamente, la pendiente de una curva de calibración. Algunas realizaciones de la presente invención proporcionan una pendiente mayor de aproximadamente 0,1, mayor de aproximadamente 0,2, mayor de aproximadamente 0,5 o mayor de aproximadamente 0,8. Algunas realizaciones preferidas de la presente invención proporcionan una pendiente de aproximadamente 1/1.

Por ejemplo, algunas realizaciones de la presente invención proporcionan una relación capaz de detectar una diferencia de menos de aproximadamente 2 veces, una diferencia de menos de aproximadamente 1 vez, una diferencia de menos de aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 %, una diferencia de menos de aproximadamente un 30 %, una diferencia de menos de aproximadamente un 20 %, una diferencia de menos de aproximadamente un 10 %, una diferencia de menos de aproximadamente un 5 % o una diferencia de menos de aproximadamente un 1%. Dichas sensibilidades pueden corresponder a la identificación de cambios pequeños en la expresión génica.

En algunas realizaciones, pueden detectarse una o más de estas diferencias en aproximadamente 1.000 moléculas o menos del ácido nucleico en la muestra, por ejemplo, en aproximadamente 800, en aproximadamente 600 o en aproximadamente 400 moléculas. En algunas realizaciones, pueden detectarse una o más de estas diferencias en aproximadamente 100 moléculas o menos (por ejemplo, en aproximadamente 60 moléculas), en aproximadamente 10 moléculas o menos (por ejemplo, en aproximadamente 6 moléculas) o en aproximadamente 1 molécula o menos de ácido nucleico en una muestra. En algunas realizaciones, pueden detectarse una o más de estas diferencias en menos de aproximadamente 10.000.000, menos de aproximadamente 5.000.000, menos de aproximadamente 1.000.000, menos de aproximadamente 500.000, menos de aproximadamente 100.000, menos de aproximadamente 50.000, menos de aproximadamente 10.000, menos de aproximadamente 8.000, menos de aproximadamente 6.000, menos de aproximadamente 5.000, o menos de aproximadamente 4.000 moléculas de ácido nucleico en una muestra.

Algunas realizaciones, como se describe anteriormente, valoran ácidos nucleicos dentro de un intervalo de concentraciones, por ejemplo, valoran la expresión génica dentro de uno o más órdenes de magnitud de expresión génica. En algunas de dichas realizaciones, la valoración detecta una diferencia de menos de aproximadamente 2 veces dentro del intervalo. En algunas realizaciones, la valoración detecta una diferencia de menos de aproximadamente 1 vez dentro del intervalo, una diferencia de menos de aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 %, una diferencia de menos de aproximadamente un 30 %, una diferencia de menos de aproximadamente un 20 %, una diferencia de menos de aproximadamente un 10 %, una diferencia de menos de aproximadamente un 5 % o una diferencia de menos de aproximadamente un 1 % dentro del intervalo.

Las sensibilidades descritas en la presente memoria pueden conseguirse por algunas de las realizaciones de la presente invención.

F. Reproducibilidad.

En realizaciones preferidas, los procedimientos para valorar un ácido nucleico son reproducibles. Algunas realizaciones, por ejemplo, proporcionan un coeficiente de variación de menos de aproximadamente un 25 % entre muestras de un ácido nucleico. En algunas realizaciones, el coeficiente de variación es de menos de aproximadamente un 50 %, menos de aproximadamente un 30 %, menos de aproximadamente un 25 %, menos de aproximadamente un 20 %, menos de aproximadamente un 15 %, menos de aproximadamente un 10 %, menos de aproximadamente un 5 % o menos de aproximadamente un 1 % entre 2 o más muestras de ácido nucleico. Dichos coeficientes de variación pueden obtenerse en algunas realizaciones en que se amplifican y/o se valoran las 2 muestras en diferentes momentos, por ejemplo, en diferentes días; en el mismo o diferentes experimentos en el mismo laboratorio y/o en diferentes experimentos en diferentes laboratorios y/o cuando las muestras se obtienen a partir de diferentes sujetos y/o diferentes especies. Las realizaciones preferidas de la presente invención proporcionan reproducibilidad tanto intra- como interlaboratorio (M.T. Vondracek, D.A. Weaver, Z. Sarang *et al.*, *Int. J. Cancer* 99, 776-782 (2002)) que es suficiente para detectar diferencias de menos de 2 veces en la expresión

génica. Por ejemplo, en algunas realizaciones, la correlación de varianza interlaboratorio era de 0,48, por ejemplo, a partir de medidas de expresión génica que usan una muestra de ADNc de A459 tomada en diferentes laboratorios en diferentes momentos, que se extienden casi un año. En algunas realizaciones, por ejemplo realizaciones que usan electroforesis capilar de microcanales, la correlación de varianza se reducía a 0,26. Se proporcionan en el Ejemplo IV siguientes detalles adicionales de un estudio para evaluar la reproducibilidad.

En algunas realizaciones, la reproducibilidad entre muestras permite el uso de menos tubos de dilución. En algunas realizaciones, puede usarse un solo tubo, simplificando los procedimientos y permitiendo la evaluación de muchas muestras diferentes a la vez.

En algunas realizaciones, incluyendo patrones internos de molde competitivo en una mezcla estandarizada común usada en diferentes medidas, pueden controlarse una o más fuentes de variación. Las fuentes de variación incluyen, por ejemplo, variación por la carga de ADNc, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y/o eficacia de amplificación intramuestra. Por ejemplo, algunas realizaciones que usan Agilent 2100 Bioanalyzer proporcionan un CV de reproducibilidad intralaboratorio de menos de aproximadamente 0,25 rutinariamente, con una sensibilidad comparable a la electroforesis en gel plano.

La Figura 22 tabula una serie de fuentes de variación y procedimientos de control

La variación de la carga de ADNc puede ser el resultado de la variación en el pipeteado, toma de alícuotas, cuantificación y/o transcripción inversa. Por ejemplo, pueden aparecer errores cuando se toman alícuotas de material de ARN en recipientes para efectuar la transcripción inversa. Aunque la eficacia de la transcripción inversa puede variar de una muestra a otra, la representación de un ácido nucleico a otro en una muestra no tiene que variar entre diferentes transcripciones inversas.

Por ejemplo, la eficacia de la transcripción inversa puede variar de aproximadamente 5 a aproximadamente 90 % (Simmonds *et al.*, 1990). La variación en la eficacia de la transcripción inversa, sin embargo, puede afectar a diferentes transcritos de la misma o sustancialmente la misma manera (Willey *et al.*, 1998; Loitsch *et al.*, 1999). En un experimento, por ejemplo, se midió la expresión génica en 5 transcripciones inversas diferentes de una muestra dada de ARN de la línea celular de carcinoma no microcítico SW900. El nivel medio de expresión obtenido fue de 3.600 moléculas/10⁶ moléculas de β -actina con un CV de 0,26, no mayor que si se hubieran realizado medidas duplicadas en ADNc resultante de una sola transcripción inversa. Sin embargo, si las reacciones de transcripción inversa y amplificación se llevan a cabo en diferentes recipientes, pueden aparecer errores cuando se pipetea ADNc de la transcripción inversa a los recipientes de PCR individual. Es decir, sin limitarse a una teoría y/o hipótesis particular, el efecto de la variación en la transcripción inversa puede ser el mismo que si se cargaran diferentes niveles de ADNc en una reacción PCR. Controlar la carga de ADNc puede controlar entonces la variación de la eficacia de la transcripción inversa.

La variación de la eficacia de amplificación intraácido nucleico puede ser el resultado, por ejemplo, de la variación de ciclo a ciclo, por ejemplo, cuando diferentes ciclos de amplificación muestran diversas fases temprana lenta, logarítmica lineal y/o tardía lenta de meseta, como se describe anteriormente. Cuando se mide la expresión génica, la eficacia de amplificación intraácido nucleico puede hacer referencia a la eficacia de amplificación intragén, concretamente por ejemplo, la variación en amplificaciones repetidas de ADNc correspondiente a un gen dado.

La variación de la eficacia de amplificación interácido nucleico puede hacer referencia a la eficacia de amplificación intergén, por ejemplo, cuando la eficacia a la que se amplifica un gen dado difiere de aquella a la que se amplifica un gen diferente. Dichas diferencias pueden estar causadas, por ejemplo, por diferencias en los cebadores usados para amplificar los diferentes genes medidos en la misma y/o diferentes muestras. Por ejemplo, la eficacia de un par de cebadores, por ejemplo como se define por el umbral mínimo de detección (LDT), no puede ser predecible, y puede variar más de aproximadamente 100.000 veces (desde menos de aproximadamente 10 moléculas a aproximadamente 10⁶ moléculas) en algunas realizaciones.

También un lote defectuoso (por ejemplo, en que ha aparecido degradación de cebadores y/o moldes competitivos) o una concentración inapropiada de cebadores causaría la variación de la amplificación por PCR de un ácido nucleico respecto a otro. En algunas realizaciones, la concentración de molde competitivo es pequeña (por ejemplo, en el intervalo femtomolar) de modo que cualquier cambio en el número de moléculas presentes en la reacción puede introducir una gran fuente de error. La presencia de un inhibidor podría alterar la eficacia de amplificación por PCR de un ácido nucleico, por ejemplo, un gen, en comparación con otro.

La variación en la eficacia de amplificación interespecimen puede estar causada, por ejemplo, por la presencia variable de un inhibidor (por ejemplo, un inhibidor de PCR) en diferentes especímenes. Los inhibidores de reacciones PCR incluyen, por ejemplo, hemo. Akane, A., Matsura, K., Nakamura, H., Takahashi, S. y Kimura, K. (1994) "Identification of the heme compound co purified with deoxyribonucleic acid (DNA) from blood stains, a major inhibitor of polymerase chain reaction (PCR) amplification". *J. Forensic Sci.* 39, 362-372; Zhu, Y. H., Lee, H. C., y Zhang, L. (2002) "An examination of heme action in gene expression: Heme and heme deficiency affect the expression of diverse genes in erythroid K562 and neuronal PC12 cells". *DNA Cell Biol.* 21, 333-346. Además, la

eficacia de amplificación para diferentes genes puede afectarse en diferentes grados en diferentes muestras y/o especímenes. Meijerink, J., Mandigers, C., van de Locht, L., *et al.* (2001) "A novel method to compensate for different amplification efficiencies between patient DNA samples in quantitative real-time PCR". *J. Mol. Diagn.* 3, 55-61; Giulietti, A., Overbergh, L., Valckx, D., *et al.* (2001) "An overview of real-time quantitative PCR: applications to quantify cytokine gene expression". *Methods* 25, 386-401. Dichas diferencias pueden dar como resultado la variación de la medida del mismo o diferentes ácidos nucleicos (por ejemplo, el mismo o diferentes genes) en el mismo o diferentes especímenes y/o muestras. Por ejemplo, un inhibidor de PCR dado puede tener poco efecto sobre la amplificación de un gen poco expresado, por ejemplo GSTM3. El mismo inhibidor de PCR puede tener un mayor efecto, por ejemplo un efecto significativamente mayor, sobre la amplificación de un gen más altamente expresado, por ejemplo, ERBB2, incluyendo, por ejemplo, evitar la amplificación o reducir la amplificación a niveles no detectables.

La variación en la amplificación intermuestra puede hacer referencia a la variación interreacción o a la variación de pocillo a pocillo en las medidas repetidas del mismo o diferentes ácidos nucleicos (por ejemplo, el mismo o diferentes genes). La variación de la eficacia de amplificación intermuestra puede ser el resultado, por ejemplo, de la presencia variable de un inhibidor (por ejemplo, un inhibidor de PCR) en diferentes recipientes de reacción, de la variación de la temperatura de ciclado entre las diferentes regiones de un bloque termociclador, de la calidad variable de uno o más reactivos de PCR o de las concentraciones variables de uno o más reactivos de PCR (por ejemplo, cebadores).

Una o más de estas fuentes de variación pueden reducir la eficacia de amplificación por PCR en un pocillo hasta el punto en que no pueda observarse producto de PCR en ese pocillo. Algunas realizaciones de la presente invención permiten reconocer este tipo de error, por ejemplo realizaciones que usan una mezcla estandarizada que comprende molde competitivo aproximadamente 10^{-17} M del ácido nucleico que se busca amplificar. En un volumen de reacción PCR de 10 μ l, aproximadamente 10^{-17} M representa aproximadamente 60 moléculas. Con aproximadamente 60 moléculas de patrón interno presentes en la reacción PCR y los componentes de la reacción PCR funcionando apropiadamente, si un ácido nucleico no está presente en una muestra, se observará el producto amplificado del molde competitivo, pero no el producto amplificado del ácido nucleico. Esto puede indicar que había menos de aproximadamente 6 moléculas (aproximadamente 10 veces menos que el número de moléculas de molde competitivo) de ácido nucleico en la muestra. Por otro lado, si no es detectable el producto amplificado ni del ácido nucleico ni de su molde competitivo, puede determinarse que la eficacia de la reacción PCR era subóptima.

La variación de la amplificación intramuestra puede hacer referencia a la variación intrarreacción, por ejemplo, la eficacia de amplificación variable en una reacción dada usando una muestra dada. La variación de la eficacia de amplificación intramuestra puede ser el resultado, por ejemplo, de la variación de la eficacia del termociclador en diversas posiciones dentro del termociclador, y puede introducir variación cuando se miden cantidades del mismo o diferentes ácidos nucleicos (por ejemplo, expresión del mismo o diferentes genes) en la misma o diferentes muestras y/o especímenes.

Algunas realizaciones para medir ácidos nucleicos controlan la variación causada por una o más de las fuentes de variación seleccionadas de carga de ADN, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y eficacia de amplificación intramuestra. Por ejemplo, en algunas realizaciones, el uso de una mezcla estandarizada y/o una serie de mezclas estandarizadas diluidas en serie puede proporcionar el control.

Algunas realizaciones preferidas controlan una o más fuentes de variación sin el uso de medidas a tiempo real obtenidas usando análisis cinético (por ejemplo, medidas de PRC a tiempo real). Por ejemplo, obtener un "cociente de cocientes" en algunas realizaciones no implica tomar medidas a tiempo real. Algunas realizaciones preferidas controlan una o más fuentes de variación sin generar una o más curvas patrón. Por ejemplo, obtener un "cociente de cocientes" en algunas realizaciones no implica generar una curva patrón. En realizaciones más preferidas, se controlan una o más fuentes de error usando procedimientos que no implican medidas a tiempo real ni la generación de una curva patrón. En realizaciones aún más preferidas, se controlan dos o más, tres o más, cuatro o más, cinco o más o seis fuentes de variación sin medidas a tiempo real ni la generación de una curva patrón.

La Figura 23 ilustra el control de una o más de estas fuentes de error en algunas realizaciones en comparación con la RT-PCR a tiempo real en dos especímenes diferentes en cuatro experimentos diferentes. En la Figura 23, se hace referencia a los ácidos nucleicos que se están midiendo como molde nativo (NT), se hace referencia al molde competitivo para cada uno como CT y el segundo ácido nucleico sirve como ácido nucleico de referencia.

La Figura 23 ilustra el producto amplificado de molde nativo y molde competitivo para un primer y un segundo ácidos nucleicos que se amplifican por PCR simultáneamente durante el número indicado de ciclos. Los productos amplificados en el punto final se separan por electroforesis, por ejemplo, en presencia de tinte intercalante fluorescente, y se cuantifican densitométricamente. En la realización ilustrada, el producto de PCR de CT más corto migra más rápido que el producto de PCR de NT y se representa por una banda de CT por debajo de la banda de NT. Como entenderá un especialista en la materia, si hay más producto de NT que producto de CT, la banda de CT emitirá más luz fluorescente. Se mide el producto de PCR fluorescente a tiempo real en cada uno de los 35 a 40 ciclos. La Figura 23 ilustra cómo se verían las reacciones si se midiera en cada ciclo a tiempo real y el C_T de la curva

a tiempo real se representa por la línea negra perpendicular.

La Figura 23a ilustra que el cociente de NT/CT presente al inicio de la PCR permanece (sustancialmente) constante a lo largo de la PCR hasta el punto final. Como se describe anteriormente, no es necesario monitorizar la reacción de amplificación a tiempo real para asegurar que la reacción está en la fase logarítmica lineal en algunas realizaciones de la presente invención.

La Figura 23a ilustra un experimento que usa una primera muestra de un primer espécimen. En la primera muestra, existen un número aproximadamente equivalente de moléculas de NT y CT del segundo ácido nucleico presentes al inicio de la reacción PCR (por ejemplo, como se describe anteriormente, cuando se usa una dilución de ADNc equilibrada). Por tanto, después de la electroforesis del producto amplificado del segundo ácido nucleico, las bandas de NT y CT son aproximadamente equivalentes, y durante la medida a tiempo real, la intensidad fluorescente para NT será aproximadamente la misma que para CT. El cociente de NT/CT es el mismo en un ciclo temprano que en un ciclo tardío (punto final), aunque la intensidad de banda tanto para NT como CT es baja en el ciclo temprano en comparación con el ciclo tardío. De forma similar, la banda de NT y la banda de CT del primer ácido nucleico son aproximadamente equivalentes, y el valor a tiempo real para NT es aproximadamente el mismo que para CT. La ΔC_T entre el segundo y el primer ácidos nucleicos en las medidas a tiempo real es de aproximadamente 10.

La Figura 23b ilustra además los controles de carga de una muestra a otra. En la Figura 23b, se reanaliza el primer espécimen usando una menor cantidad de partida de ácido nucleico, por ejemplo, menos ADNc cargado, debido a una variación en el pipeteado, por ejemplo, en la toma de alícuotas de una segunda muestra del primer espécimen en un recipiente diferente. El cociente de NT/CT para el segundo ácido nucleico es menor. Sin embargo, debido a que la concentración relativa de los moldes competitivos está fijada y a que la representación relativa de cada ácido nucleico está fijada, el cociente de NT/CT para el primer ácido nucleico cae proporcionalmente. Por consiguiente, el "cociente de cocientes" (cociente de posibilidades) de NT/CT del primer ácido nucleico dividido entre el NT/CT del segundo ácido nucleico permanece igual que en la Figura 23a. En este caso, la ΔC_T del análisis a tiempo real está también inalterada.

La Figura 23c ilustra los controles de carga y variación de la eficacia de amplificación. En la Figura 23c, se reanaliza de nuevo el primer espécimen, pero tanto con (1) una mayor cantidad de ADNc cargado debido a variación en el pipeteado (que conduce a una variación de la cantidad de partida de molde nativo) como (2) una eficacia de amplificación reducida del segundo ácido nucleico, como podría causar un inhibidor en el pocillo, que afecta a la amplificación de este ácido nucleico más que al otro, o concentraciones inapropiadas de cebadores para el segundo ácido nucleico.

La Figura 23c ilustra que, con medidas a tiempo real, esto reduce la ΔC_T de 10 a 6, y el valor del primer ácido nucleico es inapropiadamente alto. En medidas a tiempo real, la inhibición selectiva del gen está asociada a una ΔC_T reducida y a una medida errónea.

En contraposición, usando ciertas realizaciones descritas en la presente memoria, debido a que la eficacia de amplificación de los NT para cada uno de los dos ácidos nucleicos se afecta de la misma manera que su correspondiente CT, el cociente de NT/CT está inalterado en las Figuras 23a y 23c para el primer o el segundo ácidos nucleicos. También, con la mayor cantidad de ADNc cargada, el cociente de NT/CT del primer ácido nucleico y el cociente de NT/CT del segundo ácido nucleico aumentan proporcionalmente. En consecuencia, el "cociente de cocientes" (cociente de posibilidades) de NT/CT del primer ácido nucleico dividido entre NT/CT del segundo ácido nucleico queda igual entre las Figuras 23a y 23c.

La Figura 23d ilustra además controles para cargar una muestra de un segundo espécimen, en que el primer ácido nucleico está más altamente expresado. Aunque el primer ácido nucleico se expresa a un mayor nivel en comparación con el segundo ácido nucleico, las medidas a tiempo real dan una ΔC_T de aproximadamente 7.

En contraposición, usando ciertas realizaciones descritas en la presente memoria, el cociente de cocientes indica el mayor nivel de expresión. Ya que se carga menos ADNc en la reacción PCR, hay menos copias del segundo ácido nucleico de NT que copias de CT presentes al inicio de la reacción PCR, en comparación con la Figura 23a. A lo largo de la medida a tiempo real, el valor de fluorescencia de NT es menor que el de CT y, al final de la PCR, la banda de NT del segundo ácido nucleico sigue siendo menor que la banda de CT. Sin embargo, aunque se cargó menos ADNc en la reacción PCR en comparación con la primera muestra, la banda de NT del primer ácido nucleico es más densa que la banda de CT del primer ácido nucleico debido a su mayor expresión, y el valor de fluorescencia de NT del primer ácido nucleico durante la medida a tiempo real es mayor a lo largo de la PCR. En consecuencia, el "cociente de cocientes" (cociente de posibilidades) de NT/CT del primer ácido nucleico dividido entre NT/CT del segundo ácido nucleico proporciona un mayor valor en la Figura 23d que en la Figura 23a.

Por tanto, la RT-PCR a tiempo real puede controlar la carga midiendo el primer y segundo ácidos nucleicos en la misma reacción PCR (Figuras 23a, 23b, 23d). Los C_T (para cada ácido nucleico representado por una línea negra que interseca con el eje X) del primer y segundo ácidos nucleicos podrían variar ambos de un experimento a otro, pero la ΔC_T no varía. Sin embargo, el tiempo real no controla la variación en presencia de inhibidores ni la calidad de los reactivos de PCR.

II. Procedimientos de preparación de composiciones para valorar ácido nucleico

Se describen también en la presente memoria procedimientos para preparar composiciones para valorar un ácido nucleico en una muestra.

A. Preparación de mezclas estandarizadas

5 Se describe también en la presente memoria un procedimiento para preparar una mezcla estandarizada de reactivos. Como se usa en la presente memoria, "reactivo" puede hacer referencia a un componente usado en una mezcla, incluyendo disolvente y/o soluto. Por ejemplo, los reactivos incluyen ácidos nucleicos y/o agua, por ejemplo en el caso de mezclas acuosas. Se describe también que la mezcla estandarizada de reactivos comprende suficientes cantidades de molde competitivo para valorar las cantidades de un número de ácidos nucleicos en un número de muestras, por ejemplo, más de aproximadamente 10^6 muestras. Se describe además que la mezcla estandarizada permite la comparación directa de las cantidades entre al menos 2 de las muestras. Preferiblemente, permite la comparación directa de cantidades valoradas en al menos aproximadamente 5 muestras, al menos aproximadamente 10 muestras, al menos aproximadamente 50 muestras, al menos aproximadamente 100 muestras, al menos aproximadamente 500 muestras, al menos aproximadamente 1.000 muestras, al menos aproximadamente 5.000 muestras, al menos aproximadamente 10.000 muestras, al menos aproximadamente 50.000 muestras, al menos aproximadamente 100.000 muestras, al menos aproximadamente 500.000 muestras, al menos aproximadamente 1.000.000 muestras, al menos aproximadamente 5.000.000 muestras o al menos aproximadamente 10.000.000 muestras. Se describe además que la mezcla estandarizada permite la comparación directa de cantidades valoradas en un número ilimitado de muestras.

20 Se describe además en la presente memoria que la mezcla estandarizada comprende suficientes reactivos para valorar cantidades de un ácido nucleico. La mezcla estandarizada comprende suficientes reactivos para valorar cantidades de más de 1 ácido nucleico, por ejemplo, al menos aproximadamente 50, al menos aproximadamente 96, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 300, al menos aproximadamente 500, al menos aproximadamente 800, al menos aproximadamente 1.000 o al menos aproximadamente 5.000, al menos aproximadamente 10.000, al menos aproximadamente 50.000 o al menos aproximadamente 100.000 ácidos nucleicos. Se describe además que la mezcla estandarizada comprende suficientes reactivos para valorar cantidades de menos de aproximadamente 100.000, menos de aproximadamente 500.000 o menos de aproximadamente 1.000.000 ácidos nucleicos. Se describe también que diferentes ácidos nucleicos corresponden a diferentes transcritos génicos.

30 Se describe además en la presente memoria que los reactivos incluyen al menos un cebador de codificación y/o un cebador inverso capaces de cebar la amplificación de un molde competitivo en la mezcla. Se describe también que al menos un molde competitivo, cebador de codificación y/o cebador inverso comprende una secuencia referenciada en la Tabla 4, ilustrada en la Figura 4.

35 Se describe además que se diseñan un cebador de codificación y/o un cebador inverso para que tengan sustancialmente la misma temperatura de asociación que otro cebador de codificación y/o cebador inverso en la mezcla estandarizada. Diseñar cebadores con la misma o sustancialmente la misma temperatura de asociación puede permitir a las reacciones de amplificación conseguir aproximadamente la misma eficacia de amplificación en condiciones idénticas o sustancialmente idénticas. En dichos casos, si hay variación de la eficacia de amplificación, la eficacia de amplificación de un ácido nucleico y su molde competitivo pueden afectarse de forma idéntica (o sustancialmente idéntica), de modo que el cociente de producto amplificado del ácido nucleico y su correspondiente molde competitivo puede no variar o puede no variar sustancialmente. Se describe además que un cebador de codificación e inverso tienen la misma o sustancialmente la misma temperatura de asociación que cada uno de los otros cebadores de codificación e inversos en una mezcla estandarizada dada. Por ejemplo, la temperatura de asociación puede ser de aproximadamente 40 °C, aproximadamente 44 °C, aproximadamente 50 °C, aproximadamente 55 °C, aproximadamente 57 °C, aproximadamente 58 °C, aproximadamente 59 °C, aproximadamente 60 °C, aproximadamente 65 °C, aproximadamente 70 °C, aproximadamente 75 °C o aproximadamente 85 °C.

50 Se describe también que puede prepararse un molde competitivo patrón interno para evaluar un número de ácidos nucleicos, incluyendo ácidos nucleicos que pueden servir como uno o más ácidos nucleicos de referencia. Los moldes competitivos pueden clonarse entonces generando cantidades de ácido nucleico suficientes para valorar en más de aproximadamente 10^4 muestras, en más de aproximadamente 10^5 muestras, en más de aproximadamente 10^6 muestras, en más de aproximadamente 10^7 muestras, en más de aproximadamente 10^8 muestras, en más de aproximadamente 10^9 muestras, en más de aproximadamente 10^{10} muestras, en más de aproximadamente 10^{11} muestras, en más de aproximadamente 10^{12} muestras, en más de aproximadamente 10^{13} muestras, en más de aproximadamente 10^{14} muestras o en más de aproximadamente 10^{15} muestras.

Los moldes competitivos pueden cuantificarse cuidadosamente y mezclarse entonces conjuntamente formando una mezcla estandarizada. En algunas realizaciones, el cebador de codificación y/o cebador inverso pueden seleccionarse para que permitan la detección del ácido nucleico para medir aproximadamente 10^{-10} , aproximadamente 10^{-11} , aproximadamente 10^{-12} , aproximadamente 10^{-13} , aproximadamente 10^{-14} , aproximadamente

10^{-15} , aproximadamente 10^{-16} , aproximadamente 10^{-17} , aproximadamente 10^{-18} M o menos. Por ejemplo, el cebador de codificación y/o cebador inverso puede permitir la detección de aproximadamente 600 moléculas, aproximadamente 60 moléculas o aproximadamente 6 moléculas de ácido nucleico.

5 Se describe también que una mezcla estandarizada de la presente invención puede medir y/o dar cuenta de menos de aproximadamente 1.000 moléculas de ácido nucleico en una muestra, por ejemplo aproximadamente 800, aproximadamente 600 o aproximadamente 400 moléculas. Se describe además que pueden medirse y/o dar cuenta de menos de aproximadamente 100 moléculas (por ejemplo aproximadamente 60 moléculas), preferiblemente menos de aproximadamente 10 moléculas (por ejemplo aproximadamente 6 moléculas) o más preferiblemente menos de aproximadamente 1 molécula de un ácido nucleico en una muestra. En algunas realizaciones, una mezcla
10 estandarizada de la presente invención puede medir y/o dar cuenta de menos de aproximadamente 10.000.000, menos de aproximadamente 5.000.000, menos de aproximadamente 1.000.000, menos de aproximadamente 500.000, menos de aproximadamente 100.000, menos de aproximadamente 50.000, menos de aproximadamente 10.000, menos de aproximadamente 8.000, menos de aproximadamente 6.000, menos de aproximadamente 5.000 o menos de aproximadamente 4.000 moléculas de ácido nucleico en una muestra.

15 Además, se describe en la presente memoria que los reactivos para medir cantidades de ácidos nucleicos son estables. Por ejemplo, los cebadores y/o moldes competitivos de una mezcla estandarizada pueden comprender moléculas de ácido nucleico estables tales como ADN. Los reactivos pueden ser estables durante al menos aproximadamente 20 años, al menos aproximadamente 50 años, al menos aproximadamente 100 años, al menos aproximadamente 500 años o al menos aproximadamente 1.000 años. Además, se describe que una mezcla
20 estandarizada de la presente invención puede proporcionar reactivos para medir suficientes ácidos nucleicos correspondientes a las medidas de expresión génica que se esperan realizar durante al menos aproximadamente 20 años, al menos aproximadamente 50 años, al menos aproximadamente 100 años, al menos aproximadamente 500 años o al menos aproximadamente 1.000 años, por ejemplo a la tasa actual de medida de la expresión génica (que se estima que es de aproximadamente mil millones de ensayos al año ("Una previsión económica del mercado de expresión génica". <http://www.researchandmarkets.com/reports/5545>)).
25

Se describe también que el almacenamiento a largo plazo de reactivos y/o muestras que comprenden ADN puede conseguirse a -20 °C. Además, se describe que los reactivos y/o muestras que comprenden ARN son estables durante años congelados en forma de precipitado con EtOH y/o en agua exenta de ARNasa. Se describe además que los moldes competitivos se congelan establemente durante más de 6 años. Se describe adicionalmente que las
30 muestras de ADNc son estables durante más de 2 años congeladas a -20 °C.

Puede prepararse una mezcla estandarizada según algunas realizaciones de la presente invención efectuando uno o más de los procedimientos descritos en la presente memoria. Por ejemplo, como se describe anteriormente, usando una mezcla estandarizada puede valorarse un ácido nucleico respecto a uno o más de otros ácidos nucleicos (por ejemplo, que pueden servir como controles para el ADNc cargado en la reacción). También como se
35 detalla anteriormente, puede valorarse un ácido nucleico respecto a su molde competitivo respectivo proporcionado en la mezcla estandarizada.

Se describe también en la presente memoria que la mezcla estandarizada puede permitir la detección con una o más de las sensibilidades, una o más de las precisiones, uno o más de los límites de detección y/o con uno o más de los coeficientes de variación enseñados en la presente memoria. Resultarán evidentes rasgos adicionales de la
40 mezcla estandarizada preparada para un especialista en la materia basándose en la divulgación de la presente memoria.

B. Preparación de series de mezclas estandarizadas diluidas en serie

Se describe además en la presente memoria un procedimiento para preparar una serie de mezclas estandarizadas diluidas en serie. Se describe además que una o más de las series de mezclas estandarizadas comprende
45 cantidades suficientes de moldes competitivos para valorar las cantidades de un número de ácidos nucleicos en un número de muestras, por ejemplo, más de aproximadamente 10^6 muestras. Se describe adicionalmente que la mezcla estandarizada permite la comparación directa de cantidades entre al menos 2 de las muestras. Preferiblemente, permite la comparación directa de cantidades valoradas en al menos aproximadamente 5 muestras, al menos aproximadamente 10 muestras, al menos aproximadamente 50 muestras, al menos aproximadamente 100
50 muestras, al menos aproximadamente 500 muestras, al menos aproximadamente 1.000 muestras, al menos aproximadamente 5.000 muestras, al menos aproximadamente 10.000 muestras, al menos aproximadamente 50.000 muestras, al menos aproximadamente 100.000 muestras, al menos aproximadamente 500.000 muestras, al menos aproximadamente 1.000.000 muestras, al menos aproximadamente 5.000.000 muestras o al menos aproximadamente 10.000.000 muestras. Se describe además en la presente memoria que la mezcla estandarizada
55 permite la comparación directa de cantidades valoradas en un número ilimitado de muestras.

La serie de mezclas estandarizadas diluidas en serie puede obtenerse mediante dilución en serie de una mezcla estandarizada, por ejemplo una mezcla estandarizada descrita anteriormente. Por ejemplo, una o más de las series puede contener suficientes reactivos para valorar diversos números de ácidos nucleicos y/o para valorar diversos números de muestras, por ejemplo como se detalla anteriormente. De forma similar, una o más de las series de

mezclas estandarizadas diluidas en serie pueden comprender cualquiera de los reactivos de algunas realizaciones de las mezclas estandarizadas descritas anteriormente.

Se describe también que se diluye una mezcla estandarizada de modo que el molde competitivo de un primer ácido nucleico esté a una serie de concentraciones respecto al molde competitivo de un segundo ácido nucleico. Se describe además que se diluye en serie una mezcla estandarizada 10 veces, proporcionando diluciones en serie de 10 veces del molde competitivo del primer ácido nucleico respecto al molde competitivo del segundo ácido nucleico. Se describe también en la presente memoria que al menos dos de las series de concentraciones se extienden aproximadamente 1 orden de magnitud, aproximadamente 2 órdenes de magnitud, aproximadamente 3 órdenes de magnitud, aproximadamente 4 órdenes de magnitud, aproximadamente 5 órdenes de magnitud, aproximadamente 6 órdenes de magnitud, aproximadamente 7 órdenes de magnitud o más. Se describe además en la presente memoria que la serie de concentraciones incluye al menos 2, al menos 3, al menos 4, al menos 5 o 6 concentraciones seleccionadas de aproximadamente 10^{-10} M, aproximadamente 10^{-11} M, aproximadamente 10^{-12} M, aproximadamente 10^{-13} M, aproximadamente 10^{-14} M, aproximadamente 10^{-15} M y aproximadamente 10^{-16} M.

Se describe también en la presente memoria que una o más de las series de mezclas estandarizadas pueden permitir la detección con una o más de las sensibilidades, una o más de las precisiones, uno o más de los límites de detección y/o con uno o más de los coeficientes de variación enseñados en la presente memoria, dentro de diversos intervalos de órdenes de magnitud, por ejemplo, cualquiera de los órdenes de magnitud discutidos en la presente memoria.

III. Composiciones para valorar ácido nucleico

Se describen también en la presente memoria composiciones para valorar un ácido nucleico en una muestra, por ejemplo, composiciones que comprenden una mezcla estandarizada o una serie de mezclas estandarizadas diluidas en serie, por ejemplo, como se describe anteriormente. Además, se describen bases de datos, por ejemplo, bases de datos que comprenden datos obtenidos usando algunas realizaciones de los procedimientos de la presente invención.

A. Base de datos de valores numéricos

Por tanto, se describe una base de datos de valores numéricos correspondientes a cantidades de un primer ácido nucleico en un número de muestras.

Se describe en la presente memoria que los valores numéricos son directamente comparables entre un número de muestras. Por ejemplo, los valores numéricos son directamente comparables entre al menos aproximadamente 5 muestras, al menos aproximadamente 10 muestras, al menos aproximadamente 50 muestras, al menos aproximadamente 100 muestras, al menos aproximadamente 500 muestras, al menos aproximadamente 1.000 muestras, al menos aproximadamente 5.000 muestras, al menos aproximadamente 10.000 muestras, al menos aproximadamente 50.000 muestras, al menos aproximadamente 100.000 muestras, al menos aproximadamente 500.000 muestras, al menos aproximadamente 1.000.000 muestras, al menos aproximadamente 5.000.000 muestras o al menos aproximadamente 10.000.000 muestras. Se describe también que la comparación directa implica comparar los valores numéricos con otro sin usar un recurso bioinformático. Se describe además que puede usarse un recurso bioinformático, por ejemplo un recurso bioinformático sencillo.

La Figura 24 ilustra el desarrollo y uso de una base de datos de valores numéricos. En la etapa 2401, se obtienen las cantidades medidas mediante cualquier procedimiento de las diversas realizaciones de la presente invención descritas en la presente memoria, proporcionando valores numéricos. Por ejemplo, como ilustra la etapa 2401, puede valorarse un ácido nucleico respecto a un número conocido de moléculas de molde competitivo del ácido nucleico que se han combinado en una mezcla estandarizada. Dichas realizaciones pueden facilitar el informe de medida de ácido nucleico como valor numérico. Por ejemplo, el valor numérico puede obtenerse calculando un "cociente de cocientes" como se describe anteriormente. Se describe allí que cada valor en la base de datos se ha realizado respecto a un patrón interno dentro de una mezcla estandarizada de patrones internos.

Se describe además en la presente memoria que los valores numéricos corresponden a números de moléculas de un ácido nucleico dado en una muestra. Se describe además que los valores numéricos pueden proporcionarse en unidades de (moléculas de un primer ácido nucleico)/(moléculas de un segundo ácido nucleico), por ejemplo, en que el segundo ácido nucleico sirve como ácido nucleico de referencia. Se describe además en la presente memoria que las medidas se proporcionan en unidades de (moléculas de ADNc de un primer ácido nucleico)/(10^6 moléculas de ADNc de un segundo ácido nucleico). Los valores numéricos en algunas realizaciones, por ejemplo, pueden corresponder a menos de aproximadamente 1.000 moléculas de ácido nucleico en una muestra, por ejemplo, a aproximadamente 800, a aproximadamente 600 o a aproximadamente 400 moléculas. Además, se describe que los valores numéricos pueden corresponder a menos de aproximadamente 100 moléculas (por ejemplo, a aproximadamente 60 moléculas), menos de aproximadamente 10 moléculas (por ejemplo, a aproximadamente 6 moléculas) o menos de aproximadamente 1 molécula de un ácido nucleico en una muestra. Se describe además que los valores numéricos pueden corresponder a menos de aproximadamente 10.000.000, menos de

aproximadamente 5.000.000, menos de aproximadamente 1.000.000, menos de aproximadamente 500.000, menos de aproximadamente 100.000, menos de aproximadamente 50.000, menos de aproximadamente 10.000, menos de aproximadamente 8.000, menos de aproximadamente 6.000, menos de aproximadamente 5.000 o menos de aproximadamente 4.000 moléculas de un ácido nucleico en una muestra.

5 La base de datos descrita en la presente memoria puede comprender valores numéricos que varían dentro de un intervalo. Por ejemplo, los valores numéricos pueden variar dentro de un intervalo de menos de aproximadamente 1 orden de magnitud, más de aproximadamente 1 orden de magnitud o más de aproximadamente 2 órdenes de magnitud. Se describe además en la presente memoria que los valores numéricos de las cantidades medidas de los diferentes ácidos nucleicos, por ejemplo, los niveles de ARNm expresado a partir de dos o más genes diferentes, 10 pueden variar dentro de un intervalo de aproximadamente 3 o más órdenes de magnitud, aproximadamente 4 o más órdenes de magnitud, aproximadamente 5 o más órdenes de magnitud, aproximadamente 6 o más órdenes de magnitud o aproximadamente 7 o más órdenes de magnitud, por ejemplo, extendiéndose el intervalo de aproximadamente 7 unidades logarítmicas de expresión génica a aproximadamente 10^{-3} , aproximadamente 10^{-2} , aproximadamente 0,1, aproximadamente 1, aproximadamente 10, aproximadamente 10^2 , aproximadamente 10^3 y 15 aproximadamente 10^4 copias/célula. Se describe también en la presente memoria que los valores numéricos de las cantidades medidas de diferentes ácidos nucleicos pueden variar dentro de un intervalo de aproximadamente 8 o más, aproximadamente 9 o más o aproximadamente 10 o más órdenes de magnitud, por ejemplo, extendiéndose el intervalo de aproximadamente 10 unidades logarítmicas de expresión génica a aproximadamente 10^{-3} , aproximadamente 10^{-2} , aproximadamente 0,1, aproximadamente 1, aproximadamente 10, aproximadamente 10^2 , 20 aproximadamente 10^3 , aproximadamente 10^4 , aproximadamente 10^5 o aproximadamente 10^6 copias/célula. Dichos intervalos de expresión génica pueden ser importantes para detectar agentes de guerra biológica, por ejemplo.

Se describe además que los valores numéricos de la base de datos corresponden a una diferencia de menos de aproximadamente 2 veces en un ácido nucleico entre 2 de las muestras. Se describe también que los valores numéricos corresponden a una diferencia de menos de aproximadamente 1 vez, una diferencia de menos de 25 aproximadamente un 80 %, una diferencia de menos de aproximadamente un 50 %, una diferencia de menos de aproximadamente un 30 %, una diferencia de menos de aproximadamente un 20 %, una diferencia de menos de aproximadamente un 10 %, una diferencia de menos de aproximadamente un 5 % o una diferencia de menos de aproximadamente un 1 %.

30 Sin limitarse a una hipótesis y/o teoría dada, puesto que los datos en algunas realizaciones están estandarizados frente a una mezcla común de moldes competitivos patrones internos, son posibles comparaciones directas. Por ejemplo, como se discute anteriormente, en algunas realizaciones los valores numéricos son directamente comparables entre un número de muestras, por ejemplo, muestras obtenidas a partir de diferentes sujetos y/o de diferentes especies. En algunas realizaciones, los valores numéricos son directamente comparables entre un número de muestras medidas en diferentes laboratorios y/o en diferentes momentos. En realizaciones preferidas, 35 dichas comparaciones son posibles sin el uso de una muestra calibradora (por ejemplo, una muestra calibradora no renovable).

Puede describirse que dos valores son "directamente comparables" cuando, por ejemplo, los valores numéricos de cada uno describen las cantidades respecto a un patrón común. Como analogía fácilmente comprensible, 10 °C es directamente comparable con 50 °C, ya que ambos valores se proporcionan respecto al punto de ebullición del agua 40 (100 °C). Usando algunas realizaciones proporcionadas en la presente memoria, se mide el número de moléculas de ADNc que representa un gen en una muestra dada respecto a su correspondiente molde competitivo en una mezcla estandarizada, en lugar de comparándolo con otra muestra. El uso de una mezcla estandarizada común puede proporcionar el patrón común y puede facilitar las comparaciones directas.

45 En contraposición, al usar técnicas tales como RT-PCR a tiempo real y/o análisis de micromatrices (distintas de aquellas en combinación con algunas realizaciones de la presente invención), los ácidos nucleicos que se están midiendo se escalan diferentemente. Por ejemplo, las diferencias en las temperaturas de fusión por hibridación entre ADNc y polinucleótidos unidos (micromatrices) o sondas fluorescentes (RT-PCR a tiempo real) causan que las medidas se escalen diferentemente. Por consiguiente, las cantidades relativas de diferentes ácidos nucleicos en un espécimen y/o entre especímenes pueden no ser directamente comparables, por ejemplo, puede no ser posible 50 comparar la diferencia de expresión entre muchos genes en una muestra. Además, la RT-PCR a tiempo real y/o análisis de micromatrices pueden no proporcionar información directa en cuanto al número de moléculas de ácido nucleico presentes en una muestra.

Las cantidades valoradas pueden corregirse también por una o más fuentes de variación, por ejemplo, de acuerdo con diversas realizaciones de las enseñanzas proporcionadas en la presente memoria. Se describe además que los valores en la base de datos muestran un coeficiente de variación de menos de aproximadamente un 50 %, menos de aproximadamente un 30 %, menos de aproximadamente un 25 %, menos de aproximadamente un 20 %, menos de aproximadamente un 15 %, menos de aproximadamente un 10 %, menos de aproximadamente un 5 % o menos de aproximadamente un 1 % entre 2 o más muestras. Se describe también que los valores numéricos no comprenden un número estadísticamente significativo de falsos positivos. Además, se describe también que los valores numéricos no comprenden un número estadísticamente significativo de falsos negativos. Además, se describe que los valores numéricos no comprenden falsos negativos.

Se describe también que la base de datos comprende además valores numéricos correspondientes a cantidades de un número de otro(s) ácido(s) nucleico(s) en las muestras, en que dichas cantidades son directamente comparables. El número de otros ácidos nucleicos para los que están incluidos datos en la base de datos puede ser de al menos aproximadamente 5, al menos aproximadamente 10, al menos aproximadamente 50, al menos aproximadamente 100, al menos aproximadamente 500, al menos aproximadamente 1.000, al menos aproximadamente 5.000, al menos aproximadamente 10.000, al menos aproximadamente 50.000, al menos aproximadamente 100.000, al menos aproximadamente 500.000, al menos aproximadamente 1.000.000, al menos aproximadamente 5.000.000 o al menos aproximadamente 10.000.000 otros ácidos nucleicos.

La base de datos descrita en la presente memoria puede servir como banco de datos común, en que las cantidades medidas de ácidos nucleicos (por ejemplo, medidas de expresión génica) se reseñan como valores numéricos que permiten la comparación interexperimento directa. La etapa 2402 ilustra dicha base de datos. Se describe además que la base de datos establece un experimento múltiple virtual en continua expansión (concretamente, pueden introducirse los datos de un número en continua expansión de ácidos nucleicos, muestras y/o especímenes en una base de datos dada y compararse directamente con los demás datos dentro de la base de datos). Esto puede conducir a aumentos sinérgicos de conocimiento, por ejemplo, conocimiento respecto a la relación entre los patrones de expresión génica y el fenotipo.

Se describe además en la presente memoria proporcionar un lenguaje común para la expresión génica. La expresión génica puede medirse a nivel de ARNm, proteína o funcional, preferiblemente a nivel de ARNm. Por ejemplo, se regula la expresión de ARNm principalmente mediante el número de transcritos disponibles para traducción. Debido a que la expresión de ARNm está relacionada principalmente con el número de copias, puede desarrollarse un patrón interno para cada gen y/o establecer una unidad común para la medida de expresión génica. Por ejemplo, en un estudio multiinstitucional, los datos generados por procedimientos discutidos en la presente memoria eran suficientemente sensibles y reproducibles para apoyar el desarrollo de una base de datos de expresión génica significativa, que sirve como lenguaje común para la expresión génica.

Además, se describe proporcionar un lenguaje común para la expresión génica entre especies. Por ejemplo, pueden identificarse cebadores que amplifican por PCR ácidos nucleicos correspondientes a genes tanto humanos como de ratón, por ejemplo, para al menos aproximadamente un 20 %, para al menos aproximadamente un 30 %, para al menos aproximadamente un 50 %, para al menos aproximadamente un 80 % o para al menos aproximadamente un 90 % de los genes comunes a ser humano y ratones. También pueden desarrollarse cebadores para obtener una aplicación más amplia entre especies, por ejemplo, para amplificar ácidos nucleicos correspondientes a dos o más especies diferentes. Por ejemplo, pueden identificarse cebadores que amplifican ácidos nucleicos correspondientes a dos o más de genes de ser humano, rata, cerdo, caballo, oveja, mono, planta, mosca de la fruta, pez, levadura, bacteria y/o virus.

Se describe también que la base de datos está basada en la web. Se describe además que la base de datos de la invención encuentra uso en investigación experimental, diagnósticos clínicos y/o desarrollo de fármacos. La etapa 2403 ilustra este uso. Por ejemplo, la base de datos puede usarse para estudios avanzados sobre rutas de control transcripcional y/o servir como base para investigación mecanística. Por ejemplo, los análisis bivariados de valores numéricos de expresión génica individual para genes de factor de transcripción y genes controlados por estos factores de transcripción pueden mejorar la comprensión de la regulación de la expresión génica. Esto puede aumentar el entendimiento del control de la expresión génica, por ejemplo, en células normales y malignas.

Se describe también que los valores numéricos de una base de datos descrita en la presente memoria pueden usarse en una, dos o más etapas del desarrollo de fármacos. Las etapas del desarrollo de fármacos pueden incluir, por ejemplo, cribado del fármaco diana, identificación del compuesto principal, evaluación preclínica (por ejemplo, bioensayo y/o estudio animal), ensayo clínico y tratamiento de pacientes. Dichas aplicaciones se describen con más detalle a continuación.

B. Base de datos de índices numéricos

Se describe además en la presente memoria una base de datos que comprende índices numéricos. Los índices numéricos pueden obtenerse mediante cómputo matemático de 2 o más valores numéricos, en que los valores numéricos corresponden a cantidades de ácidos nucleicos en un número de muestras. Se describe también que la base de datos incluye uno o más índices numéricos proporcionados en las Figuras 1, 2 y/o 4.

Se describe adicionalmente que los índices numéricos son directamente comparables entre las muestras. Por ejemplo, los índices numéricos son directamente comparables entre al menos aproximadamente 5 muestras, al menos aproximadamente 10 muestras, al menos aproximadamente 50 muestras, al menos aproximadamente 100 muestras, al menos aproximadamente 500 muestras, al menos aproximadamente 1.000 muestras, al menos aproximadamente 5.000 muestras, al menos aproximadamente 10.000 muestras, al menos aproximadamente 50.000 muestras, al menos aproximadamente 100.000 muestras, al menos aproximadamente 500.000 muestras, al menos aproximadamente 1.000.000 muestras, al menos aproximadamente 5.000.000 muestras o al menos aproximadamente 10.000.000 muestras. Se describe además que la comparación directa implica comparar los índices numéricos entre sí sin un recurso bioinformático. Se describe también que puede usarse un recurso

bioinformático, por ejemplo, un recurso bioinformático sencillo. Se describe además que cada medida en la base de datos se ha realizado respecto a un patrón interno dentro de una mezcla estandarizada de patrones internos.

Como se discute anteriormente, las medidas de ácido nucleico pueden reseñarse como valores numéricos. Los valores numéricos pueden combinarse en índices numéricos mediante cómputo matemático, proporcionando un índice numérico, por ejemplo, permitiendo la interacción matemática entre los valores numéricos. Por ejemplo, se calcula un índice numérico dividiendo un numerador entre un denominador, correspondiendo el numerador a la cantidad de uno de los dos ácidos nucleicos y correspondiendo el denominador a la cantidad del otro de los dos ácidos nucleicos. Se describe también que se calcula un índice numérico mediante una serie de una o más funciones matemáticas. Por ejemplo, puede calcularse un índice numérico mediante la fórmula $(\text{gen } 1 + \text{gen } 2)/(\text{gen } 3 - \text{gen } 4)$. Puede describirse un índice numérico como equilibrado, por ejemplo, cuando está computado por una fórmula que tiene iguales números de valores numéricos en el numerador que el en denominador. Se detallan a continuación procedimientos para obtener índices numéricos que indican un estado biológico, por ejemplo, que pueden actuar como biomarcadores correlacionándolos con un fenotipo de interés.

Se describe además que los índices numéricos son directamente comparables entre un número de muestras, por ejemplo, muestras obtenidas de diferentes sujetos y/o de diferentes especies. Se describe también que los índices numéricos son directamente comparables entre un número de muestras medidas y/o recontadas en diferentes laboratorios y/o en diferentes momentos.

Se describe además que la base de datos de la presente invención puede servir como banco de datos común, en que las cantidades medidas de ácidos nucleicos (por ejemplo, medidas de expresión génica) se combinan matemáticamente proporcionando índices numéricos que permiten la comparación interexperimento directa. Se describe también que la base de datos establece un experimento múltiple en continua expansión (concretamente, pueden usarse los datos de una serie de ácidos nucleicos, muestras y/o especímenes en continua expansión para calcular los índices numéricos que se introducen en una base de datos dada y se comparan directamente con otros datos dentro de la base de datos).

Como se discute anteriormente, cualquier ácido nucleico o combinación de ácidos nucleicos medida, incluyendo todos los ácidos nucleicos medidos, puede usarse como gen de referencia y los datos calculados usando un primer ácido nucleico de referencia pueden recalcularse respecto a los de otro ácido nucleico de referencia. En el caso de índices numéricos, la diferencia en el valor obtenido después de convertir desde un ácido nucleico de referencia a otro puede depender de cuántos valores numéricos estén en el numerador y cuántos estén en el denominador. Por ejemplo, cada valor numérico en un índice numérico puede convertirse en una nueva referencia calculando el índice. Por ejemplo, cuando hay números iguales de valores numéricos en el numerador y el denominador, la conversión a una nueva referencia puede no tener efecto sobre el índice numérico relativo entre muestras y/o especímenes.

En el caso de índices numéricos equilibrados en que los valores numéricos corresponden a medidas de expresión génica, el efecto de un ácido nucleico de referencia que varía en expresión de una muestra y/o espécimen a otro puede neutralizarse. Esto puede ocurrir también al hacer análisis bivariados. Se describe también que, por ejemplo, cuando hay números no iguales de valores numéricos en el numerador y el denominador, el índice numérico relativo entre muestras y/o especímenes puede cambiar de acuerdo con la diferencia en los valores numéricos relativos para los ácidos nucleicos de referencia entre las muestras y/o especímenes.

Se describe también que la base de datos está basada en la web. Se describe además que la base de datos de la invención encuentra uso en investigación experimental, diagnóstico clínico y/o desarrollo de fármacos. Por ejemplo, la base de datos puede usarse para estudios avanzados sobre rutas de control transcripcional y/o servir como base para investigación mecanística. Por ejemplo, al menos un índice numérico indica un estado biológico. Los índices numéricos pueden correlacionarse mejor con un estado biológico dado, por ejemplo un fenotipo dado, que un valor numérico correspondiente a un ácido nucleico individual (por ejemplo, a un gen individual). Por ejemplo, los índices numéricos de una base de datos descrita en la presente memoria pueden usarse en una, dos o más etapas del desarrollo de fármacos. Dichas aplicaciones se describen con más detalle a continuación.

IV. Aplicaciones

Se describen en la presente memoria además procedimientos de uso de valores y/o índices numéricos en investigación, diagnóstico y/u otras aplicaciones.

A. Identificación de biomarcadores

Por tanto, se describen procedimientos para la obtención de índices numéricos. Se describe además que el índice numérico obtenido indica un estado biológico. Un "estado biológico", como se usa en la presente memoria, puede hacer referencia a un estado fenotípico, por ejemplo, un fenotipo clínicamente relevante u otra condición metabólica de interés. Los estados biológicos pueden incluir, por ejemplo, un fenotipo patológico, una predisposición a un estado patológico o un estado no patológico; una respuesta a fármaco terapéutico o predisposición a dicha respuesta, una respuesta adversa a fármaco (por ejemplo toxicidad a fármaco) o predisposición a dicha respuesta, una resistencia a fármaco o predisposición a mostrar dicha resistencia, etc. Se describe además que el índice numérico obtenido puede actuar como biomarcador, por ejemplo, correlacionándolo con un fenotipo de interés. El

fármaco puede ser un fármaco antitumoral. Además, se describe que el uso de realizaciones de la presente invención descritas en la presente memoria puede proporcionar una medicina personalizada.

Se describe además que un procedimiento para obtener un índice numérico que indica un estado biológico comprende proporcionar 2 muestras correspondientes a cada uno de un primer estado biológico y un segundo estado biológico; medir y/o dar cuenta de la cantidad de cada uno de los 2 ácidos nucleicos en cada una de las 2 muestras; proporcionar las cantidades como valores numéricos que son directamente comparables entre un número de muestras; computar matemáticamente los valores numéricos correspondientes a cada uno del primer y segundo estados biológicos y determinar un cómputo matemático que discrimine los dos estados biológicos.

El primer y segundo estados biológicos como se usan en la presente memoria corresponden a dos estados biológicos para comparar, tal como dos estados fenotípicos para distinguir. Los ejemplos incluyen, por ejemplo, tejido no patológico (normal) frente a tejido patológico; cultivo que muestra respuesta a fármaco terapéutico frente a cultivo que muestra menos respuesta a fármaco terapéutico; sujeto que muestra respuesta adversa a fármaco frente a sujeto que muestra respuesta menos adversa; grupo tratado de sujetos frente a grupo no tratado de sujetos, etc.

Puede usarse un índice numérico que discrimina un estado biológico particular, por ejemplo, una afección patológica o metabólica, como biomarcador para la afección y/o afecciones dadas relacionadas con el mismo. Por ejemplo, el estado biológico indicado puede ser al menos uno de una afección relacionada con la angiogénesis, una afección relacionada con antioxidantes, una afección relacionada con la apoptosis, una afección relacionada con el sistema cardiovascular, una afección relacionada con el ciclo celular, una afección relacionada con la estructura celular, una afección relacionada con citocinas, una afección relacionada con la respuesta de defensa, una afección relacionada con el desarrollo, una afección relacionada con la diabetes, una afección relacionada con la diferenciación, una afección relacionada con la replicación y/o reparación de ADN, una afección relacionada con células endoteliales, una afección relacionada con receptores hormonales, una afección relacionada con receptores de folato, una afección relacionada con la inflamación, una afección relacionada con el metabolismo intermedio, una afección relacionada con el transporte de membrana, una afección relacionada con la neurotransmisión, una afección relacionada con el cáncer, una afección relacionada con el metabolismo oxidativo, una afección relacionada con la maduración de proteínas, una afección relacionada con la transducción de señal, una afección relacionada con la respuesta al estrés, una afección relacionada con la estructura de tejido, una afección relacionada con los factores de transcripción, una afección relacionada con el transporte o una afección relacionada con el metabolismo xenobiótico.

Por ejemplo, se han identificado los índices numéricos que indican cáncer de pulmón (E.L. Crawford, K. A. Warner, S. A. Khuder *et al.*, Biochem. Bioph. Res. Co. 293, 509-516 (2002); E.L. Crawford, S.A. Khuder, S.J. Durham *et al.*, Cancer Res. 60, 1609-1618 (2000); J.P. DeMuth, C.M. Jackson, D.A. Weaver *et al.*, Am. J. Respir. Cell Mol. Biol. 19, 18-24 (1998)), sarcoidosis pulmonar (M.G. Rots, R. Pieters. G.J. Peters *et al.*, Blood 94, 3121-3128 (1999)) fibrosis quística (J.T. Allen, R.A. Knight, CA. Bloor y M.A. Spiteri, Am. J. Respir. Cell. Mol. Biol. 21, 693-700 (1999)) y quimiorresistencia en leucemias infantiles (S. Mollerup, D. Ryberg, A. Hewer *et al.*, A. Cancer Res. 59, 3317-3320 (1999)). Además, se han evaluado genes de enzimas del metabolismo antioxidante y xenobiótico en células epiteliales bucales humanas; se ha asociado la expresión génica de células endoteliales microvasculares con la progresión de escleroderma; se ha estudiado la expresión de genes del transporte de membrana en modelos de insuficiencia cardíaca congestiva en rata; se ha estudiado la resistencia inmunitaria en tejidos primarios humanos; se ha estudiado el control de la transcripción de la expresión de receptores hormonales y se han asociado los patrones de expresión génica con resistencia a carboplatino y/o taxol en carcinoma de ovario y con resistencia a gemcitabina en múltiples tumores humanos. Otros ejemplos específicos incluyen, por ejemplo, identificación de los índices numéricos para predecir la sensibilidad del cáncer de colon ante 5-FU y para indicar una o más etapas diferentes del carcinoma de vejiga. Las realizaciones de las invenciones descritas en la presente memoria pueden acelerar el descubrimiento de asociaciones entre los patrones de expresión génica y estados biológicos de interés, conduciendo a mejores procedimientos para prevenir, diagnosticar y tratar diversas afecciones.

La Figura 25 ilustra el uso de índices numéricos en la identificación de un estado biológico.

La medición de las cantidades de ácido nucleico puede efectuarse mediante cualquier procedimiento conocido en la materia y/o descrito en la presente memoria. Preferiblemente, el procedimiento usado puede medir y/o dar cuenta de menos de aproximadamente 10.000 moléculas, menos de aproximadamente 8.000, menos de aproximadamente 6.000 o menos de aproximadamente 4.000, preferiblemente menos de aproximadamente 1.000, menos de aproximadamente 800, menos de aproximadamente 600 o menos de aproximadamente 400 moléculas de un ácido nucleico dado en una muestra dada. Se describe además que las medidas corresponden a medidas de expresión génica, por ejemplo, pueden medirse niveles de transcritos de ARNm. Se describe también que pueden usarse los niveles de transcritos, en particular los niveles de transcritos de 2 o más genes, para indicar un estado biológico. Por ejemplo, el análisis de micromatrices ha identificado conjuntos de genes que están asociados a estados patológicos y/o a respuestas a fármacos (D.A. Wigle, I. Jurisica, N. Radulovich *et al.*, Cancer Res. 62, 3005-3008 (2002); M.E. Garber, O.G. Troyanskaya. K. Schluens *et al.*, Proc. Natl. Acad. Sci. USA 98, 13784-13789 (2001); A. Bhattacharjee, W.G. Richards, J. Staunton *et al.*, Proc. Natl. Acad. Sci. USA 98, 13790-13795 (2001); I. Hedenfalk, D. Duggan, Y. Chen *et al.*, New Engl. J. Med. 344, 539-548 (2001); T. Sorlie, C.M. Perou, R. Tibshirani *et al.*, Proc. Natl. Acad. Sci. USA 98, 10869-10874 (2001); C.M. Perou, S.S. Jeffrey, M. van de Rijn *et al.*, Proc. Natl. Acad. Sci. USA 96, 9212-

9217 (1999)). Se logran proporcionar las cantidades medidas y/o recontadas como valores numéricos preferiblemente mediante procedimientos descritos en la presente memoria, en que los valores numéricos son directamente comparables para el número de muestras usadas.

Se describe además que uno o más de los ácidos nucleicos para medir están asociados con uno de los estados biológicos en un mayor grado que el (los) otro(s). Por ejemplo, uno o más de los ácidos nucleicos para evaluar están asociados a un primer estado biológico y no a un segundo estado biológico. Puede decirse que un ácido nucleico “está asociado a” un estado biológico particular cuando el ácido nucleico está asociado positiva o negativamente al estado biológico. Por ejemplo, puede decirse que un ácido nucleico está “asociado positivamente” a un primer estado biológico cuando el ácido nucleico aparece en mayores cantidades en un primer estado biológico en comparación con un segundo estado biológico. Como ilustración, puede decirse que los genes altamente expresados en células cancerosas en comparación con células no cancerosas están positivamente asociados al cáncer. Por otro lado, puede decirse que un ácido nucleico presente en menores cantidades en un primer estado biológico en comparación con un segundo estado biológico está negativamente asociado al primer estado biológico.

El ácido nucleico que medir y/o dar cuenta puede corresponder a un gen asociado a un fenotipo particular. La secuencia del ácido nucleico puede corresponder a las regiones transcritas, expresadas y/o reguladoras del gen (por ejemplo, una región reguladora de un factor de transcripción, por ejemplo, un factor de transcripción para coregulación).

Se describe también que se miden y usan las cantidades expresadas de más de 2 genes para proporcionar un índice numérico indicativo de un estado biológico. Por ejemplo, en algunos casos, se usan patrones de expresión de aproximadamente 50 a aproximadamente 100 genes para caracterizar un estado fenotípico dado, por ejemplo, un fenotipo clínicamente relevante. Véase, por ejemplo, Heldenfalk, I. *et al.* *NEJM* 344: 539, 2000. Se describe además que pueden medirse las cantidades expresadas de al menos aproximadamente 5 genes, al menos aproximadamente 10 genes, al menos aproximadamente 20 genes, al menos aproximadamente 50 genes o al menos aproximadamente 70 genes y usarse para proporcionar un índice numérico indicativo de un estado biológico. Se describe también que pueden medirse las cantidades expresadas de menos de aproximadamente 90 genes, menos de aproximadamente 100 genes, menos de aproximadamente 120 genes, menos de aproximadamente 150 genes o menos de aproximadamente 200 genes y usarse para proporcionar un índice numérico indicativo de un estado biológico. Los ejemplos específicos incluyen, por ejemplo, identificación de patrones de expresión génica asociados a cáncer de pulmón (Crawford, E. L. *et al.* “Normal bronchial epithelial cell expression of glutathione transferase P1, glutathione transferase M3, and glutathione peroxidase is low in subjects with bronchogenic carcinoma”. *Cancer Res.*, 60: 1609-1618, 2000; DeMuth, *et al.*, “The gene expression index c-myc x E2F-1/p21 is highly predictive of malignant phenotype in human bronchial epithelial cells”. *Am. J. Respir. Cell Mol. Biol.*, 19: 18-24, 1998); sarcoidosis pulmonar (Allen, J. T., *et al.*, “Enhanced insulin-like growth factor binding protein-related protein 2 (connective tissue growth factor) expression in patients with idiopathic pulmonary fibrosis and pulmonary sarcoidosis”. *Am. J. Respir. Cell Mol. Biol.*, 21: 693-700, 1999); fibrosis quística (Allen, *et al. supra*) y quimiorresistencia en leucemias infantiles (Rots, M. G., *et al.*, “Circumvention of methotrexate resistance in childhood leukemia subtypes by rationally designed antifolates”. *Blood*, 94(9): 3121-3128, 1999; Rots, M. G., *et al.*, “mRNA expression levels of methotrexate resistance-related proteins in childhood leukemia as determined by a competitive template-based RT-PCR method”. *Leukemia*, 14: 2166-2175 (2000)).

Computar matemáticamente los valores numéricos puede hacer referencia a usar cualquier ecuación, operación, fórmula y/o regla para interaccionar valores numéricos, por ejemplo, una suma, diferencia, producto, cociente, potencia logarítmica y/u otro cómputo matemático. Como se describe anteriormente, se calcula un índice numérico dividiendo un numerador entre un denominador, en que el numerador corresponde a una cantidad de un ácido nucleico y el denominador corresponde a una cantidad de otro ácido nucleico. Se describe además que el numerador corresponde a un gen asociado positivamente a un estado biológico dado y el denominador corresponde a un gen asociado negativamente al estado biológico. Se describe además que pueden usarse más de un gen asociado positivamente al estado biológico que se está evaluando y más de un gen asociado negativamente al estado biológico que se está evaluando. Por ejemplo, puede derivarse un índice numérico que comprende valores numéricos para los genes asociados positivamente en el numerador y valores numéricos para un número equivalente de genes asociados negativamente en el denominador. Como se menciona anteriormente, en dichos índices numéricos equilibrados, los valores numéricos del ácido nucleico de referencia se anulan. Un ejemplo de índice numérico equilibrado incluye un índice numérico para predecir resistencia antifolato entre leucemias infantiles (Rots, M. G., Willey, J. C., Jansen, G., *et al.* (2000) “mRNA expression levels of methotrexate resistance-related proteins in childhood leukemia as determined by a standardized competitive template-based RT-PCR method”. *Leukemia* 14, 2166-2175. Se describe también que los valores numéricos equilibrados pueden neutralizar los efectos de la variación en la expresión del (de los) gen(es) que proporciona(n) el (los) ácido(s) nucleico(s) de referencia. Se describe además que se calcula un índice numérico mediante una serie de una o más funciones matemáticas.

Puede conseguirse la determinación de cuál cómputo matemático usar para proporcionar un índice numérico indicativo de un estado biológico mediante cualquier procedimiento conocido en las técnicas, por ejemplo, en las técnicas matemática, estadística y/o computacional. Se describe también que determinar el cómputo matemático implica el uso de un software. Por ejemplo, puede usarse un software de aprendizaje automático.

- Se describe además que puede proporcionarse más de una muestra correspondiente a cada estado biológico. Por ejemplo, pueden proporcionarse al menos aproximadamente 5 muestras, al menos aproximadamente 10 muestras, al menos aproximadamente 50 muestras, al menos aproximadamente 100 muestras, al menos aproximadamente 500 muestras, al menos aproximadamente 1.000 muestras, al menos aproximadamente 5.000 muestras, al menos aproximadamente 10.000 muestras, al menos aproximadamente 50.000 muestras, al menos aproximadamente 100.000 muestras, al menos aproximadamente 500.000 muestras, al menos aproximadamente 1.000.000 muestras, al menos aproximadamente 5.000.000 muestras o al menos aproximadamente 10.000.000 muestras.
- Se describe además que pueden compararse más de 2 estados biológicos, por ejemplo distinguirse. Por ejemplo, pueden proporcionarse muestras de un intervalo de estados biológicos, por ejemplo, correspondientes a diferentes etapas de la progresión de la enfermedad, por ejemplo, diferentes etapas del cáncer. Las células en diferentes etapas del cáncer incluyen, por ejemplo, una célula no cancerosa frente a una célula cancerosa no metastatizante frente a una célula metastatizante de un paciente dado en diversos momentos durante el curso de la enfermedad. Pueden usarse células cancerosas de diversos tipos de cáncer incluyendo, por ejemplo, cáncer de vejiga, cáncer de hueso, cáncer de cerebro, cáncer de mama, cáncer de colon, cáncer del sistema endocrino, cáncer gastrointestinal, cáncer ginecológico, cáncer de cabeza y cuello, leucemia, cáncer de pulmón, linfoma, metástasis, mieloma, tejido neoplásico, cáncer pediátrico, cáncer de pene, cáncer de próstata, sarcoma, cáncer de piel, cáncer de testículo, cáncer tiroideo y cáncer del tracto urinario. Se describe también que pueden desarrollarse biomarcadores para predecir cuál agente quimioterapéutico puede funcionar mejor para un tipo dado de cáncer, por ejemplo, en un paciente particular.
- Una célula no cancerosa puede incluir una célula de hematoma y/o tejido cicatricial, así como parénquima morfológicamente normal de pacientes sin cáncer, por ejemplo, relacionada con pacientes sin cáncer o no relacionada con pacientes con cáncer. Las células no cancerosas pueden incluir también parénquima morfológicamente normal de pacientes con cáncer, por ejemplo, de un sitio cercano al sitio del cáncer en el mismo tejido y/o el mismo órgano, de un sitio alejado del sitio del cáncer, por ejemplo, en un diferente tejido y/u órgano en el mismo sistema de órganos o de un sitio aún más alejado, por ejemplo, en un diferente órgano y/o un diferente sistema de órganos.
- Los índices numéricos obtenidos pueden proporcionarse como una base de datos. Los índices numéricos y/o bases de datos de los mismos pueden encontrar uso en diagnósticos, por ejemplo, en el desarrollo y aplicación de ensayos clínicos, como se describe a continuación.
- B. Cribado de micromatrices**
- Se describen también en la presente memoria técnicas de cribado, por ejemplo, técnicas de cribado que usan hibridación en fase sólida tal como análisis de micromatrices. Por ejemplo, pueden identificarse patrones de expresión génica relevantes mediante cribado de expresión génica por micromatrices y pueden seguir a esto ensayos adecuados para el análisis de un subconjunto de genes.
- La Figura 26 ilustra el proceso global respecto al uso de cribados de micromatrices. La Figura 26a ilustra esquemáticamente el descubrimiento de genes de interés usando micromatrices. Puede cribarse una población de genes para determinar un subconjunto de genes de interés, por ejemplo, genes correspondientes a ácidos nucleicos asociados a un primer estado biológico, pero no a un segundo estado biológico. Por ejemplo, puede encontrarse que un subconjunto de que comprende aproximadamente 30, aproximadamente 50, aproximadamente 80, aproximadamente 100, aproximadamente 120, aproximadamente 150, aproximadamente 200, aproximadamente 250 o aproximadamente 300 genes está asociado a un fenotipo clínicamente relevante, por ejemplo, estados patológicos frente a no patológicos, o a cualquier otro estado biológico para distinguir, como se discute anteriormente. Los análisis de micromatrices usados pueden usar cualquier micromatriz y técnica de micromatrices conocidas en la materia y/o descritas en la presente memoria. Pueden evaluarse entonces uno o más de los ácidos nucleicos así identificados de acuerdo con algunas realizaciones descritas en la presente memoria.
- La Figura 26b ilustra esquemáticamente evaluaciones de genes de interés según algunas realizaciones descritas aquí. Brevemente, como ilustra la Figura 26b, puede extraerse el ARNm correspondiente a uno o más genes (por ejemplo 3 genes) y transcribirse de forma inversa, por ejemplo, como se discute con detalle anteriormente. De nuevo como se discute anteriormente, puede equilibrarse cuantitativamente una muestra de ADNc y combinarse con una mezcla estandarizada apropiada, por ejemplo, que comprende moldes competitivos de cada uno de los genes para evaluar. Los moldes nativos de cada uno de los 3 genes pueden coamplificarse con su correspondiente molde competitivo en un recipiente dado. La amplificación por PCR puede estar seguida por electroforesis, proporcionando un electroferograma. Pueden usarse las áreas bajo la curva para obtener un "cociente de cocientes" como se detalla anteriormente. Se proporcionan las medidas de la expresión de cada uno de los 3 genes como un valor numérico.
- Puede usarse cualquier otra realización o variación de estos procedimientos, por ejemplo como se dan a conocer en la presente memoria, por ejemplo, para permitir la detección con una o más sensibilidades, una o más precisiones, uno o más límites de detección y/o uno o más coeficientes de variación enseñados en la presente memoria. En realizaciones preferidas, los procedimientos empleados pueden mejorar el umbral de detección, sensibilidad y/o coeficiente de variación en comparación con micromatrices. Por ejemplo, el análisis que usa algunas realizaciones

de la presente invención puede evitar, reducir y/o controlar las diferencias en las temperaturas de fusión entre ADNc de cada gen y el oligonucleótido o ADNc localizado en la matriz; las diferencias en la cantidad de muestra cargada; el tiempo de hibridación; el rigor del lavado y/o parámetros usados para calibrar la intensidad de fluorescencia. Se proporcionan en el ejemplo V siguiente detalles de experimentos que comparan procedimientos de la presente invención con análisis de micromatrices.

Se describe además en la presente memoria que se evalúan ácidos nucleicos correspondientes a genes de interés en muestras correspondientes a uno o más estados biológicos. Por ejemplo, puede usarse una muestra correspondiente a un primer estado biológico y una muestra correspondiente a un segundo estado biológico. Se describe también que pueden evaluarse y/o dar cuenta de las cantidades de ácidos nucleicos correspondientes a cada uno de dichos dos estados biológicos, por ejemplo, proporcionando datos representativos de los dos estados biológicos.

La Figura 26c ilustra esquemáticamente el cómputo matemático de los valores numéricos obtenidos para los genes de interés. Pueden usarse los valores numéricos obtenidos de 2 o más ácidos nucleicos para determinar uno o más índices numéricos. Por ejemplo, de nuevo como se detalla anteriormente, pueden combinarse matemáticamente los valores numéricos correspondientes a cada uno de un primer y segundo estados biológicos. Puede determinarse un cómputo matemático que indica el estado biológico de interés, por ejemplo, discriminando el primer y segundo estados biológicos. La Figura 26c ilustra el uso de software para efectuar los cálculos matemáticos, de nuevo como se proporciona con detalle anteriormente.

La Figura 26d ilustra esquemáticamente el uso de dichos índice numéricos en un entorno clínico, por ejemplo, como biomarcador para diagnósticos y/o pronósticos, como se discute con más detalle a continuación. Cuando se analizan muestras clínicas que son de tamaño limitado, es probable que sea más económico medir solo aquellos genes que contribuyan a la información sobre los perfiles de expresión que definen el estado biológico de interés. En consecuencia, en lugar de medir la expresión de una gran población de genes, por ejemplo de aproximadamente 40.000 a aproximadamente 80.000 genes, puede evaluarse un subconjunto menor en muestras clínicas. Usando algunas realizaciones para evaluar ácidos nucleicos proporcionadas en la presente memoria, pueden hacerse medidas multigénicas en el subconjunto menor y usarse los datos en pruebas de biomarcador, por ejemplo, como índices numéricos indicativos del estado biológico.

C. Aplicaciones de diagnóstico

Se describe también en la presente memoria un procedimiento de identificación de un estado biológico. Se describe además que el procedimiento comprende medir y/o dar cuenta de una cantidad de cada uno de 2 ácidos nucleicos en una muestra, proporcionar las cantidades como valores numéricos, y usar los valores numéricos para proporcionar un índice numérico, con lo que el índice numérico indica el estado biológico. Además, se describe que el índice numérico comprende un índice numérico proporcionado en la Figuras 1, 2 y/o 4.

Puede determinarse un índice numérico que indica un estado biológico como se describe anteriormente. La muestra puede obtenerse a partir de un espécimen, por ejemplo un espécimen recogido de un sujeto para tratar. El sujeto puede estar en un entorno clínico incluyendo, por ejemplo, un hospital, un consultorio de un prestador de servicios médicos, clínica y/u otra instalación de servicios médicos y/o investigación. Pueden entonces medirse y/o dar cuenta de las cantidades de ácido(s) nucleico(s) de interés en la muestra.

La valoración de las cantidades de ácido nucleico puede efectuarse mediante cualquier procedimiento descrito en la presente memoria. Preferiblemente, el procedimiento usado puede medir y/o dar cuenta de menos de aproximadamente 10.000 moléculas, menos de aproximadamente 8.000, menos de aproximadamente 6.000 o menos de aproximadamente 4.000, preferiblemente menos de aproximadamente 1.000, menos de aproximadamente 800, menos de aproximadamente 600 o menos de aproximadamente 400 moléculas de un ácido nucleico dado en una muestra dada. En casos en que se van a medir varios genes en una muestra y/o espécimen, pueden practicarse realizaciones preferidas usando pequeñas cantidades de material celular de partida, por ejemplo, usando las cantidades de material obtenidas a partir de una muestra de biopsia de diagnóstico, por ejemplo, mediante los procedimientos descritos con más detalle anteriormente y/o conocidos en la materia. En realizaciones más preferidas, pueden evaluarse más de un gen al mismo tiempo, y en realizaciones altamente preferidas, cuando se van a evaluar un número dado de genes, pueden obtenerse simultáneamente los datos de expresión para ese número dado de genes. Por ejemplo, en algunas realizaciones, pueden ensayarse los datos obtenidos a partir de tejido de cáncer de pulmón primario. Al comparar el patrón de expresión de ciertos genes con los de una base de datos, puede determinarse el agente quimioterapéutico al que responderá más probablemente un tumor con ese patrón de expresión génica.

En algunas realizaciones, pueden usarse procedimientos de la invención para evaluar simultáneamente tanto un gen informador exógeno como un gen constitutivo endógeno, tal como ARN de GAPDH, en una célula transfectada, *in vitro* o *in vivo*. En algunas realizaciones, por ejemplo, pueden medirse las cantidades relativas por célula de gen exógeno regulador de la conductancia transmembrana de la fibrosis quística (CFTR). Aunque se ha reseñado que numerosas mutaciones diferentes en el gen CFTR están asociadas a la enfermedad, la mutación asociada a la enfermedad más común es una delección de 3 bases en la posición 508. Es posible preparar cebadores que den

como resultado la amplificación de uno u otro del gen 508 anormal eliminado o del gen CFTR normal usando los procedimientos descritos, por ejemplo, en Cha, R. S., Zarbl, H., Keohavong, P., Thilly, W. G., "Match amplification mutation assay (MAMA): application to the c-Ha ras gene, PCR methods and applications", 2:14-20 (1992). Estos pueden usarse con ciertas realizaciones de la presente invención para medir las cantidades de ácido nucleico de CFTR normal exógeno y/o las cantidades de gen CFTR mutante endógeno.

De forma similar, en algunas realizaciones pueden usarse procedimientos de la invención para cuantificar el gen de distrofina exógeno normal en presencia del gen endógeno mutado. En el caso de distrofina, la enfermedad es el resultado de deleciones relativamente grandes. Usando cebadores que cubran la región eliminada, puede amplificarse selectivamente y cuantificar la expresión de un gen normal transfectado y/o un gen constitutivo anormal de distrofina. Como se apreciará por los especialistas en la materia, pueden evaluarse también de manera similar otros genes asociados a otras enfermedades y/o afecciones.

En algunas realizaciones, pueden usarse los procedimientos descritos en la presente memoria para determinar los niveles de expresión normales, por ejemplo, proporcionando valores numéricos correspondientes a niveles de expresión de transcrito génico normales. Dichas realizaciones pueden usarse para indicar un estado biológico normal, al menos con respecto a la expresión del gen evaluado.

Niveles de expresión normales puede hacer referencia al nivel de expresión de un transcrito en condiciones no asociadas normalmente a una enfermedad, traumatismo y/u otro ataque celular. En algunas realizaciones, los niveles de expresión normales pueden proporcionarse como un número, o preferiblemente como un intervalo de valores numéricos correspondientes a un intervalo de expresión normal de un gen particular, por ejemplo dentro de \pm un porcentaje de error experimental. Puede compararse la comparación de un valor numérico obtenido para un ácido nucleico dado en una muestra, por ejemplo un ácido nucleico correspondiente a un gen particular, con valores numéricos normales establecidos, por ejemplo, mediante comparación con datos en la base de datos proporcionada en la presente memoria. Como los valores numéricos pueden indicar números de moléculas de ácido nucleico en la muestra, esta comparación puede indicar si el gen se está expresando a niveles normales o no.

Se describe además un procedimiento para identificar un estado biológico que comprende valorar una cantidad de ácido nucleico en una primera muestra y proporcionar dicha cantidad como valor numérico, en el que dicho valor numérico es directamente comparable entre un número de otras muestras. Se describe también que el valor numérico es directamente comparable con al menos aproximadamente 5, al menos aproximadamente 10, al menos aproximadamente 100, al menos aproximadamente 1.000, al menos aproximadamente 5.000 o al menos aproximadamente 10.000 otras muestras. Además, se describe que el valor numérico potencialmente es directamente comparable con un número ilimitado de otras muestras. Las muestras pueden evaluarse en diferentes momentos, por ejemplo, a diferentes días, en los mismos o diferentes experimentos en el mismo laboratorio y/o en diferentes experimentos en diferentes laboratorios.

Se describe también que el estado biológico corresponde a un nivel de expresión normal de un gen. Cuando el estado biológico no corresponde a niveles normales, por ejemplo cae fuera del intervalo deseado, puede indicar un estado no normal, por ejemplo una afección patológica, como se discute anteriormente.

EJEMPLOS

EJEMPLO I

El siguiente ejemplo compara un enfoque no "en dos etapas" con uno "en dos etapas" de acuerdo con algunas realizaciones de la presente invención.

Reactivos

Se obtuvo el tampón 10xPCR para Rapidcycler (Tris 500 mM, pH 8,3, BSA 2,5 mg/ μ l, MgCl₂ 30 mM) de Idaho Technology, Inc. (Idaho Falls, Id.). Se obtuvieron el tampón Thermo 10x (KCl 500 mM, Tris-HCl 100 mM, pH 9,0, 1,0 % de Triton X-100), polimerasa taq (5 U/ μ l), cebadores de oligo-dT, RNasin (25 U/ μ l), marcador de tamaño pGEM y dNTP de Promega (Madison, Wis.). Se obtuvieron transcriptasa inversa de M-MLV (200 U/ μ l) y tampón de primera hebra 5x (Tris-HCl 250 mM, pH 8,3, KCl 375 mM, MgCl₂ 15 mM, DTT 50 mM) de GibcoBRL (Gaithersburg, Md.). Se obtuvieron agarosa NuSieve y SeaKem LE de FMC BioProducts (Rockland, Me.). Se obtuvo TriReagent de Molecular Research Center (Cincinnati, Ohio). Se obtuvo agua exenta de ARNasa de Research Genetics (Huntsville, Ala.). Se obtuvo el kit DNA 7500 Assay que contiene tinte, matriz y patrones de Agilent Technologies (Palo Alto, Calif.). Se adquirió la línea celular de adenocarcinoma de pulmón A549 en la American Type Culture Collection (Rockville, Md.). Se obtuvo el medio de cultivo celular RPMI-1640 de Sigma (St. Louis, Mo.). Se obtuvo ARN de referencia humano universal de Stratagene (La Jolla, Calif.). Se sintetizaron a medida los cebadores oligonucleotídicos por Biosource International (Menlo Park, Calif.). Se proporcionaron amablemente los kits de expresión génica sistema 1 y sistema 1a de G.E.N.E. por Gene Express National Enterprises, Inc. (Huntsville, Ala.). Todos los demás productos químicos y reactivos eran de pureza de biología molecular.

Extracción de ARN y transcripción inversa

Se extrajo el ARN total de células crecidas en monocapa según el protocolo del fabricante TriReagent. Se precipitó el ARN de referencia humano universal según el protocolo del fabricante. Se transcribió de forma inversa aproximadamente 1 µg de ARN total usando transcriptasa inversa de M-MLV y un cebador de oligo-dT.

Enfoque no de dos etapas

- 5 Se efectuaron las medidas de expresión génica usando procedimientos publicados anteriormente (no de dos etapas) (véanse, por ejemplo, Willey, J. C. *et al.*, *Am. J. Respir. Cell Mol. Biol.* 19: 6-17,1998; Manual de instrucciones del sistema 1 de Gene Express, Gene Express National Enterprises, Inc. www.genexnat.com 2000) con el kit de expresión génica sistema 1 o sistema 1a de G.E.N.E. (Gene Express National Enterprises, Inc.). Brevemente, se preparó una mezcla maestra que contenía tampón, MgCl₂, dNTP, ADNc, mezcla de moldes competitivos (CT) del sistema 1 o sistema 1a de G.E.N.E. y polimerasa taq y se tomaron alícuotas en tubos, por ejemplo de una microplaca de 384 pocillos que contienen cebadores específicos de gen, y se ciclaron en un Rapidcycler (Idaho Technology, Inc.) o un ciclador térmico Primus HT Multiblock (MWG-BIOTECH, Inc., High Point, N.C.) o un termociclador de bloque PTC-100 con tapa calentada durante 35 ciclos.

- 15 En cada protocolo de este ejemplo, la temperatura de desnaturalización era de 94 °C, la temperatura de asociación era de 58 °C y la temperatura de alargamiento era de 72 °C. Para el Rapidcycler, el tiempo de desnaturalización era de 5 segundos, el tiempo de asociación era de 10 segundos, el tiempo de alargamiento era de 15 segundos y la pendiente era de 9,9. Para el Primus HT Multiblock, los tiempos de desnaturalización, asociación y alargamiento eran cada uno de 1 minuto, la temperatura de tapa era de 110 °C y la presión de tapa era de 150 N. Se evaluaron los productos de PCR en un gel de agarosa o en Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.) como se describe a continuación.

Enfoque en dos etapas

- 25 En este ejemplo, se obtuvieron medidas de expresión génica para 9 genes. Se amplificaron las reacciones de PCR en dos rondas. En la primera ronda, se estableció una reacción que contenía tampón, MgCl₂, dNTP, una mezcla preparada anteriormente de ADNc y mezcla de moldes competitivos (1:1 de ADNc de p85 de A549 y una de las mezclas de moldes competitivos del sistema 1a de G.E.N.E., la mezcla D), polimerasa taq y pares cebadores para los 9 genes. Se cicló esta reacción durante 5, 8, 10 o 35 ciclos. La mezcla D del sistema 1 de G.E.N.E. contenía CT de β-actina 10⁻¹² M y CT de los otros genes 10⁻¹⁵ M. La concentración de cada cebador en la mezcla de cebadores era de 0,05 µg/µl. Después de esta amplificación, se diluyó este producto PCR con agua para uso como molde en la ronda 2.

- 30 En la ronda 2, se tomaron alícuotas de una mezcla estandarizada que contenía tampón, MgCl₂, polimerasa taq y un par cebador específico de un gen dado en tubos que contenían 1 µl de cada una de las siguientes diluciones de producto de PCR de la primera ronda: no diluido, 1/5, 1/10, 1/50, 1/100, 1/1.000, 1/10.000, 1/100.000 y 1/1.000.000. Se ciclaron estas reacciones 35 veces y se detectaron en un gel de agarosa o Agilent 2100 Bioanalyzer como se describe a continuación. Se seleccionaron los pares cebadores usados en esta ronda de entre los pares cebadores usados en la ronda 1. No se añadió ADNc ni mezcla de moldes competitivos adicionales a la reacción PCR en la ronda 2, en este ejemplo.

- 40 Para controlar las reacciones no en dos etapas, se diluyó en serie la mezcla de ADNc y mezcla de moldes competitivos preparada para uso en la ronda 1 de los 9 genes antes de la amplificación: no diluido, 1/5, 1/10, 1/50, 1/100, 1/1.000 y 1/10.000. Se combinó una alícuota de 1 µl de cada una de estas diluciones con una alícuota de 9 µl de una mezcla estandarizada que contiene tampón, MgCl₂, polimerasa Taq y un par cebador específico de un gen dado (0,05 µg/µl de cada cebador). Se amplificaron estas reacciones con solo una ronda de 35 ciclos.

- 45 Se obtuvieron también las medidas de expresión génica para otros 96 genes usando un enfoque en dos etapas de algunas realizaciones. Se usaron en estos experimentos muestras de ADNc derivado de ARN de referencia humano universal de Stratagene y mezclas de moldes competitivos del sistema 1 de G.E.N.E. (que contienen CT de 96 genes). Se incluyó en las reacciones de la primera ronda una solución que contenía cebadores de cada uno de los 96 genes representados por CT en el sistema 1 de G.E.N.E. Se diluyó esta mezcla de cebadores de 96 genes de modo que la concentración de cada cebador fuera de 0,005 µg/µl. Se cicló cada reacción de la ronda 1 35 veces. Se diluyeron entonces los productos de PCR de la ronda 1 100 veces (se diluyó 1 µl de producto de ronda 1 en 99 µl de agua). Se usó 1 µl de producto de PCR de ronda 1 diluido en cada reacción de ronda 2 junto con cebadores para un gen dado seleccionado de entre aquellos amplificados en la ronda 1, y se ciclaron 35 veces.

- 50 Se realizaron reacciones no en dos etapas de control usando muestras de ADNc derivado de ARN de referencia humano universal de Stratagene y mezclas de moldes competitivos del sistema 1 de G.E.N.E. como se describe anteriormente. Para estos experimentos, no se realizó dilución del ADNc ni mezcla de moldes competitivos antes de la amplificación.

55 *Electroforesis y cuantificación*

Electroforesis en gel de agarosa:

Después de la amplificación, se cargaron directamente los productos de PCR en geles de agarosa al 4 % (NuSieve:SeaKem 3:1) que contenían bromuro de etidio 0,5 µg/ml. Se sometieron a electroforesis los geles durante aproximadamente 1 hora a 225 V. Se enfrió el tampón de electroforesis y se recirculó durante la electroforesis. Se visualizaron los geles con un sistema de análisis de imágenes Foto/Eclipse (Fotodyne, Hartland, Wis.). Se almacenaron las imágenes digitales en un ordenador Power Mac 7100/66 y se empleó el software Collage (Fotodyne) para análisis densitométrico (o se analizaron usando Agilent 2100 Bioanalyzer (como se discute a continuación)).

Se determinó la cuantificación de la expresión génica. En primer lugar, se calculó el cociente de molde nativo/molde competitivo (NT/CT) de un gen de referencia de β-actina, así como los cocientes de NT/CT para cada uno de los genes para medir. Debido a que la concentración inicial de molde competitivo añadida a la reacción PCR era conocida, pudo determinarse la concentración de NT inicial. Puesto que cada cociente de NT/CT estaba basado en la tinción con un tinte intercalante (bromuro de etidio) de los productos de PCR y la intensidad de tinción está afectada tanto por el número de moléculas presentes como por la longitud de las moléculas en pares de bases, se corrigieron arbitrariamente los NT al tamaño del producto molde competitivo antes de tomar el cociente de NT/CT. Los heterodímeros (HD), cuando eran medibles, se corrigieron al tamaño del molde competitivo y se dividieron entre 2. Se añadió la mitad del valor de HD al NT y se añadió la otra mitad al molde competitivo antes de tomar el cociente de NT/CT, puesto que una hebra del HD viene del NT y la otra viene del CT. En segundo lugar, se dividió el número calculado de moléculas de NT para un gen dado entre el número calculado de moléculas de NT de β-actina para corregir las diferencias de carga.

Para realizaciones que usan el enfoque en dos etapas, se midieron los genes detectados en cada condición (variando la dilución y/o número de ciclos de la ronda 1) frente a β-actina detectada en las mismas condiciones. Por ejemplo, la ronda 1 de un proceso en dos etapas contenía cebadores para 9 genes incluyendo β-actina y c-myc, que pueden usarse como ácidos nucleicos de referencia. Se hizo una dilución 1/100.000 de la reacción PCR de la ronda 1 y se usó en la ronda 2. Se usó una alícuota de esta dilución en la ronda 2 para amplificar tanto β-actina como c-myc. En estas condiciones, se midió c-myc como $3,40 \times 10^4$ moléculas/10⁶ moléculas de β-actina/cuando se cicló 35 veces en la ronda 1 y 35 veces en la ronda 2.

La Figura 29 ilustra los resultados de estos experimentos. Brevemente, se amplificaron las reacciones de PCR en Rapidcycler. En la ronda 1, se preparó una mezcla de reacción de 10 µl que contenía tampón, MgCl₂, dNTP, una mezcla preparada anteriormente de ADNc y mezcla de moldes competitivos (1:1 de ADNc de p85 de A549 y mezcla D del sistema 1 de G.E.N.E.), polimerasa Taq y 1 µl de una solución madre 10 x de 9 pares cebadores (concentración de 0,05 µg/µl). Se cicló esta reacción 5, 8, 10 o 35 ciclos. Después de la amplificación de ronda 1, se diluyeron los productos de PCR para uso como moldes en la ronda 2. En la ronda 2, se prepararon 10 µl de reacción PCR disponiendo 9 µl de mezcla maestra que contenía tampón, MgCl₂, polimerasa Taq y un par cebador específico de un gen en tubos que contenían 1 µl de cada una de las siguientes diluciones de producto de PCR de la ronda 1: no diluido, 1/5, 1/10, 1/50, 1/100, 1/1.000, 1/10.000, 1/100.000 y 1/1.000.000. Se ciclaron estas reacciones 35 veces. Se seleccionaron los pares cebadores usados en la ronda 2 de entre los pares cebadores usados en la ronda 1. No se añadieron ADNc ni mezcla de moldes competitivos adicionales a la reacción PCR en la ronda 2. Para las reacciones no en dos etapas, se diluyó en serie la mezcla de ADNc y mezclas de moldes competitivos preparadas para uso en la ronda 1 antes de la amplificación: no diluido, 1/5, 1/10, 1/50, 1/100, 1/1.000 y 1/10.000. Se amplificaron estas reacciones en solo una ronda de 35 ciclos. Se combinó una alícuota de 1 µl de cada dilución con una alícuota de una mezcla que contenía tampón, MgCl₂, polimerasa Taq y un par cebador específico de un gen (0,05 µg/µl de cada cebador). Se determinó la cuantificación de la expresión génica.

Electroforesis microcapilar en Agilent 2100 Bioanalyzer:

Después de la amplificación, se cargó 1 µl de cada 10 µl de reacción PCR en un pocillo de un chip preparado según el protocolo del fabricante para DNA 7500 Assay. Brevemente, se cargaron 9 µl de matriz de gel-tinte en el chip en un pocillo y se pusieron a presión los chips durante 30 segundos. Se rellenaron dos pocillos adicionales con matriz de gel-tinte y se cargaron los pocillos restantes cada uno con 5 µl de marcador de peso molecular. Se cargó 1 µl de marcador de tamaño de ADN en un pocillo de marcador de tamaño y se cargó 1 µl de producto de PCR en cada pocillo de muestra. Se agitó con vórtex el chip y se dispuso en Agilent 2100 Bioanalyzer. El programa DNA 7500 Assay que se manejó aplica una corriente secuencialmente a cada muestra para separar productos. Se detectó ADN por fluorescencia del tinte intercalante en la matriz de gel-tinte. Se calcularon los cocientes de NT/CT a partir del área bajo la curva para cada producto de PCR y se realizó una corrección de tamaño puesto que, como en la electroforesis en gel de agarosa teñida con bromuro de etidio, se usó un tinte intercalante para detectar el ADN.

Se realizaron todos los análisis estadísticos usando SPSS versión 9.0 para Windows. Se realizó una prueba de correlación Pearson de dos colas sobre datos transformados logarítmicamente para comparar los valores de expresión génica obtenidos usando un enfoque no en dos etapas con aquellos obtenidos usando uno en dos etapas, de acuerdo con algunas realizaciones de la presente invención. La correlación se consideró estadísticamente significativa si el vapor de p era menor de 0,05.

Resultados: Enfoque no en dos etapas de amplificación de 9 genes

La Figura 30 ilustra los resultados de experimentos que comparan enfoques no en dos etapas (30A y 30C) con enfoques en dos etapas (30B y 30D), según algunas realizaciones de la presente invención. Las Figuras 30A-D ilustran los resultados representativos de usar un proceso en dos etapas frente a uno no en dos etapas. La Figura 30A ilustra que, en una reacción no en dos etapas que usa cebadores de β -actina, una dilución de la mezcla de reacción PCR no amplificada de la ronda 1 de más de 100, seguida de una ronda de 35 ciclos de amplificación por PCR con un par de cebadores no producía ningún producto detectable. Los carriles son como sigue: carril 1, marcador de tamaño pGEM; carril 2, reacción PCR que contenía ADNc no diluido en que el NT de β -actina es equivalente a 300.000 moléculas iniciales de CT de β -actina; carril 3, reacciones PCR que contienen ADNc diluido 1:5/mezcla de CT, 60.000 moléculas; carril 4, ADNc diluido 1:10/mezcla de CT, 30.000 moléculas; carril 5, ADNc diluido 1:50/mezcla de CT, 6.000 moléculas; carril 6, ADNc diluido 1:100/mezcla de CT, 3.000 moléculas; carril 7, ADNc diluido 1:1000/mezcla de CT, 300 moléculas; carril 8, ADNc diluido 1:10.000/mezcla de CT, 30 moléculas.

La Figura 30B ilustra productos de PCR obtenidos a partir de un enfoque en dos etapas, usando una alícuota de producto de PCR de ronda 1 y cebadores de β -actina. Carril 1, marcador de tamaño pGEM; carril 2, 1/500 de 10 μ l del producto de PCR de ronda 1 (1 μ l de dilución 1:50), equivalente a 600 moléculas iniciales de CT de β -actina; carril 3, 1/1000 de producto de PCR de ronda 1, 300 moléculas; carril 4, 1/10.000 de producto de PCR de ronda 1, 30 moléculas; carril 5, marcador de tamaño pGEM; carril 6, 1/10.000 de producto de PCR de ronda 1, 30 moléculas; carril 7, 1/100.000 de producto de PCR de ronda 1, 3 moléculas; carril 8, 1/1.000.000 de producto de PCR de ronda 1, 0,3 moléculas; carril 9, 1/10.000.000 de producto de PCR de ronda 1, 0,003 moléculas.

La Figura 30C ilustra una reacción no en dos etapas que usa cebadores de catalasa, en que una dilución de la mezcla de reacción PCR no amplificada de ronda 1 de más de 100, seguida de una ronda de 35 ciclos de amplificación por PCR con un par cebador no producía ningún producto detectable. Es decir, diluir la mezcla de reacción PCR no amplificada más de 1.000 veces seguido de una ronda de 35 ciclos de PCR no producía producto, como era el caso con β -actina (Figura 30A). Los carriles son como sigue: carril 1, marcador de tamaño pGEM; carril 2, reacción PCR que contenía ADNc no diluido y mezcla de moldes competitivos, equivalente a 3.000 moléculas de CT de catalasa; carril 3, ADNc diluido 1:5/mezcla de CT, 600 moléculas; carril 4: ADNc diluido 1:10/mezcla de CT, 300 moléculas; carril 5, ADNc diluido 1:50/mezcla de CT, 60 moléculas; carril 6, ADNc diluido 1:100/mezcla de CT, 30 moléculas; carril 7, ADNc diluido 1:1.000/mezcla de CT, 3 moléculas; carril 8, ADNc diluido 1:10.000/mezcla de CT, 0,3 moléculas.

La Figura 30D ilustra productos de PCR obtenidos a partir de un enfoque en dos etapas, usando una alícuota de producto de PCR de ronda 1 y productos de PCR de cebadores de catalasa para la segunda ronda. Los carriles son como sigue: carril 1, marcador de tamaño pGEM; carril 2, 1/100 de 10 μ l de producto PCR de ronda 1 (1 μ l de dilución 1:10), equivalente a 30 moléculas de CT de catalasa; carril 3, 1/500 de producto de PCR de ronda 1, 6 moléculas; carril 4, 1/1.000 de producto de PCR de ronda 1, 3 moléculas; carril 5, 1/10.000 de producto de PCR de ronda 1, 0,3 moléculas; carril 6, 1/100.000 de producto de PCR de ronda 1, 0,03 moléculas; carril 7, 1/1.000.000 de producto de PCR de ronda 1, 0,003 moléculas; carril 8, 1/10.000.000 de producto de PCR de ronda 1, 0,0003 moléculas.

Enfoque en dos etapas de amplificación de 9 genes

Después de 35 ciclos de amplificación en ronda 1 con pares cebadores de 9 genes, se diluyeron alícuotas de los productos de PCR y se amplificaron con cebadores de uno de los 9 genes. Se observaron bandas brillantes y marcadas para cada gen. Por tanto, se usó en enfoques en dos etapas para obtener 9 medidas de expresión génica la misma cantidad de ADNc y mezcla de moldes competitivos que se usa una reacción típica no en dos etapas para medir un gen en una ronda de amplificación.

Además, el producto de PCR de la ronda 1 puede diluirse más de 100.000 veces para c-myc, 1.000.000 veces para β -actina o 10.000.000 veces para catalasa y seguir cuantificándose después de amplificación con pares cebadores de un gen en la ronda 2 (Figura 29). En contraposición, cuando se diluía el ADNc y la mezcla de moldes competitivos usados en la ronda 1 más de 1.000 veces antes de la amplificación (o más de 100 veces para β -actina) y se amplificaba entonces con un solo par cebador de uno cualquiera de estos genes en una sola ronda de 35 ciclos, no se observaba producto detectable. La cantidad de producto amplificado que podía diluirse antes de la ronda 2 y seguir produciendo producto detectable después de la ronda 2 estaba directamente relacionada con el número de ciclos de la ronda 1.

Aumentar el número de ciclos usados en la ronda 1 aumentaba la cantidad de producto de PCR que podía diluirse antes de la ronda 2 y seguir siendo detectable después de la amplificación de la ronda 2. Por lo tanto, pueden hacerse más medidas de expresión génica usando una muestra cuando se amplifica en un enfoque en dos etapas (por ejemplo, usando 35 ciclos en cada ronda) que cuando se usan menos ciclos (por ejemplo, 5, 8 y 10 ciclos) en la ronda 1, o cuando no se usa segunda ronda. Se muestran en la Figura 29 detalles de cada gen y cada condición. Se muestran en la Figura 30 los geles representativos del control de reacciones no en dos etapas y reacciones en dos etapas.

Enfoque en dos etapas de amplificación de 96 genes

Se muestran en la Figura 4 los valores de expresión génica obtenidos mediante reacción no en dos etapas y en dos etapas de 96 genes usando ADNc derivado de ARN de referencia humano universal de Stratagene. Aunque se incluyeron 96 pares de cebadores en las reacciones en dos etapas, se reseñan valores de expresión génica de solo 93 genes porque 1) cada valor de expresión génica se reseña como moléculas de un gen dado/ 10^5 moléculas de β -actina, así que no se reseñan los valores de β -actina, 2) aunque se incluyen dos conjuntos de reactivos para medir la expresión génica de GAPD (GAPD CT1 y GAPD CT1) en el kit sistema 1 de G.E.N.E., solo se midió GAPD CT1 en esta muestra y 3) los reactivos de un gen, BAX- α , proporcionado en el kit no pasaron el ensayo de control de calidad realizado por G.E.N.E., Inc., así que no se valoró este gen en este estudio. Los análisis bivariados de los valores de expresión génica usando los dos enfoques revelaron una correlación positiva ($r= 0,993$) altamente significativa ($p= 0,001$). Los dos enfoques eran reproducibles, no mostrando diferencias significativas en las medidas para más de un 90 % de los genes ensayados (Figura 29).

La Figura 31 es una gráfica que muestra la correlación de los valores de expresión génica de 96 genes obtenidos por enfoques en dos etapas o no en dos etapas. Se combinaron muestras de ADNc derivado de ARN de referencia humano de referencia de Stratagene con mezclas de CT (se usaron mezclas B, C, D, E y F del sistema 1 de G.E.N.E.) y se amplificaron por StaRT-PCR sencilla o por StaRT-PCR múltiple de 96 genes con pares cebadores de todos los genes del sistema 1 de G.E.N.E. Se presentan los valores medios en la Figura 4 para los 93 genes que pudieron evaluarse. De estos, 79 se midieron por los enfoques tanto no en dos etapas como en dos etapas y pudieron compararse. Se presentan los valores de expresión génica como moléculas de ARNm por 10^6 moléculas de ARNm de β -actina. Se representan los valores obtenidos por los procedimientos no en dos etapas frente al eje X y se representan los valores obtenidos por procedimientos en dos etapas frente al eje Y.

Medidas por enfoque en dos etapas de muestras pequeñas.

Las Figuras 1 y 2 indican los datos de expresión génica obtenidos a partir de pequeñas cantidades de materiales. La Figura 1 muestra los datos recogidos a partir de una biopsia de aspiración con aguja fina de células de tejido primario de carcinoma pulmonar no microcítico (NSCLC). Se midieron todos los datos usando las mezclas de moldes competitivos del sistema 1 de GENE por 18 PCR múltiple.

La Figura 2 muestra los datos recogidos a partir de células de tejido primario de un donante de pulmón que no tenía enfermedad pulmonar. Se recogió también la expresión génica usando PCR múltiple de 96 genes con las mezclas de moldes competitivos del sistema 1 de GENE.

EJEMPLO II

El siguiente ejemplo proporciona detalles adicionales del proceso global de evaluación de las medidas de expresión según algunas realizaciones de la presente invención.

Materiales

1. Los reactivos de RT-PCR estandarizada, incluyendo cebadores y mezclas estandarizadas, se adquieren en Gene Express, Inc. (GEI, Toledo, OH).
2. Tampón para Idaho Rapidcycler air thermocycler: Tris-HCl 500 mM, pH 8,3, BSA 2,5 $\mu\text{g}/\mu\text{l}$, MgCl_2 30 mM (Idaho Technology, Inc., Idaho Falls, ID).
3. Tampón para termocicladores de bloque, Thermo 10 X, KCl 500 mM, Tris-HCl 100 mM, pH 9,0, Triton X-100 al 1,0 % (Promega, Madison, WI).
4. Polimerasa taq (5U/ μl), transcriptasa inversa del virus de leucemia de murino Moloney (MMLV), tampón de primera hebra de MMLV RT 5X: Tris-HCl 250 mM, pH 8,3, KCl 375 mM, MgCl_2 15 mM, ditiotritol 50 mM, cebadores de oligo-dT, Rnasin, marcador de tamaño pGEM y trifosfatos de desoxinucleótidos (dNTP) se obtienen también de Promega.
5. TriReagent se obtiene de Molecular Research Center, Inc. (Cincinnati, OH).
6. El agua exenta de ribonucleasa (ARNasa) y los kits de clonación TOPO TA se obtienen de Invitrogen (Carlsbad, CA). La calidad del agua exenta de ARNasa puede ser importante para la extracción eficaz de ARN intacto. Por ejemplo, un tratamiento con DEPC inadecuado y/o una retirada inadecuada del DEPC después del tratamiento puede inhibir la transcripción inversa y la PCR.
7. Los kits de preparación de plásmido GigaPrep se adquieren en Qiagen (Texas).
8. Los chips Caliper AMS 90SE se obtienen de Caliper Technologies, Inc. (Mountain View, CA).
9. Las columnas de purificación de ADN se obtienen de QiaQuick (Qiagen, Valencia, CA).

Extracción de ARN y transcripción inversa

Extracción de ARN: Las suspensiones celulares pueden sedimentarse, escurrirse el sobrenadante, disolverse el sedimento en TriReagent y extraerse (según las instrucciones del fabricante y los procedimientos descritos anteriormente, véase, por ejemplo, Bustin, S. A. (2000) "Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays". *J. Mol. Endocrinol.* 25, 169-193. El sedimento de ARN puede almacenarse en etanol a -80 °C o suspenderse en agua exenta de ARNasa y congelarse a -80 °C. Puede almacenarse en estas condiciones durante años. La calidad del ARN puede evaluarse en Agilent 2100 usando el chip de ARN, según las instrucciones del fabricante.

Transcripción inversa 1: Puede transcribirse de forma inversa 1 µg de ARN total usando MMLV RT y un cebador de oligo-dT como se describe anteriormente. Véase, por ejemplo, Willey, J. C., Coy, E. L., Frampton, M. W., *et al.* (1997) "Quantitative RT-PCR measurement of cytochromes p450 1A1, 1B1, and 2B7, microsomal epoxide hydrolase, and NADPH oxidoreductase expression in lung cells of smokers and non-smokers". *Am. J. Respir. Cell Mol. Biol.* 17, 114-124. Para cantidades pequeñas de ARN (por ejemplo, menos de aproximadamente 100 ng), puede mejorarse la eficacia de la transcripción inversa usando Sensiscript™ en lugar de transcriptasa inversa de MMLV, por ejemplo, puede obtenerse una transcripción inversa eficaz de aproximadamente 50 ng de ARN con Sensiscript™. La reacción puede incubarse a 37 °C durante 1 h.

Síntesis y clonación de moldes competitivos

Los moldes competitivos patrones internos (CT) pueden construirse basándose en procedimientos anteriormente descritos. Véanse, por ejemplo, Willey, J. C., Crawford, E. L. y Jackson, C. M. (1998) "Expression measurement of many genes simultaneously by quantitative RT-PCR using standardized mixtures of competitive templates". *Am. J. Respir. Cell Mol. Biol.* 19, 6-17; Crawford, E. L., Peters, G. J., Noordhuis, P., *et al.* (2001) "Reproducible gene expression measurement among multiple laboratories obtained in a blinded study using standardized RT (StaRT)-PCR". *Mol. Diagn.* 6, 217-225; y/o Celi, F. S., Zenilman, M. E. y Shuldiner, A. R. (1993) "A rapid and versatile method to synthesize internal standards for competitive PCR". *Nucleic Acids Res.* 21, 1047.

Diseño de cebador de molde nativo

Antes de construir un CT de un gen, puede diseñarse un par cebador que amplifique (preferiblemente que amplifique eficazmente) ADNc nativo correspondiente al gen expresado. Por ejemplo, pueden diseñarse cebadores con una o más de las siguientes características: (1) la capacidad de amplificar de aproximadamente 200 a aproximadamente 850 bases de la región de codificación de genes de interés; (2) una temperatura de asociación de aproximadamente 58 °C (tolerancia de ± 1 °C). Puede usarse Primer 3 del software Primer 3.1 (Steve Rozen, Helen J. Skaletsky, 1996, 1997) para diseñar los cebadores (código disponible en http://www-genome.wi.net.edu/genome_software/other/primer3.html) en algunas realizaciones. Los cebadores se diseñaron inicialmente usando el software Primer 3.1 para amplificar de aproximadamente 200 a aproximadamente 800 bases de la región de codificación de genes diana con una temperatura de asociación de aproximadamente 58 °C (tolerancia de ± 1 °C). Esto permitió que las reacciones PCR de este ejemplo operaran en condiciones idénticas o casi idénticas y permite además la automatización y aplicaciones de alto rendimiento, incluyendo electroforesis en gel capilar de microfluidos. Por ejemplo, las secuencias cebadoras y los números de acceso a Genbank para ciertos genes están disponibles en www.geneexpressinc.com. Pueden diseñarse también los cebadores para amplificar de aproximadamente 20 a aproximadamente 2.000 bases, en otros ejemplos.

Ensayo de cebador de molde nativo

Los cebadores diseñados pueden sintetizarse y usarse para amplificar molde nativo de ADNc correspondiente al (a los) gen(es) de interés. La presencia de una sola banda fuerte después de 35 ciclos de PCR puede verificar que los cebadores son suficientemente eficaces y/o específicos para algunas realizaciones. Por ejemplo, pueden ensayarse los cebadores usando ARN transcrito de forma inversa de una variedad de tejidos y/o clones de ADNc conocidos por representar el (los) gen(es) de interés. En algunas realizaciones, pueden rediseñarse los pares de cebadores que no consiguen amplificar el gen diana en ningún tejido ni clon de ADNc individual, por ejemplo menos de aproximadamente un 10 % del tiempo, y repetirse el proceso.

Diseño de cebador de molde competitivo

Puede prepararse un cebador de CT según procedimientos descritos anteriormente y/o como se ilustran en la Figura 3. La Figura 32a ilustra que pueden usarse cebadores de codificación (barra rayada) e inverso (barra negra) (aprox. 20 pb de longitud) que cubren una región de 150-850 pb para amplificar el molde nativo (NT) de ADNc. La polimerasa Taq puede sintetizar ADN a partir de estos cebadores (líneas discontinuas) usando el NT.

La Figura 32b ilustra que después de ensayar que los cebadores de molde nativo funcionan, puede diseñarse un cebador de CT para que sea un cebador de aproximadamente 40 pb con la secuencia del cebador inverso (barra negra) en el extremo 5' y una secuencia de 20 pb homóloga de una secuencia de molde nativo interna (barra blanca) en el extremo 3', colineal con la secuencia del cebador inverso. El extremo 3' de este cebador de 40 pb puede diseñarse para que sea homólogo de una región de aproximadamente 50 a aproximadamente 100 pb interna del cebador inverso. El extremo 5' de este cebador de aproximadamente 40 pb puede hibridar con la región homóloga del cebador inverso, mientras que el extremo 3' puede hibridar con la secuencia interna. Además, la polimerasa Taq

puede sintetizar ADN usando los cebadores unidos al extremo 3' (línea discontinua) y no el cebador unido al extremo 5'.

La Figura 32c ilustra que, en el siguiente ciclo de PCR, el ADN recién sintetizado usando el cebador de aproximadamente 40 pb hibridado con la secuencia interna puede unirse al cebador de codificación (barra rayada) y puede sintetizarse una hebra homóloga. La Figura 32d ilustra que esto puede generar un CT bicatenario con la secuencia de cebador inverso aproximadamente 100 pb más cerca del cebador de codificación de lo que aparece naturalmente en el NT. Véanse, por ejemplo, Chomczynski, P. y Sacchi, N. (1993) "Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction". *Anal. Biochem.* 62, 156-159; Celi, F. S., Zenilman, M. E. y Shuldiner, A. R. (1993) "A rapid and versatile method to synthesize internal standards for competitive PCR". *Nucleic Acids Res.* 21, 1047).

Ensayo de cebador de molde competitivo

El CT preparado puede ensayarse. Por ejemplo, el cebador de CT puede emparejarse con el cebador de codificación diseñado y usarse para amplificar el CT a partir de ADNc nativo. Antes de construir cada molde competitivo en este ejemplo, se ensayó cada par cebador de este ejemplo usando ARN transcrito de forma inversa a partir de una variedad de tejidos o clones de ADNc individuales conocidos por representar el gen de interés como control de calidad. Para los pares cebadores que fallaban (aproximadamente un 10 % del tiempo), se diseñaron nuevos y se repitió el proceso. Para cada gen, se preparó entonces un cebador de molde competitivo (un oligo de fusión de aproximadamente 40 pb). El extremo 3' de cada cebador de fusión consistía en una secuencia de aproximadamente 20 bases homóloga de una región de aproximadamente 50 a aproximadamente 100 bases en dirección 3' al cebador inverso. El extremo 5' era el cebador inverso de 20 pb.

Producción de molde competitivo-patrón interno

Para cada uno de un número de genes para ensayar, pueden establecerse 5 reacciones PCR de 10 μ l, usando el cebador de codificación de NT diseñado y el cebador de CT, y amplificarse durante 35 ciclos. Los productos de las 5 reacciones PCR pueden combinarse, someterse a electroforesis en gel NuSieve al 3 % en 1xTAE, cortarse la banda del tamaño correcto del gel y extraerse usando un procedimiento QiaQuick (Qiagen, Valencia, Calif). Los productos de PCR purificados pueden clonarse en el vector PCR 2.1 usando los kits de clonación TOPO TA (Invitrogen, Carlsbad, Calif.), que pueden transformarse entonces en HS996 (una variante de DH10B resistente al fago T1). Después de clonación y transformación, pueden sembrarse colonias en placas LB que contienen X-Gal, IPTG y carbenicilina y seleccionarse 3 colonias blancas aisladas. Pueden prepararse minipreparaciones de plásmido, efectuarse la digestión con EcoRI y someterse a electroforesis los productos digeridos en agarosa SeaKem al 3 %. Para aquellos clones que muestran un inserto basándose en la digestión con EcoRI, puede confirmarse que el inserto es el deseado mediante secuenciación de la misma preparación de plásmido no digerido usando cebadores específicos de vector. Pueden usarse clones con homología con la secuencia génica correcta y que tienen un 100 % de coincidencia con las secuencias de cebador en la preparación de CT a gran escala, y pueden incluirse en mezclas estandarizadas. Por ejemplo, aquellos que pasan esta valoración de control de calidad pueden usarse en las siguientes etapas.

Se prepararon entonces plásmidos de cada clon de calidad asegurada en cantidades suficientemente grandes (aproximadamente 1,5 l) para permitir aproximadamente 1.000 millones de ensayos (aproximadamente 2,6 mg). Se purificaron los plásmidos de las células cosechadas resultantes usando el kit Qiagen GigaPrep. Se valoraron los rendimientos de plásmido usando un fluorómetro Hoeffer DyNAQuant 210.

En este ejemplo, se secuenció de nuevo una alícuota de cada preparación de plásmido como control de calidad. Para cada molde competitivo que pasaba las etapas de control de calidad esquematizadas en este ejemplo, se valoró la sensibilidad del CT clonado y los cebadores efectuando reacciones PCR en diluciones en serie y determinando la concentración limitante que seguía produciendo un producto de PCR. En este ejemplo, se permitieron incluir en mezclas de moldes competitivos estandarizadas solo aquellas preparaciones y cebadores que permitían la detección de 60 moléculas o menos (por ejemplo, un producto obtenido con 10^{-16} CT en un volumen de reacción PCR de 10 μ l). En este ejemplo, la mayoría de ensayos que se desarrollaron tenían una sensibilidad de aproximadamente 6 moléculas o menos (por ejemplo, más de un 80 % de los CT que se desarrollaron tenían una sensibilidad de 6 moléculas o menos o CT 10^{-17} M).

Preparación de mezclas estandarizadas

Se mezclaron plásmidos de preparaciones de calidad asegurada en mezclas de moldes competitivos que representan 24 o 96 genes. La concentración de moldes competitivos en las mezclas estandarizadas de 24 genes eran 4×10^{-9} M para CT de β -actina, 4×10^{-10} M para GAPD (CT1), 4×10^{-11} M para GAPD (CT2) y 4×10^{-8} M para cada uno de los otros CT en este ejemplo.

Las mezclas de moldes competitivos de 24 genes pueden linealizarse por digestión con NotI antes de la preparación de una serie de mezclas estandarizadas diluidas en serie descritas a continuación. Por ejemplo, pueden incubarse las mezclas con enzima NotI a una concentración de 1 unidad/ μ g de ADN de plásmido en aproximadamente 15 ml de tampón a 37 °C durante 12-16 horas. Se combinaron 4 mezclas de moldes competitivos de 24 genes linealizados

en iguales cantidades, produciendo mezclas de moldes competitivos de 96 genes con concentraciones de 10^{-9} M para β -actina, 10^{-10} M para GAPD (CT1), 10^{-11} para GAPD (CT2) y 10^{-8} M para los otros CT. Estas mezclas pueden diluirse entonces en serie con una mezcla de CT de genes de referencia, por ejemplo, que comprende β -actina 10^{-9} M, mezcla de GAPDH (CT1) 10^{-10} M, GAPDH (CT2) 10^{-11} produciendo una serie madre a concentraciones de 10^{-9} M para β -actina, 10^{-10} M para GAPD CT1, 10^{-11} M para GAPD CT2 y 10^{-8} , 10^{-9} , 10^{-10} , 10^{-11} , 10^{-12} y 10^{-13} M para los otros CT usados en este ejemplo.

Estas concentraciones madre pueden diluirse 1.000 veces, proporcionando diluciones de trabajo, por ejemplo, produciendo una serie de 6 mezclas estandarizadas diluidas en serie (A-F) a concentraciones de 10^{-12} M para β -actina, 10^{-3} M para GAPD CT1, 10^{-14} M para GAPD CT2 y 10^{-11} (A), 10^{-12} (B), 10^{-13} (C), 10^{-14} (D), 10^{-15} (E) y 10^{-16} M (F) para los otros CT usados en este ejemplo.

A continuación, se ilustra el uso de una serie de mezclas estandarizadas diluidas en serie de acuerdo con algunas realizaciones de la presente invención. En este ejemplo, "SMIS" hace referencia a una mezcla estandarizada de patrones internos, preparada de acuerdo con realizaciones de la presente invención.

Puede combinarse un volumen de muestra de ADNc (diluida a un nivel en equilibrio con la cantidad de moléculas de CT de β -actina en 1 μ l de SMIS (6×10^5 moléculas)) y mezclarse con un volumen igual de SMIS A-F apropiada, de tal modo que el cociente de NT/CT de un ácido nucleico que se está midiendo sea mayor de aproximadamente 1/10 y menor de aproximadamente 10/1. Por ejemplo, si entre las muestras previas se ha expresado un gen dentro de un intervalo de 10^1 - 10^3 moléculas/ 10^6 moléculas de β -actina, el gen se medirá usando la SMIS E. En contraposición, si entre las muestras previas se ha expresado un gen dentro de un intervalo de 10^5 - 10^7 moléculas/ 10^6 moléculas de β -actina, el gen se medirá usando la SMIS B. Si no es conocida la SMIS apropiada para un gen particular en una muestra de un tipo particular de tejido, puede medirse la expresión usando las SMIS tanto C como E. Esto permite la medida dentro de 4 órdenes de magnitud. Para las raras muestras que expresan el gen fuera de los intervalos esperados, puede efectuarse un análisis de seguimiento con la mezcla de CT apropiada. Por ejemplo, para los pocos genes expresados a muy alto o muy bajo nivel, pueden repetirse los análisis con la SMIS A o F.

Puede usarse un volumen de 1 μ l de mezcla de ADN/SMIS para cada ensayo de expresión génica para efectuar y puede combinarse con los demás componentes de la mezcla de reacción PCR, por ejemplo, tampón, dNTPs, Mg⁺⁺, polimerasa Taq, H₂O). Pueden prepararse tubos o pocillos con un par cebador de un solo gen para medir. Si los productos se van a analizar por el dispositivo PE 310, los cebadores pueden marcarse con el flúor apropiado. Pueden disponerse alícuotas de esta mezcla de reacción PCR en tubos individuales que contienen cada uno cebadores de un solo gen. Usando este enfoque, se fija simultáneamente el cociente de CT para cada gen en la mezcla respecto a su correspondiente NT en el ADNc. Cuando se transfieren alícuotas de esta mezcla a recipientes de reacción PCR, aunque puede haber variaciones en los volúmenes de carga resultantes del pipeteo, la variación está controlada en el cociente de NT/CT para cualquier gen respecto al cociente de NT/CT para un gen de referencia. Este enfoque posibilita también la medida de la expresión estandarizada.

35 *Amplificación por PCR*

Cada mezcla de reacción puede ciclarse en un termociclador de aire (por ejemplo, Rapidcycler (Idaho Technology, Inc., Idaho Falls, ID)) o termociclador de bloque (por ejemplo, ciclador térmico en bloque PTC-100 con tapa calentada, MJ Research, Inc., Incline Village, NV) durante 35 ciclos. En cualquier termociclador, la temperatura de desnaturalización es de 94 °C, la temperatura de asociación es de 58 °C y la temperatura de alargamiento es de 72 °C.

Separación y cuantificación de los productos de PCR de NT y CT

a. *Gel de agarosa.* Después de la amplificación, todo el volumen del producto de PCR (típicamente 10 μ l) puede estar en pocillos de geles de agarosa al 4 % (NuSieve: Sea Kem 3/1) que contienen bromuro de etidio 0,5 μ g/ml. Los geles pueden someterse a electroforesis durante aprox. 1 h a 225 V en tampón enfriado continuamente y visualizarse y cuantificarse entonces con un analizador de imágenes (productos disponibles en Fotodyne, BioRad). Después de la electroforesis, puede determinarse la cantidad relativa de NT y CT mediante cuantificación densitométrica de bandas que se han teñido con un tinte intercalante (por ejemplo, bromuro de etidio).

b. *Dispositivo de PE Prism 310 Genetic Analyzer.* Los productos de PCR pueden amplificarse con cebadores marcados con flúor. Puede combinarse 1 μ l de cada reacción PCR con 9 μ l de formamida y 0,5-01 μ l de marcador de tamaño ROX. Pueden calentarse las muestras a 94 °C durante 5 min y enfriarse instantáneamente en una suspensión densa de hielo. Pueden cargarse muestras en la máquina y someterse a electroforesis a 15 kV, 60 °C durante 35-45 min usando polímero POP4 y el conjunto de filtros D. Los parámetros de inyección pueden ser 15 kV, 5 s. Puede usarse software de análisis de fragmentos GeneScan (Applied Biosystems, Inc., Foster City, CA), para cuantificar las alturas de pico que se usan para calcular los cocientes de NT/CT. No tiene que efectuarse corrección de tamaño cuando cada molécula de ADNc se marcaba con un marcador fluorescente de un cebador marcado.

c. *Dispositivo de CE Agilent 2100 Bioanalyzer Microfluidic.* Pueden usarse el kit DNA 7500 o DNA 1000 LabChip. Después de la amplificación, puede cargarse 1 μ l de cada reacción PCR de 10 μ l en un pocillo de un chip preparado

según el protocolo suministrado por el fabricante. Puede realizarse el ensayo de ADN, que aplica una corriente a cada muestra secuencialmente para separar NT de CT. El ADN puede detectarse mediante fluorescencia de un tinte intercalante en la matriz de gel-tinte. Los cocientes de NT/CT pueden calcularse a partir del área bajo la curva (AUC) y pueden hacerse una o más correcciones de tamaño.

5 d. *Dispositivo de CE Caliper AMS 90 Microfluidic*. Las reacciones PCR pueden establecerse en pocillos de una microplaca de 96 o 384 pocillos. Después de la amplificación, puede disponerse la microplaca en un Caliper AMS 90 y seguirse el protocolo recomendado por el fabricante. El AMS 90 puede retirar y someter a electroforesis una muestra de cada pocillo secuencialmente cada 30 s. Los productos de PCR de NT y CT pueden separarse y cuantificarse. Cuando la detección es mediante tinte intercalante fluorescente, no es necesaria la corrección de tamaño.

10 e. *Separación por MALDI-TOF*. Se ha descrito recientemente un procedimiento para separar productos de PCR. Ding, C. y Cantor, C. R. (2003) "A high-throughput gene expression analysis technique using competitive PCR and matrix-assisted laser desorption ionization time-of-flight MS". *Proc. Natl. Acad. Sci. USA* 100, 3059-3064. Este procedimiento puede usarse para cuantificar productos resultantes de la amplificación de ADNc en presencia de SMIS.

Cálculo de la expresión génica- cálculo del número de moléculas de NT presentes al inicio de la PCR para cada gen

Las etapas realizadas para calcular la expresión génica pueden estar basadas en valores de medida densitométrica para productos de PCR de NT y CT separados electroforéticamente, tales como aquellos presentados en la Figura 33. Los cálculos siguientes están basados en el ejemplo de la Figura 33, midiendo la expresión del gen GST respecto a β -actina en una muestra de células epiteliales bronquiales (BEC) reales. Se incluyó un volumen de SMIS que contiene 600.000 moléculas de moldes competitivos de β -actina y 6.000 moléculas de moldes competitivos de GST al inicio de la reacción PCR. Para cada gen, el NT y el molde competitivo amplifican con la misma eficacia. Por tanto el cociente de producto de PCR de NT/CT del gen de β -actina permite la determinación del número de copias de NT de β -actina al inicio de la PCR y el cociente de NT/CT del gen diana permite la determinación del número de copias de NT del gen diana al inicio de la PCR, como se detalla en las etapas siguientes:

1. Corregir el área bajo el pico (AUP) del producto de PCR de NT a la longitud del ADN de CT.
2. Determinar el cociente de AUP de NT corregido a AUP de CT.
3. Multiplicar el valor de NT/CT por el número de moléculas de CT al inicio de la PCR.

Se esquematiza a continuación un cálculo de las moléculas de β -actina usando el protocolo anterior:

- 30 1. $416/532$ (pb de CT de β -actina/pb de NT) x 42 (AUP de NT) = 33 (valor de NT corregido).
2. Corregir el AUP de NT de β -actina dividido entre el AUP de CT de β -actina= 0,37.
3. $0,37$ (NT/CT de β -actina) x 600.000 (número de moléculas de CT de β -actina al inicio de la PCR)= 222.000 moléculas de NT al inicio de la PCR.

Se esquematiza a continuación un cálculo de las moléculas de GST usando el protocolo anterior:

- 35 1. $227/359$ (pb de CT de GST/pb de NT) x 1,5 (AUP de NT) = 0,95 (AUP de NT corregido).
2. $0,95$ (AUP de NT de GST corregido) dividido entre 4,4 (AUP de CT de GST)= 0,22.
3. $0,22$ (NT/CT de GST) x 6000 (número de moléculas de CT de GST al inicio de la PCR) = 1290 moléculas de NT de GST al inicio de la PCR.

40 El cálculo de las moléculas de GST/ 10^6 moléculas de β -actina es de 1290 moléculas de NT de GST/222.000 moléculas de NT de β -actina = 580 moléculas de GST/ 10^6 moléculas de β -actina.

EJEMPLO III

El siguiente ejemplo proporciona detalles adicionales de un enfoque no en dos etapas para evaluar la expresión génica según algunas realizaciones de la presente invención.

Extracción de ARN

45 Se diluyeron desoxirribonucleótidos purificados obtenidos de Pharmacia (Piscataway, NJ.) a una solución madre 10 mM. Se obtuvieron ADN polimerasa recombinante de *Thermus aquaticus* (polimerasa Taq), transcriptasa inversa de virus de mieloblastosis aviar (AMV) e inhibidor de ribonucleasa (RNasin) de Promega (Madison, Wis.). Se obtuvo la enzima EcoRI de USB (Cleveland, Ohio). Se prepararon los cebadores en un sintetizador modelo 391 PCR-Mate EP™ de Applied Biosystems. Se efectuó la PCR en el ciclador térmico de ADN 480 de Perkins, Elmer, Cetus. Los

demás tampones y soluciones usados eran de diversas fuentes y de pureza de biología molecular.

Se efectuaron estudios sobre la línea celular epitelial bronquial humana inmortalizada en papilomavirus humano (BEP2D) (Willey *et al.*, Cancer Res. 51: 5370-5377, 1990). El aislamiento de ARN fue como sigue: se aisló el ARN basándose en el procedimiento descrito por Chomczynski y Sacchi (Analytical Biochemistry 162: 156-159, 1987). Se retiró el medio de cultivo de matraces que contenían la línea celular BEP2D. Inmediatamente, se dispuso tampón GIT (tiocianato de guanidinio 4,0 M, Tris-Cl 0,1 M, pH= 7,5, 1 % de β -mercaptoetanol) en las células (aproximadamente 500 μ l por 5-10 millones de células BEP2D). Se transfirieron entonces cada 500 μ l de tampón GIT que contenía las células lisadas a un tubo de microcentrifuga de 1,5 ml. Se añadió secuencialmente a cada tubo de microcentrifuga 50 μ l de acetato de sodio 2 M, pH 4, 500 ml de fenol saturado con agua y 100 ml de mezcla de cloroformo-alcohol isoamílico (49:1). Se agitaron concienzudamente entonces los tubos, se dispusieron sobre hielo durante 15 min y se microcentrifugaron durante 20 min a 14.000 rpm y 4 °C. Se transfirió la fase acuosa de cada tubo a un tubo reciente y se repitió la extracción anterior. De nuevo, se transfirió la fase acuosa de cada tubo a un tubo reciente con isopropanol (500 μ l), y se dispuso a -70 °C durante 15 min. Se microcentrifugaron entonces los tubos durante 20 min a 14.000 RPM y 4 °C. Se lavó el ARN dos veces con etanol al 70 % y se secó a vacío. Se suspendió el ARN en pirocarbonato de dietilo (DEPC) al 0,1 %, se trató con H₂O y se cuantificó por espectrofotometría (Gilford Instrument Spectrophotometer 260).

Transcripción inversa

Se realizó la transcripción inversa como sigue: se dispuso el ARN extraído en un tubo de microcentrifuga estéril. Por cada 1 μ g de ARN, se añadieron 0,2 mg de oligo-dT. Se calentó esto a 65 °C durante 5 min y se dispuso sobre hielo durante 1 min. Se añadió a esto 2 μ l de dNTP 1 mM, 2 μ l de tampón de transcriptasa inversa (RT) (Tris 500 mM, KCl 400 mM y MgCl₂ 80 mM), 0,5 μ l de RNasin y 1 μ l de transcriptasa inversa de AMV (9.500 unidades/ml). Se incubó esto a 42 °C durante 1 hora y se calentó a 80 °C durante 10 min, deteniendo la reacción. Se almacenó el ADNc resultante a -20 °C.

Preparación de cebadores y CT, amplificación por PCR y electroforesis en gel

La preparación de cebadores y moldes competitivos fue como sigue: se identificaron las secuencias adecuadas usando el software de análisis de cebador Oligo™ (National Biosciences, Hamel, Minn.). Se hicieron los cebadores usando un sintetizador de ADN modelo 391 PCR-Mate de Applied Biosystems. Las secuencias de cebador se describen a continuación.

Glutation peroxidasa (GSH-Px) (Chada *et al.*, Genomics 6: 268-271, 1990)

Se usaron los cebadores "externos" para amplificar tanto el ácido nucleico para medir como su molde competitivo y dar como resultado una longitud de producto de 354 pares de bases. Los cebadores "externos" son

SEC ID N° 1 (Chada *et al.*, Genomics 6: 268-271, 1990) Pos. 241 **5'-GGGCCTGGTGGTGCTTCGGCT-3'** (secuencia de codificación), que corresponde a las bases 241-261 de la secuencia clonada, y SEC ID N° 2 (Chada *et al.*, Genomics 6: 268-271, 1990) Pos. 574 **5'-CAATGGTCTGGAAGCGGCGGC-3'** (secuencia anticodificación), que se asocia con las bases 574-594.

Los cebadores "internos" usados para sintetizar el molde competitivo mutado retiran un sitio de reconocimiento de la endonucleasa de restricción EcoRI (GAATTC) cambiando un par de bases de ADNc nativo (bases en negrita). Los cebadores "internos" son:

SEC ID N° 3 (Chada *et al.*, Genomics 6: 268-271, 1990) Pos. 309 **5'-ATTCT GATTC CCTCAAGTACGTCGGCCT-3'** (sentido codificante)

SEC ID N° 4 (Chada *et al.*, Genomics 6: 268-271, 1990) Pos. 309 **3'-TAAGA CTAAGGGAGTTCATGCAGGCCGGA-5'** (sentido anticodificante).

Ambos cebadores corresponden a las bases 309-338 de la secuencia clonada. La mutación es el resultado de la sustitución con una T de la A nativa en posición 316 de la hebra codificante. La digestión con endonucleasa de restricción de GSH-Px nativa da productos de 280 y 74 pares de bases.

Gliceraldehído-3-fosfato deshidrogenasa (GAPDH) (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985)

Los cebadores "externos" usados para amplificar tanto el molde nativo como el mutado dan como resultado una longitud de secuencia de 788 o 790 pares de bases. Los cebadores "externos" son:

SEC ID N° 5 (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985) Pos. 46 **5'-GGTCGGAGTCAACGGATTGGTCG-3'** (sentido codificante) correspondiente a las bases 9-32 de la secuencia clonada, y SEC ID N° 6 (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985) Pos. 812

5'- CCTCCGACGCCTGCTTCACCAC-3' (sentido anticodificante) que asocia las bases 777-798.

Los cebadores "internos" usados para sintetizar el molde mutado crean un sitio de reconocimiento de la endonucleasa de restricción EcoRI (GAATTC) cambiando un par de bases de ADNc nativo (bases en negrita). Los cebadores "internos" son:

5 SEC ID N° 7 (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985) Pos. 234
5'-TGATCAATG GAATTC CCATCACCA-3' (sentido codificante)

SEC ID N° 8 (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985) Pos. 234
3'-ACTAGTTAC CTTAAG GGTAGTGGT-5' (sentido anticodificante).

10 Ambos cebadores corresponden a las bases 199-222 de la secuencia clonada. La mutación es el resultado de la sustitución por una T de la A nativa en posición 211 de la hebra codificante. La digestión con endonucleasa de restricción de la GAPDH mutada da productos de 588 y 200 pares de bases.

Se efectuaron varios experimentos usando un molde de GAPDH mutado diferente. Este molde tenía introducido un sitio de restricción de BamHI novedoso.

15 Los cebadores "externos" usados para amplificar tanto los moldes nativo como mutado dan como resultado una longitud de producto de 634 pares de bases. Los cebadores "externos" son:

SEC ID N° 9 (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985) Pos. 200
5'- CATGGCACCGTCAAGGCTGAGAAC-3' (secuencia de codificación) correspondiente a las bases 165-188 de las secuencias clonadas, y SEC ID N° 10 (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985) Pos. 813
5'- CCTCCGACGCCTGCTTCACCAC-3' (sentido anticodificante) que se asocia con las bases 777-798.

20 Los cebadores "internos" usados para sintetizar el molde mutado crean un sitio de reconocimiento de la endonucleasa de restricción BamHI (GGATCC) cambiando un par de bases de ADNc nativo (bases en negrita). Los cebadores "internos" son:

SEC ID N° 11 (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985) Pos. 368
5'-CAGGGG GGATCC AAAAGGGTCATCAT-3' (sentido codificante).

25 SEC ID N° 12 (Tso *et al.*, Nucleic Acids Res. 13: 2485-2502, 1985) Pos. 368
3'-GTCCCC CCTAGG TTTTCCCAGTAGTA-5' (sentido anticodificante).

Ambos cebadores corresponden a las bases 333-358 de la secuencia clonada. La mutación es el resultado de la sustitución por una T de la G nativa en posición 342 de la hebra codificante. La digestión con endonucleasa de restricción de esta GAPDH mutada produce productos de 460 pares de bases y 174 pares de bases.

30 Se prepararon los moldes competitivos de patrón interno mutados mediante mutagénesis dirigida a sitio como se describe en Higuchi *et al.*, Nucleic Acids Res. 16: 7351-7367, 1988. Estas mutaciones de bases individuales dieron como resultado la ganancia (GAPDH) o pérdida (GSH-Px) de un sitio de reconocimiento de endonucleasa de restricción EcoRI. (Los experimentos se realizaron también usando una GAPDH mutada con un sitio BamHI introducido). Para cada producto mutado, dos reacciones en cadena de la polimerasa iniciales que usan un cebador "externo" y un cebador "interno" con desalineamiento de una sola base producen dos fragmentos de ADN superpuestos. (Los cebadores 1 y 4, 2 y 3 para GSH-Px y los cebadores 5 y 8, 6 y 7 para GAPDH). Se sometieron a electroforesis estos fragmentos de ADN superpuestos en gel de agarosa al 3 % de Nusieve y 1 % de LE teñido con bromuro de etidio. Se recortaron las bandas y se purificaron usando un filtro de 0,45 µm Millipore Ultrafree MC (Nihon Millipore Kogyo K.K., Yonezawa, Japón). Se precipitó con etanol el ADN purificado, se lavó, se secó a vacío y se suspendió en 100 µl de H₂O estéril. Se amplificó 1 µl de cada uno de los dos fragmentos de ADN superpuestos por PCR usando solo los cebadores externos. Se efectuó el primer ciclo de PCR sin cebadores para permitir la formación de heterodímeros. Se formó por tanto el producto mutado completo y se amplificó. Se purificó en gel el producto de PCR mutado como se describe anteriormente y se reamplificó formando producto bruto. Se purificó en gel el producto bruto y se midió espectrofotométricamente. Se diluyeron los productos mutados al intervalo attomolar para uso como moldes competitivos. Se usó ADN de esperma de arenque (Lofstrand, Bethesda, Md.) 1 µg/ml como portador. Se efectuó la digestión con endonucleasa de restricción en muestras de cada molde mutado para asegurar la falta de contaminación.

50 Las condiciones de PCR fueron las siguientes: se estandarizaron las condiciones de PCR para cada experimento usando una mezcla maestra que contenía 1x tampón de PCR (KCl 50 mM, Tris-HCl 10 mM, pH 9,0, MgCl₂ 1,5 mM), 25 pmoles de cebadores que codifican GSH-Px y GAPDH, dNTP 0,2 mM (A,T,C,G) y cantidades constantes de ambos patrones internos por mezcla de reacción de 100 µl. Se añadió ADN polimerasa Taq (2,5 unidades) a cada reacción de 100 µl antes de la amplificación. Se diluyó en serie el ADNc obtenido de la línea celular BEP2D y se

añadió a tubos de PCR de muestra. En todos los experimentos, se amplificaron tubos de control que no contenían molde, solo ADNc nativo o solo moldes competitivos mutados, para comprobar la contaminación o digestión enzimática completa.

- 5 Se llevó a cabo la amplificación por PCR durante 35 ciclos a 94 °C durante 1 min, a 60 °C durante 1 min y a 72 °C durante 1 min. Después de la amplificación, se calentaron los productos de PCR durante 10 min para maximizar la formación de heterodímero.

10 La cuantificación de los productos fue como sigue: se digirieron muestras (40 µl) de cada tubo de PCR con endonucleasa de restricción EcoRI durante 12-16 horas (los experimentos realizados usando GAPDH mutada con el sitio de restricción de BamHI novedoso se digirieron también con endonucleasa de restricción BamHI durante 4-5 horas). Se aislaron estos productos por electroforesis en gel de agarosa al 3 % de Nusieve y 1 % de LE teñido con bromuro de etidio durante 2-3 horas a 60 V. Se tomó una fotografía negativa del gel usando película instantánea positiva/negativa Polaroid 665.

15 Se sometió la fotografía negativa a densitometría (Zeineh Soft Laser Scanning Densitometer, modelo SLR 2D/1D usando el software Zeineh ID Autostepover Videophoresis Program, Biomed Instruments, Fullerton, Calif.). Como alternativa, se evaluó densitométricamente el gel teñido directamente usando una cámara digital, o se evaluó en un gel de secuenciación automatizada tal como el ofrecido por Applied Biosystems, Inc.). Se calcularon las áreas bajo la curva y se usaron para cuantificación. Se hicieron correcciones por los tamaños de banda relativos y la formación de heterodímero. Se expresaron los datos como cocientes relativos de GSH-Px a GAPDH.

20 En un segundo conjunto de experimentos, se prepararon reacciones en cadena de la polimerasa con transcriptasa inversa competitivas múltiples (MC RT-PCR) con moldes competitivos mediante el procedimiento de Celi, evaluando el gen del citocromo p450 (CYP 1A1) en células BEP2D expuestas a β-naftoflavona. Se evaluó la inducción de la expresión del gen CYP1A1 usando tanto MC RT-PCR con moldes competitivos de Celi como análisis Northern. Se prepararon los moldes competitivos tanto para genes CYP1A1 como GAPDH. Los cebadores usados para preparar el molde competitivo de GAPDH fueron:

25 SEC ID N° 13 (Tokunaga *et al.*, Cancer Res. 47: 5616-5619, 1990) Pos. 75
5'-GGT CGG AGT CAA CGG ATT TGG TCG-3' Pos. 94 y SEC ID N° 14 (Tokunaga *et al.*, Cancer Res.
 47: 5616-5619, 1990) Pos. 822. V Pos. 636 Pos. 842
5'-CCT CCG ACG CCT GCT TCA CCC CAT CAC GCC ACA GTT TCC C-3' Pos. 616.

30 El cebador externo inferior usado junto con la SEC ID N° 13 para amplificar tanto el molde competitivo como el nativo fue:

SEC ID N° 15 (Tokunaga *et al.*, Cancer Res. 47: 5616-5619, 1990) Pos. 842
5'-CCT CCG ACG CCT GCT TCA CC-3' Pos. 822. Los cebadores usados para preparar el molde competitivo para CYP1A1 fueron:

35 SEC ID N° 16 (Jaiswal *et al.*, Science 228: 80-83, 1989) Pos. 1241
5'-CAT CCC CCA CAG CAC AAC AAG-3' Pos. 1262 y:

SEC ID N° 17 (Jaiswal *et al.*, Science 228:80-83, 1989) Pos. 1555. V Pos. 1428
 Pos. 1575 **5'-ACA GCA GGC ATG CTT CAT GGG TCT CAC CGA TAC ACT TCC G-3'** Pos. 1448.

El cebador externo inferior usado junto con la SEC ID N° 18 para amplificar tanto el molde competitivo como el nativo fue:

40 SEC ID N° 18 (Jaiswal *et al.*, Science 228: 80-83, 1989) Pos. 1575
5'-ACA GCA GGC ATG CTT CAT GG-3' Pos. 1555.

Las condiciones de amplificación por PCR eran las mismas que se describen para experimentos que usan los moldes competitivos preparados para GAPDH y GSHPx mediante el procedimiento de Higuchi, excepto porque la temperatura de asociación era de 55 °C y la amplificación se llevó a cabo durante 38 ciclos.

45 Debido a que los moldes nativo y competitivo se separan sin digestión con endonucleasa de restricción anterior, se tomaron directamente las muestras del tubo de reacción PCR y se aplicaron a geles de agarosa de 3 % de Nusieve, 1 % de LE teñidos con bromuro de etidio. Fue entonces posible cuantificar los productos tomando una fotografía negativa del gel usando una película instantánea positiva/negativa Polaroid 665, sometiendo la fotografía negativa a densitometría.

50 Se sometió a electroforesis ARN de células BEP2D incubadas durante tiempos variables con β-naftoflavona (10 µM) en gel de 1 % de LE desnaturante con formaldehído para análisis Northern o se amplificó MC RT-PCR, como se describe anteriormente. Para el análisis Northern, después de transferencia del ARN a GeneScreen, se hibridaron

los filtros con ADNc de CYP1A1 marcado con ³²P.

El procedimiento usado para la cuantificación por PCR es como sigue: se coamplificaron diluciones en serie de ADNc de BEP2D (que representa de 0,25 µg a 0,05 µg de ARN total) con cantidades constantes de cada molde competitivo de patrón interno mutado en una sola base (10 attomoles cada uno) y se analizaron entonces como se describe anteriormente.

La Figura 34 ilustra fotografías negativas de los geles analizadas por densitometría para cuantificar cada banda. Partiendo del área bajo la curva obtenida por la evaluación densitométrica de las bandas, se calcularon los cocientes de producto amplificado de molde nativo/competitivo como sigue. Se hicieron correcciones para los tamaños de banda relativos (se multiplicó el molde competitivo de GAPDH por 788/588 cuando se compara con el ácido nucleico nativo de GAPDH y se multiplicó el GSH-Px nativo por 354/280 cuando se compara con el molde competitivo de GSH-Px).

Durante la PCR, en condiciones en que el cebador es limitante, pueden asociarse hebras individuales de ADN con homología de secuencia, formando heterodímeros (Gilliland, G., Perrin, S., Blanchard, K. y Bunn, H. F. (1990) Proc. Natl. Acad. Sci. 87: 2725-2729). Cuando las hebras heterólogas difieren en solo un par de bases, como en este ejemplo particular, las hebras heterólogas pueden reasociarse aleatoriamente (Gilliland *et al.*, *supra*; Thompson, J. D., Brodsky, L, and Yunis, J. J. (1992) Blood 79: 1629-1635), como se muestra en el cuadro de Punnett siguiente:

	N	M
N	NN	NM
M	MN	MM

en que N= la proporción de productos nativos monocatenarios antes de la reasociación, M= la proporción de productos mutados monocatenarios antes de la reasociación, NN (o N²)= la proporción de productos nativos bicatenarios después de la reasociación, 2NM= la proporción de heterodímero formado después de la reasociación y MM (o M²)= la proporción de productos mutados bicatenarios después de la reasociación.

Los heterodímeros se representaron indirectamente porque no se cortaron, en este ejemplo, por la enzima de restricción y tenían la misma movilidad electroforética que el homodímero no digerido. Por lo tanto, los heterodímeros se leyeron densitométricamente junto con el homodímero no digerido. Para cuantificar los productos basándose en la distribución del cuadro de Punnett, se promovió la formación de heterodímero aleatorio después de la PCR. Esto se realizó (según los procedimientos descritos en Gilliland *et al.*, *supra*, y Thompson *et al.*, *supra*), calentando los productos a 100 °C durante 10 min, seguido de enfriamiento lento. Después de promover la formación de heterodímeros, se determinó la cantidad de cada producto mediante análisis de los datos densitométricos usando una fórmula cuadrática, ya que la formación de heterodúplex sigue una distribución binomial en estas condiciones (Gilliland *et al.*, Proc. Natl. Acad. Sci. 87: 2725-2729 (1990), Becker-Andre *et al.*, Nucleic Acids Res. 17: 9437-9446 1989).

Para GAPDH, en este ejemplo, no se escindieron ni el producto nativo (NN) ni el heterodímero (NM) por EcoRI. Por lo tanto, la banda mayor representaba tanto homodímero de GAPDH nativo (NN) como heterodímero NM. Esta banda se presentaba aritméticamente por N² +2NM, según el cuadro de Punnett, mientras que la proporción de la banda resultante de escisión por EcoRI se representaba por el valor M². Por lo tanto, cuando la cantidad de molde nativo (N) y mutado (M) es igual (1:1) antes de la PCR, después de aleatorizar la formación de heterodímero, el cociente aparente será de 3:1 [N² +2NM): M²]. Para ilustrar además esto, se muestran en la Figura 35 los datos densitométricos brutos del carril de la primera muestra (mostrado en la Figura 34) y se procesan matemáticamente hasta los cocientes finales siguientes:

El valor de M² es conocido (2.214), así como el valor de N² + 2NM (10.095). A partir de esta información, se calcula M (47,05) y al resolver N da como resultado la ecuación cuadrática (aX²+bX+c= 0): N²+2N(47,05)-10.095= 0.

Se usa la fórmula cuadrática (N=-b±√[(b²-4ac)/2a] para resolver N. En este caso, a= 1, b= 94,1, c= 10.095 y por tanto N= 63,89. La información buscada es el cociente de N/M, que es 63,89/4710 o 1,36/1. (Aunque se resuelven las proporciones de ADN monocatenario presentes después de la PCR, serán idénticas a aquellas del correspondiente ADN bicatenario presente antes de la PCR, en este ejemplo).

Puesto que los valores densitométricos son relativos, es posible evitar la inconveniencia de usar la fórmula cuadrática asignando a la bandas valores densitométricos proporcionados que cuando se suman son = 1 o (N²+2NM) + M² = 1. Resolviendo esta ecuación: (N²+2NM) + M²= (N+M)² = 1 y por lo tanto N+M= 1.

Las fracciones relativas de 1 asignadas a cada una de la bandas se determinan por sus valores densitométricos respectivos (Figura 35). Puesto que el valor densitométrico total de ambas bandas es de 12.309 (10.095+2.214), la proporción relativa de la banda mayor (N²+2NM) es 0,82 (10.095/12.309) y la proporción relativa de la banda menor (M²) es 0,18 (2.214/12.309). Por tanto, la proporción de homodímero de GAPDH mutado (M²) es 0,18, y la proporción de GAPDH (M) mutado monocatenario es 0,424. Puesto que N+M =1, la proporción de GAPDH (N) nativo

monocatenario es 1-0,424 o 0,576, y el cociente de producto nativo a mutado es 0,576/0,424 o, como se calcula anteriormente, 1,36/1.

A continuación, se llevan a cabo en este ejemplo los mismos cálculos usando los valores densitométricos para GSH-Px nativo y mutado del mismo carril que los valores de GAPDH anteriores (Tabla 1):

5 $N^2 = 0,558$, $N = 0,747$ y $M = 1-0,747 = 0,253$

Se obtienen cocientes de nativo/mutado:

GSH-Px nativo/mutado = $0,747/0,253 = 2,95/1$

GAPDH nativo/mutado = $0,576/0,424 = 1,36/1$

Los valores finales se expresan como cociente de probabilidades (por ejemplo, "cociente de cocientes"):

10 GSH-Px nativo/mutado: GAPDH nativo/mutado = $2,95/1,36 = 2,17/1$

Como ilustra la Figura 19, la relación entre la cantidad de producto nativo (en unidades densitométricas arbitrarias) y el ARN de partida total no permanecía lineal a lo largo de la amplificación por PCR para GSH-Px ni GAPDH.

15 Como ilustra la Figura 20, sin embargo, la relación de los cocientes de GSH-Px nativo/molde competitivo y GAPDH nativo/molde competitivo a ARN total de partida era lineal para ambos genes. Promediando el cociente de molde nativo/competitivo de GSH-Px a molde nativo/competitivo de GAPDH obtenido de tubos de muestra (2,17:1, 2,14:1, 2,00:1, 1,76:1, 2,46:1, 2,71:1 y 1,92:1), se obtuvo un valor de 2,17:1 con una DE de 0,33. En este ejemplo, ningún valor variaba más de un 25 % de la media.

20 Para valorar la variabilidad de esta técnica, se repitió el experimento usando diferentes diluciones de patrones mutados (molde competitivo) y mezcla maestra. Promediando el cociente de molde nativo/competitivo de GSH-Px a GAPDH de molde nativo/competitivo obtenido de cada tubo de muestra en este ejemplo (1:9,09, 1:8,13, 1:9,43, 1:8,13, 1:6,62, 1:8,77, 1:7,69, 1:10,00, 1:7,58 y 1:7,04), se obtuvo un valor medio de 1:8,25 con una DE de 1,07. En este ejemplo, ningún valor variaba más de un 22 % de la media.

25 Para valorar la variabilidad entre muestras usando las mismas mezcla maestra y diluciones de patrones mutados (usando GAPDH mutado con sitio de restricción de BamHI novedoso), se extrajo independientemente ARN de BEP2D de tres matraces separados y se transcribió de forma inversa a ADNc. Se efectuaron diluciones 5x del ADNc. Se procesaron en cada estudio 4 tubos de PCR. Los cocientes obtenidos de molde nativo/competitivo de GSH-Px a molde nativo/competitivo de GAPDH era 15,01:1, 17,69:1 y 21,76:1. (media = 18,15, DE = 3,40). En este ejemplo, los 3 valores estaban dentro del 20 % de la media.

30 Como ilustra la Figura 36, se observó un aumento similar en la expresión génica del gen CYP1A1 tanto en análisis Northern como en algunas realizaciones de los procedimientos dados a conocer en la presente memoria. La Figura 36A ilustra el análisis Northern de ARN obtenido de células BEP2D que se trataron con DMSO al 0,1% como control o β -naftoflavona en un esfuerzo por inducir citocromo p450 IAI (CYP1A1). La Figura 36B ilustra ADN amplificado por PCR de diluciones en serie de ADNc de las mismas células usadas en la Figura 36SA. Se coamplificó el ADNc en presencia de moldes competitivos de GAPDH y CYP1A1, según algunas realizaciones de la presente invención.

35 Comparando las bandas del gen de referencia GAPDH que representan ADNc de molde nativo y competitivo, se indica que se cargó aproximadamente la misma cantidad de ADNc en el carril con 1 μ l de espécimen de las células de control y el carril con 3 μ l de espécimen de células expuestas a β -naftoflavona. Por ello, la banda que representa el gen CYP1A1 nativo está mucho más fuertemente representada en el carril que contiene ADN de células expuestas a β -naftoflavona en comparación con células de control.

40 Para valorar la variabilidad de esta técnica, se efectuó una repetición del experimento anterior usando diferentes diluciones de moldes competitivos y mezcla maestra. Promediando el cociente obtenido de cada tubo de muestra (1:9,09, 1:8,13, 1:9,43, 1:8,13, 1:6,62, 1:8,77, 1:7,69, 1:10,00, 1:7,58 y 1:7,04), se obtuvo un valor medio del cociente de molde nativo/competitivo de GSH-Px a molde nativo/competitivo de GAPDH de 1:8,25 con una DE de 1,07. En este ejemplo, ningún valor variaba más de un 22 % de la media, indicando la precisión de esta técnica y la variabilidad introducida por las nuevas mezclas maestras que contienen nuevas diluciones de patrones competitivos.

45 Para valorar la variabilidad entre muestras usando las mismas mezcla maestra y diluciones de moldes competitivos, se extrajo independientemente ARN de BEP2D de 3 matraces separados y se transcribió de forma inversa a ADNc. Solo se efectuaron diluciones bajas (5 veces) del ADNc. Se procesaron 4 tubos de PCR para cada estudio. Los cocientes obtenidos de molde nativo/competitivo de GSH-Px a molde nativo/competitivo de GAPDH eran 15,01:1, 17,69:1 y 21,76:1 (media = 18,15, DE = 3,40). En este ejemplo, los 3 valores estaban todos dentro del 20 % de la media, indicando la precisión de esta técnica cuando se comparan muestras que se han transcrito de forma inversa independientemente pero amplificadas con las mismas mezcla maestra y diluciones de patrón interno. El análisis Northern de ARN de BEP2D revela un cociente de ARNm de GSH-Px/GAPDH de aproximadamente 1:8.

EJEMPLO IV*Estudio interlaboratorio con anonimato para evaluar la reproducibilidad*

5 En un primer estudio, participaron 6 laboratorios en la medida por triplicado de 5 genes en ADNc derivado de una muestra de tejido de carcinoma broncogénico 16009T. Se usaron una variedad de procedimientos de electroforesis y programas de software de imágenes en los diferentes laboratorios para analizar el producto amplificado. El estudio 1, laboratorio 2, usó Agilent 2100 Bioanalyzer. El CV medio intralaboratorio para todas las medidas de expresión génica era de 0,36, que es comparable con el reseñado anteriormente (Willey *et al.*, 1998; Rots *et al.*, 1999; Rots *et al.*; 2000; Mollerup *et al.*, 1999; Loitsch *et al.*, 1999). La variación interlaboratorio mostró un CV medio de 0,71.

10 En un segundo estudio, se usó electroforesis en gel plano y el software NIH Image para medir la expresión de 10 genes (lo 5 medidos anteriormente más 5 genes adicionales) en ADNc de A549. 4 de los laboratorios originales pudieron participar en el segundo estudio. El CV medio combinado de los 9 genes que pudieron medirse era de 0,27 y 0,48 para la comparación intralaboratorio e interlaboratorio, respectivamente. Para TNF- α , cada laboratorio determinó que la expresión era demasiado baja para cuantificarse. De los 4 laboratorios, 3 laboratorios pudieron cuantificar HNF3 α , mientras que el cuarto laboratorio no pudo. Se estableció el límite inferior de detección de un producto de PCR por encima del fondo para el segundo estudio como un valor densitométrico arbitrario en imágenes de NIH de 5 por encima del fondo. Aunque el cuarto laboratorio observó los productos de PCR de NT y CT para HNF3 α , estaban por debajo del nivel de corte de 5 y por lo tanto no se incluyeron en el análisis. Se usó una mezcla de CT que contribuía con 60 moléculas de ácido nucleico de CT (mezcla F) para detectar HNF3 α .

EJEMPLO V20 *Comparación con micromatriz de oligonucleótidos*

En un primer estudio, se usaron cada uno de matrices de oligonucleótidos Affymetrix y una realización del enfoque en dos etapas de la presente invención para medir la expresión de un total de 22 genes enzimáticos o antioxidantes metabólicos xenobióticos en células epiteliales orales humanas (HOE). Se comparó la expresión en células HOE normales con la expresión en la línea celular epitelial bucal inmortalizada SVpgC2a. Se comparó la diferencia de expresión entre HOE y SVpgC2a para cada gen. Se compararon entonces las diferencias detectadas por micromatriz con las diferencias detectadas por StaRT-PCR.

El procedimiento de chip de ARNm da resultados basados en señales relativas que se basan en coincidencias perfectas y desapareamientos de hibridaciones de chip, en combinación con análisis del software de imágenes, derivando resultados de las intensidades de hibridación. La técnica permite el análisis de hasta 12.000 genes simultáneamente y es semicuantitativo en términos de intensidades de señal de un experimento de hibridación en comparación con otras hibridaciones. En este estudio, se efectuaron tres conjuntos de hibridaciones que utilizan muestras de células HOE normales o células SVpgC2a, y se compararon entonces los datos de cada medida en células SVpgC2a con los datos correspondientes a células HOE. Se efectuó la primera de las tres hibridaciones con el chip HuFL 6800.

35 Las medidas de expresión génica estaban en estrecha concordancia y demostraban que los niveles de expresión de varios transcritos del metabolismo de fase I y II eran similares en queratinocitos normales e inmortalizados. De los genes CYP analizados, la mayoría se expresaba a bajos niveles, concretamente, menores de 20 moléculas de ARNm/ 10^6 moléculas de ARNm de β -actina. En este estudio, la concentración de ADNc permitió la cuantificación de niveles de ARNm de 4 moléculas/ 10^6 moléculas de β -actina.

40 Los procedimientos que comprenden realizaciones de la presente invención eran más sensibles. Por ejemplo, los transcritos que se encontraron expresados a bajos niveles no se detectaron con el procedimiento de chip, concretamente, transcritos expresados a niveles por debajo de unos cientos de moléculas/ 10^6 copias de β -actina no eran detectables con el procedimiento de hibridación de chip. Usando los procedimientos de la presente invención, se obtuvo un valor de expresión génica tanto en células normales como inmortalizadas para 14 genes de los 22 genes evaluados (para los 7 genes restantes, la expresión era demasiado baja para cuantificar en las células normales o inmortalizadas). De estos 14 genes, se obtuvo también un valor de expresión génica mediante análisis de micromatriz tanto en células normales como inmortalizadas para solo 5 de ellos. Se comparó la diferencia de expresión de estos 5 genes en células normales e inmortalizadas. Se presentan los resultados en la Figura 37.

50 En un segundo estudio, se midió la expresión génica para ARN de referencia humano de Stratagene. El ARN de referencia humano de Stratagene comprende ARN de 10 líneas celulares mezclado conjuntamente, para representar los niveles de transcripción de una gran fracción de genes del genoma humano. Se evaluaron los genes usando micromatriz de oligonucleótidos (Affymetrix U95 versión 2 y chips genéticos HuGenFL) y usando realizaciones de la presente invención.

55 Usando realizaciones de la presente invención, se obtuvieron datos de 163 de los 192 genes representados en mezclas de CT de los sistemas 1 y 2 de G.E.N.E. a partir del ARN de referencia humano de Stratagene. Los genes restantes representados en los sistemas 1 y 2 se expresaban a un valor demasiado bajo para cuantificarse (menos de 6 moléculas/ 10^6 moléculas de β -actina). De los 163 genes medidos, 85 se representaban en el chip genético

HuGenFL. De estos 85 genes, fue posible asignar un valor de expresión a todos los genes medidos de acuerdo con realizaciones de la presente invención, pero solo a 41 genes basándose en el análisis de micromatriz.

5 Análisis de correlación bivariada: La Figura 38 muestra la correlación de Pearson para algunos valores de realizaciones en dos etapas frente a micromatriz. La correlación de Pearson tenía un valor de r^2 de 0,373, que era altamente significativo ($P < 0,001$).

10 Sensibilidad: La Figura 3 compara también la sensibilidad entre algunas realizaciones en dos etapas y micromatrices. Entre los 41 genes para los que se obtuvieron valores de expresión mediante ambos, los valores oscilaban dentro de aproximadamente dos unidades logarítmicas en el análisis de micromatriz y aproximadamente tres unidades logarítmicas para algunas realizaciones de análisis en dos etapas. Figura 38. Algunas realizaciones de análisis en dos etapas eran aproximadamente 10 veces más sensibles que el análisis por micromatriz.

EJEMPLO VI

El siguiente ejemplo detalla la medida de expresión génica en el procedimiento comercial proporcionado como servicio.

15 Preparación automatizada de reacciones: Se usa un manipulador robótico de líquidos PerkinElmer (Boston, Massachusetts, EE.UU.) para preparar reacciones PCR de 10 μ l en microplacas de 96 o 384 pocillos. En primer lugar, se programa el manipulador de líquidos para distribuir 1 μ l de cebadores de genes para medir a pocillos de las microplacas. En segundo lugar, se prepara para cada ADNc un volumen suficiente de mezcla de PCR para el número previsto de medidas de expresión génica, que contiene tampón, polimerasa Taq, dNTP, ADNc y patrones internos. Entonces, el robot distribuye 9 μ l de esta mezcla de reacción PCR en cada pocillo. Por tanto, en cada pocillo están presentes los moldes competitivos de patrones internos de cada gen y ADNc al mismo cociente. Sin embargo, debido a que solo está presente un par de cebadores en cada pocillo, solo se amplifica en cada pocillo un solo gen y su molde competitivo de patrón interno respectivo. Después de 35 ciclos de PCR, se transfiere cada microplaca a un dispositivo microfluídico de alto rendimiento AMS 90 SE30 (Caliper/Zymar, Hopkinton, Massachusetts, EE.UU.) para análisis.

25 *Diseño de la medida de expresión génica de alto rendimiento*

30 Etapa 1. Amplificación de 96 genes: Pueden combinarse los moldes competitivos en grupos de 96 y nombrarse como "sistemas" secuenciales. Por tanto, la primera mezcla de CT que representa 96 genes se llamó sistema 1 y así. Puede combinarse una mezcla de pares cebadores específicos de cada uno de los 96 genes en un sistema y diluirse a una concentración de 0,05 μ g. Por tanto, para cada uno de los sistemas 1-4 de mezclas de CT que representan 96 genes, puede haber una mezcla de cebadores correspondiente a los mismos 96 genes.

35 La muestra de ADNc puede diluirse de modo que esté en equilibrio con, concretamente calibrada con, aproximadamente 600.000 moléculas de β -actina de la mezcla D de cada sistema. Se amplificará por PCR entonces la cantidad apropiada de ADNc en presencia de la mezcla de cebadores de uno de los sistemas y la mezcla B, C, D, E o F del correspondiente sistema. De este modo, pueden generarse productos de PCR en cada una de las 20 reacciones PCR separadas de 10 μ l, a saber, sistema 1, mezclas B, C, D, E y F; sistema 2, mezclas B, C, D, E y F; sistema 3, mezclas B, C, D, E y F y sistema 4, mezclas B, C, D, E y F.

Hasta aquí, se ha consumido la cantidad de ADNc y mezcla de CT para las 20 reacciones. A continuación, pueden usarse los productos de PCR incluidos en estas 20 reacciones PCR para medir los 384 genes incluidos en los sistemas 1-4.

40 Etapa 2: Amplificación por PCR del gen individual inicial a partir de productos de PCR de la mezcla D: Los productos de PCR generados anteriormente en la ronda 1 pueden diluirse hasta 10.000 veces y seguir siendo productos de PCR cuantificables en una segunda ronda de amplificación por PCR. Debido a que se incluyó un CT patrón interno en cada reacción PCR, y debido a que la eficacia de amplificación del CT patrón interno es la misma que la del NT, el valor de expresión génica obtenido después de una segunda ronda de 35 ciclos de amplificación por PCR puede ser el mismo que el obtenido después de los primeros 35 ciclos. Por ejemplo, pueden diluirse 2 μ l de la primera ronda de amplificación 100 veces para uso en la segunda ronda. Debido a que había 10 μ l en el volumen de reacción PCR de la primera ronda, esto constituye una dilución de 1.000 veces. Puede mezclarse entonces la mitad de esta reacción PCR de la ronda 1 diluida de la mezcla D de cada sistema (100 μ l) con la cantidad apropiada de dNTP, taq y H₂O y tomarse alícuotas en cada uno de los 96 pocillos, conteniendo cada pocillo un par diferente de cebadores (que representan un gen del correspondiente sistema) secados en el fondo. Los 100 μ l restantes pueden apartarse para la etapa 3. Por tanto, puede prepararse una mezcla de reacción PCR de etapa 2 que contiene producto de PCR de mezcla D de la etapa 1 diluido para cada uno de los 4 sistemas, y distribuirse en alícuotas de 10 μ l en cada uno de 96 pocillos de la microplaca de 384 pocillos.

55 Después de la amplificación por PCR, puede transferirse 1 μ l de producto de PCR de cada reacción PCR a un pocillo de un chip DNA 1000 montado en Agilent 2100. Puede someterse a electroforesis cada muestra. El producto de PCR restante puede recuperarse en tubos de microcentrifuga, proporcionando material de respaldo en el caso de problemas con la electroforesis.

Etapa 3: Selección del producto de la etapa 1 de mezcla de CT más apropiado para cuantificar cada uno de los 96 genes, basándose en los resultados de la etapa 2: Se espera que solo aproximadamente un 50 % de los NT y CT de los diversos genes estén en equilibrio. Los genes no equilibrados pueden reevaluarse usando una mezcla que contiene diferentes concentraciones de CT para esos genes respecto al CT de β -actina. En la ronda 3, puede evaluarse la expresión de cada gen usando una reacción PCR de etapa 1 que contenía una concentración apropiada de CT de esos genes respecto a CT de β -actina. Puede prepararse una mezcla de reacción PCR de 3.900 μ l que contiene 3,9 μ l de producto de PCR de la etapa 1 y volúmenes apropiados de polimerasa Taq, dNTP y H₂O destilada.

Pueden prepararse por adelantado microplacas de 384 pocillos con cebadores para diferentes genes individuales secados en el fondo de cada pocillo. Dichas placas pueden almacenarse a 4 °C durante meses sin pérdida ni reducción de la función cebadora. No obstante, no se esperaría que una reducción de la eficacia de cebador con el tiempo cambiara los valores numéricos de expresión génica, porque cada reacción PCR contiene un CT patrón interno. Sin embargo, podría reducirse la cantidad de producto amplificado tanto para NT como CT, reduciendo la señal para cuantificar en Agilent 2100.

Análisis del coste de proporcionar realizaciones de procedimiento comercial como servicio

Usando termocicladores de microplaca de 384 pocillos, pueden cribarse dos muestras de ADNc en 8 reacciones de etapa 1 de 96 genes diferentes. 4 de estas reacciones pueden contener ADN de muestra 1 y mezcla D de los sistemas 1-4, respectivamente, y 4 reacciones pueden contener ADNc de muestra 2 y mezcla D de los sistemas 1-4, respectivamente. Pueden diluirse los productos de PCR de cada una de estas reacciones 100 veces. Pueden prepararse 8 mezclas de reacción PCR, suficientes cada una para 96 reacciones PCR, con un 10 % en volumen representado por uno de los 8 productos de PCR de la etapa 1. Pueden dispensarse entonces las mezclas de reacción PCR que incluyen la muestra 1 a los 96 pocillos apropiados de una de las microplacas de 384 pocillos, y pueden dispensarse las mezclas de reacción PCR que incluyen la muestra 2 a los 96 pocillos apropiados de la otra placa de 384 pocillos. Como alternativa, cuando se usan los sistemas 5-8, pueden usarse ambas microplacas de 384 pocillos para cribar 768 genes en la muestra 1 con la mezcla D.

Preparar las 8 reacciones PCR de 96 genes diferentes puede requerir un total de 8 μ l de ADNc y 8 μ l de mezcla D de CT. El coste primario puede ser la polimerasa Taq y los cebadores para 768 reacciones PCR, aproximadamente 25,00 \$ y 30,00 \$, respectivamente. El coste primario del análisis puede ser los chips Agilent, que cuestan 12 \$/chip. Al aplicar 4 productos de PCR/canal, pueden analizarse 48 genes/chip. Por tanto, el coste de materiales por ensayo puede ser de aproximadamente 0,32 \$. Los costes de laboratorio pueden incluir un día de trabajo de aproximadamente 14,00 \$/hora= 112,00 \$, sumando aproximadamente 0,15 \$/ensayo para dar aproximadamente 0,47 \$/ensayo. Puede haber aproximadamente 2 días más de entrada de datos y análisis, llevando el coste total a 0,77 \$/ensayo. Basándose en estos cálculos, la tarifa puede ser de aproximadamente 1,00 \$/medida de expresión génica.

Secuencias

ADN

Homo sapiens

aataatctct cccattctg g

ADN

Homo sapiens

tgtggttgag aatgagcatg t

ADN

Homo sapiens

tgtggttgag aatgagcatg tggataccac ctctctaga gt

ADN

Homo sapiens

atgacacaga gctggtagcc

ADN

Homo sapiens

aaccagaaaa tacgagccct

ADN

Homo sapiens

aaccagaaaa tacgagccct tctccatcta ccacaggcac

5 ADN

Homo sapiens

aaaactgcta aggccaaggt

ADN

Homo sapiens

10 **tcagcaacct ctttctcac**

ADN

Homo sapiens

tcagcaacct ctttctcac tctgattcat ctgtctgcc

ADN

15 *Homo sapiens*

accattgtcc agccatcagc

ADN

Homo sapiens

accctctgct gtccgtgtct

20 ADN

Homo sapiens

accctctgct gtccgtgtct ctgaaggagg atggagtctg

ADN

Homo sapiens

25 **ttttaggaga ccgaagtccg**

ADN

Homo sapiens

agccaacgtg ccatgtgcta

ADN

30 *Homo sapiens*

agccaacgtg ccatgtgcta cctctgttcc ttccctctac

ADN

Homo sapiens

gtaccggcgg gcattcagtg

35 ADN

Homo sapiens

agagtgagcc cagcagaacc

ADN

Homo sapiens

5 **agagtgagcc cagcagaacc cgttctctg gatccaaggc**

ADN

Homo sapiens

tacgcagcgc ctccctccac

ADN

10 *Homo sapiens*

ctgttctcgt cgtttccgca

ADN

Homo sapiens

ctgttctcgt cgtttccgca accttggggg ccttttcatt

15 ADN

Homo sapiens

tacgacaagg atagaagcgg

ADN

Homo sapiens

20 **aggacatgac gctctttctg**

ADN

Homo sapiens

aggacatgac gctctttctg gatttccttt ttgttttct cg

ADN

25 *Homo sapiens*

ggatacaaag aaggagtg

ADN

Homo sapiens

ccaactcagg acaaggtaca

30 ADN

Homo sapiens

ccaactcagg acaaggtaca tcttctgggc tcttggggc

ADN

Homo sapiens

35 **ccagaagaaa gcggtaaga**

ADN

Homo sapiens

aaccttcatt ttcccctggg

ADN

5 *Homo sapiens*

aaccttcatt ttcccctggg ccagtgatga gcgggttaca

ADN

Homo sapiens

ggccttgcca gagcttttg aatacc

10 ADN

Homo sapiens

agccatttc atccaagtt ttgaca

ADN

Homo sapiens

15 **agccatttc atccaagtt ttgacagtg attaattct gaatcccat gg**

ADN

Homo sapiens

gggtggaca tggggctgt tggct

ADN

20 *Homo sapiens*

tggcttggc agctgacatc caggt

ADN

Homo sapiens

tggcttggc agctgacatc caggtctag taagctgtct cctgc

25 ADN

Homo sapiens

tgccggttgt caaatccctt

ADN

Homo sapiens

30 **tccgatgcag ctcaagtaccg**

ADN

Homo sapiens

tccgatgcac agtaccgatg tggattggaa cgctgat

ADN

35 *Homo sapiens*

agcagtcctt tggagtgacc agcaacttg

ADN

Homo sapiens

catgcaatga agctgaacat gaccgtagtt

5 ADN

Homo sapiens

catgcaatga agctgaacat gaccgtagtt tctgtgatga gttttaaaaa

ADN

Homo sapiens

10 **gctgcaggcc ctgaagga**

ADN

Homo sapiens

ccccgacggt ctctcttc

ADN

15 *Homo sapiens*

ccccgacggt ctctcttcag ttctgagctt tcaagg

ADN

Homo sapiens

tgtggtacaa ccacgaacag

20 ADN

Homo sapiens

agatatttcc gcagcaacag

ADN

Homo sapiens

25 **agatatttcc gcagcaacag atgccacagc caggactaat**

ADN

Homo sapiens

caggagctaa aggccaaga

ADN

30 *Homo sapiens*

ccaggctgac ctcggggac

ADN

Homo sapiens

ccaggctgac ctcggggacg acctccaggg acgccatc

35 ADN

Homo sapiens

tgaagggtg gggaagcatt a

ADN

Homo sapiens

ttacaccaca agccaaacga c

5 ADN

Homo sapiens

ttacaccaca agccaaacga ctgatgcaat ggtctcctga ga

ADN

Homo sapiens

10 **ccaagaggac caggagaata tcaa**

ADN

Homo sapiens

ggataatcaa gagggaccaa tggg

ADN

15 *Homo sapiens*

ggataatcaa gagggaccaa tgggtgtaac gcaggctgtt tact

ADN

Homo sapiens

ggcgcgtctc cggcacgatg

20 ADN

Homo sapiens

gcagccagca aaaaagaaca gactc

ADN

Homo sapiens

25 **gcagccagca aaaaagaaca gactcaaagt ttcagtcaa gatcc**

ADN

Homo sapiens

tggaaftcig ttcggtgitt aagccagca

ADN

30 *Homo sapiens*

caatatggga tagcgggtct ttaagtcga

ADN

Homo sapiens

caatatggga tagcgggtct ttaagtcgat ccaagaggac tctcccggag gttccaa

35 ADN

Homo sapiens

catccccac agcacaaca g

ADN

Homo sapiens

5 **acagcaggca tgcttcatgg t**

ADN

Homo sapiens

acagcaggca tgcttcatgg gtctaccga tacactccg

ADN

10 *Homo sapiens*

accccagtc tcaatctaa c

ADN

Homo sapiens

cgttcgggct gaggctggg c

15 ADN

Homo sapiens

cgttcgggct gaggctggg ccgtaacag gaaccgcag gc

ADN

Homo sapiens

20 **ggaacttcgg aaatccaagg**

ADN

Homo sapiens

ccaigtggag caggtagggtg

ADN

25 *Homo sapiens*

ccaigtggag caggtagggtg gtgtcccca ggaaagtatt

ADN

Homo sapiens

ccttctctct gctgggtgcc

30 ADN

Homo sapiens

gccggatgtc cttccaggta

ADN

Homo sapiens

gccggatgic cttccaggta atcaccacca tgcgctgctg cga

ADN

Homo sapiens

ggggaagaga agcattgagg

5 ADN

Homo sapiens

gcctggtggt cgtggacgct

ADN

Homo sapiens

10 **gcctggtggt cgtggacgct cgggaatctg gggctagga**

ADN

Homo sapiens

agaaagagaa cagcttcgca

ADN

15 *Homo sapiens*

cacattgatt cattggctga

ADN

Homo sapiens

cacattgatt cattggctga tttctcca ggattctcc

20 ADN

Homo sapiens

gctggatgcc catgagagag g

ADN

Homo sapiens

25 **catgggaaca gctctgagg a**

ADN

Homo sapiens

catgggaaca gctctgagg aatctgttt tcttcatgc tc

ADN

30 *Homo sapiens*

ggggacgcag tagccgagat

ADN

Homo sapiens

tcacttcagc atcacctcca

35 ADN

Homo sapiens

tcacttcagc atcacctcca taaaggaag agccgagtcg

ADN

Homo sapiens

5 **cggatgggaa tgcgtttgg**

ADN

Homo sapiens

gggggtctcg cctcgggact

ADN

10 *Homo sapiens*

gggggtctcg cctcgggact acttgactgg ggtaaggtgg

ADN

Homo sapiens

acaaaagaag atgccacagc

15 ADN

Homo sapiens

tgagcaaca aaaacacagt

ADN

Homo sapiens

20 **tgagcaaca aaaacacagt tcctagggag tgaataagg c**

ADN

Homo sapiens

tgatacccca actccctcta

ADN

25 *Homo sapiens*

aaagcaggag ggaacagagc

ADN

Homo sapiens

aaagcaggag ggaacagagc actgcagggg cccacagg

30 ADN

Homo sapiens

tgcccagcta ctgctaccta

ADN

Homo sapiens

cccagttcag gtccagga

ADN

Homo sapiens

cccagttcag gtccaggatg tcataccgag tcttct

5 ADN

Homo sapiens

gccctgggac tgatagcaag

ADN

Homo sapiens

10 **agacgaagca gaggggcaaa**

ADN

Homo sapiens

agacgaagca gaggggcaaa ggggagtcc aaaacacctg

ADN

15 *Homo sapiens*

ccagaaatgg gtcagaatgg acaa

ADN

Homo sapiens

catctgccgg ggtaggagaa agcc

20 ADN

Homo sapiens

catctgccgg ggtaggagaa agcctgtctg ctgcagagcc tggc

ADN

Homo sapiens

25 **ctggagcccc gaggaagc**

ADN

Homo sapiens

cactgggggt ttccttig

ADN

30 *Homo sapiens*

cactgggggt ttccttgaa ggccagatct tctctt

ADN

Homo sapiens

agtgcattct catgtccgc tacta

35 ADN

Homo sapiens

cgatgttctt aacgtggtgc atcaa

ADN

Homo sapiens

5 **cgatgttctt aacgtggtgc atcaacaggc tgtggctgc ttgt**

ADN

Homo sapiens

cctcctgcag tcccagctct c

ADN

10 *Homo sapiens*

ggtttctccc cgccgttctc a

ADN

Homo sapiens

ggtttctccc cgccgttctc atgagcaaat aatccattct ga

15 ADN

Homo sapiens

ggagctcccc tgtggtcatc

ADN

Homo sapiens

20 **tttgaactg tggaaggaac**

ADN

Homo sapiens

tttgaactg tggaaggaac aggggctcgc tcttcigatt

ADN

25 *Homo sapiens*

aaagtacttg gagtctgcag gtgcg

ADN

Homo sapiens

tgcaattgac ctccagtgaa gtcca

30 ADN

Homo sapiens

tgcaattgac ctccagtgaa gttcagtgcc ccacacagga aaatagtctc

ADN

Homo sapiens

35 **cactgggacg aaggggaa**

ADN

Homo sapiens

gtcataagcc cggccaat

ADN

5 *Homo sapiens*

gtcataagcc cggccaatag acacagctgc catcct

ADN

Homo sapiens

aagcccaccg acccatct

10 ADN

Homo sapiens

tcaggcgctt cacaaagc

ADN

Homo sapiens

15 **tcaggcgctt cacaaagcag tgctggggta ggtgaa**

ADN

Homo sapiens

ggtcggagtc aacggatttg g

ADN

20 *Homo sapiens*

cctccgacgc ctgcttcacc a

ADN

Homo sapiens

cctccgacgc ctgcttcacc agaggggcca tccacagtct tc

25 ADN

Homo sapiens

gaggagcgag gactggagcc a

ADN

Homo sapiens

30 **gcacctccat gggtcgaaat t**

ADN

Homo sapiens

gcacctccat gggtcgaaat ttgtcagtgg gtctctaata aa

ADN

Homo sapiens

tcgccgagga gagcaagtc

ADN

Homo sapiens

5 **ggaacagcgc ggcctgtaa**

ADN

Homo sapiens

ggaacagcgc ggcctgtaa gaataatcca aaagaccaga

ADN

10 *Homo sapiens*

cgtaccctgt gccatccaa

ADN

Homo sapiens

taaacagcca gacagatgca

15 ADN

Homo sapiens

taaacagcca gacagatgca atacggggaa taaaccacgt

ADN

Homo sapiens

20 **gtcggggct tctgctga**

ADN

Homo sapiens

aacccttgag tgtagccca

ADN

25 *Homo sapiens*

aacccttgag tgtagcccag atgtgcatat tcacctc

ADN

Homo sapiens

gctgctggcc gaaaacttgc

30 ADN

Homo sapiens

gtctgccctc gttgctccca

ADN

Homo sapiens

gtctgccttc gttgctcca ttcttcctt gtagcacag

ADN

Homo sapiens

gcagagccgg ggacaagaga a

5 ADN

Homo sapiens

ctgctcttc tctccattga c

ADN

Homo sapiens

10 **ctgctcttc tctccattga cgctcttct gtagtcatt ca**

ADN

Homo sapiens

gggacgctcc tgattatgac

ADN

15 *Homo sapiens*

gcaaaccatg gccgctccc

ADN

Homo sapiens

gcaaaccatg gccgctccc ttctcaaaa tctccacag

20 ADN

Homo sapiens

gtgcgagtcg tctatggtc

ADN

Homo sapiens

25 **agttgtgtgc ggaaatccat**

ADN

Homo sapiens

agttgtgtgc ggaaatccat tgcctgggt gatctgttc

ADN

30 *Homo sapiens*

tccgctgcaa atacatctc

ADN

Homo sapiens

tgttcccgt tgccattgat

ADN

Homo sapiens

tgttccccgt tgccattgat taggacctca tggatcagca

ADN

5 *Homo sapiens*

gctctacctg gacctgctgt

ADN

Homo sapiens

ggaacacagg gaacatcacc

10 ADN

Homo sapiens

ggaacacagg gaacatcacc tagagcagga tggccacact

ADN

Homo sapiens

15 **atgtgaacca gccagatgtt**

ADN

Homo sapiens

ctctgggttc tctgccgtag

ADN

20 *Homo sapiens*

ctctgggttc tctgccgtag aggaggagg tgggctgag

ADN

Homo sapiens

gctctacgtt gcccgccagc ctg

25 ADN

Homo sapiens

gtttggggcc gctttgtag taa

ADN

Homo sapiens

30 **gtttggggcc gctttgtag taatttata tgctgtgac ggtttg**

ADN

Homo sapiens

ttttgggagg gggttgtgcc

ADN

Homo sapiens

ggccacacca gcagcatcca

ADN

Homo sapiens

5 **ggccacacca gcagcatcca taaccaactt ctgaggaact**

ADN

Homo sapiens

agttgtggcc ttacagcag

ADN

10 *Homo sapiens*

tgtgtgcccc aagtaatttt

ADN

Homo sapiens

tgtgtgcccc aagtaatttt gcacggacaa ttttaaggg

15 ADN

Homo sapiens

cctaccagct ccagacctt g

ADN

Homo sapiens

20 **tggcttcgtc agaatcacgt t**

ADN

Homo sapiens

tggcttcgtc agaatcacgt tcccagtatt actgcacacg tc

ADN

25 *Homo sapiens*

ccttactgtg agtctggggt ga

ADN

Homo sapiens

tgggttttct gctttctgat at

30 ADN

Homo sapiens

tgggttttct gctttctgat atgtcccttt acagcagtca tgtg

ADN

Homo sapiens

tggcccagct caaacagaa

ADN

Homo sapiens

cctcttcccc tcctgtta

5 ADN

Homo sapiens

cctcttcccc tcctgttac ttgtaaactg cgaggtg

ADN

Homo sapiens

10 **ctcaaggatg ccaggaacaa**

ADN

Homo sapiens

acactgagcc caccacctag

ADN

15 *Homo sapiens*

acactgagcc caccacctag ccactgccat atccagagga

ADN

Homo sapiens

tgccctggac acgggggtct

20 ADN

Homo sapiens

ttgcccttct gaatagtccc

ADN

Homo sapiens

25 **ttgcccttct gaatagtccc catggatgcc gtctaattgc**

ADN

Homo sapiens

ttcagcgaga gcagcgacac

ADN

30 *Homo sapiens*

cagaaccaac agggagaacc

ADN

Homo sapiens

cagaaccaac agggagaacc tcttgatgct ggtgctggaa

ADN

Homo sapiens

tgacatcgag gtggagagcg

ADN

5 *Homo sapiens*

ccccatcga aggcagaaat

ADN

Homo sapiens

tgacatcgag gtggagagcg cacacacacc agcaagatat

10 ADN

Homo sapiens

cctgctgaag tggctgcca a

ADN

Homo sapiens

15 **aacttggtt gatgctgtc c**

ADN

Homo sapiens

aacttggtt gatgctgtc ctttctct ggagactcaa aa

ADN

20 *Homo sapiens*

tgtccgagt tgatggccag agaca

ADN

Homo sapiens

acttgattac cgcagacagt gatga

25 ADN

Homo sapiens

acttgattac cgcagacagt gatgaacaac cggttgaggt cctgataat

ADN

Homo sapiens

30 **tgctgaagaa cggagggatg t**

ADN

Homo sapiens

tttgcattt tctgctct c

ADN

Homo sapiens

ttggcattt tctgctct ccatccea aaaaagtctt cg

ADN

Homo sapiens

5 **caggagtgc cagtcaagat**

ADN

Homo sapiens

acaggatgtt ctccggcttg

ADN

10 *Homo sapiens*

acaggatgtt ctccggcttg atctggcttg cttccgactc

ADN

Homo sapiens

acaattgact ctggccttc

15 ADN

Homo sapiens

tagacaatgg ccagcgaac

ADN

Homo sapiens

20 **acaattgact ctggccttc acgatcag acgtcagcgt**

ADN

Homo sapiens

gagagcccgg acatcaagta

ADN

25 *Homo sapiens*

acttcccttt gccctgtag

ADN

Homo sapiens

acttcccttt gccctgtag actgagacat cttccctcca

30 ADN

Homo sapiens

cgctcatcgt gggctccta a

ADN

Homo sapiens

agaagtcctg ggcaatgtcg g

ADN

Homo sapiens

agaagtcctg ggcaatgtcg gcaggtcggc caggcatac tc

5 ADN

Homo sapiens

cctgctccgc tgctccttg

ADN

Homo sapiens

10 **catgcccaac actcccctcc**

ADN

Homo sapiens

catgcccaac actcccctcc cctctctaa cacctcagca

ADN

15 *Homo sapiens*

aggtacagct ccccaccagc

ADN

Homo sapiens

cttcagcca gggcctgagc

20 ADN

Homo sapiens

cttcagcca gggcctgagc aggaatggtt accgtttgcc

ADN

Homo sapiens

25 **ggagcccaac tgcgccgacc**

ADN

Homo sapiens

ccttcgggta ctgatgatct aa

ADN

30 *Homo sapiens*

ggagcccaac tgcgccgacc cccgtggacc tgctgag

ADN

Homo sapiens

gcgctgcagg ttatgaaact t

ADN

Homo sapiens

agccccggtt gcctgcatca g

ADN

5 *Homo sapiens*

agccccggtt gcctgcatca gacccggagg tggccttctt tg

ADN

Homo sapiens

gcatcccgac gccctcaacc

10 ADN

Homo sapiens

gatgtccacg aggtcctgag

ADN

Homo sapiens

15 **gcatcccgac gccctcaacc gatgtccacg aggtcctgag**

ADN

Homo sapiens

gctgtccctc ccccttgtct

ADN

20 *Homo sapiens*

tgctccgctg ctaatcaaag

ADN

Homo sapiens

tgctccgctg ctaatcaaag tactccccca tcatatcccc

25 ADN

Homo sapiens

cgggacttgg agaagcactg

ADN

Homo sapiens

30 **tagaagaatc gtcgggtgca**

ADN

Homo sapiens

tagaagaatc gtcgggtgca tgacatcctg gctctcctgc

ADN

Homo sapiens

agaccggcgc acagaggaag

ADN

Homo sapiens

5 **cttttggac ttcaggtggc**

ADN

Homo sapiens

cttttggac ttcaggtggc cctcattcag ctctcggaac

ADN

10 *Homo sapiens*

cgaaggctac gaaggctatt aca

ADN

Homo sapiens

tggggagaag aaggggacca cga

15 ADN

Homo sapiens

tggggagaag aaggggacca cgaaggaatc ctgggagata caagaa

ADN

Homo sapiens

20 **gctccagcgg tgtaaacctg ca**

ADN

Homo sapiens

cgtgcaaatt caccagaagg ca

ADN

25 *Homo sapiens*

cgtgcaaatt caccagaagg catcaacttc attcatagt ctga

ADN

Homo sapiens

ctttcggttt tcagggggaa

30 ADN

Homo sapiens

tggctcacag ttctgcaggc

ADN

Homo sapiens

tggctcacag ttctgcagca ggcaattcct tatggcgcac agg

ADN

Homo sapiens

gggggtcccgc tcatcaagta

5 ADN

Homo sapiens

aactgccaca tcctttgcgt

ADN

Homo sapiens

10 **aactgccaca tcctttgcgt ccgcatgaag atgggagctc**

ADN

Homo sapiens

gcttccaagc ccgacctgat g

ADN

15 *Homo sapiens*

acggtggaaa tggtagtagg a

ADN

Homo sapiens

acggtggaaa tggtagtagg actccagccc tgagggtcc

20 ADN

Homo sapiens

ggccagaatt tagcaagaca

ADN

Homo sapiens

25 **tgactatggg cctagagcag**

ADN

Homo sapiens

tgactatggg cctagagcag ggcttcttct ttccactgg t

ADN

30 *Homo sapiens*

ccgaccagat caccttcc

ADN

Homo sapiens

gcttccgcac gtagacct

35 ADN

Homo sapiens

gcttccgcac gtagacctag ccccgtctcc gcatca

ADN

Homo sapiens

5 **ttcagaagg tctgccaaca ccaa**

ADN

Homo sapiens

gtgtccacca aggtcctgag atcc

ADN

10 *Homo sapiens*

gtgtccacca aggtcctgag atcccatttc tgccagtttc tgctgaaa

ADN

Homo sapiens

aggtgggcaa agggaagtaa

15 ADN

Homo sapiens

tagagcccct gagaagagcc

ADN

Homo sapiens

20 **tagagcccct gagaagagcc cagagaactg acagtccgtg**

ADN

Homo sapiens

ccagccactg ttgcagcatg a

ADN

25 *Homo sapiens*

aggcaaatgg gactcataca c

ADN

Homo sapiens

aggcaaatgg gactcataca cgggctggtg ctggagtgac ta

30 ADN

Homo sapiens

gctctgagcg agattgagac

ADN

Homo sapiens

caggatcaca cagcagatga

ADN

Homo sapiens

caggatcaca cagcagatga tccacatagt ctaccgctg

5 ADN

Homo sapiens

tgtgcacaaa tccatcaacc

ADN

Homo sapiens

10 **gaaaggctcc agggtaggt**

ADN

Homo sapiens

gaaaggctcc agggtaggt cacggatccg catggccatc

ADN

15 *Homo sapiens*

ccacgctctt ctgcctgctg

ADN

Homo sapiens

ctggtaggag acggcgatgc

20 ADN

Homo sapiens

ctggtaggag acggcgatgc gggttgcta caacatgggc

ADN

Homo sapiens

25 **gccgcctact tggtgctaac**

ADN

Homo sapiens

cgctgctgct cctgccttat

ADN

30 *Homo sapiens*

cgctgctgccc cctgccttag aggagtcca agttctatt

ADN

Homo sapiens

gattcctatg tgggcgacga g

ADN

Homo sapiens

ccatctcttg ctcgaagtc

ADN

5 *Homo sapiens*

ccatctcttg ctcgaagtc gccagccagg tccagacgca

REIVINDICACIONES

1. Un procedimiento de valoración de la cantidad de un primer ácido nucleico en una primera muestra, que comprende:
 - 5 proporcionar una mezcla estandarizada que comprende un molde competitivo de dicho primer ácido nucleico y un molde competitivo de un segundo ácido nucleico en dicha primera muestra, en el que dichos moldes competitivos están a concentraciones conocidas uno respecto a otro;
 - combinar dicha primera muestra con dicha mezcla estandarizada;
 - coamplificar dicho primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico, produciendo un primer producto amplificado del mismo sin coamplificar dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico;
 - 10 diluir dicho primer producto amplificado;
 - coamplificar además dicho primer producto amplificado diluido de dicho primer ácido nucleico y de dicho molde competitivo de dicho primer ácido nucleico, produciendo un segundo producto amplificado del mismo;
 - coamplificar dicho segundo ácido nucleico, en la que dicho segundo ácido nucleico sirve como primer ácido nucleico de referencia, y dicho molde competitivo de dicho segundo ácido nucleico en dicho primer producto amplificado de dicho primer ácido nucleico, produciendo el primer producto amplificado del mismo;
 - 15 y
 - obtener una primera relación comparando el segundo producto amplificado del ácido nucleico diana con el segundo producto amplificado del molde competitivo del ácido nucleico diana; y
 - 20 obtener una segunda relación comparando el primer producto amplificado del ácido nucleico de referencia con el primer producto amplificado del molde competitivo del ácido nucleico de referencia; y
 - comparar la primera y segunda relaciones, determinando así la cantidad del primer ácido nucleico.
2. El procedimiento como se indica en la reivindicación 1, que comprende además:
 - 25 diluir dicho producto amplificado de dicho primer ácido nucleico y dicho molde competitivo de dicho primer ácido nucleico; y
 - coamplificar además dicho producto amplificado diluido, produciendo un producto amplificado adicional del mismo.
3. El procedimiento según una cualquiera de las reivindicaciones 1 o 2, que comprende además:
 - 30 diluir dicho producto amplificado de dicho segundo ácido nucleico y dicho molde competitivo de dicho segundo ácido nucleico, y
 - coamplificar además dicho producto amplificado diluido, produciendo un producto amplificado adicional del mismo.
4. El procedimiento como se indica en una cualquiera de las reivindicaciones 1 a 3, en el que dicho primer ácido nucleico de referencia es un control para la carga de ADNc en la reacción.
5. El procedimiento como se indica en una cualquiera de las reivindicaciones 1 a 4, en el que dicho primer ácido nucleico de referencia corresponde a al menos un gen seleccionado de GADH, ACTB y β -actina.
- 35 6. El procedimiento según una cualquiera de las reivindicaciones 1 a 5, en el que dicho primer ácido nucleico comprende una secuencia correspondiente a un gen referenciado en la Fig. 1 o 2.
7. El procedimiento según una cualquiera de las reivindicaciones 1 a 6, en el que dicho molde competitivo de dicho primer o dicho segundo ácido nucleico comprende una secuencia referenciada en las Fig. 4A-4B.
- 40 8. El procedimiento según una cualquiera de las reivindicaciones 1 a 7, en el que dicho molde competitivo de dicho primer ácido nucleico está a una serie de concentraciones respecto a dicho molde competitivo de dicho segundo ácido nucleico.
9. El procedimiento como se indica en la reivindicación 8, en el que dicha serie de concentraciones proporciona diluciones en serie de 10 veces de dicho molde competitivo de dicho primer ácido nucleico respecto a dicho molde competitivo de dicho segundo ácido nucleico.
- 45 10. El procedimiento según una cualquiera de las reivindicaciones 1 a 9, en el que obtener dicha primera o dicha segunda relaciones comprende el uso de electroforesis capilar de microfluidos, matriz de oligonucleótidos,

espectrometría de masas o cromatografía.

- 5 11. El procedimiento según una cualquiera de las reivindicaciones 1 a 10, en el que dichas mezclas estandarizadas controlan al menos dos fuentes de variación seleccionadas de carga de ADNc, eficacia de amplificación intraácido nucleico, eficacia de amplificación interácido nucleico, eficacia de amplificación interespecimen, eficacia de amplificación intermuestra y eficacia de amplificación intramuestra.
12. El procedimiento según una cualquiera de las reivindicaciones 1 a 11, en el que dicho procedimiento está informatizado y en el que dicha informatización comprende instruir a un manipulador robótico para seleccionar dicha primera de las mezclas estandarizadas diluidas en serie para combinación.
- 10 13. El procedimiento como se indica en la reivindicación 12, en el que dicha informatización comprende obtener dicha primera relación.
14. El procedimiento como se indica en la reivindicación 13, en el que obtener dicha primera relación implica determinar un área bajo la curva.
- 15 15. El procedimiento según una cualquiera de las reivindicaciones 12 a 14, en el que dicha informatización comprende instruir a un manipulador robótico para seleccionar dicha segunda de las mezclas estandarizadas diluidas en serie basándose en dicha primera relación.
16. El procedimiento según una cualquiera de las reivindicaciones 1 a 15, en el que dicho ácido nucleico comprende una molécula de ARN.
17. El procedimiento según una cualquiera de las reivindicaciones 1 a 15, en el que dicho ácido nucleico comprende una molécula de ADN.
- 20 18. El procedimiento de la reivindicación 1, en el que la etapa de proporcionar la mezcla estandarizada comprende: preparar una mezcla estandarizada de reactivos, comprendiendo dichos reactivos suficiente molde competitivo para valorar las cantidades de una serie de ácidos nucleicos en más de aproximadamente 10^6 muestras, en el que dicha mezcla estandarizada permite la comparación directa de dichas cantidades entre dos de dichas muestras.
- 25 19. El procedimiento según la reivindicación 18, en el que dichos reactivos comprenden además un cebador directo y/o un cebador inverso para cebar la amplificación de dicho molde competitivo de dicho número de ácido(s) nucleico(s).

NSCLS primarias				
	Media	SD ²	% de la media ⁴	
BCL-2	ND³			
Catalasa	1.83E+05	3.80E+04	20	5
GSTP1	8.96E+08	5.80E+08	65	3
SOD1	7.50E+05	2.21E+05	30	3
DNASE1	3.55E+04	1.04E+04	29	4
GSTM3	1.19E+04	5.80E+03	49	5
GSTM1-5	1.58E+04	2.50E+04	157	4
ICAM1	7.33E+04	3.89E+04	53	5
BCLX-<i>lg</i>	1.83E+04	2.70E+03	15	5
BCLX-<i>sm</i>	9.89E+04	2.18E+04	22	5
ERCC1	5.05E+05	1.31E+05	26	3
MLH1	2.15E+04	9.40E+03	44	5
MSH2	2.39E+04	3.70E+04	155	3
MSH6	9.70E+04	2.69E+04	28	2
ERCC4	NE⁵			
p53	5.64E+02	2.10E+02	37	3
E2F-1	6.18E+04	1.50E+04	24	4
p21	NE⁵			
GJA1	1.13E+05	1.41E+04	13	5
GSR	NE⁵			
MSK2	NE⁵			
TPGS	5.32E+04	1.48E+04	28	4

¹ Los valores de expresión génica (media) se reseñan como moléculas por 10677 moléculas de actina

²SD= desviación estándar

³ND= no detectable

⁴ Se determinaron los valores mediante la prueba t de Student entre células H460 y H1435, números en negrita= significativos, p< 0,05

⁵ NE= no evaluado

Figura 1

Medida de la expresión génica en células epiteliales de vías aéreas de donante de pulmón por Start-PCR múltiple

sección pulmonar	ARN1	ESLX	BC1X	CA10	CAIN	CTP1	ERBB2	MPSA	CA11	M	MUT1	NSE	PCKA	BARA	RB	SPX	TRIM2	TM7B1A
F0N3H	1.10E+04	1.07E+03	1.02E+03	NA	9.10E7	NA	8.00E	1.3E+03	1.4E+04	4.00E	4.00E	9.1E+04	1.6E+04	2.10E+03	4.90E	6.00E	4.60E	8.00E+04
U1.3	1.10E+04	3.2E+04	2.9E+03	1.9E+04	61.00E	NA	4.4E+03	4.7E+03	1.10E+04	4.60E	4.60E	1.7E+03	8.4E+04	1.9E+04	4.90E	4.90E	4.90E	5.3E+04
U1.4	6.10E+03	1.97E+04	2.9E+03	1.41E+04	4.30E+03	2.64E+01	8.8E+03	2.6E+03	6.0E+03	4.00E	1.6E+03	6.6E+04	7.2E+04	5.9E+03	4.1E+03	4.90E	4.90E	9.4E+03

1 Valores de expresión para cada gen diana presentado como moléculas/10⁶ moléculas de β actina
 2 NA= no valorado
 3 La valoración no se completó para los valores citados como < o > , pero pudieron aproximarse los valores mediante la medida de CT o NT

Figura 2

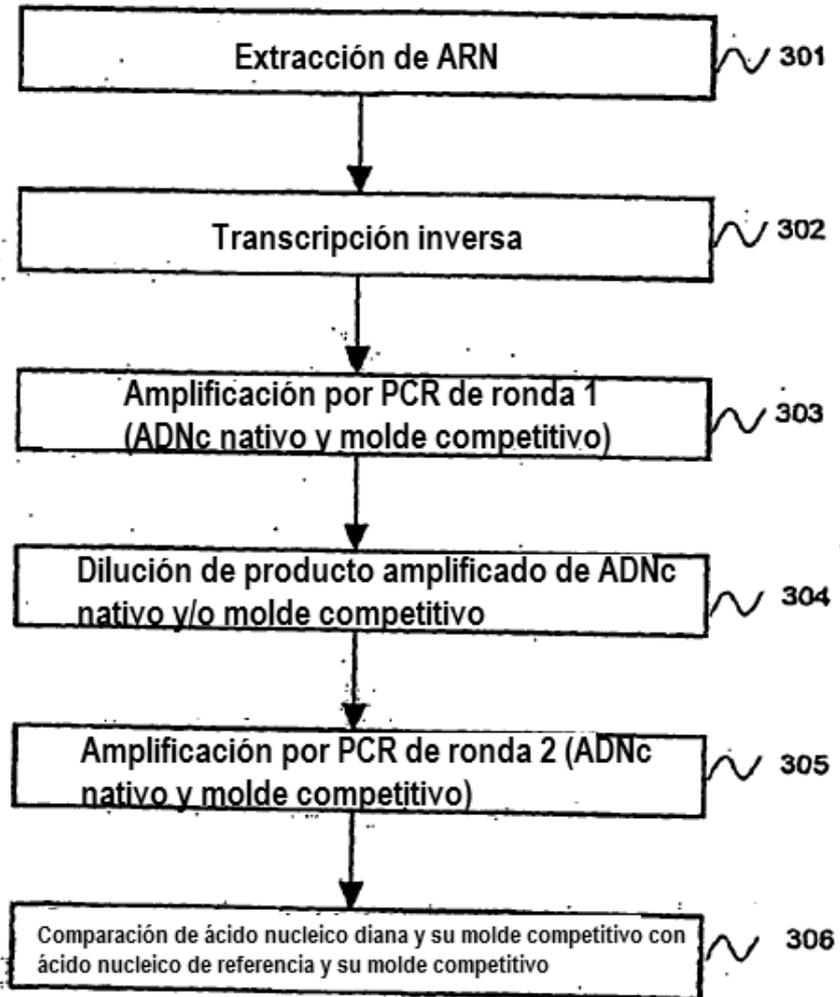


Figura 3

Expresión génica media en ADNc de referencia universal medida mediante los procedimientos dados a conocer en la presente memoria

Gen	Acceso a Genbank	Expresión media ¹		Seq. ID. No.		
		Dos etapas	No dos etapas	R.	CT.	
HSD11	M06664/M58487	ND ²	ND	1	2	3
ACHE	M35040	140	NA ³	4	5	6
ALAS1	X56351	7300	17000	7	8	9
ARNT	M59238	8300	NA	10	11	12
BCL2	M14745	410	530	13	14	15
BCL2L1	Z23115/Z23116	3200	12000	16	17	18
BCL2L1	Z23115/Z23116	6900	2300	19	20	21
MYC	V00568	81000	47000	22	23	24
CALB2	X56667	NA	NA	25	26	27
CBRI	J04056/X1818	6700	34000	28	29	30
CAT	X04076	5400	7900	31	32	33
CDC2	X05360	43000	40000	34	35	36
CDK7	X79193	6000	17000	37	38	39
CDK8	X85753	1200	NA	40	41	42
GJA1	X52947	20000	13000	43	44	45
JUN	J04111	60000	210000	46	47	48
CICN3	X78520	2600	4600	49	50	51
COE3A	X14420	25	25	52	53	54
SOD1	X02317/X00065	220000	130000	55	56	57
CCNA2	X51688	4000	6700	58	59	60
CCNG2	U47414	3600	3300	61	62	63
CCNB	U11791	2100	3400	64	65	66
CYP1A	K03191	540	240	67	68	69
CYP1B	U03688	1900	800	70	71	72
CYP2B	M29874/J02864	ND	ND	73	74	75
CYP2E	J02840	48	18	76	77	78
CYP2F1	J02906	ND	ND	79	80	81
DAO	X13227	96	59	82	83	84
SULT2	U08004	2600	4700	85	86	87
DNAH8	M55903	baja ⁴	540	88	89	90
TDFP2	U78422	13000	20000	91	92	93
DPY5	D78011	57	22	94	95	96
E2F1	M06377	2400	6300	97	98	99
E2F2	L22846	1400	720	100	101	102
E2F4	U15641	12000	15000	103	104	105
E2F5	U15642	3200	2000	106	107	108
ERCC1	M13194	61000	25000	109	110	111
ERCC4	U64315	650	170	112	113	114
EDN1	NM_001935	770	610	115	116	117
FNI	X02761/K00799	360000	140000	118	119	120
GCH	U35206	18000	14000	121	122	123
FPGS	M09045	4100	2600	124	125	126
FOSL1	X16707	40000	92000	127	128	129
GAPD	M03197	230000	350000	130	131	132
GILC	M00636	380	1700	133	134	135
GLI2	NM_005270	5000	1700	136	137	138
GLUCT	M57899	baja	ND	139	140	141
GLUCT	M57931	530	1000	142	143	144
GSR	X15722	24000	12000	145	146	147

Figura 4a

Expresión génica media en ADNc de referencia universal medida mediante los procedimientos dados a conocer en la presente memoria

GPX3	D06532	2500	NA	148	149	150
GSTM1	J03817/0463509	16000	19000	151	152	153
GSTM3	J05459	14000	17000	154	155	156
GSTP1	X06547	210000	230000	157	158	159
GSTT1	X79989	NA	180	160	161	162
ERBB2	X03363	84000	66000	163	164	165
ENF3A	U09840	360	730	166	167	168
HSPD1	X0112182	130000	NA	169	170	171
HSPA9	L15189	17000	13000	172	173	174
ICAM1	X06990	4300	6100	175	176	177
IVL	M03902	baja	baja	178	179	180
JUNB	X51345	1300	11000	181	182	183
KRT3	M09723	ND	baja	184	185	186
IL16	M90391	baja	690	187	188	189
MAD	L06825	1200	1500	190	191	192
MAX	X60287	ND	ND	193	194	195
MLH1	U07943	13000	6000	196	197	198
MSE2	L47581	9200	9600	199	200	201
MSE6	U28946	180000	57000	202	203	204
MSE2	AF074715	4300	2000	205	206	207
MUC1	X52229	ND	ND	208	209	210
NADH	M28703	170000	24000	211	212	213
PCR	S90469	29000	15000	214	215	216
NSE	M27610	76000	62000	217	218	219
RBL2	X74594	1700	5000	220	221	222
CDKN2	L27211	ND	baja	223	224	225
CDKN2	U17074	9300	17000	226	227	228
CDKN2	U40343	1800	3500	229	230	231
CDKN1	U03105	4000	9200	232	233	234
CDKN1	U10906	9000	8600	235	236	237
TP53	R03199	7200	3300	238	239	240
SAM68	M28108	190000	130000	241	242	243
PCNA	J04719	120000	54000	244	245	246
PMS2L	D38437	1400	3900	247	248	249
SULT1	L19999	6100	4000	250	251	252
SULT1	U08032	23000	21000	253	254	255
RAF1A	M22995	NA	38000	256	257	258
RARA	X06614	4200	3200	259	260	261
RBI	M15400	3900	3600	262	263	264
SPARC	R03040	320000	110000	265	266	267
SPR1B	M284757	18	62	268	269	270
STX1A	L37792	1400	2400	271	272	273
TGSM2	M55153	NA	3200	274	275	276
TNF	X02910/0301159	ND	ND	277	278	279
TNFRF	M582860/030480	2400	7600	280	281	282

² ND = Ninguna detección. No se detectó NT cuando se usó la mezcla F

³ NA = No valorado

⁴ baja= los niveles de expresión son menores de 600 ARNm/10⁶ARNm de beta-actina

Se reseña baja actividad cuando no se observa NT cuando se usó mezcla

E y no eran mensurables ni NT ni CT cuando se usó mezcla F

Figura 4B

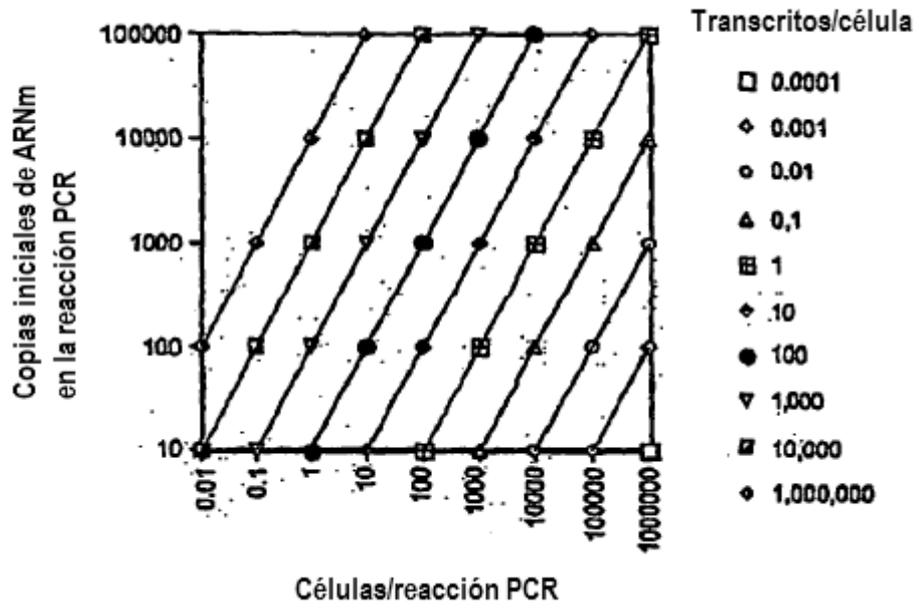


Figura 5

Mezclas estandarizadas de patrones internos

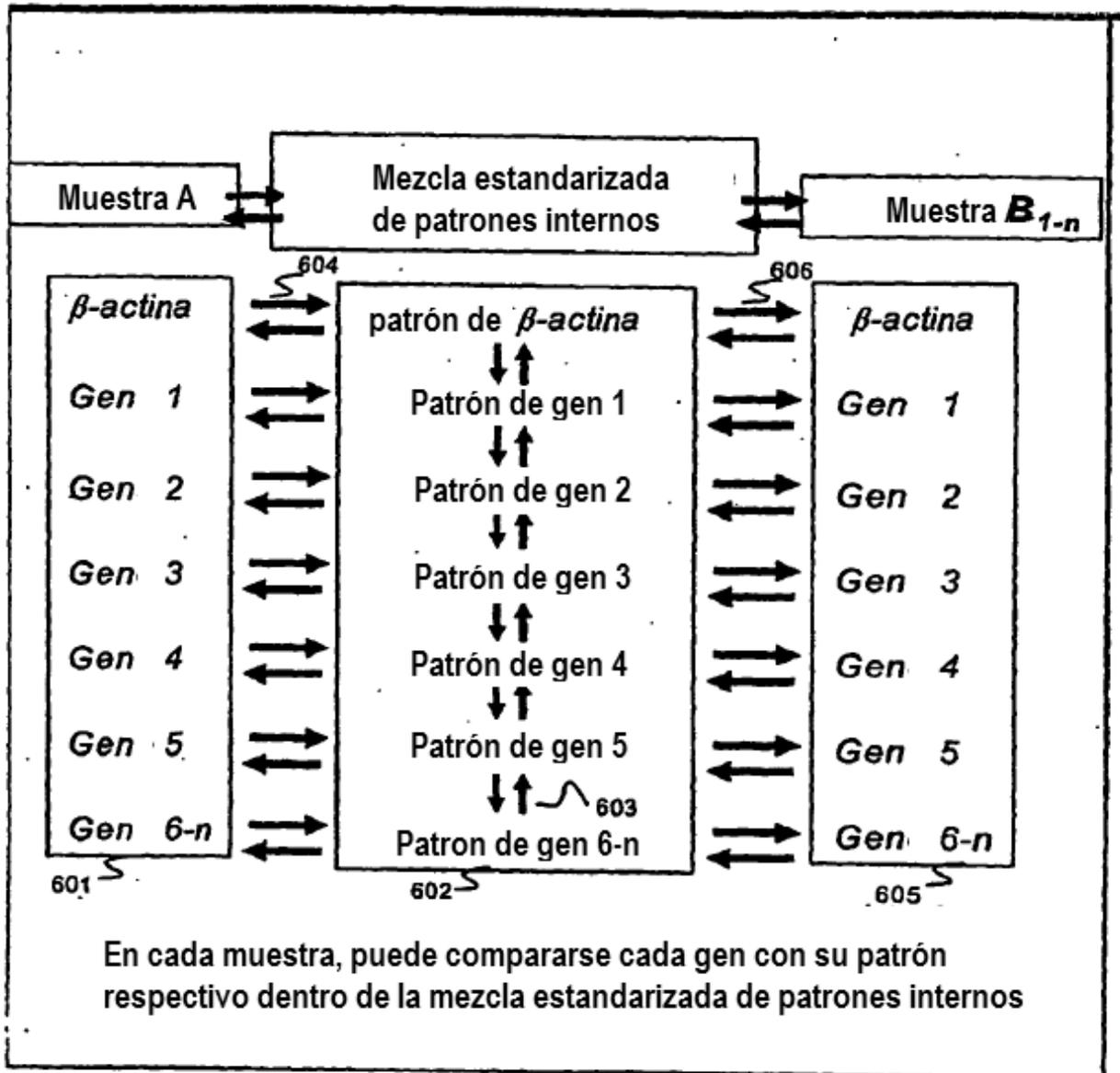
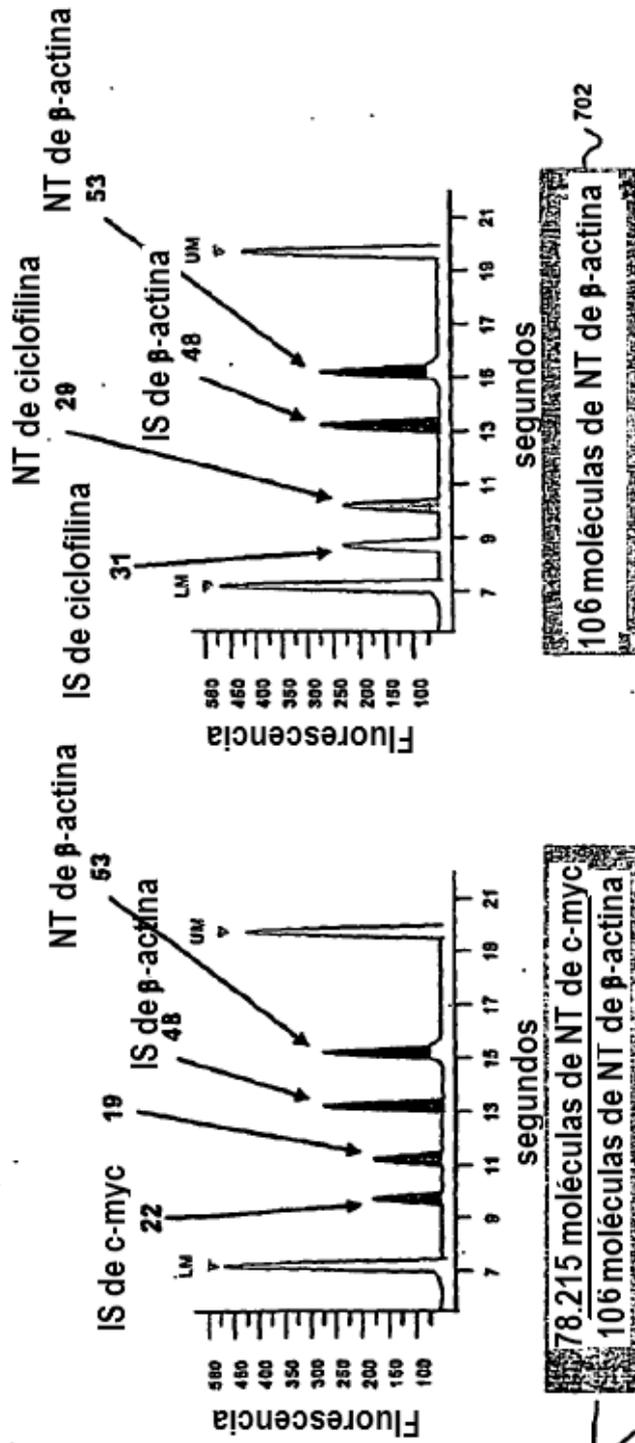


Figura 6

Normalización flexible



701

78.215 moléculas de NT de c-myc
106 moléculas de NT de beta-actina

703

beta-actina : 53/48 (NT/IS) x 600.000 moléculas = 662.500 moléculas de NT de beta-actina

704

78.215 moléculas de c-myc	X	106 moléculas de beta-actina	=	923.184 moléculas de c-myc
106 moléculas de beta-actina		84.723 moléculas de ciclofilina		106 moléculas de ciclofilina

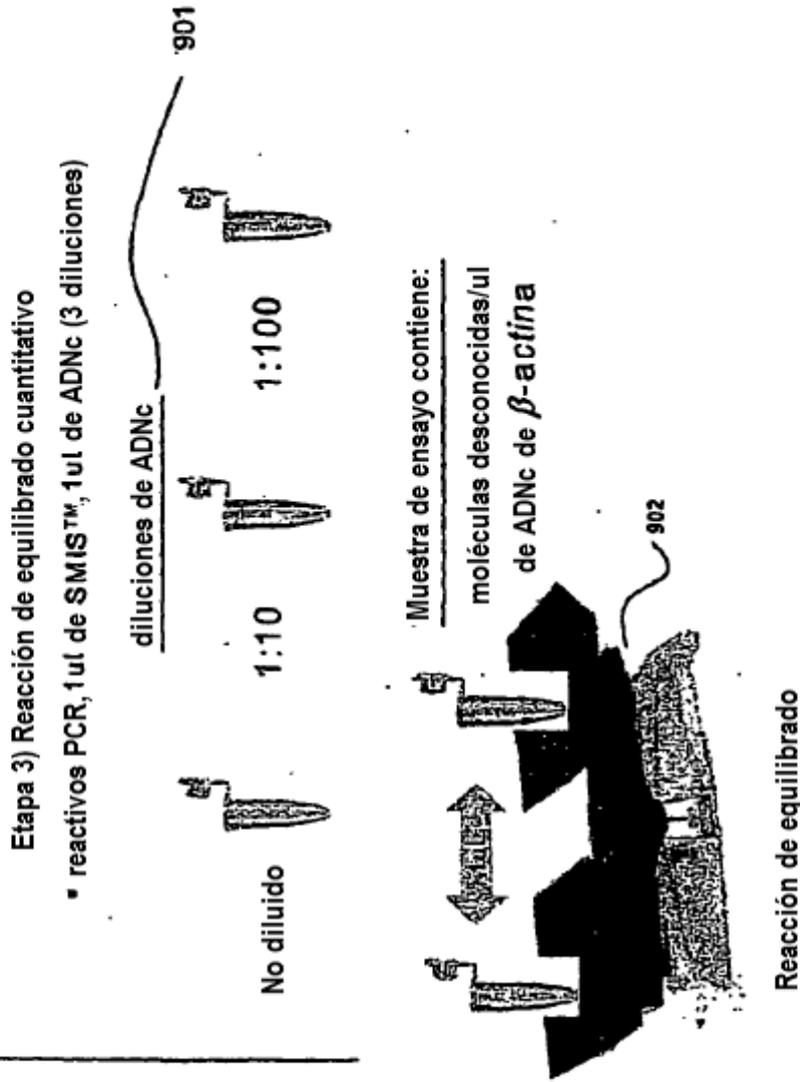
Figura 7

SMIS™	Muestra A	Muestras estandarizadas de patrones internos	Muestra B
Mezcla A	β -actina	Patrón de β -actina	β -actina
600.000 moléculas de b-actina/ul	Gen 1	Patrón de gen 1	Gen 1
6.000.000 de moléculas de otros genes/ul	Gen 2	Patrón de gen 2	Gen 2
Mezcla B	β -actina	Patrón de β -actina	β -actina
600.000 moléculas de b-actina/ul	Gen 3	Patrón de gen 3	Gen 3
600.000 moléculas de otros genes/ul	Gen 4	Patrón de gen 4	Gen 4
Mezcla C	β -actina	Patrón de β -actina	β -actina
600.000 moléculas de b-actina/ul	Gen 5	Patrón de gen 5	Gen 5
60.000 moléculas de otros genes/ul	Gen 6	Patrón de gen 6	Gen 6
Mezcla D	β -actina	Patrón de β -actina	β -actina
600.000 moléculas de b-actina/ul	Gen 7	Patrón de gen 7	Gen 7
6.000 moléculas de otros genes/ul	Gen 8	Patrón de gen 8	Gen 8
Mezcla E	β -actina	Patrón de β -actina	β -actina
600.000 moléculas de b-actina/ul	Gen 9	Patrón de gen 9	Gen 9
600 moléculas de otros genes/ul	Gen 10	Patrón de gen 10	Gen 10
Mezcla F	β -actina	Patrón de β -actina	β -actina
600.000 moléculas de b-actina/ul	Gen 11	Patrón de gen 11	Gen 11
60 moléculas de otros genes/ul	Gen 12	Patrón de gen 12	Gen 12

Figura 8

Equilibrado cuantitativo

- 1.) Extracción de ARN
- 2.) Transcripción inversa
- 3.) Equilibrado cuantitativo

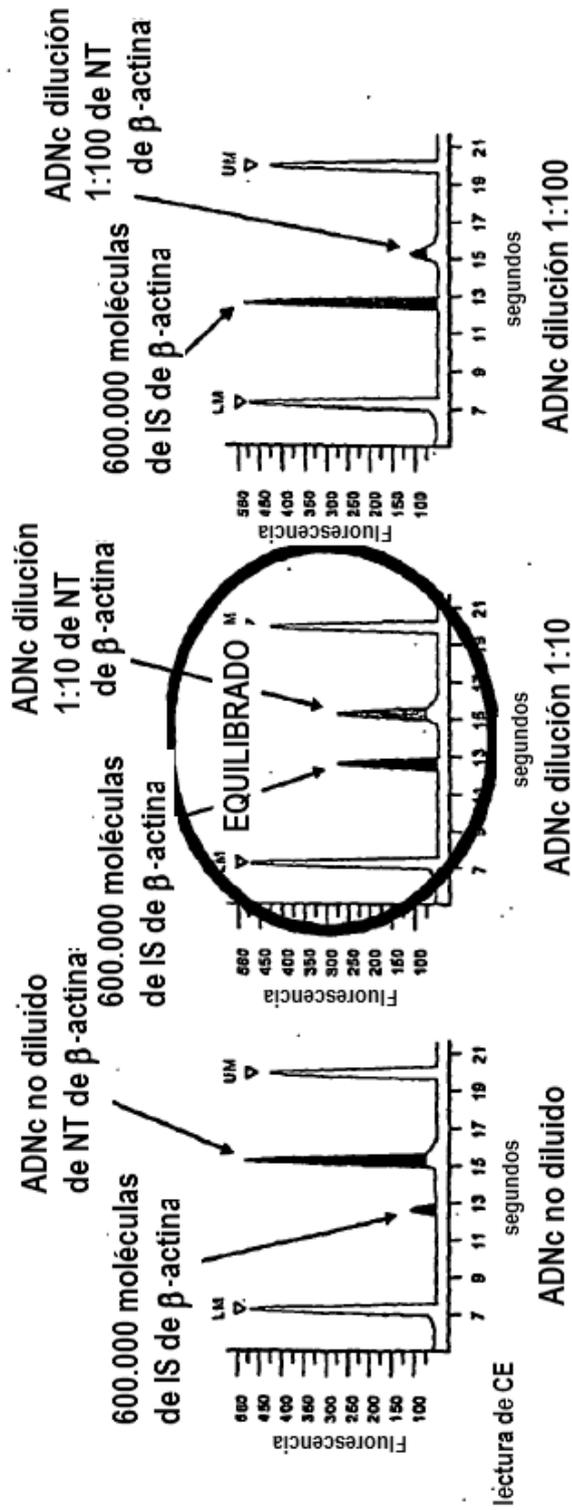


SMIS™ contiene:
600.000 moléculas/ul de patrón interno de β -actina

Figura 9

Equilibrado cuantitativo
Control de carga

Determinar la dilución a la que NT de β -actina = IS de β -actina = 600.000 moléculas de NT de β -actina



En este caso, el gen de referencia de β -actina en ADNc a dilución 1:10 está en equilibrio con 600.000 moléculas del patrón interno de β -actina en SMIS™

Figura 10

Medida de los genes diana

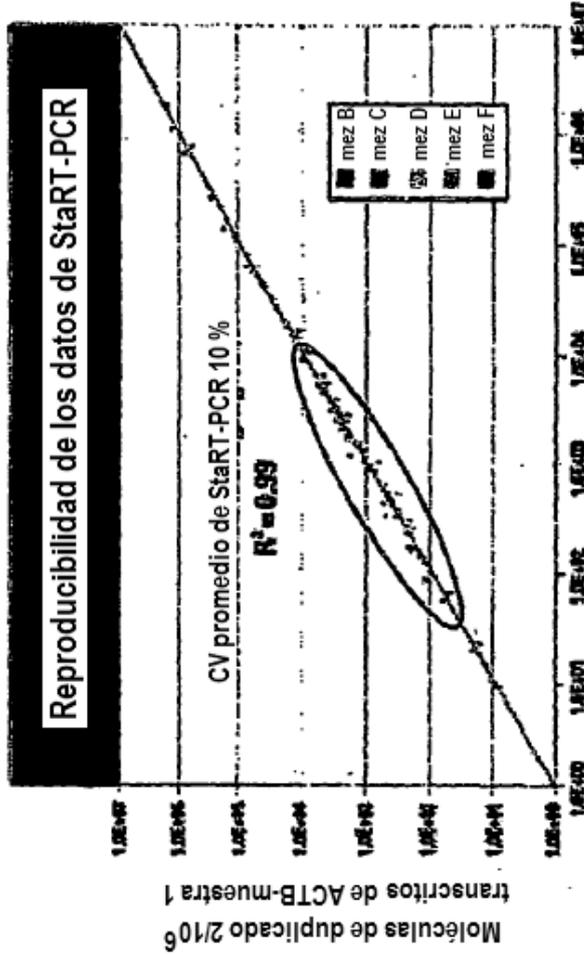
Selección de la mezcla correcta

Para niveles de expresión desconocidos, se usa inicialmente la mezcla "E"

Mezclas estandarizadas de patrones internos TM (SMIS TM)

- Mezcla A**
600.000 moléculas de ACTB/ul
6.000.000 moléculas diana/ul
- Mezcla B**
600.000 moléculas de ACTB/ul
600.000 moléculas diana/ul
- Mezcla C**
600.000 moléculas de ACTB/ul
60.000 moléculas diana/ul
- Mezcla D**
600.000 moléculas de ACTB/ul
6.000 moléculas diana/ul
- Mezcla E**
600.000 moléculas de ACTB/ul
600 moléculas diana/ul
- Mezcla F**
600.000 moléculas de ACTB/ul
60 moléculas diana/ul

Para niveles de expresión desconocidos, se usa inicialmente la mezcla "E"



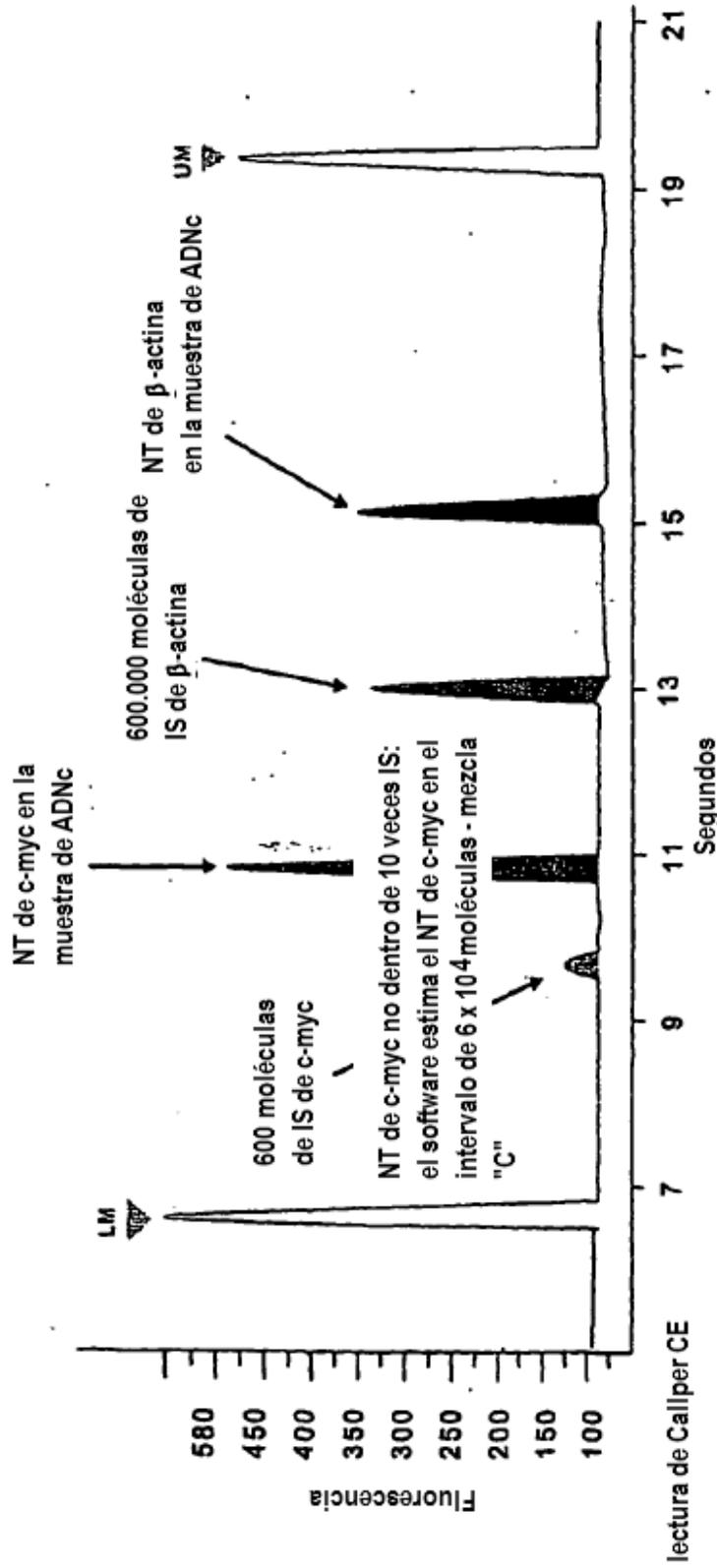
Moléculas de duplicado 1/10⁶
transcritos de ACTB- muestra 1

Figura 12

Figura 11

Medida de genes diana

usando la mezcla "E"



Basándose en el área bajo la curva, el software determina que el NT del gen diana (c-myc) en la muestra de ADNc no está dentro del intervalo de 10 veces el patrón interno de c-myc, sino que amplía al intervalo a 10^4 y por lo tanto recomienda la mezcla "C" de SMIS™

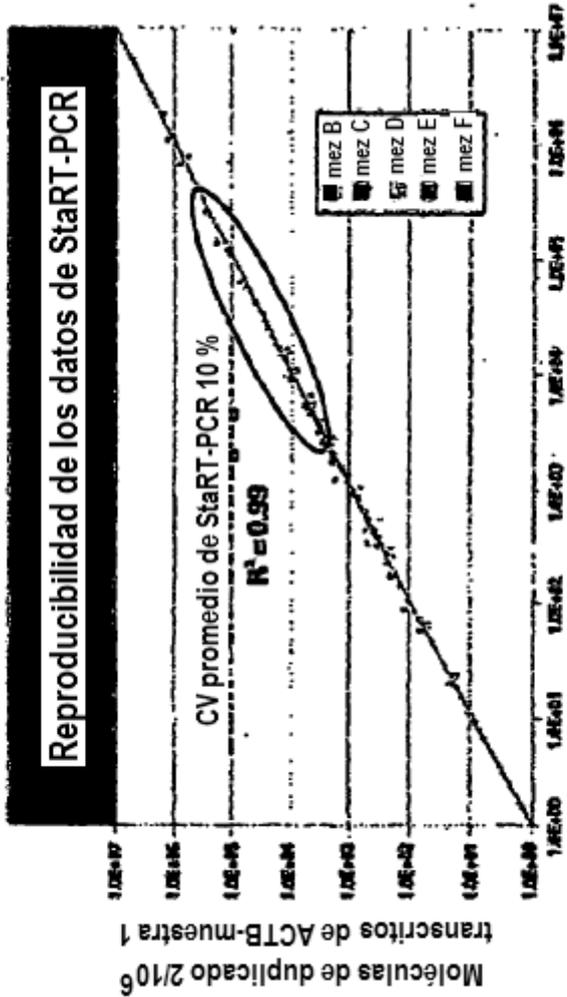
Figura 13

Medida de los genes diana
Selección de la mezcla correcta

Si se requiere otra mezcla de SMIS, el software comunica automáticamente la mezcla correcta al robot

Mezclas estandarizadas de patrones internos TM (SMIS TM)

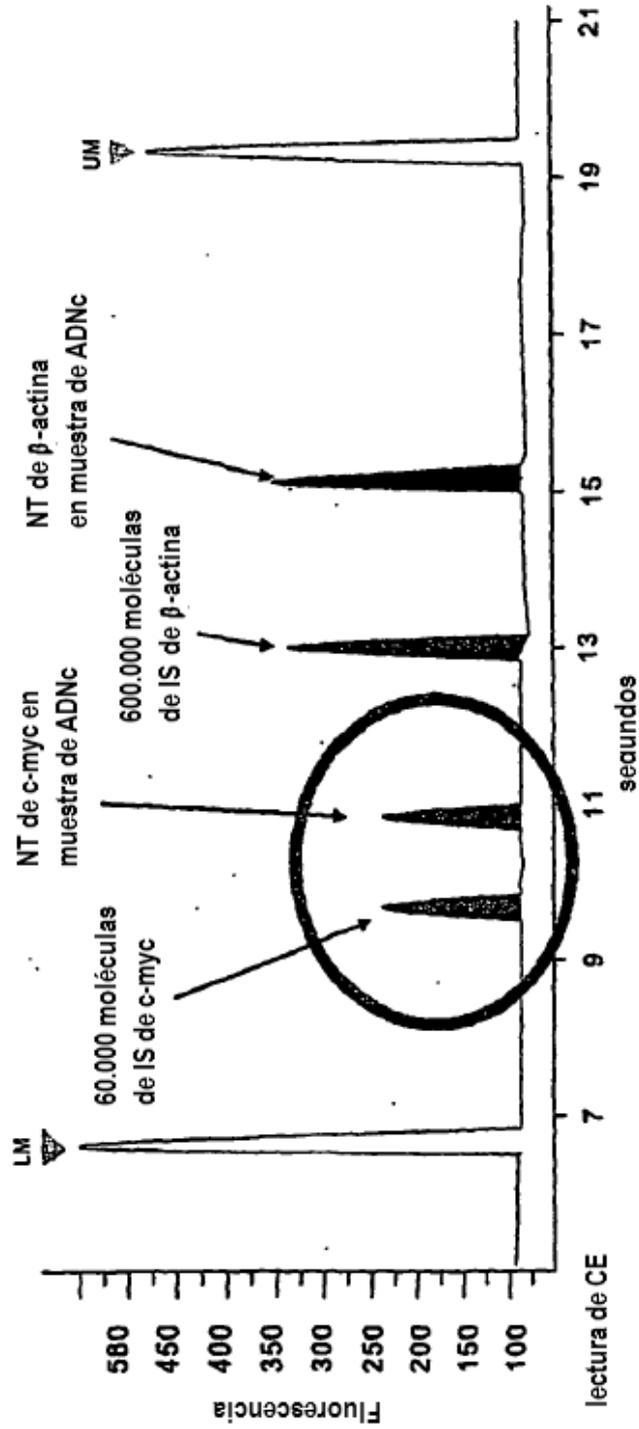
- Mezcla A**
600.000 moléculas de ACTB/ul
6.000.000 moléculas diana/ul
- Mezcla B**
600.000 moléculas de ACTB/ul
600.000 moléculas diana/ul
- Mezcla C**
600.000 moléculas de ACTB/ul
60.000 moléculas diana/ul
- Mezcla D**
600.000 moléculas de ACTB/ul
6.000 moléculas diana/ul
- Mezcla E**
600.000 moléculas de ACTB/ul
600 moléculas diana/ul
- Mezcla F**
600.000 moléculas de ACTB/ul
60 moléculas diana/ul



Moléculas de duplicado $1/10^6$
transcritos de ACTB- muestra 1

Figura 14

Medida de los genes diana
usando la mezcla "C"



Basándose en el área bajo la curva, el software determina la abundancia de transcritos para el gen diana c-myc en el formato de: n° de moléculas de c-myc/10⁶ moléculas de β-actina

Figura 15

Medida multitranscrito

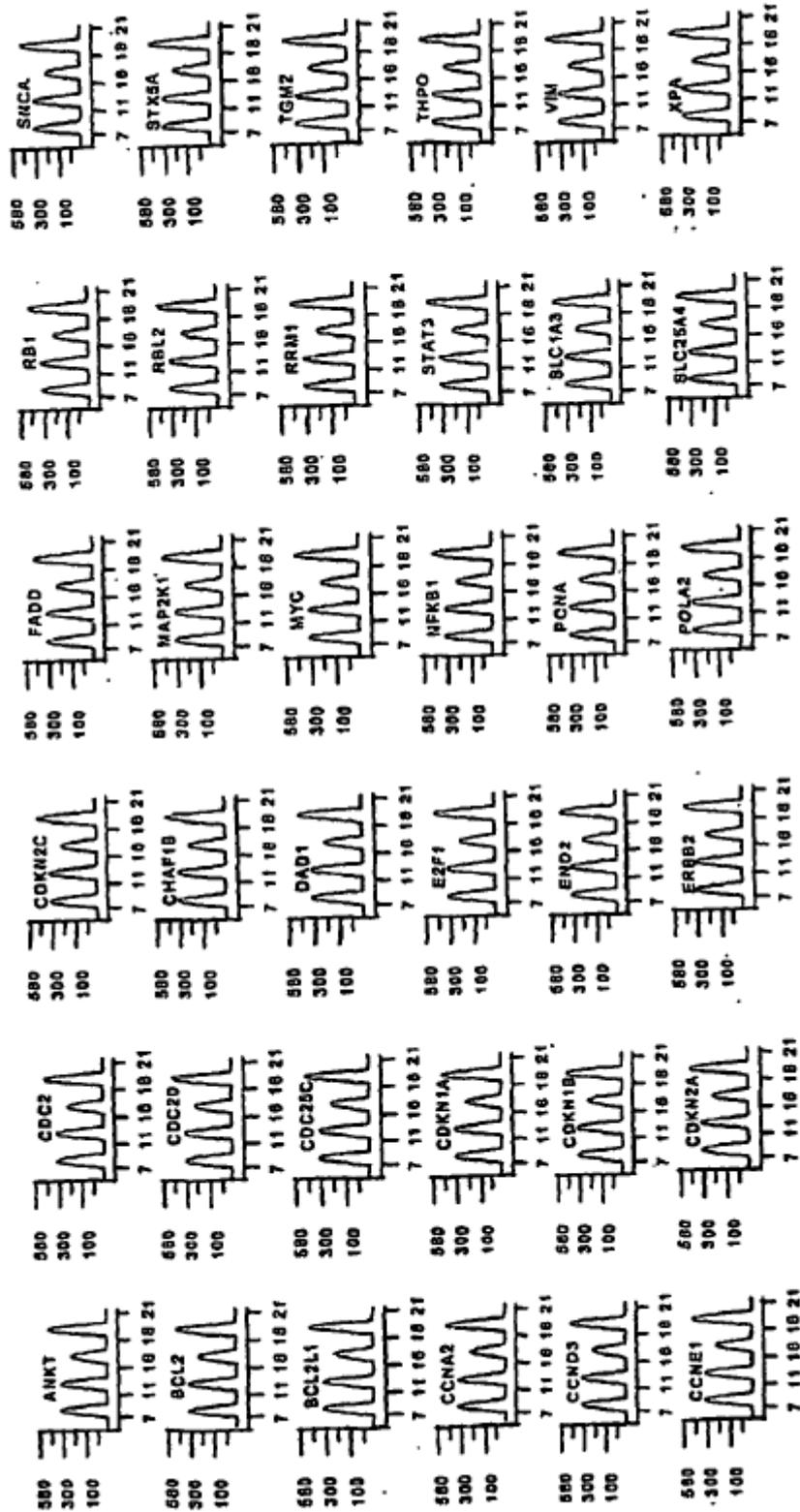


Figura 17

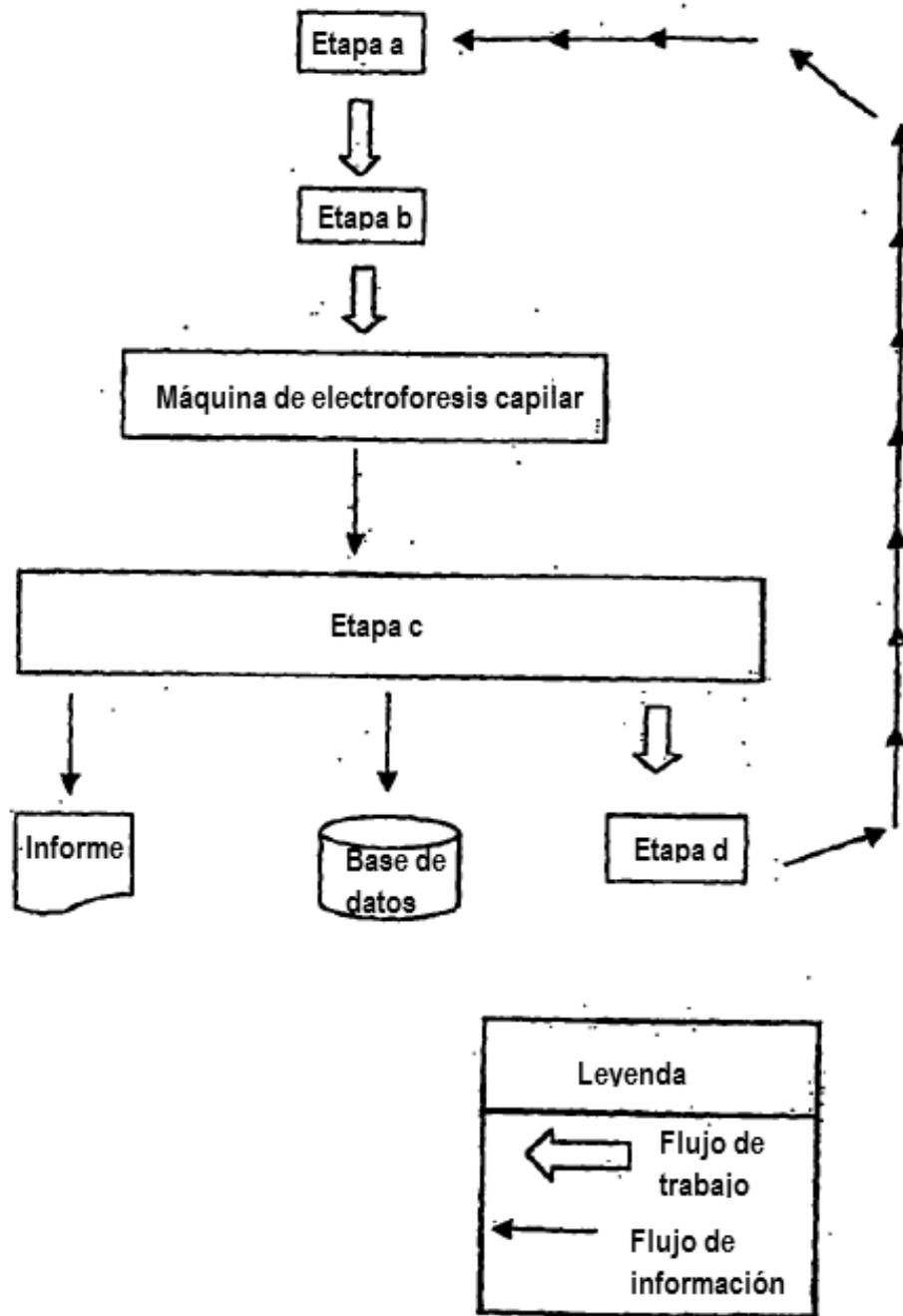


Figura 18

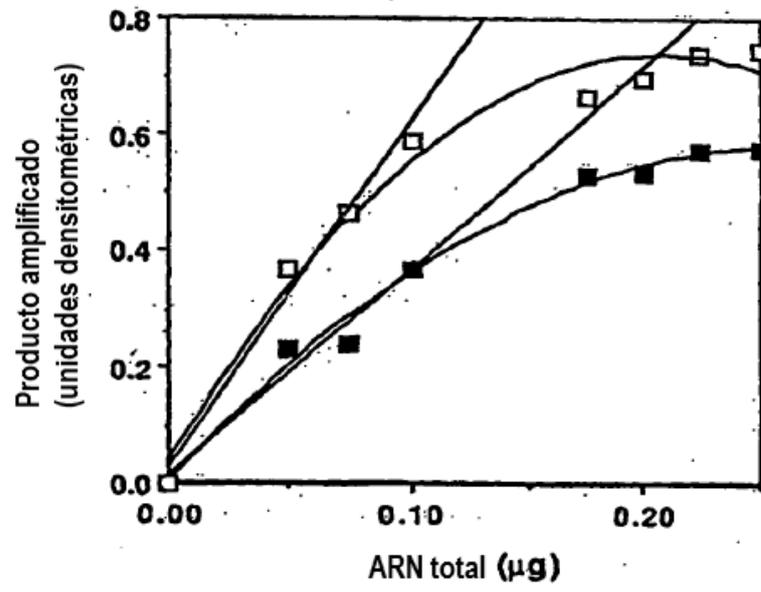


Figura 19

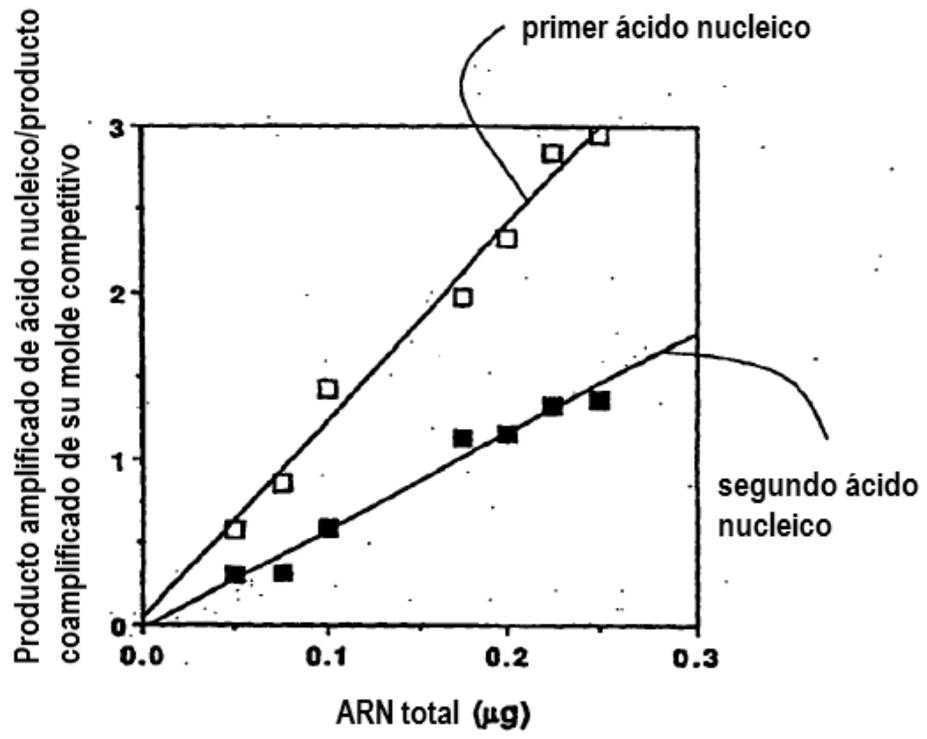


Figura 20

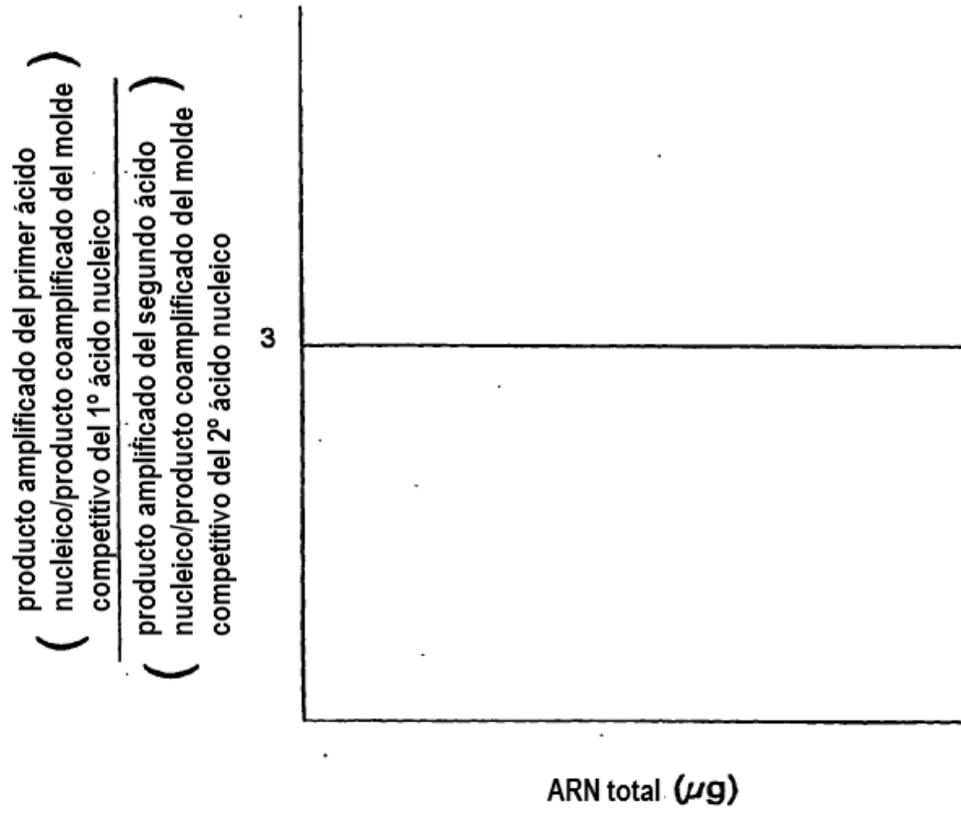


Figura 21

Fuentes de variación en la medida de expresión génica por RT-PCR cuantitativa y procedimientos de control

Fuente de variación	Realizaciones de procedimientos descritos en la presente memoria	Tiempo real
Carga de ADNc: Debido a la variación en el pipeteado, cuantificación o transcripción inversa.	Amplificación múltiple con gen de referencia (por ejemplo, β -actina)	Amplificación múltiple con gen de referencia (por ejemplo, β -actina)
Consecuencia: comparación poco fiable de la expresión del mismo gen en dos muestras diferentes		
Variación de ciclo a ciclo de la eficacia de amplificación intraácido nucleico: fases de meseta temprana lenta, logarítmica lineal y tardía lenta	CT de patrón interno para cada gen en una mezcla estandarizada de patrones internos (SMIS)	Medida a tiempo real
Consecuencia: comparación poco fiable de la expresión del mismo gen en diferentes muestras		
Eficacia de la amplificación interácido nucleico: en la eficacia de los cebadores	CT de patrón interno para cada gen en una SMIS	Curva patrón externa para cada gen medido
Consecuencia: comparación poco fiable de la expresión de diferentes genes en la misma o diferentes muestras		
Eficacia de amplificación interespecimen: presencia variable de un inhibidor de PCR	CT de patrón interno para cada gen en una SMIS	Curva patrón de la muestra de referencia en comparación con la muestra de ensayo ²
Consecuencia: comparación poco fiable de la expresión del mismo o diferentes genes en la misma o diferentes muestras		
Eficacia de amplificación intermuestra: en la calidad y/o concentración de reactivos de PCR (por ejemplo, cebadores) en presencia de un inhibidor de PCR	CT de patrón interno para cada gen en una SMIS	Ninguno ²
Consecuencia: comparación poco fiable de la expresión del mismo o diferentes genes en la misma o diferentes muestras		
Eficacia de amplificación intramuestra: en la eficacia del termociclador	CT de patrón interno para cada gen	Ninguno ²
Consecuencia: por ejemplo, comparación poco fiable de la expresión del mismo o diferentes genes en la misma o diferentes muestras		

Figura 22

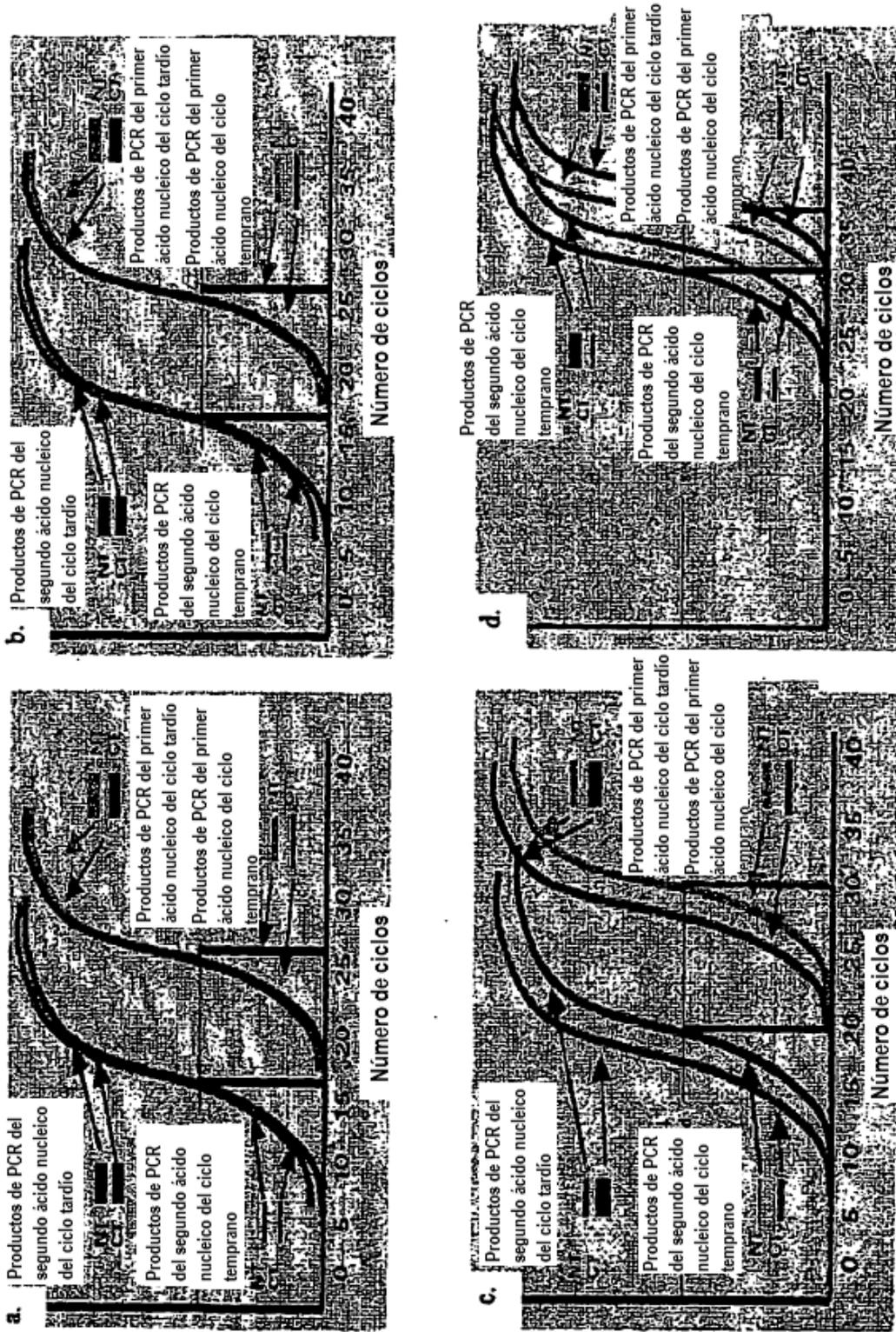


Figura 23

Datos estandarizados

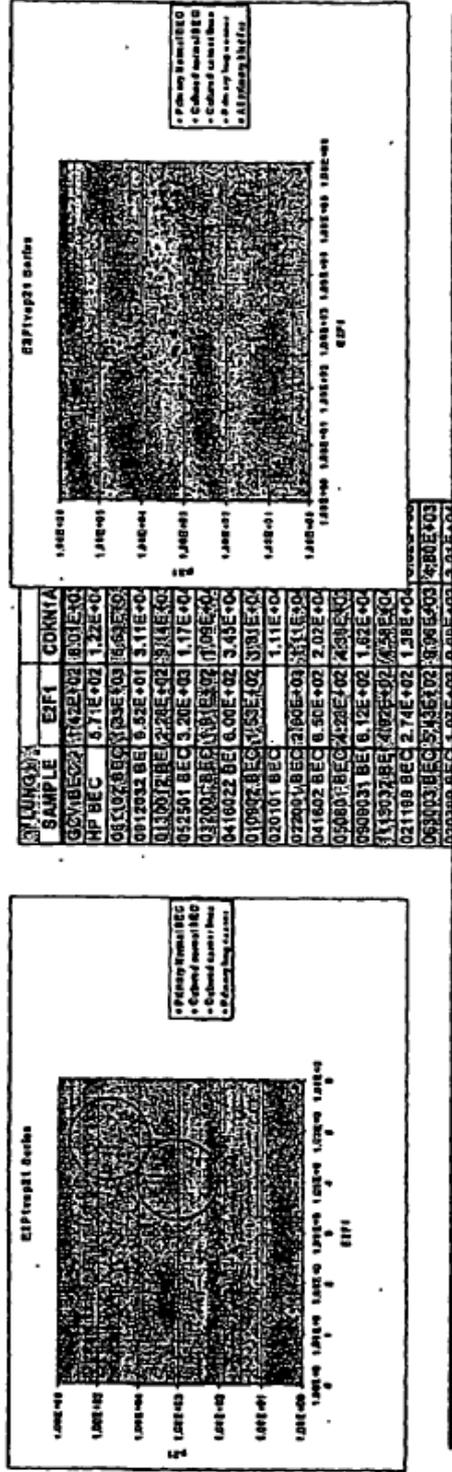
que posibilitan un análisis múltiple virtualmente infinito

Ejemplo:

- un ensayo de expresión génica para aumentar el diagnóstico de cáncer de pulmón en muestras citomorfológicas
- *c-myc* x *E2F1/p21*
- sensibilidad de diagnóstico de 100 % y especificidad de 96 %

Análisis bivariado de genes del ciclo celular en células epiteliales bronquiales humanas normales y malignas

La relación *E2F1/p21* observada en tejido pulmonar no se observa en tejido de vejiga



Pueden compararse todos los conjuntos de datos de diferentes experimentos, diferentes laboratorios y diferentes días en la base de datos de expresión estandarizada

Figura 25

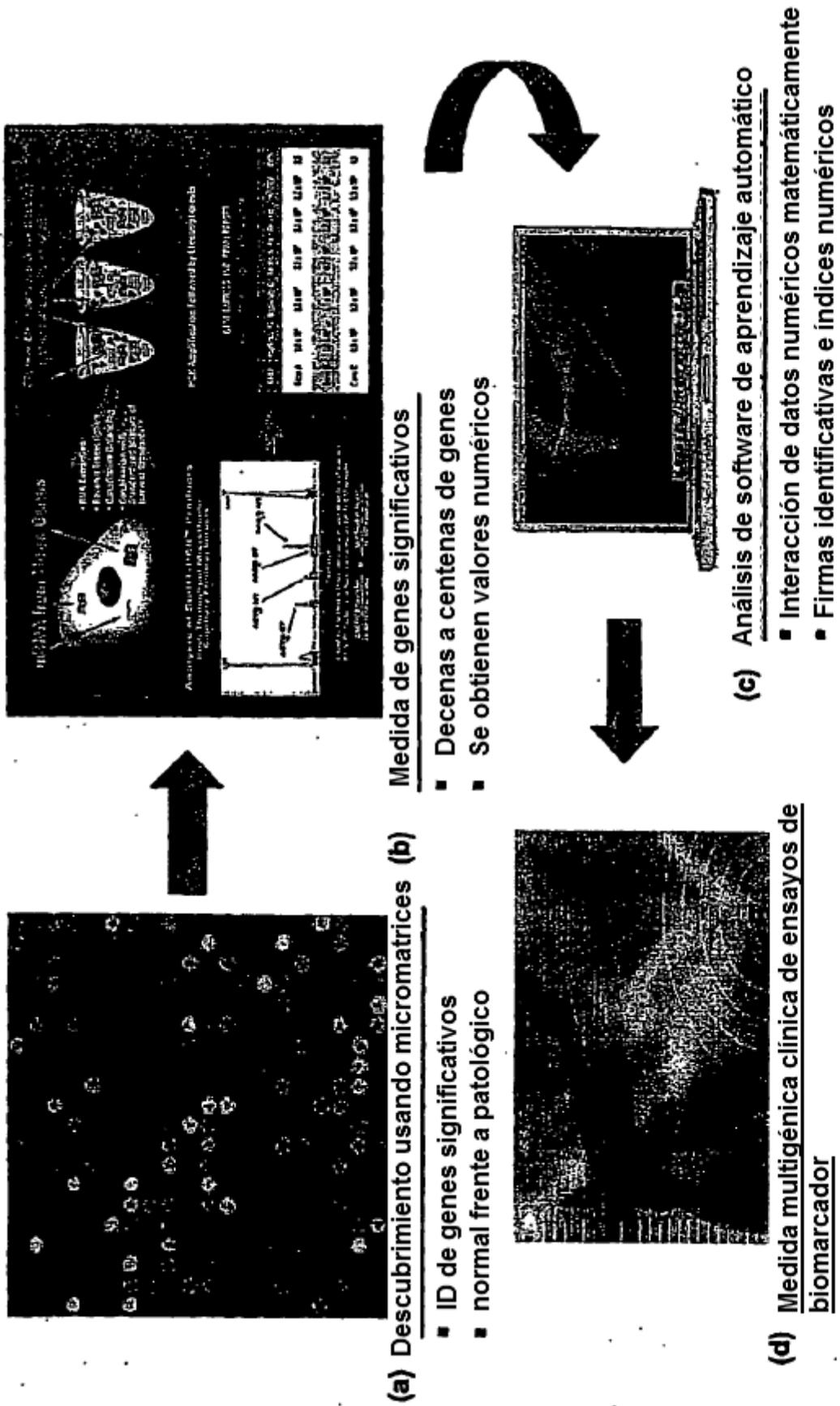


Figura 26

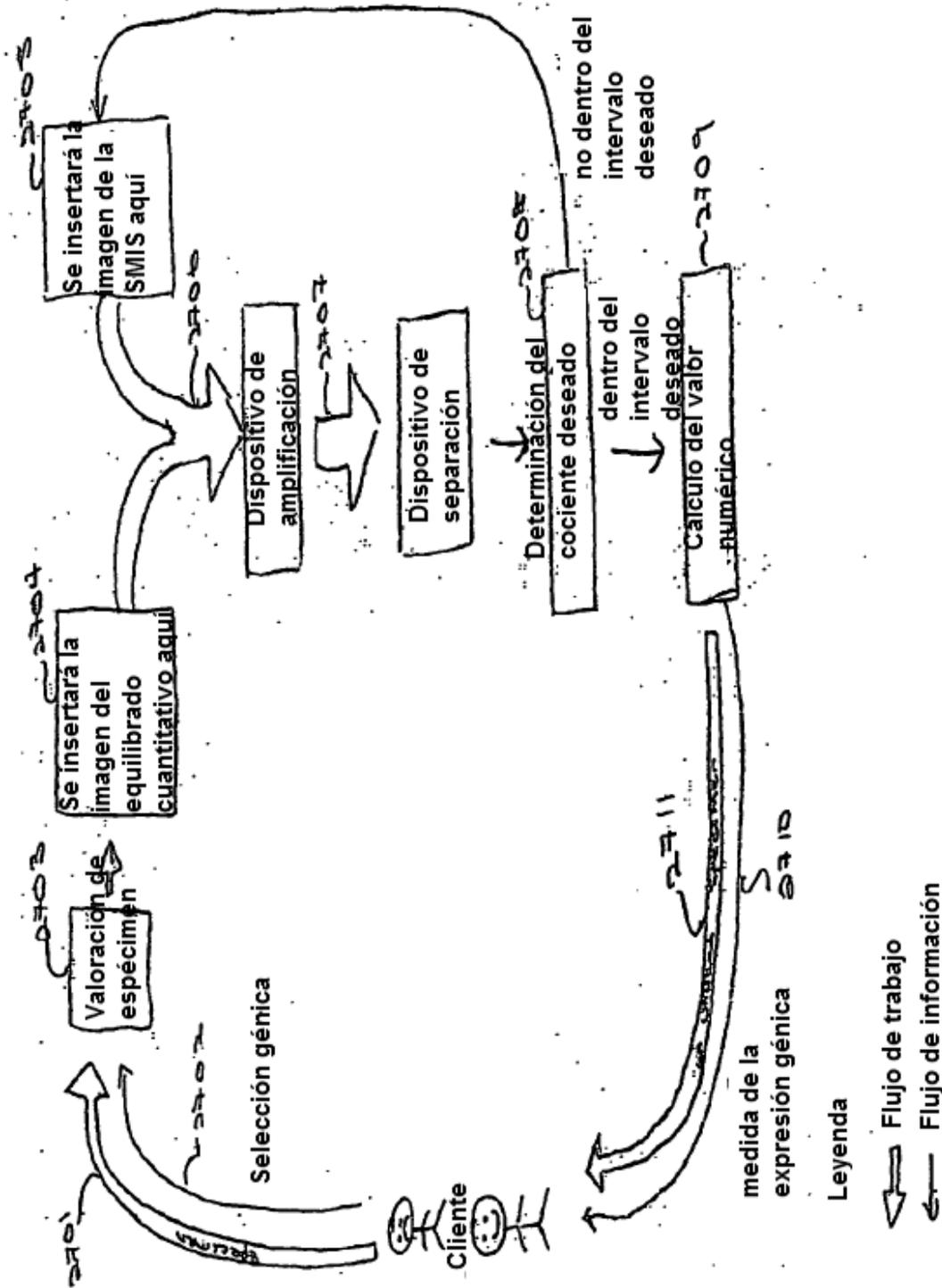
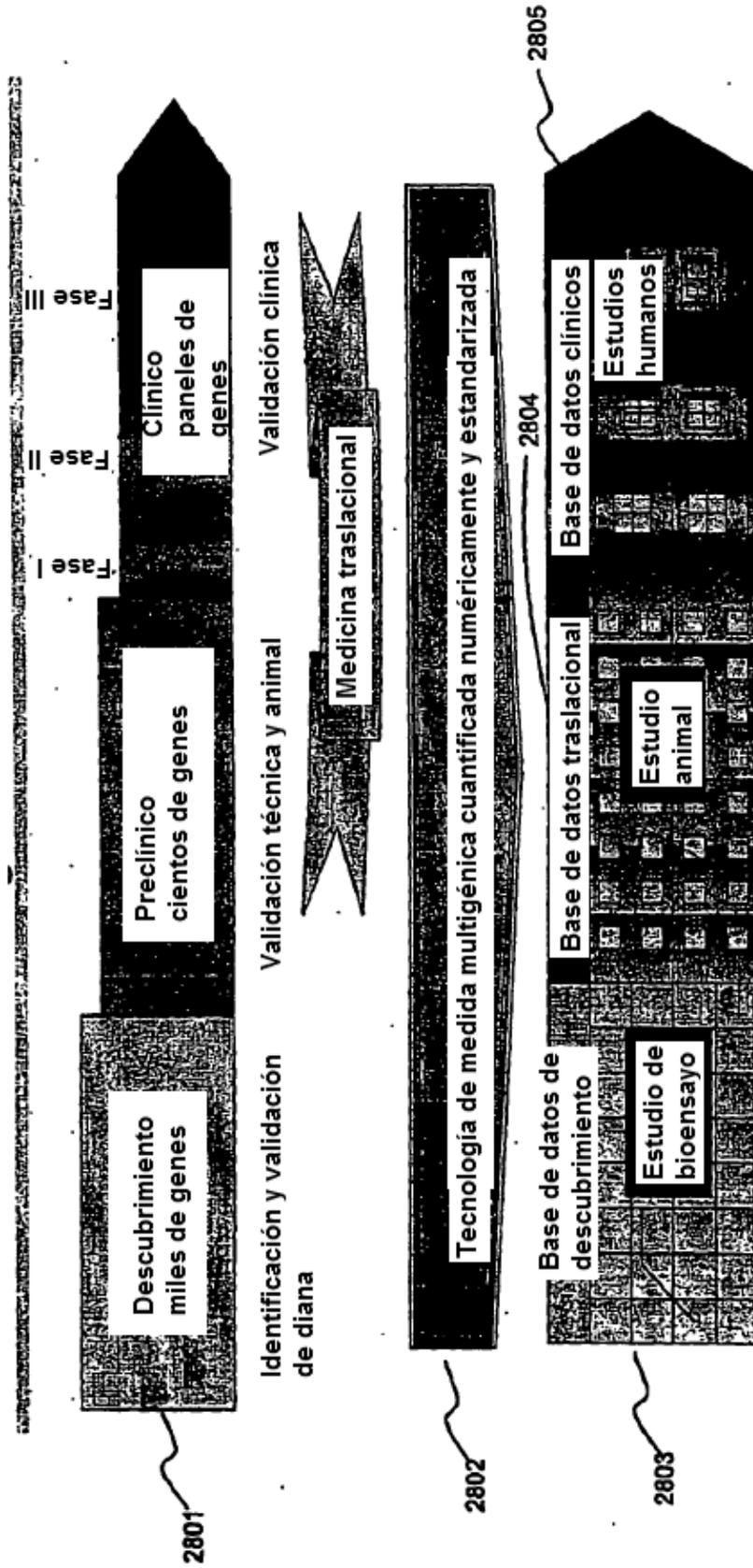


Figura 27

Los datos estandarizados pueden comunicarse a lo largo de todo el proceso



La(s) base(s) de datos de StaRT-PCR permanece(n) "vivas" y son directamente comparables a lo largo del transcurso del descubrimiento de fármacos y la campaña de desarrollo, las comunicaciones mejoran entre grupos funcionales

Figura 28

Gen	Ciclos de ronda 1	Ciclos de ronda 2	Dilución máxima ¹	Valor de expresión génica ²
<i>β-actina</i>	35	0	1/10	NA ³
	5	35	1/100	NA
	8	35	1/100	NA
	10	35	1/1,000	NA
	35	35	1/100,000	NA
<i>c-myc</i>	35	0	1/1,000	8.99E+04
	5	35	1/1,000	8.24E+04
	8	35	1/10,000	7.02E+04
	10	35	1/10,000	3.30E+04
	35	35	1/1,000,000	3.40E+04
<i>catalasa</i>	35	0	1/1,000	1.40E+04
	5	35	1/1,000	3.75E+04
	8	35	1/10,000	8.76E+03
	10	35	1/10,000	1.20E+04
	35	35	1/1,000,000	9.25E+03

¹ Dilución máxima a la que podían detectarse los productos de PCR cuantificables. Para reacciones individuales (ciclos de ronda 2= 0), la dilución mostrada es la de la mezcla de partida de ADNc y mezcla de CT. Para las reacciones múltiples, las diluciones mostradas son las de productos de ronda 1 (por ejemplo 1/100= 1 ml de los 10 ml del producto de ronda 1 + 9 ml de agua)

² Los valores de expresión génica se reseñan como moléculas por 10⁶ moléculas de *β-actina*

³ NA= no aplicable

Figura 29

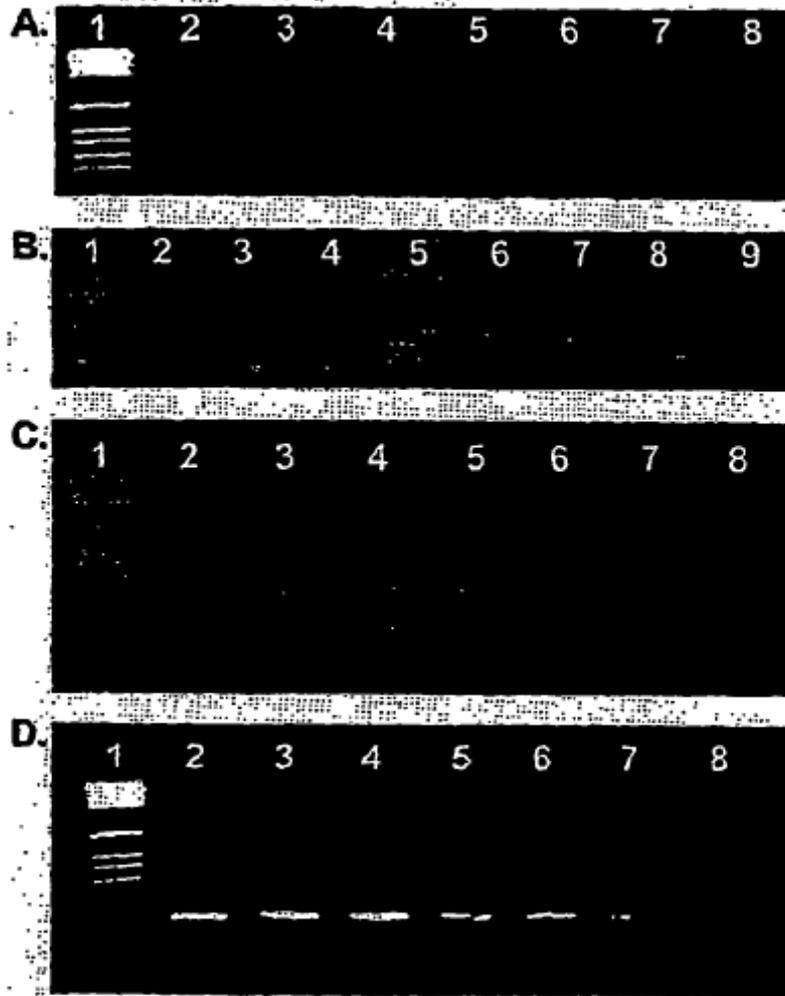


Figura 30

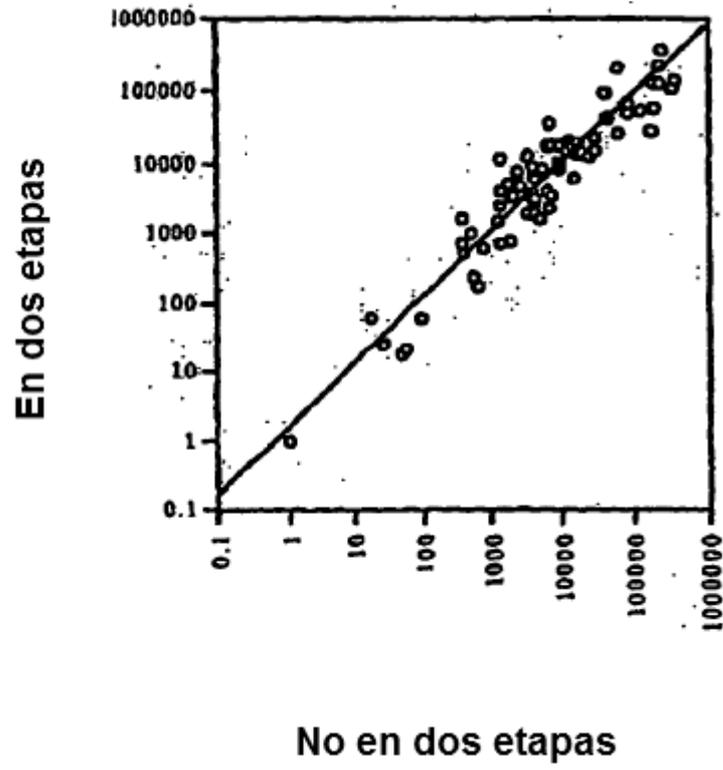


Figura 31

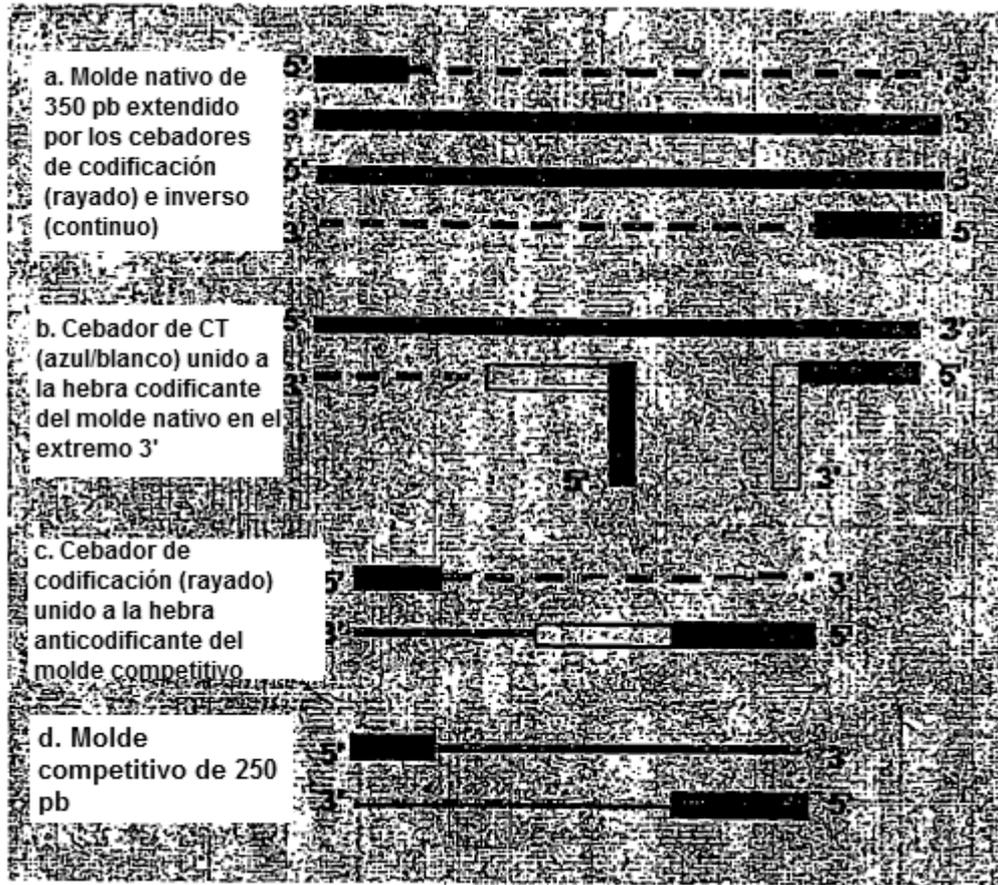


Figura 32

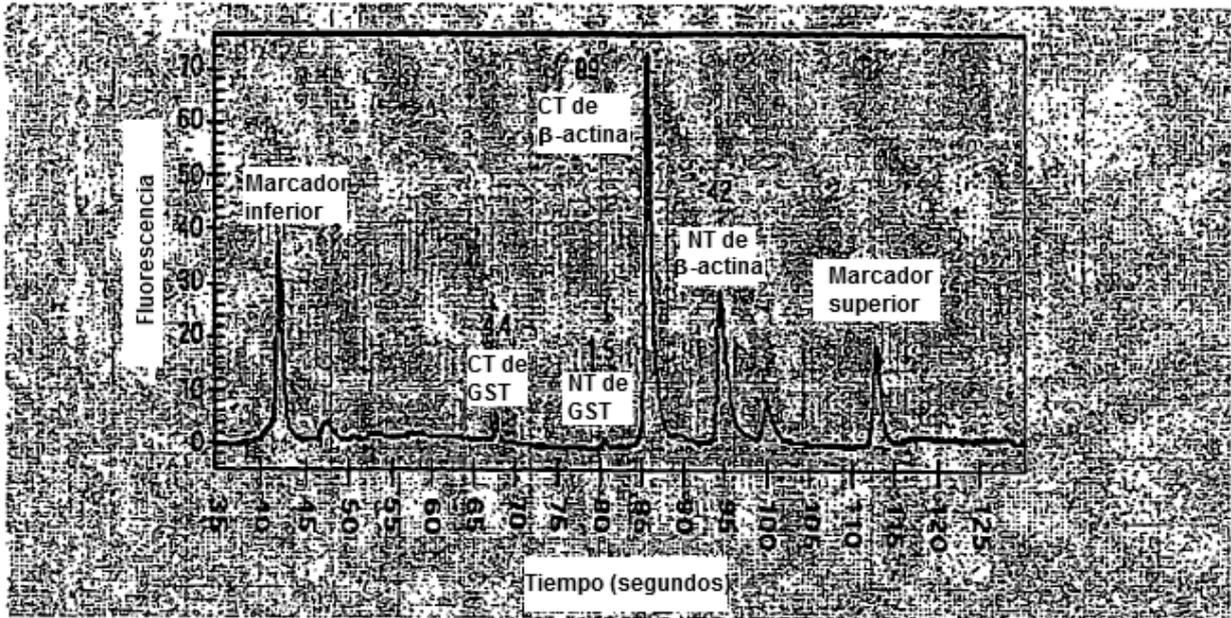


Figura 33

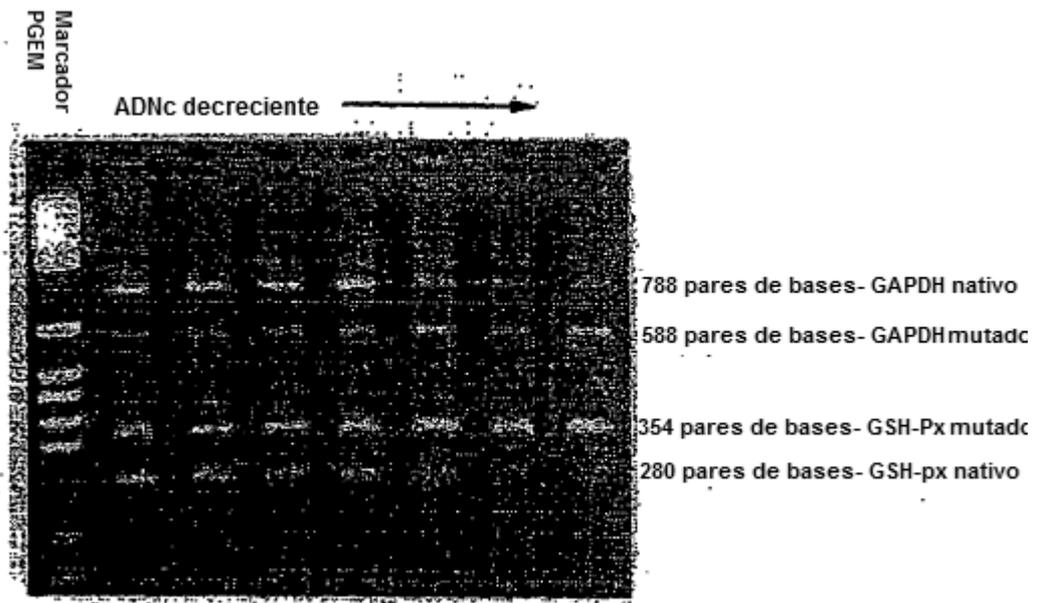


Figura 34

<u>Datos densitométricos del primer carril de muestra de la Fig. 34</u>						
			Longitud (pares de bases)	Valores densitométricos Zeineh	Corrección de tamaño	Valores densitométricos relativos
<u>(A) GAPDH</u>						
GAPDH nativo heterodímero (N ² +2NM)	y		788	10.095		0,820
GAPDH mutado (M ²)			588	1.652	X788/588= 2214	0,180
<u>(B) GSH-Px</u>						
GSH-Px mutado heterodímero (N ² +2M)	y		354	6.709		0,442
GSH-Px (N ²)			280	6.662	X354/280= 8461	0,558

Figura 35

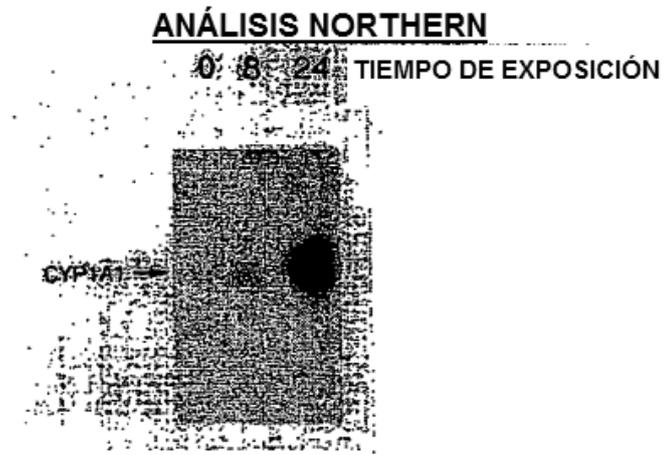


Figura 36A

ANÁLISIS DE RT-PCR EN DOS ETAPAS

TIEMPO DE EXPOSICIÓN (h) 0 24

ADNc en la reacción PCR (μl)

 1 3 5 1 3 5

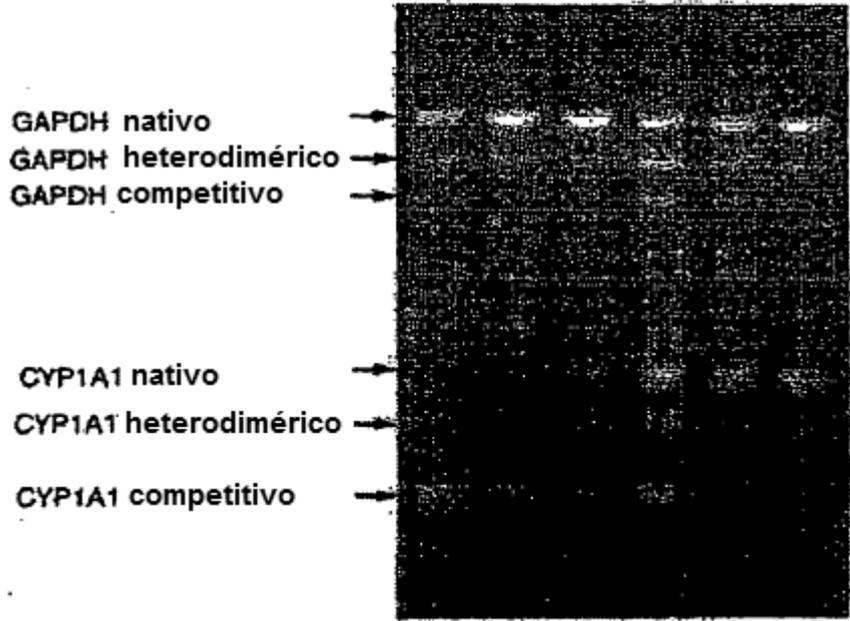


Figura 36B

GEN	Datos de expresión usando algunas realizaciones dadas a conocer en la presente memoria			Datos de expresión de micromatriz
	SVpgC2a	HOE	Diferencia en veces: SVpgC2a/HOE	Diferencia en veces: SVpgC2a/HOE
AHR	420	125	3,3	+3,2
GSTM1,2,4,5	260	30	8,6	+9
GSTP1	10200	26.000	-2,5	-1,8
SOD2	2300	420	5,5	5,1
GPX1	17500	27.000	-1,5	-2,4

La comparación de los datos de expresión génica por micromatriz (técnica de chip de ADNc) en células epiteliales orales humanas normales o inmortalizadas confirmó el concepto de que las matrices génicas son procedimientos de cribado valiosos mientras que otras técnicas, como los procedimientos descritos en la presente memoria, permiten un análisis de expresión génica más detallado

5

Figura 37

El menor umbral de detección, intervalo dinámico lineal más amplio y 100% de sensibilidad posibilitan la identificación de diferencias más significativas

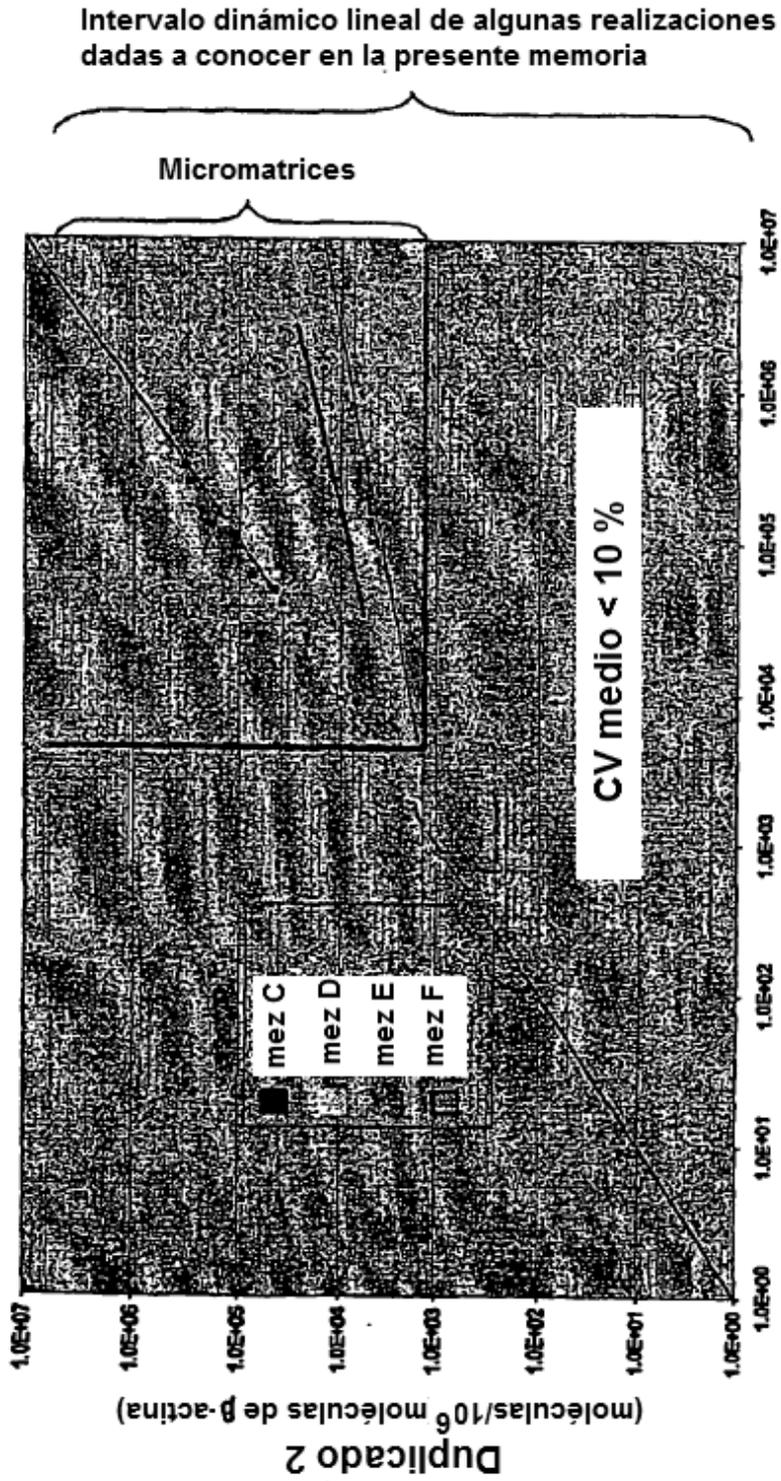


Figura 38