

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 763 931**

51 Int. Cl.:

C12Q 1/68 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.05.2013 PCT/US2013/042157**

87 Fecha y número de publicación internacional: **28.11.2013 WO13177245**

96 Fecha de presentación y número de la solicitud europea: **22.05.2013 E 13794236 (3)**

97 Fecha y número de publicación de la concesión europea: **11.12.2019 EP 2852689**

54 Título: **Genes Nano46 y métodos para predecir el resultado del cáncer de mama**

30 Prioridad:

22.05.2012 US 201261650209 P
17.01.2013 US 201361753673 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
01.06.2020

73 Titular/es:

NANOSTRING TECHNOLOGIES, INC. (20.0%)
530 Fairview Avenue North, Suite 2000
Seattle, WA 98109, US;
THE UNIVERSITY OF NORTH CAROLINA AT
CHAPEL HILL (20.0%);
UNIVERSITY OF UTAH RESEARCH FOUNDATION
(20.0%);
WASHINGTON UNIVERSITY (20.0%) y
BRITISH COLUMBIA CANCER AGENCY BRANCH
(20.0%)

72 Inventor/es:

FERREE, SEAN M.;
STORHOFF, JAMES JUSTIN;
PARKER, JOEL S.;
PEROU, CHARLES M.;
ELLIS, MATTHEW J.;
BERNARD, PHILIP S. y
NIELSEN, TORSTEN O.

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 763 931 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

Genes Nano46 y métodos para predecir el resultado del cáncer de mama

5 Campo de la invención

Esta descripción se refiere generalmente al campo de la biología del cáncer, y específicamente, a los campos de detección e identificación de fenotipos específicos de células cancerosas y la correlación con las terapias apropiadas.

10 Antecedentes de la invención

Los enfoques actuales para tratar temprano el cáncer de mama, que incluye la terapia adyuvante, han mejorado de hecho la supervivencia y reducido la recurrencia. Sin embargo, el riesgo de recurrencia puede estar subestimado en algunos pacientes, pero sobreestimado en otros.

15 Si bien el riesgo de recurrencia disminuye algo con el tiempo, en muchos estudios se ha observado un riesgo continuo, algunos de los cuales involucran a decenas de miles de pacientes con cáncer de mama. De hecho, algunos de los pacientes que experimentaron recurrencia después de cinco años en estos estudios habían sido considerados previamente como de "bajo riesgo", por ejemplo, su cáncer no se había diseminado a los ganglios linfáticos en el momento del diagnóstico inicial, o el estado de su receptor de estrógeno fue positivo. En uno de estos estudios, un número sustancial de recurrencias se produjeron más de cinco años después del tratamiento. Así, existe la necesidad en la técnica de determinar el riesgo de recurrencia y determinar las terapias que reducen ese riesgo y mejoran la supervivencia general.

20 El documento de patente núm. WO2009/158143 y Parker Joel S y otros: "Supervised risk predictor of breast cancer based on intrinsic subtypes", J. of Clinical Oncology, America Soc. of Clinical Oncology, 27: 8, págs. 1160-1167, 10 de marzo de 2019 describe los métodos para clasificar y evaluar el pronóstico de un sujeto que tiene cáncer de mama. El modelo de predicción se basa en el perfil de expresión de genes de un conjunto específico de genes. El documento de patente núm. EP1641810 A2 proporciona conjuntos de genes, cuya expresión es importante en el diagnóstico y/o pronóstico del cáncer, en particular del cáncer de mama. "Affymetrix Genechip Human Genome U133 Array Set HG-U133A" GEO proporciona sondas para todos los genes de codificación de la célula humana en una micromatriz.

25 Resumen de la invención

30 La presente invención proporciona un método para determinar un riesgo de recurrencia bajo o alto (ROR) en un sujeto que tiene cáncer de mama que comprende: determinar el tamaño del tumor de cáncer de mama del sujeto; determinar un valor de correlación de un tumor de cáncer de mama para cada uno de al menos cuatro subtipos intrínsecos, que incluye un tipo Basal, Luminal A, Luminal B, o enriquecido con HER-2, midiendo la expresión de ARN de al menos 40 de los genes en la lista de genes intrínsecos NANO46 de la Tabla 1; determinar una puntuación de la proliferación midiendo la expresión de ARN de cada uno de los subconjuntos de 18 genes de los genes intrínsecos NANO46 seleccionados de ANLN, CCNE1, CDC20, CDC6, CDCA1, CENPF, CEP55, EXO1, KIF2C, KNTC2, MELK, MKI67, ORC6L, PTTG1, RRM2, TYMS, UBE2C, y UBE2T; calcular una puntuación de riesgo de recurrencia (ROR) con el uso de la siguiente ecuación: $ROR-PT = -0,0067 \cdot Basal + 0,4317 \cdot Her2 + -0,3172 \cdot LumA + 0,4894 \cdot LumB + 0,1981 \cdot Puntuación\ de\ la\ Proliferación + 0,1133 \cdot Tamaño\ de\ tumor$; en donde la puntuación de riesgo de recurrencia es un valor entero en una escala de 0-100, en donde el nivel de expresión de ARN para al menos 40 de los genes en la lista de genes intrínsecos NANO46 de la Tabla 1 se establece en la Tabla 3; determinando de ese modo si el sujeto tiene un riesgo de recurrencia bajo o alto en función de la puntuación de riesgo de recurrencia (ROR) y la interpretación de la puntuación ROR con el uso de la clasificación de riesgo por intervalo de riesgo de recurrencia (ROR) y estado nodal.

35 El método de la presente invención puede incluir determinar al menos uno de, una combinación de, o cada uno de, los siguientes: grado de tumor, ploidía de tumor, expresión del receptor de estrógenos, expresión del receptor de progesterona, y expresión de HER2/ERBB2.

40 El riesgo puede ser la supervivencia específica del cáncer de mama, la supervivencia libre de eventos, o la respuesta al tratamiento.

45 La expresión de ARN de los miembros de la lista de genes intrínsecos NANO46 puede determinarse con el uso del sistema de código de nanoreporter (sistema de análisis nCounter®).

50 El método puede utilizar un muestreo de células o tejidos. La muestra puede ser un tumor. El tejido puede obtenerse de una biopsia. La muestra puede ser un muestreo de fluidos corporales. El líquido corporal puede ser sangre, linfa, orina, saliva o aspiración de pezón.

Breve descripción de las figuras

55 La Figura 1 es un mapa de calor de los subtipos intrínsecos de cáncer de mama y los genes intrínsecos de la Tabla 1.

La Figura 2 muestra las curvas de supervivencia de Kaplan Meier de una cohorte de pacientes con cáncer de mama no tratadas.

La Figura 3 muestra las curvas de supervivencia de Kaplan Meier de una cohorte de pacientes con cáncer de mama ER+, con ganglios negativos, tratados con tamoxifeno.

5 La Figura 4 muestra una probabilidad de evento en 10 años en función de la Puntuación ROR en pacientes con cáncer de mama con ganglios negativos, ER+ tratados con tamoxifeno. El gráfico muestra la subpoblación subtipada como Luminal A o B dentro de esta población. RFS = Supervivencia libre de recurrencia; DSS = supervivencia específica de la enfermedad

La Figura 5 es un esquema del ensayo de subtipaje intrínseco de cáncer de mama.

10 La Figura 6 es un esquema del proceso del algoritmo.

La Figura 7 es una ilustración que muestra la hibridación del CodeSet a ARNm.

La Figura 8 es una ilustración que muestra la eliminación del exceso de reporteros.

La Figura 9 es una ilustración que muestra la unión de los reporteros a la superficie de un cartucho.

La Figura 10 es una ilustración que muestra la inmovilización y la alineación de un reportero.

15 La Figura 11 es una ilustración de la recopilación de datos.

La Figura 12 es una ilustración del proceso de ensayo de la prueba de cáncer de mama del sistema de análisis nCounter.

La Figura 13 es una ilustración de la estación de preparación nCounter.

La Figura 14 es una ilustración del Analizador Digital nCounter.

20 Descripción detallada de la invención

La invención se define por las reivindicaciones adjuntas. La descripción presenta un método para predecir el resultado en un sujeto que tiene cáncer de mama que comprende: proporcionar una muestra de tumor del sujeto; determinar la expresión de los genes en la lista de genes intrínsecos NANO46 de la Tabla 1 en la muestra de tumor; determinar el subtipo intrínseco de la muestra de tumor basándose en la expresión de los genes en la lista de genes intrínsecos NANO46, en donde el subtipo intrínseco consiste en al menos tipo Basal, Luminal A, Luminal B o enriquecido con HER2; determinar una puntuación de proliferación basada en la expresión de un subconjunto de genes de proliferación en la lista de genes intrínsecos NANO46; determinar el tamaño del tumor, calcular una puntuación de riesgo de recurrencia con el uso de una suma ponderada de dicho subtipo intrínseco, puntuación de proliferación y tamaño del tumor; y determinar si el sujeto tiene un riesgo bajo o alto de recurrencia en función de la puntuación de recurrencia. En una modalidad, una puntuación baja indica un resultado más favorable y una puntuación alta indica un resultado menos favorable.

35 Los genes intrínsecos se seleccionan estadísticamente para tener baja variación en la expresión entre réplicas de muestras biológicas del mismo individuo y alta variación en la expresión entre muestras de diferentes individuos. Así, los genes intrínsecos se usan como genes clasificadores para la clasificación del cáncer de mama. Aunque la información clínica no se usó para derivar los subtipos intrínsecos de cáncer de mama, esta clasificación ha demostrado tener importancia pronóstica. El tamizaje de genes intrínsecos puede usarse para clasificar los cánceres de mama en cinco subtipos intrínsecos moleculares distintos, Luminal A (LumA), Luminal B (LumB), enriquecido con HER2 (Her-2-E), tipo Basal y tipo Normal (Perou y otros Nature, 406 (6797):747-52 (2000); Sorlie y otros. PNAS, 98 (19): 10869-74 (2001)).

40 Un ensayo de expresión de genes NANO46, como se describe en la presente descripción, puede identificar el subtipo intrínseco de una muestra biológica, por ejemplo, un tejido de tumor embebido en parafina fijado con formalina estándar. Los métodos usan un algoritmo supervisado para clasificar muestras de sujetos de acuerdo con el subtipo intrínseco de cáncer de mama. Este algoritmo, denominado en la presente descripción como modelo de clasificación NANO46, se basa en el perfil de expresión de genes de un subconjunto definido de genes intrínsecos que se ha identificado en la presente descripción como superior para clasificar los subtipos intrínsecos de cáncer de mama. El subconjunto de genes, junto con los cebadores de secuencias específicas de la diana utilizadas para su detección, se proporciona en la Tabla 1. La Tabla 1A proporciona las secuencias de las secuencias de sonda específicas de la diana para detectar cada gen usado en la Tabla 1. Las secuencias proporcionadas en la Tabla 1A son meramente representativas. El experto en la técnica puede usar cualquier sonda específica de secuencia de la diana para detectar cualquiera de (o cada uno) de los genes en la Tabla 1.

Tabla 1

55

GEN	NÚMERO DE ACCESO REPRESENTATIVO DE GENBANK	CEBADOR DIRECTO	SEQ ID NO:	CEBADOR INVERSO	SEQ ID NO:
ACTR3B	NM_020445	AAAGATTCCTGGG ACCTGA	1	TGGGGCAGTTCGTGTA TTACTTC	47
	NM_001040135				
ANLN	NM_018685	ACAGCCACTTTCA GAAGCAAG	2	CGATGGTTTTGTACA AGATTTCTC	48

60

65

ES 2 763 931 T3

GEN	NÚMERO DE ACCESO REPRESENTATIVO DE GENBANK	CEBADOR DIRECTO	SEQ ID NO:	CEBADOR INVERSO	SEQ ID NO:
BAG1	NM_004323	CTGGAAGAGTTGA ATAAAGAGC	3	GCAAATCCTTGGGC AGA	49
BCL2	NM_000633	TACCTGAACCGGC ACCTG	4	GCCGTACAGTTCCAC AAAGG	50
BLVRA	BX647539	GCTGGCTGAGCAG AAAG	5	TTCCTCCATCAAGAG TTCAACA	51
CCNE1	BC035498	GGCCAAAATCGAC AGGAC	6	GGGTCTGCACAGAC TGCAT	52
CDC20	BG256659	CTGTCTGAGTGCC GTGGAT	7	TCCTTGTAATGGGGA GACCA	53
CDC6	NM_001254	GTAATCACCTTC TGAGCCT	8	ACTGGGATATGTGA ATAAGACC	54
CDCA1	NM_031423	GGAGGCGGAAGA AACCAG	9	GGGAAAGACAAAG TTCCA	55
CDH3	BC041846	GACAAGGAGAAT CAAAAGATCAGC	10	ACTGTCTGGGTCCAT GGCTA	56
CENPF	NM_016343	GTGGCAGCAGATC ACAA	11	GGATTTCTGGTGGG TTC	57
CEP55	AB091343	CCTCACGAATTGC TGAACCT	12	CCACAGTCTGTGATA AACGG	58
CXXC5	BC006428	CATGAAATAGTGC ATAGTTTGCC	13	CCATCAACATTCTCT TTATGAACG	59
EGFR	NM_005228	ACACAGAATCTAT ACCCACCAGAGT	14	ATCAACTCCCAAAC GGTCAC	60
ERBB2	NM_001005862	GCTGGCTCTACA CTGATAG	15	GCCCTTACACATCGG AGAAC	61
ESR1	NM_001122742	GCAGGGAGAGGA GTTTGT	16	GACTTCAGGGTGCTG GAC	62

ES 2 763 931 T3

GEN	NÚMERO DE ACCESO REPRESENTATIVO DE GENBANK	CEBADOR DIRECTO	SEQ ID NO:	CEBADOR INVERSO	SEQ ID NO:
FXO1	NM_130398	CCCATCCATGTGA GGAAGTATAA	17	TGTGAAGCCAGCAA TATGTATC	63
FGFR4	AB209631	CTTCTTGACCTT GGCG	18	TATTGGGAGGCAGG AGGTTTA	64
FOXA1	NM_004496	GCTACTACGCAGA CACG	19	CTGAGTTCATGTTGC TGACC	65
FOXC1	NM_001453	GATGTTTCGAGTCA CAGAGG	20	GACAGCTACTATTCC CGTT	66
GPR160	AJ249248	TTCGGCTGGAAGG AACC	21	TATGTGAGTAAGCTC GGAGAC	67
HSPC150 (UBE2T)	NM_014176	GGAGATCCGTCAA CTCCAAA	22	AGTGGACATGCGAG TGGAG	68
KIF2C	NM_006845	TGGGTCGTGTCAG GAAAC	23	CACCGCTGGAAACT GAAC	69
KNTC2	NM_006101	CGCAGTCATCCAG AGATGTG	24	CGTGCACATCCATGA CCTT	70
KRT14	BC042437	ACTCAGTACAAGA AAGAACCG	25	GAGGAGATGACCTT GCC	71
KRT17	AK095281	GTTGGACCAGTCA ACATCTCTG	26	GCCATAGCCACTGCC ACT	72
KRT5	M21389	TGTGGCTCATTAG GCAAC	27	CTTCGACTGGACTCT GT	73
MAPT	NM_001123066	GACTCCAAGCGCG AAAAC	28	CAGACATGTTGGTAT TGCACATT	74
MDM2	M92424	CCACAAAATATTC ATGGTTCTTG	29	AGGCGATCCTGGGA AATTAT	75
MELK	NM_014791	CCAGTAGCATTGT CCGAG	30	CCCATTGTCTGTCT TCAC	76

ES 2 763 931 T3

GEN	NÚMERO DE ACCESO REPRESENTATIVO DE GENBANK	CEBADOR DIRECTO	SEQ ID NO:	CEBADOR INVERSO	SEQ ID NO:
MIA	BG765502	GTCTCTGGTAATG CACACT	31	CTGATGGTTGAGGCT GTT	77
MKI67	NM_002417	GTGGAATGCCTGC TGACC	32	CGCACTCCAGCACCT AGAC	78
MLPH	NM_024101	AGGGGTGCCCTCT GAGAT	33	TCACAGGGTCAAAC TTCCAGT	79
MMP11	NM_005940	CGAGATCGCCAAG ATGTT	34	GATGGTAGAGTTCC AGTGATT	80
MYC	NM_002467	AGCCTCGAACAAT TGAAGA	35	ACACAGATGATGGA GATGTC	81
NAT1	BC013732	ATCGACTGTGTAA ACAAC TAGAGAA GA	36	AGTAGCTACATCTCC AGGTICTCTG	82
ORC6L	NM_014321	TTTAAGAGGGCAA ATGGAAGG	37	CGGATTTTATCAACG ATGCAG	83
PGR	NM_000926	TGCCGCAGAACTC ACTTG	38	CATTGCCCGTCCTTC ATCG	84
PHGDH	AK093306	CCTCAGATGATGC CTATCCA	39	GCAGGTCAAAACTC TCAAAG	85
PTTG1	BE904476	CAGCAAGCGATGG CATAGT	40	AGCGGGCTTCTGTAA TCTGA	86
RRM2	AK123010	AATGCCACCGAAG CCTC	41	GCCTCAGATTTCAAC TCGT	87
SFRP1	BC036503	TCGAACTGAAGGC TATTTACGAG	42	CTGCTGAGAATCAA AGTGGGA	88
SLC39A6	NM_012319	GTCGAAGCCGCAA TTAGG	43	GGAACAAACTGCTC TGCCA	89
TMEM45B	AK098106	CAAACGTGTGTTC TGGAGG	44	ACAGCTCTTTAGCAT TTGTGGA	90

ES 2 763 931 T3

GEN	NÚMERO DE ACCESO REPRESENTATIVO DE GENBANK	CEBADOR DIRECTO	SEQ ID NO:	CEBADOR INVERSO	SEQ ID NO:
TYMS	BQ56428	TGCCCTGTATGAT GTCAGGA	45	GGGACTATCAATGTT GGGTTC	91
UBE2C	BC032677	GTGAGGGGTGTCA GCTCAGT	46	CACACAGTTCAGTGC TCCACA	92

Tabla 1a. Sondas para detectar los genes NANO46

Nombre del Gen	Acceso RefSec	Secuencia diana	SEQ ID NO:
ACTR3B	NM_001040135.1	CCAGAAGAAGTTTGTATAGACGTTGGTTACG AAAGATTCTGGGACCTGAAATATCTTTCAC CCGGAGTTTGCCAACCCAGACTTTATGGAGTC CATC	140
ANLN	NM_018685.2	CGTGCCAGGCGAGAGAATCTTCAGAGAAAA TGGCTGAGAGGCCACAGCAGCTCCAAGGTC TATGACTCATGCTAAGCGAGCTAGACAGCCA CTTCAG	141
BAG1	NM_004323.3	CTTCATGTTACCTCCAGCAGGGCAGCAGTGA ACCAGTTGTCCAAGACCTGGCCAGGTTGTTG AAGAGGTCATAGGGGTTCCACAGTCTTTTCAG AAAC	142
BCL2	NM_000633.2	CCAAGCACCGCTTCGTGTGGCTCCACCTGGAT GTTCTGTGCCTGTAAACATAGATTTCGCTTCC ATGTTGTTGGCCGGATCACCATCTGAAGAGCA GACG	143
BLVRA	NM_000712.3	TTCTGAAAAAAGAAGTGGTGGGAAAGACC TGCTGAAAGGGTCGCTCCTTTCACAGCTGGC CCGTTGGAAGAAGAGCGGTTTGGCTTCCCTGC ATTCA	144
CCNE1	NM_001238.1	GAGAACTGTGTCAAGTGGATGGTTCCATTTGC CATGGTTATAAGGGGAGACGGGGAGCTCAAAA CTGAAGCACTTCAGGGGCGTTCGCTGATGAAG ATGCAC	145
CDC20	NM_001255.1	CCCGAGTGGGCTCCCTAAGCTGGAACAGCTA TATCCTGTCCAGTGGTTCACGTTCTGGCCACA TCCACCACCATGATGTTCCGGTAGCAGAACA CCATGT	146
CDC6	NM_001254.3	GGGGAAGTTATATGAAGCCTACAGTAAAGTC TGTCGAAACAGCAGGTGGCGGCTGTGGACC AGTCAGAGTGTGTCACCTTCAGGGCTCTTG GAAGCC	147
CDCA1	NM_145697.1	GCCTGGCGGTGTTTCGTCTGTGCTCAGCGGTG GGAGGAGGCGGAAGAAACCAGAGCCTGGGA GATTAACAGGAAACTTCCAAGATGGAACCTT TGTCCTT	148

ES 2 763 931 T3

5
10
15
20
25
30
35
40
45
50
55
60
65

Nombre del Gen	Acceso RefSec	Secuencia diana	SEQ ID NO:
CDH3	NM_001793.3	CCCTCGACCGTGAGGATGAGCAGTTTGTGAG GAACAACATCTATGAAGTCATGGTCTTGGCCA TGGACAATGGAAGCCCTCCCACCACTGGCAC GGGAAC	149
CENPF	NM_016343.3	AGAAAATCTTGCAGAGTCCTCCAAACCAACA GCTGGTGGCAGCAGATCACAAAAGGTCAAAG TTGCTCAGCGGAGCCCAGTAGATTCAGGCAC CATCCTC	150
CEP55	NM_018131.3	GTACTACCGCATTGCTTGAACAGCTGGAAGA GACAACGAGAGAAGGAGAAAAGGAGGGAGCA GGTGTGAAAAGCCTTATCTGAAGAGAAAAGAC GTATTGAA	151
CXXC5	NM_016463.5	AGCTGCCCTCTCCGTGCAATGTCACCTGCTCGT GTGGTCTCCAGCAAGGGATTCCGGCGAAGAC AAACGGATGCACCCGTCTTTAGAACCAAAAA TATTCT	152
EGFR	NM_005228.3	GCAGCCAGGAACGTACTGGTGAAAACACCGC AGCATGTCAAGATCACAGATTTTGGGCTGGCC AAACTGCTGGGTGCGGAAGAGAAAAGAATACC ATGCAG	153
ERBB2	NM_004448.2	TGAAGGTGCTTGGATCTGGCGCTTTTGGCACA GTCTACAAGGGCATCTGGATCCCTGATGGGG AGAATGTGAAAATTCCAGTGGCCATCAAAGT GTTGAG	154
ESR1	NM_000125.2	AGGAACCAGGGAAAATGTGTAGAGGGCATGG TGGAGATCTTCGACATGCTGCTGGCTACATCA TCTCGGTTCCGCATGATGAATCTGCAGGGAGA GGAGT	155
EXO1	NM_006027.3	TGGCCACAAAGTAATTAAGCTGCCCGTCT CAGGGGGTAGATTGCCTCGTGGCTCCCTATGA AGCTGATGCGCAGTTGGCCTATCTTAACAAAG CGGG	156
FGFR4	NM_002011.3	CCCACATCCAGTGGCTGAAGCACATCGTCATC AACGGCAGCAGCTTCGGAGCCGACGGTTTCC CCTATGTGCAAGTCCTAAAGACTGCAGACATC AATAG	157
FOXA1	NM_004496.2	TGGATGGTTGTATTGGGCAGGGTGGCTCCAG GATGTTAGGAACTGTGAAGATGGAAGGGCAT GAAACCAGCGACTGGAACAGCTACTACGCAG ACACGCA	158
FOXC1	NM_001453.1	TTGAGTACAGAGGATCGGCTTGAACAACCT CTCCAGTGAACGGGAATAGTAGCTGTCAAAT GGCCTTCCCTTCCAGCCAGTCTCTGTACCGCA CGTCCG	159
GPR160	NM_014373.1	GGATTTCAGTCCTTGTCTTATGTTTTGGGAGAC CCAGCCATCTACCAAAGCCTGAAGGCACAGA ATGCTTATTCTCGTCACTGTCCTTCTATGTCA GCAT	160
UBE2T	NM_014176.1	GTGTCAGTCTCAGTGCATCCCAGGCAGCTCTTA GTGTGGAGCAGTGAAGTGTGTGGTTCCCTTC TACTTGGGGATCATGCAGAGAGCTTCACGTCT GAAG	161

ES 2 763 931 T3

Nombre del Gen	Acceso RefSec	Secuencia diana	SEQ ID NO:
KIF2C	NM_006845.2	GTTGTCTACAGGTTACAGCAAGGCCACTGGT ACAGACAATCTTTGAAGGTGGAAAAGCAACT TGTTTTGCATATGGCCAGACAGGAAGTGGCA AGACAC	162
KNTC2	NM_006101.1	AAAAGGTCATAAGCATGAAGCGCAGTTCAGT TTCCAGCGGTGGTGTGCTGGCCGCCTCTCCATGC AGGAGTTAAGATCCCAGGATGTAATAAACA AGGCCT	163
KRT14	NM_000526.3	GCAGTCATCCAGAGATGTGACCTCCTCCAGCC GCCAAATCCGCACCAAGGTCATGGATGTGCA CGATGGCAAGGTGGTGTCCACCCACGAGCAG GTCCTT	164
KRT17	NM_000422.1	CTGACTCAGTACAAGAAAGAACCGGTGACCA CCCCTCAGGTGCGTACCATTGTGGAAGAGGT CCAGGATGGCAAGGTCATCTCCTCCCGCGAG CAGGTCC	165
KRT5	NM_000424.2	CTGGTCTCTTGCTCCACCAGGAACAAGCCAC CATGTCTCGCCAGTCAAGTGTGTCTTCCGGA GCGGGGGCAGTCGTAGCTTCAGCACCGCCTCT GCCA	166
MAPT	NM_016835.3	GCCGGTCCCTCAACTCAAAGCTCGCATGGTC AGTAAAAGCAAAGACGGGACTGGAAGCGATG ACAAAAAGCCAAGACATCCACACGTTCCCTC TGCTAA	167
MDM2	NM_006878.2	GGTGAGGAGCAGGCAAATGTGCAATACCAAC ATGTCTGTACCTACTGATGGTGTGTGTAACCAC CTCACAGATTCCAGCTTCGGAACAAGAGACC CTGGTT	168
MELK	NM_014791.2	AGAGACAGCCAACAAAATATTCATGGTTCCTT GAGTACTGCCCTGGAGGAGAGCTGTTTGACT ATATAATTTCCAGGATCGCCTGTGAGAAGAG GAGACC	169
MIA	NM_006533.1	CCGGGGCCAAGTGGTGTATGTCTTCTCCAAGC TGAAGGGCCGTGGGCGGCTCTTCTGGGGAGG CAGCGTTCAGGGAGATTACTATGGAGATCTG GCTGCT	170
MKI67	NM_002417.2	GCTTCCAGCAGCAAATCTCAGACAGAGGTTTC CTAAGAGAGGAGGAGAAAGAGTGGCAACCTG CCTTCAAAAAGAGAGTGTCTATCAGCCGAAAGT CAACATG	171
MLPH	NM_024101.4	GAGGAAGTCAAACCTCCCAGATATTCTCCCTC GAGTGGCTGGGAAACTTGGCAAGAGACCAGA GGACCCAAATGCAGACCCTTCAAGTGAGGCC AAGGCA	172
MMP11	NM_005940.3	AGCAGCCAAGGCCCTGATGTCCGCCTTCTACA CCTTTCGCTACCCACTGAGTCTCAGCCCAGAT GACTGCAGGGGCGTTCAACACCTATATGGCC AGCCC	173

ES 2 763 931 T3

Nombre del Gen	Acceso RefSec	Secuencia diana	SEQ ID NO:
MYC	NM_002467.3	CACCGAGGAGAATGTCAAGAGGGCGAACACAC AACGTCTTGGAGCGCCAGAGGAGGAACGAGC TAAACCGGAGCTTTTTTGCCTGCGTGACCAG ATCCCG	174
NAT1	NM_000662.4	AGCACTTCCTCATAGACCTTGGATGTGGGAGG ATTGCATTCACTAGTTCTGGTTGCCGGCT GAAATAACCTGAATTCAAGCCAGGAAGAAGC AGCAA	175
ORC6L	NM_014321.2	GACTGTGTAACAACACTAGAGAAGATTGGACA GCAGGTCGACAGAGAACCTGGAGATGTAGCT ACTCCACCACGGAAGAGAAAGAAGATAGTGG TTGAAGC	176
PGR	NM_000926.2	GGGATGAAGCATCAGGCTGTCATTATGGTGTG CTTACCTGTGGGAGCTGTAAGGCTTCTTTAA GAGGGCAATGGAAGGGCAGCACAACTACTTA TGTGC	177
PHGDH	NM_006623.2	GCGACGGCTTCGATGAAGGACGGCAAATGGG AGCGGAAGAAGTTCATGGGAACAGAGCTGAA TGAAAAGACCCTGGGAATTCTTGGCCTGGGC AGGATTG	178
PTTG1	NM_004219.2	CACCAGCCTTACCTAAAGCTACTAGAAAAGGC TTGGGAAGTGTCAACAGAGCTACAGAAAAAG TCTGTAAAGACCAAGGGACCCCTCAAAACAAA AACAGCC	179
RRM2	NM_001034.1	TTCTTTTGGACCGCCGAGGAGGTTGACCTCT CCAAGGACATTCAGCACTGGGAATCCCTGAA ACCCGAGGAGAGATATTTTATATCCCATGTT TGGCT	180
SFRP1	NM_003012.3	GTGGGTACACACACGCACTGCGCCTGTCAGT AGTGGACATTGTAATCCAGTCGGCTTGTCTT GCAGCATTCCCGCTCCCTCCCTCCATAGCCA CGCT	181
SLC39A6	NM_012319.2	GATCGAACTGAAGGCTATTTACGAGCAGACT CACAAAGAGCCCTCCCACTTTGATTCTCAGCAG CCTGCAGTCTTGAAGAAGAAGAGGTCATGA TAGCTC	182
TMEM45B	NM_138788.3	CTGGCTGCCCTCAGCATTGTGGCCGTCAACTA TTCTCTTGTTTACTGCCTTTTACTCGGATGAA GAGACACGGAAGGGGAGAAATCATTGGAATT CAGA	183
TYMS	NM_001071.1	TGCTAAAAGAGCTGTCTTCCAAGGGAGTGAAA ATCTGGGATGCCAATGGATCCCGAGACTTTTT GGACAGCCTGGGATTCTCCACCAGAGAAGAA GGGGAC	184
UBE2C	NM_007019.2	GTCTGCCCTGTATGATGTCAGGACCATTCTGC TCTCCATCCAGAGCCTTCTAGGAGAACCCAAC ATTGATAGTCCCTTGAACACACATGCTGCCGA GCTC	185

La Tabla 2 proporciona secuencias seleccionadas de los genes NANO46 de la Tabla 1.

Tabla 2

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_020445	<p>CAGCGGCGCTGCGGC3GCTCGCGGGAGACGCTGCGCGCGGGGCTAGCGGGCGGCGGAGCGGGCGAGC GGGCGCTICGCGGCTGCGCGCGGGGCGAGCGCC3CGCGTCCCGAGCATGGCA3GCTCCCTGCGCTCCCTG CGTGGTGGACTGTGGCACC3GGTATACCAAGCTTGGCTACGCAGGCAACACTGAGCCCGAGTTCATTAT CCTTCATGTATTGCCATCAGAGAGTCAGCAAGGTAGTTGACCAAGCTCAAAG3AGAGTGTGAGGGGAG TTGATGACCTTGACTTTTTCAATAGGAGATGAAGCCATCGATAAACCTACATATGCTACAAAAGTGGCCGAT ACGACATGGAATCATTGAAGACTGGGATCTTATG3AAAGTTCAAGGAGCAAGTGGTTTTTAAATATCTT CGAGCTGAACCTGAGGACCATTTATTTTTAATGACAGAACTCCACTCPATAACACAGAAAACAGAGAGT ATCTTGAGAAAATATGTTTGAATCAITTTAACGTACCAGGACTCTACATTGCAGTTCAGGCAGTGTCTGGC CTTGGCGGCATCTTGGACATCTCGACAAGTGGGTGAACGTACGTTAACGGGGATAGTCAITGACAGCGGA GATGGAGTCACCCATGTTATCCAGTGGCAGAAAG3IITATGTAATGGAGGCTGCATCAAACACATCCCGA TTGCAGGTAGAGATATTACGTATTTCAITCAACAGCTGCTAAGGGAGAGGGAG3TGGGAATCCCTCCTGA GCAGTCACTGGAGACCCCAAAGCCATTAAGGAGAAATACITGTTACATTTGCCCGGATATAGTCAAAGGAA TTTGCCAAGTATGATGIGGATCCCGGAAAGTGGATCAAACAGTACACGGGTATCAATGGGATCAACCAGA AGAAGTTTGTATAGACGTTGGTTACGAAAGATTCCTGGGACCTGAAATATCTTTACCCCGGAGTTTGC CAACCCAGACTTTATGAGTCCATCTCAGATGTTGTTGATGAAGTAAACAGAACTGCCCGCATCGATGTG CGCGCCCGCTGTATAAGAAATGCTCACTCTCAG3AGGCTCCACCATGTTTCAGGATTTCCGACGCGCAG TGCAGAGG3ATTTGAAGAGAGTGGTGGATGCTAGGCTGAGGCTCAGCGAGGAGCTCAGCGCGGGAGGAT CAAGCCGAAAGCCTGTGAGGTTCCAGTGGTCAACGATCAGATGCAGCGCTACGCGGTGTGGTTCGGAGGC TCCATGCTGGCCTCGACTCCCGAGTCTTTTCAGGTCTGCCACACCAAGAGGACTATGAACAGTACGGCC CCAGCATCTGCCCGCCAAACCCCGCTTTTGGAGTCAATGCTTAGTCTGCTGCTGAAACCGGTGCTTCGATG GTGTCACGTTGGGGAACAAGTGTCTTCAGAACCCAGAGAAGGCCCGCTTCTGTAATAGCGACGTCGG TGTGCTGCCAGCAG3CGTGTCTGCAITGGCCGTTGCATGAGGCGCGCGCGGGCCCTTCAGTAAAAGCCA TTTATCCGTTGCGCAGCCGCTGCTGCGAGCCCTCCCTCTCTCCCGCCCTCCCTCACCCCTGGCTCTCCCTC CTCTCTCTCTCCGAGCTGCTAGCTGACAAATACAATTCGAAAGGAATCCAAATGTGACTTTGAAAATTTG TTAGAGAAAACAACATTAGAAAATGGCGCAAAATCGTTAGGTCAGGAGAGAAAGTGGGGCGCAAAACC CTTTTCTCCAGCCATTTTTTGTAAATAAAATGTTTAAACTTGAATACAAATCGATGTTTATTTTCC TATCATTTTGTATTTTATGGTATTTGGTACAAC3GCTGATACTAAGCAGCAATAGATATTGATGTTATG GAGTGTGTAATCCAAAGTTTTTAACTGTGAGGCAITGTTCTGATATGTTTATAGGCAAAACAATAAAACA GCAAACTTTTTGGCCACATGTTTGGCAGAAAATGATTTACTTTATTTGGAGTACATGAAGTTTGAACAC TAAACAGTAATGATGAGAATTACTACAGATACATGATCTTTTGTGTTTTTTTGTGAACTTCTGTGA GCTGTTTTATAGAAGATGATGGTTTGTTCGGT3AGTGTGGATGAAATACTTCTTGCACCATGTGTA TAAAAGCTGTTAGAATATTTGTAATATC</p>	93
NM_001040135	<p>CAGCGGCGCTGCGGC3GCTCGCGGGAGACGCTGCGCGCGGGGCTAGCGGGCGGCGGAGCGGGCGAGC GGGCGCTICGCGGCTGCGCGCGGGGCGAGCGCC3CGCGTCCCGAGCATGGCA3GCTCCCTGCGCTCCCTG CGTGGTGGACTGTGGCACC3GGTATACCAAGCTTGGCTACGCAGGCAACACTGAGCCCGAGTTCATTAT CCTTCATGTATTGCCATCAGAGAGTCAGCAAGGTAGTTGACCAAGCTCAAAG3AGAGTGTGAGGGGAG TTGATGACCTTGACTTTTTCAATAGGAGATGAAGCCATCGATAAACCTACATATGCTACAAAAGTGGCCGAT ACGACATGGAATCATTGAAGACTGGGATCTTATG3AAAGTTCAAGGAGCAAGTGGTTTTTAAATATCTT CGAGCTGAACCTGAGGACCATTTATTTTTAATGACAGAACTCCACTCPATAACACAGAAAACAGAGAGT ATCTTGAGAAAATATGTTTGAATCAITTTAACGTACCAGGACTCTACATTGCAGTTCAGGCAGTGTCTGGC CTTGGCGGCATCTTGGACATCTCGACAAGTGGGTGAACGTACGTTAACGGGGATAGTCAITGACAGCGGA GATGGAGTCACCCATGTTATCCAGTGGCAGAAAG3IITATGTAATGGAGGCTGCATCAAACACATCCCGA TTGCAGGTAGAGATATTACGTATTTCAITCAACAGCTGCTAAGGGAGAGGGAG3TGGGAATCCCTCCTGA GCAGTCACTGGAGACCCCAAAGCCATTAAGGAGAAATACITGTTACATTTGCCCGGATATAGTCAAAGGAA TTTGCCAAGTATGATGIGGATCCCGGAAAGTGGATCAAACAGTACACGGGTATCAATGGGATCAACCAGA AGAAGTTTGTATAGACGTTGGTTACGAAAGATTCCTGGGACCTGAAATATCTTTACCCCGGAGTTTGC CAACCCAGACTTTATGGAGTCCATCTCAGATGTTGTTGATGAAGTAAACAGAACTGCCCGCATCGATGTG CGCGCCCGCTGTATAAGCCGAGTCTTTTCAGGTCTGCCACACCAAGAGGACTATGAAGAGTACGGCC CCAGCATCTGCCCGCCAAACCCCGCTTTTGGAGTCAATGCTTAGTCTGCTGCTGAAACCGGTGCTTCGATG GTGTCACGTTGGGGAACAAGTGTCTTCAGAACCCAGAGAAGGCCCGCTTCTGTAATAGCGACGTCGG TGTGCTGCCAGCAG3CGTGTCTGCAITGGCCGTTGCATGAGGCGCGCGGGCCCTTCAGTAAAAGCCA TTTATCCGTTGCGCAGCCGCTGCTGCGAGCCCTCCCTCTCTCCCGCCCTCCCTCACCCCTGGCTCTCCCTC CTCTCTCTCTCCGAGCTGCTAGCTGACAAATACAATTCGAAAGGAATCCAAATGTGACTTTGAAAATTTG TTAGAGAAAACAACATTAGAAAATGGCGCAAAATCGTTAGGTCAGGAGAGAAAGTGGGGCGCAAAACC CTTTTCTCCAGCCATTTTTTGTAAATAAAATGTTTAAACTTGAATACAAATCGATGTTTATTTTCC TATCATTTTGTATTTTATGGTATTTGGTACAAC3GCTGATACTAAGCAGCAATAGATATTGATGTTATG GAGTGTGTAATCCAAAGTTTTTAACTGTGAGGCAITGTTCTGATATGTTTATAGGCAAAACAATAAAACA GCAAACTTTTTGGCCACATGTTTGGCAGAAAATGATTTACTTTATTTGGAGTACATGAAGTTTGAACAC TAAACAGTAATGATGAGAATTACTACAGATACATGATCTTTTGTGTTTTTTTGTGAACTTCTGTGA GCTGTTTTATAGAAGATGATGGTTTGTTCGGT3AGTGTGGATGAAATACTTCTTGCACCATGTGTA TAAAAGCTGTTAGAATATTTGTAATATC</p>	94

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_018685	CTCGGCGCTGAAATTCAAATTTGAACGGCTGCAGAGGCGGAGTCCGTCACCTGGAAGCCGAGAGGAGAGGA CAGCTGGTIGTGGGAGAGTTCACCCCGCTCAGACTCCTGGTTTTTCCAGGAGACACACTGAGCTGAGAC TCACITTTCTCTTCCTGAATTTGAAACCACCGTTTTCCATCGTCTCGTAGTCCGACCCCTGGGGCGATGGAT CCGTTTTACGGAGAAACTGCTGGAGCGAACCCCGTGCCAGGCGAGAGAATCTTCAGAGAAAAATGGCTGAGA GGCCACAGCAGCTCCAAAGTCTATGACTCATGTCTAAGCGAGCTAGACAGCCACTTTCAGAAAGCAAGTAA CCAGCAGCCCTCTCTGGTGGTGAAGAGAAAATCTTGTACAAAACCATCGCCATCAAAAAACCGCTGTTCT GACAACACTGAAGTAGAAGTTTCTAACTTGGAAAAATAACAACCCAGTTGAGTCGACATCTGCAAAATCTT GTTCTCCAAGTCTGTGTCTCCTCAGGIGCAGCCACAAGCAGCAGATACCATCAGTGTACTGTGTGCTGT CCGGCATCACTGCTGGGCATGAGGAGAGGGCTGAACCTAAGATGGAAAGCAACIGCAGCCCTCCCTCAGTT AAAACACGATGCAAAAACCTTGCAAGCAACCGCCCGTTGGGATAATGATGATATGACAGATGACATTC CTGAAAGCTCCTCTCTCACCACATGCCATCAGAGGAAAAGGCTGCTTCCCCTCCAGACCTCCGCTTTC AAATGCCTCGGCAACTCCAGTGGCAGAAAGGGCCGCTGGCCAATCTTGTGCAACTATTTGCTCCIGG GAAGATGATGTAATCACTCATTTGCAAAAACAAAACAGTGTACAAGAACAGCCCTGTTACCCGCTGTTTTAT CCAAAATTTTCTCTGCAAGTGGAGCATCTGTAGGATCAATAGCAGCAGTGTAAAGCAGGAAGCTACATT CTGTTCCTCAAGGATGGCGAATGCTCTTTGAATAAAGCCCTATCCTCAAGTGTGATGATGCTGTTTTG GTTAATGCCTCAATTTCCAGCTCTGTGAAAGCTACTTCTCCAGTGAAATCTACTACATCTATCACTGATG CTAAAAAGTTGTGAGGACAAAATCCTGAGCTACTTCCAAAAACTCCTATTAGTCTCTGAAAAAGGGGGT ATCGAAACCAATTTGTGAAGTCAACTTTATCCAGACAGTTCATCCAGGGAGAAATTAAGTAGAGAAAT TGCTGCAATCTCAATCTAAAGACAAAATCTACGACACCAGGAGGAACAGGAATTAAGCCTTTCTGGAAC GCTTTGGAGAGCGTTTCTCAAGAACATAGCAAAGAAAGTCCAGCTCGTAGCACACCCCAAGACCCCCAT TATTACTCCAAATACAAAAGGCCATCCAAGAAAAGATTTATCAAGCAAGACACATCTTCATCTACTACCCAT TTAGCACACAGCTCAAGCAGGAACGCTCAAAAAGAACTAGCATGCTTTCGTGGCCGATTTGACAAGGGCA ATATATGGAGTGCAGAAAAGGGCGGAAACTCAAAAAGCAAAACACTAGAAAACCAACAGGAAACTCACTG TCAGAGCACTCCCTCAAAAACACCAAGGTGTTCAAAACACTCAGTCACTCCAGTAACAGAAAAGGTG ACCGAAAACAGATACAGCCAAAAATTTAGTACAGAACCTAAAGGTTTCACTGAATGCGAAATGACGA AATCTAGCCCTTTGAAAAAACAATGTTTTAGAAAGGACAAAACCTTAAAGTAAGCATCAGACCCAAA GGTGAGCAGAAAATGAAGTGTATCGGAAATTTGAGATGAGTGGGATGATGATGATATCAAACTGTTG AAAGTAATTAATGACCTCTCAGTGTATGCTTAGAGGAAGGTGAACCTAGATATGGAGAAGGCCAAGAGG AGATGGATCAAGCATAGCAGAAAGCAGCGAAGAACAGGAAGATGCATGAATACTCCCTCAATGCTCTTT ACTTGCACCATTTGGCACAACAGTTGGTGTGGTAAGTCCAGAGAGTTTAGTGTCCACACCTAGACTGGAA TTGAAAGACACCAGCAGAAGTGTAAAGTCCAAAACAGGAAATTCAAAAGAACTCGTGTCCCTCGAG CTGAATCTGGTGATAGCCTTGGTTCTGAAGATCGTGATCTTCTTACAGCATGATGATATAGATCICA AAGATTCAAAGAAACAGAACGTCATCAATAAAGCAGGIGATTGTTCCGAAAGGAAGATGTTACTCAAAA CTGGATGAAAAAATAATGCTTTCTCTGCAAGTTAATAICAAAACAGAAAATGCAGGAACCTCAATAACG AAATAAATATGCAACAGACAGTATCAAGCTAGCCAGGCTCTTAACTGCTGTGTTGATGAAGAACA TGGAAAAGGGTCCCTAGAAGAAGCTGAAGCAGAAAGACTTCTTCAATTGCAACIGGGAAGAGAACACTT TTGATTGATGAATGAATAAATTTGAAGAACGAAGGACCTCAGAGGAAGAATAAGGCTAGTCCCAAAAGT AATTTATGCCATCCAAAGGATCAGTTACTTTGTGAGAAATCCGCTTGCCTCTAAAAGCAGATTTTGTCTG CAGTACGGTTCAGAAACAGATGCAGCAAAATTAATTAATTAATACTAAAAGCAGGAGCTGAAAAAT ATGGTAGCCACACCAATAGCAAGTACTTCAAACCTCTTTAAGCGTGATGCTCTGACATTTACTACTACAT TTACTCTGCAAGATGATCCAAATGACTTTGAAATAAATAITGAAGTTTACAGCTTGGTGCAAAAGAAAGA TCCCTCAGGCCCTTGATAAGAAGAAAAACATCCAAAGTCCAAAGGCTATTACCCAAAGCGACTCCTCACA TCTATAACCACAAAAAGCAACATTCATCTCAGTCAATGGCCAGTCCAGGAGGCTTAGTGTCTGTGGCGAA CCAGCAACTTCGCCCTTGTGGATCTTACACATTAATCAATGCTTCCAGTAGGAAATACTAAGTTGTTCT GGACAAGGICCCCTTTTTATCTTCTTTGAAAGGTCATATTTAATTTAAAATAAATGTCAGTGAATICC AGTGTGAAAGAAAGAGTTTTCTAACCATATTTGAAGATGTTAGTGGTTTTGGTCCCTGGCATCGAAGAT GGTGTGTCTTCTGAAAACGTATATCTTATTTGGACTTATCCAGATGATGAGAAACCGCAAGAAATCCCAT AGGAAGGATTAATCTGGCTAATTTGATACCAGTCTGATAGAAACCGCCAAAGAGAAATTTGTGCAAGA CGCAACACTTTGAAATTAATTAAGTCCGACCCAAAGAGAAAGTGAACCGAGACTCTGTGACGCCAAT GCAGGGACACACTCTGTGTACCAAGAACTGGCTGTCTGCAGATACAAAAGAAAGCGGGGATCCCTGGAT GCAAAAACICAAATCAAGTTCTGTGTGATTTGCTCTGGCAACCTGATGCTTGTACAAACCTATTGGA	95

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
	GTTCATCCCTCAGAATTCAACTCACATGICTCTTGCATGGAGAACCCTCACCTACTGTGTTGAGACCCTG TCCAGCCCCAGGTGGGATCCCTCTCGACTTCCCATACATTTCTTCCACAGCATTTACATAGTCCATGA TAGTTTTACTTGTGGGATTATTTGGTTTAACTCTTGCCTTTAACACCAGGGTTCCCTTGGGTGAAGGAGCTTC TTTATCTTGGTAACAGCATTATTTCAAGCATAAAGTTGTAATATAGTATATATACATATATAACATATATA ATATAACATAACATATATAACATATATAACAAGCATAAAGTTGTTATATAGTCTTGTATATAGTAAGACC TCAATAAAATATTTGGAGAACAACAAAAA	
NM_000633	TTTCTGTGAAGCAGAAGTCTGGGAATCGATCTGGAAATCCCTCTAATTTTTACTCCCTCTCCCGGGGACT CCTGATTCATTGGGAAGTTTCAAATCAGCTATAACTGGAGAGTGCCTGAAGATTGATGGGATCGTTGCCCTT ATGCATTTGTTTTGGTTTACAAAAAGGAAACTTGACAGAGGATCATGCTGTACTTAAAAAATACAACAT CACAGAGGAAGTAGACTGATTTAACAATACTACTAATAATAACGTGCCTCATGAAATAAGATCCGAA AGSAATTGAATAAAAAATTTCCCTGCATCTCATGCCAAGGGGAAACACCAGAATCAAGTGTTCGCCGGA TTGAAGACACCCCTCTGCCAAGAAAGCAAGACATCCAATAAAATAGCTGGATATAAATCCCTCTCTTCT TTTCTCTGGGGGCTGTGGGTGGGAGCTGGGGCGAGAGGTGCCGTGGCCCCGTGTGCTTTTCCCTCTGGGA AGSATGGCGACCTTGGGAGAACAGGGIACGATAACCGGGAGATAGTGTGAAGTACATCCATTATAAGC TGTCCGAGAGGGGCTACGAGTGGGATGGGGAGATGTGGGGCCCGGCCCGCCCGGGGGCCCGCCCGCACC GGGCATCTTCTCCTCCAGCCCGGGCACACGCCCATCCAGCCGCATCCCGGGACCCGGTCCGAGGACC TCGCCGTGCGAGACCCTGGCTGCCCGGGCCCGCCCGGGGGCTGGCTCAGCCCGTGGCCACTGTGG TCCACCTGACCTCCGCCAGGGCCGGCAGACTTCTCCCGCCGCTACCGCCGCGACTTCCCGGAGATGTC CAGCCAGCTGCACCTGACGCCCTTACCGCGCGGGGACGCTTTGCCACGGTGGTGGAGGAGCTCTTCAGG GACGGGTGAACCTGGGGAGGATTTGGCTTCTTTGAGTTCGGTGGGTGCATGTGTGTGGAGAGCGTCA ACCGGGAGATGTGCCCTGGTGGACAACATCGCCCTGTGGATGACTGAGTACCTGAACCCGGCCTGCA CACTGGATCCAGGATAACCGGAGGCTGGGATGCCCTTTGTGGAACTGTACGGCCCCAGCATCCGGCCCTGTG TTTGATTTCTCCTGGCTGTCTGTAAGACTCTGCTCAGTTGGCCCTGGTGGGAGCTTGCATCACCCCTGG GTGCCATCTGGGCCACAAGTGAAGTCAACATGCCCTGCCCAACAATAATGCAAAAGGTTTACATAAAGC AGTAGAAAATATGCAATTTGTCAGTGAIGTACCATGAAACAAAGCTGCAGGCTGTTAAGAAAAAATAAC ACACATATAAACAATCACACACACAGACACACACACACACAACAATTAACAGTCTTCAGGCAAAAACG TCGAATCAGCTATTTACTGCCAAAGGGAATATCATTTATTTTTACATTATAAGAAAAAAGATTTAT TTATTTAAGACAGTCCCATCAAAACTCCTGTCTTTGGAAATCCGACCACTAATTGCACAGCACCCTTCCG TGTGGCTCCACCTGGATGTCTGTGCCCTGTAACATAGATTCGCTTCCATGTTGTGGCCGGATCACC TCTGAAGCCAGCAGCGATGGAAAAAGGACCTGATCATTTGGGGAAGCTGGCTTTCTGGGTCTGGAGGCTG GGSAGAAGGTGTTCACTACTTGCATTTCTTTGCCCTGGGGCTGTGATATAACAGAGGGAGGTTCCCT GTGGGGGAAAGTCCATGCCCTCCCTGGCCCTGAAGAAGAGACTTTCGATATGACTCAGATGATGCATACC TGGTGGGAGGAAAGAGTTGGGAACCTCAGTGGACCTAGTACCCACTGAGATTTCACCGCCGAAAGGACA GCGATGGGAAAAATGCCCTTAAATCATAGGAAAGTATTTTAAAGCTACCAATTTGTGCCGAGAAAACCA TTTTAGCAATTTATACAATATCATCCAGTACCTTAAGCCCTGATTTGGTATATTCATATATTTGGATAC GCACCCCAACTCCCAATACCTGGCTCCTGTCTGAGTAAAGAAACAGAAATCCTCTGGAATCTGAGGAAGTGA ACATTTCCGTGACTTCCGCATCAGGAAGGCTAGAGTTACCCAGAGCATCAGGCCCCACAAGTGCCTGCT TTTAGGAGACCGAAGTCCCGAGAACCCTGCCCTGTGCCAGCTTGGAGGCTGGTCTTGGAACTGAGCCGG GGCCCTCACTGGCTCCTCCAGGATGATCAACAGGGCAGTGTGGTCTCCGAAATGTCTGGAAGCTGATGG AGCTCAGAATCCACTGTCAAGAAAGAGCAGTAGAGGGGTGTGGCTGGCCCTGTACCCCTGGGGCCCTCC AGGTAGGCCCGTTTTCAGTGGAGCATGGGAGCCACGACCCTTCTTAAGACATGTATCACTGTAGAGGGA AGAAACAGAGGCCCTGGGCCCTTCCATCAGAAAGACATGGTGAAGGCTGGGAAGCTGAGGAGAGGCA GGCCACGGCCCATTTTGGCTGTAGCACATGGCACGTGGCTGTGGCCCTGGCCACCCTGTGAGTTTAA AGCAAGGCTTTAAATGACTTTGGAGAGGGTACAAAATCTAAAAGAAGCATTGAAGTGAAGTGTGATGGA TTAATGACCCCTGTCTATGGAATACATGTAAAACATTAICTGTCACTGTAGTTGGTTTTATTTGAA AACCTGACAAAAAAGAGTTCCAGGTGIGGAATATGGGGTTATCTGTACATCCCTGGGGCATTAATAAAA AAATCAATGGTGGGGAACATAAAGAAGTAAACAAAAGAGTACATCTTCAGCAAAATAACTAGGAAATTT TTTTTTCTTCCAGTTTGAATACAGCTTGAACATGTATGGAATAACTCTGTGGCATTTATGCAATTTATA TACCATTTATCTGTATTAATTTGGAATGACTCTGTTCATGTTTAAATGCTGTGGTTGATATTCGAAA GCTGCTTTAAAAAATACATGCATCTCAGCGTTTTTTTGTTTTTAATTTGATTTTAGTTAGGCTATACA CTATTTGTGAGCAAAAGTGTATCGTTTCTGTGTTGACATTTTATCTCTTGAATCTTCAAAAGCATTTCTGA GAAGGTGAGATAAGCCCTGAGTCTCAGCTACCTAAGAAAAACCTGGATGTCACTGGCCACTGAGGAGGCTT TGTTCACCAAGTCACTGCTATTTCCAGTCAACAGAAATGTTTATTTGTGACAGTTATATCTGTGTGCC CTTTACCTTGTCTTGAAGGTTTCCCTGCTCCCTGGGCAATCCGCATTTAATTCATGTTATTCAGGAT TACATGCATGTTTGGTTAAACCCATGAGATTCATTTCAGTTAAAAATCCAGATGGCAAAATGACCAGCAGAT TCAAACTATGGTGGTTTGACCTTTAGAGAGTTGCTTTACGTGGCCCTGTTTCAACACAGACCCACCCAGA GCCCTCTGCCCTCCTTCCGGGGGGCTTCTCATGGCTGICCTTCAGGGTCTTCTGAAATGCAGTGGT GCTTACGCTCCACCAAGAAAGCAGGAAACCTGTGGTATGAAGCCAGACCTCCCGCGGGGGCTCAGGGAA CAGAATGATCAGACCTTTGAATGATTTAAATTTTAAAGCAAAATATATTTTATGAAAGGTTTACATTTG CAAAGTGAATATGAAATATCCAAATCTGTGCTGCTATCTGCCAAATCATTTTAAITGGAGTCAGTT TGCAGTATGCTCCACGTTGGTAAGATCCCAAGCTGCTTTAGAAGTAAACAAGAAGACCTGGACGTTTT TAATAAAGCCTGTTTTGTCTTTTGTGTTTGTTCAAAACGGGATCCACAGATATTTGAAAAATGTATAT ATATTAAGAGGTCACGGGGCTAATGCTGGCTGGCTGCCCTTTGCTGTGGGTTTTGTATCCCTGGTTTT AATAACAGTAAATGTGCCAGCCTCTGGCCCCAAGCTGTACAGTATTTGTGGCTGCACCTTGCCTAAGA GTAGTTGATGTTGCTATTTCCCTTATTTGTTAAAAACATGTAGAAGCAATGAATGTATATAAAGGCTCAA CTAGTCATTTTTTCTCCTCTCTCTTTTTTTCATTATATCTAATATTTTGCAGTTGGGCACAGAGAAAC CATCCCTATTTTTGTATGAAAGAGGATTCACATCTGCATCTAACCTGCTCTTTAATGAAATGAAAAACAGT CCTCTGATGTACTCCTCTTTACCTGGCCAGGGTCAGAGTTAAATAGAGTATATGCACTTTTCCAAATTTG	97

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25 30	GGGACAAGGGCTCTAAAAAAGCCCAAAGGAGAAGAACATCTGAGAACCCTCCGCGCCCTCCAGTCC CTCGCTGCACAAATCTCCGCAAGAGAGGCCAGAAATGACAGCTGACAGGGTCTATGGCCATCGGGTCTC TCCGAAGATTTGGCAGGGGCAGAAAACCTCTGGCAGGCTTAAGATTTGGAAATAAGTACAGAAATTAAGGA AGCACCTCAATTTAGTTCAAAACAAGACGCCAACATTTCTCCACAGCTCACTTACCTCTCTGTGTTCAGA TGTGGCCTTCCATTTATATGTGATCTTTGTTTTATAGTAAATGCTTATCATCTTAAGAGTGTAGCTCTGG CCCAGTGGGAAAAATTAGGAAGTGAATATAAATCSAGAGGAGTTATAAATAAACAAGATTAATTAATAA ATCAGGGCAATCCCAACACATGTCTAGCTTTCCACCTCCAGGATCTATTGAGTGAACAGAAATTCGAAATAG TCTCTATTTGTAAATGAACCTTATCCATAAACAAATAGTTTATAAATGTGAACCTAAACTCTAATTAATTC CAACTGTACTTTTAAAGCAGTGGCTTTTTAGACTTTCTTATCACTTATAGTTAGTAATGTACACCTAC TCTATCAGAAAAAACAGGAAAGGCTCGAAATACAAGCCATTCTAAGGAAATTAGGGAGTCAAGTTGAAAT TCTATTTCTGATCTTATCTGTGGTGTCTTTTGCAGCCAGACAAATGTGGTTACACACTTTTTAAGAAAT ACAATTTCACTTTGCAAGCTTATGAAGTTCCAAATCAGATCTTTATTTGTTATCAATTTGGACTTTTCA GGGATTTTTTTTTTAAATTTATTTGGGACAAAGGACATTTGTTGGAGGGGTGGGAGGGAGGAAATTTT TAAATGTAAGCAATTTCCCAAGTTTGGATCAGGGAGTGGAAAGTTTTCAGAAATAACAGAACTAAGGGTAT GAAGGACCIGTATTGGGTCGATGTGAIGCCCTCTGCCAAGAACCTTGTGTGACAAATGACAAACATTTTG AAGTTTGTGGTACGACCTTTAGATTCAGAGACATCAGCATGGCTCAAAGTGCAGCTCCGTTTGGCAGTG CAATGGTATAAATTTCAAGCTGGATATGTCTAATGGGTATTTAAACAATAAATGTGCAAGTTTTAACTAAC AGGATATTTAATGACAACCTTCTGGTGGTGGGACATCTGTTTTCTAAATGTTTTATATGTACAATAACAG AAAAAAATTTTTATAAATTAAGCAATGTGAAACTGAATGGAGAGTGATAATACAAGTCTTTTAGTCTTA CCCAGTGAATCATCTCTTCCATGTCTTTGGACAACCATGACCTTGGACAATCATGAAATATGCATCTCA CTGGATGCAAAAGAAAATCAGATGGAGCATGAAATGGTACTGTACCGGTTTCATCTGGACTGCCCCAGAAAA TAACCTCAAGCAAAACATCTTACAACAACAGGTTGTTCTGCATACCAAGCTGAGCACAGAAAGTGGGAA CACTGGTGGAGGATGGAAGGCTCGCTCAATCAAGAAAATCTGAGACTATTAATAAATAAGACTGTAGT GTAGATACTGAGTAAATCCATGCACCTAAACCTTTTGGAAAATCTGCCGTGGGCCCTCCAGATAGCTCAT TTCATTAAGTTTTCCCTCCAAGGTAGAATTTGCAAGAGTGACAGTGGATTGCATTTCTTTTGGGGAAGC TTTCTTTGGTGGTTTTGTTTTATATACCTTCTTAAGTTTTCAACCAAGGTTTGCCTTTTGTTTTGGATTA CTGGGTTATTTTTGTTTTTAAATAAAAAAATAGTGTACAATAAGTGTTTTTGTATTGAAAGCTTTTGT CAAGATTTTCATACTTTTACCTCCATGGCTCTTTTAAAGATTGATACCTTTAAGAGTGGCTGATATTC TGCAACACTGTACACATAAAAAATACGGTAAAGGATACCTTTACATGGTTAAGGTAAGTAAAGTCTCCAGTT GCGCCACTTAGCTATAATGGCACTTTGTTTTGTGTGTTGGAAAAGTACATTTGCCATTAACCTTTCTCT TGCTGTCTAGTTAATATTGTGAAGAAAAATAAAGTACAGTGTGAGATACTG	
35 40 45 50 55 60	AATGAGGGTATTTATAAACTACTTAAATATAAAAAAGAAATGAGACATCAGACTTACAGTITTTGGATACTA ATTTTTTTCACCTAACCTTCAATATGTGATAGGAGTTTTCCATCCTATTATACCGCTGTGGGATCTGATC TTGGGCACGTTAAACCAACCTCTGTGCTCGATTTTCTCACCTGTAAAAGTGGGGTAAATCAATAGCT TACTTAGTAGGATAGCCCTGAAGAATTAAGTGAATTTAGCGAACATAAATAGCTTACAATAGGGTCTTCAGC ATGGGAAGGATTCAGTAAATGTTAGCTGTCTATCATCACCACCTACAAGGAAAGCAATAGCTGGTGAAG TTTTTCCATCATTAATGTAATTTCTATAGTACGATTTCCCAAGAGATATTAATAATATGGAATAAAGGT ATTGGTATATTCCTAATTTATTTCCATAAAGATTGTATTGATAAATATGCTCATCTTCCCTTAAACGGGAT GCAATCCAGAAAAACAAGTCAAATGTAGACAAGATATCAGAAGGAAATTTCTGTAGCCAGAGAGCTTAA AATTACAATAGGGTCTCTAATTTATACTTCACTTTTTTAGGAATAATTTCTCAGTGTGTTTTCCACATTT CATATGTAATTTTTTTTTTTTTTTTTTTTGGAGACAGAGCCCTCGCCCTGTCCACAGGCTGGAGTACAGTG GCGCGATCTCCGCTCAGTCAACTTCCACCTGCTGGTTCAAGCAATTTCTTCACTTCACTTCACTTCACT CCGCCTCGGCCCTCCAAAGTGTGGGATTTAACAAGGCGTGGCATGAGTCAACCAGCCGCGCCGCGCCATCTT ACTTTTTTATCTTTTGTACCCCTGCCTATCCAGTTAGCATGTGATTAAGTCAAAGATTTGCCACTTTG GGCCACATCTATTAATTTTCACTTTTGTATAAATTTGATTTAGTTTTGATCTACACTGCTTTATCTCC CAGTCAATTTTTATAGAACTGAAAATCTGGTAAAAATCTCAAAATGCAGTCACTTCTATGTAGAGCGCA CACTCCATCAGAACCCGTGGGCTGACAGGGAAATCCCACTGTGCAGGAGCTGCGCCATTTTCATTTCTGAT TCTCTTTGGCGTATCCAGGACTCTGATGACATGATCATAATTTATCAGTAGTAAACAGGTTGGGCCATTT GTTTTTTGGTAAATCATATATTAAGATTTTAGAAATAAGTTGATAGCCATGTATTTGGAAATTTGAA AAAGCATTGCTATCTCAGCTTCAAATTAAGCTTTAATCAAATAGTAAACTTTCCATTAATGGACAGT GTATACCTTTTTGTGATTTAAAAAATAAACAACCTGAATATAGTGCCTTTGTGACAGGGGAGCTTGGTTC CTGACAATGTCTCTTACGCTTTTTTTTTTTTTTTGAGATGGAGTCTCACTGTGTCAACCCAGGCTGGAGT GCAGTGGCCCATCTGGCTCACTGCAACCTCCGCCCTGGGTCAAGTGAATTTCTCATCTCCCTGCTC CTAAGTAGCTGGGATACAGGCACGCCACCATTGACCAGCTAATTTTTACTTTTTAGTAGAGACAGGG TTTTGGCAATGTTGGCTAGGTTGGTCTCGAATCTTGCACCTCAAGTAAATCCACCCCAATGGCCCTCCCAA AGTGTGGGATTAAGGCGTGAAGCCATTTACCCCGCCCTCTCTTCCGCTTTTGAAGTGTGAGGAAATAGC TACATTCATGAGCTGCTAGATCTGCCITATGGTCAAGAAATGAAGTTGAAGTCTCAGGAACAGTGAAT ATATACACACTGATATTTCCAAAGTACAATGCCCAATTTGATCCACAAGGAAATTAAGGTCATTTGCAA CAAAATCAGAAATAGTAAACAATAAATAGRAGATAAATATGGCCAGGGATGCTGCAAACTGATATACCTG CCAAGTTTATCAGTTGGAAATCCCAACAGTGAAGAAATGAAAGGAAATTTAAGGAGACTTTTT ATAGAAGAGTGGGAAAGGATTTGGAGGACCAACAAGTGTGGTGGGCACACAGGGAAGAGCTTCAGTGG CACCTCCCTCTCTGTTTTGAAGGGGTAGGGAGGGACAGAGCTGGGAGGAGGGGGTGGAAATCTGCT TGAGGAGGACCTCCCTTCCAGACCTGCTGTGGCCATCAGAGAAATGACGCCACTGCCAGAGCAGCAGCCC GAGGAACAGGCGAGGGGAGCACAAGTACCTAGCCTCTCTCTTTCTGTTTTCTGCTTCCGATCTCTCT CACTGGCTAAACCCAGCTGGATGCTAAGAGTACAGTCAAGCTGCTGCTGAGGAGGAGCCACCCAGGAC ACCATCAGCAAGGGATCCAATGCTTTCTGCTCTGCAAGAAATGAAGGTTGGGGCGGGGGGGCTCTAC TTTCTAGGGATATTGGGAAATAAAGGAAATAGGCAAAAAATGTTTTGAAAACCAAGCACAATCTG GCACCCGTGGGCCACTACTGCTTTTGAACCCCTGGCTCTGTTTTCAAGAGTAAATGTGCTGATCTCTTT	98

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25 30	TTAGGTGCTACAGGATTTCTTTAGGTTTGTCTGTCCACCATATTTCAACTCATGTGIGCTGTTTGT GTGCTAAAACAAATATTTGCTGATGCCIGAGTGAATAGTTCAATATTTATATAAGTCAAATTTATACGT AATGATTTTTCTTTGTAACCTAGCCGTTTCTCTTTTACAAAACCTCAGAAAACCTCAGACTTTGAAAAGGCC TGAAGTTCCCTCACCTGAAATCTGAGAAGCTGGAGGCGCTTAAAAAATCTAAAGGAAAACAAAACAGTGAA AGAAATGATATAGTCAGTGTAGAGAATAPAAATATTTATGTAATTAATATGAGGATGCGAGAAGACACA TTGTGAAAATCTTGTCTGTAATAAATCTCGATCTGCTGAAAGAAAGATGTTCTCTCTAGAGATCTTTGAAAG CATAATTAATGAGCTTTTAAAAATGTTAGAAAACAAAAGTTAGACCCACACATATCTGGCGTGTGGAAAGAT TTGCATTCCTTCCCTGCCCCGCCCCGCCCCACACTTGTGAGTTGTGCCTGTGTACGCAGTTCTGTGAC ACTCGGCTGGGCAGAAATCATCTTTCAGCACTAAGGGAACATAGTTATGATCTGGACCTTCTGGGAGTGG TCAGTCCCCAAGAACAGGTATGGGACTCCAGAAAATTTCTGCTCTCAACCCATTTTGGAAAATAGAGTTACA CATTTGTTCTACAAATTAATGAGTTAATAAGCAGCTCTTTTCAAACGTTGATATGCCCCCTCCAAAGTTTAAA TACACTAGACTTTAGTGAAGTAATGACCTCATCTCATTTCTCTCCCTGTTATATTAAGATCACATTTTACAG TAAAAGGTAGAAGCTTTTGAAGTGGTGGAGGAGGAGTGTAGAGGAGGACATAGAGCAGATAGGGGCTGGAA AGTGGGGTGGAGAAAGAGTGGCTTCTCTTTGGCAGAGTACCAAGGAAAAGCCCTATCTGTACAGAACTCT TTGTGCTGGGAACTTGTGGTGCACCTGAGCCTCAACCTAGTTGCTTGGGGAGCAAGAGAGAAGC TAAAAACCTTCAAGTAAACCAAGCCAGACCAAGAAAAGTTAAACCGAAAAGAGAACCCCCACCCCCGCA AAAAAAGAAAGTAAAGTGGGTAAAGTGATATCATGTTAGCACAGAAAAGAGAAATAAGGGTCACTTAAG TTCACTGCCCCCTCTCTATTTCAAGTTGCAAAAACCTAAGGCACAAGGACCCCGTGTCTCTTAAATGTC TCACATAGCTAGTGGTGGCAAGCCAGGTTAGAACTCTGTTCTCTGGGGTACAGGCTGGCTCTTCACTC CCTCTAGAGAGATAGCTCATCTGTGTGCACCTGAGCCCGTTGTGTTTCCGGAGTCAAAGCAATAAAGGCT CAAAACCTCAAAGACTGTTTTGCAGACCCGCTGCAAGTAGATAAGGGGGAGGAGAAAACCTGCTTTAAATGTC TTCAAGCAAGTTGTTTTGCAAGGTTGTTGACTTTTTTCTTTCAACTTCTAGTGAGTCACTGCAGCCTG AGCTGTTATTTGCTATTTGCAATAATTCAGGAACCTCAAGATCTCTCTTTTAAATATTTGTTTAA TTTAGAGACAGAGTCTTGTCTGTGGTGGCCAGGCTGGAGTGCAGTGGTGTGATCTCGGCTCACTGCAGCCT CTGCTCTGGGTTCAGCAATTTCTCAIGTCTCAGCCTCCCGAAATAGCTGGTATTTGCAGGCTCTGTGCCAC CACCCCTGCTAATTTTTGTAATTTTAGTGGAGACACGGTTTCGCCATGTTGGCCGGGCTCGTCTTGGAGC TCTGGCCCTCAGGTGATCCGCCCCGCTGGCCCTCCCAAAGTGTGGGATTCAGCCCTGAGCCCTCCACAC CCGCCCTATTTATTTTAAATTTGGCTGCTCTTAGAAAGGCATACCATGTTTCTGGATGGGAAGGCT TATTAATTCACCTAATTTAATGTAATAATTTGATGCAATCATAGTCCAGTCCAGTGGAAATTTTTAA CTTGGTAAAGATGTTCTAAAATTAATGAGAGAACTTGAATTACCAAGGATTTGAAAACACTGTAAAGCCCAA TCATGTAACAGTATGTTATAACCAATGGGAATAGAGGCTCTGTGACACAGCAGAAAAGAGTCAAAAAAGA ATAACCTGATTCATAAAAAATTTAAATGIGGAGTCACTGGGGGAAAGGATTAATATTCGATAATGTAGAA ACAACTCAACTATTTGGAGAAATGTAATTTAGAGCCTTATCTCATGCCATATACAAAATACTATTTAG ATTTGATTAATAAATAAAAAAATAAAAAA	
35 40 45 50 55	BC035498 GCGCCGCCAGCGCGTGTAGGGGGCAGGCGGGATCCCGCCACCGCCGCGCTCGGCCCGCGACTCC CGCCCGCCCGCCGCACTGCCGTCCCGCCCGCCCGCTCCCGGGACTGGAGCGCCCGCTCGCCCGCGGAC AAGACCCTGGCCCTCAGCCCGGAGCAGCCCATCATCCCGAGGGAGCGCAGGGAGCGGGAIGCGAAGGAGC GGGACACCATGAAGGAGGACGGCCGCGGAGTCTCGGCTCGCTCCAGGAAGGAAAGCAAGCTGAC CGTTTTTTGCAAGGATCCAGATGAAGAAATGGCCAAAATCGACAGGACGGCGAGGGACCGTGGGAGC CAGCCTTGGGACAAATGTCAGTCTGTGCAGACCCTGCTCCCTGATCCCCACACCTGACAAAAGAGATG ATGACC GGTTTACCRAACTCAAGTGCAGCCTCGGATTAATGACCATCCAGAGGCCTCCCGCTGCC TGTACTGAGCTGGGCAATAGAGAGGAAGTCTGGAAAATCATGTTAAACAAGGAAAAGACATACTTAAGG GATCAGCACTTTCTTGGCAACACCCCTTCTGACGCCAAAATGCGAGCAATCTCTGGATTTGGTTAA TGGAGGTGTGAGTCTATAAACTTCCAGGGAGACCTTTTACTTGGCACAAGATTTCTTTGACCCGTA TATGGCAGACAAAATGTTGTAATAAATCTTTTACAGCTTATGGGATTTCTATCTTATTTATTGCA GCCAACTTGGGAAAATCTATCTCCAAAGTTGCACCAAGTTTGGGTATGTGACAGATGGAGCTTGTTCAG GAGATGAAATCTCACCATGGAATTAATGATTAATGAAGCCCTTAAAGTGGCGTTTAAAGTCCCTGACTAT TGTGCTCTGGCTGAATCTATACATGCAGGTTGCATATCTAAATCACTTACATGAAGTGCCTACTGCCGCAG TATCCCCAGCAATCTTTATACAGATTCAGAGCTGTTGGATCTCTGTGCTCTGGATGTTGACTGCTTGA AATTTCTTATGGTACTTGTCTGCTCGGCCCTTGTATCATTTCTCTGTCATCTGAATTTGATGCAAAAGGT TTCAGGGTATCAGTGTGCGACATAGAGAATGTTGTAAGTGGATGTTTCCATTTGCCATGTTATRAAGG GAGACGGGGAGCTCAAACTGAAGCACTTCAGGGGCGTCCGTGATGAAGATGCACACAACATACAGACCC ACAGAGACAGCTTGGATTTGCTGGACAAGCCCGAGCAAGAAAAGCCATGTTGCTGAAACAAAATAGGGC TTCTCTCTCCCGAGTGGGCTCTCACCDCGCCACAGAGCGTAAGAAGCAGAGCAGCGGGCCGGAATG GCGTGACCACCCCATCTCTCCCAAAAGACAGTTGGCGGCTGCTCCACGTTCTCTCTGTCTGTTG AGCGGAGGCGTGGCTTTTACAGATATCTGAATGGAAGAGTGTCTTCCACAACAGAAGTATTTCT TGTGGATGGCATCAACAGGGCAAAGTGTTTTTATGAAATGCTTATAGGTTTTTTTTTAAATAAGTGGGT CAAATACACCAGCCCTCCAGACCCAGTGCCTGCTCCCGATGCTGCTATGGAAGTGCCTACTGACCT AAGGACTCCCAACAACAAAAGCTTGAAGCTGTGAGGGCCACGGTGGCGTGGCTCTCTCCAGGTTG TTCTGGGCTCCGTTGTACCAAGTGGAGCAGGTTGTTGGGGCAAGCGTGTGACAGAGCCATAGCCAGCT GGCAGGGGGCTGCCCTCTCCACTTATCAGTTGACAGTGTACAAATGCTTTGATGAACTGTTTTGTAAG TGCTGCTATATCTATCCATTTTTTAAATAAGATAAATCTGTTTTTGAATAAATAAATAAATAAATAA AA	99

ES 2 763 931 T3

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
BG256659	<p>GAGGGCACGGGCTCCGTAGGCACCAACTGCAAGGACCCTCCCCCTGCGGGCGCTCCCAITGGCACAGTTC GCGTTTCGAGAGTGACCTGCACCTCGCTGCTTCAGCTGGATGCACCCATCCCCAATGCACCCCCCTGCGCGCT GGCAGCGCAAAGCCAAAGGAAGCCGAGGCCCGCCCGCCCTCACCCATGCGGGCCGCCAACCGATCCCACAG CGCCGGCAGGACTCCGGGCCGAACTCCTGGCAAATCCAGTTCCTCAAGGTTTCAGACCCTCTAGCAAACCT GCGGTGACCGCTAATATCCCCATCGCAGTGTGCCAGATGGAGGTGGCCAGCTTCCTCCTGAGCAAGG</p> <p>AGAACCAGCCTGAAAACAGCCAGACGCCACCAAGAAGGAACATCAGAAAGCCTGGGCTTTGAACCTGAA CGSTTTTGATGTAGAGGAAGCCAAAGATCCTTCGGCTCAGTGGAAAAACCACAAAAATGCGCCAGAGGGTT ATCACGAACAGACTGAAAGTACTCTACAGCCAAAAGGCCACTCCTGGCTCCAGCCGGAAGACCTGCCGTT TACATTCCCTCCCTGCCAAGACCGTATCCTGGATGCGCCTGAAATCGAATGACTATTAACCTGAACCTGTG GGACTGGCAGTCCGGGGAATGTCCGGCCGGGCCACGGCCACGAGGTGTTCCTGTGGAGTGCRAAGCTGG GACACACCGTGCCTTGTGCACAGGGCCACGCGGGGAAATAATCCGGGGCGCCAAAAGCGGCACTGGC GAGAGCCGCACGGGCCGGTGTGGGGGTGGTACAACAGGCCAAAAACAACACAAAGGCCAACAGACATA CGCGCGCTGACACCCAGGTGCAAAAGCGCTCAGACGAGTAGTAACCGGCACTGTGGTTGCTGCCCTCCCCAC CTCTCCCGCTCTCAGCGTAAGATAAAAAGAAAGAGCAAAAAGCAAAGAAAGAAAGACGAGACGAGACAC ACAGGAACGAACAGTAAAGCAAGCTAAAGCAAACGCAAGACCAGACAACAGAAAATAGAAAGAACCAACAG AGAGGAGACAGAAACAGGACGCCAGCAACATAGCAACAAACGAACAGAAAGAGGCCTAAACAAAAGCAGC AGCAAGACGAGACAGAGAGAGAGGAGGAGGAGGGCCGAGCGAGCAGGGAGCGCGAGCAGCGAGGCGAAG CAGCAGACAAGGGCAGCCGAAAGGGCAACGAGAGGAGGCCACACAAAAAGGAGAGGGGACAGGAGAAGC AGCGAGAGAAGCGGAGCAGCAACAGAGGAGAAAGGAGAGGGAGAGGAGGGAGAGAGCGGAAAGGAGGA AGAAAACAGCAGAGGCGACGAAGGGGGAGACGCGGGGGCAGGAAAAGACACAGGAAGGCAGCGCGGAGG AGSAGAAGGGGAAGCACGAAGGAGACGGGAAAGGAGAAGGGAGAGGACAGCGCAAGAGAGCGCGCGCGC GACAGCGAGGGACGGAGCGAGAGAGGAAACGGAAAGCGAGAGGGAAGAGGAGGCAACCGCAGCGAAC CAAACGAAAACAGCAAAAAGAGAGGAGAGGAGCGCGCAAGAGGCAAGCGCAAGCAGCAGGAAAACGAAG CGAGAGACGAGAAGCCGTTGACGAGCAGGAGAAAGGGAAAGGCAGGAGACAGGACAGGCGGAAAGAGACA CGCGAGACGCAAAAGTIGAGCAGAACGAAGCGAAAGAGCAACGCACGAGAGAAACGAC</p>	100

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_001254	GAGCGGGGCTGGAGTTTIGCTGCTGCCGCTGTGCAGTTTGTTCAGGGGCTTGTGGTGGTGGAGTCCGAGAGG CTGCGTGTGAGAGACGTGAGAAGGATCCCTGCACTGAGGAGGTGGAAAGAAGAGGATTGCTCGAGGAGGCC TGGGGTCTGTGAGGCGCGGAGCTGGGGTGAAGGCTGCGGGTTCGGGCGAGGCGCTGAGCTGTGCTGTGCTC ATGCTCAAACCCGATCCCGAGGCACAGGCTACAATCAGTTTTCCAAAAAGGAAGCTGTCTCGGGCATTGA ACAAGCTAAAACTCCAGTGTATGCCAACTAGAACCACAAATGCCAAACCGTAAACCTGTCTCCCTCG TGTAAAAGCCCTGCCTCTCAGCCCCAGGAAACGCTGTGGGCGATGACAACCTATGCAACACTCCCCATTTA CCTCCTTGTCTCCACCAAGCAAGGCAAGAAAGAGAATGGTCCCTCACATACACTTAAGGGAC GAAGATTGGTATTTGACAACTCAGCTGACAATTAAGTCTCCCTAGCAAAAGAGAAGCTAGCCAAAGTTCACCA AAACAAAATACTTCTTCAGTTAGAAAAGTCAAGAGATCACAACAAATTCGAGCAGAGATGCCACTG AAGAAAAGAACTGTGATGTGTGAGACTATTCAAGCAAGAAGGCACTTGTACACAGCAAGCAAGCTGTGCTC TGAACACAGCTGTCCAGATCGGCTGCCAGGAAAGGGAGATGGATGTCATCAGGAATTTCTGTAG GGAACACATCTGTGGAAAAAAGCTGGAAGCCTTTACCTTCTGGTGTCTCCGAACTGGAAAACTGCC TGCTTAAGCCGGATTCTGCAAGACCTCAAGAAGGAACTGAAAGGCTTTAAAATATCATGCTGAATTCCA TGTCTTGGAGACTGCCAGGCTGTATCCCGAGCTATTGCTCAGGAGATTTGTGAGGAAGAGGATCCAG GCCAGCTGGGAAGGACATGATGAGGAAATGGAAAAACATATGACTGCAGAGAAGGGCCCATGATTGTG TTGGTATTGGAGCAGATGGATCAACTGGACAGCAAGGCCAGGATGTATTGTACACGCTATTGAATGGC CATGGCTAAGCAATTTCTCACTTGGTGTGATTTGGTATTGCTAATACCCTGGATCTCACAGATAGAATCT ACCTAGGCTTCAAGCTAGAGAAAAATGTAAGCCACAGCTGTTGAACTTCCACCTTATACCAAGAAATCAG ATAGTCACTATTTTGCAAGATCGACTTAATCAGGTATCTAGAGATCAGGTTCTGGACAAGTCTGCAGTTC AATTCTGTGCCCGAAAGTCTCTGCTGTTTCAGGAGATGTTCGCAAGCACGATGATGTTTCAGGAGAGC TATTGAAATTTAGAGTCAAGTGTCAAAGCCAGACTATTCTCAAACCACTGTCTGAATGTAATCACTT CTGTAGCCCTGTGATTTCCCAAGAGGCTGGTCTTATTCACATATCCCAAGTCACTCAGAAAGTGTATGTA ACAGGATGACCTTGAAGCAAGGAGGACACAAGATTCCTTCCCTCTTCAGCAGAAGATCTTGGTTGCTC TTTGATGCTCTTGTATCAGCCAGTTGAAAATCAAAGAGGTCACCTCGGGGAAGTTATATGAAGCCCTACAGT AAAGTCTGTCCGAAACAGCAGGTGGCGGCTGTGGACCAGTCAAGTGTGTCACCTTCAGGGGCTCTTGG AAGCCAGGGGCTTTTGGATTAAAGAGAAACAAGGAAACCCGTTTGACAAAGGTGTTTTCAAGATTGA AGAGAAAGAAATAGAACATGCTCTGAAAGATAAAGCTTTAATTGGAAATATCTTAGCTACTGGATGTCCT TAAATCTTCTCTTACACCCACCCGAAAGTATTCAGCTGGCATTTAGAGAGCTACAGTCTTCAATTTTAG TGCTTTACACATTCGGGCTGAAAACAAATATGACCTTTTTTACTTGAAGCCAATGAATTTAATCTATA GATTCTTTAATATTAGCACAGAAATAATCTTTGGGCTTTACTATTTTACCCATAAAAAGTGACCAGTAA GACCCCTTTTAATTACATTCACCTACTTCTACCACTTGTGTATCTTACCCAAATGTGCTTGCAGGTGTACA GATCTGTGTAGAGGAATGTGTATATTTACCTCTTCGTTTGTCTCAAACATGAGTGGGATTTTTTTTGT TGTTTTTTTTGTGTTGTTTTTTGAGGCGGCTCACCCCTGTGCCAGGCTGGAGTGCATTTGCCCG TTCTCTGCTCACTACAGCACCCGCTTCCCAGGTTGAAGTGAATCTTGGCTCAGCCCTCCCGAGTAGCTG GGATTACAGGTGCCCAACCCGCGCCAGCTAATTTTTAATTTTAGTAGAGCAGGGTTTTACCATGT TGCCAGGCTGTCTTGAACCTCTGACCCCTCAAGTGTCTGCCACCTTGGCTCCCTAAGTGTGGGAT TATAGGCGTGTAGCCCAATGCTCAGCCATTAAGGTATTTGTTAAGAATTTAAGTTTAGGGTAAGAAGA ATGAAAATGATCCAGAAAAATGCAAGCAAGTCCACATGGAGATTTGGAGGACACTGGTTAAAGAAATTTAT TTCTTTGTATAGTATACTATGTTCACTGGTGCAGATACTACAACATTTGTGGCAATTTAGACTCGTGTAGTT TCTTGGGCACTCCCAAGGGCGTTGGGGICATAAGSAGACTATAACTCTACAGATTGTGAATATATTTATT TTCAAGTTGCATTTTGTCTTTTTAAGCAATCAGATTTCAAGAGAGCTCAAGCTTTCAGAAAGTCAATGT GAAAATTCCTTCTAGGCTGTCCACAGTCTTTGCTGCCCTTAGATGAAGCCACTTGTTCAGATGACT ACTTTGGGGTTGGGTTTTTCATCAAAACATTTTTCCAGTCTTATAGATAAATTAGTCCATACTGGTTGG TTAATCAAGAGCCCTTCGGGTTTGGTTGGTGGCATTAATGG	101
NM_031423	GCGGAAATGGGGCGGGACTTCCAGTAGGAGGCGGCAAGTTTGAAGTGTATGACGGTTGACGTTTGTGTAT TTTTGAATTTGCTGTAGCTGTCCCGAATCGCGGCTTCTGTGCGGCGCCGGCACTGTAGATTAAAC AGGAACTTCCAAGATGAAACTTTGTCTTCCCGAGATATAATGTAGCTGAGATTTGTATTCAATATTCCG CAATAAGATCTTAACAGGAGCTGATGGTAAAAACCTCACCAGAATGATCTTTATCCAAATCCAAAGCC	102

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
	<p>GAAGTCTTGCACATGATCTACATGAGAGCCTTACAAAATAGTATATGGAATTCGACTGGAACATTTTTTACA TGATGCCAGTGAACCTCGAAGTCATGATCCACATTTAATGGAAGGCTTCTTACCATTTCAGCAATTTAGT TACTCATCTGGACTCATTTTTGCGCTATCTGCGGGTGAATGACTTTGAGACTGCTGATATTTCTATGTC AAAGCAAAACGGACAAAGTCGGTTTTTAAAGTGGCATTTCAACTTTATTCACTTCAGAGAAGCATGCCGTG AAACCTATATGGAATTTCTTTGGCAATATAAATCCCTCTGCCGACAAAATGCAACAGTTAAACGCCGAC CCAGAGGCATTAATGAACTGGAGACTTTGATTTCTGTTCCAGTTGAAGAGCAAGAAGATTCAAGCAG CTTTCAGATGGAATTCAGGAGCTACAACAATCACTAAATCAGGATTTTCATCAAAAAACGATAGTGTCTG AAGAGGAAAATTCACAAAAGAACTCAAAATTTTCAGAGAAAACCAAGCGTTTGAATGAACTAAAATTTCT GGTGGTTTTCTTTGAAAGAAATACAGAGAGTTTGAACAAAATTTGTTGATTTCTCCAGAGAAGTTAAAG AATTATAAAGAAAAATGAAAGATACGGTCCAGAAGCTTAAAAATGCCAGACAAGAAGTGGTGGAGAAAT ATGAAATCTATGGAGACTCAGTTGACTGCCTGCCTTCATGTCAGTTGGAAGTGCAGTTATATCAAAAGAA AATACAGGACCTTTTCAGATAATAGGAAAAAATAGCCAGTATCTTAAAGGAGAGCCTGAACTTGGAGGAC CAATTTGAGACTGATGAGTCAAGAAATGAAAGACTGAAGAAAATTCGTTCAAAAGACTGATGA TTGTGAAGAAAGAAAAACTTGCCACAGCACAAATCAAAAATAAATAAGAAAGCATGAAGATGTTAAGCAATA CAAAACGCACAGTAATTCAGGATTGCAATAAAGTTCAAGAAAAAAGAGGTGCTGCTATGAACAGGTAACC ACAATTAATCAAGAAATCCAAAAATTAACCTTGAATTCACAACTAAAAGATGCTGCTGAAAAGGAGA AACTGAAGTCCAGGAAATTTTCAAACTTGAAGAACTGCTTTGGAGAAAATACCACCGGATTTGAAA GGCAGCAGAGACTCCTATGCTAAGATAGATGAGAAGACAGCTGAACCTGAAGAGGAAGATGTTCAAAATG TCAACCTGATTAACAAAATTCATGCTTTTTTGTAAATGGCTTGCCTCTTTTAAATTTCTATTAGAAA GAAAAGTTGAAGCGAATGGAAGTATCAGAAGTACCAAAATTAATGTTGGCTTCATCAGTTTTTATACACT CATAGTAGTTAATAAGATGAATTTAATGAGGCTTTTATTAATTTATAATTAATAAATAAATTCGCGAGCT ATTCATGCTCTACTCTGCCCTTGTGTAAATAGTTTGAAGTAAAACAAAACCTAGTTACCTTTGAAATAT ATATATTTTTTCTGTACTAIC</p>	
BC041846	<p>GGCTAGCGGGGAGGTGGAGAAAGAGGCTTGGGCGCCCGCTGTAGCCCGGTGTGGGAGGACGCACGGG CCTGCTTCAAGGCTTTGGGATAACAGCGCCTCCGGGGGATAATGAATGCGGAGCCTCCGTTTTCAGTCA CTTTCAGATGTGCTCCACTTTTTTCCGCTGTAGCCGCAAGGCAAGGAAACATTTCTCTTCCCGTACTGAG GAGGCTGAGGAGTGCATGGGGTGTCTTTTTCTCTTAACCCAGAACTGCGAGACAGAGGCTGAGTCCCT GTAAGAAACAGCTCCAGAAAAGCCAGGAGAGCGCAGGAGGGCATCCGGGAGGCGAGGAGGGTCTCGCTGG GCCTTCAACCGCACCCACATCGGTCCACCTGCGAGGGGGGGGACCTCTGTCGCTGAGCAATTCAGCA CCCACCTGCGCTCACCTGGCCCTCTCCCGCTGGCTCCCGGGGGCTGCGGTGCTCAAAGGGGCAAGAGCTG AGCGGAACACCGGCCCGCGCTGCGGCGAGCTGCTTCACCCCTCTCTCTGCGAGCCATGGGGTCCCTCGTG GACCTCTCGCTCTCTCTCTCTCTCTCCAGGTTTGTGGCTGCGAGTGGCGGCCCTCCGAGCGCTCCCGGG GGTCTTCAGGGAGGCTGAAGTGACCTTGGAGCGGAGGCGCGGAGCAGGAGCCCGGCGAGGCGCTGGGG AAAGTATTCATGGGTGCCCTGGGCAAGGCCAGCTCTGTTTAGCACTGATAATGATGACTTCACTGTGCG GGAATGGCGAGACAGTCCAGGAAAGAAAGTCACTGAAGAAAGGAAATCCATGAAAGATTTTCCCATCCAA ACSTATCTTACGAAGACACAAGAGAGATTGGTGGTGTCTCCATATCTGTCCCTGAAAATGGCAAGGTT CCCCTCCCCCAGAGACTGAATCAGCTCAAGTCTAATAAAGATAGAGACCAAGATTTTCTACAGCATCA CGGGCCGGGGGCAGACAGCCCTTGGGGTGTCTTCGCTGTAGAGAAAGGAGACAGCTGTTGTGTT GAATAAGCCACTGGACGGGAGGAGATTGCCAAGTATGAGCTCTTTGGCCACGCTGTGTTCAGAGAAATGGT GCCTCAGTGGAGGACCCCATGAACATCTCCATCATAGTACCCGACCAAGATGACCCACAGCCCAAGTTTA CCCAGGACACCTTCCAGGGAGTGTCTTAGAGGGAGTCTTACCAGGTACTTCTGTATGACAGTGAACAGC CACAGATGAGGATGATGCCATCTACACCTACAATGGGTGGTGTCTTACTCCATCCATAGCCAGAGACCA AAGGACCCACAGCCTCATGTTCACAATTCACCGGAGCACAGGCACCATCAGCGTCACTTCCAGTGGCC TGGACCGGAAAAAGTCCCTGAGTACACACTGACCATCCAGGCCACAGACATGGATGGGACCGCTCCAC CACCACGGCAGTGGCAGTAGTGGAGATCTTGTATGCCAATGACAAATGCTCCCATGTTTACCCCCAGAG TACGAGGCCATGTGCCTGAGAATGCAGTGGCCATGAGGTCAGAGGCTGACGCTCACIGATCTGGAGC CCCCCAACCTACCCAGCTGGCGTGCCACCTACCTTATCATGCGCGGTGACGACGCGGACCATTTTACCAT CACCACCCACCTGAGAGCAACCCAGGCGATCTGACAACAGGAAAGGTTTGGATTTTGGAGCCAAAAAC CAGCACACCTGTACSTTGAAGTGACCAACAGGAGCCCTTTTGTGCTGAAGCTCCCAACCTCCACAGCCA CCATAGTGGTCCACGTGAGGATGTGAATGAGGCACCTGTGTTTGTCCACCCCTCCAAAGTCGTTGAGGT CCAGGAGGGCATCCCACTGGGGAGCCTGTGTGTGTCTACTACTGCAAGACCCCTGACAAGGAGAATCAA AAGATCAGCTACCGCACTCTGAGAGACCCAGCAGGTTGGCTAGCCATGGACCCAGACATGGGAGGTCA CAGCTGTGGGCACCTTCGACCGTGGAGTGAAGGATTTGTGAGGAACAACAATCTATGAAGTCACTGGTCTT GGCCATGGACAATGGAAGCCCTCCCACTGGCACGGGAACCCCTCTGCTAAACTGATGATGTCAAC GACCATGGCCAGTCCCTGAGCCCGTCAATCAACATCTGCAACCAAAAGCCCTGTGGCCAGGTTGCTGA ACATCAGGACAAAGGACCTGTCTCCACACCTCCCTTTCCAGGCCAGCTCAGAGATGACTCAGACAT CTACTGGACGGCAGAGCTCAACGAGGAAGGTGACACAGTGGTCTTGTCCCTGAAGAGTTCTTGAAGCAG GATACATATGACGTGCACCTTTCTCTGCTGACCATGGCAACAAAGAGCAGCTGACGGTGAATCAGGGCCA CTGTGTGCGACTGCCAIGGCCATGTGCAAACTGCTGAGCCCTGAAAGGAGGTTTCACTCCCTCCCTGT GCTGGGGGTGCTCTGCTGCTGCTTCTCTGCTGCTGCTGCTTTTGTGTTGGTGAAGAAAGAGCGGAA ATCAAGGAGCCCTCTACTCCAGAAAGATGACACCCCTGACAACGCTCTCTACTATGGCGAAGAGGGGG GTGGCGAAGAGGACAGGACTATGACATCACCAGCTCCACCGAGGCTGGAGGCCAGGCGGAGGTTGGT TCTCCGCAATGACGTGGCACCACCATCATCCGACCCATGTACCCTCTAGCCAGCCCAACCCAGAT GAAATCGGCAACTTTAATAATGAGAACCTGAAGCGGCTAACACAGACCCCAACGCCCCCGCCCTACGACA CCTCTTGGTGTTCAGACTATGAGGGCAGCGGCTCCGACGCGCGTCCCTGAGCTCCCTACCTCTCTCCG CTCCGACCAAGACCAAGATTACGATTAICTGAACAGTGGGCGAGCCGCTTCAAGAAGCTGGCAGACTG TACGGTGGCGGGGAGGACGACTAGGCGGCTGCTGACAGGCTGGGGACCAACCTGAGGCCACAGAGCA TCTCAAGGGGTCTCAGTTCCCCCTCAGCTGAGSACTTCGGAGCTTGTCAAGAAAGTGGCGGTAGCAACT</p>	103

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
	TGCCGGAGACAGGCTAIGAGTCTGACGTTAGAGTGGTTGCCITCCTTAGCCTTTCAGGATGGAGGAATGTG GGCAGITTTGACTTCAGCACTGAAAACCCCTCCACCIGGGCCAGGGTGCCTCAGAGGCCAAGTTCCAGA AGCCCTTACCTGCCSTAAAAIGCTCAACCCCTGTSTCCTGGGCCCTGGCCCTGCTGTGACTGACCTACAGT GGACTTTCCTCTGGAAATGGAACCTTCTTAGGCCCTCCTGGTGCAACTTAATTTTTTTTTTAACTGCTATC TTCAAACGTTAGAGAAAGTTCTTCAAAGTGCAGCCACAGCTGCTGGGCCACTGGCCCTCCTGCATT TCTGGTTCCAGACCCCAATGCCTCCCATTCGGATGGATCTCTGCGTTTTTATACTGAGTGTGCTAGGT TGCCCTTATTTTTTATTTTTCCCTGTGCGTTGCTATAGATGAAGGGTGGAGCAATCGTGTATATGTAC TAGAACTTTTTTATAAAGAACTTTTCCCAAAAAAAAAAAAAAAAAA	
NM_016343	GAGACCAGAAGCGGGCGAATGGGCACCGGTGGCCGCTGCGGGCAGTTTGAATTAGACTCTGGGCTCCAG CCCGCCAAGCCGCGCCAGAACTGTACTCTCCGAGAGGTCGTTTTCCCGTCCCGCAGAGCAAGTTTATTT ACAAATGTTGGAGTAATAAAGAGGCCAGAACAAAATGAGCTGGGCTTTGGAAGAATGGAAAGAGGGGCTG CCTCAAGAGCTCTTCAGAAAAATCAAGAGCTTGAAGGACAGCTGACAAAATGAAGAGGAAAAAGCAGC AAAGGCAGITTCAGCTTIGACAGTCTCGAGGCTGCCTGCAGAAGCAAAAACAGAAGTTGAAAAAGAAAA AACCCGAGGTTACAAACCTGAAAAGGGAGAAATCAAAGATTGATGGAATATGAAAGTCTGGAGAAAACT AAGCAGAAGATTTCTCATGAACTTCAAGTCAAGGAGTCAACAAGTGAATTTCCAGGAAGGCAACTGAATT CAGGCAAAAAACAATAAGAAAACTGGAACAGGAATTAAGGCTGTAATCTGAGCTGAAAGAGCCCA ACAAGTGGCAGTCTGCAGATGTCTCTGTAATCCATGCAATACACCACAAAAAATTTTTTACAACCTCCA CTAACACCAAGTCAATATTATAGTGGTTCCAAGTATGAAGATCTAAAAAGAAAAATATAAIAAGAGGTTG AAGAACGAAAAAGATTAGAGGCAGAGGTTAAGGCTTGCAGGCTAAAAAGCAAGCCAGACTCTCCACA AGCCACCATGAATCACCGGACATTCGCCGGCATCAGGCTTCATCATCTGTGTTCTCATGGCAGCAAGAG AAGACCCCAAGTCACTTTCTCACTAATTTCTCAAAGAACTCCAATTAGGAGAGATTTCTCTGCCACTTACT TTCTGGGGAACAGAGGTGACTCCAAGTCCATCAACTTTGCAAAATAGGAAAAACAGATGCTAATAGCAG TTCTTTGACAATCTAGCAGTCTCACTTTTTGATCAATTAAGAGCGCAGAACTCAAGAGCTAAGAAAC AAGATTAAAGTGGAACTACGCCCTGCAAGGACATGAAAAAGAAATGAAAGGCCAAGTGAATAAGTTTCT AAGAATCCAACTCCAATGGAGAAAGCAAAAGTGAATTAATGAAAAAGAGAAAGTTTGAACAATG TAGGATGAATAGTGAACAACAGCACAATACGACAGGCGTCAACCAAGTATCTGCATTTGAAACG AAATGAAAAATTTGACGGAAAGTTTGAAGTTGTGACGGACAAAAAGCAGAAAGTCCAGATGTTCTCTGG AACAGAAAAATTAAGGAAAAAGAAAGGAGTTTCAAGAGGAGCTCTCCCGTCAACAGCGTCTCTCCAAAC ACTGGACCCAGGAGTGCATCCAGATGAAGGCCAGACTCACCCAGGAGTTACAGCAAGCCAGAAATGAC AACGTCCTGCAGGCTGAATGGATAAACTCACATCAGTAAAGCAACAGCTAGAAAAACAATTTGGAAGAT TTAAGCAAAAGTTGTGCAGAGCTGAACAGGCGTTCAGGCGAGTCAAGTCAAGAGAAATGAGCTGAGGAG AAGCATGGAGGAATGAAGAAAGGAAACAACTCCTTAAGAGTCACTCTGAGCAAAAGGCCAGAGAACTC TGCCACTGGAGGCAAGCTCAAGAACATCAAAACAGTGTAAAAAGCAGCCAGAAATTTGCAGAGAAAA TGAAGCGCAAGAAATCCTCTCAGGAAACCATGTTAAGAGATCTCAAGAAAAATAAATCAGCAAGAAAA CTCTTTGACTTTAGAAAAACTGAAGCTTGCTGTGGCTGATCTGGAAAAAGCAGCGAGATTTGTTCTCAGAC CTTTTGAAGAAAAAGAACATCACATGAACAATTAATGATAAGTTAAGCAAGCAGAGAAAGAGTCCA AAGCCTTGCTGAGTGTCTTAGAGTTAAAAAGAAAGAAATGAAAGAAATGAAAGAAAGAAAACTCTGTT TTCTTGTGGAAAGTTGAAACGAAAACTTTTAACTCAGATGGAATCAGAAAAAGGAAACTTTGAGAGAT AAAATTAATCACTTGGAACTTGCTTGAAGCACAGCAAAATAAAAAGTCAATGAATACAAACAGAGAGTAA GAACGCTGGAGATGGACAGAGAAAACTAAGTGTGAGATCAGAAAACTTCAACAAGCTGTAGACAGTAA GTCACTGGAGGTTAGACCCAGAACTAGCTTATATGAGCTACAGCAGAAAGCTGAGTTCTCAGATCAG AAACATCAGAAGGAAATAGAAAAATGIGTTTTGAAGACTTCTCAGCTTACTGGGCAAGTGAAGATCTAG AACACAAGCTTCAGTTACTGTCAAAATGAATTAATGGCAAAAGACCGGTGTTACCAAGACTTGCATGCCGA ATATGAGAGCCTCAGGGATCTGCTAAAATCCAAAGATGCTTCTCTGGTGACAAATGAAGATCAACAGAA AGTCTTTTGGCTTTTTGATCAGCAGCTGCGATGCACTTCTTTGCAAAATATAATGGAGAACAGGAA CCAATGCTTCAGAGAGGAGTGAATGTCGTTTAGAAGCAGACCAAGTCCGAAAAATTTGCCACTCTACA AAAATAGAGTTGATTCACTTGAATTTTCATTAGAGTCTCAAAAACAGATGAATCAGACTGCAAAAAGCAG TGTGAAGAGTTGGTGCAAAATCAAAGGAGAAATAGAAAGAAATCTCATGAAAGCAGAACAGATGCATCAAA GTTTTGTGGCTGAAACAAGTCAAGCTTGTAGTAAAGTTACAGGAAGACACTTCTGCTCACCAGAAATGTTG TGCTGAAACCTTAAAGTCCCTTGAGAACAAAGGAAAAAGAGCTGCAACTTTTAAATGATAAGGTAGAACT GAGCAGGACAGATTCAGAAATTAAGAAAGGCAACCTTCTACTGAAAGACTCTCTAAAGGAGCTACAAC TTTTATCCGAAACCTTAAGCTTGGAGAAGAAAGAAATGAGTTCCATCATTTCTCTAAATAAAAGGAAAT TGAAGAGCIGACCCAGAGAAATGGCACICTTAAAGAAATTAATGCATCTTAAATCAAGAGAGATGAAC TTAATCCAGAAAGTGAAGTTTTCGAAACTATATAGATGAAAGGGAGAAAAAGCATTTCCAGAGTTATCTG ATCAGTACAAGCAAGAAAAAATTTATTTTACTACAAGATGTGAAGAAACCCGAAATGCAATAGAGGATCT TAGTCAAAAATACAAAGCAGCAGGAAAAAGAAATTTCAAATAGAAATGCTTGTAAATGAATGCACTAGT CTTTGTGAAAAATAGGAAAAATGAGTTGGAACAGCTAAAGGAAGCATTTGCAAAAGGAAACCAAGAAATCT TAACAAAAATAGCAATTTGCTGAAGAAAGAAATCAGAATCTGATGCTAGAGTTGGAGACAGTGCAGCAAGC TCTGAGATCTGAGATGACAGATAACCAAAACAAATCTAAGAGCGAGGCTGGTGGTTTAAAGCAGAAATC ATGACTTTAAAGGAAAGCAAAAACAAAATGCAAAAGCAAGTTAAAGACTTATTAACAAGAAATGAACAG TGATGAAGGTAATGAAGACTAAACATGAATGCAAAAATCAGAAACAGAACCAATTAGGAACCTCTGTGAA AGAAAGAGAGAGTGAAGAAATCAATGTAATTTTAAACCTCAGATGGATCTTGAAGTTAAAGAAATTTCT CTAGATAGTTATAATGCCAGITGGTCAATTAGAAGCTAIGCTAAGAAATAAGAAATTAACCTCTCAGG AAAGTGAAGAGGAAAGGAGTGCCTGCAGCATGAATACAGACAATTAGAGGAGATCTTGAACCCAGCAA TTTGCAAGACTGCACTCACAAGAAATTAGTGGCTTAAAGACTGTGAATAGATCGGGAAGAAAGAT ATTTCAAGGCTCATGAGTTGTCACAAGTCAAAAGCAATGCAACCTTCACTGCTCTGCAAAACAA CAATGAACAAGCTGAATGAGCTAGAGAAATATGTGAAATCTGCAGGCTGAAAAAGTGAACCTGTAAC TGAGCTGAATGATTCAGAGTCAAGATGATCACAGCAACTAGGAAAAATGGCAGAAAGGATAGGGAACCTA	104

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25 30 35 40 45 50 55	CTAAATGAAGTTAAATATTAATGATGACAGTGGTCTTCTCCATGGTGAAGTTAGTGGAGACATACCAG GAGGTGAATTTGGTGAACCAACCAAAATGAACAGCACCCCTGTCTCTTTGGCTCCATTGGACGAGAGTAATTC CTACGAGCCTTGACATTTGCAGACAAAGAAGTTCAAATGCACCTTTGCCAAATGGCAAGAGAAATTTCTTA TCTTTACAAAGTGAACACAAAATTTTACATGATCAGCACCCTGCAGATGAGCTCTAAATATGCAGAGCTGC AGACCTATGTTGACATTAATAAGGCCGAAATTTGGTCTTTGCAACGAAATCGAGAAACTTTCAAGGTGA CTTGGTGAAGGACATGCAGCTGGCTTGGAGGAGGCTCTTCCATCCCTGTCTATCTCTTGTGTGCTT GACAGCTCTAGTCTTAGCAGTTTGGGAGACTCCTCCTTTTACAGAGCTCTTTAGAACAGACAGGAGATA TGTCTCTTTTGGTAAATTTAAGAGGGGCTCTTTACAGCAACAGTGCAGTGTAGATGAAGTATTTTGCAG CAGTCTGCAGGAGGAAATCTGACCAGGAAAGAAACCCCTTCGGCCCCAGCGAAGGGTGTGAAGAGCTT GAGTCCCTCTGTGAGGTGTACCGGAGTCCCTCGAGAAGCTAGAAGAGAAAATGGAAGTCAAGGGATTA TGAAAAATAAGGAAATTCAGAGCTCGAGCAGTTTAAAGTTCTGAAAGGCAAGAGCTTACTGCCITAG GAAGCAGTATTTGCAGAAAATGAAACAGTGGCAACAGAAAGCTGCAAGCCCTGACTCTGAGATGGAGTCC AAGTTGGCGGCGAGAAAAGAAACAGACGGAACAACCTGCTACTTGAGCTGGAAGTAGCAGACTCCAGTAC AAGGTCGGACTTAAAGTCTCGGCTTGGCTTGGCATCGACACAGAAGTGTATTCAGGCGCAAGATGA GAGCTGAGCATATCAAAAAGAACATACTTCAGAAAATACAGAAAAGAACCAAGCAATGATGTTCAICAG ATTTGTTGATAAAGATGCTCAGCAGGACCTCAATCTAGACATTTGAGAAAATAACTGAGACTGGTGCAGTGA AACCCACAGGAGAGTGTCTTGGGAAACAGTCCCGCAGATACCAATTAAGAGCTCCAGGGAGATAGAAAC CCAGGGCTCTTCAGAAATGCATTTCTGAATTTGTCATTTCTGGTCTAATGCTTTGGTACCTATGGATTT CTGGGGAATCAGGAAGATACTCAATACTTCAACTGCGGGTAAAAGAGACATCAAAATGAGAAATTTGAGAT TACTTCATGTGATAGAGGACCGTGCAGAAAAGTTGAAAGTTTGTAAATGAAATGAAAGAAATTAGACTC AAAACTCCATTTACAGGAGGTACAACATAAGCAAAAATTAAGCAATGCTAGAAATGGAAAAAATAGTT GGGAACTTAAGAAAGAAAATCAGATTTAAGTGA AAAATTTGGAATATTTTTCTTGTGATCCAGGAGT TACICAGAGAGTAGAAACTCTGGAAGGCTCAATTTCTGATTTAGAAAATGCATGCAGATAAATCATCAG TGAAGATATGGAGATAATGTGGCCAAAGTGAATGACAGCTGGAAGGAGAGATTTCTTGTGTGGAAAAAT GAGCTGAGTAGATCAGATCGAGAAAAGCTAGCATTGAGCATGAAGCCCTCTACCTGGAGGCTGACTTAG AGGTAGTTCAAACAGAGAAAGCTAAGTTTGA AAAAGACAAATGAAAAAAGCAGAAAGTTATTTGCTGCT TGAAGAAGAACTCTCAGTGGTCAACAGTGCAGAAAACAGCTTCTGTGGAGAAATAGATACTATGTCAAAA AAAACCAGGCACTGGATCAGTTTGTCTGAAAAAATGAAGGAGAAAACACAGAGCTTGGATCTCATCAAA GTGAGTGTCTCCATTTGATTTGAGTGGCAGGACAGAGTGAAGGAAAAGCAGGAACTCCTTCAGACTTT GTCTCTGATGTGAGTGAAGCTGTTAAAAGACAAAATCACTCCAGGAAAAGCTGCAGAGTTTGGAAAAAG GACTCAGAGGCACTGTCTTGCAGAAAATGTGAGCTGGAAAACCAAAATGCACAACTGAAATAAGAGAAAAG AATTTGCTTGTCAAGGAAATCTGAAAGCTGAGGACAGACTGAGTGAATCAGATATGAAAAGCTGAAATGT CTCCAAAGGCTTTGGAGGCGGACITGGTGGAGAAAGGTGAGTTCCGCAITGAGGCTGAGCTCAACAGGAG GAAGTGCATCAGCTGAGAAGAGGCTCGACAAAATGAGAGTTCCGCAITGAGGCGCATGAAAAGAGAGCGC TGCACATCGCAGAGAAACTGAAAGAAACGCGAGCGGAGAAATGATTCATTAAGGATAAAGTTGAGAACT TGAAAAGGAAATTCAGAGTGTGAGAAAAGAAACAGGAGCTAGTGAATTTGATGCGGAGAAATTCCAAAGCA SAAGTAGACACTCTAAAACCAAAATGAAAGAGATGGCCAGAAAGCTGAAAGTTTTGAAATAGACCTTTC TCAGCTTAAGGCTGAAAAGAAATCTGACAAAACAAATCAAGAAAACAAAGGTCAGTTGTGAGAACT AGACAAGTTACTCTCTTCAATTAAGAGTCTGTGAGAAGAAAAGGAGCAAGCAGAGATACAGATCAAGAGAA GAACTAAAATGCACTGAGAGTGTCTCACAATCAGTTAAAGGAGCTAAATGAGGAGTAGCAGCCCTGTG GTGGTGCACCAAGAAATATGAAGGCCACAGAACAGAGTCTAGACCCCAATAGAGGAAGAGCATCAGCT GAGAAATAGCATTTGAAAGCTGAGAGCCCCCTTAGAAGCTGATGAAAAGAGCAGCTCTGTGCTTACAA CAACTGAAGGAAAGTGCAGCATGCAAGATTTACTTAAGGTTAGAGTGGAGAACTTGAAGAGAGAGTGTAG AGATAGCCAGGACAAAACAGAGCATGCACCTCTTGGAGCAGAGAAATTCCAAAGGAGGAGTAGACCCCT AAAAGCAAAAATAGAGGGATGACCCAAAGCTGTGAGAGTCTGGAAATAGATGTTGTTACTATAAGGTCA GAAAAAGAAAATCTGCAGAAATGAATTACAAAAAGAGCAAGAGCGAAATCTGAAATAGAAAATAAATAAT CATCATTTGAAAATATTTTGCAGAAAAGAGCAAGAGAAAAGTACAGATGAAAAGAAAATCAAGCAGCTGC CATGGAGATGCTTCAACCAAAATTAAGAGAGTCAATGAGAGATGGCAGCCCTGCATAATGACCAAGAA GCCITGAAGCCAAAGAGCAGAAATCTAGTAGTCAAGTAGAGTGTCTTGAATTTGAGAAGGCTCAGTTGC TACAAGGCTTGTGAGGCCCCAAAATAATATATTTGTTTTGCAATCTCAGTGAATGGCCCTCAITCAAGA AGTAGAAGATGGCAAGCAGAAACTGGAGAAAGAGGATGAAGAAATCAGTACTGAAAATCAAAATCAAA GACCAAGAGCAGCTTGTCTTAAACTGTCCAGGTTGGAAGGAGAGCCCAACTTTGGAAGGAGGCAAAACI TAGAATGAGAAATCTGACAGTGGAAATGGAGCAGAAATCCAAGTGTCAAAATCCAAAATGCTCTTT GCAGGACACATTAAGAGTGTGAGAGTCTTACAAGAACTTAGAGAAATGAGCTTGAATGACAAAAATG GACAAAATGCTCTTTGTTGAAAAGTAAACAAAATGACTGCAAGGAAACTGAGCTGCAGAGGAAAATGC ATGAGATGGCAGAAAACAGCAGAGCTGCAAGAAAGAACTCAGTGGAGAGAAAATAGGCTAGCTGAGAGA GTTGCAGTTACTGTTGGAAGAAAATAAGAGCAGCAAAAGTCAATTAAGGAGCTCACACTAGAAAA TAGT GAAITGAAGAAGAGCTTAGATTCATGCACAAAAGACCGGTTGAAAAGGAAAGGAAAGTGAAGAGGAAAA TAGCTGAATATCAGCTACGGTCTCATCAAGCTGAAAAGAAAACAGGCTTTGCTTTTGCACACAAAACAA ACAGTATGAAGTGA AAAATCCAGACATACCCAGAGAAATGACTTCTAAAAGAAAGTGTCTCAGTTACAG AAGCTGGAGATAGACTTTTAAAGTCTAGTAAAAGAGGCTCAATAATTCATGAAAAGCTACTACTCAGA TTTTGGAGAATTAAGAAAACCAAGATGCACAAATCAAAAATGTAATCAGTTCAAGAAAGGAAAATGA ACGTCCCGAGGGA AAAATGAAGTTGTGATCAAACTCTGTAACAGCTGGAAGGAAAAGGAGATACTG CAGAAAAGAACTCTCTCAACTCAAGCTGCACAGGAGAAAGCAAAAACAGGTACTGTTATGGATACCAAGG TCGATGAATTAACAACCTGAGATCAAGAACTGAAAAGAACTCTTGAAGAAAACCAAGGAGGAGATGA ATACTTGGATTAAGTACTTCTCTTGCITTAAGCCATGAAAAGTGAAGAAAGCTAAAAGATGTTAGAG ACACAAGTGGCCATCTGTGTTACAGCAATCTAAAACAAGATTCAGGAGGTTCTCTTTGTAGGTCCAG TTGTTCCAGCACCATCTCAATCCCTTCTCTTACTGAAAAGAGTTATCATCTGGCCAAAATAAGCTTC	

60

65

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
	AGGCAAGAGGCAAAGATCCAGTGGAAATATGGGAGAATGGTAGAGGACCAACACCTGCTACCCACAGAGAGC TTTTCTAAAAAAGCAAGAAAGCAGTCATGAGTGGTATTCACCCCTGCAGAAAGACACGGAAAGGTACTGAGT TTGAGCCAGAGGGACTTCCAGAAGTTGTAAGAAAGGGTTTGCTGACATCCCGACAGGAAAGACTAGCCC ATATATCCTGCGAAGAAACAACCATGGCAACTCGGACCCAGCCCCCGCTGGCTGCACAGAAGTTAGCCGCTA TCCCCACTGAGTCTCGGCAAGAAAACTTGCAGAGTCTCCAAACCAACAGCTGGTGGCAGCAGATCAC AAAAGGTCAAAGTTGCTCAGCGGAGCCAGTAGATTCAGGCACCATCTCCGAGAACCCACCAGAAATC CGTCCAGTCAATAATCTTCCGAGAGAAGTCCGACTGACAGCCCCAGAGAGGGCTGAGGGTCAAGCGA GGCCGACTTGTCCCGAGCCCAAGCTGGACTGGAGTCCAACGGCAGTGAAGACTGTAAGGTCCAGTGAA GGCACCTTTGTGTGTCAGTACCCCTGGGAGGTGCCAGTCATTTGAAATAGATAAGGCTGTGCCCTACAGGACTT CTCCTTAGTCAGGGGCACTTTTATTAGTAGGAGAAAACAATTCCTTAGAAGCTTTAAAATATACTGTACT CTTTAGATCTCCCATGTGTAGTATGAAAAGTTTGGAAAGCACTGATCAGCTGTITAGCATTGCCATTCC TCTACTGCAATGTAAATAGTATAAAGCTATGTATATAAAGCTTTTGGTAATATGTTACAATTAATAATGA CAAGCACTATATCACAATCTCTGTTTGTATGTGGTTTTTACACTAAAAAATGCAAAAACACATTTTATTC TTCTAATTAACAGCTCTAGGAAAAATGTAGACTTTTGTCTTATGATATTTATCTGTAGTATGAGGCATG GAATAGTTTTGTATCGGAAATTTCTCAGAGCTGAGTAAAAIAGAAGAAAGCAATGTTATGTTTTTTAAG GAAAAATGTGCACACATATACATGTAGGAGTGTATCTTTCTCTACAACTGTGTTTAGACATCTTTGCT TATGAAACCTGTACATATGTGTGTGGGTATGTTTATTTCCAGTGAGGGCTGCAGGGCTTCCTAGAGG TGTGCTATACCATGCGTCTGTGCTGTGTTTTTCTGTTTTTAGACCAATTTTTITACAGTCTCTGGTA AGCATTGTGCTATCTGCTGATGGATTAACATATAGCCTTTGTTTTCTAATAAAAATAGTCGCTCTGGTTTT CTGTAATAAAAAAAAAAAAAAAAAAAAAA	
AB091343	GGCACGAGGGCCGACCGGAGGCGCGCTTCGCTTCAGCTGCTAGCTGGCCCAAGGGAGGCGACCGCGG AGGGTGGCAGGGCCGCCAGGACCCCGAGCCCCGGGGCCGGTCCGGACCGCCAGGGAGGGCCAGG TCAGTGGGCGAGTCCGCTCCCGGGATTCATCTCTCCCGCTCTGATAACAGTCCCTTTCCCTGGCCCT CACTTGTGCTGGCACCCGGCTGGGCGCTCAAGACCGTTGCTCTTCGATCGCTTCTTTGGACTTGGC GACCATTTCAGAGATGCTTCCAGAAGTACCAAGATTTAATTAAGTAAAGTGGGATCGAAGCTAGT AACTCCAAATCCGAAACTACATTAAGAAAAATTAAGCGGAGAAATGCACACTTAAAGACATCAGTGGATG AAATCACAAGTGGGAAAGGAAAGCTGACTGATAAAGAGAGACACAGACTTTGGAGAAAAATTCGAGTCT TGAGGCTGAGAAGGAGAAATGCTTATCAACTCACAGAGAAGGACAAGAAATACAGCGACTGAGAGAC CAACTGAAGGCCAGATATAGTACTACCGATTGCTTGAACAGCTGGAAGAGACAAACGAGAGAAAGGAGAAA GGAGGGAGCAGTGTTCAGAGCCTTATCTGAAGAGAAAAGCAGTATGAAAACAACAGTTGCTGTGCAAC CTCAGCAATGCTGAACCTGAAAGCAAACAATACACTCCGTTTATCACAGACTGTGGCTCCAAACTGC TTCAACTCATCAATAAATAATATTCAATGAAATGAAATACAGCTGAAAGATGCTCTGGAGAAAAATCAGC AGTGGCTCGTGTATGATCAGCAGCGGGAAGTCTATGTAAGAGGACTTTTAGCAAGATCTTTGAGTTGGA AAAGAAAACGGAAACAGCTGCTCATTCACTCCACAGCAGACAAAAAGCCTGAAATCAGAAGGTATCTT CAAGAAGAGAAGCAGAAATGTACAAAGTCTCTTGGCAAGTCAAAAAAAGATCTTGAGGTTGAACGAC AAACCATAACTCAGCTGAGTTTTGAAGTGTGAAATTCGAAGAAAAATGAAAGAAACCCAAAAAGAGT TCACAAATTAATCAGCTGTGTTGATTCACAAGAAAGGGCAGATGCAACATCTGGAAGATGATAGGCAT AAAACAGAGAAGATACAAAACTCAGGGAAGAGAAATGATTTGCTAGGGGAAAACCTTGAAGAAGAGAAGA AGAGATCCGAAGAGCTCTTATCTCAGGTCCAGTTTCTTTACACATCTCTGCTAAAGCAGCAAGAAGAACA AACAAAGGTAGCTCTGTTGGAACAACAGATGCAGGCATGTACTTTAGACTTTGAAAAATGAAAAACTCGAC CGTCAACATGTGCAGCATCAATGCAATGTAATTTCTAAGGAGCTCCGAAAAGCAAGAAATCAAAATAACAC AGTTGGAATCTTTGAAACAGCTTCAATGAGTTTGCATCACAGAGCCATTAGTCACTTTCCAAAGGAGAGAC TGAAAAACAGAGAAAAAGTTGCCGCTCACCAAAAAGTCCCCTGCTGCACTCAATGAAAGCCTGGTGGAA TGTCCCAAGTGAATATACAGTATCCAGCCACTGAGCATCGCGATCTGCTTGTCCATGTGGAATCACTGTT CAAAGTAGCAAAATAAGTATTTGTTTTGATATTAAGATTCAACTACTGATTTTTCTGTIAGCTTGTGGG CATTTTTGAATATATATTTACATTTTGCATAAAACTGCCATCTACCTTTGACACTCCAGCATGCTAGT GAATCATGTATCTTTTAGGCTGCTGTGCATTTCTCTTGGCAGTGATACTCCCTGACATGGTTCATCATC AGSCTGCAATGACAGAAATGTGTTGAGCAGCGTCTACTGAGACTACTAACATTTTCCACTGTCAAATACT TGGTGAAGAAAAGATAGCTCAGGTTATTTGCTAATGGGTTAATGCACCAGCAAGCAAAATATTTTATGTTT TGGGGTTTTGAAAAATCAAAGATAATTAACCAAGGATCTAACTGTGTTGCAATTTTTTATCCAAAGCACT TAGAAAACCTACAATCTCAATTTTGTATGCTCAATTTGTTAAGAGGTTGGTATAGATACTATTTTTTTTICA TATTTGTATAGCGGTTATTAGAAAAGTTGGGGATTTTTCTGTATCTTATTGCTGCTTACCATTGAAACTTA ACCAGCTGTGTTCCCAACTCTGTTCTGCGCACGAAACAGTATCTGTTTGAAGCATAAATCTTAAAGTGC CACACACAATGTTTTCTTATGTTATCTGGCAGTAACTGTAACCTGAAATACATTAGCACATCTGTGTT AGCTAAAAATGTTAAAATAAACTTTAATAAACCAATGTAGCCCTCTCATTTGATTGACAGTATTTTACTT ATTTTTGGCAATCTTAAAGCTGGGCAATGTAATGATCAGATCTTTGTTTTGTCTGAACAGGATTTTTATA CATGCTTTTTGTAACCAAAAACCTTTAAATTTCTTCAGGTTTTCTAACATGCTTACCCTGGGCTACTG TAAATGAGAAAAGAAATAAAATATTTAATGTTTTAAAAAATAAAAAA	105

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_005228	CCCCAGGCGAGCGCGCCGCGAGCAGCCCTCCGCCCCCGCAGCGGTGTGAGCGCCGACGCGCGCCGAGGCGG CCGGAGTCCCGAGCTAGCCCCGGCGCGCCGCCCGCCAGACCGGACGACAGGCCACCTCGTCGGCGTCC GCCCGAGTCCCCGCCCTCCGCCGCAACGCCACAACCACCGCGCACGGCCCCCTGACTCCGTCAGTATIGA TCGGGAGAGCCCGAGCGAGCTCTTCGGGGAGCAGCGATGCGACCCCTCCGGGACGCCCCGGGCGAGCGCTCC TGGCGCTGCTGGCTGCGCTCTGCCCGGCGAGTCCGGCTCTGGAGGAAAAGAAAGTTGCCAAGGCACGAG TAACAAGCTCAGCCAGTGGGCACTTTGAAGATCATTTTCTCAGCCCTCCAGAGGATGTICAAATAACTGT GAGGTGGTCTTGGGAATTTGGAATTACCTATGTGACAGGAAATATGATCTTTCCTTCTTAAAGACCA TCCAGGAGGTGGCTGTTATGTCCTCAITGCCCTCAACACAGTGGAGCGAATTCCTTTGGAACCCCTGCA GATCATCAGAGGAAATATGTACTACGAAAATTCCTATGCCITAGCAGTCTTATCTAACTATGATGCAAAAT AAAACCGGACTGAAGSAGCTGCCCATGAGAAATTTACAGGAAATCCTGCATGGCCCGTGGGTTACGCA ACAACCCCTGCCCTGTGCAACGTGGAGAGCATCCAGTGGCGGGACATAGTCAGCAGTACITTTCTCAGAAC CATGTCGATGGACTTCCAGAACCCTGGGCAGCTGCCAAAAGTGTGATCCAAAGCTGTCCCAATGGGAGC TGCTGGGGTGCAGGAGAGGAACTGCCAGAACTGACCAAAATCATCTGTGCCAGCAGTGTCCCGGGC GCTGCCGTGGCAAGTCCCCAGTGTGCTGCCAACAACAGTGTGCTGCAGGCTGCACAGGCCCCCGGGA GAGCGACTGCCCTGGTCTGCCGCAAAATCCCGAGACGAAGCCAGTGCAGGACACCTGCCCCCACTCATG CTCTACAACCCCAACAGTACAGATGGATGTGAACCCCGAGGGCAAAATACAGCTTTGGTGGCCACTCGG TGAAGAAGTGTCCCCSTAATTTATGTGGTGCAGATCAGGCTCGTGCCTCCGAGCCTGTGGGCGGACAG CTATGAGATGGAGGAACGCGGCTCCGCAAGTGTAAAGAAGTCCGAAGGGCCCTGCCGCAAAAGTGTGTAAC GGAATAGGTATTGGTGAATTTAAAGACICACTCTCCATAAATGCTACGAATATTAACACITCAAAAAT GCACCTCCATCAGTGGCGATCICCACAICCTGCCGGTGGCATTITAGGGGTGACTCCTTCACACATACICC TCCTCTGGATCCACAGGAACGGATATCTGAAAACCGTAAAGGAAATCACAGGGTTTTTIGCTGATTCAG GCTTGGCCGTGAAAACAGGACGGACCTCCATGCCCTTIGAGAACCTAGAAATCATACCGGGCAGGACCAAGC AACATGGTCAGTTTTCTCTTGCAGTGCICAGCCTGAACATAACATCTCTGGGATTACGCTCCCTCAAGGA GATAAGTGTGAGATGTGATAATTTACGAAACAAAAATTTGTGCTATGCAAAATACAAATAAACTGGAAA AAACCTGTTGGGACCTCCGGTCAGAAAACCAAAATATAAGCAACAGAGGTGAAAACAGCTGCAGGCGCA CAGGCCAGTCTGCCATGCCCTGTGCTCCCGGAGGGCTGCTGGGGCCCGAGCCAGGGACTGCGTCTC TTGCCGAATGTGAGCCGAGGCAAGGAAATGCCGTGACAAAGTGCACCTTCTGGAGGTGAGCCAGGGAG TTTGTGGAGAACTCTGAGTGCATACAGTCCACCCAGAGTGCCTGCCTCAGGCCATGAACATCACCTGCA CAGGACGGGGACAGACAACTGTATCCAGTGTGCCACTACATGACGGCCCCACTGCGTCAAGAACCTG CCCGGCAGGAGTCAITGGAGAAAACAACACCCCTGCTCTGGAAGTACGACAGCCCGGCCATGTGTGCCAC CTGTGCCATCCAAACTGCACCTACGGATGCCTGGCCAGGTCTTGAAGGCTGTCCAAACGAATGGGCCATA AGATCCCGTCCATCGCCACTGGGATGGTGGGGCCCTCCTCTTGTGCTGCTGGTGGTGGGCCCTGGGGATCGG CCTCTTCAITGGCAAGCGCCACATCGTTCGGAAGCCACGCTGCCGAGGCTGCTGCAGGAGAGGGAGTT GTGGAGCCCTTACACCCAGTGGAGAAGCTCCCAACCAAGCTCTCTTGAGGATCTTGAAGGAAACTGAAT TCAAAAAGATCAAAGTCTGGGCTCCGGTCCGTTCCGGACCGGTGTATAAGGGACTCTGGATCCAGAAAGG TGAGAAAGTTAAATTCCTCGTCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAAACAGGAA ATCCTCGAIGAAAGCTACGTGATGGCCAGCGTGGACAACCCCAAGTGTGCGCCCTGCTGGGCACTCTGCC TCACCTCCACCGTGCAGCTCATCAGCAGCTCATGCCCTTCGGCTGCCTCCGGACTATGTCGGGAAACA CAAAGCAAATATTGGCTCCAGTACCTGCTCAACTGGTGTGTGAGATCGCAAGGGCAATGAACACTGTG GAGGACCGICGCTTGTGACCCGCACTGGCAGCCAGGAACGTACTGGTGAACACCCGACGATGICA	107

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25 30 35 40	GAACCCGCACCAAGCTCTGCTCCACACTGCCAACCGGCCAGAGGACGAGTGTGTGGCGAGGGCCTGGCC TGCCACCAGCTGTGCCCGGAGGGCACTGCTGGGSI CCAGGGCCCCCAGTGTGTCAACTGCAGCCAGT TCCTTCGGGGCCAGGAGTGGGTGGAGAAATGCCGAGTACTGCAGGGGCTCCCCAGGGAGTATGTGAATGC CAGGCACCTGTTTGGCGTGGCCACCTGAGTGTGAGCCCCAGAATGGCTCAGTGAACCTGTTTGGACCGGAG GCTGACCAGTGTGTGGCCTGTGCCACTATAAGGACCCCTCCCTTCTGGTGGCCGCTGCCCGAGCGGTG TGAACCTGACCTCTCTACATGCCATCTGGAAGTITCCAGATGAGGAGGGCCGATGCCAGCCTTGCCC CATCAACTGCACCCACTCCTGTGTGGACCTGGATGACAAAGGGCTGCCCGCCGAGCAGAGAGCCAGCCCT CTGACGTCCATCATCTCTGGCGTGGTGGCATTCTGTGTGGTCTGTGGTCTTGGGGTGGTCTTGGGAICC TCAITCAAGCGACGGCACCAGAAGATCCGGAAGTACACGATGCCGAGACTGTGTCAGGAAACGGAGCTGTG GGAGCCCTGACACCTAGCCGAGCCGATGCCAACCCAGGCCAGATGCCGATCCTCAAAGACAGCCGAGCTG AGGAAGGTGAAGGTGCTTGGATCTGGCGCTTTGGCACAGTCTACAAGGGCATCTGGATCCCTGTATGGGG AGAATGTGAATTTCCAGTGGCCATCAAAGTGTGAGGAAAAACACATCCCCAAAGCCAAACAAAGAAAT CTTAGACGAAGCATACCTGATGGCTGGTGTGGGCTCCCCATATGCTCTCCCGCTTCTGGGCATCTGCCCTG ACATCCACGGTGCAGCTGGTGCACAGCTTATGCCCTATGGCTGCCCTTAGACCATGTCCGGGAAAAACC GCGGACGCCCTGGGCTCCAGGACCTGCTGAACTGGTGTATGACAGATGCCAAGGGGATGAGCTACCTGGA GGTGTGGGCTCGTACACAGGGACTTGGCCGCTCGGAACGTGCTGGTCAAGAGTCCCAACCATGTCAAA ATTACAGACTTCCGGCTGGCTGGCTGCTGGACATGACGAGACAGAGTACCATGAGATGAGGGGCAAGG TGCCCATCAAGTGGATGGCGCTGGAGTCCATCTCCGCGGGCGGTTACCCACCAGAGTGTGATGTGGAG TTAGGTGTGACTGTGTGGGAGCTGTGACTTTTGGGGCCAAACCTTACGATGGCATCCAGCCCGGGGAG ATCCCTGACCTGCTGAAAAAGGGGAGCGGCTGCCCGAGCCCGCCATCTGCACCATTTGATGTCTACATGA TCAITGGTCAAATGTGGATGATGACTCTGAATGTGGCCAAAGATCCCGGGAGTGGTGTCTGAATTTCTC CCGCATGGCCAGGACCCCCAGCGCTTGTGGTCAITCCAGAAAGAGGACTTGGGCCAGCCAGTCCCTTG GACAGCACCTTCTACCCCTCACTGCTGGAGGACGATGACATGGGGGACCTGGTGGATGCTGAGGAGTATC TGGTACCCCGAGCAGGGCTTCTCTGTGCCAGACCTGCCCGGGCGCTGGGGGCAIGTGTCCACCACAGGCA CCGCAGCTCATCTACCAGGAGTGGCGTGGGGACCTGACACTAGGGCTGGAGCCCTCTGAAGAGGAGGCC CCCAGGTCTCCACTGGCACCCCTCCGAAGGGGCTGGCTCCGATGATTTGATGGTACCTGGGAATGGGG CAGCCAAAGGGGCTGCAAGGCTCCCCACACATGACCCAGCCCTCTACAGCGGTACAGTGAAGACCCAC AGTACCCCTGCCCTCTGAGACTGATGGCTACGTGCCCGCTGACCTGCAGCCCGAGCCTGAATATGTG AACCCAGCCAGATGTTCCGGCCAGCCCTTCCCGCCGAGAGGGCCCTCTGCCCTGTGCCCGACCTGCTG GTGCCACTCTGGAAGGCCCAAGACTCTCTCCCGAGGGAAGAAATGGGGTCTGTCAAAGAGCTTTTTGCCCT TGGGGGTGCCGTGGAGAACCCCGAGTACTTGCACCCCGAGGAGGAGCTGCCCTCAGCCCCACCCCTCT CCTGCCCTCAGCCAGCCTTCGACAACTCTATTACTGGACCCAGGACCCACAGAGCCGGGGGCTCCAC CCAGCACCTTCAAAGGACACCTACCGCAGAGAACCCAGAGTACCTGGGTCTGGAGCTGCCAGTGTGAAC CAGAAAGGCCAAGTCCGAGAAAGCCCTGATGTGCTCAGGGAGCAGGGGAGGCCCTGACTTCTGCTGGCAT CAAGAGGTGGGAGGGCCCTCCGACCACTTCCAGGGGAACCTGCCATGCCAGGAACCTGTCTTAAGGAACC TTCTTCTGCTGTGAGTTCAGATGCTGGAAGGGTCCAGCCCTGTTGGAAAGGAAACAGCACCTGGGG AGTCTTTGGGATCTGAGGCCCTGCCCAATGAGACTCTAGGGTCCAGTGGATGCCACAGCCAGCTTGG CCCTTCCCTCCAGATCCTGGTACTGAAAGCCTTAGGGAGCTGGCTGAGAGGGGAAAGCGGCCCTAAG GGAGTGTCTAAGAACAAAAAGCCACCCATTCAGAGACTGTCCCTGAAACCTAGTACTGCCCCCCATGAGGA AGGAACAGCAATGTTGTCAGATATCCAGGCTTGTACAGAGTGTCTTCTGTTTAGTTTTTACTTTTTTTT TTTTGTTTTTTAAAGATGAAATAAAGACCCAGGGGAGAAATGGGTGTGTATGGGAGGCCAAGTGTGGG GGGTCTTCTCCACACCCACTTTGTCCATTTGCAAAATATATTTTGGAAAAACAGCTA	
40 45 50 55 60	NM_001122742 ATGGTCAATAACAGCCTCCTGTCTACGACTCAGAACGGATTTTACCAAACTGAAAATGCAGGCTCCATG CTCAGAAGCTCTTTAACAGGCTCGAAAAGTCCATGCTCCTTTCTCCTGCCATCTCTATAGCATAAAGAA CAGTCTCTGAGTGATAATCTTCTCTCAAGAAGAAGAAAAGTAGGAAGGAGTAAGCACAAAGATCTCTTC ACATTTCCGGGACTGCGGTACCAAAATCAGCACAGCCTTCTTGAAGAAGGATGATAGATTTTAAATCTG AACTTTGAACCATCACTGAGGTGGCCCGCGGTTTCTGAGCCTTCTGCCCTGCGGGACACGGTCTGCAC CCTGCCCGGGCCACGACCATGACCATGACCCCTCCACACCAAAAGCATCTGGGATGGCCCTACTGCATCA GATCCAAGGAACGAGCTGGAGCCCTGAAACCGTCCCGAGCTCAAGATCCCCCTEGAGCGGCCCTTGGG GAGGTGTACTGGACACGACGCAAGCCCGCGGTGTACAAC TACCCCGAGGGCGCCCTACGAGTCTCAAG CCGCGCCCGCCCAACCGCGAGGCTACCGTACAGCCGCGCCCTCCCTACGGCCCGGGTCTGAGGCTGC GCGTTCGGCTCCAACCGCCCTGGGGGTTTCCCGCCACTCAACAGCGGTCTCCGAGCCCGCTGATGCTA CTGCACCCCGCCGAGCTGTGCCCTTTCTGACGCCCCACGGCCAGCAGGTGCCCTACTACCTGGAGA ACAGCCAGCGGCTACACGGTGGCGAGGCGCGCCCGCGGCACTTCTACAGGCCAAATTCAGATAATCG ACGCCAGGGTGGCAGAGAAAGATTTGGCCAGTACCAATGACAAAGGGAAGTATGGCTATGGAATCTGCCAAG GAGACTCGCTACTGTGCAGTGTGCAATGACTATGCTTACGGCTACCATATGGAAGTCTGGTCTGTGAGG GCTGCAAGGCCCTTCTCAAGAGAAGTATTCAAGGACATAACGACTATATGTGTCCAGCCACCAACCAAGT CACCATTGATAAAAACAGGAGGAAGAGCTGCCAGGCTGCGCGCTCCGCAAAATGCTACGAAGTGGGAATG ATGAAAGGTGGCATACCAAAAGCCAAAGAGGAGGAGGAGCAATGTTGAAACACAAGCCCGCAGAGATCATG GGGAGGGCAGGGTGAAGTGGGCTCTGCTGGAGACATGAGAGCTGCCAACCTTTGGCCAAAGCCCGCTCAT GATCAAACGCTCTAAGAAGAACAGCCCTGGCCCTGTCCCTGACGGCCGACCAGATGGTCAAGTGCCTTGTG GATGCTGAGCCCCCACTACTCTATTCGAGTATGATCCTACAGACCCCTTCAAGTGAAGTCTGATGATGG GCTTACTGACCAACCTGGCAGACAGGAGCTGGTTCACATGATCAACTGGCGGAAGAGGTTGCCAGGCTT TGTGGATTTGACCTCCATGATCAGGTCACCTTCTAGAAATGTGCCCTGGCTAGAGATCCTGATGATGGT CTCTCTGGCGCTCCATGGACACCCAGGGAAGTACTGTTTGTCTTACTTCTCTTGGACAGGAAC AGGAAAAATGTTAGAGGGCAAGTGGAGATCTTCGACATGCTGTGGCTACATCATCTCGGTTCGCGCAT GATGAATCTGCAGGAGCAGGAGTTTGTGTCCCTAAATCTATTAATTTTGTCTTAAITCTGGAGTGTACACA TTTCTGTCCAGCACCTGAAGTCTCTGGAAGAGAGGACCATATCCACCAGTCCCTGGACAGATCACAG	109

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25 30 35 40 45 50 55	ACACCTTGTATCCACCTGATGGCCAAAGGCAAGGCTGACCCCTGCAGCAGCAGCACCAGCGGCTGCCCCAGCT CCTCCTCATCCTCTCCCACATCAGGCACATGAGTAACAAGGCATGGAGCATCTGTACAGCATGAAGTGC AAGAAGTGGTGGCCCTTATGACCTGCTGCTGGAGATGCTGGACGCCACCGCCTACATGCGCCCACTA GCCGTGGAGGGGCATCCCTGAGGAGACGGACCAAAGCCACTTGGCCACTGCGGGCTCTACTTTCATCGCA TTCCCTGCAAAAAGTATTACATCAGGGGGAGGCGAGGGTTTCCCTGCCACGGTCTGAGAGCTCCCTGGC TCCCACACGGTTCAGATAATCCCICTGCATTTTACCCTCATCATGCACCACITTAGCCAAATTTCTGCT CCTGCATACACTCCGGCATGCATCCACACCAATGGCTTTCTAGATGACTGGCCATTCATTTGCTTGCCTC AGTTCCTAGTGGCAGCATCTCTGCTCTCTGTTGGGAACAGCCAAAGGATCCCAAGGCTAAATCTTTGT ACAGCTCTCTTTCCCTTGTCTAIGTTACTAAGCGTGAGGATCCCGTAGCTCTTCACAGCTGAACCTCAG TCTATGGGTGGGGCTCAGATAACTCTGTGCATTTAAGCTACTTGTAGAGACCCAGGCCCTGGAGAGTAGA CATTTTGCCCTCTGATAAGCACITTTTAAATGGCTCTAAGAATAAGCCACAGCAAAGAATTTAAAGTGGCT CTTTAAATGGTGACTTGGAGAAAGCTAGGTCAAGGGTTTATTATAGCACCCCTCTTGTATTCCTATGGCA ATGCATCCITTTATGAAAGTGTACACCTTAAAGCTTTTATATGACTGTAGCAGAGTATCTGGTGTATT CAATTCATTTCCCTATAGSAAATACAAGGGGCACACAGGGAAGGCAGATCCCTAGTTGGCAAGACTATT TTAACITGATACACTGCAGATTCAGATGTGCTGAAAGCTCTGCCCTCTGGCTTTCCGGTGTAGGGTTCCAG TTAATTCATGCCCTCCCATGGACCTATGGAGAGCAGCAAGTTGATCTTASTAAAGTCTCCCTATATGAGGG ATAAGTTCCTGATTTTGTGTTTTATTTTGTGTACAAAAGAAAGCCCTCCCTCCCTGAACTTGCAGTAA GGTCAGCTTCAGGACCTGTCCAGTGGGCACCTGACTTGGATCTTCCCGCGGTGTGTGCTTACACAG GGGTGAACCTGTTCACTGTGGTGTAGCATGATGAGGGTAAATGGTAGTTGAAAGGAGCAGGGGGCCCTGGT TTGCATTTAGCCCTGGGGCATGGAGCTGAACAGTACTTGTGCAGGATCTGTGGCTACAGAGAACAAG AGGAAAGTAGGGCAGAACTGGATACAGTCTGAGGCACAGCCAGACTTGTCCAGGGTGGCCCTGCCAC AGGCTGCAGCTACCTAGSAACTTCTTGCAGACCCCGCATTGGCTTTGGGGTGGCCCTGGGATCCCTG GGTAGTCCAGCTCTTTCTTCAITTTCCAGCGTGGCCCTGGTTGGAAAGAGCAGCTGTACAGCTGCTGTGA GACAGCTGTGTTCCCTACAAATTTGCCCGACCCCTGGGGCAGGGAGAGGGTGGGGACCGTTGCTGTAC TACTCAGGCTGACTGGGGCTGTICAGATTACGTATGCCCTTGGTGGTTAGAGATAATCCAAAATCAGG GTTTGGTTTGGGGAAGAAAATCCICCCCTTCTCCCGCGCCCTTCCCTACCGCTCCACTCTCTGCCA GCTCATTTCCCTCAATTTCCCTTGCCTATAGGCTAAAAAGAAAGGCTCATTTCCAGCCACAGGGCAGCC TTCCCTGGGCCCTTGTCTCTCAAGCACAAATTTAGGGTACTTCCCTTTTCTTAAACAAAAGAAATGTTG ATTTCCCTGGGTGACTTATTGCTGTAATGAAACCTTATTGAGAGTGTATGCTGTGTTAGCCAAATG ACCAGGTTGAGTGTCTCGGCTTCTTGGTATGCTTGTGTTGGAAAAGTGGATTTCAITTCATTTCTGAT TTCCAGTAAATTTTGGACAGTACCTAAAGGACTGAGAACTCGGGAGGGCAAAAAAAAAAAAAAAAAAGTTG CACTTAAATTTGGGGACAAATTTATGATCTGTGTTAAGGATGTTTAAAGAACATAATTTCTTTGTTG TGTITGTTAAGAAGCACCTTAGITTTGTTAAGAAGCACCTTATATAGTATAATATATATTTTTTGGAA TTACATTTGCTTGTATCAGACAATTTGAATGATGTAATTTCTGCTTGGATTTAATTTGACTGGGTTAAAC TGCAAAACCAAGGAAAPATAATTTAGTTTTTTTTTTTTTTTTTTGATACCTTTCAAGCTACCTTGTGATG GTACAGTCAATTTAGCCTTAAAGCCTGGTGTATTTCAATTTAAATGAAGTACACATTTATATCAACTTT TGTATCCACAGTAGACAAAATAGCACTAATCCAGATGCCATTTGTTGGTACTGAATGACAGACAATCTT ATGTAGCAAAGATTTAGCCTGAAAAGGAAAATTTATCAGGGCAGCTAATTTCTGCTTTTACCAAAATATCA GTAGTAATATTTTTGGACAGTAGCTAATGGGTGAGTGGGTTCTTTTTAAAGTTTATACCTTAGATTTTCTT TTAATAAAATTTAAATAAAACAAAATAAAATTTCTAGGACTAGACGATGTAATACCAGCTAAAGCCAAAC AATTTACAGTGGAAAGGTTTACATTTATTCATCCAATGTGTTTCTATTCAIGTTAAGTACTACTACTACT TGAAGTGGGCAGAGAACATCAGATGATGAAATGTTCGCCAGGGGCTCCAGCAACTTTGGAAATCTCT TTGTATTTTTACTTGAAGTGGCATAATGGACAGCAGATATTTCTGGCTGATGTTGGTATTGGGTGTAG GAACATGATTTAAAAAAAACITCTGCTCTGCTTTCCCCACTCTGAGGCAAGTTAAAAATGTAAGAAT GTGATTTATCTGGGGGGCTCAGTATGGTGGGAAAGTGGATTCAGGAACTGGGGAAATGGCAAAATATTT AAGAAGATATTGAAAGTATTTGGAGGAAAATGGTAAATCTGGGTGTGCACCAGGGTTCAGTAGAGTCC ACTTCTGCCCTGGAGACCACAAAICAACTAGCTCCATTTACAGCCATTTCTAAAATGGCAGCTTCAGTTC TAGAGAAGAAAGAACACATCAGCAGTAAAGTCCATGGAATAGCTAGTGTGCTGTTCTTTTCCGCAT TGCCTAGCTTGCCTAATGATTCIATAATGCCATCATGCAGCAATTTATGACAGGCTAGGTCAATCCAAAGA GAAGACCCTATCAATGTAGGTTGCCAAAATCTAACCCTAAGGAAAGTGCAGTCTTTGATTTGATTTCCCTA GTAACCTTGCAGATATGTTAACCAGCCATAGCCCATGCCCTTTGAGGGCTGAACAAATTAAGGGACTTA CTGATAATTTACTTTTGTACAIITAAGGTGTTCTCACCTTGAATCTTAAACACTGAAATGGCCATTGA TTTAGGCCACTGGCTTAGAGTACICCTTCCCTGATGACACTGATTACAAAATCTTTCCATTTCACTACT TTCCAATTTAGATGGACTGTGGTACTGGGAGTGAICTACTAACCCATAGTAATGTCTAATATTACACA GGCAGATCTGCTTGGGAAAGCTAGTTATGTGAAAGGCAAAATAGASTCATAACAGTAGCTCAAAGGCAACC ATAATCTCTTTGGTGCAGGCTTGGGAGCGTGTCTAGATTACACTGCACCATTCCCAAGTTAATCCCC TGAATACTACTCTCAACTGGAGCAAAATGAACTTTGGTCCCAATAATCCATCTTTTCACTAGCGTTAAT ATGCTGTGTTTCCAACTGCATTTCCCTTTCCAAATGAAATTAAGTGTGGCCCTCGTTTTTGTGATTTAAAA TTGTTTTCTAAGTAAATGCTGCCICTATTATGGCACTTCAATTTGCACTGCTTTTGTGATTTCAAGAAA AATTTCTATTCTTTTTTTGCACTCAATTTGGCTGAACTTTTAAATAATGTAATGCTGCCATGTTCCA AACCCATCGTCACTGTGTGTTTGAAGCTGTGCACCCTAGAAACAACATAATGTTCCATGAGCAGGTGC CTGAGACACAGACCCCTTTGCATTCACAGAGAGGTCAATGGTTATAGAGACTTGAATTAATAAGTGCAT TATGCCAGTTTCTGTTCTCAGAGGTGATAAAACAATGCTTTTTTGTGCACIACACTACTCTCAGTGTAGA GCTCTGTTTTTATGGGAAAAGGCTCAAAATGCCAAATTTGTTGATGGAATAATATGCCCTTTGCCGAT GCATACTATTACTGATGTGACTCGGTTTTGTGCGAGCTTTGCTTTGTTTAAAGAAACACACTTGTAAACC TCTTTTGCCTTTGAAAAAAGAAATCCAGCGGATGCTCGAGCAGCTGTAACAATTTTCTCAACTATTG ATGTTCAAATAAAGAAATTAACCTAAA	110
60	NM_130398 AAATGAAAGGTCAGCCTTTCGCGCGTGTGTAGGCAAGTTACCCGTGTTCTGCGTGGCCGGCTGGT	110

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25 30 35 40 45	GCTCTGGCCACAGTGAAGTTAGGGGCGTCGGAGCGGGTTTCTCCAACCGCAATCGGCTCCGCTCAAGGGGA GGAGGAGAGTCCCTTCTCGAAGCCCTAAGGAAACGTGTCTGCTGGAATGGGCTTGGGGGCCACGCCCTGC ACATCTCCGCGAGACAGAGGGGATAAAGTGAAGATGCTGCTGTTATGTTACTCGAGTGCACATGCGAC CTCTGAGATAIGTACACAGTCAATCTTACTATCGCACTCAGCCATCTTACTACGCTAAAAGAAATAA TTATTCGAGGATATTTGGCTGGCCCAAGAAACTTATGTAATTTTCATGAACATTTATATCCGTTTTC TCGGAGTGAGAGAAAACCTTTTTAGATATCATCTGAGAGAACTAGTGAATCCCACTGCTGAGTGGAGT TGAGAGTCAAGAACCCTGAAATTTGAGAACTGCTGGACCAGAGCCCTTTAGAGCTCTGATAAGGTGICA ACAGGGTAGTTAATTTGGCCACCATGGGGATACAGGGATTGCTACAATTTATCAAGAAGCTTCAGAACCC ATCCATGTGAGGAAATATAAAGGGCAGGTAGTAGCTGTGGATACATATTGCTGGCTTACAAAAGGAGCTA TTGCTTGTGCTGAAAACCTAGCCAAAGGTGAACTACTGATAGGTATGTAGGATTTTGTATGAAATTTGT AAATATGTTACTATCTCATGGGATCAAGCCTATCTCGTATTTGATGGATGACTTTTACCTTCAAAAAG GAAGTAGAGAGATCTAGAAGAGAAAGACGACAAGCCAACTTCTTAAAGGAAAGCAACTTCTGTGAGG GGAAGTCTCGGAAGCTCGAGAGTGTTCACCCGCTCTATCAATATCACACATGCCATGGCCCAAAAGT AATTAAGTTCGCCGCTCTCAGGGGTAGATTGCCCTCGTGGCTCCCTATGAAGCTGATGCGCAGTGTGCC TATCTTAAACAAGCGGCAATTTGTGCAAGCCATAATACAGAGGACTCGGATCTCTGAGCTTTGGCTGTA AAAAGTAAATTTAAAGATGGACCAGTTTGGAAATGGACTTGAATTTGATCAAGCTCGGCTAGGAATGTG CAGACAGCTTGGGGATGTAATTCACGGAAGAGAAAGTTTCGTTACATGTTGATTTCTTACAGTTTGTGACT CTGTCACTCAGCTGCGTGGATTTGGATTAGCAAGGCATGCAAGTCCCTAAGACTAGCCAAATAATCCAGATA TAGTAAAGTTATCAAGAAAATTTGCATTTATCTCAAGATGAATATCAGGTTACAGAGGATATCATCAA CGGTTTATTCGGGCCAAACAATACCTTCTCTATCAGCTAGTTTTGATCCCATCAAAAGGAAACTTATT CCTCTGAACGCCATAGAGATGATGTTGATCCTGAAACATAAGCTACGCTGGGCAATAIGTTGATGATT CCATAGCTCTTCAAAATAGCACITGGAAATAAAGATATAAAATCTTTGAACAGATCGATGACTACAATCC AGACACTGCTATGCTGCCATTCAGAAAGTCAAGTGTGGGATGACAAAACATGTCAAAAGTCAAGTAAAT GTTAGCAGCAATTTGGCATAGGAATTACTCTCCCAAGACAGAGTGGGACTGTTTTCAGATGCCCCACAAT TGAAGGAAAATCCAAGTACTGTGGGAGTGGAAACGAGTGTAGTACTAAAGGGTTAAATCTCCCAAGGAA ATCATCCATTTGTAAAAGACCAAGAAGTGCAGAGCTGTCAGAAGATGACCTGTTGAGTCAATCTCTCT TCATTTACGAAGAAGACCAAGAAAATAGCTCTGAAGGCAATAAATCATTGAGCTTTTCTGAGTGTGTTG TGCCCTGACCTGTAATGGACCTACTTAAACAAGAGTGTAAAGCACTCCACTAGCAGAGAAAATAAAT TGCAACATTTTTCAAAAGGAAAATGAAGAAAGTGTGCAAGTGTGGTTCCAGGACAGCAAGCAGGTTTT TTTTGCAGTTGAGTTTACTGACTGTGTATCAACAAGGAGCATCCAGCCTCTGGATGAAACTGCTG TCACAGATAAAGAGAACATCTGCATGAATCAGAGTATGGAGACCAAGAAGGCCAAGACTGGTTGACAC AGATGTAGCAGCTAATTTCAAGTGTGACATTTCCGAATAATCATAATCCAGGTGATCATAATCCAGACAAG GCAACAGTGTTTACAGATGAAGAGTCTTACTCTTTTGAAGCAGCAAAATTTACAAGGACCAATTTCCAC CCACTTTGGGAACACTAAGAAGTGTGTTTAGTTGGTCTGGAGGCTTTGGAGATTTTCAAGAAGCCCGAG CCCCCTCCAAGCACAGCATTTGCAGCAGTTCGGAAGAAAGAGCCGATTCGCCCACTCTTTGCTCGAAT AATATGCTGTGATGTTCCGAGTTAAAGAGCCAGGAGTCCAGTGCAGATGAGTCTCATCCCTTACGAGAA AGGCATGTTCTTACACTCCCAGGAAAGTGGAGAATTTCTACTGCAGAGTTCAAATGCATCAAAGCTTCT TCAGTCTCTAGTAAGGACTCTGATTTAGAGGAATCTGATGCAATATTAAGTACTTGACAGCTAAAGT GACCGAGCTTCCAAGCTACGTTTATCTCATTCTCAAAAAGACACACCTCTAAGGAACAAAGTTCCTG GGCTATATAAGTCCAGTCTGCAAGCTCTCTTTCTACAACCAAGATCAAACTCTAGGACCTGGCCAGAGC CAGTGGGCTGAGCAAGAGCCGCAAGCATCCAGAAGAGAAAGCATCATAATGCCGAGAAGCAAGCGGGG TTACAGATCAAACCTCAATGAGCTCTGGAAAACCTTTGGATTTAAAAAAGATCTGAAAAGCTTCTCTCT GTAAGAAAACCCCTGTCCCAGTCAAGATAACATCCAATCACTTCAAGAGCGAAGGAGATAATTTTAA CAAACCTGAAATGTGGCCGTGTTCAAAGAGCAATATTCAGTAAATGACAGCTGCTGCAAAAGCTTTGGCT GCAAGAAATCTGATCAATTTGAAGTCCCTGTTTGGGAATGAGGCACTTATCAGCATGAAGAAATTTTCT TCATTCGTGGCCATTTTAAAAATACAATACATTTGTATATTAACCTTTATAAATGGGTTGTGGTTTTTT GCTCAGCTTTTTATATTTTATAAGAAGCTAAATAGAAGAATAATGATCTCTGACAGGTTTTTGGAGG TTTTAGTGTTAATGGGAAAATCCCTCGGAGTTTATAAAGTCTACTCTAAATATTTCTGTAATGTTGTC AAGTAGAAAGATAGTAAATGGAGAACTACAAAAAATAAAAAAAAAA	
50 55 60	CCATGACCTGCCTTGAGAAGGGGCAGGGGAGCCAGATGGACTGGAAGTGGAGTGGCAGTGACCAAGGAG GAGGAGGTGTATAGGCTTCCCACGCGAGGTAGATCCAGAGACACCAGTGCACCCATAGGCCCCCTAGCA CTGCAAGTGGTCAACCAGTTCCITTTGTCAGCTGAGACTCAGTTCTGAGTGTCTATTTTGGGAAACAGA GGCGTCTTGGTAGCATTTGGAAGAGGATAGCCAGCTGGGGTGTGTGTACATCACAGCCTGACAGTAACA GCATCCGAACAGAGGTTGACTGGCTAAGGGCAGACCCAGGGCAACAGGTTAACCGTTCTAGGGCCGGGCA CAGGGAGGAGAACATTTCAACACTCTGTGTGCCAGTGCACGCAAGTCTCTCTTTTATCCCTCAAAAC AGTCTATGAGGATATAAGCCAGAGAGACAGACAGAAAGGAAATACAAGTGTGTGAGAGTCAAGATTG AACTTGGCTCTGGCAGATGAAAAATTTAGGGTCTGTATTTTACAAAACCGTGTGTGCCICAGATGGAGT TGTTGCATAAACAAGCAGAGGTATCCAGGGTGCAGGCTGCTTGGCCAGGAAAGGGCCGCTTGTGAGTT GTGACCACCCAGCCCTGGAATGTGAGTAAATGCTGTAAAGGAGTGGGGATCGGATCAGATGCCATCCAGAT GCTGAAGTTTACCTTGTGCTATTTTCACTTTCTTTTTTGGCTCTCTGCAATCAATTTATTTATTTAG CAAAAAAGAAAATATGTTGTCGCGAGAGCATGCAGAAGATATGCTCCGCTCTCTGCTTCCCTCAAAAAA GAATCCAAAACCTGCTTCTGTGAACGTTGTGCCAGGTCACAGCAGGACTCAGGGAGAGCAGGAAGCCCA GCCCAGACCCCTTGCACAACCTACCGTGGGGAGGCTTAGGCTCTGGCTACTACAGAGCTGGTCCAGTC TGCCTGCCACAGCCTGGCCAGGACTTGGACACAICTGCTGGCCACTTCTGCTCAGTTTTCTTATCT GCAAAAATAAGGAAAAGCCCCACAAAAGTGCACGTGTAGCAGGAGCTTTTCCCTCCCTATTTTAGGA AGGCAGTTGGTGGGAACTCCAGCTTGGTCCCTGAGAGCTGTGAGAAGGAGATCCGGCTGCTGCTGGCC TGTGGGGTCCCTGCTGAGTGGCCGGCCCTCCAGTCTGTCCCTGGAGCCCTGAGGAAGTGGAGCT TGGTATGGCTTCTGAGTGGGAGAGGGTGGCAGGGTGGGAAGAGTGGGCACAGAGGGGGCTGCTGGG	111

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25 30 35 40 45 50	<p>CTGAGCAAAGCTGAAAAGGATCCTTGCACAGGCCCTGAGAAGTGGCGGCAGGGCAGGGCTCAACCCTG AGACTCAGTCAGTGCCTGGCTTCCAGCAAGCATTCATCTAICACTGTCTCTCGAGAGAGGACTGGCCCT GCAGGGCCGAGGGCCCTAAGCTGGGCTGCAGAGCTGGTGGGAGCTCCTTGCCTGGGTGTGTGTGGCGT GTGTGTGTGTCTGTGCACCTGGGTGTGTGACCTAGGAGTCCAGGCAGCATGTGTGGTAAAGCATTTATG AGGGTGATATGCCCCGGTGCAGCATGACCTGTATGTGGCACCAACAGCATGTGCCTTGTGTGTGTGTGT GTCCGTATGT AAATGCTGT TGTGTGGGTGTCAAGGAGTGGTGGCTCCTTACAGCATGCGTGGCAAGTGTGTGTGTGTGTGTGTGTGTGT GTGTCTCTGTACACAGGAGGTGCCTCAGATGGGGCTGCGGGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT AGAGCCCTGCCTGGCTCCAGCCTGGAGCAGCAAGAGCAGGAGCTGACAGTAGCCCTTGGGCAGCCTGTG CGCTGTGTGTGTGGGCGGCTGAGCGTGGTGGCCACTGGTACAAGGAGGGCAGTGCCTGGCAGCTGTGT GCCGTGTACGGGGCTGGAGGGCCGCCITAGAGATTGCCAGCTTCCCTACCTGAGGATGCTGGCCGTACT CTGCCTGGCAGCAGGCTCCATGATCGTCTCTGCAGAACTCACCTTGATTACAGGTGACTCCTTGCCTCC AGCAACGATGATGAGACCCCAAGTCCCATAGGGACCTCTCGAATAGGCACAGTTACCCCAAGCAAGCT AGTAGTCTCCAAAGGACTTGTGTCCCGCTGCTGCTCATCTGATCACTGAGAAAGAGGAGCCTGTGTGGG AACACACGGTCACTTCTAGGGGCTTCCCTGCCCCCAGCAGCCCTACTGGACACACCCCAAGCCATGGA GAAGAATGCTGCTGCTACCTTCCGGGGAAACACCTCAAGTTCGGCTGTCCAGCTGCAGGCAACCCCAAG CCACCATCCGCTGGCTTAAAGGATGGACAGGCTTTCATGGGGAGAACCAGCATGGAGCATTGGCTGCG GCCATCAGCACTGGAGTCTCTGTGATGGAGAGCTGGTGTCTCGGACCGGGGCACATACACTGTCTGTGT AGAGAACGCTGTGGGCGAGCATCCGTTATAACTACCTGTAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT ATCTGCAGGCGGGCTCCCGCCACACCCACAGCCGTGGTGGGAGCGAGCTGGAGCTGCTGTGCAAGG TGTACAGGATGCCAGCCACATCCAGTGGCTGAAGCACATCGTCACTAACGGCAGCAGCTTCCGAGC CGACGGTTCCTTATGTGAAGTCTAAAGACTGCAGACATCAATAGCTCAGAGGTGGAGGTCTGTGAT CTCGGAACGCTGCAGCCGAGGACGCGAGCTACACCTGCCTCGCAGGCAATCCCATCGGCCCTCTCTCT ACCACTGTGCCTGGCTCAGCGT TCTGGGCCAGCAGTGGGGCTGTGGCCGTGTGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT GGGGCAGTGTGTGATTTGTGGGTTTGTGAGCTGTATGACAGCCCTCTGTGTGTGTGTGTGTGTGTGTGT CATGTGACCGT CACATATGTTGGGAGCTGGGAGGGACTGAGTTAGGGTGCACGGGGCGGCCAGTCTCACAGATGACCAGT TGTCTGTCTGT CAGGTATACGGACATCATCTGTACGCGTCCGGCTCCCTGGCTTGGCTGTGTGTGTGTGTGTGTGTGTGT CTGTATCGAGGGCAGCGCTCCACCGCCGACCCCGCCGCGCCAGCTGTGCAGAAAGCTCTCCGCT TCCCTTGGCCGACAGTCTCCCTGGAGTCAAGGCTCTTCCGGCAAGTCAAGCTCATCCCTGTGTGTGTGT CGT CTATGGGAGTTCCTCCGGGACAGGCTGGTGTGTGGAAAGCCCTAGGCGAGGGCTGTGTGTGTGTGTGT TACGTGCAGAGGCTTTGGCAAGGACCTGCCCCGCTGACCAAGCCAGCAGTGTGTGTGTGTGTGTGTGT CAAAGCAACCGCTCTGACAAGGACCTGGCCGACCTGGTCTCGGAGATGGAGGTGATGAAGTGTATCCG CGACACAAGAACATCATCAACTGCTTGGT GCCTCCGCAAGGAAACCTGCGGGAGTTCCTGCGGGCCCGGGCCCGCCAGGCCCGCAGCTCAGCCCGA CGGTCTCGGAGCAGTGGGGCCGCTCTCTCTCCAGTCTGGTCTCTGTGTGTGTGTGTGTGTGTGTGTGT GGCATGCAGTATCTGAGTCCCGAAGTGTATCCACCGGGACCTGGCTGCCCGCAATGTGTGTGTGTGT AGSACAATGTGTGAAGATGT AAACAGCAACCGCCCTGCTGTGAAGT CAGAGTGCAGT GCATCCCGGTGGAGGAGCTGT CCAGAGCTGTACGGGCTGT CTGT GACCTTATCCCGCTGT CCCTTGCATTTGGGATCCAGCTCTTCCCTTCCGGTGTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGT GT ATAGCATGGGGCCCTGGCCAGAGT CGT ATCCCGGGT GGTCTGGGCTGTGAACCCCTTCCCGCACCTCCCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT GCATTTGGCCCTGAGCCAGAGAAGTGGAAAGCTGCGGAAACAGGAGCAATGGCGTTTTATAAATA TTTTTTTGAAT</p>	112
55 60	<p>TAAGATCCACATCAGCTCAACTGCACTTGCCTCGCAGAGGCAGCCGCTCACTTCCCGGGAGGCGCTCC CCGGCCCGCGCTCCCGGCAGCCGCTGCCCCCGCGTGTCCCCCGCCCGCCCGCCCGCGCGCCCGCC GCACAGCCCGCCCGCCGAGCTCTGGCTTCTCTTCCCGGGTGGCGTTGGCCCGCGCGGGCGCTCG GGTACTGCAGCTGTCACTCCCTCCCGCCCGCCCGCGCGCGCCCGCCCGCTCGCTTCCGCTTCCGACAGGG TGGATGGT CGACTGGAACGCTACTACGCAGACAGCAGGAGGCTACTTCTCCGTCGGGTGAGCAACATGAACCTCA GGCTGGGTCCATGAATCCATGAACCTTACATGACCATGAACCATGACTACGAGCCGCAACATGA CCCCGGCTCTTCAACATGCTCTTGCCTATGCCAACCCCGGCTAGGGGCCCGCCGTGAGTCCCGCGCAGTAC CGCATGCCGGGGGCTCGGGGGGCCATGAACAGCATGACTGCGCGCGCGGTGACGGCCATGGGTACG GCCTGAGCCCGAGCGCATGGGGCCATGGGTGCGCAGCAGGCGGCTCCATGAATGGCTGGGCGCT ACCGGGCCCGCATGAACCCGTGCATGAGCCCATGGCGTACGCGCCGCTCAACCTGGGGCCGCGCGC GGCGGGCGGGCGGACGCCAAGAGCTCAAGCGCAGCTACCCGACGCCAAGCCCTACTCTGTACATC</p>	112

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
	GAAGTTACCTATTCTCCACCTAAATCTCTGAAAAATGGAGAAACCCCTCTGACTAGTCCATGTCAAATTTT ACTAAAAGTCTTTTTGTTAGATTTATTTTCTGAGCATCTCTGCAAAATGTACTATATAGCCAGGTTT GCTTTGAGGCTAGTAAAAAGATATTTTCTAAACAGATTGGAGTGGCATATAAACAAAATACGTTTTCTC ACTAATGACAGTCCATGATTCGGAAATTTAAGCCCATGAATCAGCCCGGTCTTACCACGGTATGCCT GTGTGCCGAGAGATGGGACTGTGGCCAGATATGCACAGATAAATATTTGGCTTGTGTATCCATATAA AATTGCAGTGCATATATACATCCCTGTGAGCCAGATGCTGAATAGATATTTTCTATTTTTCAGTCTT TTATAAAAGGAAAAATAAACAGTTTAAATGTATGTATATAAATCTCCCAATTTACAATCCCTTCATG TATTACATAGAAGGATTTGCTTTTTAAAAATATACTGCGGGTTGGAAGGGATAITTAATCTTTGAGAAA CTATTTAGAAAATATGTTTGTAGAACAATATTTTGAAGAAATTTAAAGCAATAACAAGAGGAAGG CGAGAGGAGCAGAACAATTTGGTCTAGGGTGGTTTTCTTTTAAACCAATTTTCTTTGTTAATTTACAGTT AAACCTAGGGGCAATCCGGATTGGCCCTCCCCCTTTTGAATAAACCCAGGAAATGTAATAAATTCATT ATCTTAGGGTATCTGCCCTGCCAATCAGACTTTGGGAGATGGCGATTGATTACAGACGTTTCGGGGG GTGGGGGGCTTGCAGTTTGTTTTGGAGATAATACAGTTTCTGTCTATCTGCCCTCCCTATCTAGAGGCAA CACTTAAGCAGTAATTCCTGTTCCTTTGTGTCAAATTTGATCATTTGTTAAAGGATTCGTCAAAATAAAT ACACITTAATTTTCAGTCAAAA	
AJ249248	GTGGCCTCGAGGTGGTGGCAGGGCCGCCCTGTCAGTCCGGAGACGAACGCACGACCGGGCCCTCCGAG GCAGGTTCCGCTGGAAGGAACCGCTCTCGCTTCGCTTACACTTGGCAGAAATGCTCCGAGCTTACTCAC ATAGCATATTTGTTATATCAAAATGAAATGCAAGGAACCAAAAATAACATAATTGAAGGCAGTAAAAGTGA AATTAATAGGAAGATCATCAGTCAAGGAAGACCCACTGGAGAGGACAGAAAATCAAGCAGTGTTTTATC ATGTGTATTTTCAGCAGGCTCTCTTGAATTTAACTAAAAATATGACTGCTCTCTCTTCAGAGAACTGCTC TTTTCAGTACCAGTTACGTCAAAACAACCCAGCCCTAGACGTTAACTATCTGCTATTCTIGATCATACTT GGGAAAATATTTAAATATCCCTTACACTAGGAATGAGAAGAAAAAACACCTGTCAAATTTTATGGAAAT ATTTTTCAGTTCAGTACGATTCGTTGATCTTTTACTTTTGGTAAACATTTCCATTTATTTGATTTTCAG GGATTTGTACTTTTAAAGCATTAGGTTCACTAAATACCACTCTGCTTATTACTCAAAATTTTCTCTTT ACTTATGGCTTTTGCATATCCAGTTTCTGACAGCTTGTATAGATTATTTGCCGAAATTTCTTAAAA CAACCAAGCTTTTCAATTAAGTGTCAAATAATTTTATTTCTTTACAGTAATTTTAAATTTGGATTTCCAGT CCTTGTCTTATGTTTTGGGAGACCCAGCCATCTACCAAGCCCTGAAGGCACAGAAATGCTTATTTCCGTAC TGTCTTTCTATGTCAGCATTTACAGTTACTGGCTGCTATTTTTCATGGTGTATTTTATTTGTAGCTT TCATAACCTGTTGGGAAGAAGTTACTACTTTGGTACAGGCTATCAGGATAACTCCCTATATGAATGAAAC TACTTATAATTTTCCTTTTTCACTCCACATCCAGTTTACTGTGAGATCTAAAAAATATCTTATCCAAAG CTCATTGCTGTTTTCTCAGTACCTGGTACCATTTGTACTACTCCAGTAACTATTGTTTACTTAAAG TTCAGATTCCAGCATAATTTGAGATGAATATTTCCCTGGTTATACTTTGCTCAATAGTTTTCTCATTTGCTAC AGTGTATTTGGTTTTAATTTGTACAAGCTTAATTTAAAAGACATTTGGATTACCTTTTGGATCCATTTGTCAAC TGGAACTGCTGCTTCAATCCACTTACAATCTTAATCTTGAGCAAATTTGAAAAGCCCTATATCAATAATGA TTTGTTAATATTTAATTAATAAAGTTTACAGCTGTCAATAAGATCAATAATTTTATGAACAGAAAGAACTCAG GACATAATAAAAAATAAATGAACATAAACAACATTTTGCCTCCCTGACTGATAGCATTTCAGAAATGTGCT TTTGAAGGGCTATACCAGTTATTAATAGTGTTTTATTTTAAAAACAAAATAATTTCCAAAGGTTTTTAT AGTTATTTCAAGGACACTATATACAAATATTTCTTTGTTTATTAACACAAAAGTATAAGAGTTAAACATT TGGCTATACTGATGTTTGTGTTACTCAAAAAAATCTACTGGATGCAAACTGTATGTAAAATCTGAGATTTT ACTGACAACTTTAAGATATCAACCTAAACATTTTATTAATGTTCAAATGTAAGCAAGAAAAA	114
NM_014176	AGTCAGAGGTCGCGCAGGCGCTGGTACCCCGTTGTCGCGCGGTTGCTGCGTTGIGAGGGGTGTCAGCTC AGTGCATCCCAGGCACTCTTAGTGTGGAGCAGTGAACGTGTGTGGTTCCCTTACTTTGGGATCAITGC AGAGAGCTTCACGCTCTGAAGAGAGAGCTGCACATGTTAGCCACAGAGCCACCCCGAGGATCAGATGTTG GCAAGATAAAGACCAATGGATGACCTGCGAGCTCAAAATATTAGGTGGAGCCAAACACCTTATGAGAAA GGTGTTTTTAAAGCTAGAAGTTATCACTTCTGAGAGGTACCCATTTGAACCTCCTCAGATCCGATTTCTCA CTCCAATTTATCATCCAACATTTGATCTGCTGGAAGGATTTGCTGATGTTCTCAAATTTGCCACAAA AGTGTCTTGGAGACCATCCCTCAACATCGCAACTGTGTGACCTCTATTTCAGCTCTCATGTCCAGAACCC AACCTGATGACCCGCTCATGGCTGACATATCTCAGAAATTAATAATAAAGCCAGCCCTTCCCTCAAGA ATGCCAGACAGTGGACAGAGAAGCATGCAAGACAGAAACAAAAGGCTGATGAGGAAGAGATGCTTGATAA TCTACCAGAGGCTGGTACTCCAGAGTACACAACCTCAACACAGAAAAGGAAGGCCAGTCAGCTAGTAGGC ATAGAAAAAGAAATTTCACTCTGATGTTTGGGGACTGTCTGTTTCACTTTAGTTAATGTGTTCTTTG CAAGGTGATCTAAGTTGCCCTACCTTGAAATTTTTTTTTAAATATACTTGATGACATAATTTTTGTGTAGTT TATTTATCTGTACATATGATTTTGAATCTTTTAAACCIGAAAAATAATAGTCATTTAATGTTGAAA AAAAAAAAAAAAAAAAAAAAAA	115

(continúa)

5
10
15
20
25
30
35
40
45
50
55
60
65

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_006845	<p>AC3CTTGC3GCGGGGATTTAAACTGCGGGGTTTACGCGGGCGTTAAGACTTCGTAGGGTITAGCGAAAITG AG3TITCTTGGTATTGCGGGTITCTGTTCCCTTGCTGACTCICCGAATGGCCATGGACTCGTCCGTTTCAGG CCCGCTGTTTCCCGTCTCGCTATCAAGATCCAACCGCAGIAATGGTTTAAATTCACAGTGCCTAATGTAAG GACTGTGAACITGGAGAAAATCCTGTGTTTTCAGTGSAAATGGGCAGAAAGGAGGTGCCACAAAAGGGCAACAG ATTGATTTTGGATGATGTGGCTGCAATAAACCCAGAACTTTACAGCTTCTTCCCTTACATCCGAAGGACA ATCTGCCCTTGCAGGAAAATGTAACAATCCAGAAAACAAAACGGAGATCCGTCAACTCCAAAATTCCTIGC TCCAAAAGAAAAGTCTTCGAAGCCGCTCCACTCGCATGTCCACTGTCTCAGAGCTTCGCATCACGGCTCAG GAGAAATGACATGGAGGTGGAGCTGCCTGCAGTGCAAACTCCCGCAAGCAGTTTTCAGTTCCTCCTGCC CCCTTAGGCTTCCCTGCCCTGCAGTGGCTGAAATACCATTGAGGATGGTCAGCGAGGAGATGGAAGACA AGTCCATTTCCATCCGAGGCGAGCTCTTCIGCAAACCCCTGTGAAGTCAAGTTCGGAGGAAAATCATGTCTGTG AAGAAAGTGGAAAAATGAAGAACAAGCGAGAAAGAGAAGAGGCCAGAAC TCTGAAATGAGAAATGAAGA GAGCTCAGGAGTATGACAGTGTCTTCCAAACTGGGAATTTGCCCGAATGATTAAGAAATTTCCGGCTAC TTTGGAATGTCTCCACTTACTATGACIGATCCTATCGAAGAGCACAGAATATGTCTGTGTAGGAAA CGCCCACTGAATAAGCAAGAAATTTGCCAAGAAAAGAAATTTGATGTGATTTCCATTCCTAGCAAGTGTCTCC</p> <p>TCTTGGTACATGAACCCAAAGTTGAAAGIGGACTTAAACAAAGTATCTGGAGAACCAGCAITCTGCTTGA CTTTGCATTTGATGAAACAGCTTCGAAATGAAGTTTGTCTACAGGTTCCACAGCAAGGCCACTGGTACAGACA ATCTTTGAAGGTGGAAAAGCAACTTGTCTTTCATATGCGCAGACAGGAAGTGGCAAGACACATATG GCGGAGACCTCTCTGGAAAAGCCAGAAATGCATCCAAGGGATCTATGCCATGGCCCTCCCGGGACGCTTT CCTCCTGAAGAATCAACCCCTGCTACCCGAAGTTGGCCCTGGAAAGTCTATGTGACATTTCTCCGATCTAC AATGGGAAGCTGTTTACCTGCTCAACAGAAAGGCCAAGCTGCCGCTGCTGGAGCACGGCAAGCAACAG TGCAAGTGGTGGGGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGG CAGCGCCTGCAGAACCTCTGGGCAGACATTTGCCAACTCCAATTCCTCCCGCTCCACCGCGTGTCTCCAA ATTATTTCTCGAGCTAAAGGGGAGAAATGCATGGCAAGTTCTCTTTGGTAGATCTGGCAGGGAAATGAGCCAG GC3CGGACACTTCCAGTGTGACCCGGCAGACCCGCATGGAGGGCGCAGAAAATCAACAAGAGTCTCTTAGC CCTGAAGGAGTGCATCAGGGCCCTGGGACAGAAACAGGCTCACACCCCGTTCCGTGAGAGCAAGCTGACA CAGGTCTGAGGGACTCCTTTCATTTGGGGAGAACTCTAGGACTTGCATGATTGCCACGATCTCACACAGCA TAAGCTCCTGTGAATAIAC TTTAAACACCCTGAGATATGCAGACAGGGTCAAGGAGCTGAGCCCCACAG TGGCCCAAGTGGAGAGCAGTTGATTTCAAATGGAACAGAAAGAGATGGAAAGCTCTCTAACGGGGCGCTG ATCCAGGCAATTTATCCAAGGAAGAGGAGAACTGTCTTCCAGATGTCCAGCTTTAACGAAGCCATGA CTCAGATCAGGAGACTGAGAGGAGAAAGGCTATGGAAGAGCTCAAGGAGATCATACAGCAAGGACCAAGCTG GCTTGAAGTCTCTGAGATGACCGAGCAGCCAGACTATGACCTGGAGACCTTTGTGAACAAGCGGAAATCT GCTCTGGCCAGCAAGCCAAAGCATTTCACAGCCCTGCGAGATGTGATCAAGGCCITGCGCCTGGCCATGC AGTGGAAAGAGCAGGCTAGCAGACAAATAAGCAGCAAGAAAACGGCCCAAGTACGACTGCAAAATAAAAAAT CTGTTTGGTTTGCACCCAGCCTCTTCCCTGGCCCTCCCGAGAGAACTTTGGGTACCTGGTGGGTCTAGG CAGGGTCTGAGCTGGACAGGTTCTGGTAAATGCCAAGTAIGGGGGCATCTGGGCCAGGGCAGCTGGGG AG3GGGTGAGAGTACATGGGACACTCCTTTTCTGTCTCCTCAGTTGTCCGCTCACAGAGGAAAGGAGCT CTTAGTACCCCTTTTGTGTGCCCCCTCTTTCCATCAAGGGGAATGTCTCAGCATAGAGCTTTCTCCGCA GCATCTGCCTGCGTGGACTGGCTGCTAATGGAGAGCTCCCTGGGGTTGTCTGCTCTGGGAGAGAGA CG3AGCCCTTTAGTACAGCTATCTGCTGGCTTAAACCTTTCTACGCCCTTTGGGCCAGCACTGAAATGCTTT GTACTTTAAAAAAATTTTTCTGAGACCTTTCTACTTTACTGTCTCCCTAGAGATCCTAGAGGATCCCT ACTGTTTTCTGTTTTTATGTGTTTTATACATTTGATGTAACAATAAAGAGAAAAATAAATCAGCTGTTTTAA GTGTGTGAAAAAATAAAAAAAAAA</p>	116

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_006101	<p>ACTGCCGCGCTCGTGCCTAATGACGTCAGCGCCGCGGAGAAATTCAAAATTCGAACGGCTTTGGCGGGCC GAGGAAGGACCTGGTGTITTTGATGACCGCTGTCTGTCTAGCAGATACTTGCACGGTTTACAGAAATTCG GTCCCTGGTTCGTGTGAGAACTGGAAAAAGGICATAAGCATGAAGCGCAGTTCAGTTTCCAGCGGTG GTGCTGGCCGCTCTCCATGCAGGAGTTAAGATCCCAGGATGTAATAAACAAGCCCTCTATACCCCTCA AACCAAAGAGAAACCAACCTTTGGAAAGTTGAGTATAAACAAACCGACATCTGAAAGAAAAGTCTCGCTA TTTGGCAAAAGAACTAGTGGACATGGATCCCGGAATAGTCAACTTGGTATATTTTCCAGTTCTGAGAAAA TCAAGGACCCGAGACCCTTAATGACAAAGCATTCATTAGCAGTGTATTCGACAACCTCGTGTGTTTCT TACAGAAAATGGTATGCACATAAATGTGTCATGAAATCTCTACAAGCTCCCTCTGTATAAAGACTTCTGT AAGATCTTCACATTTCTTTATGGCTTCTGTGCCCTCATACGAACTTCTGACACAAAATTTGAAGAAG AGSTTCCAAGAATCTTTAAAGACCTTTGGTATCCTTTTGCACATCCAAAAGCTCCATGTACACAGTGGG GGCTCCTCATACATGCCTCACATTTGGCAGCCTTAGTTTGGCTAATAGACTGCATCAAGATACATACT GCCATGAAAGAAAAGCTCACCTTTATTTGATGATGGCCAGCCTTTGGGGAGAAAAGTGAAGATGAAATTA TGCATAATAAGTTGTTTTTGGACTACCCATAAAAATGCTAIGAGAGTTTTATGAGTGGTCCGACAGCTT TGAATGAGATGAATGCAGAGCTGCAGTCAAACCTGAAGGATTTATTTAATGTGGATGCTTTTAAAGCTGAA CATTAGAAGCAAAAACAGAGCATTTGAATGAACAGATTCGAAGATGGAAACAAGAAAAGAAAAGAAC CGAATCGTCTAGAGTCTGTGAGAAAACGAGGCTTCTTACAAGGAGATGTTCAAAGTATCAGGCATA CATGAGCAATTTGGAGTCTCATTCAGCCATCTTGACCAGAAAATTAATGTTCTCAATGAGGAAATTTCT AGAGTAGAACTAGAACTGTAAAACAATAAACACAGGAGAACACTCGACTACAGAATATCATGCACACAGA AGTACTCAGTTCCAGACATTTAGCGAATAAATCATGAAAAGAAATGAATTCAGCAGACTATTAATAAAT AACCAAGGACCTGGAACCTGAACAACAGAAAGTTGTGAATGAGGAGTTAAAATATGCCAGAGGCAAGAA GCGATGAAAACACAATTAGCAGAGTATCACAAATGGCTAGAAAATTAAACTTATTCCTAAAGGTGCTG AGAATTCAAAAGGTTAAGACTTTGAAATTAAGTTTAAATCCCGAGGCTGGTGCCTACTGCTCAAATA CAGGGCTCAAGTTTATGTACCTCTTAAGGAACTCCGAAATGAAACTGAAGAAGAAATTAATAAAGCCCTA AATAAAAAAATGGGTTTGGAGGATACTTTAGAACAAATGAATGCAATGATAACGAAAAGCAAGAGAAGTG TGAAACTCTGAAAGAAAGGTTCAAAGCTGGATGATCTTTCAACAAAATAAATTAAGGAAAGCAGAGGA AGAGGATGAAAAATGTCCAGTGTGAGTGTGAGTCTTTGAGAAAACACAAGCACCTGTAGAAAAGTACTGTT AACCGGGGCTCAGTGAAGCTATGAATGAATTAGATGCTGTTACGCGGGAATACCAACTAGTTGTGCAAAA CCACGACTGAAGAAAGACGAAAAGTGGGAAATAAATGCAACGCTGTTAGAGATGGTGTGACACATGT TGGGCTGTAGAGAAAATCTTTGAGGAGCAGATTTCTAAAAGTTGATAGAGAATATGAGAAGATGCATGICA GAAGATCTCTCGGAAAATATAAGAGATTTAGAGATAAGTATGAGAAGAAAAGCTACTCTAATTTAAGTCTT CTGAAGAAATGAAGATAAATGTTGATCATGTATATATATCCATAGTGAATAAATTTGCTCAGTAAAGTGT TAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</p>	117
BC042437	<p>CTCCCTCCTGACACCTGACTACCTGACGCGCCAGTTCACCTCCTCCAGTCCATGAAGGGCTCCTGC GGCATCGGGGGGGCAICGGGGGGCGCTCCAGCGCATCTCCTCCGCTCGCGGGAGGGTCTGCGCGCG CCCCAGCACCTACGGGGGGCGCTGTCTGTCTCATCCTCCCGCTTCTGCTCTGGGGAGCCATGGGT GGGGGGGCGCTATGGCGGTGGCTTCAGCAGCAGCAGCAGCTTTGGTAGTGGCTTTGGGGGAGATA GGTGGTGGCCTTGGTGTGGCTTGGTGGTGGCTTTGGTGGTGGCTTTGCTGGTGGTGGTGGTGGTGG TGGCAGTGAGAAAGTGCATGACGAGAACCTCAACGACCGCCTGGCCTCCTACCTGGACAGGTGGCTGC TCTGGAGGAGGCAACGCGACCTGGAAGTGAAGATCCGTGACTGGTACCAAGGAGGAGGCGCTGCTGAG ATCAAAGACTACAGTCCCTACTTCAAGACCAATTGAGGACCTGAGGAAACAAGATTTCCACAGCCACAGTGG ACAATGCCAATGTCTTCTGACAGATTGACAAATGCCCGTCTGGCCGGGATGACTTCCGACCAAGTATGA GACAGAGTTGAACCTGCGCATGAGTGTGGAAGCCGACATCAATGGCCTGCGCAGGGTGCITGGACGAACTG ACCCTGGCCAGAGCTGACCTGGAGATGCAGATTGAGAGCCIGAAAGGAGGAGCTGGCTTACCTGAAGAAGA ACCACGAGGAGGAGATGAATGCCCTGAGAGGCCAGTGGGIGGAGATGTCATGTGGAGATGGACGCTGC ACCTGGCCTGGACCTGAGCCGATCTGAAACGAGATGCGGTGACCAAGTATGAGAAGATGGCAGAGAAGAAC CGCAAGGATGCCGAGSAATGGTCTTCCACCAAGACAGAGGAGCTGAACCGGAGGTGGCCACCAACACGG AGTGTGTCAGAGCGGCAAGAGCGAGATCTCGGAGCTCCGGCGCACCATGCAGAACCTGGAGATTGACT GCAGTCCCAGCTCAGCATGAAAGCATCCCTGGAGAACAGCCTGGAGGAGACCAAAGGTTCGCTACTGCATG CAGCTGGCCAGATCCAGGAGATGATTTGSCAGCTGGAGGAGCAGCTGGCCAGCTCCGCTGCGAGATGG AGCAGCAGAACCCAGGAGTACAAGATCCIGCTGGACGTGAAGACGCGGCTGGAGCAGGAGATCGCCACCTA CC3CCGCTGTGGA3GGCGAGGACGCCCACCTCTCCTCCTCCAGTTCTCCTCTGGATCGCAGTATCC AGAGATGTGACCTCCTCCAGCCGCAAAATCCGCACCAAGGTCATGGATGTGCACGATGGCAAGGTGGTGT CCACCCACGAGCAGTCTTCCGACCAAGAACTGAGGCTGCCAGCCCGCTCAGGCTAGGAGGCCCC CGTGTGGACACAGATCCCCTGGAAGATCCCCTCTCCTGCCAAGCACTTCCAGCTGGACCCCTGCTTCA CCTCACCCCTCCTGGCAATCAATACAGCTTCATTATCTGAGTTGCATAAAAAAAAAAAAAAAAAAAAAA AA AA</p>	118

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
AK095281	CTCCTTTTGAGGGGGCCGTTCCCTCGGGGACATGACGCTGGCTCCTGCACAGATCCTGCTCCCTCTGTTGGCCCTT CCTGGGCTGCCCTCCCTCCTCCGGGACTGCTCTGGACTGACACTGCTCAGGTTCGGATTCCTCCAAAGA CTTTGGGAGACAAAGACTTGGTCCCTTTTACAACAAGGGAACGGAGGCTCTAGAACTGACTTCTCTGAA AGGCTTGGATCCAAAGCTCCCTCAGTTCAGCGGCCACGCTATTTCCCTCAGACACAGGGGATCCTTGAAC CTSTGGGCTGTATCTCCCGCGGACTTGGAAAGATCCCAAGAGAGTGGGGCTCCACAGGCTGGAGTGCA ATSGTGTGATCTCGGCTCACTGCAACCTCCACTCCCAGGTTCAAGCTATTTCTCTGCCCTCAGCCCTCTG AGTAGCTGGGATACAGATCCTGGTGGCTGGTTCGGTAATTCAGCTTCGTGCTGGCTACAGGTGGATG ATSCCCACTGGCTGCCGATGACCTTGCACCAAGTGGGCTGGTCTCTGGAGCTGCCCCAGGGGCTGG ACAAGCTGACCTGCCCCGGGCAACCTGGAGATGCAGATTGGAACTCAAGGAGGACCTGGCTTACCT GAAGAAGAACCCACAAGCAGGAAATGAACGTCCTTTGAGGTCAGGTGGATGAGGATGTCAGTGTGAAGATG GACACTGTGCCCTGGAGTGAACCTGAGCTGCATCCTGAATGAGATGCTGACCAGGACAAAGACATTTGGTGG AGAAGAGCTGCAAGGATGCCGAGGGCTGGTTCCTCAGCATGGTGGGTGGCCGTGCGTAAGCAGGTGTGTA CACGTGTGGGCACATSTGCTGCATGCTGGTGCAGCTGGAGCACTGGCAGATCCACAGGCTGTCCAGTTG GAAGGACTTTTGGAAACCAGTTGGACCAGCCCTCATGTTTATGATGTAATAACCTGAGGCTCAGAGAGGA CTCAGCTCACACAGCCCTTCACTGTGGCCTGCAAAAATAGATCCAGGCTCTACAAGTCTGGTCTGGGT TTCCACACAGCTGTTTACAGGATGTCGGTATTTGAATACATATGTATACCTTTGGCAAGCACAGGCTGA GTATCTCCGGTATCCTAGGGACAGCAACAGGCGCAAAAAGAATAACACCCAGTGCTGTCTTTGAGGTGCT GCAGTTAGTGGAAAAAGAAATGCAAAATGACCCGACAGCAGGCTGAATTCCTCCAAGTCCAAATGTGGG TGCAGAGGCTCTCTGTGTGCAGAAAGAGGGGCTGAACCTGCGAGGTGGCCACCAACACAGAGCCCTGCAG AGTGGCTGGATAGAGATATGGAGCTTACGCTCTGTGTCAGAACCTGAGCCGTCCCAGCTCAGCAAGAAA GCATCGCTGGAGGGCAGCTGGTGGAGATGGAGGTGTGTACAGGACCTGCCGCCAGCTGCAGGGGC TTAACAGAAAGCATGGAGCAGCAGCTGTGCGAGCTCTGTGCGACACGGAGCACCAGGACCACAGGCACAG GTCTTCTGGACGTGAAGACGTGGCTGGAGCAGGAGATCGCCACTACCCCGCTTGTCTGGAGGTTGAGG ACGCCAGAGGTGATACTGACGATGCAGGCTGGAGTCTGGCTGAGGAGCCTTGAATGCCAAGTAAAGCG TCTGACTAGATCACGTAGGCAATGGGGAGCCATGGAGGGATTTGGAGCAGGAGAGTGAATGAACATCA AGAGATTTTGAACATTCACCTCTGGCTGCAGAGGAGAAAATGGATCAGAGGGGTCAGGGCGGGGCCAGAG AGATGTGTCAAGGGGCTGGAGCAGGGAGTCTGGCCAGAGAAGTCCCGTGGGTGGTGGGTAGTGGGGCAG GGAAGGAAAGTGGTSCAGCAGAAAGAGAGGTTATAGCTCAAAACAGCGGGACTGGATGCTGGATCTCG GGTAAAGCATGGCTCACAGTCAAGACTCAGTAAGTGTGGGAGAACACATGAAGGAGCAGGCAATGATGG CCCTGGGTTCCTGGTCTGATGACTGTGTAGTGGTGAAGAGCAAGTGGGTGGTGGTGGTGGTGGTGGTGGT TGGGAAGGGTGTACAGCCTTCAGCTGAGAGTGTCCCGAGTCTCCATGCTTAGTCACAGCTTGCAGCTT TTTGCTCCCGGAAATGGTGAAGTCCATCTATAGTCTAACACAGTCTCTCTCTCTTTAATTGGGTCTAT TTGTGGGCCCTCTGGTTATGGAAAAACCACTTGTCTCAGCTTCTCTTGTAAATTCCTGGTGGTGGTGGT ACAGAGTGGCCAGACCTACGCTGTGCTGTTTTCTTTCTTCTCTCTGCTGTGCTGAACCCCTGCCCC TTCATTTCTGGGCTSCGCTAATTTCTGTGATTCCCAACCTGATTTTTTACCAATTTAGGGGAACCTC CTCTGCCAGGGCTACTTCTCCCCAGCAGTGTCTGAGGTGCTGGGCTGGCTGGCATCCTGGGCTGAT GGGTGTCTCTCCCTGCAGGCTGGCCACTCAGTACTCTTGTCTCCCTGGCCCTGCAGCCCAACCCGGGAG CCACAGTACCAGCCACCAGGTGTGCCATCTGGAGGAAGTCCAGGTTGGAGAGTGGTCTCTCTCTGTG AGCAGGTCCACTTCTCCACCCACTGAGACCCCTTTCTGTCTGGCAGACCCCACTCGAGGGCCACGGCA CAGCCATCAGCTCCAGCTCCAGCATGCTACTGCCAGCCCGAGTGTCCGCTGGGGCCCCGGTGCATGG CCTGTGTCTTTCTGTATCTACTTTCTGCAGCCCTCAGTGGAGGGCTCCTGGGTTTTGCTCAGTGGCT ACTATTAAGCTTTGCTCCAAGTTT	119
M21389	GCATCCTTTTGGGCTGCTCACAGCCCCAGCCTCTATGGTGAAGACATACTTGTAGCAGGTCACCAA CTTGCTGCCAAGAGATCAGTGTGCAAGGCAAGGTIATTTCTAAGTGAAGAGGCTGCCAGGAAGAAAG CGTTTGCACCCACACACAGTGTGCAAGGTGTGACCCGCTGAGCTCAGCTGCCCCAGGCAATGCCAGCC CACTTAATCATTACAGCTCGACAGCTCTCTGCCCAAGCCAGTCTGGAAAGGATAAAAAGGGGGCATT ACCGTTCCCTGGGTAACAGAGCCACTTCTGGCTCTGCTGAGCTCTGTTCTCTCCAGCACCTCCCAACCC ACTAGTGCCTGGTCTCTTGTCTCCACCAGGAACAGCCACCAATGCTCGCCAGTCAAGTGTGCTCTCCG GAGCGGGGGCAGTGTAGCTTCCAGCACCCCTCTGCCATCACCCTGCTGTCTCCCGCACCAGCTTCCAC TCCGCTGCCCGTCCGCGGGTGGCGGTGGTGGTGGCTTCCGCAAGGTCAGCCTTCCGGGTCTGTGGAG TGGTGGCTATGGCAGCCGAGCCCTTACAACCTGGGGGCTCCAAAGAGGATATCCATCAGCACTAGAGG	120

(continúa)

5 10 15 20 25	NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
		AGGCAGCTICAGGAACCGGTTGGTGTGGTGTGGAGGCGGCTATGGCTTTGGAGGTGGTGGCGGTAGT GGATTTGGTTTCGGCGGTGGAGCTGGTGGTGGCTTTGGGCTCGGTGGCGGAGCTGGCTTTGGAGGTGGCT TCGGTGGCCCTGGCTTTCCCTGTCTGCCCTCCTGGAGGTATCCAGAGGTCACCTGTCAACCAGAGTCTCCT GACTCCCCCAACCTGCAAAATCGACCCAGCATCCAGAGGGTGAGGACCGAGGAGCGGAGCAGATCAAG ACCCTCAACAATAAGTTTGGCTCCTTCATCGACAAGGTGCGGTTCCTGGAGCAGCAGAACAGGTTCTGG ACACCAAGTGGACCCCTGCTGCAGGAGCAGGGCACCAGACTGTGAGGCAGAACCTGGAGCCGTTGTTCGA CGAGTACATCAACAACCTCAGGAGGCTGGACAGCATCGTGGGGAAACGGGGCCGCTGGACTCAGAG CTGAGAAACATGCAGGACCTGGTGGAGACTTCAAGAACAGTATGAGGATGAAATCAACAGCGCTACCA CTGCTGAGAATGAGTTTGTGATGCTGAAGAGGATGTAGATGCTGCCTACATGAACAAGGTGGAGCTGA GGCCAAAGTTGATGCATGATGGATGAGATTAACCTTCATGAAGATGTTCTTTGATGCGGAGGTGTCCCA ATGCAGACGCATGTCTCGACACTCAGTGGTCTCTCCATGGACAACACCGCAACCTGGACCTGGATA GCATCATCGGTGAGGTCAAGGCCAGTATGAGGAGATTGCCAACCCGAGCCGGACAGAAAGCCGAGTCTG GTATCAGACCAAGTATGAGGAGCTGCAGCAGACAGCTGGCCGGCATGGCGATGACCTCCGCAACACCA CATGAGATCAGAGACTCAGTGGAGAGGAGTTGGACCACTCAACATCTCTGTTGTCAAGCAGTGTTCCTC GGCCTAATCTGCAGAACGCCAATGGCGGATGCCGAGCAGCGTGGGGAGCTGGCCCTCAAGGATGCCAGGA CAAGCTGGCCGAGCTGGAGGAGGCCCTGCAGAAAGSACAAGCAGGACATGGCCCGCTGCTGGCTGAGTAC CAGGAGCTCATGAACACCAAGCTGGCCCTGGACGTGGAGATCGCCACTTACCGCAAGCTGTGGAGGGCG AGGAATGCAGACTCAGTGGAGAGGAGTTGGACCACTCAACATCTCTGTTGTCAAGCAGTGTTCCTC TGGATATGSCAGTGGCAGTGGCTATGGCGTGGCCCTGGTGGAGTCTTGGCGGCGGCCCTGGTGGAGGT CTTGCCGGAGGTAGCAGTGAAGCTACTACTCCAGCAGCAGTGGGGGTGTGGCCCTAGGTTGGTGGCTCA GTGTGGGGGCTCTGGCTTCAGTGCAGCAGTGGCCGAGGGCTGGGGGTGGGCTTTGGCAGTGGCGGGGG TAGCAGCTCCAGCGTCAAATTTGTCTCCACCCTCTCTCCCGAAGAGCTTCAAGAGCTAAGAACCT GCTGCAAGTCACTGCCCTTCCAAGTGCAGCAACCCAGCCCATGGAGATTGCCCTCTCTAGGCAGTTGCTCA AGCCATGTTTTATCCTTTCTGGAGAGTAGTCTAGACCAAGCCAATTCAGAACACACATCTCTTGGTTC CAGGAGAGCCCCATTCCAGCCCTGGTCTCCCGTCCCGCAGTCTATATTTCTCTCAAAATCAGCCCTC AGGTTTCCCACAGCATGGCCCTGTGACACGAGAACCCAAAGTTTCCCAAATCTAAATCATCAAAACA GAATCCCCAACCCAAATCCAAATTTGTTTTGGTTCTAACTACCTCCAGAAATGTTTCAATAAAATGCTT TTATAATAI	
30 35 40 45 50 55 60	NM_001123066	GGACGGCCGAGCGGCGAGGGGCTCGCGCGGCCACTAGTGGCCGGAGGAGAAGGCTCCCGCGGAGGCCG CGCTGCCCGCCCTCCCTGGGGAGGCTCGCGTTCCCGCTGCTCGCGCTGCGCGCCCGCCGGCCCTCA GGAACGGCCCTCTTCGCCGGCGCGGCCCTCGCAGTCAACGCCACCCACAGCTCCGGCACCACAGCA GCSCCGCTGCCACCGCCACCTTCTGCCCGGCCACACAGCCACCTTCTCTCCCGCTGTCTCTCC CGTCTCCGCTCTGTCCACTATCAGGTGAACCTTTGAACCAAGGATGGCTGAGCCCGCCAGGATTCGAAG TGTGGAAAGTACAGCTGGGACGTACGGGTGGGGGACAGGAAAGATCAGGGGGGCTACACCATGCACCA AGACCAAGAGGGTGCACCGGACGCTGGCTGAAAGAAATCTCCCTGCAGACCCCACTGAGGACCGTGT GAGGAAACGGGCTCTGAAACCTCTGATGCTAAGAGCAGTCCCAACAGCGGAAAGATGTGACAGCAGCTT TGGATGAGGGAGTCCCGGCAAGCAGGCTGCCGCGCAGCCACACGGAGATCCAGAAAGGACACACAGC TGAAGAAGCAGGCAATTCGAGACACCCCGCCCTGGAAGACGAAAGCTGCTGGTCAAGTCAAGCAAGCC GAAAGTGGTAAAGTGGTCCAGGAAGGCTTCTCCGAGAGCCAGGCCCCCGAGGTCGAGCCACAGCTCA TGTCCGGCATGCCGTGGGGCTCCCTCTGCTGAGGGCCCGCAGAGGGCCACACGCCAACCTTCCGGGAC AGGACCTGAGGACACAGAGGGCGCCGCGCCAGCCCTGAGCTGCTCAAGCACCAGCTTCTAGGAGACTG CACCAGGAGGGGCGCGCTGAAGGGGGCAGGGGGCAAAAGAGAGGCGGGGAGCARGGAGGAGGTGGATG AAGACCGCAGCTCGATGATCTCCCTCCCAAGACTCCCTCCCTCCCAAGGCTCCCGAGCCCAAGATGG GCGGCTCCCGCAGACGCCCGCCAGAGAAGCCACAGCATCCAGGCTTCCAGCGGAGGGTGGCTCCCG CTCCCTGTGGATTTCTCTCCAAAGTTTCCACAGAGATCCAGCCCTCAGAGCCGACGGGCCAGTGTAG GGGGGCCAAAGGGCAGGATGCCCCCTGGAGTTTACGTTTACGTTGGAAATCACACCCAACTGCAGAA GGAGCAGGCGCACTCGGAGGAGCATTTGGGAAGGGCTGCATTTCCAGGGGCCCTGGAGAGGGGCCAGAG GCCCGGGGCCCTCTTTGGGAGAGGACACAAAAGAGGCTGACCTTCCAGAGCCCTGTGAAAGCAGCCCTG CTGCTGCTCCCGGGGGAAGCCCGTCAAGCGGGTCCCTCAACTCAAAGCTCGCAIGGTCAAGTAAAGCAA AGACGGGACTGGAAGCATGACAAAAGCCAAAGACATCCACACGTTCTCTGCTAAAACCTTGAATAAT AGGCCTTGGCTTAGCCCCAAACACCCACTCTGTTAGCTCAGACCCCTCTGATCCAAACCTCCAGCCCTG CTGTGTGCCAGAGCCACTTCTCTCTCTAAATACGTTCTTCTGTCACTTCCCGAATGGCAGTCTGTG AGCAAAAGGAGATGAAATCAAGGGGGCTGATGGTAAAACGAAAGTCCGCCACCCCGGGGAGCAGCCCT CCAGGCCAGAAAGGCCAGGCCACCGCACCAGGATTCAGCAAAAACCCCGCCCTCCAAAGACACCCAC CCAGCTCTGCCACTAACAAAGTCCAGAGAAAGCCACCCCTGCGAGGGCCAGATCTGAGAGGTTGAACC TCCAAAATCAGGGGATCGCAGCGGCTACAGCAGCCCGGCTCCCGAGGCACTCCCGGACCGGCTCCCGC ACCCGCTCCCTTCCAAACCCACCCACCCGGGAGCCCAAGAGGTGGCAGTGGTCCGTACTCCACCCAAAGT CGCGCTCTCCCGCAAGAGCCGCTGCAGACAGCCCGCTGCCATGCCAGACCTGAAAGAAATGTCAAGT CAAGATCGGCTCCACTGAGAACCTGAAGCACCAGCCGGGAGGCGGGAAGGTGCAGATAAATTAAGAGG CTGGATCTTAGCAACCTCCAGTCCAAAGTGTGGCTCAAAGGATAAATCAAAACAGTCCCGGGAGCGGCA GTGTCCAAATAGTCTACAAACCAAGTTGACCTGAGCAAGTACCTCCAAAGTGTGCTCAITAGGCAACAT CCATCATAAACAGGAGGTGGCCAGTGAAGTAAAATCTGAGAAGCTTGAATCAAGGACAGAGTCCAG TCGAAGATTTGGTCCCTGGACAATAACCCACGTCCTGGCGGAGGAAATAAAAGATTGAACCCCA AGCTGACCTTCCCGGAGAACGCCAAAGCCAAAGACAGACCACGGGGCGGAGATCGTGTACAAGTCCGCA GGTGTCTGGGACACGCTCCACGGCACTCAGCAATGTCTCTCCACCGGCAAGTCCAGATGTTAGAC TCGCCCGAGCTCCCGACGCTAGCTGACGAGGTGTCTGCTCCCTGGCCAAAGCAGGTTTGTGATCAGCC CCTGGGGCGGTCAATAATTTGGAGAGGAGAGAATGAGAGAGTGTGGAATAAAAAAAAAAATAATGACCCCG CCCCGGCCCTGCCCCAGCTGCTCTCGCAGTTGGTTAATGGTTAATCACTTAACCTGCTTTGTCT	121

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25 30 35 40 45 50	ACTCGGCTTTGGCTCGGGACTTCAAATCAGTGTGGGAGTAAGAGCAAATTTTCATCTTICCAAATTTGAT GGGTGGGCTAGTAATAAATAATTTAAAAAACAATTCAAAACATGGCCACATCCAACATTTCTCTCAGG CAATTCCTTTTTCATCTTTTCTTCCCTCCATCTAGAGAGGGAGAAGGACAGGCTCTGAAAGCTGC TTCTGGGGGATTTCAAGGGACTGGGGGTGCCAACCCCTCTGGCCCTGTTTGGGGGTGCACAGAGGCA GTGGCAGCAACAAGGATTTGAAACTTGGTGTGTTCTGGAGCCACAGGCAGCGATGCAACCTTGTGT GAGTGTGACGGGGTGGGGTGGGGCGGGAGGCCACGGGGAGGCCGAGGCAGGGCTGGGCAGAGGGGA GAGGAAGCACAAGAAGTGGGAGTGGGAGAGGAAGCCACGTGCTGGAGAGTAGACATCCCTCTCTGGCCG CTGGAGAGCCAAAGCCATGGCCCTGCAGCGTCTGAGCGCCGCTGCTCTTGGTGGCCGGGGTGGG GGCTGCTGTGGTCACTGTGCCACCTCTGCAGGCAGCCCTGTTGGGAGAAGGGACAGCGGGTAAAAAGA GAAGGCAAGCTGGCAGAGGGTGGCACTTCTGGATGACCTCTTAGAAAAGACTGACCTTGAATGCTGT AGAGCGCTGGCCCTCTTCCCTCCCTGCAGGGTAGGGGGCTGAGTTGAGGGGCTCCCTCTGCTCCAC AGAAACCCCTGTTTATGAGTCTGAAGGTGGAACTGCTGCCATGATTTGGCCACTTGGCAGACCTGG GACTTTAGGGCTAACCACTTCTCTTGTAAAGACTTGTGCTCTTGGGAGACGCTCCACCCGTTTCCAAGC CTGGCCACTGGCATCTCTGGAGTGTGGGGGGTCTGGGAGGCAGGTCCCGAGCCCTCTCTCTCCAC GGCCACTGCAGTCAACCCGCTCTGGCCGCTGTCTTGTCTGGCGTGAGAGCCCAATCACTGATCTGATTCG GTCCCTGCTCTCTCTCTCCCTCAGAGTGTGAGCCAGGGCACTGCTCAGCTGTGACCCTAGGTGTTCCT GCCTTGTGACATGGAGAGAGCCCTTCCCTGAAAGGCTTGGCCCTTCTCTGTGAGCCACAGCA GCAGGCTGGGTGTCTTGGTGTGACGTGGTGGCACCAGGATGGAAGGGCAAGGCCACCCAGGGCAGGCCAC AGTCCCGCTGTCCCCACTTGCACCCCTAGCTTGTAGCTGCCAACCTCCAGACAGCCAGCCCGCTGCTC AGCTCCACATGCATAGTATCAGCCCTCCACACCCGACAAAGGGGAACACACCCCTTGAATGGTTCCTT TCCCCCACTCCCAAGCTGGAAAGCCATGCTGTCTGTCTCTGGAGCAGCTGAAATATACATAGATGTG CCTTGGCCCTCCCCATCTGGACCCCTGTGAGTGTGATGTTGGATTTGTCTGTTTATGCTTGGATTCCACAGA GTGACTATGATAGTGAAGAAGAAAAAAGGACGCATGTATCTTGAATGCTTGAAGAG GTTCTAACCCACCCCTCAGAGGTGTCTCTCACCACACTGGGACTCGTGTGGCTGTGTGGTGGCCAC CCTGCTGGGGCTCCCAAGTTTGAAGGCTTCTCTCAGCAGCTGGGACCCAAAGAGACAGCTTCTAG CAGCTAAGGAGGCGCTTACGCTGTGACGAAGGCTGAAGCACAGGATTAGGACTGAAGCGATGATGTCCT CTCTCTACTTCCCTTGGGGCTCCCTGTGTGAGGACAGACTAGGCTTGTGGCTGGCTGGCTTGG GCGCAGGATGTTCTCTTGGTCAAGCCCAAGTCTCATGGCAGTCCCAAGGAGGCTTACAACCTCT GCATCACAAAGAAAAAGAACCCACTGCCAGCTGGGGGATCTGCAGCTCCCAAGCTCCGTGAGCCCTCA GCCACCCCTCAGACTGGTCTCTCCAAGCTCGCCCTCTGGAGGGCAGCGCAGCTCCCAAGGGC CCTGGACACAGCAGGAGGATGGGATGAATGGCTGTCTGGATCTGCTTAGAGGCCAAGCTGCTCTG CTGAGGAAGGATGACTTGACAAGTCAAGGACACTGTTCCCAAGGCTTGAACAGAGCACTCAGCCCGC TGACCTTGACAAACTCCATCTGTGCCATGAGAAAAGGGAAAGCCGCTTTGCAAAACAATTTGCTGCTTAA AGAACTCAGCAGCCTCAGGCCCAATCTGCCACTTCTGGTITGGGTACAGTTAAAGGCAACCCCTGAGG ACTTGGCAGTAGAAATCCAGGGCCCTCCCTGGGGCTGGCAGCTTCTGTGTCAGCTAGAGCTTTACCTGAA AGAAGTCTCTGGGGCCAGAACTCTCCACCAAGAGCCCTCCCTGGCTTGGTGTAGTCCCAAGATCTCTC TAAGTGAAGGATCTGAGAAGGAGAAGGAATGTGGGTAGATTTGGTGGTGTGTAGAGATAAGCCCTC CTACTTACGCCAAGCTTTCGGCTGCATTTCTTCAAGCAGCTGGTTCCTCTCTCTGATGCTTGTG CCTGCTCTCAGCAGCATGGGCTTCTTATACGGAAGGCTCTGGGATCTCCCTTGTGGGGCAGGCTT TGGGGCCAGCCTAAGATCATGGTTAGGGTATCAGTGTGGCAGATAAATGAAAAGGCGACTGGCTT GTGATCTTAAATGAGGACAATCCCCAGGGCTGGGCACTCTCCCTCCCTCACTTCTCCCAAGCTGCA GAGCCAGTGTCTTGGTGGGTAGATAGGATATACTGTATGCCGGCTCCTCAAGCTGTGACTCACTT TATCAATAGTTCATTTAAATAGACTTCAAGTGTGAGACTGTATCTGTTTGGTATTGCTTGTGCTA TGGGGGAGGGGGAGCAATGTGTAAGATAGTTAACATGGGCAAGGGAGATCTTGGGGTGCAGCACTTA AACTGCCTCGTAACCTTTTCATGATTTCAACCAATTTGCTAGAGGGAGGAGCAGCCAGGAGTGA GGCCCTTGGGGTTCCTTTTCCACTGACAGGCTTTCCAGGCAGCTGGCTAGTTTCAATTCCTCCCAAGC CAGGTGAGGCTAGGAAATAGGACATCTGTTGCTTTGGCTGCTGCCCTCTTTCAGGGTCTTAAGCC CACAATCATGCCCTCCTAAGACTTGGCATCTTCCCTCTAAGCCGTTGGCAGCTCTGTGCCACTCTCA CACTGGCTCCAGACACAGCCCTGTGCTTTTGGAGCTGAGATCACTCGCTTCAACCTCTCATCTTTGTT CTCCAAGTAAAGCCACAGGTGCGGGCGAGGGCAGAGTGATCACTCGCTGCCATCTACAGACCTGC AGCTTCATAAAACTTCTGATTTCTCTCAGCTTTGAAAAGGGTTACCTGGGCACTGGCCTAGAGCCCTCA CCTCCTAATAGACTTACCCCATGAGTTTGGCATGTGAGCAGGACTATTTCTGGCACTTGCAGTCCCA TGATTTCTTCGGTAACTCTGAGGTGGGGGAGGACATGAAATCATCTAGCTTAGCTTTCTGTCTGTG AATGCTATATAGTGTATTTGTTGTTTAAACAATGATTTACACTGACTGTGTGTAAGGATGAATTTG GAAATAAGTTATTAICTGATTA	122
55 60	GCACCGCGGAGCTTGGCTGCTTCTGGGGCTGTGTGGCCCTGTGTGTCGGAAAGATGGAGCAAGAAGCC GAGCCCGAGGGGGCGCCGCAACCCCTCTGACCCAGATCCTGCTGCTTTCGCGAGCCAGGAGCAGCGTCCCT CCCCAGATTAGTGGCTACGAGGCGCCAGTGCCTTGGCCCGGASAGTGAATGATCCCCAGGGCCAGGGC GTCTGCTTCCGAGTAGTCACTCCCGTGAAGGAAACTGGGAGTCTTGGAGGACCCCGACTCCAAGC GCGAAAACCCCGGATGGTGGAGCAGGCAAAATGTGCAATACCAACATGCTGTACCTACTGATGGTGT GTAACCACTCACAGATCCAGCTTCCGAAACAAGAGACCCTGGTTAGACCAAGCCATTTGGTTTGAAGT TATTAAGCTCTGTGGTGCACAAAAGACACTATACTATGAAAGAGGTCTTTTATCTTGGCCAGTA TATTAAGCTCTGTGGTGCACAAAAGACACTATACTATGAAAGAGGTCTTTTATCTTGGCCAGTA GATTTGTTGGGTGCCAGCTTCTGTGAAAGAGCACAGGAAAAATATACCATGATCTACAGGAACT TGGTAGTCAATCAGCAGGAATCACTGGACTCAGGTACATCTGTGAGTGAAGCAGGTGTACCTTGA	122

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
5 10 15 20 25	AGGTGGGAGTGCATCAAAGGACCTTGTACAAGAGCTTCAGGAAGAGAAACCTTCATCTTCACATTTGGTT TCTAGACCCTACCTCATCTAGAAGGAGAGCAATTAGTGACACAGAGAAAATTCAGATGAATATATCTG GTSAACGCACAAAGAAAACGCCACAAATCTGATAGTATTTCCCTTTCCTTTGATGAAAGCCTGGCTCTGTG TGTAAATAAGGGAGATAITGTGTGAAAAGAACAGTAGCAGTGAATCTACAGGGACCCCATCGAATCCGGAT CTTGATGCTGGTGTAAAGTGAACATTCAGGTGATGGTTGGATCAGGATTCAGTTTCAGATCAGTTTAGTG TAGAATTTGAAGTTGAATCTCICGACTCAGAAGATATAGCCTTAGTGAAGAAGCACAAGAACCTCAGA TGAAGATGATGAGGTATCAAGTTACTGTGATCAGGCAGGGGAGAGTATACAGATTCATTTGAAGAA GATCCTGAAATTTCCCTTAGCTGACTATGGAAATGCATTCATGCAATGAAATGAATCCCCCCTTCAT CACATTCACAACAGATGTGGGCCCTTCGTGAGAATTTGGCTTCCTGAAGATAAAGGGAAAGATAAAGGGGA AATCTCTGAGAAAGCCAACTGGAAAACCAACACAAGCTGAAGAGGGCTTTGATGTTCTCGATTTGTA AAAAATAAGTATAGTGAATGATCCAGAGAGTCATGTTTGGAGAAAATGATGATAAAAATACACAAGCTCAC AATCACAAGAAAGTGAAGACTATTCAGCCATCAACTTCIAGTAGCATTATTTATAGCAGCCAAAGAA TGTGAAAGAGTTTGAAGGGAAAGAAACCAAGACAAAGAGAGGTGTGAATCTAGTTTGGCCCTTAAT GCCATTTGAACCTTTGTGTGATTTGTCAAGTTCAGCTTAAATATGGTTGCATTTGCCATGGCAAAACAGC ATCTTATGGCCCTGCTTACATGTGCAAGAAAGCTTAAAGAAAAGGAATAAGCCCTGCCAGTTAGACACA ACCAATTCAAATGATTTGTCTAACTTAITTCCTTAGTTGACCTGTCTATAAGACAATATATATTTCTA ACTATATAACCCCTAGGAAATTTAGACAACCTGAAATTTATTCACATATATCAAGTGAAGAAAATGGCTCAA TTTACATAGATTTCTCTCTTTAGTATAAATGACCTACTTTGGTAGTGAAATAGTGAATCTTACTATATA TTTGACTTGAAATAGTAGCTCATCTTTACCCAACCTCCTAATTTTAAATAATTTCTACTCTGCTTAAA TGAGAAGTACTTGGTTTTTTTTTCTTAAATATGTATATGACATTTAAATGTAACCTTATTTTTTTTTT AGACCGAGTCTTCTCTCTTACCCAGGCTGGAGTGCAGTGGGTGATCTTGGCTCACTGCAGCCCTGCC TCCCGGGTTCGCACCATTCTCTGCTCAGCCTCCCAATTAGCTTGGCTACAGTCTCTGCCACCACA CCTGGCTAATTTTTTGTACTTTTAGTAGAGACAGGGTTTACCCTGTTAGCCAGATGGCTCGATCTCC TGACCTCGTGATCCGCCCTCGGCCCTCCAAAAGTGTGGGATACAGGCATGAGCCACCG	
30 35 40 45 50 55	NM_014791 GAGATTTGATTCCTTGGGGGGCGGAAGGGGCCAACCCGGCGATCGAAAAGATTCCTAGGAACGCCGT ACCAGCCGGCTCTCAGGACAGCAGGCCCCCTGTCTCTCTGTCGGGGCGCCCTCAGCCGTGCCCTCCGCC CCTCAGGTTCTTTTTCTAATTCCAAATAAACTTGCAAGAGGACTATGAAAGATATGATGAACCTCTCAA ATATATGAAATACATGAAACTATGGGACAGGTGGCTTTGCAAAAGTCAAACCTTGCCTGCCATATCTCT ACTGGAGAGATGGTAGCTATAAAAACTCAITGGATAAAAAACACTAGGGAGTGATTTGCCCCGGATCAAAA CGGAGATTTGAGGCCCTTGAAGAACCCTGAGACATCAGCATATATGTCAACTTACCATGTGCTIAGAGACAGC CAACAAAATATTCATGGTTCTTGTAGTACTGCCCTGGAGGAGAGCTGTTTACTATATAAATTTCCAGGAT CGCCTGTGAGAAGAGSAGACCCGGGTGTCTTCCGTCAGATAGTATCTGTCTTCTTAITGTGCACAECC AGGGCTATGCTCACAGGGACCTCAAGCCAGAAAATTTGCTGTTTGTATGAATATCATAAATTAAGCTGAT TGACTTTGGTCTCTGTGCAAAAACCAAGGGTAACAGGATTAACATCTACAGACATGCTGTGGGAGCTGTG GCTTATGACGACCTGAGTTAATACAAGGCAAAATCATATCTTGGATCAGAGGCAGATGTTGGAGCATGG GCATACTGTTATATGTTCTTATGTGTGGATTTCTACATTTGATGATGATAATGTAATGGCTTTATACAA GAAGATATGAGAGGAAAATATGATGTTCCCAAGTGGCTCTCTCCCAAGTACCTTCTGCTTCTTCAACAA ATGCTGCAGGTGGACCCAAAGAAAGGATTTCTATGAAAAATCTATGAAACCATCCCTGGATCATGCAAG ATTACAACATCTCTGTTGAGTGGCAAGCAAGAACTCTTTTATTCAACCTCGATGATGATGGCTAACAGA ACTTTCTGTACATCACAGAAAACAAGGCAACAATGGAGGATTTAATTTCACTGTGGCAGTATGATCAC CTCACGGCTACCTATCTTCTGCTTCTAGCCAAGAAGGCTCGGGGAAAACAGTTCGTTTTAAGCTTTCTTT CTTTCTCCIGTGGACAAGCCAGTGCACCCCATTCACAGACATCAAGTCAAATAATTTGGAGTCTGGAAGA TGTGACCGCAAGTGATAAAAAATTTGTGGCGGGATTAATAGACTATGATTTGGTGTGAAGATGATTTATCA ACAGGTGCTGCTACTCCCGAACATCACAGTTTACCAAGTACTGGACAGAAATCAAATGGGGTGGAACTA AATCATTAACCTCCAGCTTATGCAAGACCTGCAAAATAAATAAAGAAACAAAGAAAATGTATATCTCC TAAGTCTGCTGTAAGAAATGAAGAGTACTTTATGTTTCTGAGCCAAAGACTCCAGTTAATAAAGAACCG CATAAGAGAGAAAATCTCACTACGCCAAATCGTTACACTACCCCTCAAAGCTAGAAAACAGTGCCTGA AAGAACTCCATTAATAATACAGTAAATCAACAGGAACAGACAAGTTAATGACAGGTGTCATTAGCCC TGAGAGGGCGTCCGCTCAGTGGAAATGGATCTCAACCAAGCACATATGGAGGAGACTCCAAAAGAAAAG GGAGCCAAAGTGTTTGGAGCCCTGAAAAGGGGGTTGGATAAGGTTATCACTGTGCTCACCAGGAGCAAAA GGAAGGGTTCGCCAGAGACGGGCCAGAAAGACTAAAGCTTCACTATAACGTGACTACAACATAGATTAGT GAATCCAGATCAACTGTGAATGAATAATGTCTATTTCTTCAAAGAAGCAATGTTGACTTTGTACAAAAG GGTATACACTGAAGTGTCAAACACAGTCAGATTTTGGGAAAGTGAACATGCAATTTGAATTAGAAGTGT GCCAGCTTCAAACCCGATGTTGGTGGTATCAGGAGGCGAGCGGCTTAAGGGCGATGCCGGGTTTACAA AAGATTAGTGGAAAGACATCTATCTAGCTGCAAGSIATAATGATGGATTCTCCATCCGCGCGGATGAG TGTGGGTGTGATACAGCTTACATAAAGACTGTTATGATCGCTTTGATTTTAAAGTTTCAITGGAACTACCA ACTTGTTTCTAAAGAGCTATCTTAAAGACCAATATCTCTTTGTTTTTAAACAAAAGATATATTTGTGTA TGAATCAAATCAAGCCCATCTGATTAATGTTACTGCTTTTTTAAATCATGTGTTTTGTATATTAATA ATTGTTGACTTTCTTAGATTCATTCATATAGTGAATGTAAGCTCTTAACTATGCTCTTTGTAATGTT AATTTCTTTCTGAAATAAAACCATTTGGAATATAG	123

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
BG765502	GCAGCGGAGGAGCCAGTCCACGATGGCCCGTCCCTGGTGTGCCTTGGTGTCAICATCITGCTGTCTGC CTTCICCGGACCTGGTGTGTCAGGGGTGGICCTATGCCCAAGCTGGCTGACCGGAAGCTGTGTGCGGACCAG GAGTGCAGCCACCCCTAICTCCATGGCTGTGGCCCTTCAGGACTACATGGCCCCGACTGCCGATTCCTGA CCATTACCCGGGGCCAAGTGGTGTATGTCTTCTCCAAGCTGAAGGGCCGTGGGCGGCTCITCTGGGGAGG CAGCGTTCAGGGAGATTACTATGGAGATCTGGCTGCTCGCCTGGGCTATTTCCCAGTAGCATTGTCCGA GAGGACCAGACCCCTGAAACCTGGCAAAGTCGATGTGAAGACAGACAAATGGGATTTCTACTGCCAGTGAG CTCAGCCTACCGCTGGCCCTGCCGTTTCCCTCCTTGGGTTTATGCAAATACAATCAGCCCAAGTCCAAAA AAAAAAAAAAAAAAAAAAACITCGGAGAAGAGATAGCAACAAAAGGCCGCTTGTGTGAAGCGCCAAAA GTTTTCGCCCAAGAGACCTTCGGCCTCCCCAGGCGCGCGCAAAGGCCCTTGTTTTGACAACCTCTTG GACAACCGGAGGGGCTACCGCCCGGAGACCCCTGTGGTGGACCCCCGGGCAACCCGGTGTGACAGGGTA CTCACCCACCGGCTTTGTGGGGTCCCACCAAGGCCCAAGAGGCCTTTCAAGGCCTATTTCCTT GTTGTAGACCTTGTGTGTCACAGGCGCAAAGAAACCTCGGGGGCTAACAAACGCACGTGCTTGCA GCTCCGAGAAGGCTCTCTCCACCCGAGGGGTGGACGCAACAGGGGAAAGGGCCATCATATTGTGCCC CCGGTGGGCACCAACTCTTTTCCCCATAGAGAGGCCTTAGCACACTAIGTGGGACGTTATTGCCC CTAGAGAAACCGAGCGCCAGAAAAATTCGAAGGGGGGGCGCTTCTCATCATTTTGGCAAAAACCCCTT GTGGGAGTATGCCCGAACTCCTCTGGAACACACAAGCGACACTTGCAGGGGCTCGCAAAAACCTCCT GTTGGGAAGCCGGCTTCACN	124

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_002417	TACCGGGCGGAGGTGAGCGCGGGCGGGCTCTCTGCGGGCGGACTTTGGGTGCGACTTGACGAGCGGTG GTTCGACAAGTGGCCCTTGGCGGGCGGGATCGTCCCAGTGGAAAGAGTGTAAATTTGCTTTCGGCCCTCCCC TACGGATTATACCTGGCCCTCCCCCTACGGATTATACTCAACTTACTGTTTAGAAAATTTGGCCACGAGA CGCCTGGTTACTATCAAAAGGAGCGGGGTCGACGGTCCCCACTTTCCCTGAGCCTCAGCACCTGCTTGT TTGAAAGGGGTATTGAATGTGACATCCGTATCCAGCTTCCTGTTGTGTCAAACAACATTCGAAAATTTGA AATCCATGAGCAGGAGGCAATATTACATAATTTTCAGTTCACAAATCCAACACAAAGTAAATGGGTCTGTT ATTTGATGAGCCTGTACGGCTAAAACATGGAGATGTAATAACTATTATTGATCGTTCCCTCAGGATGAAA ATGAAAGTCTTCAGAAATGGAAGGAGTCAACTGAAATTTCCAAGAAAAATACGTGAACAGGAGCCAGCAGC TCGTGTCTCAAGATCTAGCTTCTCTTCIGACCCTGATGAGAAAGCTCAAGATTCCAAGGCCATTTCAAAA ATCACTGAAGGAAAAATTTTCAGGAAATCCTCAGGTACATATCAAGAAATGTCAAAACAAGACAGTACCCGAG ATGACTCAAAAGACAGTGTGTCTCAGGGAACTAAATGTTCATCTCCAGAACATGCTGGACGTAATGG CAGAAATGCAGCTGATCCATTTTCGGGGATTTTAAAGAAAATTTCCAGCGTTAAATTAGTGGCCGTTAT GGAGAATTGAAGTCTGTTCCCACTACACAATGTCTTGACAATAGCAAAAAAATGAATCCTCCCTTTTGGGA AGCTTTATGAGTCAGTGAAGAAAGAGTTGGATGTAAAAATCACAAAAAGAAAATGTCCTACAGTATTGTAG AAAATCTGGATTACAACTGATTACGCAACAGAGAAAAGAAAGTGTGATGGTTTACAGGGGGAGACCCAA CTGTTGGTCTCGCGTAAGTCAAGACCAAAATCTGTTGGAGCGGCCACGCTGTGCGAGAGCCTGCTTCAC CTGAACAAGAGCTTGACAGAAACAGGGGAGGGGAGAGACGCTGGATCTGTTCAAGCTCCCGACGACGC TGTGGCGCCAGCTTTCCCTCTCTATGAGCCGGCTAAAATGAAGACCCCTGTACAATATTACAGCAACAA AATTCCTCCAAAAACATAAGAACAAGACCTGTATACTACTGGTAGAAGAGAACTGTGAACTGGGTA AAAAGTGAAGGCTTCAAGGCTGGTGTAAAACCTTACTCCAGGAAAGCTTCAACTAGAAAATCGAACACC AGCTAAAGTTGAAGATCCAGCTGACTTCIGCCACTAAGCCAGAAAAATCTCTCTTCCAAAAACAGAGGAAGT ATTCTACAGATGTGAAGTTCTGCCTACGGAACTGAAATTCACAATGAGCCATTTTTAACTCTGTGGC TCACTCAAGTTGAGAGGAAGATCCAAAAGGATTCCTCAGCAAGCCTGAGAAAATTTGGCCACTACAGCTGG ACAGATGTGCTCTGGTTTACCTGGTCTTAGTTTCAGTTGATATCAACAACCTTTGGTGTATCCATTAATGAG AGTGAGGGAATACCTTTGAAAAGAGGCGGTGTCTCTTTGGTGGGCACCTAAGACCTGAATATTGTATG AAAACTTGCCCTCTAATACGCCTCTCAAAAGGGGAGAGCCCAACCAAAAGAAAGTCTCTGGTAAATGCA CACTCCACCTGTCCCTGAAGAAAATCATCAAGGAAACAGCCTCAACCATCAGGAAAACAAGAGTCAGGTTCA GAAATCCATGTGGAAGTGAAGGCCAACAAGCTTGGTTATAAGCCCTCCAGCTCCTAGTCCTAGGAAAACCTC CAGTTGCCAGTGATCAACGCCGTAGGTCCTGCAAAACAGCCCTGCTTCCAGCAGCAAACTCAGACAGA GGTTCCTAAGAGAGGAGGAGAAAAGAGTGGCAACCTGCCTTCAAAGAGAGTGTCTATCAGCCGAAGTCAA CATGATATTTTACAGATGATATGTTCCAAAAGAAAGAGTGGTGCCTCGGAAGCAAACTGATGTTGCAA AATCATGGGCAGATGTAGTAAAACCTTGGTGCAAAACAACAACAATAAAGTCAATAAAAATGTTGCTTCA AAGGTCATGAACAAAAGGCAAGAAAGACCTGCTACTCCAAAGAACCTGTGGGCGAAATTCACAGTCAA TTTAGTACAGGCCACCAAACTCTCTTGTACCATAAATAAGGAAAGCTCATACTGAAAAGTACATG TGCTGTCTCGACCTACAGAGTGTCTCAACAACCTTCAATTTCCAACCAAAAAATGGACTTTAAGGAAGATCT TTCAGGAATAGCTGAAATGTTCAAGACCCAGTGAAGGAGCAACCGCAGTTGACAAGCACATGTCACATC GCTATTTCAAATTCAGAGAAATTTGCTTGGAAAACAGTTTCAAGGAACGTATCAGGAGAAAGAACTCTG TCCCAACCTCAGAGAGTTTTGGAGGAAATGTGTCTTTCAGTGCACAGAAATGCAGCAAAAACAGCCATCIGA TAAATGCTCTGCAAGCCCTCCCTTAAAGACGGCAGTGTATTAGAGAAAATGGAAACGTAGCAAAAACGCC AGGAACACCTACAAAATGACTTCTCTGGAGACAAAACCTCAGATCTGAGACAGAGCCCTTCAAAAACAG TATCCACTGCAAAACAGGTCAGGAAGGCTTACAGAGTTCAGGAATATACAGAAGCTAACCTGTGGAAAATIA GAGTGAAGAAAACAAATACAGAAATTTGTGAGTGCATCTTAAAAGAGGTCAGAAAGCAACACTACTACAA CAAAGGAGAGAGGAGAGATGAAGGAAATAGAAAAGACCTTTTGGAGACATATAAGGAAAATATTGAATTA AAGAAAACGATGAAAAGATGAAAGCAATGAAGAGATCAAGAACTTGGGGGAGAAAATGTGCACCAATGTC TGACCTGACAGACCTCAAGAGCTTGCCTGATACAGAACTCATGAAAGACACGGCAGTGGCCGAAATCTC CTCAAAACCAAGATCATGCCAAGGCCAACAAGAGTGAAGAAAGGCAAAATCACTAAAATGCCCCTGCCAGT CATTACAACCAGAACCAATAAACACCCCAACACACAAAACAACAGTTGAAGGCATCCCTGGGGAAGAT AGGTGTGAAAGAAAGACTCCTAGCAGTGGCAAGTTCAACGGACGCTCAGGGGAGACCACGCACACGGCAG AGAGAGCCAGCAGGAGATGGCAAGGACATCAGAACCTTTAAGGAGTCTCCAAAAGCAGATCCTGGACCCAG CAGCCCGTGTAACTGSAATGAAGAAGTGGCCAAAGAACGCTTAAAGGAAAGGCCCCAGTCACTAGAAAGCCT GGTGGCTTCAAAGAGCTCTCCAGACACAGGTCCTCTGAGGAAATCAATGACTGATGAGAAAATACC AAAATAGCCTGCAAACTCCACCACCAAGATCAGTGGACACTCCAACAAAGCACAAGGAAATGGCCATAAGA GAAGTCTCAGGAAAAGCAGATGTAGAGGAAGAATTTCTAGCACTCAGGAAAATTAACACCAATCAGCAGGGAA AGCCATGCTTACGCCCAAAACAGCAGGAGGATGAGAAAAGACATTAAGCATTTATGGGAACCTCCAGTGT CAGAACTGGACCTGSCAGGAACCTTACCTGGCAGCAAAAGACAGCTACAGACTCTTAAGGAAAAGGCC AGCCTCTAGAAGACCTGGCTGGCTTTAAAGAGCTCTTCCAGACTCTGGTCAACCCGAGGAATTAGTGGC TGCTGTTAAAACCACTAAAATACCCCTGCGACTCTCCACAGTCAGACCCAGTGGACACCCCAACAGCACA AAGCAACGACCCAAAGAAAGTATCAGGAAAGCAGATGTAGAGGGAGAACTCTTACGCTGACAGGAATCTAA TGCCATCAGCAGGCAAGGCAATGCACACGCCCTAAACCATCAGTAGGTGAAGGAAAGACATCATATATT	125

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
	AACATACTCAATCCCTGTCGCCAGGCGGATCCCTGGAAGCCCTTTTCGCAGCACGCTATCCCTCCAAAGCCATTGTAATGTGTGTACAGTGTGTATAAACCTTCTTCTCTTTTTTTTTTTTTTAAACTGAGGATTGTC	
NM_002467	GACCCCGAGCTGTGTGTGCTCGCGGGCCGCCACCGCCGGGCCCGCGCTCCCTGGCTCCCTCCCTGCGCTCAGAGGGCCAGGGCTTCTCAGAGGCTTGGCGGAAAAAGAACCGGAGGGGATCGCGCTAGATATAAAA GCCGGTTTTTCGGGGCTTTATCTAACTCGCTGTAGTAATCCAGCGAGAGGCAGAGGGAGCGAGCGGGCGG CCGGCTAGGGTGGAAAGAGCCGGGGCAGCAGAGCTGCGCTGCCGGGCTCCTGGGAAGGGAGATCCGGAGCG AATAGGGGGCTTCGCCCTCGGCCAGCCCTCCCGCTGATCCCCAGCCAGCGGTCCGCAACCCCTTGGCGC ATCCACGAAACTTTGCCATAGCAGCGGGCGGGCCTTTGCACTGGAACCTTACAACACCCGAGCAAGGAC GCGACTCTCCCGACGCGGGGAGGCTATCTGCCCAITTTGGGGACACTTCCCGCGCGCTGCCAGGACCCCG TTCTCTGAAAGGCTCTCCTTGCAGCTGCTTACAGCTGCTGCCCCACCCCGCCCTTGTCCCTACGCCCGCTC TCCCGCAGCATGCCCTCAACGTTAGCTTACCAACAGGAACATGACCTCGACTACGACTCGGTGCAG CCGTATTTCTACTGCCAGGAGGAGAACTTCTACCAGCAGCAGCAGCAGAGCCAGCTGCCAGCCCCGG CGCCACGCGAGGATATCTGGAAGAAATTCGAGCTGCTGCCACCCCGCCCTTGTCCCTACGCCCGCTC CGGCTCTGCTCGCCCTCCTACGTTGGGTCACACCCCTTCICCTTCGGGGAGACAACGACGGCGGTCC GGAGCTTCTCCAGCGCCGACCGCTGGAGATGGTGACCGAGCTGCTGGGAGGACATGGTGAACCAAGA GTTTCATCTGCGACCCGAGCAGCAGACCTTCAATAAAACATCATCATCCAGGACTGTATGTGACGGCG CTTCTCGGGCCCGCCAGCTCGTCTCAGAGAAGCTGGCCCTCCTACCAGGCTGCGCGCAAGACAGCGGC AGCCCCGAAACCCCGCCCGGGCCACAGCGTCTGCTCCACTCCAGCTGTACTGCAGGATCTGAGCGCGG CCGCTCAGAGTGCATCGACCCCTCGGIGGCTTCCCTACCCTCTCAACGACAGCAGCTCGCCCAAGTC CTGCGCTCGCAAGACTCCAGCGCCCTTCTCTCCGCTCCGGATCTCTGCTCTCTCGACGGAGTCCCTC CCGACGGCAGCCCCGAGCCCTGGTGCCTCATGAGGAGACACCGCCACCACCAGCAGGACTCTGAGG AGCAACAAGAAATGAGGAAGAAATCGATGTTGTTTTCTGTGAAAAAGAGGAGGCTCCCTGCAAAAAGTC AGACTCTGGATCACTTCTGCTGGAGGCCACAGCAAACTCCTCAGACCCACTGGTCCCAAGAGGTGC CAGTCTCCACAGATCAGCAAACTACGAGCGCCCTCCCTCCACTCGGAAGGACTATCTGTCGCAAGA GGGTCAAGTTGGACAGTGTGAGGTCGAGGACAGATCAGCAACACCGAAAATGCACCGCCCGAGGTC CTCGGACACCGAGGAAATGTCAAGAGGCGAACAACACACTCTTGGAGCGCCAGGAGGAAACGAGCTA AAACGGAGCTTTTTTSCCTGCGGTACAGATCCCGGAGTGGAAAAAATGAAAGGCCCCCAAGGTAG TTATCCTAAAAAAGCCACAGCATACATCCTGTCCGTCCAAGCAGAGGAGCAAAAAGCTCATTTCTGAAGA GGACTTGTGGGGAACGACGAGAACAGTTGAAACACAACTTGAACAGCTACGGAACCTTTTGGCGTAA GGAAAAAGTAAGGAAAAAGTCTCCTTCTAACAGAAATGTCCTGAGCAATCACCTATGAACCTTGTCTCAA GCATGATCAAATGCAACCTCACAACTTGGCTGAGTCTTGGACTGAAAGATTTAGCCATAATGTAACCT GCCTCAAATTTGACTTTGGGCATAAAGAACTTTTTATGCTTACCATCTTTTTTTTTTCTTTAACAGAT TTGTATTTAAGAAATTTTAAAAAAATTTAAGATTTACACAACGTTTTCTCTGIAAATATTGCCATIAA ATGTAATAACTTTAATAAAACGTTTTATAGCAGTTACACAGAAATTCAACTCTAGTATAATAGTACCTAGT ATTATAGGTACTATAAACCCCTAATTTTTTTTTTAAAGTACATTTGCTTTTTAAAGTTGATTTTTTCT ATTGTTTTTAGAAAAATAAAAAAATGCGCAATATATCAITGAGCCAAATCTTAAAAAATAAAAAA	128
BC013732	GTGGGAGGATTGCATTAGTCTAGTTCCTGGTTCGGGCTGAAATAACCTGCTCTCCAAAATGTCACAA AAGTACTTAAAGTACAGTTCCCCAAACCCAGACACCAAGACAAAGAAATCCATGTGTGTGACTGAAGGAA GTCTGGGAGAGCCCGCAGCTGCAGCTGGATGTGAACCTGCAACTCCAAAGTGTGTCAGACTCAAGGCAA GGCAGCTAGGCTTCCAGACCCCTACTAAGTCAATGATCCAGCACTGCCCTGCCAGGACATAAATCCCT GGCACCTCTTGCCTCTGCAAGGAGGGGCAAGCAGCTTCCAGGAGCCCTTGGGAGTCCCTCCAAAGAGAGT CTAGGGTACAGGTCCGAAATGAGAAGACACAGAAAGGCAAGCCAGGGGCACGTGAGATGGTAAAAAGAGA TCTGAAGGGATCCAGAAATCAAGCCAGGAAGAAAGCAAACTGCTCTTCTGGATTAAAACCTGAAGATCAA CCTACTTTCAACTTACTAAGAAAGGGGATCATGGACATTTGAAGCATATCTTGAAGAAATGGCTATAAGA AGTCTAGGAACAAATGGACTTGGAAACATTAACCTGATATCTTCAACACCAGATCCGAGCTGTCCCTTT TGAGAACCTTAACATCCATTTGGGGATGCCATGGACTTAGGCTTAGAGGCCATTTTTGATCAAGTTGTG AGAAGAAATCGGGGTGCATGGTGTCTCCAGGTCATCATCTTCTGTACTGGGCTCTGACCCTATTTGCTT TTGAGACCCAGATGTTGGGAGGGTATGTTTACAGCACTCCAGCCAAAAAATACAGCACTGGCATGATTCA CCTTCTCCTGCAAGGTGACCATTTGATGGCAGAACTACATTTGCTGATGCTGGGTTTGGAGCTCATACCAG ATGTGGCAGCCCTGAGTAAATTTCTGGGAAGGATCAGCCCTCAGGTGCCTTGTGCTCTCCGTTGACGG AAGAGAATGGATTTGCTATCTAGACCAAAATCAGAAGGGAACAGTACATCCAAATGAAGAAATTTCTTCA TTCTGATCTCCTAGAAAGCAGCAAAATACCGAAAAATCTACICCTTACTCTTAAGCCCTCGAACAAATGAA GATTTTGGTCTATGAATACATACCTGCAGACATCTCCATCATCTGTGTTTACTAGTAAATCATTTTTCTT CCTTGCAGACCCCGAGTGGGGTTCAGTGTGGTGGGCTTCCCTCCATAGGAGATCAATATAA GGACAAATACAGATCTAATAGAGTTCAAGACTCTGAGTGAAGGAAATAGAAAAAGTGTGAAAAATATA TTAATATTTCTTGCAGAGAAAGCTTGTGCCAAACATGGTGAAGATTTTTTACTATTTAGAATAAGG AGTAAAACAACTGTCTATTTGTATCCAGCTCACCAGTTATCAACTGACGACCTATCATGTATCTTCT GTACCCCTACCTTATTTTGAAGAAATCTAGACATCAAAICATTTACCTATAAAAAATGTCAATATATA TAATTAACAGCTTTTTAAGAAACATAACCAAACTTTTCAAAATAATAATAATAATAATA ATGCTTTTTAAAGATGCCCTGTTGTTTACTTGGAAATTTGGTATTTATGCTAGAAAGCTTTTTAATGTTGG TTTTATTGTTGAATTCCTAGAAAAGTTTTATGGGTAGATGAGTAAATAAAATATTGAAAAAACTTATTG TCTATAAAGTATATTAACATTTGTTGCTAATATAAAAAAATAAAAAA	129

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_014321	<p>GCSCGCGGGTTTCGTTGACCCGCGGGGTTTCACGGGAATTGTTTCGCTTTAGTGCCGCGGCCATGGGGTCGG AGCTGATCGGGCCCTAGCCCCGCGCCIGGGCCCTCGCCGAGCCCGACATGCTGAGGAAAGCAGAGGAGTA CTTGCGCCCTGTCCCGGTTGAAAGTGTGTGCGCCCTCTCCGCACGCACCACGGAGACCAGCAGTGCAGTCAATG TGCCITGGACCTTGCAGCTTCCITGGATGAAGTGCCTTGGACAGGGCTTATTTAATTAACCTTCTGTGTT TGAACAAGGAGACATATCAGAGCTGTCTTAAATCTTTTGTAGTGTCTACTGGGCGTGAATTCAAATATTGG AATAAGAGACCTAGCTGTACAGTTTAGCTGTATAGAAGCAGTGAACATGGCTTCAAAGATACTAAAAAGC</p> <p>TATGAGTCCAGTCTTCCCCAGACACAGCAAGTGGATCTTGACTTATCCAGGCCACTTTTCACTTCTGTCTG CACTGCTTTCAGCATGCAAGATTCTAAAGCTGAAAGTGGATAAAAAACAAAATGGTAGCCACATCCGGTGT AAAAAAAGCTATATTTGATCGACTGTGTAAACAACCTAGAGAAGATTGGACAGCAGGTGACAGAGAACCCT GGAGATGTAGCTACTCCACCAGGAAGAGAAAGAAAGATAGTGGTGAAGCCCAAGCAAAGGAAATGGAGA AGGTAGAGGAGATGCCACATAAACACAGAAAGATGAAGATCTGACACAGGATATGAAGAATGAAAAAG AAAAAATTTGGAAAAATGCTGCCAGTGTCTCAAAGGCTACAGCAGAGTGAATTCAGCTTCCAAACTGGIAT ACATTCAAACTGATAGTACATTCGCAICTCCAGGAAGACTTGACGGCTTTGGGATTTTGTITAAACITTT TATAATAAGGATCCTAAGACTGTTCCTTTAAATAGCAAAGCAGCCTACCTGGAGGCTAAGTCTGGGCAG TGGCTGGCCCCCTGGTGTGAGCATTAGACCAGCCACAGTGCCTGATGGGTATAGCCTTAAGTGTCTTCTCT ACAAAAATGGAAATGGAGGCCGGCCGAGTGGCTCAGCCTGTAACTCCAGCACTTTGGGAGGCCAAGGTG GGTGGATCACCTGAGTGCAGGAGCTCGAGACCAGCCTGGCCAACATGGTGAAACCCCATCTCTACTAAAA ATACAAAAATTAGCCAGGTGTGATGGTGCATGCCTGTAATCCCAGCTCCTCAGTAGGCTGAGACAGGAGC ATCACTTGAACGTGGGAGGCAGAGTTCAGTGCAGCCGAGATTGCACCACCGCACTCCAGCCTGGGTGAC AGAGCGAGACTTATCTCATAAATAAATAGATAGATACTCCAGCCTGGGTGACAGAGCGAGACTTATAGAT AGATAGATAGATAGATGGATAGATAGATAGATAGATAGATAGATAGATAAACGGAATTGGAGCCATTTTG CTTTAAAGTGAATGGCAGTCCCTTGTCTTATTCAGAAATATAAAATTCAGTCTGAAATGGCAICTTACAGATT TTACTTCAATTTTGTGTACGGTATTTTTTATTTTACTAAATCAATATATTGTACAGCCIAAGTTAATAA ATGTTATTTATATATCAAAAAAATAAAAAAAAAA</p>	130

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
NM_000926	AGTCCACAGCTGTCACTAATCGGGGTAAGCCTTGTGTATTGTGCGTGTGGGTGGCATTCCTAATGAGA ACTAGCTTCACTTGTCAATTTAGTGAATCTACAACCCGAGGCGGCTAGTGTCCCGCACTACTGGGATC TGAGATCTTCGGAGATGACTGTCCGCCGAGTACGGAGCCAGCAGAAAGTCCGACCCCTTCTGGGAATGGG CTGTACCGAGAGGTCCGACTAGCCCCAGGGTTTTAGTGAGGGGGCAGTGGAACTCAGCGAGGGACTGAGA GCTTACACGATGACGAGTTTGTATGCCAGAGAAAAGTTCGGGAGATAAAGGAGCCGGGTGTCTACTAAT TGCCGTCCGAGCCGAGCCACTCAAGTGCCGGACTTGTGAGTACTTGTGCTCCAGTCTCCGGACAGAA GTTGGAGAACTCTTTGGAGAACTCCCCGAGTTAGGAGACGAGATCTCCTAACAAATTACTACTTTTTCTT GCCTCCCCACTTGCCTCGCTGGGACAAAACGACAGCCACAGTTCCTGACGACAGGATGGAGGCCAA GGCAGGAGCTGACCAAGCCGCCCTCCCCGCCCCGACCCAGGAGGTGGAGATCCCTCCGGTCCAGCC ACATTCAAACCCACTTTCTCTCCCTCTGCCCTATATTCGCCAAAACCCCTCTCTCTCTCTCTCTCTCT TCTCTCTGGAGACGGGGAGGAGAAAAGGGGAGTCCAGTCTGACTGAGCTGAAGGCAAAAGGGTCC CGGGTCCCCAGTGGCGGGGGGGCCCTCCCCGAGGICGGATCCCCACTGCTGTGTCTGCGCCAGCCG CAGTCCGTTCGGGGGAGCCAGACTCGGACACCTTGCCTGAAGTTTCGGCCATACCTATCTCCCTCGA CGGGTACTCTTCCCTCGGCCCTGCCAGGGACAGGACCCCTCCGACAAAAGACCGAGGACCCAGCTCG CTGTGGACGTGGAGGGCCATATTTCCAGAGCTGAAGCTACAAGGGGTGCTGGAGCCAGCAGTCTACTC CCCCAGAAAAGGACAGCGGACTGCTGGACAGTGTCTTGGACACTCTGTTGGCGCCCTCAGGTCCCGGCA GAGCCAAACCCAGCCCTCCCGCTTGGAGGTACACAGCTCTTGGTGGCTGTTTGGCCCGAACTTCCCGAA GATCCACCGGCTGCCCGCCACCCAGCGGGTGTGTGCCCGCTCATGAGCCGGTCCGGTGCAGAGGTTG GAGACAGCTCCGGGACCGCAGTGCCTATAAAGTCTGCCCCGGGGCTGTCAACAGCCCGCAGCTCTCT GCTCCCGGCTCTGAGAGCCCTCACTGGTCCGGGGCCCGAGTGAAGCCGTCTCCCGAGCCCGCTGCGGTG GAGGTTGAGGAGGAGATGGCTCTGAGTCCGAGGAGTCTGCGGGTCCGCTTCTGAAGGGCAAACTCGGG CTCTGGGTGGCGGGCGCTGGAGGAGGAGCCGCGCTGTCCCGCCGGGGCGGCAGCAGGAGGCGTCCG CCTGGTCCCCAAGGAAGATTCCCGCTTCTCAGCGCCCGAGGGTCCCGCTGGTGGAGCAGGACGCGCGATG GCGCCCGGGCGCTCCCGCTGGCCACCCAGCGGTGATGGATTTTCATCCAGTGCCTATCCTGCTCTCAATC ACCTCTATTGGCAGCCCGACTCCGCGAGTGTGGAAGACGAAAGTTACGACGCGGGGGCGGGGCTGCG CAGCGCTTTGCCCCCGCGGAGTTCACTTGTGCTGCTGCTCCACCCCGGTGCTGTAGGCGACTTCCCC GACTGCGCGTACCCGCCCCGAGCCGAGCCCAAGGACGACGCGTACCCTCTCTATAGCGACTTCCAGCCG CCCTCTAAAGATAAAGGAGGAGGAGGAGGCGCGGAGGCTTCCCGCGCTTCCCGCGCTTCTACCTTGT GGCCGTGCCAACCCCGAGCCTTCCCGATTCCCGTGGGGCCACCGCCCGCTGCGCGCGGAGCG ACCCATCCAGACCCCGGGAAGCGCGGTGACGGCCGACCCGCAAGTGCCTCAGTCTCGCTCGCTCTCT CCTCGGGTTCGACCTTGGAGTGCATCTGTACAAAGCGGAGGGCGCGCCGCCCCAGCAGGCGCGTTCGC GCGCGCCCTGCAAGCGCCGGGCGGAGCGGCTGCTTCCCGGGGACGCGCTGCCCTCCACTCCCTCC GCCTGTGCCCGCCCGCCCGGGGGCGCCCGCGCTTACCTTGCCTCGGCTCAACGGGGTCCCGCAGC TCGGTACCGAGGCGCGCTGTCAAGGAGGGCTGCGCGAGGCTACCGCCCTATCTCAACTACCTGAG GCGGATTCAGAAGCCAGCCAGAGCCACAATACAGCTTCGAGTCAATACCTCAGAAGATTGTTAATC TGTGGGATGAAGCATCAGGCTGTCAATATGTTGCTTACCTGTGGGAGCTGAAGTCTTCTTTAAGA GGCAATGGAAGGGCAGCCAACTACTTATGTGCTGGAAGAAATGACTGCATCGTTGATAAAATCCGCG AAAAAACTGCCAGCATGTCCCTTAGAAAGTGTGTCAAGGCTGGCATGTTCTTGGAGTCCGAAAATTT AAAAGTTCAATAAAGTCAGAGTTGTGAGAGCACTGGATGCTGTGCTCTCCACAGCCAGTGGCGTTC CAAATGAAAGCCAGCCCTAAGCCAGAGATTCACTTTTTCAACAGGTCAAGACATACAGTTGATTCACCC ACTGATCAACCTGTTAATGAGCATTGAACCCAGATGTGATCTATGCAAGGACATGACAACACAAAACCTGAC ACCTCCAGTCTTTGCTGACAAGTCTTAATCAACTAGGGGAGAGGCAACTTCTTTCAGTACTAAGTGGT CTAATCAITGGCAGTTTTTCGAAACTTACATATTGATGACCAGATAACTCTCAITCAGTATTCTTGGAT GAGCTTAATGGTGTGTTGGTCTAGGATGGAGATCCTACAAACACGTCAGTGGGCAGATGCTGATTTTTC CCTGATCTAATACTAAATGAACAGCGGATGAAAGAAATCATCTATTCTATTATGCTTACCATGTGGC AGATCCACAGGAGTTTGTCAAGCTTCAAGTGTAGCCAGAAAGAGTTCCTCTGTATGAAAGTATTTGTACT TCTTAATAACAATTCCTTTGGAAGGGTACGAAAGTCAAACCCAGTTTGGAGATGAGGTCAAGCTACATT AGAGAGCTCATCAAGGCAATTTGGTTTGGGCAAAAAGGAGTTGTGTGAGCTCACAGGCTTCTATCAAC TTACAAAATCTTGTAACTTGCATGATCTTGTCAAAGAACTTCAATGTACTGCTTGAATACATTTAT CCAGTCCCGGCACTGAGTGTGAATTTCCAGAAATGATGCTGAAGTTATTGCTGCACAATACCCAAAG	131

(continúa)

5

10

15

20

25

30

35

40

45

50

55

60

65

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
	ATATTGGCAGGGATGGTGAACCCCTTCTCTTTTCATAAAAAGTGAATGTCATCTTTTTCTTTTAAAGAAAT TAAATTTTGGGGTATGCTCTTTTGGTTTGGTCAGGATTATGAGGTCCTGAGTTTTTATAATGTTCTTCTG AAAGCCCTTACATTTATAACATCAIAGTGTATAAATTTAAAAGAAAAATTTCTGAGGTCCTAATATTTCTI TTTATAAAGTATAATTAGAATGTTTAACTGTTTTGTTTACCCATATTTTTCTTGAAGAAATTTACAAGATTG AAAAGTACTAAAATTTGTTAAAGTAACTATCTTATCCATATTTATTTTCATACCATGTTAGGTTGAGGATTTI TAACTTTTGCATCTAAACAAATCAICGACTTAAGAGAAAAATCTTACATGTAATAACACAAAGCTATTAT ATGTTATTTCTAGGTAACCTCCTTTTGTGCAATTAATTTCCAAAAATGAACCTTTAAAAATGGTATGCAA AATTTTGTCTATATATTTTGTGTGAGGAGGAAATTCATAACTTTCCCTCAGATTTTCAAAGTATTTTTTA ATGCAAAAAATGTAGAAGASTTTAAACCCACTAAAATACATTGATGTTCTTCAAACCTAGSCAAAACCAAC TCAIATGTTAAGACCACTTTCCAGATTGGAAACACAAAATCTCTTAGGAAGTTAATAAGTAGATTCAATC ATTATGCAAAATAGTATTTGGGGTTTTGTAGGTTTTTAAAATAACCTTTTTTGGGGAGAGAATTTGCTCTCI AATGAGGTTATTGGCAGTGGACATAAAGAAATCAGAAATTTATGGCCTAACGTTACTCTCTTACCAACTGTGG CATGCTGAAAGTTAGTCACTCTTACTGATTTCTCAATCTCTCACCTTTGAAAGTAGTAAAATATCTTTCC TGCCAAITGCTCCTTTGGGTCAGAGCTTATTAACATCTTTTCAAATCAAAGGAAAGAAAGAGGAGGAGG AGGAGGAGGAGGTTACAAATTCACATACCTTTCTCCTCTTTATCCTCCACTATCATGAATTTCAIATATG TTTGAGCCTGCTGGAGTCTGGAGCCATATAAAGTGGTCAATAGTGTGCTGACTGAGAGTTGAATG ACAITTTCTCTGTCTGGTATTACIGTAGATTTCCGATCATTCTTIGGTACATTTCTGCATATTTCTG TACCCATGACTTTATCACTTTCTTCTCCCATGCTTTATCTCCATCAATTTATCTTCAATTTCTTTAAATTT TCCACCTTTGCTTCTACTCTTGTGAGATCTCTCCCTTTACTGACTATAACATAGAAGAAATAGAAGTGTAT TTTATGCTCTTAAAGGACAATACITTAGATTTCTTGTCTTAAAGTTTTTAAACTGAATTTAAATGGAAATTA TTTCTCCTTAAAGCAAAATTTCCACAAAACAATTTATTTCTTATGTTTATGTAGCCTTAAATTTGTTTTGTA CTGTAACCTCAGCATAAAACCTTTCTTCAATTTCAATTTCAATTTCAACAAAATATTGATTTGAATACCTGGI ATTAGCAAGAAAATTTGTCTAATAAGCCTTTATGAGAAATTTGGAGCTGAAGAAAAGACATATAACTCAGG AAAGTTACAGTCCAGTAGTAGGTATAAATACAGTGCCTGATAAATAGGCATTTTAAATTTGTACACTC AACSTACTAGTGTAGTGCAAAACAATTTACATATAATTTACTGATACCCATGCAGCAAAAGTACTA ACTTTAAATATTAATAACACCTTTATGTGTCAAGTAAATCATTGTCATTAATCTTATTTGAAAAGGCTTT CAATATATTTCCCCCAAAATGTCTATCCCAAGAAAAAGTATTTTTAACTCTCCCAAAATATAATAGTTA CAGGAAATCTACCTCTGTGAGAGTGCACACCTCTCAGAATGAACGTGTGACACAAAGAAATGAATGTAGG TCTATCCAAAAAAACCACAAAGAACAAAACAATATTTAGCCCTTTATGCTTAAAGTGTAGGACICAG GGAACAGTTGATTTGTGACCTATTTATTTATCTGATTTCTTGTACTTTGAATTTAAACCAATATTTTGTATG ATATAAATCATTCCACCAGCATATAATTTAATTTCCATAATAACTTTAAAAATTTCTAATTTCACTCAAC TATGAGGAAATAGAATGTGGTGGCCACAGTTTTGGCTTTTGTAAAATGTTTGTATCTCTGGATGTGTAT CTCIGTCTGCAATGTAGATGCTTAAACACIAGGATTTAATATTTAAGGCTAAGCTTTAAAATATAAGTAC CTTTTTAAAAGAATATGGCTTCCAAAAATGGAAAAATACCTAATTTCTAATCTTTTTCTCTACAAAGTCT CTATCTACTAATGTCTCCATCTACTATTTAGTCACTCAATACCATTTATCTTCAATTTTACATGCTCGGTCTT TCTGGTAGCTTAAATTTGACACTAAATCATAAAGAACAGGTTACATATCAGGAAATCTTTGAAGGTTAC TGAATAGATTTCTGAGTTAATGAAAATATTTTCTGTAAGAGGTTTGAAGGCCATTTGAGTCTAAAGC ATTATACCTCCATTTACAGTAGTTATGTGACAAATTTGTGTGTGTTTAAATGTTTAAAGATCTGGCAGCTTT TTAATAAGGCAATGCTATGCTATTTTTTCCCATTTAAACATTAAGATAATTTATGCTATACAGATGATAI GCAATATGATGAACAATATTTTTTTGGCAAACTATGCTTTGTAAGTAGCCATGGAAATGTCACACCTGI AACTTAAATTTCCACAGATGTCATGTGTTTGTATGATGGCAGCTGTTGGAGATAACTGACATAGGACTGTI GCCCCCCCTCTGCCCCTACTAGCIGGATGAGATTAAGCAAGTCAATTAACCTGCTGATTTAAACCTG CCTTTCCCAAGTGTCTTGAATGAATAGAATGGAAACCAAAAAACGTAACAGGCTTTAGCAAAATAG TAAITGCTACTATTTTGTCTTCAITTAAGCCATAGTTCTGGCTATAAATTTTATCAAACCTACCAGCTATAI TCTACAGTGAAGCAGGATTTAGAAAAGTCTCACTGTTTTTATTTATGTCACCATGTGCTATGATATATTT GGTGAATTTCAATTTGAAATTTAGGCTGGAAGTATTCAGTAAATTTCTTCTGCTGAAAAATACAGTGT TGAGTTTAGGGCTGTTTTATCAAAGTTCTAAAAGCCATATCACTCTTCCATTTGATAGCATTTTAAAAATA ATGACACTGATTTTAACTTTTTTAAAGTGTCTTTTTTGAACAGAGAGCCCTGACTAGAACACAGCCCTCCA AAAACCCATGCTCAAAATTTTTTIACATAGGAGCAATTTCCAAAAAGGGAACAAATGGGTTTAGAAATTA CAATGAAGTCACTCAACCAAAAAACAATCCCTATCCCTAAGAGGTTATGATATAAATGCCCCAAGAAAA TCTATGCTGCTTTAAATCTGTCTTTTATGCTTTTGAAGGATGGCTATTACATTTTCTGTTTTGCTGTG AATACCGTAGCAGTTTCTCTTCAATTTCCATCACTTATCTTCCACATCAGAAATGAGGATAGAATGAATCA TTTTAAAAACTTTTACAACCTCCAGAGCCATGTGCTAAGAAGCATTCAAACCTTGCCAAAAACATACATTT TTTTCAAATTTAAGATACCTTATTTTGTATTTCAATAGCTCAACACTGTGGTCCCCACTGATAAAGT GAAGTGGACAAGGAGACAAGTAAATGGCATAAGTTTGTTTTTCCCAAAGTATGCTGTTCAATAGCCATTTG SATGTGGGAAATTTCTACATCTCTTAAAAATTTACAGAAAATACATAGCCAGATAGTTACAAAAATTTT ACCAAGTCCATAAATGCTTATCTTACTTCACTAAGTCAATGAAATCATTTTAATGAAAAGAACATCACT AGGTTTTGTGTTTTCTTTTTTTCTTATTCATGGCTGAGTGAACCAACAATCTCTGTTTTCTCCCTAGCAT CTGIGGACTATTTAATGTACCATTATTTCCACACTCTATGGTCTTACTAAATACAAAATTTGAACAAAAG CAGTAAAACAACTGACTCTTCCCCCAATTTATAAATAAATCCAAAGCCAGATTTAGTCAACATCCAATAAG ATGAATCCAAGCTGAACCTGGSCCTAGATTTATGAGTTTCAAGTTGGATCACATCCCTATTTTATTAATAAAC TTAGGAAAGAGGCTTACAGACCATCAGTTAGCTGGAGCTAATAGAACCCTACTCTTAAAGTTCCGGCC TAGAATCAATGTGGCTTTAAAAGCTGAAAAGAAAGCAGGAAAGAACAGTTTTCTTCAATATTTTGTCCACC CTGTCACITGGAGAAAATTTAAGAAATTTGGGGGTGTGGTATGAGTTAAACACAGCACTTTCTCATGGCA GAAATTTTCAATACATACCTTCTCTGAATATCTTAAACAAAAGCAAGAAAAACCAAGGGGTTTTGT	

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
	<p>GAGGGTGC AAGAATAAAATGCCATTACATTTCCACGATACACTCGGCCAGT CAGACGATCTCATTCACCA CCATCATGACTACCATCATATICTCCATCATCACCACCACCAAAACCACCATCCICACAGICACAGCCAG CGCTACTCTCGGGAGGAGCTGAAAGATGCCGGCGTCGCCACTCTGGCCCTGGATGGTGATAATGGGTGATG GCCTGCACAATTTACGGATGGCCATAGCAATGGTGCTGCTTTTACTGAAGGCTTATCAAGTGGTTAAAG TACTTCTGTTGCTGTGTTCTGCATGAGTTGCCCTCATGAATAGGTGACTTTGCTGTCTACTAAAGGCT GGCATGACCGTTAAGCAGGCTGTCTTTATAATGCATTGTGAGCCATGCTGGCGTATCTTGGAAATGGCAA CAGGAATTTTCATTGGTCATTATGCTGAAAATGTTTCTATGTGGATATTTGCACITACTGCTGGCTTATT CATGTATGTTGCTCTGTTGATATGGTACCTGAAATGCTGCACATGATGCTAGTGACCATGGATGTAGC CGCTGGGGGTATTTCTTTTACAGAAATGCTGGGATGCTTTTGGGTTTGGAAATATGTTACTTATTTCCA TATTTGAACATAAAAATCGTGTTCGTATAAATTTCTAGTAAAGTTTAAATGCTAGAGTAGCTTAAAAAG TTGTCATAGTTTTCAGTAGGTCATAGGGAGATGAGTTTGTATGCTGTACTATGCAGCGTTTAAAGTTAGTG GGTTTTGTGATTTTGTATGAAATATGCTGTCTGTACAAAAGTCAGTTAAAGGTACGTTTAAATATTTA AGTTATCTATCTGGAGATAAAAATCTGTATGTGCAATTCACCGGATATTACAGTTTATATGTAACA GAGATTTGGCATGACATGTTCTGTATGTTTACAGGAAAATGCTTTAAATGCTTTTCAAGAACTAACAC AGTTATCTCTATCTGATTTTAGGTCCTGAGAACTGCTGGTGTTTAGGAATAAGAAITGTCATGAAG CCTAAAATACCAAGAAAGCTTATACCTGAATTTAAGCAAAGAAATAAAGGAGAAAAGAGAAAGAACTGAGA ATTTGGGGAGGCATAGATTTCTATAAAAATCACAATAATTTGTTGTAATTTAGAGGGGAGAAATTTAGAAAT AAGTATAAAAAGGCAGAAATAGTATAGAGTACATTTAACAATTTTGTGAGGATATTTCCCGTAAA AACGTAGTGAGCACTTTTCATATACCTAATTTAGTTGTACATTTAACTTTGTATAATACAGAAATCTAAAT ATATTTAAATGAATTCAGCAATATATCACTTGACCAAGAAATTTGGAATTTCAAAAATGTTGCTGGGGTAT ATACCAGATGAGTACAGTGGTGTGTTTATGATACACCAGACGCGGTTTATGCAAGTTATATATCACCA AAAGCTGTATGACTGGATGTTCTGTTTACCTGGTTTACAAAATATCAGAGTAGTAAACTTTGATATAT ATGAGGATATTAAGACTACACTAAGTATCATTTGATTCGATTCAGAAAGTACTTTGATATCTCCAGTGC TTCAGTGCATACATTTGAGCAATTTGCTTTTATATACGGTACTGTAGCCATACAGGCGCTGTCTGTGGC ATTCTCTAGATGTTCTTTTACACAATAAATTCCTTATATCAGCTTGAAAAAATTTTTTTTTT</p>	
AK098106	<p>AACGCATTTGGCGCGCGCGCGGGCTGCAGACGGCTGCGAGGCGCTGGGCACAGGTGTCCTGATGGCAAA TTCAAGGGCCACGCGCTTCCAGGAGTTTCTTCTGTATCATTTGGGCTGTGTTGGTCAAGTACCCG CTGAAGTACTTTAGCCACACCGCGGAAGAACAGCCCACTACATTACTATCAGCGTCTCGAGATCGTCGAAG CGCAATTAGGACTTTGTTTTCCTGCTACTGGGATCCTGGCAGAGCAGTTTGTTCGGATGGGCGCCACT GCACCTCTACCATGAGAACCCACTGGATAAAGTTAATGAATTTGGCAGCACAGCCATGACTTCTTT GCAGTCTCAGGAATTTGTGACATGCTCACCTATCTGGTACGCCAGTTCCTTTGGGGTGGACAGACTGG TTATGGCTGTGGCAGTATTCATGGAAGTTTTCCTCTTCTACTACCAGTCCCAACCGGCTCCGCTGGA CCAGCACATCCACTCAGTCTGCTGTATGCTGTGTTGGAGGGTGTGTTAGTATCTCCCTIAGAGGTGATC TTCCGGGACCACATTTGCTGGAACCTTTCCGAACCACTCICATCATTTTTCAGGGAACCTGGTTCTGGC AGATTTGGGTTTGTGCTTTCCACCTTTTGGAAACCCCGAATGGGACCAGAGGATGATGCCAACCTCAT GTTCATACCAATGTGCTTCTGCTGGCACTACCTGGCTGCCCTCAGCATTTGCGCGTCAACTATCTCTCT GTTTACTGCCITTTGACTCGGATGAAAGACACGGAAGGGGAGAAATCATTGGAATTCAGAAGCTGAAT CAGATGACACTTACCAGACCGCCCTTTGAGTGGCTCAGATGAGGAATGAGCCGAGATGCGGAGGGCCCA GATGTCACACTGCACAGCTGGAATGAATGGAGTTTATCCCTCCACCTGAATGCTGTGTTGGCTGTATC TTAAGGGTCTATATTTGCACCTCTCATTCAACACAGGGCTGGAGGTTCTACACAGGAAATCAGGCC TACAGCATCCTGTGTATCTGCAGTTGGGATTTTAAACATACTATAAAGTCTGTGTTGGTATAGTACCC TTCATAAGGAAAAATGAAGTAATGCCATAAGTAGCAGGCCCTTTGTGCCCTCAGTGTCAAGGAAATCAAG AGATGCTAAAAGCTTTTCAATGGAAGTGGCTCATGGATGAATCCGGGGTATGAGCCAGGAGAACGTGC TGCTTTTGGTAACTTATCCCTTTTCTCTTAAAGAAAGCAGGTACTTCTTATTAGAAATATGTAGAATG TGTAAGCAAACGACAGTGCCTTTAGAATTACAATTCATTAATTACATATTTTGAAGTAAAAATAATCA CAAGCTTTGGTATTTTAAATATTTGTTAAACATATCATAACTAATCATACCAGGGTACGCAATACCAC TGTTTATAAGTGACAAAATTAGCCAAAGGTGATTTTTTTTAAATCAGGAAGCTGGTACTGGCTCTAC TGAGAGTTGGAGCCCTGATGTTCTGATTTTCAAAGTACCCATAAAGAAAGATCTGACAGGAAAGCTGTA TAATGAGATAGAAAACCTGAGGTATGGAAGGCTTTTCAAGTTTAAATATGGCTGAAAGCAAAGGATAACGA ATTCAGAATTAGTAATGTAATACTTGCATACCTAATCTGCTTCTGGATCTGTTCTTTTTTAAAAAAA CTTCTTACCAGCCCTATAATCCTAGCACTTTGGAGGCGGAGGCAGGCAGATCACGGGGTACAGGAGAT CAAGACCATCCTGGCTAACATGGTAAACCCCGTCTCTACTGAAAATACAAAAAATAGCCGGGTGTGGT GCGGGCGCCTGTAGTTCCAGTACTCGGGAGGCTGAGGCAAGAGAATGGCATGAACCCGGTAGGGGAGC TTGAGTGAGCCAGATCATGCCACTGTACTCCAGCCTAGGTGACAGAGCAAGACTCTGTCTCAAAAAACA AGCAAAACAGACTTCTTCAACAAAATTTTATTAATATCCACTTTGCAACAGCACTGAAATGGCTGTAAAG GACTCCTGAGATATGTGCCAGCAAGGAGTTTACAGTCAAACAGGAGAGACATGCTGTAGTTACATCCA GTGTGATGGGTGCTGAGAGGCAAGTACAACCACGATG</p>	137

(continúa)

NÚMERO DE ACCESO DE GENBANK	SECUENCIA	SEQ ID NO:
BQ056428	<p>TCCCGCCGCGCCACTTCGGCTGCCTCCGTCCCGCCCGCGCCATGCCGTGTGGCCGGTCCGGAGCTG CCSCGCGGCCCCCTTGCCCCCGCCGACAGGAGCGGGACGCCGCGCTCCGCCGACGGGGAGCTGC AGTACCTGGGGCAGATCCAACACATCCCTCCGCTGCGGCGTCAGGAAGGACGCCCGCCCGGGCACCCTGAC CCTGCGGGIATTCGGCATGCAAGGCGCGCTACAGCCTGAGAGATGAATTCCTCTGCTGACAACCAACCGT GTSTTCTGGAACGGTCTTCGGAGGAGCTGCTGTGGCTTAICAGGGATCCACAAACCGCIATAGACCTGT CTTCGCCGCGCAGCAAAATCTCGGGATGCCACTGGATCCCGACACTCTCTGGACACCCCTGGGATCTTCCA CCAGAGAAGAAGCGGACTTGGGCCAGITTTGGCTCTCAGCGGAGGCTCTCTGTGGCAGAAATACATA TTCCAATCAGATCACTTCCCGGACACGGACCNIGACCAGCCTGCCAAAAAGTGGATTTCCCCCACC AGAACCCANCCCTGACGCACAGAAACCAACCCATTCGTGTGTGGCGCTTGGAAACCCCAACAGAAATC TCTCCCCCTGGCCGCGCGCCTGCCGCTGCCAATGCCCTATGGCGGCCCTTGGCCCGCACCTTCCAA TTGTGTCGCCCTGCGCAACAGCGGAGAAAACTGGCCCGCCGCTCTCCCCCGCTCCGCTACCCCACT</p> <p>TAAAGCGCTCCGTGGCATGACGCACGCGTTTGGTGTCCGCGCCGCTCTCATGTCCGCGCGGTGGACC CCGTTTTCTCTCGCGGCACATCCCCCTATTCCCTTGGCCTTTGGGGGACCCCTCTAGACCCGCGCT TCTCTTCTCGTCCGGTGGGGGACATTTGGTTTGCCTGCCGCGCGGGGGCGNTAAAAATAAAACAGCCTG TTAGCCCGCCAGTACCCCGCCCGCGCGGGCGCGCTTNGGTTTGCATTTATACCCCAACCCATAAAG CCGCGCCCTTTAGCNCCNIAACTTTTGTGGTGTGGCTCCCCCTTTTCCCGGGGAGCAGCAACGGAC ATCTGTACACTAATGCTGGCCCGGACCTTTCCAAAAACCCCGCCGCTGTCCCGTATAAATTTGGTGC CAANCTGACGNGTTCTCCCCCGCCCTCGCCCGTTGGCCCGCCGTTTAAAGCCCCCGGTGGTGGCG CGCCAACAGAGTCCACTAATGTTAANTCCACCAACACCCCGCACCTTTCTCCCGCGCGCATCTCCCC ACSTACCCCTTTTGTCCGAGATGGCCACTCCCCCCCCCTGTTGTTTAAAAACACGAGAATGGTGT GCCAACGCTGGTCTTTTCCCCCGGACCGCGACCCGCGAGGGGGAATACGTACCCATAAGCCCCCGGCC CNGCTTTTCCCCCGCTCCCGGCAATCAAGATCCGCGTCCATAGACGTTATTTTCCCCGCGATAC ACSAAAAACAGGGCCGCCATTTATACTAAATTCCTGTCGCCCGCGCGGATATGTTTCCAAAAATA CCACCCCGCCCGCATTTCTTTGCCCGCAACTCTCGCGCACCGGTGTTACACAGCTCGCGCCGC</p>	138
BC032677	<p>GGACGCGTGGGTGACCCACGCGTCCGGACCCACGCGTCCGGTCTGTCTCTCCAGTTCCTGTCTCTG CCAACGCGCCCGGATGGCTTCCAAAAACCGCCACCCAGCCCGCACTAGCGTCGCGCCCGCCCGTAAAG AGCTGAGCCGAGCGGGGGCGCCCGGGGTCCGGTGGGCAAAAGGCTACAGCAGGAGCTGATGACCCCTC ATGGTGGAGTATTAAGTGCCAGAAACCCAGCCTTCCATCCAATTTTCAGTAGCCTCCTTTTCCGTC GCTTTTTGCTAGACATAGGGTAATGTAATTTGCTCCCTCCTGGGAAAGAAGTTCATACACCCACCTA CACCATTTCTTCCAGCAGTCCCTCTCCCAATTCATCCCCCACAGGATTTATCTCGAACACTTCCCT GAAGTCATACAAGACCCCTCCCTATCCAGTGTGTCCTACTTCTTAGCCCCAACCAAGTTTACCCACCC CAACTCCCGCCCTTCTTGGTATTTAGCTATGAAATTTGGTTGCTTTATTTGGATCAGATGATGAG ATTAAGGGGAGGCTGGGCGCGGTAGCTCACACCTTATAATCCCAAAGTGTGGGATACAGGCGTGAGCC ACCGCGCCCGGCGAGCACTAATAATCTAATTTGAACAAAGCAGGATGCCAATTTACAATCTTAGAC CAAAGAGTCACTGATGCTCCACAGATAAGAGGAAAGCATCAGGCTAGGCATAGTGGCTCACACCTGTA ATCTCAGCCTTTGGGAGGCTGAGGCGAGCATCACATGAGCCAGGAGTTTGGACTGGCCCTGGGCAA CATGTTGAAACCCCTGCTCTAAAAATAAACTAAACTAAAAAACTTTTAAAAAGGCAGTGGGAGCAT CAGAACCAGCTCAACAGTTTGTCTACTGTCCGCTCCAGAGAAACTCAAGATCTAGCAAGCCCTTGTG TGGGCTTGGGTGGGACATGAGGCTGCTGCTGGAGCTTACTCTGCAACTGTTTCTCCAAATGCCAGGTA TATGAAGACCTGAGGTAAAGCTCTCGTAGAGTTCCCGAGTGGCTACCCCTTACAATGCGCCACAGTGA AGTTCTCACGCCCTGCTATCACCCCAACGTTGGACACCCAGGTAACATATGCTTGGACATCTCTGAAGGA AAAGTGGTCTGCCCTGTATGATGTAGGACCACTTCTGCTCTCATCCAGACCTTCTAGGAGAACCCAAC ATTGATAGTCCCTTGAACACACATGCTGCCGAGCTCTGAAAAACCCACAGCTTTTAAAGAAGTACCTGC AAGAAACCTACTCAAAGCAGGTACCCAGCCAGGAGCCCTGACCCAGGCTGCCAGCCCTGTCTTGTGTCG TCTTTTAAATTTTCTCTAGATGGTCTGCTCTTTTGTGATTTCTGTATAGGACTCTTTATCTTGTAGCTG TGGTATTTTGTTTTGTTTTGTCTTTTAAATTAAGCCTCGGTTGAGCCCTTGTATATTAATAAATGCA TTTTGTCCCTTTTAAAAAAAATAAAAAAAAATAAAAAAAAATAAAAAAAAATAAAAAAAAATAAAAAA A</p>	139

Al menos 40, al menos 41, al menos 42, al menos 43, al menos 44, al menos 46 o todos los 46 genes de la Tabla 1 pueden usarse en los métodos de la descripción. Preferentemente, la expresión de cada uno de los 46 genes se determina en una muestra biológica. Los perfiles prototípicos de expresión de genes (*es decir* centroide) de los cuatro subtipos intrínsecos se predefinieron a partir de un conjunto de entrenamiento de muestras de tumores de mama FFPE con el uso de análisis de agrupamiento jerárquico de datos de expresión de genes. Un mapa de calor de los perfiles prototípicos de expresión de genes (*es decir* centroides) de estos cuatro subtipos se muestra en la Figura 1, donde el nivel de expresión se ilustra mediante el mapa de calor. La tabla 3 muestra los valores reales.

Tabla 3.

Centroides de subtipo de tumor para la comparación con una muestra				
Gen diana	Tipo Basal	Enriquecido con Her2	Luminal A	Luminal B
ACTR3B	-0,2052	-0,7965	-0,2790	-0,4380
ANLN	1,0227	0,5006	-0,7289	0,1149
BAG1	-0,4676	-0,3132	0,4716	0,5879
BCL2	-0,7365	-0,7237	0,7234	0,6363

ES 2 763 931 T3

Centroides de subtipo de tumor para la comparación con una muestra					
	Gen diana	Tipo Basal	Enriquecido con Her2	Luminal A	Luminal B
5	BLVRA	-0,8761	0,2270	0,1628	0,7138
	CCNE1	1,3100	0,2201	-0,6231	-0,2729
	CDC20	1,0995	0,1445	-1,0518	-0,1173
	CDC6	0,5817	0,6601	-0,7032	0,3134
10	CDCA1	0,9367	0,1623	-0,4509	0,2692
	CDH3	0,7639	0,0144	-0,0502	-1,0229
	CENPF	1,0222	0,2944	-0,5657	0,2437
	CEP55	1,0442	0,4881	-0,6365	0,2921
15	CXXC5	-0,9732	0,1866	0,5687	0,9463
	EGFR	0,3352	-0,1326	-0,0011	-0,9755
	ERBB2	-0,7045	1,4182	0,2420	0,1978
	ESR1	-1,1847	-0,4926	0,7177	1,0101
20	EXO1	1,0546	0,4317	-0,7259	0,2559
	FGFR4	-0,2073	1,4562	0,1707	-0,2223
	FOXA1	-1,3590	0,5726	0,7131	0,7963
25	FOXC1	1,0666	-0,7362	-0,4078	-0,9877
	GPR160	-1,0540	0,5524	0,6032	0,7305
	KIF2C	0,9242	0,1104	-1,1001	-0,2771
	KNTC2	1,1373	0,2266	-0,7593	0,1656
30	KRT14	0,4759	-0,5269	0,8187	-0,8879
	KRT17	0,6863	-0,3777	0,6149	-1,1415
	KRT5	0,7136	-0,4146	0,5832	-0,9462
	MAPT	-1,1343	-0,2711	1,0957	0,8372
35	MDM2	-0,7498	-0,4855	-0,1788	0,2397
	MELK	1,0209	0,2678	-0,8016	0,1012
	MIA	1,2408	-0,5475	0,3289	-0,6320
40	MKI67	1,0446	0,4630	-0,6717	0,3161
	MLPH	-1,4150	0,4842	0,8829	0,8194
	MMP11	-0,1295	0,5220	0,3402	0,5653
	MYC	0,5639	-0,9904	-0,3015	-0,2791
45	NAT1	-0,9711	-0,2708	1,2256	0,9576
	ORC6L	1,0086	0,5152	-1,0385	-0,0336
	PGR	-0,9216	-0,5755	1,2061	0,9278
	PHGDH	0,9192	0,0322	-0,5194	-0,5371
50	PTTG1	0,9541	0,2079	-1,1207	0,1052
	RRM2	0,7895	0,6336	-0,8099	0,3228
	SFRP1	0,7694	-0,8271	0,2617	-1,0846
55	SLC39A6	-0,9992	-0,4573	0,6607	0,9222
	TMEM45B	-1,0721	0,7926	0,3190	0,2016
	TYMS	0,9823	-0,0960	-0,8593	0,1827
	UBE2C	0,8294	0,3358	-1,0141	0,0608
60	UBE2T	0,6258	0,0617	-0,8652	-0,0487

Después de realizar la prueba de subtipaje intrínseco de cáncer de mama con una muestra de prueba de tumor de cáncer de mama y la muestra de referencia proporcionada como parte del kit de prueba, un algoritmo computacional basado en una correlación de Pearson compara el perfil de expresión de genes normalizado y escalado del conjunto de genes intrínsecos NANO46 de la muestra de prueba con las firmas de expresión prototípica de los cuatro subtipos intrínsecos de cáncer de mama. El análisis del subtipo intrínseco se determina determinando la expresión de un conjunto de genes

NANO50 (que determina la expresión del conjunto de genes NANO46 e incluye además determinar la expresión de MYBL2, BIRC5, GRB7 y CCNB1) y el riesgo de recurrencia ("ROR") se determina con el uso del conjunto de genes NANO46). Específicamente, el subtipo intrínseco se identifica comparando la expresión del conjunto de genes NANO50 en la muestra biológica con los perfiles de expresión esperados para los cuatro subtipos intrínsecos. El subtipo con el perfil de expresión más similar se asigna a la muestra biológica. La puntuación ROR es un valor entero en una escala de 0-100 que está relacionado con la probabilidad de un paciente individual de recurrencia distante dentro de los 10 años para el uso previsto en la población definida. La puntuación ROR se calcula comparando los perfiles de expresión de los genes NANO46 en la muestra biológica con los perfiles esperados para los cuatro subtipos intrínsecos, como se describió anteriormente, para calcular cuatro valores de correlación diferentes. Estos valores de correlación se combinan después con una puntuación de proliferación (y opcionalmente una o más variables clínico patológicas, tal como el tamaño del tumor) para calcular la puntuación ROR. Preferentemente, la puntuación ROR se calcula comparando solo los perfiles de expresión de los genes NANO46.

La Figura 6 proporciona un esquema de las transformaciones específicas del algoritmo. A la muestra de tumor se le asigna el subtipo con la mayor correlación positiva con la muestra. Las curvas de supervivencia de Kaplan Meier generadas a partir de un conjunto de entrenamiento de pacientes con cáncer de mama no tratadas demuestran que los subtipos intrínsecos son un indicador pronóstico de supervivencia libre de recurrencia (RFS) en esta población de prueba, que incluye pacientes tanto con receptor de estrógeno positivo/negativo como pacientes con HER2 positivo/negativo, Figura 2.

Las pruebas independientes en una cohorte de pacientes con ganglios negativos, receptor de estrógeno positivo tratados con tamoxifeno muestran predominantemente pacientes del subtipo Luminal A y B con pacientes Luminal A que exhiben mejores resultados que los pacientes Luminal B, Figura 3. Se espera que el resultado de los pacientes Luminal A mejore aún más con el uso de muestras de ensayos clínicos que usan regímenes de tratamiento más modernos (es decir inhibidores de aromatasas) y tienen mejor observancia a la terapia lo que mejorará el resultado

El conjunto de entrenamiento de muestras de tumor de mama FFPE, que tenía características clínicas bien definidas y datos de resultados clínicos, se usó para establecer una puntuación continua de riesgo de recurrencia (ROR). La puntuación se calcula con el uso de coeficientes de un modelo de Cox que incluye la correlación con cada subtipo intrínseco, una puntuación de proliferación (expresión de genes media de un subconjunto de 18 de los 46 genes), y el tamaño del tumor, Tabla 4.

Tabla 4. Coeficientes para calcular ROR-PT (ecuación 1)	
Variables de la prueba	Coeficiente
Correlación de Pearson del tipo basal (A)	- 0,0067
Correlación de Pearson del enriquecido con Her2 (B)	0,4317
Correlación de Pearson del Luminal A (C)	- 0,3172
Correlación de Pearson del Luminal B (D)	0,4894
Puntuación de proliferación (E)	0,1981
Tamaño de tumor (F)	0,1133

Las variables de prueba en la Tabla 4 se multiplican por los coeficientes correspondientes y se suman para producir una puntuación de riesgo ("ROR-PT").

$$\text{Ecuación ROR-PT} = -0,0067*A + 0,4317*B + -0,3172*C + 0,4894*D + 0,1981*E + 0,1133*F$$

En estudios previos, la puntuación ROR proporcionó una estimación continua del riesgo de recurrencia para pacientes con ganglios negativos, ER-positivos, que se trataron con tamoxifeno durante 5 años (Nielsen y otros. Clin. Cancer Res., 16 (21):5222-5232 (2009)). Este resultado se verificó en pacientes con ganglios negativos, ER positivo, de la misma cohorte, Figura 4. La puntuación ROR además exhibió una mejora estadísticamente significativa sobre un modelo clínico basado en la determinación de RFS dentro de esta población de prueba, proporcionando evidencia adicional de la precisión mejorada de esta herramienta de toma de decisiones en comparación con las medidas clínico-patológicas tradicionales (Nielsen y otros. Clin. Cancer Res., 16(21):5222-5232 (2009)).

El conjunto de genes contiene muchos genes que son marcadores conocidos de la proliferación. Los métodos descritos en la presente descripción proporcionan la determinación de subconjuntos de genes que proporcionan una firma de proliferación. Los métodos pueden incluir determinar la expresión de al menos uno de, una combinación de, o cada uno de, un subconjunto de 18 genes de los genes intrínsecos NANO46 seleccionados de ANLN, CCNE1, CDC20, CDC6, CDCA1, CENPF, CEP55, EXO1, KIF2C, KNTC2, MELK, MKI67, ORC6L, PTTG1, RRM2, TYMS, UBE2C y/o UBE2T. Preferentemente, se determina que la expresión de cada uno del subconjunto de 18 genes del conjunto de genes NANO46 proporciona una puntuación de proliferación. Puede determinarse la expresión de uno o más de estos genes y puede generarse un índice de la firma de proliferación promediando las estimaciones de expresión normalizadas de uno o más

de estos genes en una muestra. A la muestra le puede asignarse una firma de proliferación alta, una firma de proliferación moderada/intermedia, una firma de proliferación baja o una firma de proliferación ultra baja. Los métodos para determinar una firma de proliferación a partir de una muestra biológica son como se describen en Nielsen y otros. Clin. Cancer Res., 16 (21):5222-5232 (2009)) y material complementario en línea.

5

Descripción de la biología del subtipo intrínseco

Subtipos luminales: Los subtipos más comunes de cáncer de mama son los subtipos luminales, Luminal A y Luminal B. Estudios previos sugieren que el luminal A comprende aproximadamente del 30% al 40% y el luminal B aproximadamente el 20% de todos los cánceres de mama, pero representan más del 90% de los cánceres de mama con receptores positivos de las hormonas (Nielsen y otros. Clin. Cancer Res., 16(21):5222-5232 (2009))). El patrón de expresión de genes de estos subtipos se asemeja al componente epitelial luminal de la mama. Estos tumores se caracterizan por la alta expresión del receptor de estrógenos (ER), el receptor de progesterona (PR) y los genes asociados con la activación de ER, tal como LIV1, GATA3, y ciclina D1, así como la expresión de citoqueratinas luminales 8 y 18 (Lisa Carey y Charles Perou (2009). Gene Arrays, Prognosis, and Therapeutic Interventions. Jay R. Harris y otros. (4ta ed.), Diseases of the breast (págs. 458-472). Filadelfia, PA: Lippincott Williams & Wilkins).

10

15

20

25

Luminal A: Los cánceres de mama Luminal A (LumA) exhiben una baja expresión de genes asociados con la activación del ciclo celular y el agrupamiento ERBB2, lo que resulta en un mejor pronóstico que el luminal B. El subgrupo Luminal A tiene el pronóstico más favorable de todos los subtipos y está enriquecido para tumores con respuesta a la terapia endocrina.

Luminal B: Los cánceres de mama Luminal B (LumB) además expresan ER y genes asociados a ER. Los genes asociados con la activación del ciclo celular se expresan altamente y este tipo de tumor puede ser HER2(+) (~20%) o HER2(-). El pronóstico es desfavorable (a pesar de la expresión de ER) y la respuesta a la terapia endocrina generalmente disminuye con relación a LumA.

30

35

Enriquecido con HER2: El subtipo enriquecido con HER2 es generalmente negativo para ER y es positivo para HER2, en la mayoría de los casos con alta expresión del agrupamiento ERBB2, incluidos ERBB2 y GRB7. Los genes asociados con la activación del ciclo celular se expresan altamente y estos tumores tienen un resultado malo.

Tipo Basal: El subtipo tipo basal es generalmente negativo para ER, casi siempre es clínicamente negativo para HER2 y expresa un conjunto de biomarcadores "basales" que incluyen las citoqueratinas epiteliales basales (CK) y el receptor del factor de crecimiento epidérmico (EGFR). Los genes asociados con la activación del ciclo celular se expresan altamente.

40

Variables Clínicas

El modelo de clasificación NANO46 descrito en la presente descripción puede combinarse además con información sobre variables clínicas para generar un predictor continuo de riesgo de recurrencia (ROR). Como se describe en la presente descripción, se conocen varios factores clínicos y pronósticos de cáncer de mama en la técnica y se usan para predecir el resultado del tratamiento y la probabilidad de recurrencia de la enfermedad. Tales factores incluyen, por ejemplo, compromiso de los ganglios linfáticos, tamaño del tumor, grado histológico, estado del receptor de la hormona estrógeno y progesterona, niveles de HER-2 y ploidía de tumor. En una modalidad, se proporciona una puntuación de riesgo de recurrencia (ROR) para un sujeto diagnosticado o sospechoso de tener cáncer de mama. Esta puntuación usa el modelo de clasificación NANO46 en combinación con factores clínicos del estado de los ganglios linfáticos (N) y el tamaño del tumor (T). La evaluación de las variables clínicas se basa en el sistema estandarizado del American Joint Committee on Cancer (AJCC) para la estadificación del cáncer de mama. En este sistema, el tamaño del tumor primario se clasifica en una escala de 0-4 (T0: sin evidencia de tumor primario; T1: <2 cm; T2: > 2 cm - <5 cm; T3: > 5 cm; T4: tumor de cualquier tamaño con propagación directa a la pared torácica o la piel). El estado de los ganglios linfáticos se clasifica como N0-N3 (N0: los ganglios linfáticos regionales están libres de metástasis; N1: metástasis a los ganglios linfáticos axilares móviles del mismo lado; N2: metástasis a los ganglios linfáticos del mismo lado fijados uno al otro o a otras estructuras; N3: metástasis a los ganglios linfáticos del mismo lado debajo del esternón). Los métodos para identificar pacientes con cáncer de mama y estadificar la enfermedad son bien conocidos y pueden incluir examen manual, biopsia, revisión del historial del paciente y/o familiar, y técnicas de imagen, como mamografía, resonancia magnética (MRI), y tomografía por emisión de positrones (PET).

45

50

55

Fuente de muestra

En una modalidad de la presente descripción, el subtipo de cáncer de mama se evalúa mediante la evaluación de patrones de expresión, o perfiles, de los genes intrínsecos enumerados en la Tabla 1 en una o más muestras del sujeto. Para propósitos de discusión, el término sujeto, o muestra de sujeto, se refiere a un individuo independientemente de su estado de salud y/o enfermedad. Un sujeto puede ser un sujeto, un participante del estudio, un sujeto de control, un sujeto de tamizaje, o cualquier otra clase de individuo de quien se obtiene y evalúa una muestra en el contexto de la descripción. En consecuencia, un sujeto puede diagnosticarse con cáncer de mama, puede presentar uno o más síntomas de cáncer de mama, o un factor predisponente, tal como un factor familiar (genético) o de historial médico (médico), para cáncer de mama, puede estar bajo tratamiento o terapia para el cáncer de mama, o similares. Alternativamente, un sujeto puede estar sano con respecto a cualquiera de los factores o criterios mencionados anteriormente. Se apreciará que el término

60

65

"saludable", como se usa en la presente descripción, es relativo al estado del cáncer de mama, ya que el término "saludable" no puede definirse que corresponda a cualquier evaluación o estado absoluto. Así, un individuo definido como sano con referencia a cualquier enfermedad o criterio de enfermedad específico, de hecho puede ser diagnosticado con otra o más enfermedades, o exhibir cualquier otro criterio de enfermedad, que incluye uno o más cánceres distintos al cáncer de mama. Sin embargo, los controles sanos están preferentemente libres de cáncer.

En modalidades particulares, los métodos para predecir los subtipos intrínsecos de cáncer de mama incluyen la recolección de una muestra biológica que comprende una célula o tejido canceroso, tal como una muestra de tejido de mama o una muestra de tejido de tumor de mama primario. Por "muestra biológica" se entiende cualquier muestreo de células, tejidos, o fluidos corporales en los que pueda detectarse la expresión de un gen intrínseco. Los ejemplos de tales muestras biológicas incluyen, pero no se limitan a, biopsias y frotis. Los fluidos corporales útiles en la presente descripción incluyen sangre, linfa, orina, saliva, aspirados de pezón, fluidos ginecológicos, o cualquier otra secreción corporal o derivado de estos. La sangre puede incluir sangre completa, plasma, suero, o cualquier derivado de la sangre. En algunas modalidades, la muestra biológica incluye células de mama, particularmente tejido de mama de una biopsia, tal como una muestra de tejido de tumor de mama. Las muestras biológicas pueden obtenerse de un sujeto mediante una variedad de técnicas que incluyen, por ejemplo, raspar o frotar un área, usar una aguja para aspirar células o fluidos corporales, o extraer una muestra de tejido (*es decir*, biopsia). Los métodos para recolectar diversas muestras biológicas son bien conocidos en la técnica. En algunas modalidades, se obtiene una muestra de tejido mamario mediante, por ejemplo, biopsia por aspiración con aguja fina, biopsia con aguja gruesa, o biopsia por escisión. Pueden aplicarse soluciones de fijación y tinción a las células o tejidos para preservar la muestra y facilitar el examen. Las muestras biológicas, particularmente las muestras de tejido mamario, pueden transferirse a un portaobjetos de vidrio para verlas con aumento. En una modalidad, la muestra biológica es una muestra de tejido mamario embebido en parafina y fijada en formalina, particularmente una muestra de tumor de mama primario. En diversas modalidades, la muestra de tejido se obtiene de una muestra de núcleo del tejido guiada por un patólogo.

Perfiles de expresión

En diversas modalidades, la presente descripción proporciona métodos para clasificar, pronosticar, o controlar el cáncer de mama en sujetos. En esta modalidad, los datos obtenidos del análisis de la expresión de genes intrínsecos se evalúan usando uno o más algoritmos de reconocimiento de patrones. Tales métodos de análisis pueden usarse para formar un modelo predictivo, que puede usarse para clasificar los datos de prueba. Por ejemplo, un método de clasificación conveniente y particularmente efectivo emplea el modelado de análisis estadístico multivariado, primero para formar un modelo (un "modelo matemático predictivo") con el uso de datos ("datos de modelado") de muestras de subtipo conocido (por ejemplo, de sujetos conocidos que tienen un subtipo intrínseco de cáncer de mama en particular: LumA, LumB, tipo basal, enriquecido con HER2 o tipo normal), y segundo para clasificar una muestra desconocida (*por ejemplo*, "muestra de prueba") de acuerdo con el subtipo. Los métodos de reconocimiento de patrones se han usado ampliamente para caracterizar muchos tipos diferentes de problemas que abarcan, por ejemplo, la lingüística, las huellas digitales, la química y la psicología. En el contexto de los métodos descritos en la presente descripción, el reconocimiento de patrones es el uso de estadísticas multivariadas, tanto paramétricas como no paramétricas, para analizar datos y, por lo tanto, clasificar muestras y predecir el valor de alguna variable dependiente en función de un intervalo de mediciones observadas. Existen dos enfoques principales. Un conjunto de métodos se denomina "sin supervisión" y estos simplemente reducen la complejidad de los datos de manera racional y además producen gráficos de visualización que pueden interpretarse por el ojo humano. Sin embargo, este tipo de enfoque puede no ser adecuado para desarrollar un ensayo clínico que pueda usarse para clasificar muestras derivadas de sujetos independientes de la población de muestra inicial usada para estudiar el algoritmo de predicción.

El otro enfoque se denomina "supervisado" mediante el cual se usa un conjunto de muestras de entrenamiento con clase o resultado conocidos para producir un modelo matemático que después se evalúa con conjuntos de datos de validación independientes. Aquí, se usa un "conjunto de entrenamiento" de datos de la expresión de genes intrínsecos para construir un modelo estadístico que predice correctamente el "subtipo" de cada muestra. Este conjunto de entrenamiento se prueba después con datos independientes (denominado conjunto de prueba o validación) para determinar la robustez del modelo basado en computadora. Estos modelos a veces se denominan "sistemas expertos", pero pueden basarse en una gama de diferentes procedimientos matemáticos. Los métodos supervisados pueden usar un conjunto de datos con una dimensionalidad reducida (por ejemplo, los primeros componentes principales), pero típicamente usan datos no reducidos, con toda la dimensionalidad. En todos los casos, los métodos permiten la descripción cuantitativa de los límites multivariados que caracterizan y separan cada subtipo en términos de su perfil de expresión de genes intrínsecos. Además es posible obtener límites de confianza en cualquier predicción, por ejemplo, un nivel de probabilidad para la bondad del ajuste. La robustez de los modelos predictivos además puede verificarse mediante validación cruzada, al excluir muestras seleccionadas del análisis.

El modelo de clasificación NANO46 descrito en la presente descripción se basa en el perfil de expresión de genes para una pluralidad de muestras de sujetos que usan los genes intrínsecos enumerados en la Tabla 1. La pluralidad de muestras incluye un número suficiente de muestras derivadas de sujetos que pertenecen a cada clase de subtipo. Por "muestras suficientes" o "número representativo" en este contexto se entiende una cantidad de muestras derivadas de cada subtipo que sea suficiente para construir un modelo de clasificación que pueda distinguir de manera confiable cada subtipo del resto del grupo. Se desarrolla un algoritmo de predicción supervisado basado en los perfiles de muestras

prototipo seleccionadas objetivamente para "entrenar" el algoritmo. Las muestras se seleccionan y subtipan con el uso de un conjunto de genes intrínsecos expandidos de acuerdo con los métodos descritos en la publicación de patente internacional WO 2007/061876 y Publicación de Patente de los Estados Unidos Núm. 2009/0299640. Alternativamente, las muestras pueden subtiparse de acuerdo con cualquier ensayo conocido para clasificar los subtipos de cáncer de mama. Después de estratificar las muestras de entrenamiento de acuerdo con el subtipo, se usa un algoritmo de predicción basado en centroide para construir los centroides basados en el perfil de expresión del conjunto de genes intrínsecos descrito en la Tabla 1.

En una modalidad, el algoritmo de predicción es la metodología centroide más cercana relacionada con la descrita en Narashiman y Chu (2002) PNAS 99:6567-6572. En la presente descripción, el método calcula un centroide estandarizado para cada subtipo. Este centroide es la expresión de genes promedio para cada gen en cada subtipo (o "clase") dividido por la desviación estándar dentro de la clase para ese gen. La clasificación de centroides más cercana toma el perfil de expresión de genes de una nueva muestra y lo compara con cada uno de estos centroides de clase. La predicción de subtipo se realiza calculando la correlación de intervalo de Spearman de cada caso de prueba con los cinco centroides, y asignando una muestra a un subtipo basado en el centroide más cercano.

Detección de la expresión de genes intrínsecos.

Todos los métodos disponibles en la técnica para detectar la expresión de los genes intrínsecos enumerados en la Tabla 1 se incluyen en la presente descripción. Por "detectar la expresión" se entiende determinar la cantidad o presencia de un transcrito de ARN o su producto de expresión de un gen intrínseco. Los métodos para detectar la expresión de los genes intrínsecos de la descripción, es decir, el perfil de expresión de genes, incluyen métodos basados en el análisis de hibridación de polinucleótidos, métodos basados en la secuenciación de polinucleótidos, métodos de inmunohistoquímica, y métodos basados en proteómica. Los métodos generalmente detectan productos de expresión (*por ejemplo*, ARNm) de los genes intrínsecos enumerados en la Tabla 1. En modalidades preferidas, se usan los métodos basados en PCR, tales como la transcripción inversa PCR (RT-PCR) (Weis y otros, TIG 8:263- 64, 1992), y métodos basados en matrices tales como micromatrices (Schna y otros, Science 270:467- 70, 1995). Por "micromatriz" se entiende una disposición ordenada de elementos de matriz hibridables, tales como, por ejemplo, sondas de polinucleótidos, sobre un sustrato. El término "sonda" se refiere a cualquier molécula que sea capaz de unirse selectivamente a una biomolécula diana específicamente prevista, por ejemplo, un transcrito de nucleótido o una proteína codificada por o correspondiente a un gen intrínseco. Las sondas pueden sintetizarse por un experto en la técnica, o derivarse de preparaciones biológicas apropiadas. Las sondas pueden estar específicamente diseñadas para ser marcadas. Los ejemplos de moléculas que pueden utilizarse como sondas incluyen, pero no se limitan a, ARN, ADN, proteínas, anticuerpos, y moléculas orgánicas.

Muchos métodos de detección de la expresión usan ARN aislado. El material de partida es típicamente ARN total aislado de una muestra biológica, tal como un tumor o línea celular de tumor, y el tejido normal o línea celular correspondiente, respectivamente. Si la fuente de ARN es un tumor primario, el ARN (*por ejemplo*, ARNm) puede extraerse, por ejemplo, de muestras de tejido (*por ejemplo*, muestras de núcleo de tejido guiadas por patólogo) congeladas o archivadas embebidas en parafina y fijadas (*por ejemplo*, fijadas en formalina).

Los métodos generales para la extracción de ARN son bien conocidos en la técnica y se describen en libros de texto estándar de biología molecular, que incluyen Ausubel y otros, Ed., Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York 1987-1999. Los métodos para la extracción de ARN de tejidos embebidos en parafina se describen, por ejemplo, en Rupp y Locker, Lab Invest. 56:A67, (1987); y De Andres y otros. Biotechniques 18:42-44, (1995). En particular, el aislamiento de ARN puede realizarse con el uso de un kit de purificación, un conjunto de tampones y una proteasa de fabricantes comerciales, tales como Qiagen (Valencia, CA), de acuerdo con las instrucciones del fabricante. Por ejemplo, el ARN total de las células en cultivo puede aislarse usando minicolumnas Qiagen RNeasy. Otros kits de aislamiento de ARN disponibles en el comercio incluyen el kit de purificación de ADN y ARN completo MASTERPURE™ (Epicenter, Madison, Wisconsin) y el kit de aislamiento de ARN de bloque de parafina (Ambion, Austin, TX). El ARN total de las muestras de tejido puede aislarse, por ejemplo, con el uso de ARN Stat-60 (Tel-Test, Friendswood, TX). El ARN total de FFPE puede aislarse, por ejemplo, con el uso de Microkit de ARN de FFPE de alta pureza, Núm. de cat. 04823125001 (Roche Applied Science, Indianápolis, IN). El ARN preparado de un tumor puede aislarse, por ejemplo, mediante centrifugación en gradiente de densidad de cloruro de cesio. Además, pueden procesarse fácilmente grandes cantidades de muestras de tejido con el uso de técnicas bien conocidas por los expertos en la técnica, tales como, por ejemplo, el proceso de aislamiento de ARN en una sola etapa de Chomczynski (Patente de los Estados Unidos Núm. 4.843.155).

El ARN aislado puede usarse en ensayos de hibridación o amplificación que incluyen, pero no se limitan a, análisis de PCR y matrices de sondas. Un método para la detección de niveles de ARN involucra poner en contacto el ARN aislado con una molécula de ácido nucleico (sonda) que puede hibridarse con el ARNm codificado por el gen que se está detectando. La sonda de ácido nucleico puede ser, por ejemplo, un ADNc de longitud completa, o una porción de este, tal como un oligonucleótido de al menos 7, 15, 30, 60, 100, 250, o 500 nucleótidos de longitud y suficiente para hibridar específicamente bajo condiciones estrictas para un gen intrínseco de la presente descripción, o cualquier ADN o ARN derivado. La hibridación de un ARNm con la sonda indica que se está expresando el gen intrínseco en cuestión.

En una modalidad, el ARNm se inmoviliza sobre una superficie sólida y se pone en contacto con una sonda, por ejemplo haciendo pasar el ARNm aislado en un gel de agarosa y transfiriendo el ARNm desde el gel a una membrana, tal como nitrocelulosa. En una modalidad alternativa, las sondas se inmovilizan en una superficie sólida y el ARNm se pone en contacto con las sondas, por ejemplo, en una matriz de chips de genes Agilent. Un experto en la técnica puede adaptar fácilmente los métodos de detección de ARNm conocidos para usar en la detección del nivel de expresión de los genes intrínsecos de la presente descripción.

Un método alternativo para determinar el nivel de producto de expresión de genes intrínsecos en una muestra implica el proceso de amplificación de ácido nucleico, por ejemplo, por RT-PCR (Patente de los Estados Unidos Núm. 4.683.202), reacción en cadena de la ligasa (Barany, PNAS USA 88: 189-93, (1991)), replicación de secuencia autosostenida (Guatelli y otros, Proc. Natl. Acad. Sci USA 87: 1874-78, (1990)), sistema de amplificación transcripcional (Kwoh y otros., Proc. Natl. Acad. Sci USA 86: 1173-77, (1989)), Replicasa Q-Beta (Lizardi y otros, Bio/Technology 6:1197, (1988)), replicación de círculo rodante (Patente de los Estados Unidos Núm. 5,854,033), o cualquier otro método de amplificación de ácido nucleico, seguido de la detección de las moléculas amplificadas con el uso de técnicas bien conocidas por los expertos en la técnica. Son especialmente útiles estos esquemas de detección para la detección de moléculas de ácido nucleico si tales moléculas están presentes en números muy bajos.

En aspectos particulares de la descripción, la expresión de genes intrínsecos se evalúa mediante RT-PCR cuantitativa. Numerosos protocolos diferentes de PCR o QPCR se conocen en la técnica y se ejemplifican en la presente descripción más abajo y pueden aplicarse directamente o adaptarse para usar con el empleo de las composiciones actualmente descritas para la detección y/o cuantificación de los genes intrínsecos enumerados en la Tabla 1. Generalmente, en la PCR, una secuencia de polinucleótido diana se amplifica por reacción con al menos un cebador de oligonucleótido o un par de cebadores de oligonucleótidos. El cebador(es) se hibrida(n) con una región complementaria del ácido nucleico diana y una ADN polimerasa extiende el(los) cebador(es) para amplificar la secuencia diana. En condiciones suficientes para proporcionar productos de amplificación de ácido nucleico basado en la polimerasa, un fragmento de ácido nucleico de un tamaño domina los productos de reacción (la secuencia de polinucleótido diana que es el producto de amplificación). El ciclo de amplificación se repite para aumentar la concentración de la secuencia única del polinucleótido diana. La reacción se puede realizar en cualquier termociclador usado comúnmente para PCR. Sin embargo, se prefieren los cicladores con capacidades de medición de fluorescencia en tiempo real, por ejemplo, SMARTCYCLER® (Cepheid, Sunnyvale, CA), ABI PRISM 7700® (Applied Biosystems, Foster City, CA), ROTOR-GENE™; (Corbett Research, Sydney, Australia), LIGHTCYCLER® (Roche Diagnostics Corp, Indianapolis, Ind), ICYCLER® (Biorad Laboratories, Hercules, CA) y MX4000® (Stratagene, La Jolla, CA).

En otra modalidad de la descripción, se usan micromatrices para el perfil de expresión. Las micromatrices son particularmente adecuadas para este propósito debido a la reproducibilidad entre diferentes experimentos. Las micromatrices de ADN proporcionan un método para la medición simultánea de los niveles de expresión de grandes cantidades de genes. Cada matriz consta de un patrón reproducible de sondas de captura unidas a un soporte sólido. El ARN o ADN marcado se hibrida con sondas complementarias en la matriz y después se detecta mediante escaneo láser. Las intensidades de hibridación para cada sonda en la matriz se determinan y se convierten en un valor cuantitativo que representa los niveles relativos de expresión de genes. Ver, por ejemplo, las patentes de los Estados Unidos Núms. 6.040.138, 5.800.992 y 6.020.135, 6.033.860, y 6.344.316. Las matrices de oligonucleótidos de alta densidad son particularmente útiles para determinar el perfil de expresión de genes para una gran cantidad de ARN en una muestra.

En una modalidad preferida, el sistema de análisis nCounter® se usa para detectar la expresión de genes intrínsecos. La base del sistema de análisis nCounter® es el código único asignado a cada diana de ácido nucleico a analizar (Publicación de solicitud de patente internacional Núm. WO 08/124847, Patente de los Estados Unidos Núm. 8.415.102 y Geiss y otros. Nature Biotechnology. 2008. 26(3): 317-325). El código se compone de una serie ordenada de puntos fluorescentes de colores que crean un código de barras único para cada diana a analizar. Se diseña un par de sondas para cada diana de ADN o ARN, una sonda de captura biotinilada y una sonda reportera que lleva el código de barras fluorescente. Este sistema además se denomina, en la presente descripción, el sistema de código de nanoreportero.

Se sintetizan reporteros específicos y sondas de captura para cada diana. Brevemente, las sondas de oligonucleótidos de ADN específicas a la secuencia se unen a moléculas reporteras específicas al código. Preferentemente, cada sonda reportera específica de secuencia comprende una secuencia específica diana capaz de hibridar a no más de un gen NANO46 de la Tabla 1 y opcionalmente comprende al menos dos, al menos tres, o al menos cuatro regiones de unión al marcador, comprendiendo dichas regiones de unión una o más monómeros del marcador que emiten luz. Las sondas de captura se realizan ligando un segundo oligonucleótido de ADN específico de secuencia para cada diana a un oligonucleótido universal que contiene biotina. Las sondas reporteras y de captura se agrupan en una única mezcla de hibridación, la "biblioteca de sondas". Preferentemente, la biblioteca de sondas comprende un par de sondas (una sonda de captura y una reportera) para cada uno de los genes NANO46 en la Tabla 1.

La abundancia relativa de cada diana se mide en una única reacción de hibridación multiplexada. El método comprende poner en contacto una muestra biológica con una biblioteca de sondas, la biblioteca comprende un par de sondas para los genes NANO46 en la Tabla 1, de manera que la presencia de la diana en la muestra crea un complejo de par de sonda - diana. El complejo después se purifica. Más específicamente, la muestra se combina con la biblioteca de sondas, y se produce la hibridación en solución. Después de la hibridación, los complejos hibridados tripartitos (pares de sondas y

diana) se purifican en un procedimiento de dos etapas con el uso de perlas magnéticas unidas a oligonucleótidos complementarios a las secuencias universales presentes en las sondas de captura y reportero. Este proceso de purificación doble permite que la reacción de hibridación se complete con un gran exceso de sondas específicas de la diana, ya que finalmente se eliminan y, así, no interfieren con la unión e imágenes de la muestra. Todas las etapas posteriores a la hibridación se manejan robóticamente en un robot de manejo de líquidos personalizado (Prep Station, NanoString Technologies).

Las reacciones purificadas se depositan por la estación de preparación en celdas de flujo individuales de un cartucho de muestra, se une a una superficie recubierta de estreptavidina a través de la sonda de captura, se somete a electroforesis para alargar las sondas indicadoras y se inmoviliza. Después del procesamiento, el cartucho de muestra se transfiere a un dispositivo de captura de datos e imágenes completamente automatizado (Analizador Digital, NanoString Technologies). El nivel de expresión de una diana se mide por la imagen de cada muestra y contando el número de veces que se detecta el código para esa diana. Los datos se generan en un formato de hoja de cálculo simple que enumera el número de conteos por diana, por muestra.

Este sistema puede usarse junto con nanoreporteros. Puede encontrarse información adicional sobre nanoreporteros en la Publicación internacional Núm. WO 07/076129 y WO 07/076132y Publicación de Patente de los Estados Unidos Núm. 2010/0015607 y 2010/0261026. Además, el término sondas de ácido nucleico y nanoreporteros puede incluir el diseño racional (*por ejemplo* secuencias sintéticas) descritas en la Publicación internacional Núm. WO 2010/019826 y Publicación de Patente de los Estados Unidos Núm. 2010/0047924.

Procesamiento de los Datos

Frecuentemente es útil procesar previamente los datos de expresión de genes, por ejemplo, abordando datos faltantes, traducción, escalado, normalización, ponderación, etc. Los métodos de proyección multivariados, tales como el análisis de componentes principales (PCA) y el análisis de mínimos cuadrados parciales (PLS), se denominan métodos sensibles al escalado. Al usar el conocimiento previo y la experiencia sobre el tipo de datos estudiados, la calidad de los datos antes del modelado multivariado puede mejorarse mediante el escalado y/o ponderación. El escalado y/o ponderación adecuados pueden revelar variaciones importantes e interesantes ocultas dentro de los datos y, por lo tanto, hacer que el modelado multivariado posterior sea más eficiente. El escalado y la ponderación pueden usarse para colocar los datos en la métrica correcta, basado en el conocimiento y la experiencia del sistema estudiado, y por lo tanto revelar patrones ya inherentemente presentes en los datos.

Si es posible, se deben evitar los datos faltantes, por ejemplo, interrupciones en los valores de las columnas. Sin embargo, si es necesario, tales datos faltantes pueden reemplazarse o "rellenarse" con, por ejemplo, el valor medio de una columna ("relleno medio"); un valor aleatorio ("relleno aleatorio"); o un valor basado en un análisis de componentes principales ("relleno de componentes principales").

La "traducción" de los ejes de coordenadas del descriptor puede ser útil. Los ejemplos de tal traducción incluyen la normalización y el centrado medio. La "normalización" puede usarse para eliminar la variación de muestra a muestra. Para los datos de micromatrices, el proceso de normalización tiene como objetivo eliminar errores sistemáticos equilibrando las intensidades de fluorescencia de los dos tintes de marcaje. El sesgo de tinte puede provenir de diversas fuentes, incluidas las diferencias en la eficiencia del marcaje de tinte, la sensibilidad al calor y la luz, así como la configuración del escáner para escanear dos canales. Algunos métodos comúnmente usados para calcular el factor de normalización incluyen: (i) normalización global que usa todos los genes en la matriz; (ii) normalización de genes de mantenimiento que usa genes de mantenimiento/invariantes expresados constantemente; y (iii) normalización de controles internos que usa una cantidad conocida de genes de control exógenos añadidos durante la hibridación (Quackenbush Nat. Gineta. 32 (Supl.), 496-501 (2002)). En una modalidad, los genes intrínsecos descritos en la presente descripción pueden normalizarse para controlar los genes de mantenimiento. Por ejemplo, los genes de mantenimiento descritos en la Publicación de Patente de los Estados Unidos 2008/0032293 puede usarse para la normalización. Los genes de mantenimiento ilustrativos incluyen MRPL19, PSMC4, SF3A1, PUM1, ACTB, GAPD, GUSB, RPLPO, y TFRC. Un experto en la técnica entenderá que los métodos descritos en la presente descripción no se unen por la normalización a cualquier gen de mantenimiento particular, y que puede usarse cualquier gen de mantenimiento adecuado conocido en la técnica.

Son posibles muchos enfoques de normalización, y frecuentemente pueden aplicarse en cualquiera de varios puntos del análisis. En una modalidad, los datos de las micromatrices se normalizan con el uso del método LOWESS, que es una función de normalización suavizado del diagrama de dispersión ponderada localmente global. En otra modalidad, los datos de qPCR se normalizan con la media geométrica del conjunto de múltiples genes de mantenimiento.

El "centrado medio" además puede usarse para simplificar la interpretación. Por lo general, para cada descriptor, se resta el valor promedio de ese descriptor para todas las muestras. De esta manera, la media de un descriptor coincide con el origen, y todos los descriptores están "centrados" en cero. En el "escalado de varianza unitaria", los datos pueden escalarse con la misma varianza. Por lo general, el valor de cada descriptor se escala mediante $1/\text{StDev}$, donde StDev es la desviación estándar para ese descriptor para todas las muestras. El "escalado de Pareto" es, en cierto sentido, intermedio entre el centrado medio y el escalado de la varianza unitaria. En el escalado de Pareto, el valor de cada

descriptor se escala en $1/\sqrt{\text{StDev}}$, donde StDev es la desviación estándar para ese descriptor para todas las muestras. De esta manera, cada descriptor tiene una varianza numéricamente igual a su desviación estándar inicial. El escalado de Pareto puede realizarse, por ejemplo, en datos sin procesar o datos centrados en la media.

5 El "escalado logarítmico" puede usarse para ayudar a la interpretación cuando los datos tienen un sesgo positivo y/o cuando los datos abarcan un amplio intervalo, *por ejemplo*, varios órdenes de magnitud. Por lo general, para cada descriptor, el valor se reemplaza por el logaritmo de ese valor. En el "escalado de intervalo igual", cada descriptor se divide por el intervalo de ese descriptor para todas las muestras. De esta manera, todos los descriptores tienen el mismo intervalo, es decir, 1. Sin embargo, este método es sensible a la presencia de puntos atípicos. En "autoescalado", cada vector de datos se centra en la media y se escala la varianza de la unidad. Esta técnica es muy útil porque cada descriptor se pondera por igual después y los valores grandes y pequeños se tratan con igual énfasis. Esto puede ser importante para los genes expresados a niveles muy bajos, pero aún detectables.

15 En una modalidad, los datos se recolectan para una o más muestras de prueba y se clasifican usando el modelo de clasificación NANO46 descrito en la presente descripción. Al comparar datos de múltiples análisis (*por ejemplo*, que comparan los perfiles de expresión para una o más muestras de prueba con los centroides construidos a partir de muestras recolectadas y analizadas en un estudio independiente), será necesario normalizar los datos en estos conjuntos de datos. En una modalidad, la discriminación ponderada a distancia (DWD) se usa para combinar estos conjuntos de datos (Benito y otros (2004) *Bioinformática* 20(1): 105-114). DWD es una herramienta de análisis multivariante que puede identificar sesgos sistemáticos presentes en conjuntos de datos separados y después hacer un ajuste global para compensar estos sesgos; en esencia, cada conjunto de datos separado es una nube multidimensional de puntos de datos, y DWD toma dos nubes de puntos y cambia una de tal manera que se superpone de manera más óptima con la otra.

25 Los métodos descritos en la presente descripción pueden implementarse y/o registrarse los resultados con el uso de cualquier dispositivo capaz de implementar los métodos y/o registrar los resultados. Los ejemplos de dispositivos que se pueden usar incluyen, pero no se limitan a, dispositivos electrónicos computacionales, que incluyen computadoras de todo tipo. Cuando los métodos descritos en la presente descripción se implementan y/o registran en una computadora, el programa de computadora que puede usarse para configurar la computadora para llevar a cabo las etapas de los métodos puede estar contenido en cualquier medio legible por computadora capaz de contener el programa de computadora. Los ejemplos de medios legibles por computadora que pueden usarse incluyen, pero no se limitan a, disquetes, CD-ROM, los DVD, ROM, RAM, y otros dispositivos de almacenamiento de memoria y computadora. El programa de computadora que puede usarse para configurar la computadora para llevar a cabo las etapas de los métodos y/o registrar los resultados además puede proporcionarse a través de una red electrónica, por ejemplo, a través de internet, una intranet, u otra red.

35 Cálculo del riesgo de recurrencia.

En la presente descripción se proporcionan métodos para predecir el resultado del cáncer de mama dentro del contexto del subtipo intrínseco y, opcionalmente, otras variables clínicas. El resultado puede referirse a la supervivencia general o específica de la enfermedad, la supervivencia libre de eventos, o el resultado en respuesta a un tratamiento o terapia en particular. En particular, los métodos pueden usarse para predecir la probabilidad de supervivencia a largo plazo, libre de enfermedad. "Predecir la probabilidad de supervivencia de una paciente con cáncer de mama" se entiende como evaluar el riesgo de que una paciente muera como un resultado del cáncer de mama subyacente. El término "supervivencia a largo plazo sin enfermedad" se entiende que significa que la paciente no muere o sufre una recurrencia del cáncer de mama subyacente en un período de al menos cinco años, o al menos diez o más años, después del diagnóstico inicial o tratamiento.

50 En una modalidad, el resultado se predice en base a la clasificación de un sujeto de acuerdo con el subtipo. Además de proporcionar una asignación de subtipo, el modelo bioinformático NANO46 proporciona una medición de la similitud de una muestra de prueba con los cuatro subtipos que se traduce en una puntuación de riesgo de recurrencia (ROR) que puede usarse en cualquier población de pacientes independientemente del estado de la enfermedad y opciones de tratamiento. Los subtipos intrínsecos y ROR además tienen valor en la predicción de la respuesta patológica completa en mujeres tratadas con, por ejemplo, taxano neoadyuvante y quimioterapia con antraciclina (Rouzier y otros, *J Clin Oncol* 23: 8331-9 (2005)). Así, en diversas modalidades de la presente descripción, se usa un modelo de riesgo de recurrencia (ROR) para predecir el resultado. Con el uso de estos modelos de riesgo, los sujetos pueden estratificarse en grupos de riesgo de recurrencia bajo, medio y alto. El cálculo de ROR puede proporcionar información pronóstica para guiar las decisiones de tratamiento y/o monitorear la respuesta a la terapia.

60 En algunas modalidades descritas en la presente descripción, el rendimiento pronóstico de los subtipos intrínsecos definidos por NANO46 y/u otros parámetros clínicos se evalúa utilizando un Análisis de modelo de peligros proporcionales Cox, que es un método de regresión para los datos de supervivencia que proporciona una estimación de la relación de peligro y su intervalo de confianza. El modelo de Cox es una técnica estadística bien reconocida para explorar la relación entre la supervivencia de un paciente y variables particulares. Este método estadístico permite la estimación del peligro (*es decir*, riesgo) de los individuos dadas sus variables pronósticas (*por ejemplo*, perfil de expresión de genes intrínsecos con o sin factores clínicos adicionales, como se describe en la presente descripción). La "relación de peligro" es el riesgo de muerte en cualquier momento dado para pacientes que muestran variables de pronóstico particulares. Ver generalmente Spruance y otros, *Antimicrob. Agents & Chemo.* 48:2787-92 (2004).

5 El modelo de clasificación NANO46 descrito en la presente descripción puede entrenarse para el riesgo de recurrencia con el uso de distancias de subtipo (o correlaciones) solo, o con el uso de distancias de subtipo con variables clínicas como se discutió *arriba*. En una modalidad, la puntuación de riesgo para una muestra de prueba se calcula con el uso de distancias de subtipo intrínseco solo con el uso de la siguiente ecuación:

10 $ROR = 0,05 \cdot \text{Basal} + 0,11 \cdot \text{Her2} + -0,25 \cdot \text{LumA} + 0,07 \cdot \text{LumB} + -0,11 \cdot \text{Normal}$, donde las variables "Basal", "Her2", "LumA", "LumB" y "Normal" son las distancias al centroide para cada clasificador respectivo cuando el perfil de expresión de una muestra de prueba se compara con los centroides construidos con el uso de datos de expresión de genes depositados con expresión de genes Omnibus (GEO).

15 La puntuación de riesgo además puede calcularse con el uso de una combinación de subtipo de cáncer de mama y las variables clínicas tamaño del tumor (T) y estado de los ganglios linfáticos (N) con el uso de la siguiente ecuación: $ROR (\text{completo}) = 0,05 \cdot \text{Basal} + 0,1 \cdot \text{Her2} + -0,19 \cdot \text{LumA} + 0,05 \cdot \text{LumB} + -0,09 \cdot \text{Normal} + 0,16 \cdot T + 0,08 \cdot N$, nuevamente cuando se comparan los perfiles de expresión de prueba con los centroides construidos con el uso de los datos de expresión de genes depositados con GEO como número de acceso GSE2845.

20 Aun en otra modalidad, la puntuación de riesgo para una muestra de prueba se calcula con el uso de distancias de subtipo intrínsecos solo con el uso de la siguiente ecuación:

25 $ROR-S = 0,05 \cdot \text{Basal} + 0,12 \cdot \text{Her2} + -0,34 \cdot \text{LumA} + 0,023 \cdot \text{LumB}$, donde las variables "Basal", "Her2", "LumA" y "LumB" son como se describen *arriba* y los perfiles de expresión de prueba se comparan con los centroides construidos con el uso de los datos de la expresión génica depositados en GEO con el número de acceso GSE2845. Aun en otra modalidad, la puntuación de riesgo puede calcularse además con el uso de una combinación de subtipo de cáncer de mama y la variable clínica de tamaño tumoral (T) con el uso de la siguiente ecuación (donde las variables son como descrito *arriba*): $ROR-C = 0,05 \cdot \text{Basal} + 0,11 \cdot \text{Her2} + -0,23 \cdot \text{LumA} + 0,09 \cdot \text{LumB} + 0,17 \cdot T$.

30 Aun en otra modalidad, la puntuación de riesgo para una muestra de prueba se calcula con el uso de distancias de subtipo intrínsecas en combinación con la firma de proliferación ("Prolif") con el uso de la siguiente ecuación:

35 $ROR-P = -0,001 \cdot \text{Basal} + 0,7 \cdot \text{Her2} + -0,95 \cdot \text{LumA} + 0,49 \cdot \text{LumB} + 0,34 \cdot \text{Prolif}$, donde las variables "Basal", "Her2", "LumA", "LumB" y "Prolif" son como descrito *arriba* y los perfiles de expresión de prueba se comparan con los centroides construidos con el uso de los datos de expresión de genes depositados en GEO con el número de acceso GSE2845.

40 Aun en otra modalidad, la puntuación de riesgo puede calcularse además con el uso de una combinación de subtipo de cáncer de mama, firma de proliferación y la variable clínica de tamaño tumoral (T) con el uso de el ROR-PT descrito junto con la Tabla 3 *arriba*.

40 Detección de subtipos

45 La inmunohistoquímica para estrógeno (ER), progesterona (PgR), HER2 y Ki67 se realizó simultáneamente en secciones en serie con el método estándar del complejo de estreptavidina-biotina con 3,3'-diaminobencidina como cromógeno. La tinción para la interpretación ER, PgR y HER2 puede realizarse como se describió anteriormente (Cheang y otros, Clin Cancer Res. 2008;14(5):1368-1376.), sin embargo, puede usarse cualquier método conocido en la técnica.

50 Por ejemplo, un anticuerpo Ki67 (clon SP6; ThermoScientific, Fremont, CA) puede aplicarse a una dilución 1:200 durante 32 minutos, siguiendo del protocolo de inmunotinción automática Ventana Benchmark (Ventana, Tucson AZ) con el acondicionador celular 1 estándar (CC1, un tampón patentado) a 98°C durante 30 minutos. Puede usarse un anticuerpo ER (clon SP1; ThermoFisher Scientific, Fremont CA) a una dilución 1:250 con una incubación de 10 minutos, después de una recuperación de antígeno por microondas de 8 minutos en citrato de sodio 10 mM (pH 6,0). El anticuerpo PR listo para el uso (clon 1E2; Ventana) puede usarse siguiendo el protocolo de CC1 como se indicó anteriormente. La tinción de HER2 puede realizarse con un anticuerpo SP3 (ThermoFisher Scientific) a una dilución 1:100 después de la recuperación del antígeno en el tampón Tris 0,05 M (pH 10,0) con calentamiento a 95°C en un vaporizador durante 30 minutos. Para el ensayo de hibridación fluorescente *in situ* (FISH) de HER2, los portaobjetos pueden hibridarse con sondas para LSI (identificador específico de locus) HER2/neu y para el centrómero 17 mediante el uso del kit de sonda de ADN HER-2 PathVysion (Abbott Molecular, Abbott Park, IL) de acuerdo con el fabricante instrucciones, con las modificaciones para el pretratamiento e hibridación como se describió anteriormente (Amplification of EMSY, a novel oncogene on 11q13, in high grade ovarian surface epithelial carcinomas. Gynecol Oncol. 2006;100(2):264-270). Los portaobjetos pueden contrateñirse después con 4',6-diamidino-2-fenilindol, el material teñido se visualizó en un microscopio epifluorescente Zeiss Axioplan y las señales se analizaron con un sistema de adquisición de imágenes Metafer (Metasystems, Altlußheim, Alemania). Después, dos patólogos calificaron la expresión de biomarcadores a partir de los ensayos de inmunohistoquímica, quienes no conocían las características clínico-patológicas y los resultados, y usaron los criterios previamente establecidos y publicados para los niveles de expresión de biomarcadores que se habían desarrollado en otras cohortes de cáncer de mama.

65

Los tumores se consideraron positivos para ER o PR si se observó inmunotinción en más del 1% de los núcleos tumorales, como se describió anteriormente. Los tumores se consideraron positivos para HER2 si la inmunotinción se calificó como 3+ de acuerdo con los criterios de HercepTest, con una relación de amplificación para la hibridación fluorescente *in situ* de 2.0 o más siendo el punto de corte que se usó para segregar los tumores equívocos de inmunohistoquímica (clasificado como 2+) (Yaziji, y otros, JAMA, 291(16):1972-1977 (2004)). Ki67 se clasificó visualmente por los dos patólogos por el porcentaje de núcleos de células tumorales con inmunotinción positiva por encima del nivel de fondo.

Pueden usarse además otros métodos para detectar subtipos. Estas técnicas incluyen ELISA, transferencias de tipo Western, transferencias de tipo Northern o análisis FACS.

Kits

La presente descripción describe además los kits útiles para clasificar los subtipos intrínsecos de cáncer de mama y/o proporcionar la información pronóstica para identificar el riesgo de recurrencia. Estos kits comprenden un conjunto de sondas de captura y/o cebadores específicos para los genes intrínsecos enumerados en la Tabla 1. El kit puede comprender además un medio legible por computadora.

En una modalidad de la presente descripción, las sondas de captura se inmovilizan en una matriz. Se entiende por "matriz" un soporte sólido o un sustrato con sondas peptídicas o de ácido nucleico unidas al soporte o sustrato. Las matrices típicamente comprenden una pluralidad de sondas de captura diferentes que están acopladas a una superficie de un sustrato en diferentes ubicaciones conocidas. Las matrices de la descripción comprenden un sustrato que tiene una pluralidad de sondas de captura que pueden unirse específicamente a un producto intrínseco de expresión de genes. El número de sondas de captura en el sustrato varía con el propósito para el cual está destinada la matriz. Las matrices pueden ser matrices de baja densidad o matrices de alta densidad y pueden contener 4 o más, 8 o más, 12 o más, 16 o más, 32 o más direcciones, pero comprenderán mínimamente sondas de captura para los 46 genes intrínsecos enumerados en Tabla 1.

Las técnicas para la síntesis de estas matrices con el uso de los métodos de síntesis mecánica se describen en, *por ejemplo*, la patente de los Estados Unidos núm. 5.384.261. El conjunto puede fabricarse en una superficie de prácticamente cualquier forma o incluso en una multiplicidad de superficies. Las matrices pueden ser sondas (por ejemplo, sondas de unión al ácido nucleico) sobre perlas, geles, superficies poliméricas, fibras tales como fibra óptica, vidrio o cualquier otro sustrato apropiado, ver las patentes de los Estados Unidos Núms. 5.770.358, 5.789.162, 5.708.153, 6.040.193 y 5.800.992. Las matrices pueden empaquetarse de tal manera que permitan diagnósticos u otra manipulación en el dispositivo. Ver, por ejemplo, las patentes de los Estados Unidos Núms. 5.856.174 y 5.922.591.

En otra modalidad, el kit comprende un conjunto de cebadores oligonucleotídicos suficientes para la detección y/o cuantificación de cada uno de los genes intrínsecos enumerados en la Tabla 1. Los cebadores oligonucleotídicos pueden proporcionarse en forma liofilizada o reconstituida, o pueden proporcionarse como un conjunto de secuencias de nucleótidos. En una modalidad, los cebadores se proporcionan en un formato de microplaca, donde cada conjunto de cebadores ocupa un pocillo (o múltiples pocillos, como en el caso de las réplicas) en la microplaca. La microplaca puede comprender además cebadores suficientes para la detección de uno o más genes de mantenimiento como se discutió *más abajo*. El kit puede comprender además reactivos e instrucciones suficientes para la amplificación de productos de expresión a partir de los genes enumerados en la Tabla 1.

Para facilitar el acceso rápido, *por ejemplo*, para la comparación, revisión, recuperación y/o modificación, las firmas moleculares/perfiles de expresión se registran típicamente en una base de datos. Más típicamente, la base de datos es una base de datos relacional accesible por un dispositivo computacional, aunque otros formatos, por ejemplo, archivos indexados de perfiles de expresión manualmente accesibles como fotografías, lecturas de imágenes analógicas o digitales, hojas de cálculo, *etc.* pueden usarse. Independientemente de si los patrones de expresión registrados inicialmente son de naturaleza analógica o digital, los patrones de expresión, los perfiles de expresión (patrones de expresión colectiva) y las firmas moleculares (patrones de expresión correlacionados) se almacenan digitalmente y se acceden a ellos a través de una base de datos. Típicamente, la base de datos se compila y mantiene en una instalación central, con acceso disponible de forma local y/o remota.

Dispositivos y Pruebas

General - el sistema de análisis NanoString nCounter ofrece mediciones directas y multiplexadas de la expresión de genes a través de lecturas digitales de la abundancia relativa de cientos de transcritos de ARNm. El sistema de análisis nCounter usa pares de sondas específicas de genes (Figura 7) que se mezclan juntas para formar un solo reactivo denominado un CodeSet. Los pares de sondas se hibridan directamente con la muestra de ARNm en solución, eliminando cualquier reacción enzimática que pueda introducir sesgo en los resultados.

Después de la hibridación, todas las etapas de procesamiento de muestras se automatizan en la estación de preparación nCounter. Primero, se eliminan el exceso de captura y las sondas reporteras (Figura 8), seguido de la unión de los complejos diana-sonda en ubicaciones aleatorias sobre la superficie del cartucho nCounter a través de un enlace estreptavidina-biotina (Figura 9).

Finalmente, los complejos sonda/diana se alinean e inmovilizan (Figura 10) en el cartucho nCounter. La sonda reportera transporta la señal fluorescente; la sonda de captura permite que el complejo se inmovilice para la recopilación de datos. Pueden combinarse hasta 800 pares de sondas, cada una específica para un gen en particular, con una serie de controles internos para formar un CodeSet.

Después de completado el procesamiento de la muestra, los cartuchos se colocan en el analizador digital nCounter para la recopilación de datos. Cada molécula diana de interés se identifica por el "código de color" generado por seis puntos fluorescentes ordenados presentes en la sonda reportera. Las sondas reporteras en la superficie del cartucho se cuentan y tabulan después, para cada molécula objetivo (Figura 11).

Reactivos y componentes de prueba - la prueba de cáncer de mama medirá simultáneamente los niveles de expresión de NANO46 más ocho genes de mantenimiento en una única reacción de hibridación con el uso de un nCounter CodeSet diseñado específicamente para los genes. Cada ensayo incluye además los controles de ensayo positivos compuestos por una valoración lineal *in vitro* de transcritos de ARN transcritos y sondas correspondientes, y un conjunto de sondas sin homología de secuencia con el elemento, número entero o etapa, o grupo de elementos, números enteros o etapas.

EJEMPLOS

Ejemplo 1. Prueba de subtipaje NANO46

La Figura 5 describe los procesos de ensayo asociados con la prueba de subtipaje intrínseco de cáncer de mama. Después del aislamiento de ARN, la prueba medirá simultáneamente los niveles de expresión de 46 genes diana más ocho genes de mantenimiento en una única reacción de hibridación con el uso de un nCounter CodeSet diseñado específicamente para los genes. Por ejemplo, los genes de mantenimiento descritos en la Publicación de Patente de los Estados Unidos 2008/0032293 puede usarse para la normalización. Los genes de mantenimiento ilustrativos incluyen MRPL19, PSMC4, SF3A1, PUM1, ACTB, GAPD, GUSB, RPLP0, y TFRC. Los genes de mantenimiento se usan para normalizar la expresión de la muestra de tumor. Cada prueba incluye además una muestra de referencia que consiste en ARN transcritos *in vitro* de los 58 objetivos con fines de normalización.

Revisión/adquisición de tejidos FFPE y extracción de ARN: La prueba de subtipaje intrínseco de cáncer de mama usará ARN extraído del tejido fijado con formalina e incluido en parafina (FFPE) que se ha diagnosticado como carcinoma invasivo de mama. Un patólogo revisa un portaobjetos teñido con H & E para identificar el área de tejido que contiene suficiente contenido de tejido tumoral para la prueba. Las secciones de tejido montadas en los portaobjetos sin teñir se procesan mediante macro-disección del área tumoral identificada en cada portaobjetos para eliminar cualquier tejido normal adyacente. Después se aísla el ARN del tejido tumoral y se elimina el ADN de la muestra.

Configuración del ensayo e inicio de la hibridación: Para cada lote de hasta 10 muestras de ARN aisladas de un tumor de mama, el usuario configurará una ejecución con el uso del programa del sistema nCounter Analysis x5, que rastrea el procesamiento de muestras, los lotes de reactivos y resultados para cada muestra. Para iniciar el ensayo, el usuario pipeteará la cantidad especificada de ARN en tubos separados dentro de un tubo de 12 tiras de reacción y agregará el CodeSet y el tampón de hibridación. Se pipetea una muestra de referencia en los dos tubos restantes con CodeSet y el tampón de hibridación. El CodeSet consta de sondas para cada gen que se especifica, sondas adicionales para la normalización endógena de los genes de "mantenimiento" y controles positivos y negativos que se agregan al ensayo. La muestra de referencia consta del ARN transcrito *in vitro* para los genes diana y los genes de mantenimiento. Una vez que los reactivos de hibridación se agregan a los tubos respectivos, el usuario transfiere el tubo con la tira a un bloque térmico de tapa calentada durante un período de tiempo específico a una temperatura establecida.

Purificación y unión en la estación de preparación: Al completar la hibridación, el usuario transferirá el tubo con la tira que contiene el conjunto de 10 ensayos y 2 muestras de referencia a las mediciones en un espectrofotómetro UV-Vis de tipo cubeta; por lo tanto, el protocolo de NanoString incluye una etapa para cuantificar el ARN total con el uso de un espectrofotómetro de bajo volumen tal como el espectrofotómetro NanoDrop™. NanoString definirá las especificaciones de rendimiento para el espectrofotómetro de modo que el intervalo de entrada de ARN recomendado para la prueba esté por encima del límite de detección del espectrofotómetro de bajo volumen y sea medible de forma reproducible.

Hibridación - Para cada conjunto de hasta 10 muestras de ARN, el usuario pipeteará la cantidad especificada de ARN en tubos separados dentro de un tubo de 12 tiras de reacción y agregará el CodeSet y el tampón de hibridación. Se pipetea una muestra de referencia en los dos tubos restantes con CodeSet y el tampón de hibridación. El CodeSet consta de sondas para cada gen que se especifica, sondas adicionales para la normalización endógena de los genes de "mantenimiento" y controles positivos y negativos. Las sondas dentro del CodeSet pertenecientes a cada uno de estos genes dentro de los cuatro grupos (genes diana, genes de mantenimiento y controles positivos y negativos) tienen asignado un código único y, por lo tanto, son identificables individualmente dentro de cada ejecución. La muestra de referencia consta del ARN transcrito *in vitro* para los genes diana y los genes de mantenimiento. Una vez que los reactivos de hibridación se agregan a los tubos respectivos, el usuario transfiere el tubo con la tira a un bloque térmico de tapa calentada durante un período de tiempo específico a una temperatura establecida.

5 Requisito de bloqueo térmico con tapa calentada para la etapa de hibridación - el ensayo nCounter incluye una hibridación durante la noche en condiciones isotérmicas. Debido a que la hibridación nocturna se realiza en un pequeño volumen a temperatura elevada, se debe tener cuidado para evitar la evaporación. Muchos termocicladores de PCR comerciales están equipados con tapas calientes que evitarán la evaporación de pequeños volúmenes de líquido. Debido a que el ensayo no requiere de ningún control preciso de la rampa de temperatura, cualquier bloque de calor con una tapa calentada programable y un bloque con dimensiones que se ajusten a los tubos NanoString funcionará con el ensayo NanoString. NanoString planea proporcionar especificaciones para los bloques térmicos que cumplan con los requisitos del ensayo.

10 Purificación y unión en la estación de preparación: al completar la hibridación, el usuario transferirá el tubo con la tira que contiene el conjunto de 10 ensayos y 2 muestras de referencia a la estación de preparación nCounter junto con los reactivos y desechables preempaquetados requeridos que se describen en la Tabla 1. Las placas de preparación contienen los reactivos necesarios para la purificación del exceso de sondas y la unión al cartucho (ver la sección IIIC más abajo para obtener una descripción detallada del proceso de purificación). Las placas de preparación se centrifugan en una centrifuga de cubeta oscilante antes de colocarlas en la plataforma de la estación de preparación. Después, un proceso de purificación automatizado se elimina el exceso de sonda de captura y reportera a través de dos etapas sucesivas de captura por perlas magnéticas impulsadas por hibridación. La estación de preparación nCounter transfiere después los complejos diana/sonda purificados a un cartucho nCounter para capturarlos en un portaobjetos de vidrio. Después de completar la ejecución, el usuario retira el cartucho de la estación de preparación y lo sella con una película adhesiva.

25 Imágenes y análisis en el analizador digital - El cartucho sellado se inserta en el analizador digital nCounter que cuenta el número de sondas capturadas en el portaobjetos para cada gen, que corresponde a la cantidad diana en solución. Después, el programa automatizado verifica después los umbrales para los genes de mantenimiento, la muestra de referencia y los controles positivos y negativos para calificar cada ensayo y garantizar que el procedimiento se realizó correctamente. Los genes de mantenimiento proporcionan una medida de la integridad del ARN, y los umbrales indican cuándo una muestra de ARN analizada está demasiado degradada para ser analizada por la prueba debido a un manejo o almacenamiento incorrecto del tejido o ARN (por ejemplo, fijación tumoral inadecuada, almacenamiento en bloque de FFPE, almacenamiento de ARN, manejo de ARN que introduce RNasa). Los controles de ensayo positivo y negativo indican una falla del proceso de ensayo (por ejemplo, error en la configuración del ensayo, tal como la mezcla de muestras con CodeSet, o el procesamiento de muestras, tal como temperatura). Las señales de cada muestra se normalizan después con el uso de los genes de mantenimiento para controlar la calidad de la muestra de entrada. Después, las señales se normalizan con la muestra de referencia dentro de cada ejecución para controlar las variaciones de ejecución a ejecución. Los datos normalizados resultantes se ingresan en el algoritmo de subtipaje intrínseco de cáncer de mama para determinar el subtipo intrínseco del tumor, la puntuación de riesgo de recaída y la clasificación de riesgo.

40 Instrumentación- el sistema de análisis nCounter se compone de dos instrumentos, la estación de preparación nCounter usada para el procesamiento posterior a la hibridación y el analizador digital usado para la recopilación y el análisis de datos.

45 Estación de preparación nCounter - La estación de preparación nCounter (Figura 13) es un robot automatizado de manejo de fluidos que procesa las muestras después de la hibridación para prepararlas para la recolección de datos en el analizador digital nCounter. Antes del procesamiento en la estación de preparación, el ARN total extraído de las muestras de tejido FFPE (fijado en formalina, embebido en parafina) se hibrida con las sondas reporteras NanoString y sondas de captura de acuerdo con el protocolo nCounter descrito anteriormente.

50 La hibridación con el ARN diana es impulsada por el exceso de sondas NanoString. Para analizar con precisión estas moléculas hibridadas, primero se purifican a partir del exceso de sondas restantes en la reacción de hibridación. La estación de preparación aísla las moléculas de ARNm hibridadas del exceso de sondas reporteras y de captura con el uso de dos etapas secuenciales de purificación con perlas magnéticas. Estas purificaciones de afinidad utilizan perlas magnéticas modificadas con oligonucleótidos personalizados que retienen solo los complejos tripartitos de las moléculas de ARNm que están unidas tanto a una sonda de captura como a una sonda reportera.

55 A continuación, esta solución de complejos tripartitos se lava a través de una celda de flujo en el cartucho de muestra NanoString. Una superficie de esta celda de flujo está recubierta con un hidrogel de polietilenglicol (PEG) que está densamente impregnado con estreptavidina covalentemente unida. A medida que la solución pasa a través de la celda de flujo, los complejos tripartitos se unen a la estreptavidina en el hidrogel a través de moléculas de biotina que se incorporan a cada sonda de captura. El hidrogel de PEG actúa no solo para proporcionar una superficie densa en estreptavidina sobre la cual los complejos tripartitos pueden unirse específicamente, sino que también inhibe la unión no específica de cualquier exceso de sondas reporteras restantes.

60 Después de que los complejos se unen a la superficie de la celda de flujo, se aplica un campo eléctrico a lo largo de la celda de flujo de cada cartucho de muestra para facilitar la identificación óptica y el orden de los puntos fluorescentes que conforman cada sonda reportera. Debido a que las sondas reporteras son ácidos nucleicos cargados, el voltaje aplicado les imparte una fuerza que las estira y orienta uniformemente a lo largo del campo eléctrico. Mientras se aplica el voltaje, la estación de preparación agrega un reactivo de inmovilización que bloquea a los reporteros en la configuración alargada

después que se retira el campo eléctrico. Una vez que los reporteros están inmovilizados, el cartucho puede transferirse al analizador digital nCounter para la recopilación de datos. Todos los componentes consumibles y reactivos necesarios para el procesamiento de muestras en la estación de preparación se proporcionan en el kit maestro nCounter. Estos reactivos están listos para cargar en la plataforma de la estación de preparación nCounter, que puede procesar hasta 10 muestras y 2 muestras de referencia por ejecución en aproximadamente 2,5 horas.

Analizador digital nCounter - El analizador digital nCounter (Figura 14) recopila datos al tomar imágenes de los reporteros fluorescentes inmovilizados en el cartucho de muestra con una cámara CCD a través de una lente objetivo de microscopio. Debido a que las sondas reporteras fluorescentes son pequeños códigos de barras de una sola molécula con características de longitud de onda más pequeña que la luz visible, el analizador digital usa imágenes de gran aumento y difracción limitada para resolver la secuencia de los puntos en los códigos de barras fluorescentes.

El analizador digital captura cientos de campos de visión (FOV) consecutivos que pueden contener cientos o miles de sondas reporteras discretas. Cada FOV es una combinación de cuatro imágenes monocromáticas capturadas a diferentes longitudes de onda. La superposición resultante puede considerarse como una imagen de cuatro colores en azul, verde, amarillo y rojo. Cada FOV de 4 colores se procesa en tiempo real para proporcionar un "conteo" para cada código de barras fluorescente en la muestra. Debido a que cada código de barras identifica específicamente una sola molécula de ARNm, los datos resultantes del Analizador Digital son una medida precisa de la abundancia relativa de cada ARNm de interés en una muestra biológica.

Programa - la estación de preparación y el analizador digital son unidades independientes que no requieren conexión a una PC externa, pero deben conectarse en red entre sí con el uso de una red de área local (LAN). El programa del sistema nCounter gestiona de forma segura las operaciones a través de cuentas de usuario y permisos. Ambos instrumentos usan asistentes de configuración y proceso en una interfaz de usuario de pantalla táctil incorporada para guiar al usuario a través del procesamiento de la muestra y las etapas de recopilación de datos del ensayo. El usuario es guiado a través del procedimiento mediante instrucciones paso a paso en la estación de preparación y el analizador digital. La pantalla táctil del instrumento usa un método sensible a la presión para controlar las operaciones y permite al usuario interactuar con el sistema tocando una selección en la pantalla. Debido a que la pantalla táctil proporciona una interfaz humana limitada para la entrada de datos, el sistema además aloja una aplicación basada en la web para la administración de cuentas de usuario, definición de lotes de muestra y seguimiento del estado de la muestra.

Cuando se procesan las muestras, el programa del sistema sigue la cuenta de usuario y los lotes de reactivos para cada muestra en un repositorio de datos centralizado. Después de que el analizador digital adquiere los datos de la expresión para una muestra, primero se analiza para garantizar que se cumplan todas las métricas de control de calidad especificadas previamente. Los datos calificados se procesan a través de un algoritmo PAM50 bloqueado para generar un informe que contiene un subtipo intrínseco y una puntuación de riesgo de recurrencia (ROR). El informe de muestra se transfiere al repositorio central donde un usuario puede acceder de forma segura para descargarlo con los permisos correctos.

Algoritmo de subtipaje intrínseco de cáncer de mama - El sistema nCounter se usará para identificar el subtipo intrínseco de un carcinoma invasivo de mama extirpada con el uso de un algoritmo clasificador de 50 genes originalmente denominado PAM50 (Parker JS, y otros, Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *Journal of Clinical Oncology*, 27: 1160-1167 (2009)). El perfil de expresión de genes asignará un cáncer de mama a una de cuatro clases moleculares o subtipos intrínsecos: Tipo Basal, Luminal A, Luminal B y enriquecido con HER2. Más abajo se proporciona una breve descripción de cada subtipo.

Subtipos luminales: Los subtipos más comunes de cáncer de mama son los subtipos luminales en la población con receptores hormonales positivos, Luminal A y Luminal B. Estudios previos sugieren que el luminal A comprende aproximadamente del 30% al 40% y el luminal B aproximadamente el 20% de los cánceres de mama² y más del 90% de los cánceres de mama con receptores hormonales positivos. El patrón de expresión de genes de estos subtipos se asemeja al componente epitelial luminal de la mama (Nielsen, TO y otros, A comparison of PAM50 intrinsic subtyping with immunohistochemistry and clinical prognostic factors in tamoxifen-treated estrogen receptor positive breast cancer. *Clinical Cancer Research*, 16:5222-5232 (2010)). Estos tumores se caracterizan por una alta expresión del receptor de estrógenos (ER), el receptor de progesterona (PR) y los genes asociados con la activación de ER² tales como LIV1, GATA3 y ciclina D1, así como la expresión de citoqueratinas luminales 8 y 18.

Luminal A: Los cánceres de mama Luminal A (LumA) exhiben una baja expresión de genes asociados con la activación del ciclo celular y el agrupamiento ERBB2, lo que resulta en un mejor pronóstico que el luminal B. El subgrupo Luminal A tiene el pronóstico más favorable de todos los subtipos y está enriquecido para tumores con respuesta a la terapia endocrina.

Luminal B: Los cánceres de mama Luminal B (LumB) expresan genes ER y asociados a ER, pero en menor medida que LumA. Los genes asociados con la activación del ciclo celular se expresan altamente y este tipo de tumor puede ser HER2(+) o HER2(-). El pronóstico es desfavorable (a pesar de la expresión de ER) y la respuesta a la terapia endocrina generalmente disminuye con relación a LumA.

Tipo basal: El subtipo tipo basal es generalmente negativo para ER, casi siempre es clínicamente negativo para HER2 y expresa un conjunto de biomarcadores "basales" que incluyen las citoqueratinas epiteliales basales (CK) y el receptor del factor de crecimiento epidérmico (EGFR). Los genes asociados con la activación del ciclo celular se expresan altamente.

5 Enriquecido con HER2: El subtipo enriquecido con HER2 es generalmente negativo para ER y es positivo para HER2, en la mayoría de los casos con alta expresión del agrupamiento ERBB2, incluidos ERBB2 y GRB7. Los genes asociados con la activación del ciclo celular se expresan altamente y estos tumores tienen un resultado malo.

10 Los valores de corte para el algoritmo de subtipaje intrínseco están predefinidos a partir de conjuntos de entrenamiento que definen a seguir: 1) centroides de subtipo intrínseco (es decir, el perfil de expresión de genes prototípico de cada subtipo), 2) coeficientes para la puntuación de riesgo de recurrencia (ROR), y 3) clasificación de riesgo (bajo/intermedio/alto). Los centroides intrínsecos de subtipo (Luminal A, Luminal B, enriquecido con Her2, tipo Basal) se entrenaron con el uso de un conjunto clínicamente representativo de muestras de tumor de mama FFPE archivadas y recolectadas de múltiples sitios. El análisis de agrupamiento jerárquico de los datos de expresión de genes de las muestras
15 de tumor de mama FFPE se combinó con la biología del tumor de mama (es decir, la expresión de genes de subtipos intrínsecos previamente definidos) para definir el perfil de expresión prototípico (es decir, centroide) de cada subtipo. Un algoritmo computacional correlaciona el perfil normalizado de expresión de 50 genes de una muestra de tumor de cáncer de mama desconocido con cada una de las firmas de expresión prototípica de los cuatro subtipos intrínsecos de cáncer de mama. A la muestra de tumor se le asigna el subtipo con la mayor correlación positiva con la muestra.

20 Se usaron 304 muestras de tumor únicas con características clínicas bien definidas y datos de los resultados clínicos para establecer la puntuación ROR. La puntuación ROR se calcula con el uso de coeficientes de un modelo de Cox que incluye la correlación de Pearson (R) con cada subtipo intrínseco, una puntuación de proliferación (P) y el tamaño del tumor (T), como se muestra en la ecuación más abajo.

25

$$\text{ROR} = aR_{\text{LumA}} + aR_{\text{LumB}} + cR_{\text{Her2e}} + dR_{\text{basal}} + eP + fT$$

30 Para clasificar las muestras de tumor en grupos de riesgo específicos (bajo riesgo/riesgo intermedio/alto riesgo) en función de su puntuación ROR calculada, se establecieron valores de corte basados en la probabilidad de supervivencia libre de recurrencia en una población de pacientes que consiste en pacientes posmenopáusicos receptores hormonales positivos, tratados solo con la terapia endocrina.

35 Uso anticipado de la prueba de cáncer de mama NanoString en la práctica clínica - los oncólogos usan actualmente una serie de pruebas para desarrollar un protocolo de tratamiento para pacientes con cáncer de mama. En estos se incluyen las pruebas IHC / FISH como ER/PR IHC y HER2 IHC / FISH, y el ensayo Agendia MammaPrint® y la prueba Genomic Health Oncotype Dx®. Estas pruebas ofrecen al oncólogo información adicional sobre el pronóstico del paciente y los regímenes de tratamiento recomendados.

40 Estas pruebas, sin embargo, tienen limitaciones. Las pruebas de ER, PgR y Her2 se realizan localmente por patólogos y laboratorios de referencia, pero los desafíos con la estandarización generalizada de las pruebas de IHC y FISH están bien documentados (Lester, J y otros, Assessment of Tissue Estrogen and Progesterone Receptor Levels: A Survey of Current Practice, Techniques, and Quantitation Methods. The Breast Journal, 6:189-196 (2000); Wolff, A y otros, American Society of Clinical Oncology / College of American Pathologists Guideline Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer. Archives of Pathology and Laboratory Medicine, 131:18-43 (2007)). La prueba
45 MammaPrint está aprobada por la FDA para su uso solo con muestras de tejido congeladas o conservadas en fresco, aunque la mayoría de las muestras de tumores recolectadas en los Estados Unidos son FFPE en lugar de congeladas en fresco. Esta prueba tampoco se distribuye y solo está disponible a través de los laboratorios de referencia de Agendia. La prueba Oncotype Dx puede usarse para predecir el riesgo de recaída en pacientes con estadio I/II, nódulo negativo, receptores de estrógenos positivos que reciben terapia adyuvante con tamoxifeno, así como la respuesta a la
50 quimioterapia con ciclofosfamida/metotrexato/5-fluorouracilo (CMF). Sin embargo, esta prueba solo se ofrece como una prueba desarrollada en laboratorio (LDT) a través del laboratorio CLIA de Genomic Health y no está aprobada por la FDA para uso pronóstico ni aprobada por la FDA para predecir la respuesta a la quimioterapia.

55 NanoString visualiza un modelo que usaría la prueba de cáncer de mama junto con otras fuentes de datos clínicos actualmente disponibles para los oncólogos para el pronóstico del cáncer de mama en segmentos seleccionados de pacientes. La prueba de cáncer de mama sería una fuente adicional de información pronóstica que agregaría un valor significativo a los parámetros clínicos establecidos (es decir, tamaño del tumor, estado nodular) usados por los oncólogos para tratar a un paciente con cáncer de mama.

60 Métodos, ensayos y kits

Los métodos, ensayos y kits de la descripción incluyen una serie de métricas de control de calidad que se aplican automáticamente a cada muestra durante el análisis. Estas métricas evalúan el rendimiento del ensayo para determinar si los resultados se encuentran dentro de los valores esperados. Tras un análisis exitoso de estas métricas de control de
65 calidad, el ensayo ofrece los siguientes resultados:

Resultado	Valores de salida
El subtipo intrínseco de la muestra de cáncer de mama	Luminal A Luminal B Enriquecido con Her2 Tipo basal
Estimación individual de la probabilidad de recurrencia distante dentro de 10 años	0-100%
Puntuación de riesgo de recurrencia (ROR)	Valor entero en una escala de 0-100
Categoría de riesgo	Bajo, intermedio, alto

Subtipos intrínsecos

Se ha demostrado que el subtipo intrínseco de un tumor de cáncer de mama está relacionado con el pronóstico del cáncer de mama en estadio temprano. En promedio, los pacientes con un tumor Luminal A tienen resultados significativamente mejores que los pacientes con tumores Luminal B, enriquecidos con HER2, o de tipo basal.

El subtipo intrínseco se identifica comparando el perfil de expresión de genes de 50 genes en una muestra desconocida con los perfiles de expresión esperados para los cuatro subtipos intrínsecos. El subtipo con el perfil más similar se asigna a la muestra desconocida.

Los subtipos más comunes de cáncer de mama son los subtipos luminales, Luminal A (LumA) y Luminal B (LumB). Estudios previos sugieren que Luminal A comprende aproximadamente 30% a 40% y Luminal B aproximadamente 20% de los cánceres de mama. Sin embargo, más del 90% de los pacientes con receptores hormonales positivos tienen tumores luminales. El patrón de expresión de genes de estos subtipos se asemeja al componente epitelial luminal del tejido mamario. Estos tumores se caracterizan por una alta expresión del receptor de estrógenos (ER), el receptor de progesterona (PR) y los genes asociados con la activación de ER, tales como LIV1, GATA3 y ciclina D1, así como la expresión de citoqueratinas luminales 8 y 18. Los cánceres de mama Luminal A exhiben una menor expresión de genes asociados con la activación del ciclo celular en comparación con los cánceres de mama Luminal B, lo que resulta en un mejor pronóstico.

Estudios anteriores sugieren que el subtipo enriquecido en HER2 (Her2E) comprende aproximadamente el 20% de los cánceres de mama. Sin embargo, los tumores enriquecidos con HER2 son generalmente negativos para ER, por lo que se encontró que solo el 5% de la población de pacientes con ER positivo probado tenían cáncer de mama enriquecido con HER2. Independientemente del estado de ER, los tumores enriquecidos con HER2 son positivos para HER2 en la mayoría de los casos con alta expresión del agrupamiento ERBB2, incluyendo ERBB2 y GRB7. Los genes asociados con la activación del ciclo celular también se expresan altamente.

Los datos publicados sugieren que el subtipo de tipo basal comprende aproximadamente el 20% de los cánceres de mama. Sin embargo, los tumores de tipo basal son generalmente negativos para ER, por lo que solo el 1% de los pacientes con receptores hormonales positivos tienen cáncer de mama de tipo basal. El subtipo de tipo basal es casi siempre es clínicamente negativo para HER2 y expresa un conjunto de biomarcadores "basales" que incluyen las citoqueratinas epiteliales basales (CK) y el receptor del factor de crecimiento epidérmico (EGFR). Los genes asociados con la activación del ciclo celular se expresan altamente.

Puntuación ROR

La puntuación ROR es un valor entero en una escala de 0-100 que está relacionado con la probabilidad de un paciente individual de recurrencia distante dentro de los 10 años para el uso previsto en la población definida. La puntuación ROR se calcula comparando los perfiles de expresión de 46 genes en una muestra desconocida con los perfiles esperados para los cuatro subtipos intrínsecos, como se describió anteriormente, para calcular cuatro valores de correlación diferentes. Estos valores de correlación se combinan con una puntuación de proliferación y el tamaño del tumor para calcular la puntuación ROR.

Probabilidad de recurrencia distante a 10 años

Las puntuaciones ROR para una cohorte de mujeres posmenopáusicas con cáncer de mama en estadio temprano con receptores hormonales positivos se compararon con la supervivencia sin recurrencia distante después de la cirugía y el tratamiento con 5 años de terapia endocrina adyuvante seguido de 5 años de observación. Este estudio dio como resultado un modelo que relaciona la puntuación ROR con la probabilidad de recurrencia distante en esta población de pacientes probados, incluyendo un intervalo de confianza del 95%.

Clasificación de riesgo

La clasificación de riesgos se proporciona además para permitir la interpretación de la puntuación ROR mediante el uso de puntos de corte relacionados con el resultado clínico en las poblaciones de pacientes probados.

5

Clasificación de riesgo por intervalo ROR y estado nodular

Estado nodular	Intervalo ROR	Clasificación de riesgo
Nódulo- Negativo	0-40	Inferior
	41-60	Producto intermedio
	61-100	Superior
Nódulo Positivo (1-3 nódulos)	0-15	Inferior
	16-40	Producto intermedio
	41-100	Superior

10

15

20

Control de calidad

Cada lote de los componentes del ensayo se prueba con el uso de especificaciones predeterminadas. Todos los componentes a nivel de kit se siguen por lote, y los componentes críticos contenidos en cada kit se prueban juntos y se liberan como un lote maestro.

25

El kit de ensayo incluye una serie de controles internos que se usan para evaluar la calidad de cada conjunto de análisis en su conjunto y cada muestra individualmente. Estos controles se enumeran más abajo.

30

Conjunto de control de lote: *Muestra de referencia de ARN transcrita in vitro*

Se incluye una muestra de referencia de ARN sintético como control dentro del kit de ensayo. La muestra de referencia está compuesta por dianas de ARN transcritas in vitro a partir del algoritmo de 50 y 8 genes de mantenimiento. La muestra de referencia se procesa por duplicado en cada análisis junto con un conjunto de hasta 10 muestras de ARN de tumor de mama desconocidas en un tubo de 12 tiras de reacción. La señal de la muestra de referencia se analiza contra los umbrales predefinidos para calificar la ejecución.

35

La señal de cada uno de los 50 genes del algoritmo de la muestra de ARN de tumor de mama se normaliza con los genes correspondientes de la muestra de referencia.

40

Conjunto de control positivo: ARN dianas transcritos *in vitro* y las correspondientes sondas de captura y reporteras

Las dianas de ARN sintético se usan como controles positivos (PC) para el ensayo. Las secuencias diana de PC se derivan de la genoteca de secuencias de ADN del Consorcio de Control de ARN Externo (ERCC). Las dianas de ARN se transcriben in vitro a partir de los plásmidos de ADN. Se incluyen seis dianas de ARN dentro del kit de ensayo en una serie de titulación de 4 veces (128 - 0,125 fM de concentración final en la reacción de hibridación) junto con las correspondientes sondas de captura y reporteras. Las PC se añaden a cada muestra de ARN de tumor de mama y a la muestra de ARN de referencia probada con el ensayo prosigna. Una muestra será descalificada de un análisis posterior si las intensidades de señal de las PC no alcanzan los umbrales predefinidos.

45

50

Conjunto de control negativo: sondas exógenas sin diana

Las secuencias diana de control negativo (NC) se derivan de la genoteca de secuencias de ADN ERCC. Las sondas diseñadas para detectar estas secuencias diana se incluyen como parte del kit de ensayo sin la secuencia diana correspondiente. Los controles negativos (NC) se añaden a cada muestra de ARN de tumor de mama y a la muestra de ARN de referencia probada con el ensayo prosigna como una medida de control de calidad. La muestra será descalificada de un análisis posterior si las intensidades de señal de los NC no alcanzan los umbrales predefinidos.

55

Conjunto de control de integridad de ARN: Genes de mantenimiento

Las sondas de captura y reporteras diseñadas para detectar 8 genes de mantenimiento y 50 genes de algoritmo se incluyen como parte del kit. Los niveles de expresión de los 8 genes de mantenimiento se analizan para determinar la calidad del ARN extraído de la muestra de tejido FFPE y la entrada en el ensayo. La muestra será descalificada de un análisis posterior si el nivel de expresión de los genes de mantenimiento cae por debajo de los umbrales predefinidos.

60

65

Los genes de mantenimiento se usan además para normalizar cualquier diferencia en la cantidad de ARN intacto en una muestra antes de la normalización de la muestra de referencia.

Definiciones

5 Para los fines de la presente descripción, "cáncer de mama" incluye, por ejemplo, las afecciones clasificadas por biopsia o histología como patología maligna. La delineación clínica de los diagnósticos de cáncer de mama es bien conocida en las técnicas médicas. Un experto en la técnica apreciará que el cáncer de mama se refiere a cualquier malignidad del tejido mamario, incluidos, por ejemplo, carcinomas y sarcomas. Las modalidades particulares del cáncer de mama incluyen carcinoma ductal *in situ* (DCIS), carcinoma lobular *in situ* (LCIS), o carcinoma mucinoso. El cáncer de mama se refiere además al carcinoma ductal infiltrante (IDC) o lobular infiltrante (ILC). En la mayoría de las modalidades de la descripción, el sujeto de interés es un paciente humano sospechoso o realmente diagnosticado con cáncer de mama.

Los artículos "un" y "una" se usan en la presente descripción para referirse a uno o más de uno (*es decir*, para al menos uno) del objeto gramatical del artículo. A manera de ejemplo, "un elemento" significa un elemento o más de un elemento.

15 A través de la descripción la palabra "comprender", o variaciones tales como "comprende" o "que comprende", se entiende que implica la inclusión de un elemento declarado, entero o etapa, o grupo de elementos, enteros o etapas, pero no la exclusión de ningún otro elemento, entero o etapa, o grupo de elementos, enteros o etapas.

EJEMPLOS

20 Ejemplo 1. Prueba de subtipaje NAN046

La Figura 5 describe los procesos de ensayo asociados con la prueba de subtipaje intrínseco de cáncer de mama. Después del aislamiento de ARN, la prueba medirá simultáneamente los niveles de expresión de 46 genes diana más ocho genes de mantenimiento en una única reacción de hibridación con el uso de un nCounter CodeSet diseñado específicamente para los genes. Por ejemplo, los genes de mantenimiento descritos en la Publicación de Patente de los Estados Unidos 2008/0032293 puede usarse para la normalización. Los genes de mantenimiento ilustrativos incluyen MRPL19, PSMC4, SF3A1, PUM1, ACTB, GAPD, GUSB, RPLP0, y TFRC. Los genes de mantenimiento se usan para normalizar la expresión de la muestra de tumor. Cada prueba incluye además una muestra de referencia que consiste en ARN transcritos *in vitro* de los 58 objetivos con fines de normalización.

Revisión/adquisición de tejidos FFPE y extracción de ARN: La prueba de subtipaje intrínseco de cáncer de mama usará ARN extraído del tejido fijado con formalina e incluido en parafina (FFPE) que se ha diagnosticado como carcinoma invasivo de mama. Un patólogo revisa un portaobjetos teñido con H & E para identificar el área de tejido que contiene suficiente contenido de tejido tumoral para la prueba. Las secciones de tejido montadas en los portaobjetos sin teñir se procesan mediante macro-disección del área tumoral identificada en cada portaobjetos para eliminar cualquier tejido normal adyacente. Después se aísla el ARN del tejido tumoral y se elimina el ADN de la muestra.

Configuración del ensayo e inicio de la hibridación: Para cada lote de hasta 10 muestras de ARN aisladas de un tumor de mama, el usuario configurará una ejecución con el uso del programa del sistema nCounter Analysis x5, que rastrea el procesamiento de muestras, los lotes de reactivos y resultados para cada muestra. Para iniciar el ensayo, el usuario pipeteará la cantidad especificada de ARN en tubos separados dentro de un tubo de 12 tiras de reacción y agregará el CodeSet y el tampón de hibridación. Se pipetea una muestra de referencia en los dos tubos restantes con CodeSet y el tampón de hibridación. El CodeSet consta de sondas para cada gen que se especifica, sondas adicionales para la normalización endógena de los genes de "mantenimiento" y controles positivos y negativos que se agregan al ensayo. La muestra de referencia consta del ARN transcrito *in vitro* para los genes diana y los genes de mantenimiento. Una vez que los reactivos de hibridación se agregan a los tubos respectivos, el usuario transfiere el tubo con la tira a un bloque térmico de tapa calentada durante un período de tiempo específico a una temperatura establecida.

Purificación y unión en la estación de preparación: Al completar la hibridación, el usuario transferirá el tubo con la tira que contiene el conjunto de 10 ensayos y 2 muestras de referencia a la estación de preparación nCounter junto con los reactivos y desechables preempaquetados requeridos. Después, un proceso de purificación automatizado se elimina el exceso de sonda de captura y reportera a través de dos etapas sucesivas de captura por perlas magnéticas impulsadas por hibridación. La estación de preparación nCounter transfiere después los complejos diana/sonda purificados a un cartucho nCounter para capturarlos en un portaobjetos de vidrio. Después de completar la ejecución, el usuario retira el cartucho de la estación de preparación y lo sella con una película adhesiva.

Imágenes y análisis en el analizador digital: El cartucho sellado se inserta en el analizador digital nCounter que cuenta el número de sondas capturadas en el portaobjetos para cada gen, que corresponde a la cantidad diana en solución. Después, el programa automatizado verificará los umbrales para los genes de mantenimiento, la muestra de referencia y los controles positivos y negativos para calificar cada ensayo y garantizar que el procedimiento se realizó correctamente. Las señales de cada muestra se normalizan después con el uso de los genes de mantenimiento para controlar la calidad de la muestra de entrada. Después, las señales se normalizan con la muestra de referencia dentro de cada ejecución para controlar las variaciones de ejecución a ejecución. Los datos normalizados resultantes se ingresan en el algoritmo de subtipaje intrínseco de cáncer de mama para determinar el subtipo intrínseco del tumor y la puntuación de riesgo de recurrencia.

Ejemplo 2: Validación clínica de la puntuación de riesgo de recurrencia (ROR) de NANO46 para predecir el riesgo residual de la recurrencia a distancia (DR) después de la terapia endocrina en mujeres posmenopáusicas con cáncer de mama temprano HR+ (EBC): Un estudio ABCSG8.

5

El objetivo del estudio es evaluar el rendimiento de la puntuación ROR en la predicción de la recurrencia distal de pacientes posmenopáusicas con cáncer de mama temprano con receptores hormonales positivos (HR+ EBC) tratados con tamoxifeno o tamoxifeno seguido de anastrozol cuando la prueba NANO46 se realiza en un laboratorio de patología del hospital de rutina. ¿La puntuación ROR agrega información pronóstica (RFS distante) más allá de la puntuación de tratamiento clínico en todos los pacientes (CTS incluye: nódulos, grado, tamaño del tumor, edad, tratamiento)? ¿Los grupos de riesgo basados en ROR en la información pronóstica (RFS distante) superan la puntuación de tratamiento clínico en todos los pacientes?

10

Resumen del estudio: 3.714 pacientes se excluyeron en un ABCSG8. Las pacientes eran mujeres posmenopáusicas con HR+ EBC (nódulo negativo y nota positiva), grado uno o dos, sin tratamiento previo. 1.671 pacientes volvieron a dar su consentimiento para un seguimiento a largo plazo o fallecieron. La mediana de seguimiento fue de 11 años. Se recogieron 1.620 bloques de FFPE. 25 tuvieron cáncer insuficiente en la revisión del bloque por el patólogo, 73 tuvieron ARN insuficiente incluido, 44 fallaron a las especificaciones de control de calidad para el dispositivo NanoString. 1.478 pacientes (91,2%) pasaron el análisis NANO46.

15

20

Métodos: Se enviaron tres secciones sin teñir de 10 micras y 1 portaobjetos de H&E para cada paciente a un laboratorio de patología académico independiente a BCCA donde se realizaron la revisión de tejidos, microdissección manual y extracción de ARN. El análisis NANO46 se realizó después en 250 ng del ARN extraído con el uso del Sistema de Análisis NanoString nCounter; se calcularon tanto el subtipo intrínseco como la puntuación ROR.

25

Resultados: La puntuación ROR añade información pronóstica estadísticamente significativa (RFS distante) más allá de CTS en todos los pacientes (prueba de razón de probabilidad $\Delta LR\chi^2 = 53,5$, $p < 0,0001$). Los grupos de riesgo basados en ROR añaden información pronóstica estadísticamente significativa (RFS distante) más allá de CTS en todos los pacientes (prueba de razón de probabilidad $\Delta LR\chi^2 = 34,1$, $p < 0,0001$). La diferenciación entre Luminal A y Luminal B añade información pronóstica estadísticamente significativa (RFS distante) más allá de CTS en todos los pacientes (Luminal B frente A: HR=2,38, IC 95%; 1,69-3,35, $p < 0,0001$). Los resultados en los subgrupos de nódulos negativos y nódulos positivos son similares a los resultados para todos los pacientes que se reportan en el estudio.

30

Conclusiones: Los resultados muestran que tanto la puntuación ROR como los grupos de riesgo basados en ROR añaden información pronóstica estadísticamente significativa más allá de la puntuación de tratamiento clínico. Los resultados demuestran que puede realizarse una prueba compleja de expresión de múltiples genes en el laboratorio de patología de un hospital y cumplir con las mismas métricas de calidad que un laboratorio central de referencia. Los resultados de los estudios TransATAC y ABCSG8 juntos proporcionan evidencia de Nivel 1 de la validez clínica de la prueba NANO46 para predecir el riesgo de recurrencia distante en mujeres posmenopáusicas con HR+ EBC tratadas solo con la terapia endocrina. Los resultados además muestran que los subtipos Luminal A tienen mejores resultados que los subtipos Luminal B en mujeres posmenopáusicas con HR+ EBC tratadas solo con la terapia endocrina.

35

40

REIVINDICACIONES

1. Un método para determinar el riesgo de recurrencia bajo o alto (ROR) en un sujeto que tiene cáncer de mama que comprende:
 - 5 determinar el tamaño del tumor de cáncer de mama del sujeto;
 - determinar un valor de correlación de un tumor de cáncer de mama para cada uno de al menos cuatro subtipos intrínsecos, que incluyen un tipo Basal, Luminal A, Luminal B o HER-2 enriquecido, midiendo la expresión de ARN de al menos 40 de los genes en la lista de genes intrínsecos NANO46 de la Tabla 1;
 - 10 determinar una puntuación de proliferación midiendo la expresión de ARN de cada uno de los subconjuntos de 18 genes de los genes intrínsecos NANO46 seleccionados de ANLN, CCNE1, CDC20, CDC6, CDCA1, CENPF, CEP55, EXO1, KIF2C, KNTC2, MELK, MKI67, ORC6L, PTTG1, RRM2, TYMS, UBE2C y UBE2T;
 - calcular una puntuación de riesgo de recurrencia (ROR) con el uso de la siguiente ecuación: $ROR-PT = -0,0067 \cdot Basal + 0,4317 \cdot Her2 + -0,3172 \cdot LumA + 0,4894 \cdot LumB + 0,1981 \cdot Puntuación\ de\ proliferación + 0,1133 \cdot Tamaño\ de\ tumor$; en donde la puntuación de riesgo de recurrencia es un valor entero en una escala de 0-100, en donde el nivel de expresión de ARN para al menos 40 de los genes en la lista de genes intrínsecos NANO46 de la Tabla 1 se muestran en la Tabla 3;
 - 15 determinar de ese modo si el sujeto tiene un riesgo de recurrencia bajo o alto en función de la puntuación de riesgo de recurrencia (ROR) y la interpretación de la puntuación ROR con el uso de la clasificación de riesgo mediante el intervalo de riesgo de recurrencia (ROR) y estado nodular.
- 20 2. El método de conformidad con la reivindicación 1, que comprende además determinar al menos uno de los siguientes: grado del tumor, ploidía del tumor, expresión del receptor de estrógenos, expresión del receptor de progesterona, y expresión de HER2/ERBB2.
- 25 3. El método de conformidad con la reivindicación 1, que comprende además determinar cada uno de los siguientes: grado del tumor, ploidía del tumor, expresión del receptor de estrógenos, expresión del receptor de progesterona, y expresión de HER2/ERBB2.
- 30 4. El método de conformidad con la reivindicación 1, en donde el riesgo es la supervivencia específica del cáncer de mama, la supervivencia libre de eventos, o la respuesta a la terapia.
- 35 5. El método de conformidad con la reivindicación 1, en donde la expresión de ARN de los miembros de la lista de genes intrínsecos NANO46 se determina con el uso del sistema de código de nanoreporteros (sistema de análisis nCounter®).

Figura 1

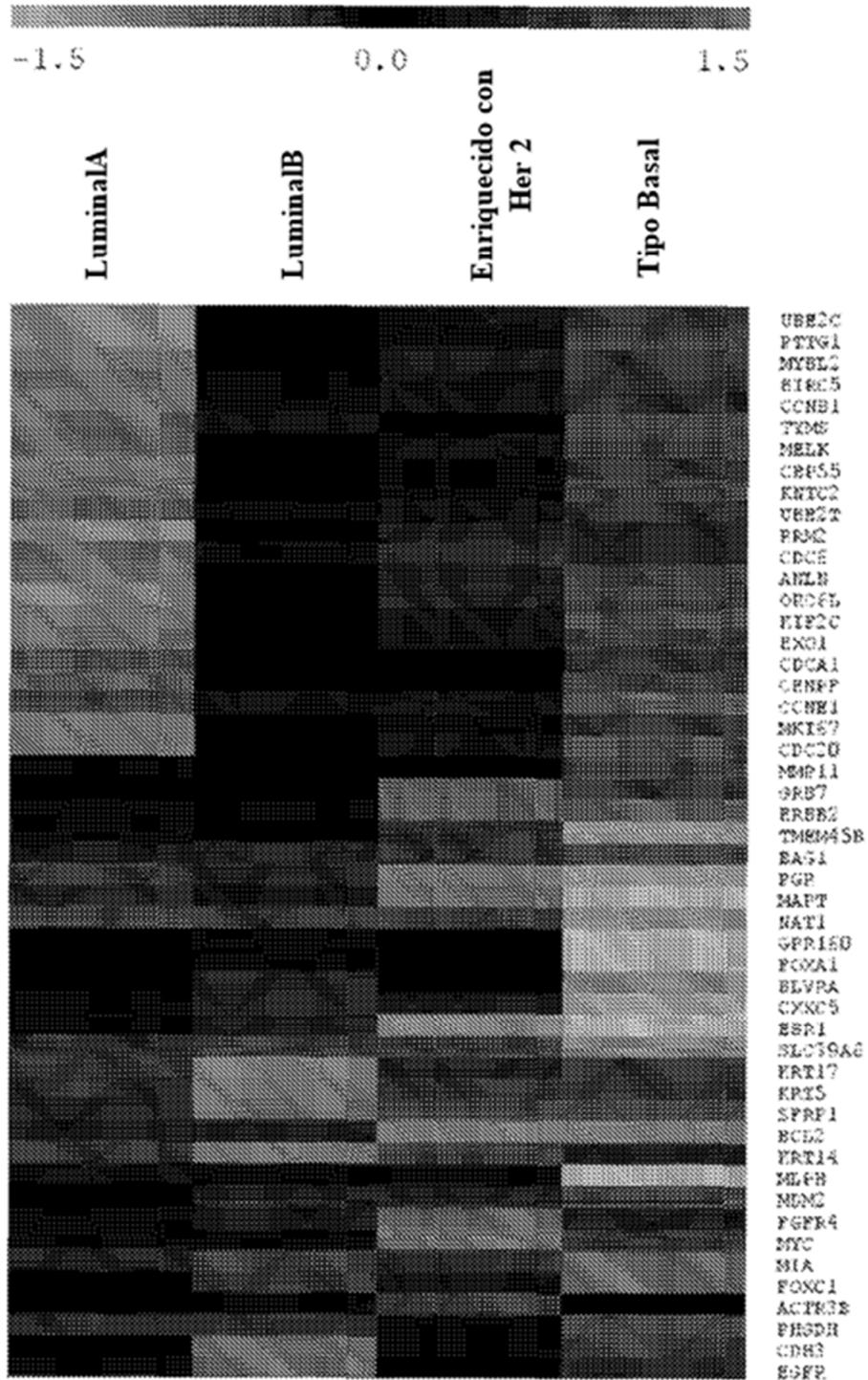


Figura 2

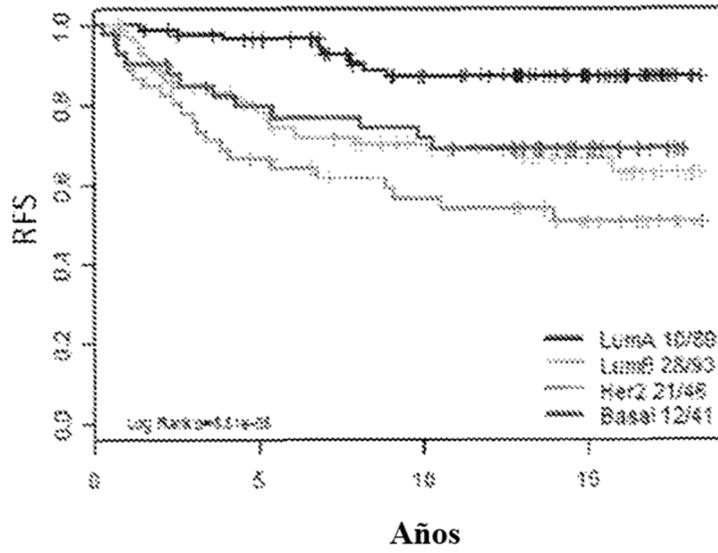


Figura 3

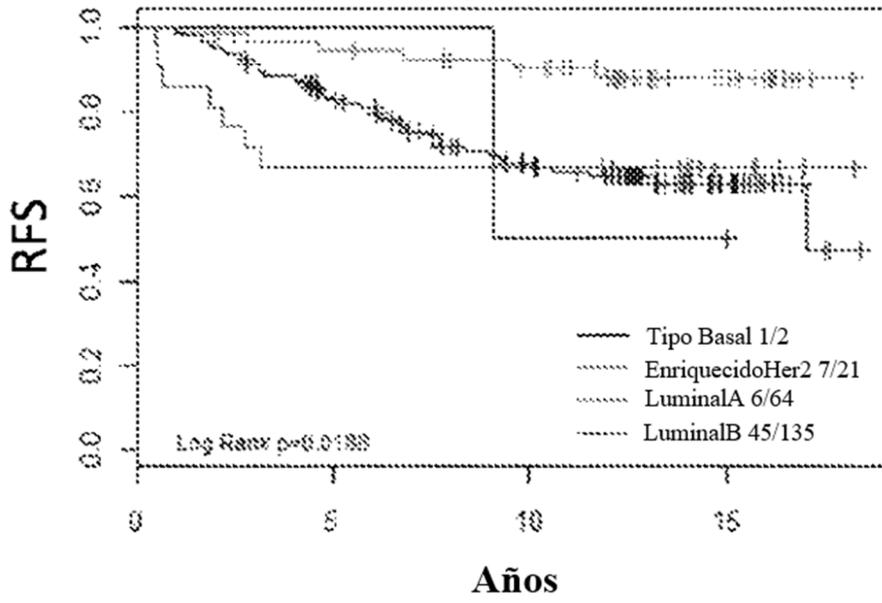


Figura 4

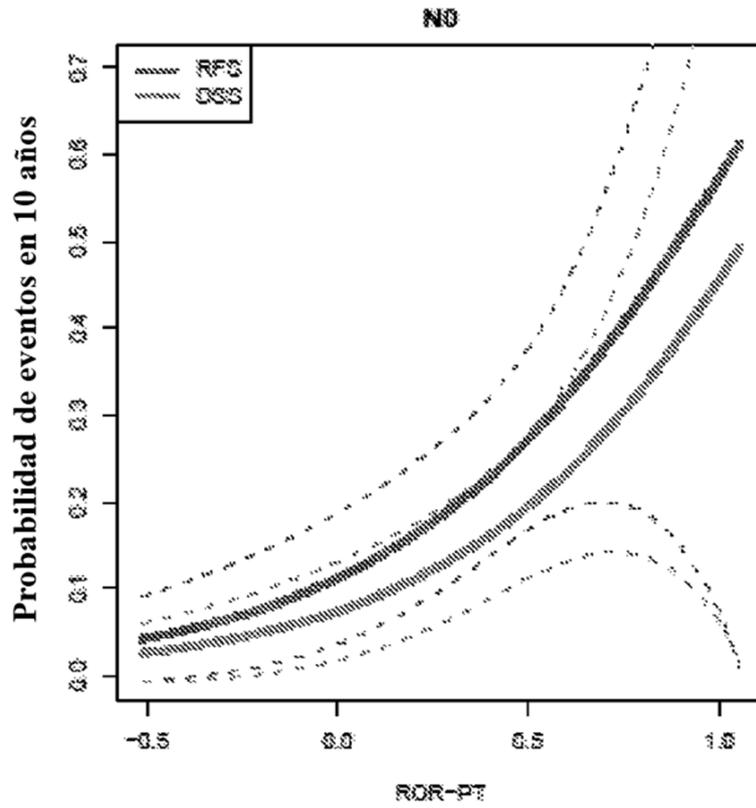


Figura 5

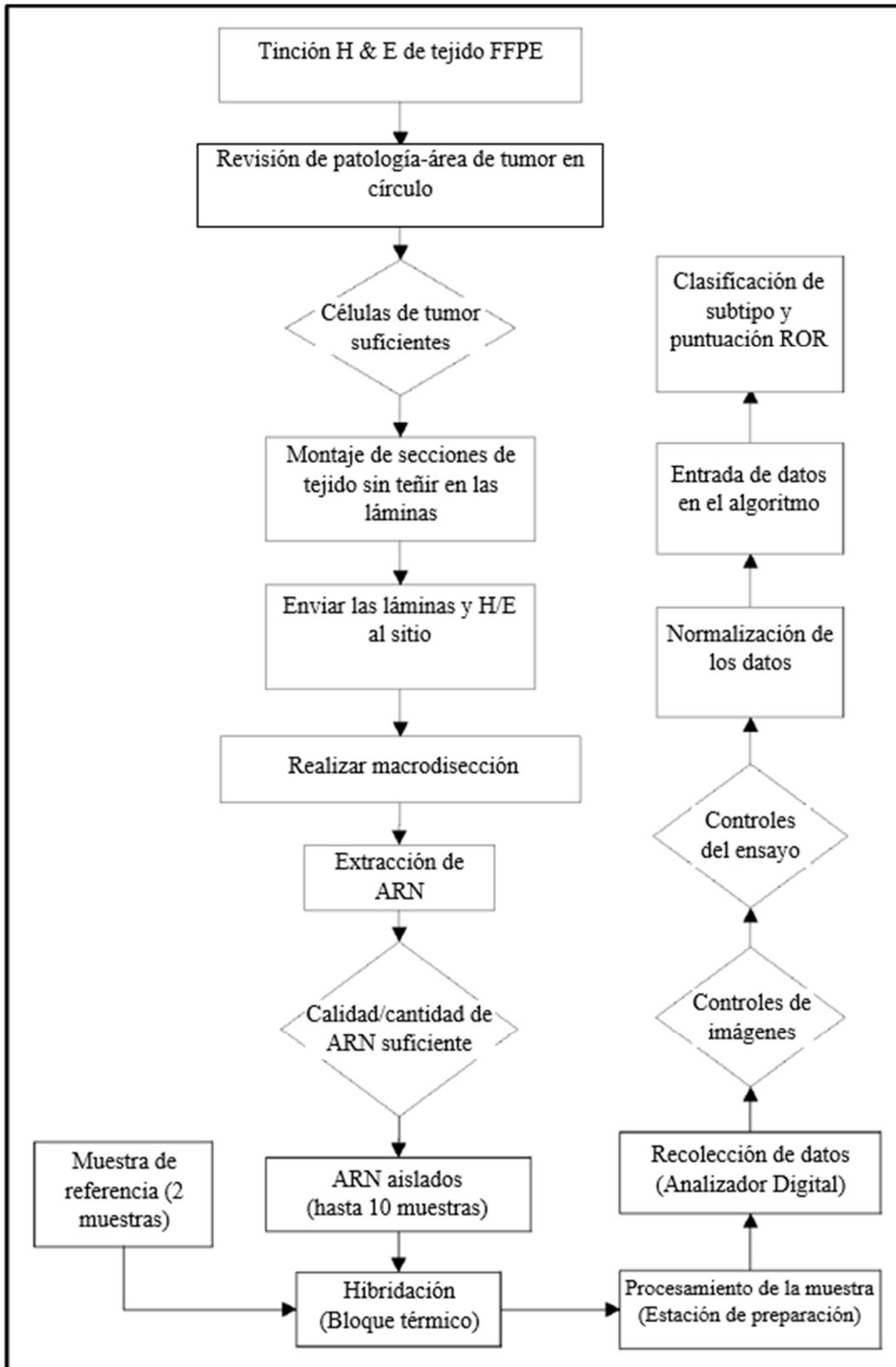


Figura 6

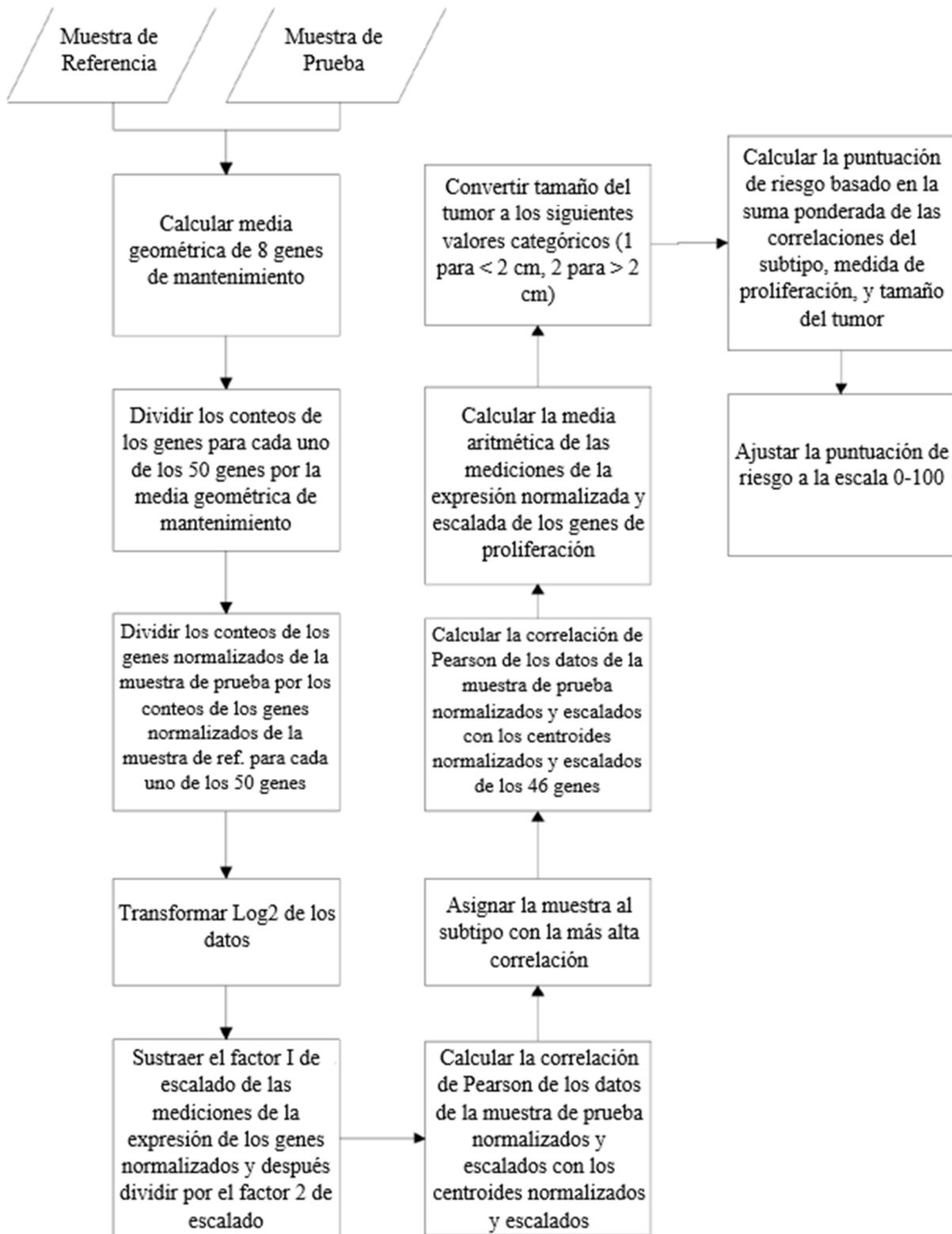


Figura 7

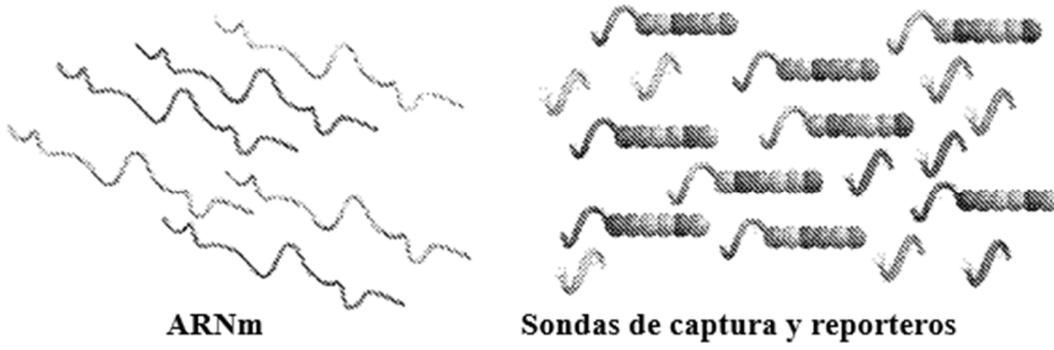


Figura 8

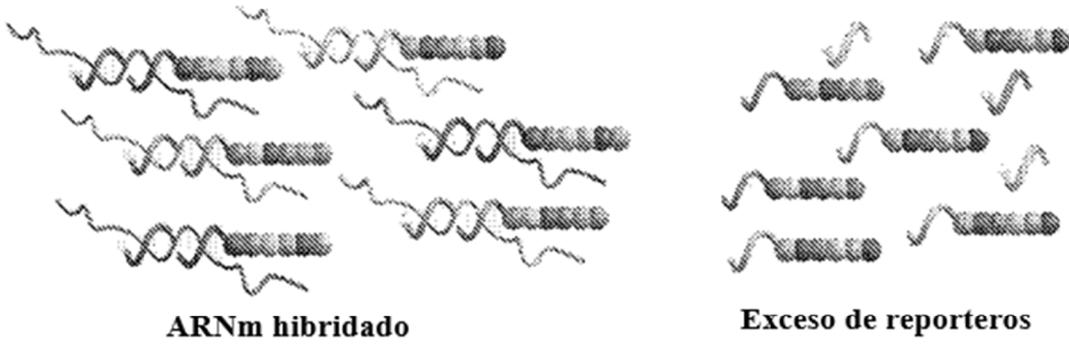


Figura 9

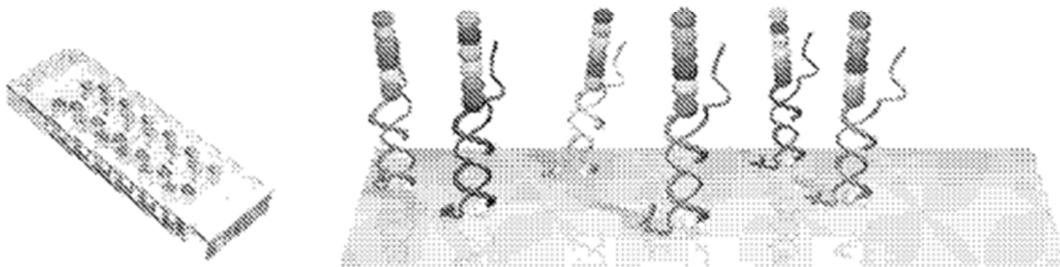


Figura 10



Figura 11

Código	Gen	Cantidad
●●●●● ●●●●● ●●●●●	x	3
●●●●●	y	1
●●●●● ●●●●●	z	2

Figura 12

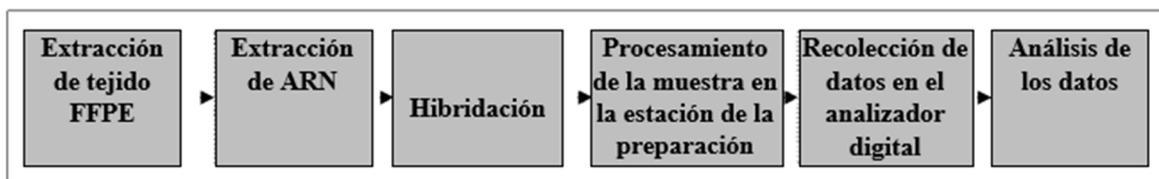


Figura 13

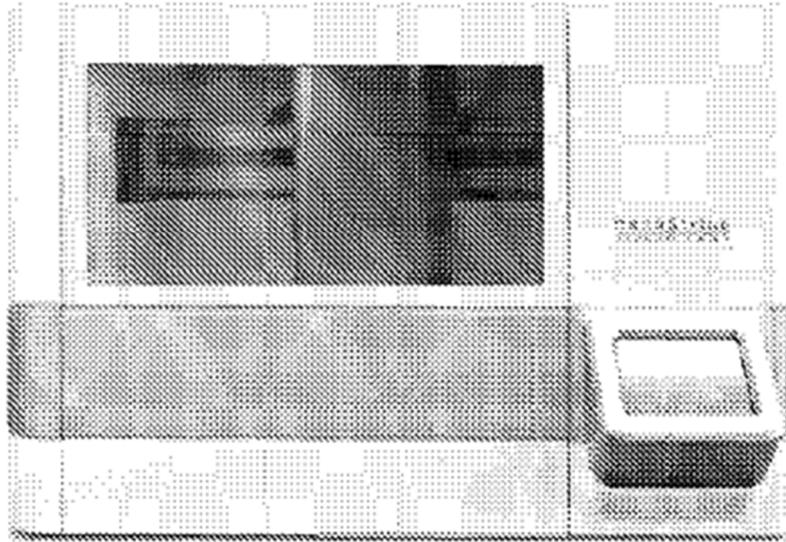


Figura 14

