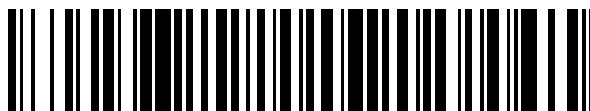


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 719 122**

51 Int. Cl.:

C12Q 1/68 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.03.2015 PCT/US2015/021853**

87 Fecha y número de publicación internacional: **24.09.2015 WO15143385**

96 Fecha de presentación y número de la solicitud europea: **20.03.2015 E 15718004 (3)**

97 Fecha y número de publicación de la concesión europea: **16.01.2019 EP 3119906**

54 Título: **Ensayo de referencia multi-copia**

30 Prioridad:

21.03.2014 US 201461968609 P

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

08.07.2019

73 Titular/es:

**LIFE TECHNOLOGIES CORPORATION (100.0%)
5823 Newton Drive
Carlsbad, CA 92008, US**

72 Inventor/es:

**MERRILL, DAVID;
BRZOSKA, PIUS;
LI, ZHENG;
LIN, WENDY;
LEE, WING y
WONG, MANDI**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 719 122 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

Ensayo de referencia multi-copia

Antecedentes

5 La información en relación con el número de copias de un objetivo de interés en el genoma de una muestra biológica puede ser deseable para varios propósitos, incluyendo la investigación básica y el diagnóstico clínico de diversas enfermedades. Una clase de enfermedades donde el número de copias de un objetivo de interés puede ser particularmente de interés es el cáncer. Numerosos cánceres se presentan con números de copias anormales de uno o más genes. En muchos casos, existe una correlación positiva entre el número de copias y la existencia y/o progresión del cáncer. Por lo tanto, puede ser útil determinar el número de copias de un objetivo de interés en una muestra de un paciente de tejido sospechoso de ser tejido canceroso, por ejemplo, en diagnóstico, tratamiento y/o monitorización del curso del cáncer del paciente.

10 Hasta la fecha, la reacción en cadena de la polimerasa cuantitativa (qPCR) se ha considerado como una técnica molecular para determinar el número de copias de un objetivo de interés. Para determinar el número de copias de un objetivo de interés, la qPCR puede requerir la amplificación simultánea tanto del objetivo de interés como de una secuencia de referencia en el genoma de la muestra. A partir de las cantidades relativas del amplicon del objetivo de interés y el amplicon de la secuencia de referencia y suponiendo que se conoce el número de copias absoluto de la secuencia de referencia, se puede determinar el número de copias del objetivo de interés.

15 Sin embargo, cuando se intenta aplicar la qPCR para determinar el número de copias de un objetivo de interés en una célula cancerosa, la técnica puede resultar difícil como resultado de una o ambas anomalías genómicas tanto aleatorias como evolutivas (p. ej., aneuploidía) encontradas en células cancerosas y las modificaciones que pueden ocurrir en los ácidos nucleicos cuando las muestras de tejido se archivan mediante fijación con formalina e inclusión en parafina (FFPE).

20 Se sabe que ocurren anomalías genómicas de varios tipos en células cancerosas y que aumentan en su cantidad, distribución y complejidad a medida que se replican las células cancerosas con el tiempo. Estas anomalías genómicas incluyen supresiones y multiplicaciones génicas, algunas de las cuales pueden estar relacionadas con una función específica en seres humanos y ser comunes a ciertos tipos de cáncer, mientras que otras pueden no tener un efecto manifiesto o asociación con la enfermedad. Como resultado, una secuencia de referencia de interés puede sufrir una eliminación y/o multiplicación en el genoma de la célula cancerosa, lo que hace que sea muy difícil determinar el número absoluto de copias de la secuencia de referencia.

25 Se sabe que las modificaciones en los ácidos nucleicos (ADN y ARN), como el entrecruzamiento de nucleótidos con ellos mismos o con proteínas, la despurinización de nucleótidos y la fragmentación de ácidos nucleicos se producen como parte del archivo de muestras de cáncer por el método FFPE. Sin embargo, las ubicaciones y el alcance de estas modificaciones son aleatorias y varían mucho a lo largo de las muestras FFPE debido, en gran parte, a uno o más factores, como la variabilidad en el tejido muestra en sí mismo, reactivos usados en la fijación e inclusión y las variaciones preferidas por el usuario en el método FFPE en diferentes laboratorios. Estos efectos también hacen que sea difícil o imposible determinar el número absoluto de copias de la secuencia de referencia.

30 Ambas observaciones proporcionan desafíos significativos a los ensayos de genética molecular actuales y las pruebas para la variación del número de copias a nivel de ADN, como la qPCR.

35 Por lo tanto, existe la necesidad de seleccionar secuencias de referencia que sean relativamente resistentes a las anomalías genómicas inducidas por el cáncer y/o a las modificaciones de los ácidos nucleicos inducidas por la FFPE. Long y cols., describieron en "Multicopy reference assay (MRef) – a superior normalizer of sample input in DNA copy number analysis", publicado en la red (<https://www.qiagen.com/nl/resources/download.aspx?id=a0c54902-07a6-493b-9352-885f9063689b&lang=en>) en el conjunto de secuencias de referencia que se proporciona en múltiples copias en cromosomas múltiples. KATSUHIRO OKUDA y cols., describieron en "Met gene copy number predicts the prognosis for completely resected non-small cell lung cancer", (CANCER SCIENCE 2008, vol. 99, n.º. 11. págs. 2280-2285) el uso de un elemento repetitivo con números de copia conocidos como una referencia.

Compendio de la descripción

40 En una realización, la presente descripción se refiere a un método que comprende cuantificar una secuencia de ácido nucleico de interés en relación con una secuencia de ácido nucleico de referencia en donde al menos un primer número mínimo de copias de la secuencia de ácido nucleico de referencia está presente en cada uno de al menos un segundo número mínimo de cromosomas del ADN genómico del sujeto y determinar un número de copias de la secuencia de interés a partir de la secuencia amplificada de interés cuantificada relativa.

45 El método comprende además amplificar la secuencia del ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto y amplificar la secuencia de ácido nucleico de referencia en la muestra antes de cuantificar.

ES 2 719 122 T3

La secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases de cualquiera de las SEQ ID Nos. 2-8 o en donde la secuencia de referencia comprende de 60-150 pares de bases de la SEQ ID N°. 1. La al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr11-114010-126106, chr19-183944-196032, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr5-180756063-180768074, chr6-170921836-170922549, chrY-26424506-27537936 o chr6-132951-145064, la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas.

10 En una realización particular, la secuencia de ácido nucleico de referencia puede tener al menos un 80% de identidad de secuencia para al menos uno de

GGCTGCTTGCAGTAGTTGTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCC (SEQ ID NO:2),

GGCTGCTTGCGGTAGTTATGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCC (SEQ ID NO:3),

GGCTGCTTGCGGTAGTTGTCTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCC (SEQ ID NO:4),

GGCTGCTTGCGGTAGTTGTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAAAAATTCC (SEQ ID NO:5),

GGCTGCTTGCGGTAGTTGTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCC (SEQ ID NO:6),

GGCTGCTTGCGGTAGTTGTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCN (SEQ ID NO:7),

GGCTGTTTGCGGTAGTAGTCTGTGTAGCAGCAGCACAATGGCCGCAGACGAG
GAAAACAGTTTCTAGGAA (SEQ ID NO:8),

AGTGCAGYRWTGYTGACTCTTCCAAGCTTAACATTTCTCASAARTCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:9),

AGTGCAGCAATGTTGACTCTTCCAAGCTTAACATTTCTCAGAAGTCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:10),

AGTGCAGCGATGCTGACTCTTCCAAGCTTAACATTTCTCACAAGTCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:11),

AGTGCAGCGTTGCTGACTCTTCCAAGCTTAACATTTCTCACAAATCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:12),

AGTGCAGTGATGCTGACTCTTCCAAGCTTAACATTTCTCACAAGTCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:13),

GTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAAACAGTTTCTAGGAATTC
CTCGTATATAATTTTATATTTTGGACAAGATTAATGACCCATGCTCC (SEQ ID
NO:14),

TGCARMGATGCTGACTCTTCCAAGCTTAACATTTCTCACAAGTCAATTAGCTT
TGTACTGGGAGG (SEQ ID NO:15),

TGCTGACTCTTCCAAGCTTAACATTTCTCACAAGTCAATTAGCTTTGTACTGG
GAGGAGGGCGTGAAGGGCTGCTTGCG (SEQ ID NO:16),

CAAGGGACAAGGAAAAATTATCCAAACATTGTTTAAAACAATCATCATTAAAT
TAGTAACACTTATCCAGGGGGGTTTTAACCTTTCCCCCACTCAASGATTATT
CTAATGTCAGAGTAGAATAAAAAATAAGTGCARMGATGCTGAC (SEQ ID
NO:17),

GGAGGAGGAAAATAGGTAGTTTTTCAAAGTTTTCAAATATGAAAAGAAG
AAATGAAATGGTACTTGGAAGAGATTGTTGAAATGGGAGAGACTATGGTGGC
(SEQ ID NO:18),

CAACTAAAAGGCAATGTCCTCCAATAATCACCAGAGTAATCAATTTGCTTA
TTGCTGTCCCTTTAAATATAGTTCTCTGG (SEQ ID NO:19),

GGAGAGACTATGGTGGCTTGTTTAGAAGCAGTTGAGATAGATCCAATTGAGA
TAGAGATATTGAGTATATAAACAAAAGAATGACAAATTAATAGTGTAATGGA
TAACTTGACTTTGGCA (SEQ ID NO:20),

ES 2 719 122 T3

GTGTAATGGATAACTTGACTTTGGCAAATATTGTGAATTTTGTGAAAGTACA
ACTAAAAGGCAATGTCACTCCAATAATCACCAG (SEQ ID NO:21),

GTAATCAATTTGCTTATTGCTGICCCCTTAAATATAGTTCTCTGGTATCAACTA
ACATGTTTTTAACTAATGATGCTTCTTAAAGAAAAGGGAAAAGACCT (SEQ ID
NO:22),

CCCTGGGCCCTCAGGGGAGTCCCTGCTGGACAGTGAGACAGAGAATGACCA
TGATGATGCTTTCCT (SEQ ID NO:23),

GGGTTTATGTTTGATATRTAATGTAATTTTCTAATGCTAAATCAAGTGGTAAT
TTTGTTAGTCAAGTTGATTTAGTGGCTTGGGAAGAAAGCT (SEQ ID NO:24),

GAGACCCCCAGGTGTTGAGGCAGGGCTGGGGTGTCCCCTTCCAACCAGGCTG
TCAAGGCCCAACTCTGGGGCAGAGGCAGTGGCAGGG (SEQ ID NO:25),

CATCCGTTTCACCTGCAGTTGAAGATCCGTGAGGTGCCCAGAAGATCATGCA
GTCAWCAGTCCCACG (SEQ ID NO:26),

GAKATAAGGAAGCTCGAGGAAGAGAAAAACAACCTGGAAGGAGAAATCATA
GATTTTTATAAAATGAMAGCTGCCTCTGAAGC (SEQ ID NO:27),

CCGTTTTGGAGGAGGAACAGATTCCATGTCCACTAGAATGGAATGAACAAGA
AATGGAGGAGGAAAATAGGTAGTTTTTCAAAGTTTTCAAAAATATGAAAAG
AAGAAATGAAATGGTACTTGGAAAGAGATTGTTGAAATGGGA (SEQ ID NO:28),

TGCTTCTTAAAGAAAAGGGAAAAGACCTTTTTCTTTCTTTCAGTCTTCAATGA
TTCAGTCTTCATCTCGCTCCACCAAAGATAAATGAAATCTACATCTCT (SEQ
ID NO:29),

CTTTCCCCACTCAASGATTATTCTAATGTCAGAGTAGAATAAAAAATAAGTG
CARMGATGCTGACTCTTCCAAGCTTAACATTTCTCA (SEQ ID NO:30), or

GGGAGGAGGGCGTGAAGGGCTGCTTGCGGTAGTTGTGTAGCAGCAGCACAAT
GGCCGCAGACAAG (SEQ ID NO:31).

La presente descripción se refiere a un kit, que comprende una primera sonda que hibrida específicamente con al menos una porción de al menos una secuencia de referencia que tiene al menos un 80% de identidad de secuencia

5 con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases en donde la al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr10-38741930-38753964, chr11-114010-126106, chr16-90239446-90251554, chr19-183944-196032, chr2-114323560-114323652, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr4-165196360-165199636, chr5-180756063-180768074, chr6-170921836-170922549, chr7-39837560-63231088, chr7-128296352-128298474, chr8-143133-150475, chr9-49679-49771, chrY-26424506-27537936 o chr6-132951-145064, la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas.

10 En otra realización más, la presente descripción se refiere a una composición que comprende una primera sonda que hibrida específicamente con al menos una porción de al menos una secuencia de referencia que tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases de cualquiera de las SEQ ID Nos 2-8 o en donde la secuencia de referencia comprende de 60-150 pares de bases de la SEQ ID N° 1. La al menos una porción puede estar presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr11-114010-126106, chr19-183944-196032, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr5-180756063-180768074, chr6-170921836-170922549, chrY-26424506-27537936 o chr6-132951-145064, la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas.

20 En aún otra realización, la presente descripción se refiere a un sistema que comprende un amplificador de ácido nucleico configurado para amplificar una secuencia de ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto y amplificar una secuencia de referencia en la muestra; un reservorio de reactivo que contiene al menos un primer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de referencia, en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases de cualquiera de las SEQ ID Nos 2-8, o en donde la secuencia de referencia comprende de 60-150 pares de bases de la SEQ ID N° 1. La al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr11-114010-126106, chr19-183944-196032, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr5-180756063-180768074, chr6-170921836-170922549, chrY-26424506-27537936 o chr6-132951-145064, la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas, y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de referencia; un detector configurado para proporcionar una primera indicación en relación con una cantidad de la secuencia amplificada de interés y una segunda indicación en relación con una cantidad de la secuencia de referencia amplificada, y un controlador configurado para cuantificar la secuencia de interés amplificada en relación con la secuencia de referencia, en función, al menos en parte, de la primera indicación y de la segunda indicación y determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa.

Breve descripción de las figuras

Los siguientes dibujos forman parte de la presente especificación y están incluidos para demostrar además ciertos aspectos de la presente descripción. La descripción se puede entender mejor haciendo referencia a uno o más de estos dibujos en combinación con la descripción detallada de realizaciones específicas presentadas en esta memoria.

45 La Figura 1A muestra el número de copias calculado con punto decimal de IRS2 según lo se determinado por qPCR con un ensayo de referencia de RNaseP.

La Figura 1B muestra el número de copias redondeado de IRS2 según lo determinado por qPCR con un ensayo de referencia de RNaseP.

50 La Figura 2A muestra el número de copias calculado con punto decimal de IRS2 según lo determinado por qPCR con un ensayo de referencia basado en las SEQ ID N°: 1-8.

La Figura 2B muestra el número de copias redondeado de IRS2 según lo determinado por qPCR con un ensayo de referencia basado en las SEQ ID N°: 1-8.

La Figura 3A muestra el número de copias calculado con punto decimal de IRS2 según lo determinado por qPCR con un ensayo de referencia basado en las SEQ ID N°: 9-13.

55 La Figura 3B muestra el número de copias redondeado de IRS2 según lo determinado por qPCR con un ensayo de referencia basado en las SEQ ID N°: 9-13.

Descripción

Varias realizaciones de la presente descripción proporcionan secuencias de referencia que son relativamente resistentes a anomalías genómicas inducidas por cáncer y/o modificaciones de ácido nucleico inducidas por FFPE. Más específicamente, la descripción proporciona secuencias objetivo en el genoma que se repiten varias veces, a lo largo de muchos cromosomas y demuestran un número de copias sustancialmente normal y resistencia a modificaciones y fragmentación severas, incluso en células cancerosas sometidas a FFPE.

En una realización, la presente descripción se refiere a un método que comprende amplificar una secuencia de ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto; amplificar una secuencia de ácido nucleico de referencia en la muestra en donde está presente al menos una copia de la secuencia de ácido nucleico de referencia en cada uno de al menos diez cromosomas del ADN genómico del sujeto; cuantificar la secuencia amplificada de interés en relación con la secuencia de referencia amplificada, y determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa.

La amplificación se puede realizar mediante cualquier técnica conocida por el experto en la técnica. De manera deseable, la amplificación se puede realizar mediante una técnica que permita la cuantificación de la secuencia de interés en relación con la secuencia de referencia. Técnicas ejemplares incluyen, pero no se limitan a, reacción en cadena de la polimerasa (PCR) (Saiki et al., (1985) *Science* 230: 1350), PCR cuantitativa a tiempo real (qPCR), PCR digital, reacción en cadena de la ligasa (LCR) (Landegren et al., (1988) *Science* 241: 1077-1080), amplificación dependiente de helicasa (HDA) (Vincent et al., (2004) *EMBO rep* 5(8): 795-800), HDA termoestable (tHDA) (An et al., (2005) *J. Biol. Chem.* 280 (32): 28952-28958), amplificación por desplazamiento de cadena (SDA) (Walker et al., (1992) *Nucleic Acids Res.* 20(7): 16916), amplificación por desplazamiento múltiple (MDA) (Dean et al., (2002) *Proc. Natl. Acad. Sci. USA* 99(8): 5261-5266), amplificación en círculo rodante (RCA) (Liu et al., (1996) *J. Am. Chem. Soc.* 118: 1587-1594), RCA con restricción asistida (Wang et al., (2004) *Genome Res.* 14: 2357-2366), amplificación isotérmica con un único cebador (SPIA) (Dafforn et al., (2004) *Biotechniques* 37(5): 854-7), amplificación mediada por transcripción (TMA) (Vuorinen et al., (1995) *J. Clin. Microbiol.* 33: 1856-1859), reacción de amplificación con enzima de corte (NEAR) (Maples et al., US2009017453), reacción de amplificación exponencial (EXPAR) (Van Ness et al., (2003) *Proc. Natl. Acad. Sci. USA* 100(8): 4504-4509), amplificación isotérmica mediada por bucle (LAMP) (Notomi et al., (2000) *Nucleic Acids Res.* 28(12): e63), amplificación de polimerasa recombinasa (RPA) (Piepenburg et al., (2006) *PloSBiol* 4(7): 1115-1120), amplificación basada en secuencia de ácido nucleico (NASBA) (Kievits et al., (1991) *J. Virol. Methods* 35: 273-286), proceso de amplificación inteligente (SMAP) (Mitani et al. (2007), *Nat. Methods* 4(3): 257-62), amplificación nanostring (Geiss et al., (2008) *Nature Biotechnology* 26: 317-325; Schwanhausser et al., (2011) *Nature* 473: 337-342; comercialmente disponible como la plataforma nCounter® de NanoString Technologies, Seattle, WA) o secuenciación de próxima generación (NGS) (Rothberg et al., (2011) *Nature* 475: 348-352; Metzker M (2010) *Nature Rev Genetics* 11: 31-46).

En una realización particular, la amplificación se realiza mediante la reacción en cadena de la polimerasa cuantitativa TaqMan (qPCR).

De manera general, los ensayos de la plataforma qPCR usan dos objetivos genómicos juntos para determinar el número de copias de un gen o región del genoma en una muestra de prueba. Uno de los objetivos del genoma es un ensayo de qPCR para el objetivo de interés (TOI) y el segundo es un ensayo de referencia de qPCR para lo que se supone que es una región del genoma normal sin modificar. Los dos ensayos se realizan simultáneamente y en paralelo en la misma muestra de prueba. Después de uno o más ciclos de la reacción en cadena de la polimerasa se pueden determinar los valores Cq de cada ensayo (indicativos de la cantidad relativa de TOI o de amplificación de referencia) mediante técnicas conocidas por el experto en la técnica y/o descritos en más detalle a continuación, y se calcula un Cq delta entre ellos. Este Cq delta calculado se puede comparar con un Cq delta que es representativo de un número de copias conocido para el TOI. Por ejemplo, el Cq delta representativo puede ser un Cq delta determinado a partir de una muestra que se sabe que es normal (es decir, que tiene un número de copias de 2, una copia de cada de uno de un par de cromosomas). Este valor de Cq delta final calculado entre la muestra de prueba y la muestra conocida se puede luego transformar en un número decimal o en un número entero que represente el número de copias del gen o región del genoma en la muestra de prueba.

Por las razones descritas anteriormente, el reto en las muestras FFPE de cáncer está en la capacidad de encontrar un objetivo genómico de referencia (objetivo de ensayo de referencia de qPCR) que es tanto normal como relativamente sin modificar por fijación e inclusión. Adicional o alternativamente, para algunas muestras, puede ser complicada la capacidad de encontrar un objetivo genómico de referencia por un estado particular de enfermedad que incluye, pero no se limita a, cáncer, donde se puede ver alterado el objetivo genómico de referencia potencial por el estado de enfermedad y se multiplica el mismo en relación con su población típica.

Cualquier secuencia de ácido nucleico a partir del ADN genómico de un sujeto y de interés para el usuario del método se puede amplificar y cuantificar de acuerdo con el método. Por ejemplo, la secuencia de interés puede ser al menos una porción de un gen que posee una asociación con una enfermedad. Como será evidente para la persona experta en la técnica, la muestra puede ser cualquier tejido probable o que contenga posiblemente la secuencia de ácido nucleico de interés en el ADN genómico.

5 El método se puede usar para amplificar y cuantificar una secuencia de interés a partir de tejido sospechoso de que sea tejido canceroso, incluyendo tejido que se ha sometido a fijación con formalina e inclusión en parafina (FFPE) antes de la amplificación de la secuencia de interés y de la amplificación de la secuencia de referencia. En tal uso, la secuencia de interés puede ser al menos una porción de un gen para el que existe una correlación entre el número de copias del gen y la presencia y/o estadio de un cáncer.

10 Se puede amplificar cualquier secuencia de ácido nucleico de referencia conocida o que se espere que esté presente en el ADN genómico de la muestra. Sin embargo, una persona con experiencia en la técnica será consciente de que en muchas realizaciones como aquellas en las que se sospecha que la muestra es tejido canceroso y, especialmente una muestra previamente sometida a FFPE, cualquier locus particular de una copia de una secuencia de ácido nucleico de referencia puede haber sufrido un evento de recombinación, un evento de aneuploidía o similar. Por lo tanto, el número de copias de la secuencia de referencia en la muestra puede diferir del esperado mediante recuento simple del número de locus de la secuencia de referencia en una muestra sin enfermedad del sujeto o un miembro de las especies del sujeto.

15 Por lo tanto, es deseable que la muestra comprenda al menos una copia de la secuencia de ácido nucleico de referencia en cada uno de al menos diez cromosomas del ADN genómico del sujeto. La presencia de múltiples copias físicamente dispersas de la secuencia de referencia puede suavizar o promediar los efectos de interrupciones individuales o duplicaciones de varios locus.

20 En una realización, la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases de cualquiera de las SEQ ID Nos 2-8 o en donde la secuencia de referencia comprende de 60-150 pares de bases de la SEQ ID N° 1. La al menos una porción puede estar presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr11-114010-126106, chr19-183944-196032, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr5-180756063-180768074, chr6-170921836-170922549, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas.

25
30 En una realización, el primer número mínimo puede ser 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 copias. Independientemente, el segundo número mínimo puede ser 5, 6, 7, 8, 9, 10 o 11 cromosomas. Los primeros y segundos números mínimos se pueden enumerar, estimar o predecir en función de cualquier genoma humano de referencia disponible.

En una realización particular, la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de

GGCTGYTTGCRGTAGTWRTSTRKSWRSMRSMRMWWSRMYGSMSRCARRSRA
RRMARWYWSTWDVWAKKMN (SEQ ID NO:1),

GGCTGCTTGCAGTAGTTGTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCC (SEQ ID NO:2),

GGCTGCTTGCGGTAGTTATGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCC (SEQ ID NO:3),

GGCTGCTTGCGGTAGTTGTCTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCC (SEQ ID NO:4),

GGCTGCTTGCGGTAGTTGTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAAAAATTCC (SEQ ID NO:5),

GGCTGCTTGCGGTAGTTGTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCC (SEQ ID NO:6),

GGCTGCTTGCGGTAGTTGTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAA
ACAGTTTCTAGGAATTCCN (SEQ ID NO:7),

GGCTGTTTGC GG TAGTAGTCTGTGTAGCAGCAGCACAATGGCCGCAGACGAG
GAAAACAGTTTCTAGGAA (SEQ ID NO:8),

AGTGCAGYRWTGYTGACTCTTCCAAGCTTAACATTTCTCASAARTCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:9),

AGTGCAGCAATGTTGACTCTTCCAAGCTTAACATTTCTCAGAAGTCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:10),

AGTGCAGCGATGCTGACTCTTCCAAGCTTAACATTTCTCACAAGTCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:11),

AGTGCAGCGTTGCTGACTCTTCCAAGCTTAACATTTCTCACAAATCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:12),

AGTGCAGTGATGCTGACTCTTCCAAGCTTAACATTTCTCACAAGTCAATTAGC
TTTGTACTGGGAGG (SEQ ID NO:13),

GTGTAGCAGCAGCACAATGGCCGCAGACAAGGAAAACAGTTTCTAGGAATTC
CTCGTATATAATTTTATATTTTGGACAAGATTAATGACCCATGCTCC (SEQ ID
NO:14),

TGCARMGATGCTGACTCTTCCAAGCTTAACATTTCTCACAAGTCAATTAGCTT
TGACTGGGAGG (SEQ ID NO:15),

TGCTGACTCTTCCAAGCTTAACATTTCTCACAAGTCAATTAGCTTTGTACTGG
GAGGAGGGCGTGAAGGGCTGCTTGCG (SEQ ID NO:16),

CAAGGGACAAGGAAAATTATCCAAACATTGTTTAAAACAATCATCATTAAAT
TAGTAACACTTATCCAGGGGGGTTTTTAACCTTTCCCCCACTCAASGATTATT
CTAATGTCAGAGTAGAATAAAAAATAAGTGCARMGATGCTGAC (SEQ ID
NO:17),

GGAGGAGGAAAATAGGTAGTTTTTCAAAGTTTTCAAATATGAAAAGAAG
AAATGAAATGGTACTTGGAAGAGATTGTTGAAATGGGAGAGACTATGGTGGC
(SEQ ID NO:18),

CAACTAAAAGGCAATGTCACTCCAATAATCACCAGAGTAATCAATTTGCTTA
TTGCTGTCCCTTTAAATATAGTTCTCTGG (SEQ ID NO:19),

GGAGAGACTATGGTGGCTTGTTTAGAAGCAGTTGAGATAGATCCAATTGAGA
TAGAGATATTGAGTATATAAACAAAAGAATGACAAATTAATAGTGTAATGGA
TAACTTGACTTTGGCA (SEQ ID NO:20),

GTGTAATGGATAACTTGACTTTGGCAAATATTGTGAATTTTTGTGAAAGTACA
ACTAAAAGGCAATGTCACTCCAATAATCACCAG (SEQ ID NO:21),

GTAATCAATTTGCTTATTGCTGTCCCTTTAAATATAGTTCTCTGGTATCAACTA
ACATGTTTTTAACTAATGATGCTTCTTAAAGAAAAGGGAAAAGACCT (SEQ ID
NO:22),

CCCTGGGCCCCTCAGGGGAGTCCCTGCTGGACAGTGAGACAGAGAATGACCA
TGATGATGCTTTCCT (SEQ ID NO:23),

GGGTTTATGTTTGATATRTAATGTAATTTTCTAATGCTAAATCAAGTGGTAAT
TTTGTTAGTCAAGTTGATTTAGTGGCTTGGGAAGAAAGCT (SEQ ID NO:24),

GAGACCCCAGGTGTTGAGGCAGGGCTGGGGTGTCCCCTTCCAACCAGGCTG
TCAAGGCCCAACTCTGGGGCAGAGGCAGTGGCAGGG (SEQ ID NO:25),

CATCCGTTTCACCTGCAGTTGAAGATCCGTGAGGTGCCCAGAAGATCATGCA
GTCAWCAGTCCCACG (SEQ ID NO:26),

GAKATAAGGAAGCTCGAGGAAGAGAAAAACAACCTGGAAGGAGAAATCATA
GATTTTTATAAAATGAMAGCTGCCTCTGAAGC (SEQ ID NO:27),

ES 2 719 122 T3

CCGTTTTGGAGGAGGAACAGATTCCATGTCCACTAGAATGGAATGAACAAGA
AATGGAGGAGGAAAATAGGTAGTTTTTCAAAAGTTTTCAAAAATATGAAAAG
AAGAAATGAAATGGTACTTGGGAAGAGATTGTTGAAATGGGA (SEQ ID NO:28),

TGCTTCTTAAAGAAAAGGGAAAAGACTTTTTCTTTCTTTTCAGTCTTCAATGA
TTCCTGCTTCATCTCGCTCCACCAAAGATAAATGAAATCTACATCTCT (SEQ
ID NO:29),

CTTCCCCCACTCAASGATTATTCTAATGTCAGAGTAGAATAAAAAATAAGTG
CARMGATGCTGACTCTTCCAAGCTTAACATTTCTCA (SEQ ID NO:30), or

GGGAGGAGGGCGTGAAGGGCTGCTTGCGGTAGTTGTGTAGCAGCAGCACAAT
GGCCGCAGACAAG (SEQ ID NO:31).

En una realización, la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-8.

5 En una realización, la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 9-13.

10 Un primer conjunto de secuencias, SEQ ID Nos: 1-8, corresponde a secuencias encontradas en el genoma humano en chr1: 121836-121905, chr1: 243203810-243203879, chr1:341419+341488, chr1:648175-648244, chr2: 243071825+243071894, chr3: 197962362+197962431, chr4: 119569113+119569182, chr5: 180768034+180768103, chr6: 132997-133066, chr6: 170922434+170922503, chr10: 38753924+38753993, chr11: 114056-114125, chr16: 90251439+90251508, chr19: 183990-184059, chr20: 62933558+62933627, chrUn_g1000227: 58864-58933, chrY: 26436540+26436609 y chrY: 27525831-27525900.

15 Un segundo conjunto de secuencias, SEQ ID Nos: 9-13, corresponde a secuencias encontradas en el genoma humano en chr1: 224126101-224126167, chr1: 228152189+228152255, chr1: 243203891-243203957, chr1: 341341+341407, chr1: 648256-648322, chr2: 243071747+243071813, chr3: 197962284+197962350, chr4: 119569035+119569101, chr5: 180767956+180768022, chr6: 133078-133144, chr6: 170922356+170922422, chr8: 143260-143326, chr10: 38753846+38753912, chr11: 114137-114203, chr16: 90251361+90251427, chr19: 184071-184137, chr20: 62933480+62933546 y chrUn_g1000227: 58945-59011.

En cualquier secuencia de ácido nucleico que figura en esta memoria, se usa la tabla IUPAC estándar de nucleótidos de origen natural y degenerados:

Símbolo	Descripción	Bases representadas
A	adenina	A
C	citocina	C
G	guanina	G
T	timina	T
U	uracilo	U
W	Débil	A, T
S	Fuerte	C, G
M	Amino	A, C
K	ceto	G, T
R	Purina	A, G
Y	Pirimidina	C, T

B	No adenina	C, G, T
D	No citosina	A, G, T
H	No guanina	A, C, T
V	No timina	A, C, G,
N	Cualquier base (no un hueco)	A, C, G, T

5 Aunque no está limitado por la teoría, el presente inventor ha encontrado que cada una del primer conjunto y cada una del segundo conjunto de secuencias están altamente repetidas (~20 copias en el genoma humano), físicamente dispersos a lo largo del genoma humano y relativamente más resistentes a la interrupción y/o duplicación por FFPE que las secuencias típicas de ADN genómico. Como resultado, una secuencia que tenga al menos un 80% de identidad con una o más de las SEQ ID Nos: 1-13 puede ser particularmente adecuada como una secuencia de referencia, especialmente en muestras sospechosas de ser tejido canceroso, en particular tejido procesado con FFPE.

10 Las etapas de amplificación generan una secuencia de interés amplificada y una secuencia de referencia amplificada. En general, siempre que la ejecución de las etapas de amplificación esté sincronizada y la amplificación no haya avanzado hasta un punto donde la cantidad de cualquier reactivo que no sea la secuencia de interés amplificada y la secuencia de referencia amplificada sea limitante de la velocidad, en cualquier punto, las cantidades relativas de la secuencia de interés amplificada y de la secuencia de referencia amplificada serán proporcionales a su número de copias en el ADN genómico de la muestra.

15 Por lo tanto, el método puede comprender cuantificar la secuencia de interés amplificada en relación con la secuencia de referencia amplificada. Se puede realizar la cuantificación mediante cualquier técnica conocida por el experto en la técnica. Por ejemplo, mediante el uso de dos sondas, cada una de las cuales comprende un resto fluorescente en un primer extremo y un quencher para este resto fluorescente en un segundo extremo, con una sonda que hibrida específicamente con la secuencia de interés y la otra que hibrida específicamente con la secuencia de referencia, en TaqMan qPCR, la escisión del quencher por la acción de la Taq polimerasa generará una señal de fluorescencia proporcional a la cantidad de sonda hibridada con la secuencia de interés o con la secuencia de referencia. Por lo tanto, en un ejemplo hipotético simple no limitante, si la señal de fluorescencia de la sonda que se hibrida con la secuencia de referencia es cinco veces más intensa que la señal de fluorescencia de la sonda que se hibrida con la secuencia de interés, la cantidad relativa de la secuencia de interés amplificada sería 0,2. (Como será evidente para la persona experta en la técnica, se pueden usar expresiones matemáticamente equivalentes alternativas para llegar a una cantidad relativa).

25 En algunas realizaciones, se puede omitir la amplificación del TOI y de la secuencia de referencia. Se conocen por los expertos en la técnica las técnicas para cuantificar secuencias de ácido nucleico no amplificadas.

30 La medida de la cuantificación relativa se puede informar usando el término "cambio de veces" que se refiere a la cantidad de producto amplificado (que se relaciona con el número de copias) en la secuencia de interés en relación con el objetivo genómico de referencia. El cambio de veces puede cuantificar usando cualquiera de los varios métodos disponibles que incluyen, pero no se limitan a, los descritos por Livak et al., (Methods, 25: 402-408 (2001)), productos comercialmente disponibles como CopyCaller™ (Applied Biosystems) o cualquier otro algoritmo adecuado para comparar cantidades de señales de fluorescencia. En muchas realizaciones, el cambio de veces se determina comparando el C_T de la secuencia de interés con el C_T del objetivo genómico de referencia. Algunos algoritmos adecuados incluyen, pero no se limitan a, los métodos descritos en la Solicitud de EE.UU. de N° de serie 13/107786, "Karyotyping Assay" presentada el 13 de mayo de 2011.

35 La etapa de cuantificación produce una secuencia de interés amplificada cuantificada relativa. El método puede entonces comprender determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa. La determinación requiere una indicación del número de copias de la secuencia de referencia. Se puede proporcionar dicha indicación mediante el análisis del genoma de una muestra sin enfermedad del sujeto o de uno o más miembros de la especie del sujeto. Esta técnica puede ser especialmente adecuada con respecto a muestras sospechosas de ser tejido canceroso para una secuencia de referencia que está una o más de muy repetida, físicamente dispersa y relativamente resistente a la ruptura y/o duplicación por FFPE. Por ejemplo, se puede esperar que el número de copias de una secuencia de referencia que tiene al menos un 80% de identidad con una o más de las SEQ ID Nos: 1-13 en una muestra sin enfermedad sea sustancialmente igual al número de copias de la secuencia de referencia en una muestra sospechosa de ser tejido canceroso.

40 Continuando con el ejemplo hipotético simple no limitante iniciado anteriormente, si el número de copias de la secuencia de referencia es 20, entonces se puede determinar que el número de copias de la secuencia de interés amplificada es $20 * 0,2 = 4$. (Como debería ser evidente, esto es un ejemplo simple basado en la sonda de un cálculo del número de copias. Es una cuestión de rutina para el experto en la técnica, que tiene el beneficio de la presente descripción, realizar un cálculo del número de copias para otras técnicas de ensayo, como la qPCR).

El número de copias determinado de la secuencia de interés se puede usar para cualquier propósito que se recomendaría al experto en la técnica. En una realización particular, el método puede además comprender diagnosticar que el sujeto tiene un biomarcador relacionado con el cáncer en función de la secuencia de interés que está asociada con el cáncer y el número de copias que es indicativo del cáncer.

- 5 La presente descripción se refiere a un kit que comprende una primera sonda que hibrida específicamente con al menos una porción de al menos una secuencia de referencia que tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases, en donde al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr10-38741930-38753964, chr11-114010-126106, chr16-90239446-90251554, chr19-183944-196032, chr2-114323560-114323652, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr4-165196360-165199636, chr5-180756063-180768074, chr6-170921836-170922549, chr7-39837560-63231088, chr7-128296352-128298474, chr8-143133-150475, chr9-49679-49771, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas.

En una realización particular, la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-31.

En una realización más particular, la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-13.

- 20 Una "sonda" como se usa en esta memoria, se refiere a un compuesto que comprende una secuencia de ácido nucleico y un resto detectable. Como tal, y para evitar dudas, cualquier "sonda" a la que se hace referencia en esta memoria no es de origen natural.

- 25 En una realización, la primera sonda comprende una secuencia de ácido nucleico configurada para hibridar específicamente con al menos la porción de la al menos una secuencia de referencia, un indicador fluorescente en un primer extremo de la secuencia de ácido nucleico y un quencher fluorescente en un segundo extremo de la secuencia de ácido nucleico.

En una realización adicional, la secuencia de ácido nucleico se configura para hibridar específicamente con la totalidad de al menos una secuencia de referencia.

- 30 Se puede determinar un porcentaje de identidad de secuencia mediante cualquier técnica conocida por el experto en la técnica. En algunas realizaciones, la secuencia de referencia tiene al menos un 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99,5% o 99,9% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-13.

Mediante el uso de técnicas conocidas por la persona experta en la técnica, la primera sonda puede permitir la detección de al menos la porción de la al menos una secuencia de referencia.

- 35 Además de la primera sonda, el kit puede comprender otros componentes. Por ejemplo, el kit puede comprender además un primer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de referencia, y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria en un segundo extremo de la al menos una secuencia de referencia.

- 40 Por "cebadores" se entiende moléculas de ácido nucleico que, en presencia de la al menos una secuencia de referencia y otro(s) reactivo(s), pueden permitir la amplificación de la al menos una secuencia de referencia.

De manera alternativa o adicional, el kit puede comprender además una segunda sonda que hibrida específicamente con al menos una porción de al menos una secuencia de ácido nucleico de interés. Aparte de la secuencia con la que hibrida específicamente, la segunda sonda puede tener las mismas características que la primera sonda descrita anteriormente.

- 45 Cualquier secuencia de ácido nucleico de interés puede ser el objetivo de hibridación de la segunda sonda. En una realización, la secuencia de interés es una porción o la totalidad de un gen asociado con un cáncer.

- 50 De manera alternativa o adicional, el kit puede comprender además un tercer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de ácido nucleico de interés, y un cuarto cebador configurado para hibridar específicamente con una secuencia complementaria con un segundo extremo de la al menos una secuencia de ácido nucleico de interés. En una realización, la presente descripción se refiere a una composición que comprende una primera sonda que hibrida específicamente con al menos una porción de al menos una secuencia de referencia que tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases, en donde la al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr10-38741930-38753964, chr11-114010-126106, chr16-90239446-90251554, chr19-

ES 2 719 122 T3

183944-196032, chr2-114323560-114323652, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr4-165196360-165199636, chr5-180756063-180768074, chr6-170921836-170922549, chr7-39837560-63231088, chr7-128296352-128298474, chr8-143133-150475, chr9-49679-49771, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas.

En una realización particular, la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-31.

En una realización más particular, la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-13.

La primera sonda puede ser sustancialmente la misma que la primera sonda del kit, descrita anteriormente.

La composición puede comprender además uno o más de (i) un primer cebador configurado para hibridar específicamente con un primer extremo de al menos una secuencia de referencia, y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de al menos una secuencia de referencia; (ii) una segunda sonda que hibrida específicamente con al menos una porción de al menos una secuencia de ácido nucleico de interés; o (iii) un tercer cebador configurado para hibridar específicamente con un primer extremo de al menos una secuencia de ácido nucleico de interés y un cuarto cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de al menos una secuencia de ácido nucleico de interés, sustancialmente lo mismo que los componentes adicionales correspondientes del kit, anteriormente descritos.

En una realización, la presente descripción se refiere a un sistema que comprende:

un amplificador de ácido nucleico configurado para amplificar una secuencia de ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto y amplificar una secuencia de referencia en la muestra;

un reservorio de reactivo que contiene al menos un primer cebador configurado para hibridar específicamente con un primer extremo de al menos una secuencia de referencia, en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases de cualquiera de las SEQ ID Nos: 2-8, o en donde la secuencia de referencia comprende de 60-150 pares de bases de SEQ ID N° 1. La al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr11-114010-126106, chr19-183944-196032, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr5-180756063-180768074, chr6-170921836-170922549, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas, y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de al menos una secuencia de referencia;

un detector configurado para proporcionar una primera indicación en relación con una cantidad de la secuencia de interés amplificada y una segunda indicación en relación con una cantidad de la secuencia de referencia amplificada; y

un controlador configurado para cuantificar la secuencia de interés amplificada en relación con la secuencia de referencia amplificada, en función, al menos en parte, de la primera indicación y la segunda indicación; y determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa.

Los expertos en la técnica conocen los amplificadores de ácido nucleico. En general, los amplificadores de ácido nucleico usan uno o más cebadores y uno o más agentes químicos o enzimáticos para copiar una secuencia de ácido nucleico molde, como una secuencia de interés o una secuencia de referencia. Dicha copia se puede hacer en múltiples ciclos para generar cantidades relativamente grandes de la secuencia de interés y de la secuencia de referencia. De manera deseable, el amplificador de ácido nucleico se configura para amplificar la secuencia de interés y la secuencia de referencia de manera simultánea y en paralelo, p. ej., añadiendo diferentes conjuntos de cebadores, uno específico a la secuencia de interés y el otro específico a la secuencia de referencia, a diferentes soluciones de reacción idénticas, como en diferentes pocillos de una placa multi-pocillo.

El sistema también comprende un reservorio de reactivo. En general, el reservorio de reactivo contiene materiales necesarios para que se produzca la reacción de amplificación, como cebadores, agentes químicos o enzimáticos, nucleótidos libres que se incorporen en copias de secuencias molde, etc. El reservorio de reactivo también puede contener una o más sondas u otros compuestos que comprenden restos detectables. Cualquiera de estos materiales se puede almacenar por separado y/o se pueden combinar dos o más de estos para su almacenamiento en el

reservorio de reactivo. Estos materiales están generalmente en solución acuosa y se pueden introducir en la(s) solución(es) de reacción mediante técnicas conocidas por la persona experta en la técnica. Dicha introducción puede darse una o varias veces antes, durante o después de un proceso de amplificación. Por ejemplo, se pueden añadir algún(os) reactivo(s) una vez por ciclo de amplificación.

- 5 En una realización, el reservorio de reactivo que contiene al menos un primer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de referencia, en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-31, como las SEQ ID N°: 1-13, y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de referencia.
- 10 La secuencia de referencia, la determinación de un porcentaje de identidad de secuencia y las SEQ ID Nos: 1-38 se describen en otra parte en esta memoria.

15 El sistema también comprende un detector. En general, el detector se puede configurar para detectar una sonda para la secuencia de interés, la secuencia de referencia, o ambas. Tras la detección, el detector puede realizar varios procesamientos de señal y/o operaciones de análisis para proporcionar una primera indicación en relación con una cantidad de la secuencia amplificada de interés y una segunda indicación en relación con una cantidad de la secuencia de referencia amplificada.

20 El sistema también comprende un controlador. El controlador se puede configurar para cuantificar la secuencia de interés amplificada en relación con la secuencia de referencia amplificada, en función, al menos en parte, de la primera indicación y de la segunda indicación; y determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa. Puede almacenar el número de copias determinado en una memoria, mostrarlo a un usuario, escribirlo en un archivo de ordenador, o similar.

En una realización adicional, el controlador se puede configurar para diagnosticar que el sujeto tiene un biomarcador relacionado con el cáncer, en función de que la secuencia de interés está asociada con el cáncer y que el número de copias es indicativo del cáncer.

25 Aunque el amplificador de ácido nucleico, el reservorio de reactivo, el detector y el controlador se han descrito por separado, pueden ser componentes de un solo aparato cualquiera de dos o más de estos componentes.

La descripción proporciona:

30 1. Un método que comprende: amplificar una secuencia de ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto; amplificar una secuencia de referencia en la muestra, en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases, en donde al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr10-38741930-38753964, chr11-114010-126106, chr16-90239446-90251554, chr19-183944-196032, chr2-114323560-114323652, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr4-165196360-165199636, chr5-180756063-180768074, chr6-170921836-170922549, chr7-39837560-63231088, chr7-128296352-128298474, chr8-143133-150475, chr9-49679-49771, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas; cuantificar la secuencia amplificada de interés en relación con la secuencia de referencia amplificada y determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa.

40 2. En el método, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-31.

45 3. En el método, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-13.

4. En el método, la muestra puede incluir tejido sospechoso de ser tejido canceroso.

5. En el método, la muestra se ha sometido a fijación con formalina e inclusión en parafina (FFPE) antes de amplificar la secuencia de interés y de amplificar la secuencia de referencia.

50 6. El método puede también incluir: diagnosticar al sujeto como si tuviera un biomarcador relacionado con el cáncer, en función de la secuencia de interés que está asociada con el cáncer y el número de copias que es indicativo del cáncer.

7. En el método, la amplificación de la secuencia de interés y la amplificación de la secuencia de referencia se pueden realizar mediante la reacción en cadena de la polimerasa cuantitativa TaqMan (qPCR).

8. Un kit que comprende una primera sonda que hibrida específicamente con al menos una porción de al menos una

- 5 10 secuencia de referencia que tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases, en donde la al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr10-38741930-38753964, chr11-114010-126106, chr16-90239446-90251554, chr19-183944-196032, chr2-114323560-114323652, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr4-165196360-165199636, chr5-180756063-180768074, chr6-170921836-170922549, chr7-39837560-63231088, chr7-128296352-128298474, chr8-143133-150475, chr9-49679-49771, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas.
9. En el kit, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-31.
10. En el kit, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-13.
- 15 11. En el kit, la primera sonda puede incluir una secuencia de ácido nucleico configurada para hibridar específicamente con al menos la porción de la al menos una secuencia de referencia, un indicador fluorescente en un primer extremo de la secuencia de ácido nucleico y un quencher fluorescente en un segundo extremo de la secuencia de ácido nucleico.
- 20 12. El kit puede además incluir: un primer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de referencia, y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de referencia.
13. El kit puede además incluir: una segunda sonda que hibrida específicamente con al menos una porción de al menos una secuencia de ácido nucleico de interés.
- 25 14. El kit puede además incluir: un tercer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de ácido nucleico de interés, y un cuarto cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de ácido nucleico de interés.
- 30 15. Una composición que incluye: una primera sonda que hibrida específicamente con al menos una porción de al menos una secuencia de referencia que tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases, en donde la al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr10-38741930-38753964, chr11-114010-126106, chr16-90239446-90251554, chr19-183944-196032, chr2-114323560-114323652, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr4-165196360-165199636, chr5-180756063-180768074, chr6-170921836-170922549, chr7-39837560-63231088, chr7-128296352-128298474, chr8-143133-150475, chr9-49679-49771, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas.
- 35 16. En la composición, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-31.
- 40 17. En la composición, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-13.
- 45 18. En la composición, la primera sonda puede incluir una secuencia de ácido nucleico configurada para hibridar específicamente con al menos la porción de la al menos una secuencia de referencia, un indicador fluorescente en un primer extremo de la secuencia de ácido nucleico y un quencher fluorescente en un segundo extremo de la secuencia de ácido nucleico.
19. La composición puede además incluir: un primer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de referencia, y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de referencia.
- 50 20. La composición puede además incluir: una segunda sonda que hibrida específicamente con al menos una porción de al menos una secuencia de ácido nucleico de interés.
- 55 21. La composición puede además incluir: un tercer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de ácido nucleico de interés, y un cuarto cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de ácido nucleico de interés.

22. Un sistema que incluye: un amplificador de ácido nucleico configurado para amplificar una secuencia de ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto y amplificar una secuencia de referencia en la muestra, un reservorio de reactivo que contiene al menos un primer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de referencia, en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases, en donde la al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr10-38741930-38753964, chr11-114010-126106, chr16-90239446-90251554, chr19-183944-196032, chr2-114323560-114323652, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr4-165196360-165199636, chr5-180756063-180768074, chr6-170921836-170922549, chr7-39837560-63231088, chr7-128296352-128298474, chr8-143133-150475, chr9-49679-49771, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas; y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de referencia; un detector configurado para proporcionar una primera indicación en relación con una cantidad de la secuencia de interés amplificada y una segunda indicación en relación con una cantidad de la secuencia de referencia amplificada; y un controlador configurado para cuantificar la secuencia de interés amplificada en relación con la secuencia de referencia amplificada, en función, al menos en parte, de la primera indicación y de la segunda indicación; y determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa.
23. En el sistema, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-31.
24. En el sistema, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-13.
25. En el sistema, la muestra puede incluir tejido sospechoso de ser tejido canceroso.
26. En el sistema, la muestra se ha sometido a fijación con formalina e inclusión en parafina (FFPE).
27. En el sistema, el controlador está además configurado para indicar que el sujeto tiene un biomarcador relacionado con el cáncer en función de la secuencia de interés que está asociada con el cáncer y el número de copias que es indicativo del cáncer.
28. En el sistema, el amplificador de ácido nucleico se configura para amplificar la secuencia de interés y amplificar la secuencia de referencia mediante la reacción en cadena de la polimerasa cuantitativa TaqMan (qPCR).
29. Un método que incluye cuantificar una secuencia de ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto en relación con una secuencia de referencia en la muestra, en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases, en donde la al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr10-38741930-38753964, chr11-114010-126106, chr16-90239446-90251554, chr19-183944-196032, chr2-114323560-114323652, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr4-165196360-165199636, chr5-180756063-180768074, chr6-170921836-170922549, chr7-39837560-63231088, chr7-128296352-128298474, chr8-143133-150475, chr9-49679-49771, chrY-26424506-27537936 o chr6-132951-145064; la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas; y determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa.
30. En el método, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-38.
31. En el método, la secuencia de referencia puede tener al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-13.
32. En el método, la muestra puede incluir tejido sospechoso de ser tejido canceroso.
33. En el método, la muestra se ha sometido a fijación con formalina e inclusión en parafina (FFPE) antes de amplificar la secuencia de interés y de amplificar la secuencia de referencia.
34. El método puede también incluir: diagnosticar al sujeto como si tuviera un biomarcador relacionado con el cáncer, en función de la secuencia de interés que está asociada con el cáncer y el número de copias que es indicativo del cáncer.

Los siguientes ejemplos se incluyen para demostrar realizaciones particulares de la descripción. Los expertos en la técnica deberían apreciar que las técnicas descritas en los ejemplos que siguen representan técnicas descubiertas por el inventor que funcionan bien en la práctica de la descripción y, por lo tanto, se puede considerar que constituyen modos particulares para su práctica. Sin embargo, los expertos en la técnica deberían, a la luz de la presente descripción, apreciar que se pueden realizar muchos cambios en las realizaciones específicas que se describen y aún se obtienen resultados parecidos o similares sin apartarse del espíritu y alcance de la descripción.

Ejemplo 1

Las SEQ ID Nos: 1-13 se identificaron mediante un proceso de dos etapas. Primero, se usó un algoritmo bioinformático para identificar objetivos candidatos en el genoma que cumplían ciertos criterios asociados con un número de copias sustancialmente normal, incluso en células cancerosas sometidas a FFPE. En general, se usaron secuencias sospechosas de ser relativamente resistentes a anomalías en el número de copias en células cancerosas y/o células sometidas a FFPE para consultar bases de datos genómicos humanos públicamente disponibles, y se consideraron como objetivos candidatos solo aquellas secuencias que dieron como respuesta múltiples resultados para pruebas adicionales.

La primera etapa identificó las SEQ ID Nos: 1-31. Las SEQ ID Nos: 2-8 y 10-13 se ubican en el genoma en los loci especificados *supra*. Las SEQ ID Nos: 14-31 se ubican en el genoma al menos en los siguientes loci:

SEQ ID N°: 14, chr10 38753941..38753964..38754039 chr11.11250s.

SEQ ID N°: 15, chr10 38753848..38753866..38753912 chr11.11350s.

SEQ ID N°: 16, chr10 38753856..38753896..38753934 chr1.5850s.

SEQ ID N°: 17, chr10 38753715..38753794..38753862 chr1.5950s.

SEQ ID N°: 18, chr10 38753145..38753206..38753248 chr16.2650s.

SEQ ID N°: 19, chr10 38753377..38753408..38753457 chr16.2850s.

SEQ ID N°: 20, chr10 38753232..38753257..38753351 chr20.2650s.

SEQ ID N°: 21, chr10 38753326..38753360..38753411 chr20.2750s.

SEQ ID N°: 22, chr10 38753413..38753454..38753513 chr20.2850s.

SEQ ID N°: 23, chr111123673..123696..123739 chr20.3850s.

SEQ ID N°: 24, chr10 38750797..38750847..38750889 chr3.150s.

SEQ ID N°: 25, chr10 38741930..38741975..38742018 chr3.1550s.

SEQ ID N°: 26, chr10 38742246..38742274..38742312 chr3.1850s.

SEQ ID N°: 27, chr10 38746651..38746676..38746733 chr3.6250s.

SEQ ID N°: 28, chr10 38753090..38753126..38753234 chr5. 2550s.

SEQ ID N°: 29, chr10 38753486..38753535..38753587 chr5.2950s.

SEQ ID N°: 30, chr10 38753797..38753835..38753885 chr5.3250s.

SEQ ID N°: 31, chr10 38753907..38753935..38753971 chr5.3350s.

Las veintitrés regiones más grandes en las que mapearon las secuencias candidatas fueron las siguientes: Región	Cromosoma	Posición inicial	Posición final
1	chr1	121790	133586
2	chr1	329448	341534
3	chr1	648129	660266
4	chr1	222643865	228172047
5	chr1	243203764	243215874
6	chr10	38741930	38753964

7	chr11	114010	126106
8	chr16	90239446	90251554
9	chr19	183944	196032
10	chr2	114323560	114323652
11	chr2	243064480	243071940
12	chr20	62921559	62933673
13	chr3	197950387	197962431
14	chr4	119557144	120325498
15	chr4	165196360	165199636
16	chr5	180756063	180768074
17	chr6	170921836	170922549
18	chr7	39837560	63231088
19	chr7	128296352	128298474
20	chr8	143133	150475
21	chr9	49679	49771
22	chrY	26424506	27537936
23	chr6	132951	145064

5 (Como es consciente el experto en la técnica, en el momento de esta publicación, no se puede predecir la resistencia de una secuencia a anomalías en el número de copias en células cancerosas y/o células sometidas a FFPE con la suficiente precisión solo a partir de los datos de secuencia. Se requieren pruebas adicionales, como las que se describen a continuación, para identificar las secuencias adecuadas para usar como ensayo de referencia multicopia).

Se diseñaron numerosos ensayos de referencia de qPCR, TaqMan usando los objetivos candidatos y se probaron en muestras de cáncer junto con un ensayo de qPCR, TaqMan de objetivo de interés (TOI). Los ensayos mapearon en las diversas regiones objetivo:

ld.	número de resultados	número de cromosomas
chr1 1.11250s.1	16	11
chr1 1.11350s.1	17	12
chr1.5850s.1	15	11
chr1.5950s.1	15	11
chr16.2650s.1	15	11
chr16.2850s.1	15	11
chr16.3250s.1	16	11
chr16.3350s.1	16	11
chr20.2650s.1	16	11
chr20.2750s.1	16	11
chr20.2850s.1	16	11
chr20.3250s.1	17	12
chr20.3350s.1	16	11
chr20.3850s.1	17	11
chr2.3250s.1	15	11
chr3.150s.1	15	12

chr3.1550s.1	16	11
chr3.1850s.1	23	11
chr3.6250s.1	18	13
chr5.2550s.1	17	12
chr5.2850s.1	15	11
chr5.2950s.1	19	13
chr5.3250s.1	17	12
chr5.3350s.1	16	11

Después de la prueba inicial, dos objetivos proporcionaron resultados sugestivos de un número de copias sustancialmente normal, incluso en células cancerosas sometidas a FFPE y se seleccionaron para una ronda final de pruebas en un panel de muestras ampliado.

- 5 La ronda final de pruebas evaluó el rendimiento de estos dos ensayos multicopia de qPCR de novo, un ensayo de referencia de copia única convencional de qPCR (RNaseP) y cinco ensayos de qPCR TOI para el gen IRS2. Los ensayos de qPCR usaron todos un ADN genómico molde extraído de 35 muestras de tejido colorectal normal sometidas a FFPE. Todos los ensayos se realizaron por duplicado, con el ensayo de referencia y el ensayo del objetivo de interés ejecutados en pocillos separados para generar valores Cq precisos.
- 10 El primer ensayo multicopia de referencia, que corresponde a las SEQ ID Nos: 1-8, usó un primer conjunto que comprendía cebador directo, cebador inverso y secuencia sonda. El segundo ensayo multicopia de referencia, que corresponde a las SEQ ID Nos: 9-13, usó un segundo conjunto que comprendía cebador directo, cebador inverso y secuencia sonda.
- 15 La expectativa era que se debería determinar un resultado de 2 copias para el gen IRS2 para muestras normales, si una prueba generaba resultados precisos. En las Figuras 1A-3B se proporciona una instantánea resumen de los resultados de cada uno de los tres ensayos de referencia: RNaseP (Figuras 1A-1B), SEQ ID Nos: 1-8 (Figuras 2A-2B) y SEQ ID Nos: 9-13 (Figuras 3A-3B). En cada figura, se representa el número de copias en el eje y y en el eje x se representan las 35 muestras en veces duplicadas de los cinco ensayos diferentes para el TOI IRS2. Cada figura contiene 350 puntos de datos.
- 20 Los resultados dependieron no solo del ensayo de referencia usado sino también del ensayo de TOI IRS2 usado. Los dos ensayos de referencia multicopia funcionaron bien, junto con los primeros 3 ensayos de TOI IRS2. En cada figura, la subfigura A tiene un número de copias calculado con punto decimal y la subfigura B tiene un número de copias redondeado.

Listado de secuencias

- 25 <110> Life Technologies Corporation
MERRILL, David C.
BRZOSKA, Pius M.
LI, Zheng
LIN, Wendy
- 30 LEE, Wing
WONG, Mandi
- <120> Ensayo de referencia multi-copia
- <130> LT00901 PCT
- <150> 61/968,609
- 35 <151> 21-03-2014
- <160> 31
- <170> FastSEQ for Windows Versión 4.0
- <210> 1
- <211> 70
- 40 <212> ADN
- <213> Homo sapiens

ES 2 719 122 T3

<220>
 <221> característica_misc
 <222> 70
 <223> n = A,T,C or G

5 <400> 1
 ggctgtytgc rgtagtwrts trkswrsmrs mmmrwsrmyg smsrcarrsr armarwyws 60
 twdvwakkmn 70

<210> 2
 <211> 70
 <212> ADN
 10 <213> Homo sapiens

<400> 2
 ggctgcttgc agtagttgtg tagcagcagc acaatggccg cagacaagga aaacagtttc 60
 taggaattcc 70

<210> 3
 <211> 70
 15 <212> ADN
 <213> Homo sapiens

<400> 3
 ggctgcttgc ggtagttatg tagcagcagc acaatggccg cagacaagga aaacagtttc 60
 taggaattcc 70

<210> 4
 20 <211> 70
 <212> ADN
 <213> Homo sapiens

<400> 4
 ggctgcttgc ggtagttgtc tagcagcagc acaatggccg cagacaagga aaacagtttc 60
 taggaattcc 70

25 <210> 5
 <211> 70
 <212> ADN
 <213> Homo sapiens

<400> 5
 30 ggctgcttgc ggtagttgtg tagcagcagc acaatggccg cagacaagga aaacagtttc 60
 taaaaattcc 70

<210> 6
 <211> 70
 <212> ADN
 <213> Homo sapiens

35 <400> 6
 ggctgcttgc ggtagttgtg tagcagcagc acaatggccg cagacaagga aaacagtttc 60
 taggaattcc 70

<210> 7
 <211> 70
 <212> ADN
 40 <213> Homo Sapiens

<220>
 <221> característica_misc
 <222> 70
 <223> n = A,T,C or G

45 <400> 7
 ggctgcttgc ggtagttgtg tagcagcagc acaatggccg cagacaagga aaacagtttc 60
 taggaattcn 70

<210> 8

ES 2 719 122 T3

<211> 70
 <212> ADN
 <213> Homo Sapiens

 <400> 8
 5 ggctgtttgc ggtagtagtc tgtgtagcag cagcacaatg gccgcagacg aggaaaacag 60
 tttctaggaa 70

 <210> 9
 <211> 67
 <212> ADN
 <213> Homo sapiens

 10 <400> 9
 agtgcagyrw tgytgactct tccaagctta acattttctca saartcaatt agctttgtac 60
 tgggagg 67

 <210> 10
 <211> 67
 <212> ADN
 15 <213> Homo sapiens

 <400> 10
 agtgcagcaa tgytgactct tccaagctta acattttctca gaagtcaatt agctttgtac 60
 tgggagg 67

 <210> 11
 <211> 67
 20 <212> ADN
 <213> Homo sapiens

 <400> 11
 agtgcagcga tgytgactct tccaagctta acattttctca caagtcaatt agctttgtac 60
 tgggagg 67

 <210> 12
 25 <211> 67
 <212> ADN
 <213> Homo sapiens

 <400> 12
 agtgcagcgt tgctgactct tccaagctta acattttctca caaatcaatt agctttgtac 60
 tgggagg 67

 30 <210> 13
 <211> 67
 <212> ADN
 <213> Homo sapiens

 <400> 13
 35 agtgcagtga tgctgactct tccaagctta acattttctca caagtcaatt agctttgtac 60
 tgggagg 67

 <210> 14
 <211> 99
 <212> ADN
 <213> Homo sapiens

 40 <400> 14
 gtgtagcagc agcacaatgg ccgcagacaa ggaaaacagt ttctaggaat tcctcgtata 60
 taattttata tttttgacaa gattaatgac ccatgctcc 99

 <210> 15
 <211> 65
 <212> ADN
 45 <213> Homo sapiens

 <400> 15

ES 2 719 122 T3

tgcarmgatg ctgactcttc caagcttaac atttctcaca agtcaattag ctttgtactg 60
 ggagg 65
 <210> 16
 <211> 79
 <212> ADN
 5 <213> Homo sapiens
 <400> 16
 tgctgactct tccaagctta acatttctca caagtcaatt agctttgtac tgggaggagg 60
 gcgtgaaggg ctgcttgcg 79
 <210> 17
 <211> 148
 10 <212> ADN
 <213> Homo sapiens
 <400> 17
 caagggacaa ggaaaaatta tccaaacatt gtttaaaaca atcatcatta attagtaaca 60
 cttatccagg ggggttttta acctttcccc cactcaasga ttattctaata gtcagagtag 120
 aataaaaaat aagtgcarmg atgctgac 148
 <210> 18
 15 <211> 104
 <212> ADN
 <213> Homo sapiens
 <400> 18
 ggaggaggaa aataggtagt ttttcaaaag ttttcaaaaa tatgaaaaga agaaatgaaa 60
 tggtagcttg aagagattgt tgaatggga gagactatgg tggc 104
 20 <210> 19
 <211> 81
 <212> ADN
 <213> Homo sapiens
 <400> 19
 caactaaaag gcaatgtcac tccaataatc accagagtaa tcaatttgct tattgctgct 60
 25 cctttaaata tagttctctg g 81
 <210> 20
 <211> 120
 <212> ADN
 <213> Homo sapiens
 30 <400> 20
 ggagagacta tggtagcttg tttagaagca gttgagatag atccaattga gatagagata 60
 ttgagtatat aaacaaaaga atgacaaatt aatagtgtaa tggataactt gactttggca 120
 <210> 21
 <211> 86
 <212> ADN
 35 <213> Homo sapiens
 <400> 21
 gtgtaatgga taacttgact ttggcaaata ttgtgaattt ttgtgaaagt acaactaaaa 60
 ggcaatgtca ctccaataat caccag 86
 <210> 22
 <211> 101
 40 <212> ADN
 <213> Homo sapiens
 <400> 22
 gtaatcaatt tgcttattgc tgtcccttta aatatagttc tctggtatca actaacatgt 60
 ttttaactaa tgatgcttct taaagaaaag ggaaaagacc t 101
 <210> 23
 45 <211> 67

ES 2 719 122 T3

<212> ADN
 <213> Homo sapiens

<400> 23
 ccctgggccc ctcaggggag tccctgctgg acagtgagac agagaatgac catgatgatg 60
 ctttcct 67

5 <210> 24
 <211> 93
 <212> ADN
 <213> Homo sapiens

<400> 24
 gggtttatgt ttgatatrta atgtaatfff ctaatgctaa atcaagtggc aattttgtta 60
 10 gtcaagttga tttagtggct tgggaagaaa gct 93

<210> 25
 <211> 89
 <212> ADN
 <213> Homo sapiens

15 <400> 25
 gagaccccc ggtggtgagg cagggctggg gtgtcccctt ccaaccaggc tgtcaaggcc 60
 ccaactctgg ggcagaggca gtggcaggg 89

<210> 26
 <211> 67
 <212> ADN
 20 <213> Homo sapiens

<400> 26
 catccgtttc acctgcagtt gaagatccgt gagtgccca gaagatcatg cagtcawcag 60
 tcccacg 67

<210> 27
 <211> 83
 25 <212> ADN
 <213> Homo sapiens

<400> 27
 gakataagga agctcgagga agagaaaaaa caactggaag gagaaatcat agatttttat 60
 aaaatgamag ctgcctctga agc 83

<210> 28
 30 <211> 145
 <212> ADN
 <213> Homo sapiens

<400> 28
 ccgttttggg ggaggaacag attccatgtc cactagaatg gaatgaacaa gaaatggagg 60
 aggaaaatag gtagtffffc aaaagffffc aaaaatatga aaagaagaaa tgaaatggta 120
 cttggaagag attggtgaaa tggga 145

35 <210> 29
 <211> 102
 <212> ADN
 <213> Homo sapiens

<400> 29
 tgcttcttaa agaaaaggga aaagacffff ttctttcttt cagtcttcaa tgattcactg 60
 40 cttcatctcg ctccaccaa gataaatgaa atctacatct ct 102

<210> 30
 <211> 89
 <212> ADN
 <213> Homo sapiens

45 <400> 30

ES 2 719 122 T3

```
ctttcccca ctcaasgatt attctaattgt cagagtagaa taaaaataa gtgcarmgat 60  
gctgactott ccaagcttaa catttctca 89
```

<210> 31

<211> 65

<212> ADN

5 <213> Homo sapiens

<400> 31

```
gggaggaggg cgtgaagggc tgcttgccgt agttgtgtag cagcagcaca atggccgcag 60  
acaag 65
```

REIVINDICACIONES

1. Un método que comprende:
 amplificar una secuencia de ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto;
 amplificar una secuencia de referencia en la muestra en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases de cualquiera de las SEQ ID Nos: 2-8, o en donde la secuencia de referencia comprende de 60-150 pares de bases de la SEQ ID N° 1, en donde la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas;
- 5
- 10
2. El método de la reivindicación 1 en donde la al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr11-114010-126106, chr19-183944-196032, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr5-180756063-180768074, chr6-170921836-170922549, chrY-26424506-27537936 o chr6-132951-145064;
- 15
3. El método de la reivindicación 1 o 2 en donde la muestra comprende tejido sospechoso de ser tejido canceroso.
- 20
4. El método de la reivindicación 3 en donde la muestra se ha sometido a fijación con formalina e inclusión en parafina (FFPE) antes de amplificar la secuencia de interés y de amplificar la secuencia de referencia.
5. El método de la reivindicación 3 que además comprende:
 diagnosticar al sujeto como que tiene un biomarcador relacionado con el cáncer en función de la secuencia de interés que está asociada con el cáncer y el número de copias que es indicativo del cáncer.
- 25
6. Una composición que comprende:
 una primera sonda que hibrida específicamente con al menos una porción de al menos una secuencia de referencia que tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases de cualquiera de las SEQ ID Nos: 2-8, o en donde la secuencia de referencia comprende de 60-150 pares de bases de la SEQ ID N° 1, en donde la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas, en donde la sonda es una sonda TaqMan.
- 30
7. La composición de la reivindicación 6 en donde la al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr11-114010-126106, chr19-183944-196032, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr5-180756063-180768074, chr6-170921836-170922549, chrY-26424506-27537936 o chr6-132951-145064;
- 35
8. La composición de la reivindicación 6 o 7 en donde la primera sonda comprende una secuencia de ácido nucleico configurada para hibridar específicamente con al menos la porción de la al menos una secuencia de referencia, un indicador fluorescente en el primer extremo de la secuencia de ácido nucleico y un quencher fluorescente en un segundo extremo de la secuencia de ácido nucleico.
- 40
9. La composición de la reivindicación 6 o 7 que además comprende:
 un primer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de referencia, y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de referencia, y opcionalmente una segunda sonda que hibrida específicamente con al menos una porción de al menos una secuencia de ácido nucleico de interés.
- 45
10. La composición de la reivindicación 9 que además comprende:
 un tercer cebador configurado para hibridar específicamente con un primer extremo de al menos una secuencia de ácido nucleico de interés, y
 un cuarto cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo
- 50

extremo de la al menos una secuencia de ácido nucleico de interés.

11. Un sistema que comprende:

un amplificador de ácido nucleico configurado para amplificar una secuencia de ácido nucleico de interés en una muestra que comprende ADN genómico de un sujeto y amplificar una secuencia de referencia en la muestra,

5 un reservorio de reactivo que contiene al menos un primer cebador configurado para hibridar específicamente con un primer extremo de la al menos una secuencia de referencia, en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una porción de ADN genómico que comprende de aproximadamente 60 a aproximadamente 150 pares de bases de cualquiera de las SEQ ID Nos: 2-8, o en donde la secuencia de referencia comprende de 60-150 pares de bases de SEQ ID N° 1, la al menos una porción está presente en al menos un primer número mínimo de copias en el genoma; y al menos una copia de la al menos una porción está presente en cada uno de al menos un segundo número mínimo de cromosomas;

y un segundo cebador configurado para hibridar específicamente con una secuencia complementaria a un segundo extremo de la al menos una secuencia de referencia;

15 un detector configurado para proporcionar una primera indicación en relación con una cantidad de la secuencia de interés amplificada y una segunda indicación en relación con una cantidad de la secuencia de referencia amplificada; y

un controlador configurado para cuantificar la secuencia de interés amplificada en relación con la secuencia de referencia amplificada, en función, al menos en parte, de la primera indicación y de la segunda indicación; y determinar un número de copias de la secuencia de interés a partir de la secuencia de interés amplificada cuantificada relativa

20 en donde el amplificador de ácido nucleico se configura para amplificar la secuencia de interés y amplificar la secuencia de referencia mediante la reacción en cadena de la polimerasa cuantitativa TaqMan (qPCR).

12. El sistema de la reivindicación 11 en donde al menos una porción está presente en chr1-121790-133586, chr1-329448-341534, chr1-648129-660266, chr1-222643865-228172047, chr1-243203764-243215874, chr11-114010-126106, chr19-183944-196032, chr2-243064480-243071940, chr20-62921559-62933673, chr3-197950387-197962431, chr4-119557144-120325498, chr5-180756063-180768074, chr6-170921836-170922549, chrY-26424506-27537936 o chr6-132951-145064.

13. El método de la reivindicación 1 o 2, la composición de la reivindicación 6 o 7 o el sistema de la reivindicación 11 o 12, en donde la secuencia de referencia tiene al menos un 80% de identidad de secuencia con al menos una de las SEQ ID Nos: 1-31 o las SEQ ID Nos: 1-13.

30 14. El sistema de la reivindicación 13 en donde la muestra comprende tejido sospechoso de ser tejido canceroso.

15. El sistema de la reivindicación 14 en donde la muestra se ha sometido a fijación con formalina e inclusión en parafina (FFPE).

35 16. El sistema de la reivindicación 14 en donde el controlador se configura además para indicar que el sujeto tiene un biomarcador relacionado con el cáncer, en función de que la secuencia de interés está asociada con el cáncer y que el número de copias es indicativo del cáncer.

FIGURA 1A

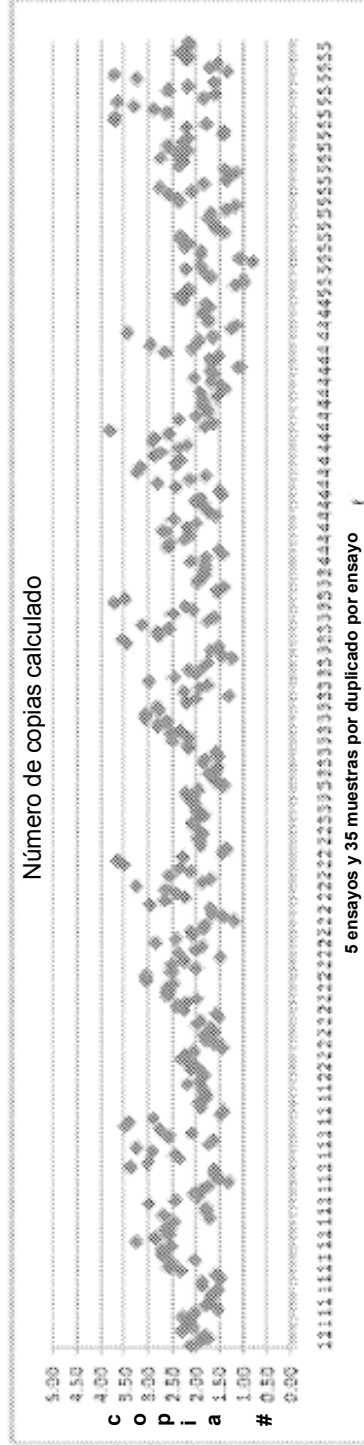


FIGURA 1B

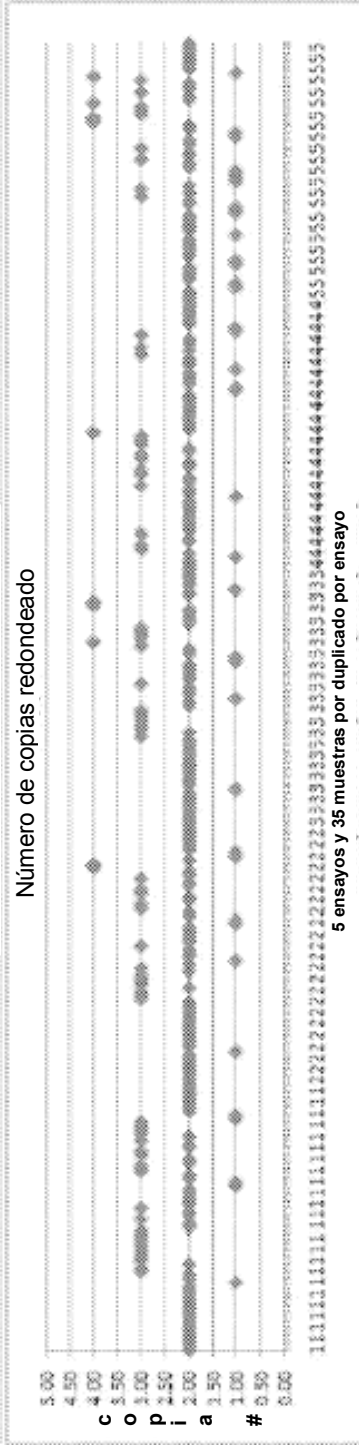


FIGURA 2A

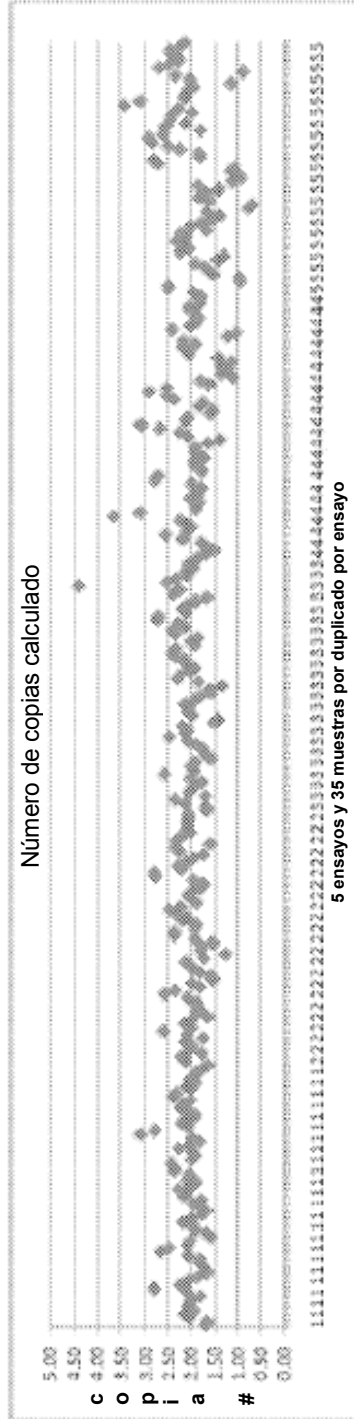


FIGURA 2B

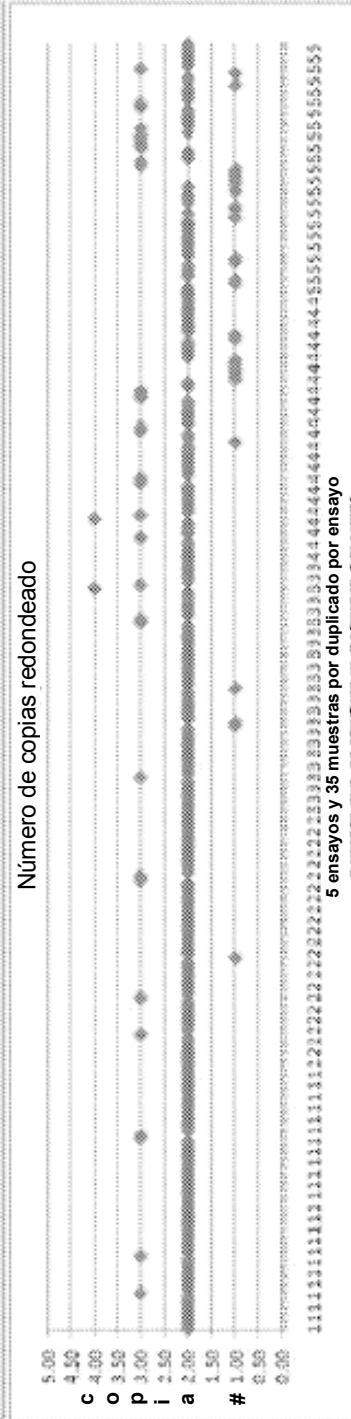


FIGURA 3A

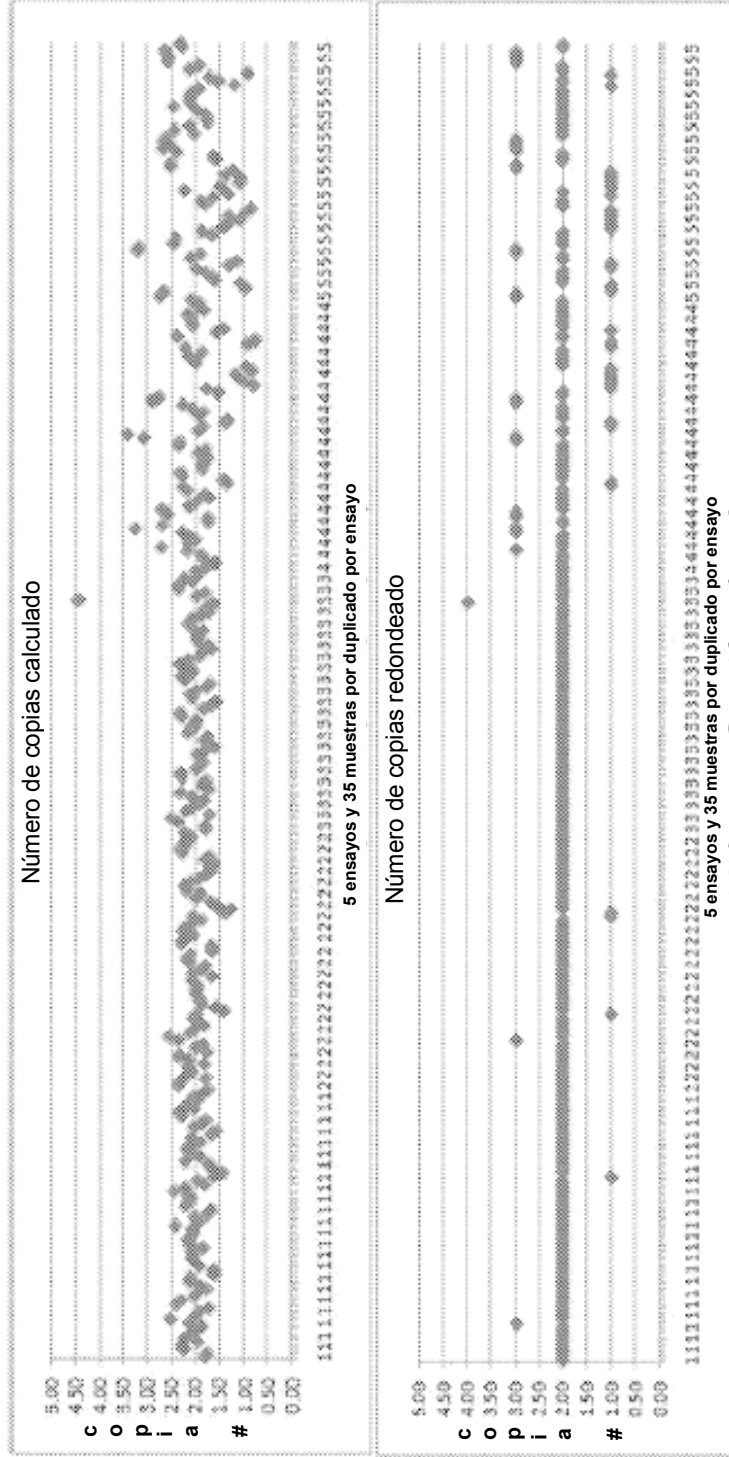


FIGURA 3B