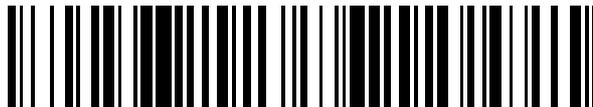


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 475 416**

51 Int. Cl.:

**C12N 15/12** (2006.01)

**C12N 15/10** (2006.01)

**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.02.1999 E 10011439 (6)**

97 Fecha y número de publicación de la concesión europea: **09.04.2014 EP 2287307**

54 Título: **Materiales y métodos para identificar y analizar marcadores de ADN de repeticiones intermedias en tándem**

30 Prioridad:

**04.02.1998 US 18584**

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

**10.07.2014**

73 Titular/es:

**PROMEGA CORPORATION (100.0%)  
2800 Woods Hollow Road  
Madison, WI 53711-5399, US**

72 Inventor/es:

**SCHUMM, JAMES W. y  
BACHER, JEFFREY W.**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

**ES 2 475 416 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

Materiales y métodos para identificar y analizar marcadores de ADN de repeticiones intermedias en tándem.

### Campo de la invención

En la presente memoria se describe la identificación y el análisis de marcadores genéticos en un sistema genómico. De manera más específica, en la presente memoria se describe la identificación de loci en el ADN, en particular en el ADN genómico, que contienen polimorfismos de longitud debido a variaciones en el número de repeticiones de secuencias intermedias (5 a 7 bases). En la presente memoria se describe la detección de tales loci polimórficos. En la presente memoria se describen métodos para identificar y distinguir individuos basándose principalmente en las diferencias de tamaños de los productos de amplificación del ADN genómico en tal locus, en los que el número de secuencias de repeticiones intermedias en tándem varían de un individuo a otro.

### Antecedentes de la invención

La tipificación del ADN se utiliza habitualmente para identificar la ascendencia de niños humanos, y para confirmar el linaje de caballos, perros, y otros animales de competición. La tipificación del ADN también se emplea habitualmente para identificar el origen de la sangre, saliva, semen y otros tejidos hallados en la escena de un crimen. Los métodos de tipificación del ADN actualmente en uso están diseñados para detectar y analizar diferencias de longitud y/o de secuencia de una o más regiones del ADN que se sabe que aparecen en al menos dos formas diferentes en una población. La tipificación del ADN se emplea también en el ámbito clínico para determinar el éxito o el fracaso del trasplante de médula ósea y la presencia de tejidos cancerosos particulares. Tal variación de longitud y/o secuencia se denomina "polimorfismo". Cualquier región (es decir, "locus") de ADN en la que se da tal variación se denomina "locus polimórfico". La mayoría de las técnicas de tipificación del ADN emplean al menos un "marcador" que contiene al menos dicho locus polimórfico. Cada marcador individual contiene un único alelo de ADN genómico derivado en última instancia de un único individuo de una población. Los métodos y materiales de la presente invención están diseñados para el uso en la detección de una clase particular de polimorfismos en el ADN caracterizados principalmente por la variación de la longitud.

Durante mucho tiempo se han buscado marcadores genéticos que sean suficientemente polimórficos con respecto a la longitud o la secuencia para el uso en las aplicaciones de identificación, tales como pruebas de paternidad e identificación de muestras de tejidos recogidas para el análisis forense. El descubrimiento y el desarrollo de tales marcadores y métodos para analizar tales marcadores ha atravesado varias fases de desarrollo a lo largo de los últimos años. En años recientes, el descubrimiento y el desarrollo de las repeticiones cortas en tándem (STRs) polimórficas como marcadores genéticos ha estimulado el progreso en el desarrollo de la cartografía de ligamiento, la identificación y la caracterización de genes afectados por enfermedades, y la simplificación y precisión de la tipificación del ADN. La expresión "repetición corta en tándem" o "STR", tal como se usa en la presente memoria, se refiere a todas las secuencias con una longitud de entre dos y siete nucleótidos que están repetidas perfectamente, o casi perfectamente en tándem en el ADN genómico de cualquier organismo. Véase, por ejemplo, la definición de "repetición corta en tándem" aplicada al ADN genómico humano en la pat. de EE.UU. N° 5.364.759, columna 4, línea 58 y sigs.

Los primeros marcadores de variantes del ADN identificados fueron sustituciones de bases simples, es decir, polimorfismos de secuencias simples, que se detectaban la mayor parte de las veces mediante ensayos de hibridación de Southern. Para ejemplos de referencias que describen la identificación de tales marcadores, diseñados para ser usados para analizar el ADN digerido con endonucleasas de restricción con sondas radioactivas, véase: Southern, E. M. (1975), *J. Mol. Biol.* 98(3):503-507; Schumm, et al. (1988), *American Journal of Human Genetics* 42:143-159; y Wyman, A. y White, R. (1980) *Proc. Natl. Acad. Sci., U.S.A.* 77:6754-6758.

La siguiente generación de marcadores fueron variantes de tamaño, es decir, polimorfismos de longitud, específicamente marcadores de "repeticiones en tándem de número variable" (VNTR) (Nakamura Y., et al. (1987), *Science* 235: 1616-1622; y Pat. de EE.UU. N° 4.963.663 expedida a White et al. (1990); Pat. de EE.UU. N° 5.411.859, continuación del documento 4.963.663 expedido a White et al. (1995)) y los marcadores de "minisatélites" (Jeffreys et al. (1985a), *Nature* 314:67-73; Jeffreys et al. (1985b) *Nature* 316:76-79., Pat. de EE.UU. N° 5.175.082 para una invención de Jeffreys). Las VNTR y los marcadores de minisatélites contienen regiones de secuencias casi idénticas repetidas en tándem. La secuencia de repetición central tiene una longitud de 10 a 70 bases, y las secuencias de repeticiones centrales más cortas se denominan repeticiones "minisatélites" y las repeticiones más largas se denominan VNTRs. Los diferentes individuos de una población humana contienen números diferentes de estas repeticiones. Estos marcadores son más polimórficos que los polimorfismos por sustitución de bases, y a veces exhiben hasta cuarenta o más alelos en un único locus genético. Sin embargo, el proceso tedioso de digestión con enzimas de restricción y el análisis de hibridación de Southern posterior todavía son necesarios para detectar y analizar la mayoría de tales marcadores.

El siguiente avance implicó la unión de la tecnología de la reacción en cadena de la polimerasa (PCR) (Pat. de EE.UU. N° 4.683.202 de Mullis, K.B.) con los análisis de loci de VNTR (Kasai K, et al. (1990) *Journal Forensic Science* 35(5):1196-1200). Se descubrieron loci de VNTR amplificables, que se podrían detectar sin necesidad de la

transferencia de Southern. Los productos amplificados se separan por medio de geles de agarosa o poliacrilamida y se detectan mediante la incorporación de radiactividad durante la amplificación o mediante tinción posterior con plata o bromuro de etidio. Sin embargo, la PCR se puede usar solamente para amplificar de manera fiable segmentos de ADN relativamente pequeños, es decir, para amplificar de manera fiable solamente segmentos de ADN con una longitud inferior a 3.000 bases, Ponce, M. y Micol, L. (1992) *NAR* 20(3):623; Decorte R, et al. (1990) *DNA Cell Biol.* 9(6):461-469). Por lo tanto, se han desarrollado muy pocas VNTRs amplificables, lo que las hace, como clase, poco prácticas para la cartografía de ligamiento.

Con el desarrollo reciente de marcadores polimórficos con repeticiones polimórficas de dinucleótidos (Litt y Luty (1989) *Am J. Hum Genet* 3(4):599-605; Tautz, D (1989) *NAR* 17:6463-6471; Weber y May (1989) *Am J Hum Genet* 44:388-396; Pat. alemana N° DE 38 34 636 C2, inventor Tautz, D; Pat. de EE.UU. N° 5.582.979 presentada por Weber, L.) y con repeticiones cortas en tándem polimórficas (STR) (Edwards, A., et al. (1991) *Am. J. Hum. Genet.* 49: 746-756.; Hammond, H.A., et al. (1994) *Am. J. Hum. Genet.* 55: 175-189; Fregeau, C.J.; y Fournay, R.M. (1993) *BioTechniques* 15(1): 100-119.; Schumm, J.W. et al. (1994) en *The Fourth International Symposium on Human Identification 1993*, págs. 177-187; y la Pat. de EE.UU. N° 5.364.759 de Caskey et al.; Pat. alemana N° DE 38 34 636 C2 de Tautz, D.) se han superado muchas de las deficiencias de los métodos previos. Los dos tipos de marcadores, los que contienen repeticiones de dinucleótidos o STR (que por definición incluyen repeticiones de 2-7 pb), se denominan en general marcadores de "microsatélites". A menudo considerados como los mejores marcadores disponibles, los loci de los microsatélites son similares a las VNTRs amplificables, ya que sus alelos se pueden diferenciar basándose en la variación de la longitud. Sin embargo, a diferencia de las VNTRs, estos loci contienen secuencias de repeticiones perfectas o imperfectas de una longitud de dos, tres, cuatro, o raramente, cinco bases. Exhiben desde solamente unos cuantos alelos a más de cuarenta en un único locus. Se pueden diseñar protocolos de amplificación para producir productos pequeños, en general de una longitud de 60 a 400 pares de bases, y los alelos de cada locus a menudo están contenidos dentro de un intervalo de menos de 50 pb. Esto permite el análisis electroforético simultáneo de varios sistemas en el mismo gel mediante el diseño cuidadoso de los cebadores de PCR de forma que todos los productos de la amplificación potencial a partir de un sistema individual no solapan con el intervalo de alelos de otros sistemas en el mismo gel.

Existen tres desventajas significativas relacionadas con el uso de los loci de microsatélites. En primer lugar, a menudo se observa tras la amplificación la presencia de artefactos por tartamudeo, es decir, uno o más fragmentos secundarios además del fragmento principal que representa cada alelo. Esta deficiencia se manifiesta de manera mucho más grave con los loci de repeticiones de dinucleótidos que con los marcadores de repeticiones de tri- o tetranucleótidos (Edwards et al., 1991. *Am J Hum Genet* 49:746-756; Edwards et al., 1992. *Genomics* 12:241-253; Weber y May, 1989. *Am J Hum Genet* 44:388-396). La presencia de estos artefactos, que se supone que son el resultado de un fenómeno relacionado con las ADN polimerasas denominado deslizamiento (Levinson y Gutman, 1987. *Mol. Biol. Evol.* 4(3):203-221; Schlotterer y Tautz, 1992. *NAR* 20:211-215), complica la interpretación del contenido alélico de los loci. A la vez que complica todas las interpretaciones, la presencia de fragmentos principales y secundarios para representar cada alelo limita especialmente la utilidad de estos marcadores en el análisis forense que a menudo requiere la determinación de si hay presente más de una fuente de ADN en la muestra. Muchos de los marcadores descritos en este trabajo representan una clase nueva de marcadores que producen significativamente menos artefacto por tartamudeo que los marcadores conocidos.

Una segunda desventaja de los sistemas de marcadores de STR y microsatélites actuales está relacionada con la dificultad de separar múltiples loci en un único gel. Esto ocurre debido a que existe una comprensión espacial de los fragmentos de tamaños diferentes en las regiones superiores de los geles usados más habitualmente para la separación de los fragmentos de ADN por los expertos en la técnica. El desarrollo de los marcadores descritos en este trabajo, basados en unidades de repetición mayores, amplía el intervalo útil en estos geles, lo que permite el análisis simultáneo de más loci.

Una tercera desventaja es que, antes de la invención descrita en la presente memoria, se habían descrito solamente unos cuantos loci de ADN del ADN genómico humano en la bibliografía con los polimorfismos de longitud basados en variaciones de un número de cinco a siete repeticiones de bases en cada locus. Véase, p.ej., Edwards et al. (1991) *Nucleic Acids Res.* 19:4791; Chen et al. (1993) *Genomics* 15(3): 621-5; Harada et al. (1994) *Am. J. Hum. Genet.* 55: 175-189; Comings et al. (1995), *Genomics* 29(2):390-6; y Utah Marker Development Group (1995), *Am. J. Genet.* 57:619-628. En 1995, Jurka y Pethiyagoda publicaron un artículo que describía un estudio en el que habían usado la base de datos GenBank para determinar la abundancia relativa y la variabilidad de repeticiones en tándem pentaméricas y hexaméricas en el genoma de primates (Jurka y Pethiyagoda (1995) *J. Mol. Evol.* 40:120-126). Sin embargo, la variabilidad se estimó solamente de manera indirecta, y no se demostraron los niveles de polimorfismo en loci individuales. *Id.* Se han desarrollado materiales y métodos para identificar y analizar loci de ADN que contienen repeticiones sumamente polimórficas de cinco a siete bases.

El documento WO95/17522 describe un método para la identificación de ADN de un fragmento que comprende un locus de repeticiones simples en tándem que comprende las etapas de: i) poner en contacto una biblioteca de ADN con al menos una sonda de hibridación para identificar una población de fragmentos de ADN enriquecida en repeticiones simples en tándem; ii) aislar y clonar dicha población; y iii) cribar la biblioteca de ADN resultante para identificar un fragmento individual que comprende un locus de repeticiones simples en tándem. También se discuten

las repeticiones simples en tándem aisladas mediante el método, caracterizadas porque se pueden amplificar al menos en parte mediante PCR con el uso de un par de cebadores especificados, junto con cebadores de amplificación y sondas específicas para las repeticiones simples en tándem así aisladas. También se discuten métodos de caracterización genética mediante el uso de las repeticiones simples en tándem, los cebadores y las sondas.

El documento US5364759 describe un ensayo de identificación de ADN para detectar polimorfismos en una repetición corta en tándem. El método incluye las etapas de extraer el ADN a partir de una muestra a ensayar, amplificar el ADN extraído e identificar los productos de extensión amplificados para cada secuencia diferente. Cada secuencia diferente se marca de manera diferencial. En el método, también se pueden usar patrones internos y externos. El método es aplicable a una amplia diversidad de muestras forenses y médicas, que incluyen sangre, semen, muestras vaginales, tejido, pelo, saliva, orina y mezclas de fluidos corporales.

Comings et al. 1995, Genomics, vol. 29, páginas 390-396 describe la secuenciación de las regiones reguladoras, intrónicas, y exónicas del gen TDO2 humano que se han secuenciado. Se identificaron doce exones. La secuencia de aminoácidos de la enzima fue un 88% homóloga a la de rata.

Bacher et al. 1998, Human identification symposium proceedings: The ninth international symposium on human identification, 8-10 de octubre de 1998, páginas 24-37, describe repeticiones de pentanucleótidos como marcadores genéticos sumamente polimórficos que exhiben un artefacto por tartamudeo mínimo.

El documento WO00/31306 describe métodos y materiales para el uso en la amplificación simultánea de al menos trece loci de ADN genómico en una única reacción multiplex, junto con métodos y materiales para el uso en el análisis de los productos de tales reacciones.

El documento WO97/30139 describe la amplificación simultánea de múltiples loci genéticos diferentes mediante el uso de PCR u otros sistemas de amplificación para determinar en una reacción multiplex los alelos de cada uno de los loci contenidos en la reacción multiplex. Los loci genéticos analizados comprenden los siguientes: HUMvWFA31, HUMLIPOL, HUMBFXIII, HUMF13A01, HUMFESFPS, HUMTHOI, HUMTPOX, HUMCSFIPO, D22S683, D20S481, D19S253, D17S1299, D17S1298, D16S753, D16S539, D16S490, D14S562, I4S548, D14S118, D13S317, DIOS1239, D9S930, D7S820, D5S818, D4S2368, D3S1539.

Mohan et al. 1992, Journal of Biomolecular Structure & Dynamics, Vol. 9(6), páginas 1195-1211 describe cálculos de mecánica molecular llevados a cabo en todas las etapas de diez pares de bases (dímeros de moléculas bicatenarias) y también varias moléculas bicatenarias triméricas y tetraméricas que los comprenden, en un intento de examinar sistemáticamente los posibles efectos de las secuencias de bases en las magnitudes de los giros de las hélices de los pares de bases en una etapa dada.

Los materiales y métodos del presente método están diseñados para el uso en la identificación y el análisis de loci polimórficos particulares de ADN de diversos tipos, que incluyen el ADN monocatenario y bicatenario de una diversidad de fuentes diferentes. La presente invención representa una mejora significativa sobre la tecnología existente, que proporciona una potencia y precisión incrementadas a la identificación de ADN para el análisis de ligamiento, la justicia penal, las pruebas de paternidad, y otros usos forenses y médicos.

### Sumario breve de la descripción

Por lo tanto, un objetivo es proporcionar materiales y métodos para la identificación y el análisis de loci de ADN con secuencias de repeticiones intermedias en tándem, en el que una "secuencia de repeticiones intermedias en tándem" es una región de ADN que contiene al menos una unidad de repetición que consiste en una secuencia de cinco (5), seis (6), o siete (7) bases repetidas en tándem al menos dos (2) veces.

Otro objetivo es proporcionar materiales y métodos para identificar marcadores de ADN de repeticiones intermedias en tándem, que producen menos artefactos cuando se usan para analizar o detectar uno o más loci de una muestra de ADN que contiene una repetición intermedia en tándem. Los métodos y materiales descritos en la presente memoria se usan preferiblemente para identificar y analizar loci de ADN genómico, cada uno de los cuales contiene una secuencia de repeticiones intermedias en tándem polimórficas. Los materiales incluyen cebadores oligonucleotídicos y marcadores de ADN para tales loci de ADN genómico humano. Los loci de repeticiones intermedias en tándem detectados mediante el uso de los métodos de la presente invención exhiben menos artefactos que los que exhiben muchos loci conocidos detectados mediante el uso de métodos similares, que incluyen las STRs cortas (es decir, repeticiones en tándem de una secuencia de ADN de dos, tres o cuatro bases).

Un objetivo particular es proporcionar un método y materiales para el análisis de loci genéticos polimórficos individuales basado principalmente en la variación de la longitud debida principalmente a diferencias en el número de unidades de repetición del ácido nucleico en una región de repeticiones intermedias en tándem del ácido nucleico. También es un objetivo específico proporcionar un método, un equipo, y cebadores para la detección y el análisis de loci polimórficos de ADN genómico, que contienen polimorfismos de repeticiones intermedias en tándem, que incluyen los polimorfismos de repeticiones en tándem de pentanucleótidos.

Una realización consiste en un método para aislar un fragmento de ADN que contiene una secuencia de repeticiones intermedias en tándem a partir de ADN genómico, que comprende: (a) proporcionar una diversidad de fragmentos de ADN en los que al menos un fragmento contiene una secuencia de repeticiones intermedias en tándem; (b) proporcionar un medio de soporte, p.ej. un medio de soporte estacionario, que tiene asociado a él al menos un oligonucleótido que comprende una secuencia de nucleótidos que es complementaria a una porción de la secuencia de repeticiones intermedias en tándem; y (c) combinar la diversidad de fragmentos de ADN con el medio de soporte en condiciones en las que el fragmento de ADN que contiene la secuencia de repeticiones intermedias y al menos otro fragmento de ADN se hibrida con el medio de soporte.

Una realización alternativa es un método para detectar una secuencia de repeticiones intermedias en tándem polimórficas que tiene una baja incidencia de artefactos por tartamudeo en el ADN genómico, que comprende: (a) proporcionar una muestra de ADN que tiene al menos una secuencia objetivo de repeticiones intermedias en tándem, y (b) detectar la secuencia objetivo de repeticiones intermedias en tándem en la muestra de ADN, en la que se observa un artefacto por tartamudeo medio como máximo de un 1,1 %.

Una realización adicional es un método para detectar una secuencia objetivo de repeticiones intermedias en tándem en una muestra de ADN mediante el uso de al menos un cebador oligonucleotídico para amplificar una secuencia de repeticiones intermedias en tándem de interés (más adelante en la presente memoria, la "secuencia objetivo de repeticiones intermedias en tándem") en el ADN de la muestra, en el que el cebador oligonucleotídico comprende una secuencia que es complementaria y que flanquea a una región de un marcador de ADN que contiene una secuencia de repeticiones intermedias en tándem (más adelante en la presente memoria, la "secuencia molde de repeticiones intermedias en tándem") en la secuencia del marcador de ADN, en la que el marcador de ADN tiene una secuencia seleccionada del grupo de secuencias que consiste en SEQ ID N<sup>os</sup>: 1 a 43.

En otra realización, se describe un equipo para la detección de al menos una secuencia objetivo de repeticiones intermedias en tándem en una muestra de ADN, y el equipo comprende un recipiente que tiene al menos un cebador oligonucleotídico para amplificar la al menos una secuencia objetivo de repeticiones intermedias en tándem, en el que el cebador oligonucleotídico comprende una secuencia de nucleótidos que es complementaria y que flanquea a una porción de una región de un marcador de ADN bicatenario que contiene una secuencia molde de repeticiones intermedias en tándem, en el que el marcador de ADN tiene una secuencia seleccionada del grupo que consiste en SEQ ID N<sup>os</sup>: 1 a 43.

En otra realización, se describe un cebador oligonucleotídico que comprende una secuencia complementaria a una cadena de un marcador de ADN bicatenario en una región del marcador que flanquea a una secuencia molde de repeticiones intermedias en tándem, en el que el marcador de ADN tiene una secuencia seleccionada del grupo que consiste en: SEQ ID N<sup>os</sup>: 1 a 6, y SEQ ID N<sup>os</sup>: 28 a 33.

Cada una de las diversas realizaciones tienen un uso específico en los campos de la identificación de seres humanos y otros organismos, el análisis forense, la determinación de la paternidad, la monitorización del trasplante de médula ósea, la cartografía de ligamientos, y la detección de enfermedades genéticas y del cáncer. La necesidad de distinguir con exactitud entre cantidades pequeñas de tejido de diferentes individuos es especialmente importante en las aplicaciones forenses, en las que muchas condenas (y exculpaciones) dependen del análisis de identificación de ADN, lo que incluye los análisis de los loci de STR.

Los objetivos, características, y ventajas adicionales serán evidentes a partir del siguiente mejor modo y de los dibujos ilustrativos.

**Breve descripción de los dibujos**

La FIG. 1 es un diagrama de flujo de un método de enriquecimiento de repeticiones intermedias en tándem mediante hibridación en filtro.

La FIG. 2 es un electroferograma de una repetición de pentanucleótidos S159.

La FIG. 3 es un electroferograma de una repetición de tetranucleótidos vWA.

La FIG. 4 es un electroferograma de una repetición de pentanucleótidos G210.

La FIG. 5 es un electroferograma de una repetición de tetranucleótidos D5S818.

La FIG. 6 es un diagrama de dispersión del % de tartamudeo de la repetición de pentanucleótidos S159.

La FIG. 7 es un diagrama de dispersión del % de tartamudeo de la repetición de pentanucleótidos G210.

La FIG. 8 es un diagrama de dispersión del % de tartamudeo de la repetición de tetranucleótidos D5S818.

La FIG. 9 es un diagrama de dispersión del % de tartamudeo de la repetición de tetranucleótidos vWA.

La FIG. 10 es una imagen imprimida por láser de los resultados del barrido mediante Fluorimager de fragmentos

amplificados marcados con fluorescencia de una repetición de pentanucleótidos S159, tras la separación mediante electroforesis en gel.

5 La FIG. 11 es una imagen imprimida por láser de los resultados del barrido mediante Fluorimager de fragmentos amplificados marcados con fluorescencia de una repetición de pentanucleótidos G210, tras la separación mediante electroforesis en gel.

Los dibujos y las figuras no están necesariamente a escala, y ciertas características pueden estar exageradas en la escala o se pueden mostrar de forma esquemática por motivos de claridad y concisión.

### Descripción detallada

#### A. Definiciones:

10 Tal como se usa en la presente memoria, la expresión "repetición intermedia en tándem" o "ITR" se refiere a una región de una secuencia de ADN que comprende una secuencia de cinco a siete bases repetida en tándem al menos dos veces. El término ITR abarca también una región de ADN en la que está repetida más de una única secuencia de cinco a siete bases en tándem o con bases interpuestas, con tal de que al menos una de las secuencias esté repetida al menos dos veces en tándem. Cada secuencia repetida al menos una vez dentro de una ITR se denomina en la presente memoria una "unidad de repetición".

Un "polimorfismo de ITR" se refiere a una ITR en el ADN genómico que varía de longitud de un cromosoma a otro en una población de individuos, debido principalmente a diferencias en el número de unidades de repetición en la misma región de cada cromosoma.

20 Las secuencias de repeticiones intermedias en tándem identificadas y analizadas se pueden dividir en dos categorías generales, perfectas e imperfectas. La expresión ITR "perfecta", tal como se usa en la presente memoria, se refiere a una región de ADN bicatenario que contiene una única unidad de repetición de cinco a siete bases repetida en tándem al menos dos veces, p.ej. (AAAAT)<sub>12</sub>. La expresión ITR "imperfecta", tal como se usa en la presente memoria, se refiere a una región de ADN que contiene al menos dos repeticiones en tándem de una unidad de repetición perfecta y al menos una repetición de una unidad de repetición imperfecta, en la que la unidad de repetición imperfecta consiste en una secuencia de ADN que podría dar como resultado una, dos, o tres inserciones, 25 deleciones, o sustituciones de bases en la secuencia de la unidad de repetición perfecta, p.ej. (AAAAT)<sub>12</sub>(AAAAAT)<sub>5</sub>AAT(AAATT)<sub>4</sub>. Cada secuencia de ITR imperfecta contiene al menos una secuencia de ITR perfecta. De manera específica, cada secuencia de ITR, ya sea perfecta o imperfecta, incluye al menos una secuencia de unidad de repetición que aparece al menos dos veces en tándem, una secuencia de unidad de repetición que se puede representar mediante la fórmula (I):

$$(A_w G_x T_y C_z)_n (I)$$

en la que A, G, T, y C representan los nucleótidos que pueden estar en cualquier orden; w, x, y y z representan el número de cada nucleótido en la secuencia, y oscilan de 0 a 7 y la suma de w + x + y + z oscila entre 5 y 7; y n representa el número de veces que la secuencia está repetida en tándem, y es al menos 2.

35 "Repetición en tándem de pentanucleótidos" se refiere a una subclase de los polimorfismos de "repeticiones intermedias en tándem" definidas anteriormente. A menos que se especifique de otra manera, la expresión "repetición en tándem de pentanucleótidos" abarca ITRs perfectas en las que la unidad de repetición es una secuencia de cinco bases, e ITRs imperfectas en las que al menos una unidad de repetición es una repetición de cinco bases.

40 "Marcador de ADN" se refiere a un fragmento de ADN que contiene una secuencia de ITR tal como un fragmento de ADN que contiene una secuencia de ITR producida mediante la amplificación de una región de ADN genómico. Cada marcador individual contiene un único alelo de ADN genómico derivado en última instancia de un único individuo de una población.

45 El término "locus" se refiere a una región específica del ADN. Cuando se usa para describir una región del ADN genómico, "locus" se refiere a una posición particular en un cromosoma. El mismo locus genómico aparece en sitios idénticos en cada par de cromosomas homólogos para cualquier individuo de una población. La secuencia de ADN en el mismo locus en dicho cromosoma, o en el mismo locus de ADN que se origina de dicho cromosoma, se denomina "alelo".

50 El término "polimorfismo", tal como se usa en la presente memoria, se refiere a las variaciones de los alelos en un locus observado en al menos dos cromosomas hallados en el ADN genómico de una población de organismos individuales de la misma especie. El término "polimorfismo" incluye las variaciones de la secuencia de ADN obtenidas del mismo locus de fragmentos de cromosomas clonados en otros vehículos, tales como vectores de ADN o el ADN cromosómico de otro organismo.

Tal como se usa en la presente memoria, "secuencia flanqueante de ITR" se refiere a la secuencia nucleotídica

adyacente a una ITR en una cadena de la secuencia de ADN que contiene una ITR. Las secuencias que incluyen la secuencia flanqueante de ITR como parte de su secuencia completa son ellas mismas secuencias flanqueantes.

La expresión "cebador oligonucleotídico", tal como se usa en la presente memoria, define una molécula compuesta de más de tres desoxirribonucleótidos o ribonucleótidos. Aunque cada secuencia cebadora no necesita reflejar la secuencia exacta del molde, cuanto mejor refleje la secuencia la complementariedad hacia un molde, mejor será la unión al molde. Su longitud y secuencia exactas dependerán de muchos factores relacionados con la función principal y el uso del cebador oligonucleotídico, que incluye la temperatura, la secuencia del cebador, y el uso del método. Cada cebador oligonucleotídico descrito en la presente memoria comprende una secuencia de ácidos nucleicos que es complementaria a la secuencia de un marcador de ADN que flanquea a una secuencia de ITR. Los cebadores oligonucleotídicos son capaces de actuar como punto de inicio para la síntesis cuando se colocan en condiciones que inducen la síntesis de un producto de extensión de los cebadores complementario a una cadena de ácido nucleico. Las condiciones pueden incluir la presencia de nucleótidos y un agente inductor, tal como una ADN polimerasa a una temperatura y pH adecuados. En la realización preferida, el cebador es un oligodesoxirribonucleótido monocatenario de longitud suficiente para actuar como cebador en la síntesis de un producto de extensión de una secuencia específica en presencia de un agente inductor. La sensibilidad y especificidad de los cebadores oligonucleotídicos están determinadas por la longitud del cebador y la singularidad de la secuencia dentro de una muestra dada de ADN molde. Los cebadores oligonucleotídicos tienen normalmente una longitud de alrededor de más de 15 bases, y preferiblemente alrededor de 20 a 40 bases.

La expresión "par de cebadores oligonucleotídicos" se refiere a un par de cebadores, y cada uno comprende una secuencia de bases desoxirribonucleotídicas o ribonucleotídicas complementaria a las cadenas opuestas del ADN bicatenario que flanquean a la misma ITR. Cada par de cebadores oligonucleotídicos se selecciona preferiblemente para detectar una única ITR. Aunque cada secuencia cebadora no necesita reflejar la secuencia exacta del molde, cuanto mejor refleje la secuencia la complementariedad hacia un molde, mejor será la unión al molde.

La expresión "producto de extensión" se refiere a la secuencia nucleotídica que se sintetiza desde el extremo 3' del cebador oligonucleotídico y que es complementaria a la cadena a la que está unido el oligonucleótido.

La expresión "sonda oligonucleotídica", tal como se usa en la presente memoria, se refiere a una molécula monocatenaria de ADN o ARN que comprende una secuencia que es complementaria a una porción de una secuencia objetivo, tal como la secuencia de repeticiones intermedias en tándem de una muestra de ADN, en la que la porción de complementariedad es de una longitud suficiente para permitir que la sonda se hibride a la secuencia objetivo.

La expresión "artefacto por tartamudeo", tal como se usa en la presente memoria, se refiere a un tipo particular de artefacto observado al detectar una o más moléculas de ADN objetivo, en las que el ADN objetivo contiene repeticiones en tándem de la misma secuencia de unidad de repetición, lo que incluye las secuencias objetivo de repeticiones intermedias en tándem detectadas y analizadas. Cuando se detecta una muestra que contiene dicho ADN objetivo tras la separación de todo el ADN de la muestra por la longitud, p.ej. mediante el uso de electroforesis en gel, cada molécula de ADN objetivo produce una señal principal (p.ej. una banda principal en un gel); pero se puede detectar una señal secundaria próxima a cada señal principal. La señal secundaria se produce en general por la detección de fragmentos de ADN que difieren en su longitud del ADN objetivo debido a la adición o delección de una o más unidades de repetición de la secuencia de ADN objetivo. Los artefactos por tartamudeo se han atribuido a un emparejamiento incorrecto de cadenas por deslizamiento durante la replicación del ADN, tanto *in vivo* como *in vitro*. Véase, p.ej., Levinson y Gutman (1987), *Mol. Biol. Evol.*, 4(3):203-221; y Schlötterer y Tautz (1992), *Nucleic Acids Research* 20(2):211-215. Tales artefactos son especialmente evidentes cuando se amplifica ADN que contiene cualquiera de tales secuencias de repetición *in vitro*, mediante el uso de un método de amplificación tal como la reacción en cadena de la polimerasa (PCR), ya que cualquier fragmento secundario presente en una muestra o producido durante la polimerización se amplifica junto con los fragmentos principales.

La expresión "% de artefacto por tartamudeo", tal como se usa en la presente memoria, se refiere a una comparación de la amplitud de una señal secundaria (es decir, artefacto) respecto de la amplitud de una señal principal (es decir, del objetivo) observadas en una muestra de ADN obtenida de una única fuente, tal como una única colonia de bacterias o un único cromosoma de ADN genómico. El % de artefacto por tartamudeo se puede determinar en un ADN que no se ha amplificado; pero preferiblemente se determina tras la amplificación de al menos una secuencia objetivo de repeticiones intermedias en tándem. La expresión "% medio de artefacto por tartamudeo" se refiere a una media del % de artefactos por tartamudeo obtenida a partir de las medidas de % de artefacto por tartamudeo detectadas a partir de una muestra representativa de al menos veinte alelos en una población.

La expresión "ADN genómico", tal como se usa en la presente memoria, se refiere a cualquier ADN derivado en última instancia del ADN de un genoma. La expresión incluye, por ejemplo, el ADN clonado en un organismo heterólogo, el ADN genómico completo, y el ADN genómico parcial (p.ej., el ADN de un único cromosoma aislado).

El ADN detectado o aislado puede ser monocatenario o bicatenario. Por ejemplo, se puede obtener ADN monocatenario adecuado para el uso a partir de bacteriófagos, bacterias, o fragmentos de ADN genómico. Se puede obtener ADN bicatenario adecuado para el uso a partir de cualquiera de varias fuentes diferentes que contienen

ADN con secuencias de repeticiones intermedias en tándem, lo que incluye bibliotecas de fagos, bibliotecas de cósmidos, y ADN genómico o plasmídico bacteriano, y ADN aislado de cualquier organismo eucariótico, lo que incluye el ADN genómico humano. El ADN se obtiene preferiblemente del ADN genómico humano. Se puede usar cualquiera de varias fuentes diferentes de ADN genómico humano, que incluyen muestras médicas o forenses, tales como sangre, semen, muestras vaginales, tejido, pelo, saliva, orina, y mezclas de fluidos corporales. Tales muestras pueden ser recientes, antiguas, secas, y/o parcialmente degradadas. Las muestras se pueden recoger de pruebas en el escenario de un crimen.

B. Método de Aislamiento de Marcadores de ADN Polimórficos que Contienen una ITR:

Una realización es un método para aislar un fragmento de ADN que contiene una ITR, mediante el uso de la selección por hibridación. El método comprende las etapas de: (a) proporcionar una diversidad de fragmentos de ADN, en los que al menos un fragmento de ADN contiene una ITR; (b) proporcionar un medio de soporte que tiene al menos un oligonucleótido asociado con él, en el que el oligonucleótido incluye una secuencia de nucleótidos que es complementaria a una porción de la secuencia de repeticiones intermedias en tándem; y (c) combinar la diversidad de fragmentos de ADN con el medio de soporte en condiciones en las que los fragmentos de ADN, que incluyen cualquier fragmento de ADN que contiene la secuencia de ITR, se hibridan al medio de soporte.

La diversidad de fragmentos de ADN proporcionados en la etapa (a) del método se puede obtener mediante fragmentación de cualquier muestra de ADN que contenga una ITR, pero se obtienen preferiblemente mediante fragmentación del ADN genómico. Véase, p.ej. Current Protocols in Human Genetics (1994), capítulo 2: Development of Genetic Markers, Construction of Small-Insert Libraries from ADN genómico, pág. 2.2.1 y sigs., que se incorpora en la presente memoria como referencia. El método más preferido para preparar una diversidad de fragmentos de ADN para el uso en la etapa (a) es según las etapas que comprenden: fragmentar una muestra de ADN, por lo que se produce una población de fragmentos de ADN en la que al menos un fragmento de ADN contiene la ITR; unir un ligante que contiene una secuencia de cebado a al menos un extremo de cada fragmento de ADN en la población de fragmentos de ADN; y amplificar cada fragmento unido al ligante mediante el uso de un cebador oligonucleotídico que comprende una secuencia que es complementaria a la secuencia de cebado. Se puede unir un ligante diferente a cada extremo de cada fragmento. Sin embargo, se une preferiblemente un único ligante a cada extremo para permitir la amplificación mediante el uso de un único cebador oligonucleotídico que tiene una secuencia que es complementaria a la secuencia de cebado del ligante. La unión del ligante se lleva a cabo preferiblemente en presencia de una enzima ligasa, tal como la ADN ligasa de T4.

Se puede usar cualquiera de varios medios diferentes para producir la diversidad de fragmentos de ADN proporcionados en la etapa (a) del método, que incluyen sonicación o fragmentación con al menos una enzima de restricción, aunque solamente el ADN bicatenario se puede fragmentar con una enzima de restricción. Cuando se usa una enzima de restricción para fragmentar una muestra de ADN bicatenario, preferiblemente es una enzima de restricción con una secuencia de reconocimiento de cuatro pares de bases, que deja salientes monocatenarios, y que no corta la muestra de ADN dentro de la región de ITR de interés. Las enzimas de restricción preferidas para el uso en la fragmentación de una muestra de ADN bicatenario incluyen *Mbo* I, *Aci* I, *Bfa* I, *Dpn* II, *Hha* I, *Hin* P11, *Hpa* II, *Mse* I, *Msp* I, *Nla* III, *Sau* 3AI, *Taq* I, *Csp* 6I, y *Tai* I.

Los fragmentos de ADN unidos a un ligante producidos como se describió anteriormente se amplifican posteriormente, mediante el uso de una reacción de amplificación, tal como una reacción en cadena de la polimerasa, (Pat. de EE.UU. N° 4.683.202 de Mullis, K.B.), amplificación basada en la secuencia del ácido nucleico (NASBA) (Kievits et al. (1991) J Virol Methods 35(3):273-286), amplificación mediada por ligadura (Volloch et al. (1994) Nucleic Acids Res 22(13):2507-2511), amplificación por desplazamiento de cadena (SDA) (Walker et al. (1992) PNAC 89(1):392-396), amplificación con un único cebador independiente de la secuencia (SISPA) (Reyes (1991) Mol Cell Probes 5(6):473-481), o reacción en cadena de la ligasa (Pat. de EE.UU. N° 5.686.272 expedida a Marshall et al.).

El medio de soporte proporcionado en la etapa (b) del presente método comprende un soporte estacionario con al menos un oligonucleótido objetivo asociado a él. El soporte estacionario comprende preferiblemente un material capaz de acoplarse con el oligonucleótido directamente o indirectamente. El material adecuado capaz de acoplarse directamente con el oligonucleótido incluye nitrocelulosa, nailon, vidrio, sílice, y látex. Los ejemplos de soportes estacionarios adecuados para el uso en esta realización preferida del presente método incluyen una membrana de nailon, un filtro incrustado con partículas de sílice, esferas de vidrio, partículas magnéticas de sílice, o una resina que contiene sílice. El material adecuado capaz de acoplarse indirectamente al oligonucleótido por medio de un primer agente de acoplamiento unido al oligonucleótido y un segundo agente de acoplamiento unido a la superficie del soporte estacionario incluye avidina y estreptavidina, o un antígeno y anticuerpo hacia él.

El al menos un oligonucleótido objetivo asociado con el soporte estacionario incluye una secuencia de nucleótidos que es complementaria a una porción de la secuencia de repeticiones intermedias en tándem del fragmento de ADN. El término "porción", tal como se usa en la presente memoria, se refiere a una secuencia de nucleótidos dentro de la región de ITR del fragmento de ADN de longitud suficiente para que un oligonucleótido que tenga una secuencia complementaria a la secuencia se hibride con ella cuando se ponga en contacto con ella. La "porción" es preferiblemente una secuencia de una longitud de al menos 20 bases, y más preferiblemente una secuencia de al

menos 40 bases. El oligonucleótido objetivo tiene más preferiblemente una secuencia caracterizada por la fórmula  $(A_wG_xT_yC_z)_n$ , en la que A, G, T, y C representan los nucleótidos que pueden estar en cualquier orden; w, x, y y z representan el número de cada nucleótido en la secuencia, y oscilan de 0 a 7 y la suma de  $w + x + y + z$  oscila entre 5 y 7; y n representa el número de veces que la secuencia está repetida en tándem, y es al menos alrededor de 4 veces, más preferiblemente al menos alrededor de 8 veces, y lo más preferiblemente al menos alrededor de 15 veces.

En la etapa (c) del método, la diversidad de fragmentos de ADN se combina con el medio de soporte en condiciones en las que el fragmento de ADN que contiene la ITR se hibrida al medio de soporte. Cuando la diversidad de fragmentos es una diversidad de fragmentos de ADN bicatenario, el ADN se desnatura antes de la hibridación al medio de soporte. Los medios adecuados para desnaturar los fragmentos de ADN bicatenarios antes de la hibridación al medio de soporte incluyen la exposición del ADN a una temperatura que es lo suficientemente elevada como para desnaturar el ADN bicatenario, o la suspensión del ADN en una disolución desnaturante. El ADN se desnatura preferiblemente mediante el uso de una disolución desnaturante que contiene un agente desnaturante, tal como una base (p.ej., hidróxido sódico o hidróxido potásico). Cuando se usa una base para desnaturar los fragmentos de ADN, el pH de la mezcla resultante se ajusta preferiblemente a alrededor de un pH neutro, preferiblemente mediante la adición de un tampón a un pH de alrededor de 4,8 a la mezcla.

Una vez que los fragmentos de ADN se han hibridado al medio de soporte, el medio de soporte se lava preferiblemente para eliminar los fragmentos de ADN y cualquier otro material presente en cualquier disolución en la que esté contenido el medio de soporte o sobre la superficie del medio de soporte que no se hayan hibridado a él. Cualquier disolución de lavado usada está configurada preferiblemente para eliminar tales materiales sin liberar los fragmentos de ADN hibridados al medio de soporte.

Los fragmentos de ADN hibridados al medio de soporte se pueden liberar del medio de soporte mediante el uso de calor o una disolución de liberación adecuada, dependiendo de la naturaleza de la asociación entre el medio de soporte y los fragmentos de ADN. Por ejemplo, se puede usar agua o una disolución acuosa de salinidad baja tal como un tampón TE (p.ej. Tris-HCl 10 mM, pH 7,5, EDTA 1 mM) para liberar los fragmentos de ADN hibridados a un medio de soporte que consta de un material de sílice. Una vez liberados del medio de soporte, los fragmentos de ADN se pueden procesar para aislar adicionalmente el ADN que contiene la secuencia de ITR de otros fragmentos de ADN presentes en la mezcla resultante de los fragmentos de ADN liberados. Las etapas de procesamiento adicionales podrían incluir la rehibridación y el cribado según el método descrito anteriormente, o la clonación en un vector de ADN y el cribado de los transformantes de los clones.

La Figura 1 ilustra una realización preferida del método de aislamiento de un fragmento de ADN que contiene una ITR, en el que se prepara una población de fragmentos de ADN, se hibrida a un medio de soporte, se amplifica, se clona, y se criba en busca de los transformantes que contienen la ITR. Cada una de las etapas ilustradas en la Figura 1 está indicada con un número romano. La Etapa I muestra una molécula de ADN bicatenario (1) que se digiere con una enzima de restricción (2), lo que produce una población de fragmentos de ADN (no mostrados) que varían en tamaño, al menos uno de los cuales incluye la ITR objetivo. La flecha entre las Etapas I y II ilustran un ligante (3) que se añade a la población de fragmentos de ADN para producir una población de fragmentos (8) unidos al ligante (3) en el extremo de cada una de dos clases diferentes de fragmentos de ADN, fragmentos con la secuencia objetivo de ITR (6) y fragmentos sin la secuencia objetivo (4). Se añade un cebador oligonucleotídico (7) que tiene una secuencia complementaria a una secuencia de cebado de cada ligante (3) a la población de fragmentos de ADN (8) en la Etapa III, y la población se amplifica por medio de una reacción de PCR, por lo que se produce una población de fragmentos de ADN amplificados (9). En la Etapa IV, la población de fragmentos de ADN amplificados (9) se coloca en un recipiente (15) con una disolución de hibridación (12) y un filtro (10) con al menos un oligonucleótido que tiene una secuencia complementaria a una porción de la secuencia objetivo de ITR asociada a él. La disolución de hibridación favorece la hibridación de los fragmentos de ADN que contienen la secuencia de ITR al filtro. En la Etapa V, el filtro (10) se extrae del recipiente (15), y se liberan los fragmentos de ADN hibridados en él. La población enriquecida resultante de fragmentos liberados se re-amplifica en la Etapa VI, mediante el uso del mismo cebador oligonucleotídico (7) usado en la reacción de amplificación de la Etapa III. Finalmente, cada fragmento de la población amplificada enriquecida en fragmentos de ADN se clona en un vector plasmídico (18) en la Etapa VII. Los vectores se muestran en la Etapa VII clonados con fragmentos con la secuencia objetivo de ITR (6) y clonados con fragmentos sin la secuencia de ITR (4).

#### C. Método para Detectar una ITR Polimórfica que Tiene un Tartamudeo Bajo:

Se observa un artefacto por tartamudeo mínimo cuando se detecta una secuencia objetivo de ITR de una muestra de ADN que tiene tal secuencia según esta realización particular del método. El artefacto por tartamudeo medio observado es preferiblemente como máximo un 1,1%, más preferiblemente como máximo un 0,9%. La secuencia objetivo de ITR puede ser una secuencia de ITR perfecta o una ITR imperfecta. La muestra de ADN detectada es preferiblemente ADN genómico.

El artefacto por tartamudeo medio se observa preferiblemente tras la amplificación de la secuencia de ITR en la muestra de ADN.

D. Cebadores, Sondas, y Marcadores

5 La presente descripción describe los marcadores de ADN identificados en la Lista de Secuencias más adelante como SEQ ID N<sup>o</sup>s: 1-43, cebadores en los que cada cebador tiene una secuencia que es complementaria a una secuencia que flanquea a una región de ITR de uno de los marcadores de ADN identificados mediante una de esas 43 secuencias, y sondas que tienen una secuencia que es complementaria a una secuencia contenida dentro de la región de ITR de uno de los 43 marcadores. Los cebadores preferidos específicos identificados en los experimentos ilustrados en los Ejemplos se enumeran a continuación en la Tabla 1.

TABLA 1

Marcador, SEQ ID N <sup>o</sup>	Número de Clon	Cebadores, SEQ ID N <sup>o</sup>	Cebador Superior y Cebador Inferior
1	C074	44	TGGCTCAGACACCTCATTG
		45	CACCACTGTATTCCCAGTTTG
2	C221	46	CACTTGCCATCCCTGCCACACA
		47	AGCGCACCCCAATTTCCGGTAT
		48	TGGGGACATGAACACACTTTGC
		49	GAGGCCCAGGACCAGATGAAAT
		50	CACCTGTCAGGCAAGGCTTAAAC
		51	CAACACTGAGCGCTTTTAGGGACT
		52	TCAGGCAAGGCTTAAACAGGGATA
		53	AACTGAGCGCTTCTAGGGACTTC
3	C221	52	TCAGGCAAGGCTTAAACAGGGATA
		54	TGAGCGCTTCTAGGGACTTCTTCA
		55	CCCTGCCCTACCCACTTG
		56	AGGCCCAGGACCAGATGA
		57	GCACCTGTCAGGCAAGGCTTAAAC
		58	CCAGCCATGAAGTGGCTGTGAG
4	C240	59	CCCGCTTCAAAGTTCCCAGTTC
		60	CCTCCCATTTCAGCCTCCTGA
5	C331	61	GTCTGCCACAGTGCTGGAACTAA
		62	GCACCCAGCCTAAGGCAATA
6	C362	63	GCATGGCGGAAGAAACAA
		64	TGGCAACAGAGCGAGACTC
7	C390	65	CCTGGGTGACAGCGAGAATCT
		66	TGTCCCTTGCCTTGTCTCACTAAA
8	G022	67	CAGCCTTGGTGACAGAGCAAA
		68	TGTGTTGAGGGTGGGGTACAT
8	G023	69	CCTGGGCAAGAGAGCAAG

ES 2 475 416 T3

Marcador, SEQ ID N°	Número de Clon	Cebadores, SEQ ID N°	Cebador Superior y Cebador Inferior
		70	CACATCCCAAAACCACCCTAC
9	G025	71	GCATTTCCCCTGCTTGTACT
		72	GATCACATTTGCTAACCACTTCTC
10	G047	73	GGCAACATATCAAGACCCCCATCTCT
		74	GAAGCTGCCCCTCACCCTACATTTT
11	G065	75	GATCACATTTGCTAACCACTTCTC
		76	TATAAATTACCCAGTCTCAGGAAG
12	G085	77	GTGATACAGCAAGCCTCATC
		78	AGAGACTCCTGGAAAGATAAAAGT
13	G132	79	GTCTGGAGAACAGTGGCCCTTGT
		80	CAGGAAGCTGAGGCAGGAGAATCT
14	G145	81	AAGGCTCCAGTGGGGTAT
		82	AAAACAAGGCAGTAGTCAATAAAG
15	G152	83	GGCATGAGAATCGCTTGAACCTG
		84	GGCCTCCATGATGTTTCCAATGAT
16	G153	85	TCAGGAGGCATGAGAATCGCTTGA
		86	GGCCTCCATGATGTTTCCAATGA
17	G158	87	CTCGCCCTCTCCTATAAGCAGTTT
		88	GCAGAGATAATTTGGAGTGGGATG
18	G181	89	CTTGGGTGCCTGTAATCC
		90	GGTAGAGCTCCCCATCT
19	G210	91	GCAGAATATTGGGGCTCATCAC
		92	AAACAAGGAAAGGAGAGGAGAGGA
	G210	93	AAGGTTGTGGGATGACTACTACA
		94	TGGTCAACACAGCAAGACATT
20	G212	95	TCCTGCCACCTGCTTGCTTTCT
		96	ATTGCACTCCAGCCTGGGTGATAC
21	G233	97	CGCTTGAGCCTTGAGATTG
		98	GAGCAGTCAGAATTCAGGAGTTGT
22	G234	99	TGGGCAACAAGAGCAAACTCCAT
		100	GGGACTTGGGCTGAGGGCTTTAC
23	G235	101	ATATCAATATCAGGCAGCCACAGG
		102	CCGTTTCAGAGCAGAGGTTTAGC

ES 2 475 416 T3

Marcador, SEQ ID N°	Número de Clon	Cebadores, SEQ ID N°	Cebador Superior y Cebador Inferior
24	G331	103	TCTCATTGGTTTCAAAGAACTTA
		104	AGACTCCATCTCAAACAAAAGA
25	G405	105	TCATGTGCATGGAGCCTGGTTCAT
		106	CCCAGCCTTGGAAGAGTGAGGT
26	G475	107	GGCGACTGAGCAAGACTC
		108	TTAAGCAAAGTAGCCTCAAACA
		109	GGGCGACTGAGCAAGACTC
		110	ACTCATTACCTTGCATGCATGATA
		107	GGCGACTGAGCAAGACTC
27	G539	111	CATTACCTTGCATGCATGATA
		112	TGGGCAACAGAGTAAGACTCA
		113	GTTTCAGTACCGTTCACCTCTTTA
		114	GTAAGACTCAGTCTCCAAAAAAAAAAAAAAG
		115	AGGAATGGTTTCTCTGTTAGTAAATGGT
28	S023	116	CAGCCTGGGCAACAAGAATGAAAC
		117	TGGCCCCTGCAGCGGAGTC
29	S071	118	GAATTCATTTGCGGAAAGATT
		119	CTAGGGAGGCTGGAGTATTCA
30	S085	120	AGAGCAAGACCCCGTCTCAT
		121	AGTCCATGGGCCTTTTAACA
31	S125	122	GAGAATCACTTGAACCCAGGAAG
		123	AGAACCAGCTGTTAGTTTCGTTGA
32	S132	124	GGTTCAGTGAGCCGAGATAAGAGT
		125	TGTGCCAGGAACCAGAAATTTACAG
33	S136	126	GGCCCAAGTTACTTTTCAC
		127	GGGCACTGCACTCCT
34	S159	128	CATGGTGAGGCTGAAGTAGGAT
		129	GTGGCGTGTCTTTTACTTTCTTTA
35	S176	130	AGGCAGCCCAGGAACAAT
		131	CCAAGATAGCGGCCAAGATAGT
36	S189	132	GAGGGCAGCTGGGATGTTACTCTT
		133	TGCCCTGTTTGGAGAAGTGTAGGT
37	S199	134	CTCCCCAGAAACAGATGTA

Marcador, SEQ ID N°	Número de Clon	Cebadores, SEQ ID N°	Cebador Superior y Cebador Inferior
		135	GTGAGCCGAGATTGTATCAT
38	S040	136	TCGGGGACAGGGCTTACTC
		137	ATCATTGTGCTGCTACTTTATCG
39	S066	138	CTACTCTACCCCATTTTCATTC
		139	GTAGAGTGGAGTGGATGAGA
40	S077	140	ATCAGGCAAAAACGAACAAAC
		141	CGGCATCCCAAAGTGAC
41	S097	142	CAGAGAGGGCAGCACCTTGGACAG
		143	GGCTTCACCTGCTCCCCTTTCAG
42	S103	144	TCTGCCCATTCGCCAGCCTCTC
		145	TACCGCGTGGCATTCAAGCATAGC
43	S110	146	TCCAGTCTGGGTGACAAA
		147	CAATCCACTCCACTCCTCTA

Los ejemplos siguientes se ofrecen a modo de ilustración, y no pretenden limitar la descripción de ninguna manera. En los ejemplos, todos los porcentajes son en peso si es para sólidos y en volumen si es para líquidos, y todas las temperaturas están en grados Celsius a menos que se indique de otra manera.

5 Ejemplo 1 Construcción de una biblioteca de PCR del genoma completo.

Las técnicas de selección mediante amplificación e hibridación particulares usadas en este Ejemplo, y en el Ejemplo 2, más adelante, son formas modificadas de un método de selección descrito en Armor, J. et al. (1994) *Hum Mol Genet* 3(4):599-605.

10 Se purificó ADN genómico humano a partir de sangre completa mezclada de 15 individuos mediante el uso de procedimientos de extracción con fenol:cloroformo habituales (Current Protocols in Human Genetics (1994), Gilber, J. ed., Apéndice).

15 Se cortaron aproximadamente 100 µg de ADN genómico con 5 unidades de enzima de restricción *Mbo* I por µg de ADN durante 16 h a 37 °C, seguido de purificación mediante extracción con fenol:cloroformo, precipitación con etanol y se resuspendieron en 100 µl de tampón TE (Tris-HCl 10 mM, EDTA 1 mM, pH 8,0) para una concentración final de alrededor de 1 µg/µl de ADN.

20 Los fragmentos de ADN que se hallaban en un intervalo de tamaños de 250-600 pb se aislaron mediante electroforesis en gel con un gel de agarosa preparativa del 1% de SeaKem GTG (FMC Bio Products, Rockland, Maine) (15x20 cm) durante 1,25 horas a 100 voltios y se recuperaron mediante electroelución (referencia). El ADN se cuantificó midiendo la absorbancia a  $A_{260}$  y se diluyó a 500 ng/µl en agua nanopura estéril y se almacenó a -20 °C.

25 Los ligantes se prepararon hibridando cantidades equimolares de oligo A (5'-GCG GTA CCC GGG AAG CTT GG-3') y 5' oligo B fosforilado (5'-GAT CCC AAG CTT CCC GGG TAC CGC-3') para una concentración final de 1.000 pmol/µl. Un µg de inserto de ADN seleccionado por tamaño (3,5 pmoles con un tamaño medio de 425 pb) se unió a 13 µg (875 pmoles) de ligantes (proporción molar 250:1 de ligante:inserto), mediante el uso de 1-3 unidades de ADN ligasa de T4 durante 16 hr a 15 °C. Los ligantes en exceso y los dímeros de ligante se separaron de los fragmentos principales mediante electroforesis en gel (agarosa del 1% de SeaKem GTG, 1,5 h a 100 voltios). Los fragmentos de ADN unidos al ligante se recuperaron del gel mediante electroelución, y se resuspendieron en 50 µl de agua estéril.

30 El ADN (50 ng) con los ligantes unidos se amplificó mediante el uso de una PCR en un volumen de reacción de 100 µl que contenía 10 µl de un tampón STR 10X (KCl 500 mM, Tris-HCl 100 mM, pH 9,0, MgCl<sub>2</sub> 15 mM, 1 % de Triton X-100, y 2 mM de cada dNTP), 1 µl de *Taq* polimerasa (5 U/µl), y 1 µM de cebador oligo A (10 µl de una disolución de reserva de 10 pmol/µl). El "oligo A" usado como cebador en esta reacción es el mismo "oligo A" usado para ensamblar el ligante de *Mbo* I, como se describió anteriormente. Las condiciones de los ciclos fueron 95°C

1 min, 67°C 1 min, 70°C 2 min; durante 30 ciclos. Los dNTPs, cebadores y dímeros de cebadores se eliminaron mediante microfiltración con Centricon-100s (añadir 2 ml de agua estéril a la muestra y cargar el Centricon-100, centrifugar 20 min a 2.000 RPM, invertir el filtro Centricon y centrifugar durante 2 min a 2.000 RPM para recuperar el ADN, resuspender en 100 µl de dH<sub>2</sub>O estéril). Se comprobó una alícuota de 5 µl de la biblioteca de PCR resultante en un gel de agarosa del 1% (1 hr a 100 voltios) para confirmar que el intervalo de tamaños estuvo entre 250 y 600 pb.

#### Ejemplo 2 Enriquecimiento de repeticiones de pentanucleótidos mediante selección por hibridación.

Los fragmentos de ADN de la biblioteca de PCR del genoma completo producidos según el Ejemplo 1 que contenían diversas repeticiones diferentes se enriquecieron mediante hibridación con el uso de diferentes mezclas de oligonucleótidos asociados a un soporte sólido. Los fragmentos que contenían repeticiones de pentanucleótidos (AAAAX)<sub>n</sub> se enriquecieron mediante selección por hibridación. Este proceso se llevó a cabo construyendo primero los oligonucleótidos para el uso en la selección por hibridación que consistieron en matrices en tándem de (AAAAC)<sub>n</sub>, (AAAAG)<sub>n</sub> y (AAAAT)<sub>n</sub> con una longitud de alrededor de 1000 pb. Estos oligonucleótidos se fijaron a membranas y se hibridaron con la biblioteca de PCR del genoma completo para seleccionar los fragmentos que contenían las repeticiones (AAAAX)<sub>n</sub>.

La matriz de oligonucleótidos se construyó como sigue: (a) se sintetizaron oligonucleótidos de 30 unidades monoméricas fosforilados en 5' de [AAAAC]<sub>6</sub>, [AAAAG]<sub>6</sub> y [AAAAT]<sub>6</sub> y sus complementos [GTTTT]<sub>6</sub>, [CTTTT]<sub>6</sub> y [ATTTT]<sub>6</sub> y se suspendieron en agua nanopura a una concentración de 1.000 pmol/µl, (b) se combinó una concentración equimolar (se usaron 10 µl o 10 nmol o 198 µg de cada uno) de oligonucleótidos que tenían secuencias complementarias, se calentó a 65°C durante 15 minutos y se dejó a 4°C durante unas horas para que hibridasen entre sí, (c) los oligonucleótidos hibridados se ligaron después entre sí mediante el uso de 1 Unidad Weiss de ADN ligasa de T4 por µg de ADN a 15°C durante la noche, (d) los concatémeros ≥200 pb se seleccionaron por tamaño en agarosa del 1% de SeaKem GTG, (e) el ADN ligado se sometió a una PCR sin cebadores para alargar las matrices en tándem, (f) los fragmentos de tamaño aparente de más de 1000 pb se recuperaron de los geles de agarosa del 1% y se purificaron mediante microfiltración. Se determinó la absorbancia a A<sub>260</sub> y se hizo una disolución de reserva de un µg/µl en agua nanopura estéril.

Después se transfirió un total de un µg de oligonucleótido (AAAAC)<sub>200</sub>, (AAAAG)<sub>200</sub>, o (AAAAT)<sub>200</sub> a trozos de 4 mm x 4 mm de filtro de membrana de nailon Hybond-Nfp (Amersham Life Sciences, Inc.), se lavó dos veces en tampón de prehibridación durante 30 minutos con agitación para eliminar los oligonucleótidos unidos débilmente, se dejaron secar al aire, se reticularon con luz UV a 1200 µJulios para unir el ADN, y después se almacenaron a -20 °C.

La selección por hibridación de la biblioteca de PCR del genoma completo al medio de soporte resultante de oligonucleótidos asociados con el filtro de nailon descrito anteriormente se llevó a cabo como sigue: (a) los filtros se prehibridaron en 1 ml de Tampón de Prehibridación [1% de BSA (Sigma B-4287), EDTA 1 mM, pH 8,0, 7% (p/v) de SDS, Na<sub>2</sub>HPO<sub>4</sub> 0,5 M] a 40°C para los filtros que contenían oligonucleótidos que tenían las secuencias de (AAAAC)<sub>n</sub> y (AAAAG)<sub>n</sub>, y a 37°C para los que contenían las secuencias (AAAAT)<sub>n</sub>. Después de 20 minutos, el tampón se elimina y se añaden 100 µl de Tampón de Prehibridación reciente, (b) se desnaturalizó el ADN de la biblioteca de PCR del genoma completo (20 µg) con álcali (KOH, concentración final 150 mM) y se neutralizó mediante la adición de 0,25 volúmenes de Tris-HCl 1 M de pH 4,8 y se añadió al tampón que contenía los filtros. La mezcla de reacción resultante se incubó durante la noche a temperaturas de prehibridación de 37°C o 40 °C, (c) los filtros con (AAAAC)<sub>200</sub> y (AAAAG)<sub>200</sub> se lavan 2X con 1 ml de Tampón de Lavado n° 1 (Na<sub>2</sub>HPO<sub>4</sub> 40 mM, pH 7,2, 0,1 % de SDS) a 40°C y 1X a temperatura ambiente durante 15 minutos con agitación. Los filtros con (AAAAT)<sub>200</sub> se lavan 1X con 1 ml de Tampón de Lavado n° 1 a 37°C y 1X a temperatura ambiente, (d) el ADN unido a cada filtro se liberó mediante calentamiento a 95°C durante 5 minutos en 100 µl de agua nanopura estéril. La muestra se extrajo mientras estaba a 95°C para prevenir la re-hibridación. Los filtros se retiraron y se reutilizaron mediante incubación en NaOH 0,4 M durante 30 minutos a 45 °C, después se transfirieron a SSC 0,1X, 0,1 % de SDS, Tris 0,2 M de pH 7,5 y se incubaron otros 15 minutos. Las membranas se secaron y se almacenaron en tubos sellados a -20 °C.

#### Ejemplo 3 Clonación de una biblioteca enriquecida en repeticiones de pentanucleótidos de fragmentos de ADN.

La población de fragmentos de ADN enriquecidos en repeticiones de pentanucleótidos según el Ejemplo 2 se re-amplificó mediante PCR. Los fragmentos reamplificados se clonaron después en el vector plasmídico pGEM-3Zf(+), como se describe más adelante. Este proceso se llevó a cabo ligando los insertos seleccionados en el vector pGEM y después transformando el plásmido circularizado en un hospedador de *E. coli* JM109.

Las ligaduras inserto-vector se llevaron a cabo como sigue: (a) 5 µl del ADN seleccionado por hibridación se reamplificaron en un volumen de reacción de 100 µl, mediante el uso de un tampón STR 1X (KCl 50 mM, Tris-HCl 10 mM, pH 9,0, MgCl<sub>2</sub> 1,5 mM, 0,1 % de Triton X-100, y 0,2 mM de cada dNTP), 1 µl de *Taq* polimerasa (5 U/µl), y 1 µM de cebador oligo A (1 µl de disolución de reserva de 100 pmol/µl). Las condiciones de los ciclos fueron 95°C 1 min, 67°C 1 min, 70°C 2 min; durante 30 ciclos. (b) El ADN reamplificado se digirió con *Mbo* I añadiendo 11 µl de Tampón C 10X de enzima de restricción de Promega y 2 µl (8 U/µl) de *Mbo* I a los 100 µl de reacción de PCR, incubando la mezcla de reacción resultante durante la noche a 37 °C, e inactivando térmicamente la enzima de restricción incubando la mezcla a 65°C durante 20 minutos. (c) El vector pGEM-3Zf(+) (20 µg o 10,6 pmol) se

preparó para la inserción del fragmento mediante digestión con *Bam*HI (5 U/μg) durante 16 horas a 37 °C, seguido de la adición de cantidades adecuadas de tampón de fosfatasa alcalina intestinal bovina 10X (Promega) y 1 μl de CIAP (Unidades/μl) e incubación durante 1 hora a 37 °C. Esta reacción se detuvo mediante la adición de EDTA 0,5 M a una concentración final de 0,02 M y después se extrajo con fenol, se precipitó con etanol y se resuspendió en tampón TE a 1 μg/μl. (d) Finalmente, se llevaron a cabo ligaduras inserto-vector de 20 μl mediante la incubación de 1 μl de ADN cortado con *Mbol* (véase la etapa b) junto con 1 μl o 200 ng de pGEM 3Zf(+) desfosforilado (véase la etapa c) y 1 μl de ADN ligasa de T4 (1 a 3 U/μl) durante 2 horas a temperatura ambiente.

Finalmente, se transformaron 10 μl de la reacción de ligadura inserto-vector en 100 μl de células JM109 competentes mediante el uso del protocolo de transformación de Promega descrito en el Boletín Técnico n° 095.

10 Ejemplo 4 Selección de clones de la biblioteca genómica de insertos pequeños que contienen repeticiones de pentanucleótidos (AAAAX)<sub>n</sub> mediante hibridación de colonias.

Los clones que contenían repeticiones de pentanucleótidos (AAAAX)<sub>n</sub> se seleccionaron mediante cribado de hibridación de colonias con el uso de los reactivos y protocolos Lightsmith II (véase el Boletín Técnico de Promega n° TM227), y se visualizaron mediante hibridación con sondas conjugadas a fosfatasa alcalina.

15 El ADN de las colonias se transfirió a membranas colocando membranas de nailon MagnaGraph (Micron Separations, Inc. Westboro, MA) en placas que contenían las colonias bacterianas, dejándolas en reposo durante 3 minutos, y después transfiriéndolas a papel de filtro seco. A continuación, las membranas se transfirieron a una serie de bandejas que contenían un 10% de SDS durante 3 minutos, después disolución desnaturante que consistía en 5 ml de NaOH + 30 ml de NaCl 5 M + 65 ml de dH<sub>2</sub>O durante 5 minutos, después disolución neutralizante que consistía en 30 ml de NaCl 5 M + 25 ml de M Tris-HCl, pH 7,4 + 45 ml de dH<sub>2</sub>O durante 5 minutos, y finalmente SSC 2X durante 5 minutos. Las membranas se secaron después a temperatura ambiente durante 30 minutos seguido de reticulación mediante luz UV con 1200 μjulios, con el uso de un Stratalink® (Stratagene, La Jolla, CA).

25 La detección de las colonias que contenían clones con repeticiones (AAAAX)<sub>n</sub> se llevó a cabo con la ayuda de sondas conjugadas a AP y quimioluminiscencia. La exposición de los filtros hibridados a sondas conjugadas con AP a una película de rayos X indicó las colonias que contenían los clones deseados. Se llevó a cabo una segunda hibridación para confirmar los resultados iniciales.

30 El procedimiento de detección utilizó el equipo Lightsmith II de Promega (véase el boletín de Promega n° TM227 para una descripción detallada del procedimiento). Brevemente, el procedimiento de detección usado consistió en las etapas de: (a) Incubar los filtros en una Disolución de Bloqueo Quantum Yield® (N° de cat. de Promega F1021) durante 45 minutos a 56°C con agitación enérgica, (b) desechar la Disolución de Bloqueo y añadir 0,05 ml de Disolución de Hibridación de Rigurosidad Alta Quantum Yield® (N° de cat. de Promega F1231) por cm<sup>2</sup> de membrana que contenía la sonda de AP e incubar 45 minutos a 56°C con agitación enérgica, (c) desechar la disolución de hibridación/sonda de los filtros y lavar los filtros dos veces con 150-200 ml de Disolución de Lavado n° 1 precalentada (SSC 2X, 0,1 % de SDS) durante 10 minutos a 56 °C, (e) combinar todos los filtros y lavar una vez con Disolución de Lavado n° 2 (SSC 1X) durante 10 minutos a temperatura ambiente, (f) equilibrar las transferencias durante 5 minutos en 200 ml de dietanolamina 100 mM, MgCl<sub>2</sub> 1 mM, (f) añadir suficiente sustrato CDP-Star 0,25 mM (Tropix, Bedford, MA) a filtros saturados y después incubar durante al menos 5 minutos a temperatura ambiente, (g) colocar los filtros saturados de sustrato sobre un protector de láminas de plástico de poliestireno en una carpeta de hibridación y cerrar la carpeta, (h) colocar la carpeta de hibridación que contiene los filtros en un casete de películas y exponer los filtros contenidos en él a una película de rayos X, y (l) revelar la película después de un periodo de al menos 1 hora de exposición a la película.

Ejemplo 5 Secuenciación de ADN y análisis.

45 Se desarrolló un método simplificado para preparar moldes de secuenciación mediante la utilización de lisados celulares para secuenciar el gran número de clones identificados en el Ejemplo 4 que posiblemente contenían insertos con al menos una secuencia (AAAAX)<sub>n</sub>. Este procedimiento consistió en transferir los clones positivos de los ensayos de hibridación de colonias a placas de microtitulación de 96 pocillos estériles (n° de cat. de Falcon 3072) que contenían 200 μl de LB/Amp (100 μg/ml) e incubarlos durante la noche a 37°C a 250 rpm. A continuación, el cultivo de la noche se dividió y se usó en tres procedimientos diferentes que implicaron crear los lisados celulares, producir filtros replicados para segundas hibridaciones para confirmar los hallazgos iniciales o hacer disoluciones de reserva en glicerol para el almacenamiento a largo plazo de los clones.

55 Los lisados celulares se hicieron tomando 2 μl del cultivo de la noche y añadiéndolos a 100 μl de agua nanopura estéril en placas de reacción de 96 pocillos (n° de cat. de Perkin Elmer N801-0560) y calentando a 100°C durante 4 minutos en un termociclador 9600. Esto se dejó enfriar, se congeló, y se almacenó a -20°C hasta que estuvo listo para su uso.

Los filtros replicados se hicieron para segundos ensayos de hibridación esterilizando en llama el replicador de 96 clavijas, sumergiendo el replicador en una placa de 96 pocillos que contenía el cultivo de la noche y

estampándolo en una membrana de nailon circular de 137 mm (MagnaGraph, MSI) sobre una placa LB / Amp (100 µg/ml) e incubando la membrana durante la noche a 37 °C.

El cultivo de la noche restante se convirtió en disoluciones de reserva en glicerol añadiendo 46 µl de glicerol al 80% a cada pocillo y colocando las placas en un agitador - incubador ajustado a 250 rpm durante unos minutos para mezclarlo, y después se almacenaron a -70 °C.

Se seleccionaron todos los clones que fueron positivos en dos ensayos de hibridación de colonias, y los clones correspondientes de las placas de lisado celular se usaron para la amplificación mediante PCR. Los productos de reacción de PCR se purificaron con placas de purificación Qiagen QIAquick 96 PCR (nº de cat. 28180) y se usaron como moldes para la secuenciación. Se usaron dos microlitros del lisado celular en una reacción de PCR de 50 µl que contenía el cebador directo M13 -47 a 2 µM (nº de cat. de Promega Q560A), el cebador inverso M13 (nº de cat. de Promega Q542A) a 2 µM, tampón STR 1X y 2,5 unidades de AmpliTaq (Perkin Elmer). Se usó el siguiente perfil de ciclos en un termociclador PE 480: 1 ciclo a 96°C / 2 min, 10 ciclos a 94°C / 1 min, 56°C / 1 min, 70°C / 1,5 min; 20 ciclos a 90°C / 1 min, 56°C / 1 min, 70°C / 1,5 min; mantener a 4 °C. Los productos de la reacción de PCR se lavaron con placas de purificación Qiagen QIAquick 96 PCR (nº de cat. 28180) siguiendo el protocolo del fabricante y se recuperaron en 70 µl de Tris-HCl 10 mM de pH 8,5 a una concentración final de alrededor de 35 ng/µl y se almacenaron a -20 °C.

La secuenciación del ADN se llevó a cabo mediante el uso de la química de secuenciación con terminadores marcados ABI y el secuenciador ABI 377. Los moldes de secuenciación se prepararon mediante el uso del equipo de terminadores marcados ABI y el protocolo del fabricante (Protocolo P/N 402078). Se usaron dos µl o aproximadamente 30 a 90 ng de producto de PCR purificado (descrito anteriormente) como ADN molde para la reacción de secuenciación. La reacción de secuenciación consistió en 8 µl de mezcla de terminadores marcados, 2 µl de ADN molde (35 ng/µl), 4 µl de cebador directo M13 -21 a 0,8 µM, y 6 µl de agua nanopura estéril. El perfil de ciclos en la secuenciación cíclica en el GeneAmp PCR System 9600 fue: 25 ciclos a 96°C / 10 s, 50°C / 5 s, 60°C / 4 minutos; mantener a 4°C. Los productos de extensión se purificaron mediante la adición de 50 µl de etanol del 95% y 2 µl de acetato sódico 3 M, pH 4,6 a cada tubo, se mezclaron mediante el uso de un vórtex, se colocaron sobre hielo durante 10 minutos, y después se centrifugaron durante 30 minutos a velocidad máxima. El sedimento se lavó con 250 µl de etanol del 70%, se secó en una centrifuga de vacío durante alrededor de 3 minutos y se almacenó seco a -20°C hasta que estuvo listo para su uso. El sedimento seco se resuspendió en 6-9 µl de tampón de carga y después se desnaturalizó durante 2 minutos a 95°C y se almacenó en hielo hasta que se cargó en el gel.

Se prepararon geles Long Ranger del cinco por ciento (FMC BioProducts, Rockland, ME) según el protocolo del fabricante y se polimerizaron durante 2 horas. El gel se pre-utilizó durante 45 minutos a 1000 voltios. Se cargaron 1,5 µl de molde en tampón de carga en el gel y se hizo funcionar en condiciones 2X o 4X durante 3,5 h o 7 h, respectivamente.

Los datos de la secuencia de ADN generados en el secuenciador ABI 377 se editaron para eliminar cualquier secuencia del vector pGEM y después se colocaron en la base de datos local creada mediante el uso del paquete informático del Genetics Computer Group Wisconsin, versión 9.0 (Madison, WI) que contiene información de secuencias para todos los clones que se estaban estudiando. A continuación, los clones se examinaron en busca de la presencia, longitud y patrones de secuencias de las repeticiones de pentámeros. Los que contenían 5 o más repeticiones se compararon después con el programa de comparación de secuencias BLAST (Altschul et al., 1990) para identificar los clones duplicados y aquellos que ya existían en la base de datos GenBank en el National Center for Biotechnology Information en Bethesda, Maryland, EE.UU. Una vez que se identificaron los clones únicos, se diseñaron cebadores para PCR con la ayuda del programa OLIGO Primer Analysis Software versión 5.0 (National Biosciences, Inc., Plymouth, MN).

Ejemplo 6 Cribado de clones en busca de los niveles de polimorfismo y determinación de la localización cromosómica.

El cribado inicial de polimorfismos se llevó a cabo en dos muestras de ADN mezcladas, una que contenía ADN genómico humano de 15 individuos aleatorios y la otra que contenía 54 individuos del CEPH del NIGMS Human Genetic Mutant Cell Repository (mezcla de ADN de la colección CEPH, nº de cat. NA13421, Coriell Cell Repositories, Camden, NJ). Se usaron cebadores de PCR marcados con fluorescencia para la amplificación mediante PCR del locus objetivo a partir de ADN genómico, y los productos de PCR se separaron en geles de poliacrilamida y se visualizaron en un escáner de fluorescencia. Esos loci con 4 alelos y un 50% de heterocigosidad se ensayaron posteriormente con 16 ADNs del CEPH individuales (102-1, 102-2, 884-1, 884-2, 1331-1, 1331-2, 1332-1, 1332-2, 1347-1, 1347-2, 1362-1, 1362-2, 1413-1, 1413-2, 1416-1, 1416-2) para determinar los valores de heterocigosidad preliminares. Los datos para los mismos loci se analizaron después adicionalmente para determinar el número de alelos, las frecuencias de los alelos y los valores de heterocigosidad (véase la TABLA 2).

Los clones que se descubrió que contenían secuencias de repeticiones pentaméricas que cumplieron los criterios de selección de  $\geq 4$  alelos y  $\geq 50\%$  de heterocigosidad se cartografiaron para determinar la localización cromosómica exacta (véase la TABLA 2). Se usaron tres métodos diferentes para la cartografía: (1) La cartografía de híbridos de células somáticas mediante el uso del panel NIGMS de 26 híbridos de células somáticas (Coriell Cell Repositories,

Camden, NJ) que representan cromosomas humanos únicos para identificar el origen cromosómico, (2) técnicas de cartografía de híbridos con radiación que utilizan el panel GeneBridge 4 RH de 93 clones RH (Schuler et al., 1996), y (3) técnicas de cartografía de ligamiento meiótico estándar y ocho familias (K102, K884, K1347, 1362, 1331, 1332, 1413, 1416) del panel de referencia de familias del CEPH y cartografiado con el programa de ligamiento multi-punto CRI-MAP (Lander y Green, 1987).

Los clones con valores de heterocigosidad que superaron el 70% en los 16 individuos del CEPH se estudiaron en cuanto a su genotipo y las frecuencias de los alelos en estudios de población mayores que contenían más de 100 individuos de las cuatro razas principales, que incluyen afroamericanos, blancos, asiáticos, y latinoamericanos. Las Figuras 10 y 11 ilustran la amplia variación de la migración de los alelos amplificados de dos loci de ITR polimórficas diferentes en muestras de ADN genómico de 24 individuos diferentes en una población (muestras de ADN S02 a S25). Véase la Tabla 1 anterior para la secuencia de los pares de cebadores usados en estos análisis. Las imágenes del gel se generaron amplificando cada locus de repeticiones de pentanucleótidos mediante el uso de cebadores marcados con fluorescencia, seguido de separación en geles de poliacrilamida y visualización mediante barrido con el escáner de fluorescencia FMBIO II (Hitachi Software Engineering America, Ltd., San Francisco, CA). Se incluyó una escalera alélica que contenía la mayoría de los alelos conocidos para cada locus ensayado en un carril a cada extremo del gel de electroforesis, en los carriles S01 y S26. Los pares de cebadores usados para amplificar cada locus tuvieron secuencias complementarias a al menos una porción de la secuencia de un marcador de ADN aislado del clon S159 o del clon G210, tal como se ilustró anteriormente en los Ejemplos. Las secuencias del par de cebadores se seleccionaron de los pares de cebadores enumerados para los Clones S159 y G210, Tabla 1, anteriormente.

Las condiciones de la PCR para los cribados de polimorfismos fueron las siguientes: reacciones de 25 µl que contenían aproximadamente 200 ng para un molde de ADN mezclado o 25 ng para ADNs del CEPH individuales, Tampón STR 1X, 1 unidad de *Taq* ADN Polimerasa, y 1 µM del par de cebadores correspondiente. La secuencia de cada par de cebadores usado para amplificar cada uno de los clones enumerados en la Tabla 2 se proporciona en la Tabla 1. Obsérvese que a cada cebador se le ha asignado el SEQ ID N° enumerado en la Tabla 1. Las condiciones de los ciclos para el termociclador Perkin-Elmer GeneAmp PCR System 9600 (Perkin-Elmer, Foster City, CA) fueron: 96°C durante 1 minuto, después 10 ciclos a 94°C durante 30 segundos, rampa de 68 segundos hasta 60 °C, mantenimiento durante 30 segundos, rampa de 50 segundos hasta 70 °C, mantenimiento durante 45 segundos; seguido por 20 ciclos de 90°C durante 30 segundos, rampa de 60 segundos hasta 60 °C, mantenimiento durante 30 segundos, rampa de 50 segundos hasta 70 °C, mantenimiento de 45 segundos, 60°C durante 30 minutos. Las muestras de PCR se prepararon mezclando 2,5 µl de cada muestra con 2,5 µl de Disolución de Carga de Azul de Bromofenol 2X, se desnaturalizaron mediante calentamiento a 95°C durante 2 minutos, se congelaron, después se sometieron a electroforesis 3 µl de cada muestra en un gel de poliacrilamida del 4% durante 50 minutos a 40 vatios. Los productos de la PCR se visualizaron mediante barrido en un escáner de fluorescencia Hitachi FMBIO y se analizaron con el soporte informático adjunto (FMBIO Analysis, versión 6.0, Hitachi Software Engineering, San Francisco, CA).

TABLA 2

SEQ ID N°	Número de Clon	Número de Acceso a GenBank	Secuencia de ITR más larga observada	N° de Alelos Observados	% de Heterocigosidad (blancos)	Localización Cromosómica
1	C074	ninguno	[TTTTG] <sub>9</sub>	6	75	1
2	C221	ninguno	[GTTTT] <sub>13</sub>	7	78	9p
3	C240	ninguno	[CAAAA] <sub>7</sub>	4	42	NA
4	C331	ninguno	[GTTTT] <sub>10</sub>	5	43	NA
5	C362	ninguno	[GTTTT] <sub>5</sub>	4	62	4
6	C390	ninguno	[CAAAA] <sub>7</sub>	5	56	NA
7	G022	ninguno	[AAAAG] <sub>6</sub>	4	63	2p
8	G023	ninguno	[AAAAG] <sub>10</sub>	12	71	16q
9	G025	ninguno	[AAAAG] <sub>6</sub>	12	86	1
10	G047	ninguno	[AAAAG] <sub>9</sub>	5	86	2p
11	G065	ninguno	[TTTTC] <sub>6</sub>	13	100	1q

ES 2 475 416 T3

SEQ ID N°	Número de Clon	Número de Acceso a GenBank	Secuencia de ITR más larga observada	N° de Alelos Observados	% de Heterocigosidad (blancos)	Localización Cromosómica
12	G085	ninguno	[AAAAG] <sub>11</sub>	8	93	10q
13	G132	ninguno	[CTTTT] <sub>15</sub>	12	100	4 qter
14	G145	ninguno	[AAAAG] <sub>13</sub>	8	33	NA
15	G152	ninguno	[AAAAG] <sub>6</sub>	5	87	8 qter
16	G153	ninguno	[AAAAG] <sub>8</sub>	5	88	8 qter
17	G158	ninguno	[AAAAG] <sub>5</sub>	8	75	5q
18	G181	ninguno	[GAAAA] <sub>14</sub>	5	72	NA
19	G210	ninguno	[CTTTT] <sub>6</sub>	9	56	8p
20	G212	ninguno	[CTTTT] <sub>9</sub>	6	100	NA
21	G233	ninguno	[AAAAG] <sub>8</sub>	12	50	10q
22	G234	ninguno	[AAAAG] <sub>12</sub>	4	80	16 qter
23	G235	ninguno	[TTTTTC] <sub>6</sub>	4	56	2p
24	G331	ninguno	[CTTTT] <sub>8</sub>	5	73	NA
25	G405	ninguno	[CTTTT] <sub>6</sub>	10	80	NA
26	G475	ninguno	[GAAAA] <sub>12</sub>	12	92	15q22.3
27	G539	ninguno	[GAAAA] <sub>12</sub>	13	100	15q26.2
28	S023	X05367	[AAAAT] <sub>6</sub>	4	50	NA
29	S071	M90078	[AAAAT] <sub>8</sub>	4	56	6q26-27
30	S085	U07000	[AAAAT] <sub>5</sub>	7	44	22q11
31	S125	Z73416	[AAAAT] <sub>13</sub>	5	64	22q11.2-qter
32	S132	Z83847	[AAAAT] <sub>10</sub>	8	69	22
33	S136	Z82250	[TTTTTC] <sub>6</sub>	11	94	22q12-qter
34	S159	AC00001 4	[GAAAA] <sub>9</sub>	12	72	21q22-qter
35	S176	AC00005 9	[GTTTT] <sub>9</sub>	4	56	7q21-7q22
36	S189	Z54073	[AAAAC] <sub>8</sub>	5	69	22q11.2-qter
37	S199	Z84475	[GTTTT] <sub>7</sub>	4	75	6q21
38	S040	X06583	[AGCCTGG] <sub>4</sub>	2	NA	NA
39	S066	M68516	[ACTCC] <sub>6</sub>	3	NA	NA
40	S077	M25718	[AATAC] <sub>12</sub>	6	NA	NA
41	S097	Z21818	[CAGGCT] <sub>3</sub>	3	NA	NA
42	S103	X15949	[ATCCC] <sub>8</sub>	3	NA	NA
43	S110	X54108	[GGA(A/G)T] <sub>32</sub>	6	NA	NA

## Ejemplo 7 Identificación de repeticiones cortas en tándem por medio de búsquedas en GenBank.

Se llevó a cabo un método alternativo para identificar secuencias repetidas en tándem buscando en GenBank en el National Center for Biotechnology Information (NCBI) la presencia de repeticiones intermedias en tándem. Se emplearon varios métodos, que incluyen la búsqueda por lotes de entradas de GenBank en CD-ROM con el paquete informático Lasergene de DNASTAR (Madison, WI), la búsqueda por lotes en GenBank con la ayuda del paquete informático del Genetics Computer Group Wisconsin, versión 9.0 (Madison, WI).

Existen  $4^5 = 1024$  palabras diferentes de cinco letras que se pueden construir a partir del alfabeto de cuatro letras (A, C, G, y T) para hacer todas las repeticiones pentaméricas posibles, y  $4^6 = 4096$  y  $4^7 = 16.384$  palabras diferentes de seis y siete letras para repeticiones de seis y siete bases. Sin embargo, el número de motivos de repeticiones únicas es considerablemente menor debido a la equivalencia de las dos cadenas complementarias (p.ej., AAAAT es equivalente a ATTTT), y debido a la equivalencia de permutaciones cíclicas (p.ej., AATAA... es equivalente a ATAAA...). En el caso de repeticiones de cinco bases, esto significa que existen 102 clases únicas de repeticiones pentaméricas si se excluyen las repeticiones de mononucleótidos  $A_5/T_5$  y  $C_5/G_5$ .

Se usaron todas las combinaciones únicas de repeticiones de 5, 6 y 7 bases con al menos tres copias consecutivas para realizar búsquedas en la base de datos del genoma humano GenBank. Se identificaron todas las regiones de repeticiones que contenían tres o más copias de una repetición, o las copias con sustituciones de bases ocasionales. Mediante el uso de los datos de secuencias existentes, se diseñaron cebadores que flanqueaban a la región de repeticiones y se amplificó mediante PCR el locus objetivo y se determinó el contenido polimórfico como se describió en el Ejemplo 6.

Cada clon que contenía una secuencia identificada mediante el uso de los cebadores ensamblados con el uso de la información de la base de datos GenBank se cribó después en función del contenido de secuencias de repeticiones como se describe en el Ejemplo 7. A la secuencia de cada clon que contenía una secuencia de ITR, es decir, un marcador de ITR, se le asignó uno de los SEQ ID N°s de 28 a 43. Véase la Tabla 1 para la secuencia de cebadores que comprenden las secuencias que flanquean a la región de ITR de cada marcador. Véase la Tabla 2 para un resumen de los resultados de analizar las características de la secuencia de cada marcador de ITR.

## Ejemplo 8 Estudio de los loci de Repeticiones Intermedias en Tándem en busca de artefactos de PCR (es decir, % de tartamudeo).

Muchos de los marcadores descritos en este trabajo representan una clase nueva de marcadores que producen menos artefactos de PCR conocidos como "tartamudeo" (véase la sección de Definiciones de la Descripción Detallada, anteriormente). La generación de estos artefactos se da durante la amplificación mediante PCR, supuestamente como resultado de un fenómeno relacionado con la ADN polimerasa denominado deslizamiento por repeticiones (Levinson y Gutman, 1987. Mol. Biol. Evol. 4(3):203-221; Schlotterer y Tautz, 1992. NAR 20:211-215). El resultado final del deslizamiento por repeticiones es la generación de productos de PCR que contienen diferentes números de unidades de repetición que el alelo auténtico. Si se da una cantidad suficiente de deslizamiento durante la PCR, el producto amplificado se visualizará en forma de una banda principal y secundaria, y la banda principal corresponde al alelo auténtico y la banda secundaria corresponde al producto alterado que contiene más o menos unidades de repetición.

Para cuantificar la cantidad de la banda tartamuda presente en loci diferentes, se analizaron productos de amplificación mediante PCR de 6 loci de ITR (C221, G023, G025, G210, S159 y una ITR adicional no descrita en esta patente, S117) y 17 loci de repeticiones en tándem de tetranucleótidos (F13A01, THO1, TPOX, F13B, FESFPS, D7S820, CSF1PO, D13S317, D8S1179, D16S539, LPL, FGA, D5S818, D3S1358, D18S51, vWA, y D21S11) en un secuenciador ABI 377 y se analizaron mediante el uso del programa GenScan (PE Applied Biosystems, Foster City, CA). Se determinaron las alturas de los picos medidas en unidades de fluorescencia relativas (UFR) para todos los picos principales y secundarios observados en las 25 a 40 muestras individuales investigadas en cada loci. Se calculó el porcentaje de UFR observado en el pico secundario (en general 5 pb más pequeño que el alelo auténtico en los pentanucleótidos o 4 pb más pequeño en las repeticiones de tetranucleótidos) respecto del pico principal del alelo auténtico (véase la Tabla 3).

Los ejemplos de electroferogramas de ABI 377 para los loci de ITR S159 (Fig. 2) y G210 (Fig. 3) y los loci de repeticiones de tetranucleótidos vWA (Fig. 4) y D5S818 (Fig. 5) muestran un tartamudeo mínimo o inexistente en los loci de ITR y un tartamudeo claramente observable para los loci de repeticiones de tetranucleótidos. De manera específica, véanse los artefactos por tartamudeo indicados mediante las flechas 14 y 15 en el electroferograma del locus de repeticiones de tetranucleótidos vWA reproducido en la Figura 3, y mediante las flechas 16 y 17 en el electroferograma del locus de repeticiones de tetranucleótidos D5S818 reproducido en la Figura 5. Compárense esos picos de artefactos diferentes con los pequeños artefactos inapreciables en los electroferogramas de las repeticiones de pentanucleótidos del ADN marcador aislado del Clon S159 (es decir, el marcador que tiene la secuencia identificada mediante SEQ ID N°:34) como se muestra en la Figura 2, y del ADN marcador aislado del Clon G210 (es decir, el marcador que tiene la secuencia identificada mediante SEQ ID N°: 19) en la Figura 4. Los electroferogramas específicos reproducidos en las Figuras 2 - 5 son las incidencias más altas de tartamudeo observado para cada uno de los loci.

Se observó cierta variabilidad en la cantidad de tartamudeo para todos los loci. En general, la tendencia fue para los alelos que contenían el número más alto de repeticiones (tal como se indica mediante su tamaño en pares de bases) para exhibir la cantidad más alta de tartamudeo. Los valores de tartamudeo en porcentaje para cada uno de los 25 a 40 individuos ensayados se muestran en gráficos de dispersión (Figuras 6, 7, 8 y 9).

- 5 En resumen, el porcentaje de la banda "tartamuda" respecto de la banda del alelo auténtico fue significativamente menor en la mayoría de los loci de ITR estudiados en comparación con los loci de repeticiones en tándem de tetranucleótidos. Esto fue cierto a pesar de que los loci de tetranucleótidos usados representasen el mejor de este tipo de marcador conocido actualmente. Por ejemplo, 13 de tales marcadores de tetranucleótidos, que incluyen varios de los marcadores de tetranucleótidos ensayados como se describió y se informa en la Tabla 3 siguiente por tener un % elevado de tartamudeo, han sido seleccionados por la Oficina Federal de Investigación de los EE.UU. para el uso en el análisis de todas las muestras de ADN para el Sistema de Índice Combinado de ADN nacional (CODIS). (Macivee, I. (1998) *Profiles in DNA* 1(3):2).

TABLA 3

Nombre del Locus o Número de Clon	Longitud de la Unidad de Repetición en Tándem	Porcentaje Medio de Tartamudeo	Porcentaje Máximo de Tartamudeo	Porcentaje Mínimo de Tartamudeo	Desviación Estándar	Número de Alelos Analizados
Clon S159	5 pb (ITR)	0,1	1,4	0,0	0,4	40,0
Clon G210	5 pb (ITR)	0,6	3,2	0,0	0,9	30,0
Clon C221	5 pb (ITR)	0,9	3,3	0,0	0,9	27,0
F13A01	4 pb	1,2	9,7	0,0	2,5	34,0
TH01	4 pb	1,7	5,2	0,0	1,7	34,0
Clon S117	5 pb (ITR)	2,0	6,9	0,0	1,7	37,0
Clon G023	5 pb (ITR)	2,3	6,6	0,0	1,7	39,0
TPOX	4 pb	2,4	5,6	0,0	1,8	34,0
F13B	4 pb	2,6	7,7	0,0	1,7	31,0
FESFPS	4 pb	3,6	10,0	0,0	2,3	34,0
D7S820	4 pb	3,8	8,2	1,6	1,6	28,0
CSF1 PO	4 pb	4,1	9,5	0,0	2,5	31,0
Clon G025	5 pb (ITR)	4,5	9,3	0,0	2,1	36,0
D13S317	4 pb	4,7	7,5	1,7	1,5	26,0
D8S1179	4 pb	5,0	8,3	2,4	1,6	27,0
D16S539	4 pb	5,1	8,6	1,7	2,0	28,0
LPL	4 pb	5,4	15,0	1,7	3,1	29,0
FGA	4 pb	5,5	11,6	3,0	1,7	36,0
D5S818	4 pb	6,1	9,0	0,0	1,9	28,0
D3S1358	4 pb	6,1	12,5	0,9	2,1	25,0
D18S51	4 pb	6,5	11,6	2,5	2,4	28,0
vWA	4 pb	6,6	11,4	3,7	1,4	28,0
D21S11	4 pb	7,5	15,7	1,9	3,5	30,0

LISTADO DE SECUENCIAS

- (1) INFORMACIÓN GENERAL:
- 5 (i) SOLICITANTE: Schumm, James W.  
Bacher, Jeffery W,
- (ii) TÍTULO DE LA INVENCION: MATERIALES Y MÉTODOS PARA IDENTIFICAR Y ANALIZAR  
MARCADORES DE ADN DE REPETICIONES INTERMEDIAS EN TÁNDEM.
- 10 (iii) NÚMERO DE SECUENCIAS: 147
- (iv) DIRECCIÓN DE CORRESPONDENCIA:  
15 (A) RECEPTOR: Promega Corporation  
(B) CALLE: 2800 Woods Hollow Road  
(C) CIUDAD: Madison  
(D) ESTADO: Wisconsin  
(E) PAÍS: EE.UU.  
(F) CÓDIGO POSTAL: 53711-5399
- 20 (v) FORMA LEIBLE POR ORDENADOR:  
(A) TIPO DE MEDIO: Disquete – 3,5 pulgadas, 1.44 Mb  
(B) ORDENADOR: PC compatible con IBM  
(C) SISTEMA OPERATIVO: Windows NT 4.0  
25 (D) PROGRAMA INFORMÁTICO: WordPerfect 7.0 (formato de texto DOS)
- (viii) INFORMACIÓN DE REPRESENTANTE/AGENTE:  
(A) NOMBRE: Grady J. Frenchick  
(B) NÚMERO DE REGISTRO: 29.018  
30 (C) NÚMERO DE REFERENCIA/EXPEDIENTE: 8976.80
- (ix) INFORMACIÓN DE TELECOMUNICACIÓN:  
(A) TELÉFONO: (608) 257-2281  
(B) NÚMERO DE FAX: (608) 257-7643  
35 (C) CORREO ELECTRÓNICO: [gfrenchick@mail.stroudlaw.com](mailto:gfrenchick@mail.stroudlaw.com)
- (2) INFORMACIÓN PARA SEQ ID NO:1
- 40 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 44 5 pb  
(B) TIPO: Ácido Nucleico  
(C) TIPO DE CADENA: Doble  
(D) TOPOLOGÍA: Circular
- 45 (ii) TIPO DE MOLÉCULA: ADN Genómico
- (iii) HIPOTÉTICO: no
- (vii) FUENTE INMEDIATA:  
50 (A) BIBLIOTECA: plásmido, pGem3Zf(+)  
(B) CLON: C074
- (viii) POSICIÓN EN GENOMA:  
55 (A) CROMOSOMA/SEGMENTO: 1
- 60
- 65

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:1:

GATCCTTTGC	ACCCAGANAG	AAGTAATTAT	TTCAACACAG	TTGGAACAGT	50
TAAAAAGATT	TAAAATTTTC	AAAAAAACAA	TCATTTTCTC	TTTTCTTTCT	100
GGCTCAGACA	CCTCATTGCT	TTCTGACTGA	CCAAGGCGCA	GCGCANTTTG	150
CAGCAGCCAT	GGGGGTCCA	GAGATTCCTG	GANAAAAACT	GGTGACAGAN	200
AGAAACAAA	AGCGCCTGGA	AAAAGATAAG	CATGAAAAAG	GTGCTCAGAA	250
AACAGATTGT	CAAAAGTAAG	TCTTACCTGT	GGCTCGCATT	ATTTGGGAGT	300
TATTAATA	TGAAAGTTTG	GCAATAACC	GGTTATCTAC	AGTCCTTTNG	350
TTTNGTTTTG	GTTTTGTTA	GTTTGGTTTT	GTTTNGTTN	GTTTGACACG	400
GAATCTCTCT	CTGTTGCCCA	AACTGGGAAT	ACAGTGGTGC	CGATC	445

5 (2) INFORMACIÓN PARA SEQ ID NO:2

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 411 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: C221

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 9p

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:2:

GATCACTTGC	CATCCCTGCC	ACACAGTTTC	CTCCTCTGGA	AACTGGGGGT	50
GATGACCCCT	GCCCTACCCA	CTTGTCATGG	CATTGGGGAC	ATGAACACAC	100
TTTGCACCTG	TCAGGCAAGG	CTTAAACAGG	GATATGCACT	GGTAATAGAA	150
AAGAGGGACT	AAGTTTTGTT	TTGTTTTGTT	TTGTTTTGTT	TTGTTTTGTT	200
TTGTTTTGTT	TTGTTTTGTT	TTGTTTTTCT	GAAGAAGTCC	CTAGAAGCGC	250
TCAGTGTGG	AATGCTCTCT	TGTAGCAGTG	GCGGCTGCTG	CTGGTTCCGG	300
GTCAGATGCC	GGAATTGGGG	GTGCGCTTGG	GTGCAGCTGC	ATTCATCTG	350
GTCCTGGGCC	TCCGTCTTGG	CTTGGAGAGG	TGCAGCTCAC	AGCCACTTCA	400
TGGCTGGGAT	C				411

(2) INFORMACIÓN PARA SEQ ID NO:3

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 354 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: C240

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:3:

GATCANCATG	GGTTCTATCT	GCCTGGCCCT	TCACCCCTA	CTCAGGGCAG	50
CTCTGAATTG	TCTNCCCCGC	TTCAAAGTTC	CCAGTTCAAC	TTCTCCCTCT	100
GCCCAATCCT	GTTTCCTTCT	CTTCCACAGG	TATTAATTTG	GCCAGNTGCA	150
GTGGCTCATG	CCTGTAATCT	CAACTTTGGG	AGGCCAAGGT	GGGAGGATTG	200
CTTGANCCCA	GAATTTTGAA	ACCANCCTCT	GAAACATANT	GANACCCCTG	250
TCTCAAAACA	AAACAAAACA	AAACAAAACA	AAACAAAAC	TANCCAGGCA	300
TGATGGTGTG	TGCCTGTGGT	CCCANCTATT	CAGGAGGCTG	AAATGGGAGG	350
ATC					353

5 (2) INFORMACIÓN PARA SEQ ID NO:4

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 317 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: C331

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:4:

GACCGTGGAA	NCCAAAAGTC	TGCCTA <del>CCGC</del>	ATCTTAGTCC	AGAGTTCCTG	50
TTTTTACTTC	TTTTTGAAGG	TCTGTGGATT	CTTTATTTTC	ATGGCACCTT	100
AGCAATACAT	TTTAAAAGCT	TGTTTTATTT	TATTCAGCAT	TTTGGTTATT	150
TCCATTGGAA	NANTCATPCA	GGGCGTTTAG	TCTGCCACAG	TGCTGGAAAC	200
TAAAGCTAGG	ATTACATGTT	TTGTTTTGTT	TTGTTTTGTT	TTGTTTTGTT	250
TTGTTTTGTT	TTGTTTTGTT	ACAGGGTCTT	GCTCTATTGC	CTTAGGCTGG	300
GGTGCAGTGT	TGTGATC				317

25 (2) INFORMACIÓN PARA SEQ ID NO:5

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 387 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: C362

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 4

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:5:

```

GATCTGGAGT GGAGAGCATT CCAGGCAGAA TGAAGAGCCA GGACCAAGAC 50
CACNAGGTGG AAACAGACTA ACAGAAAGAA AGCCANACCA CGAGGCAGAA 100

ACAGACTAAC AGAAAGAANA TCAGGTCGAC TTGCCTAAAA AGAGTGAGCT 150
AGGGAAAAGC ATGGCGGAAG AAACAANGTT GCTGAAAGCA ACTCTTATTT 200
TCTTGGCTTA GAAACCANNA AAATGCNTTT GGGTTTTATC TTAGCATAAT 250
GAAAAGACAT GTNANACTTC TGAACACGAA ATCTGACATG TTTTACAGAC 300
NTGTTTTACA TGGTTTTGTT TTGTTTNGTT TTGTTTTGGG ATGGAGTCTC 350
GCTCTGTTGC CANGCTGGGA GTGCAATGGT TGCATC 387
    
```

5

(2) INFORMACIÓN PARA SEQ ID NO: 6

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 471 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

10

(ii) TIPO DE MOLÉCULA: ADN Genómico

15

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: C390

20

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:6:

```

GATCACGAGG TCAGGAGATG GAGACCATCC TGGCTAACAT GGTGAAACCC 50
CGTCTCTACT AAAAATACCA AAAAATTAGC CGGGCATGGT GGCGGGCGCC 100
TGTAGTCCCA GCTACTCAGG AGGCTGAGGC AGGAGAATGG CGTGAACCCG 150
GGAGGCGGAG CTTGCAGTGA GCCGAGATTG CGCCACTGCG CTCCAGCCTG 200
GGTGACAGCG AGAATCTGTC TCAAAACATA ACAAACAAA ACAAACAAA 250
ACAAAACAAA ACAA AAAAGA TTTGGAATTA TG TAGGCAAA GTGGGAGAAA 300
GAGANGGACG AGGACTNAGG TAAAGATAAT ATGCAAAATA GAAAGAGCAN 350
GAAGGGGCAT GGATATGTGT AAATTCAAAAG AAAGGCAAAG TGGCTGGTGC 400
ACAAAGAGTG AGGAGAGCAA NNGTGA AAAA TGACTTTAGT GAGACAAGGC 450
AAGGGACAAA TCATGAAAAA T 471
    
```

25

(2) INFORMACIÓN PARA SEQ ID NO:7

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 367 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

30

(ii) TIPO DE MOLÉCULA: ADN Genómico

35

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G022

40

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 2p

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:7:

**GATCGCACCA CTGCACTCCA GCCTTGGTGA CAGAGCAAAA CTCNTTCTCC 50**

**AAAGAAAAGA AAAGAAAAGA AAAGAAAAGA AAAGAAAAAA AAAATCCATG 100**  
**GTGAAAGTGA CGACAGTNGA GTAGGGGATG AGCTCAAAGC AAATGCATGC 150**  
**ATGTNCCCCA CCCTCAACAC AAACACACAC ACACACACAC ACACACACAC 200**  
**ACACACACAC ACACATACTT CTTTAGAGAT ATTTAGGTGT ATATATGCTA 250**  
**ACTTAGGAAA CTTTAGAAAA CTTTGTATG ATATTATTAG TCAAAAAATA 300**  
**TTTAAGCCAC AGTTTCGCAA TTTTAAGATT GTACTACTGG TATCTGGAGT 350**  
**ATCTGAATCT CTGGATC 367**

5

(2) INFORMACIÓN PARA SEQ ID NO:8

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 295 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

10

(ii) TIPO DE MOLÉCULA: ADN Genómico

15

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf<+>
- (B) CLON: G023

20

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 16q

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:8:

**GATCACAGCA CTGCACTGCA GCCTGGGCAA GAGAGCAAGA CCCTCTCTCT 50**  
**CAGGGAAGAA AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA 100**  
**AAGAAAAGAA AGGAAGGAAA GAGAGAGGAA GGAAGGAAGG AAGGTAAGAA 150**  
**GGAAGGAAGG AAAGAAAGAA GGAAGGAAGG TAGGGTGGTT TTGGGATGTG 200**  
**AAATGCTGTC AGTCAACAAA GAGCTATGAC CACAGGTGTC ACTGAGTAGC 250**  
**AGGGGCAGCC CATCCTGCTC CCTAGCTGCA CTCACCCTGA AGATC 295**

30

(2) INFORMACIÓN PARA SEQ ID NO:9

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 361 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

35

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

40

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G025

45

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 1

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:9:

```
GATCTGATGG TTTCATAAGT GTCTGGCATT TCCCCTGCTT GTACTTCTCT 50
CCCCGGCTAC CGTGTGAAAA AGGTCCTTGC TTCCCCTTTG CCTTCCACCA 100
TGATTGTGAG CTTCTGAGG CCTCCACAGA CATGTGGAAC TGTGAGTCAA 150
TTAAACTTCT TTCCTTTATA AATTACCCAG TCTCAGGAAG TTCTTTGTAG 200
CAGTGTGAGA ATGGAGGAAG AAAGAAAAAG AAAAAAAGG AAAAGAAAAG 250
AAAAGAAAAG AAAAGAAAAG AAAGGAAGA AAGAAAGAAAG AAAGAAAAGAA 300
AGAAAGAAAG AAAGAAAAGAA AGAAAGAAAG AAAGAGAGAG AAGTGGTTAG 350
CAAATGTGAT C 361
```

5

(2) INFORMATION FOR SEQ ID NO: 10

(i) SEQUENCE CHARACTERISTICS:

- (A) LONGITUD: 318 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

10

(ii) TIPO DE MOLÉCULA: ADN Genómico

15

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G047

20

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 2p

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:10:

```
GATCACTTGA GGCCAGGGGT TCGAGGCCAG CCTGGGCAAC ATATCAAGAC 50
CCCCATCTCT ACATAAAAAG AAGAAGAAAC GAAAAGAAAA GAAAAGAAAA 100
GAAAAGAAA GAAAAGAAA GAAAAGAGTG GAAGAGTGCA GGAGCCGAGA 150
GGGAGAGAAA ATGTAGTGGT GAGGGGCAGC TTCTGGAAAG GCCCATACTA 200
CAGAGGGAGG AATCCTAATT CCTCACTATC TCTCTAACAT CAGGTAAGCA 250
TCTCATGATG CAGTTAGAAA GCACATTTCC TTCTTCAGTT TCCCCTCTGG 300
CTGTGTTGAC CCAGCCCA 318
```

30

(2) INFORMACIÓN PARA SEQ ID NO: 11

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 362 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

35

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

40

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G065

45

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 1q

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:11:

```
GATCACATTT GCTAACCACT TCTCTCTCTN TCTTTCTTIC TTTCTTTCTT 50
TCTTTCTTTC TTTCTTNCFT TCTTTCTTTC TATCTTCCTT TCTTTACTTT 100
NCTTTNCTNT TCTNTTCTAT TCCTTTANAT TTCTTTTCTT TTCTTTCTCC 150
ATTCTCACNC TGCTANAAAG AACTTCCTGA GACTGGGTAA TTTATANAGG 200
AAAGAAGTTT AATTGACTCA CAGTTCACA TGTTTGTGGA GGCCTCAGGA 250
AACTTACAAT CNTGGTGGAA NGCAAAGGGG AANCAAGGAC CTTTTTCACA 300
CGGTAGCCGG GGAATAAATT ACAANCAGGG GAAATGCCAN ACAC TTATGA 350
AACCATCAGA TC 362
```

5

(2) INFORMACIÓN PARA SEQ ID NO:12

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 297 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

10

(ii) TIPO DE MOLÉCULA: ADN Genómico

15

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G085

20

(vii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 10q

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:12:

```
GATCATGTCA TTGCACTCCA GCCTGGGTGA TACAGCAAGC CTCATCGAAA 50
GAAAAGAAAA GAAAAGAAAA GAAAAGAAAA GAAAAGAAAA GAAAAGAAAA 100
GAAAGGAAGA AAAGAAAACA AANAGATAGA AAGCAANCNN GTGGCNTGAG 150
AANTNAAATT CTTATAGGTA ACCTGGAGGA CTTTTATCTT TCCAGGAGTC 200
TCTCTCAATG CATTAGACT CAACAANGAT TTCCTTTTCT CTTGTCTCTA 250
NAAANAAATG CATTTCCTCA AAANANTGGA GGTCANATTA TGTTANAGAT 300
GGGAGAATGC ACTGAGTTNC GCTGAANGA 329
```

30

(2) INFORMACIÓN PARA SEQ ID NO:13

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 372 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

35

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

40

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G132

45

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 4 qter

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:13:

```

GATCTACCAT TCTTGGGTCT GGAGAACAGT GGCCCTTGTT TCTTTTCTTT 50
TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT CCTTTTCTTT 100
TCCTTTCTTT TCCTTTTCTT CTCTCTCTCC TTCTCTCTCT CTCTCTCTCT 150
CTCTCTCTCT CTCTCTCTCT CTCCCTCTCC CTTCCCTTCC CTTCCCTTCC 200
CTTCCTTTCC TTTCCCTTCA TTTTFTTTGA CATGGAGTTT CACTCTTGTC 250
ATCCAGGCTG GAGTACAGTA NTGTGATTTT GGCTCACTGC AACCTCTGCC 300
TCNTGGGTTT AAGAGATTCT CCTGCCTCAG CTTCTTGANT AGCTGGGATT 350
ACAGGTGCCT GCCACCATGC TT
    
```

5

(2) INFORMACIÓN PARA SEQ ID NO:14

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 350 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

10

(ii) TIPO DE MOLÉCULA: ADN Genómico

15

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G145.1

20

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:14:

```

GATCTCTTGA AGCCTCGCAN ATAAAGGCTC CAGTGGGGTA TGATTGCACC 50
ANTGCACTCC ANCCTGNGAN ACGGNAGAGA GATTCTGTCT CAAAAGAAAA 100
CAAAATAAAA GAAAANAAAA NAAAANAAAA TAAAANAAAA TANAAGAAAA 150
GAAAAGGATG CTTTAAAAAT NTGGCAAAAT GTNCCCTTTA TTGACTACTG 200
CCTTGTTTTA ATTNCTCTA TTTNTCTATT TATTTTCTCA GTGTACTTTC 250
CCATNINNCT TTNTCTCTC CTTCTTTGAA AGTAATTCTT GGCCAGGCAT 300
GGTGGTTTCA GCCTATAATC TCANCACTTN AGGGGGCTNA AGCNGGAAGA 350
    
```

25

(2) INFORMACIÓN PARA SEQ ID NO:15

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 372 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

30

(ii) TIPO DE MOLÉCULA: ADN Genómico

35

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G152

40

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 8 qter

45

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:15:

```
GACCACCTGA GGTCATGAGT TCCAGACCAG CCTGGCCAAC ATGGCAAAC 50
CCCGTCTCTA CTAAAAATAC AAAAAATAGC CGGTGTGATG GTGGGTGCCT 100
GTAATCCCAG CTA CTACTCAGGA GGCATGAGAA TCGCTTGAAC CTGGGAGGCG 150
GAGGTTGTAG TGAGCTGAGA TTGCGCCTCT GCACTCCAGC CTGAGTGATA 200
GAGTGAGACC CCATCTTGAA AGAAAAGAAA AGAAAAGAAA AGAAAAGAAA 250
AAGAAATTC A TCATTGGGAA ACATCATGGA NGGCCGNAC CAGTCAGGGG 300
AACATTTCCG AAAGCNANTT NTTCTTCAA TGCCCTATGT TNCTTCCCN 350
AAGCTTGCCA TTTTNAACCC TT 372
```

5 (2) INFORMACIÓN PARA SEQ ID NO:16

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 361 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

10

(ii) TIPO DE MOLÉCULA: ADN Genómico

15

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G153

20

(iv) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 8 qter

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:16:

```
GACCACCTGA GGTCATGAGT TCCAGACCAG CCTGGCCAAC ATGGCAAAC 50
CCCGTCTCTA CTAAAAATAC AAAAAATAGC CGGTGTGATG GTGGGTGCCT 100
GTAATCCCAG CTA CTACTCAGGA GGCATGAGAA TCGCTTGAAC CTGGGAGGCG 150
GAGGTTGTAN TGAGCTGAGA TTGCGCCTCT GCACTCCAGC CTGAGTGATA 200
GAGTGAGACC CCATCTTGAA AGAAAAGAAA AGAAAAGAAA AGAAAAGAAA 250
AGAANTTCNT CATTGGGAAA CATCATGGAG GCCGCAGCAN TCAGGGGGAAC 300
ATTTCCGAAA GCNAGTTGTC NTTCCAATGC CCTATGTTNC TTCCCNAAG 350
CNTGCCATTT T 361
```

(2) INFORMACIÓN PARA SEQ ID NO:17

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 447 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

30

(ii) TIPO DE MOLÉCULA: ADN Genómico

35

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G158

40

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 5q

45

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:17:

```
GATCGCCTGG GTACAGCAGG AAAGAAGGGG GCGGCCACGG CAAGGCAGCC 50
TCCGACTGCC CGCGGGGGA NGCCGGCGGC GGCCCCTTCT CGCCCTCTCC 100
TATAAGCAGT TTTATAAGCT TCCTGAGACT ANAAAAGGAA AAGAAAAGAA 150
AAGAAAAGAA AAGAAAATC AGTCTCTATT TTATATGCGT ATAATTTTTT 200
TTATATGCGT ATAATTTTTT TTTAACC AA ACTCNTTA TGGACAAAAC 250
AAACTACCAT CCCACTCCAA ATTATCTCTG CATCATGCTC ACAACCTCAG 300
CNCAAAATTC AATANAANTT TTATTGGGAT ATGTTTGGCT TCCATCAATT 350
GAAATTTCCC CTAATGAATA AAATTTCCCTC CCGTTTTTTT GGTAACATT 400
TCCCCTTGNA AGGCCACCT AAAAATCNC NGGNCTTTTT CCAAAGG 447
```

5

(2) INFORMACIÓN PARA SEQ ID NO:18

(i) CARACTERÍSTICAS DE LA SECUENCIA:

10

- (A) LONGITUD: 415 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

15

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

20

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G181

(viii) POSICIÓN EN GENOMA:

25

- (A) CROMOSOMA/SEGMENTO:
- (B) MAP POSITION:
- (C) UNITS:

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:18:

```
GATCCCAAGC TTCCCGGGTA CCGCGATCAC CTGAGGTCAG GAGTTCAAGA 50
CCAGCCTTCT CAACATGGCA AAACCTCAT TCTACTAAAA ATACAAAAAA 100
TTAGCTGGGC ATGGTCTTGG GTGCCTGTAA TCCAGCTAC TCAGGAGGCT 150
GAGGCAGGAG AATGTCTTGA ACCCAGGAGG CCGTGGCTGC AGTGAGGCAA 200
NATTTTGCCA GTGTNCTCCA GCCTGGGTGA CAANANTGAA ACTCCGTCTG 250
AAAGAAAGAA AGAAAAGAA AGAAAGGAAG GAAGGAAGGA AGGAAAGGGA 300
AGGAAAGAAA AGAAAAGAAA AGAAAAGAAA AGAAAAGAAA AGAAAAGAAA 350
AGAAAAGAAA AGAAAAGAAA AGAAAAGAAA TNAGATGGGG GAGCTCTACC 400
GAACTGATTC CGATC 415
```

30

(2) INFORMACIÓN PARA SEQ ID NO:19

(i) CARACTERÍSTICAS DE LA SECUENCIA:

35

- (A) LONGITUD: 444 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

40

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

45

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G210

(iv) POSICIÓN EN GENOMA:

(A) CROMOSOMA/SEGMENTO: 8p

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:19

5

```
GATCCTATCC TGACAAACTC AAGCAAATTC ACAAATACAA CCCTCTAGCC 50
GGCCCATGGC CTCCCTATTT GGGAGGAAAA AACTCAGTAT GATACTGTGA 100
CATATTTTTCAT TCATTATCTG TTAAGGTGAG CGTGGCAAAC CTGGCCGAAG 150
TGGCAGAATA TTGGGGCTCA TCACTTGGGG GAATGATTCA GGAGTGGCAT 200
CCTTCTGTGA CCGTGTGACAG CCACTTAAGG TTGTGGGATG ACTACTACAA 250
AATCCCAAAT AAAGTATATC CTAAGGCCTT TCTTTTCTTT TCTTTTCTTT 300
TCTTTTCTTT TCTCTTCTCA TCTCTTGTCT TCTCTTCTTT TCTCCTCTCC 350
CCTCCCCTCC CATCCCCTCT CCTCTCCTCT CCTTTCCTTG TTTTAAAAAC 400
AATGTCTTGC TCTGTTGACC AGGCTGGAAT GCAGTTCTGT GATC 444
```

(2) INFORMACIÓN PARA SEQ ID NO:20

10

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 321 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

15

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

20

(viii) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G212

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:20

25

```
GATCTCCTTC AGTGTACTCA GTGCATTCTC CATCTCTTAC ATAATCTGAC 50
CTCCACTCTT CCTGGAAATG CATTTCCTTT TAGAGACAAG AGAAAAGGAA 100
ATCCTTGTTG AGTCTAAATG CATTGAGANA NACTCCTGGA AAGATAAAAAG 150
TCCTCCAGGT TACCTTTAAN ACTTTCATTT CTCCTGCCAC CTGCTTGCTT 200
TCTCTCTCTT TCTTTTCTTT TCTTCCCTTC TTTTCTTTTC TTTTCTTTTC 250
TTTTCTTTTC TTTTCTTTTC TTTTCTTTTC ATGAGGCTTG CTGTATCACC 300
CAGGCTGGAG TGCAATGACA T 321
```

(2) INFORMACIÓN PARA SEQ ID NO:21

30

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 329 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

35

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

40

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G233

45

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 10q

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:21

5

```
GATCGCTTGA GCCTTGGAGA TTGAGGCTAC GGTGAGCTAT GATTGCACCA 50
CTGCACTCCA GCCTGGGTGA CAGAGTGAGA CCCTGGGAGA AAAAAAGAAA 100
GAAAAGAAAA GAAAAGAAAA GAAAAGAAAA GAAAAGAAAA GTCNTGACCT 150
TGGAAAAAAC CANAATTTCT GATGTTGTAC AACTCCTGAA TTCTGACTGC 200
TCTCTCCNCN GAAAGANGGA ATNNMTGNTC CTTGGAGGAT TCNTACTAAT 250
ATTCTTCGGT CNANACAAA ACNTGACCTC NAGCCNAGAA AACAAATTN 300
NNCCNTTCCA TAGAAAAGTT CAGGGGACA 329
```

(2) INFORMACIÓN PARA SEQ ID NO:22

10

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 412 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

15

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

20

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G234

25

(iv) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 16 qter

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:22

```
GGATCACGCC ATTGCACTCC ACTCTGGGCA ACAAGAGCAA AACTCCATCT 50
CAGAAAAAAA GAAAGAAAGA AAGAAAGAAA GAGAGAAAAG AAAACAGAAA 100
AGAAAAAGAAA AGAAAAGAAA AGAAAAGAAA AGAAAAGAAA AGAAAAGAAC 150
CCNACAGAAA GCCAAGGCAA TGGGAACAAG CTGGGGCAAG TGCCTGGAGG 200
TGTTGCTGGA AAGGCAGATA GGGCAGAGAG CACCTGGACT CTTCCAAAAC 250
ATATTAGCAT CATGGTAAAG CCCTCAGCCC AAGTCCCCCA GAACATAGCC 300
GTAGTCAACC AAGTTGAGAT TGATTACTAG CTTCTGTINA CAAGGGAGAT 350
TATNCNACA CAAGTGCCAT CTGCCTCTCC CTTACCCAG CTTGAGTTTC 400
GCTTGTAGCA CT 412
```

30

(2) INFORMACIÓN PARA SEQ ID NO:23

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 359 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

35

(ii) TIPO DE MOLÉCULA: ADN Genómico

40

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G235

45

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 2p

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:23

5

```
GATCACCAGG CCCCTGAGGA AGCAGCACAG AAAAAACACAA ATAATATCAA 50
TATCAGGCAG CCACAGGGGA AACAATGGGG CATTTCTCCG TGCTACATGC 100
ATGCTGCTAT TGTTTCAAGG GCTGGGGAAT TAATTCCACT TATTTATTTA 150
AGGCGTGTCA ACTCACTGCC TAAACCTGTT TCAGTGTCAA AATGGATAAA 200
ACTTTTATGG CTCATAAAAT ANANCCATTC ATCTCAATGT TCTTTGTGGT 250
GGGTTTTCTT TTCTTTTCTT TTCTTTTCTT TTCTTTTTTC TTTTTTTTTC 300
TGGCATACTG AGCTAAACCT CTGCTCTGAA ACGGTTACAT CTGAACCCAT 350
TGCTGCTAT 359
```

(2) INFORMACIÓN PARA SEQ ID NO:24

10

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 516 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

15

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

20

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G331

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:24

```
GACTTTCCCA CCTTCTGATG TGGGCATTTA GTGCTATAAA TTTCCCTCTA 50
AACACTGCTT TAGCTGTGTC CCANAGATTC TGGTATGTTG TGTCTTTGTT 100
CTCATTTGGT TCAAAGAACT TATTTATTTT TGCCTTAATT TTGTTATTTA 150
CCCAGTAGTC ATTCAGGAGA AGGTAGTTCA GTTCCATGT AGTTGTGAAG 200
TTTTGAGTGA GTTCTTTTCC TTTTCTTTTC TTTTCTTTTC TTTTCTTTTC 250
CTTCTTTCT TTCTTTTCTT CTTTCTTCT TTCTTTCTTT CTTTCTTCT 300
TTCTTTTGTG TGAGATGGAG TCTTACTCTG TCGCCAGTCT GGAGTGCAGT 350
GGTGTCACT CAGCTCGCTG CAACCTCCGC CTCCTGGGTT CAANAAATTC 400
CTCTGCCTCA GCCTCCCAAG TAGCTGGGTT TACAGGCACA CACCACCAG 450
CCCAGCTAAT TTTTGTATT TTANTAAAGA CAGGGTTTCA CCATGTTGAC 500
NAAAATGGTC TCGATC 516
```

(2) INFORMACIÓN PARA SEQ ID NO:25

30

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 55S pb
- (B) TIPO: Ácido Nucleico
- (A) TIPO DE CADENA: Doble
- (B) TOPOLOGÍA: Circular

35

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

40

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G405

45

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:25

```
GATCTCACAT TCTTCCTCAG AATTCCTTCTT GTTACCTCTG CAAAATTTCA 50
TCCTTCAAAC TCAAAGCTCA TTATCTTTGG ACTCTGTGAC ACTCTTCTGA 100
TTCTCATATC ACTTCTTGAT TTTCTGCAT TTCCTCACTA ACTCTCAGCT 150
CATAATCATA TAAAATCACT AAGACTCTTT TTATATGTC ATGAAGCTCA 200
GGTATTTTCA CAGATTGAAC CATTTCCTG TAGACAGCAA TGCTCAACAT 250
GAACCATTCA CATCCTTCTT CCAAAGCACA GACTCTTCTT GCCATCTGCG 300
TCATGCCCAT GCTCATGTGC ATGGAGCCTG GTTCATTATC TTCCAAAATC 350
AAGCTTCCCC CACTTGATTT CTCTTTTCTT TTCTFTCCTT TCCTTTCCCTC 400
TTTTCTTTT CCCTTCCCT TTCCTTACCT TTCCTTTCCT TTCCTTTCT 450
CTCCTCTTTT CTCTTTTCTT TTTCTTTCTT TTCTTTTCTT TTCCTTTCTT 500
TTCNTTTCTT TTATTTGCAC CTCACTCTTG CCAAGGCTGG GATGGCAGTA 550
ANCACG 556
```

10 (2) INFORMACIÓN PARA SEQ ID NO:26

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 335pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G475

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 15q22.3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:26

```
GATCAGCCA TTGCACTCCA GCCTGGGCGA CTGAGCAAGA CTCAGTCTCA 50
AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA 100
AAGAAAAGAA AATTGTAAGG AGTTTTCTCA ATTAATAACC CAAATAAGAG 150
AATTCTTTCC ATGTATCAAT CATGATACTA AGCACTTTAC ACACATGTAT 200
GTTATGTAAT CATTATATCA TGCATGCAAG GTAATGAGTA TTATTTTCTT 250
CATTTTATAA AAGAGGAAAC TGATGTTTGA GGCTACTTTG CTTAAGACCG 300
CAGAACTAGC AAAGGAAAAG AGAAGTGAAT GTATC 335
```

30 (2) INFORMACIÓN PARA SEQ ID NO:27

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 333 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA: plásmido, pGem3Zf(+)
- (B) CLON: G539

- (viii) POSICIÓN EN GENOMA:  
 (A) CROMOSOMA/SEGMENTO: 15q26.2

5

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:27

```
GATCGTGCCA CTGCACTCCA GCCTGGGCAA CAGAGTAAGA CTCAGTCTCC 50
AAAAAAAAAA AAAGAAAGAA AGAAAAAGAA AGAAAGAAAG AAAGAAAGAA 100
AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA AAGAAAAGAA 150
AAGAAAAGAA AAAGAAAAAG AAAAAATAAA GAGGTGAACG GTACTGAACA 200
GAAACTAAGA AGGCTGAGAG CCAACTCTGA GGTAACAGCT AGGAGCTGAA 250
GCAGGAAAGC TAAATCTGC CCCAGTCCCA TTGCTGATAG ACTCACCATT 300
TACTAACAGA GAAACCATT C TCCTTTTAG ATC 333
```

- 10 (2) INFORMACIÓN PARA SEQ ID NO:28

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 1011 pb  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Doble  
 (D) TOPOLOGÍA: Circular

15

- (ii) TIPO DE MOLÉCULA: ADN Genómico

20

- (iii) HIPOTÉTICO: no

- (iv) FUENTE INMEDIATA:  
 (A) BIBLIOTECA: plásmido, pGem3Zf(+)  
 (B) CLON: S023

25

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:28

```
CTGTACTGAA TTACAGCCCC AAATCTGGGT CAACTGGGGA GAGACGACGA 50
GGATTAGGGT TCCAAGGTGA AACTGTGCCA TTGCGCTCCA GCCTGGGCAA 100
CAAGAATGAA ACTCTCTTAA AATAAAATAA AATAAAATAA AATAAAATAG 150
CCTAAGGATG CATTCTCAG AACTTATCCC TGTTGTTCAA TGATGTGTGT 200
CTATACAGTG GGGCCATAAC TAAGACGTAT GTTGCCCAAG CTGGCAAGAT 250
AGCTCTGACC TTCTCTTGGG CCCCTCATTT CCCCCAACA CAGGTTGTCT 300
GCAGTCTTGA CCAATGGCTG CCAGGGCATG GACTCCGCTG CAGGGGCCAG 350
TGGGAGGCC CAGCTCAGGC AAAAGCACAG GCAGATATTT CAGGAGTCTG 400
CTAGGGCTGG CACTGAGGGC AGAGACAGAG GGGTCTCCCT GTCCTTTGGA 450
GAACCTCAG CTGCAGAAAT TCCAGACTGA ACCTTGATAC CGAGTAGGGG 500
AGGAGCTGTC TGCGGGTTTG AGCCTGCAGC AGGAGGAAGG ACGTGAACAT 550
TTTATCAGCT TCTGGTATGG CCTTGAGCTG GTAGTTATAA TCTTGCCCT 600
GGTGGCCAG GGCTACAGTC ATCCTAGCAG TCCCCGCTGA AGTGGAGCAG 650
GTACAGTCAC AGCTGTGGGG ACAGCAATGC TGGCCAAGGG TCTTCCCCCA 700
CGCTCAGTCC TGGTCAAAGG CTGCCAGACC TTTCTGAGTG CCCCCAGGGA 750
GGGGCTGGGG CGTCTCAGGG TGCCCACTGG CGAGGGAGCT GGCATCTCCA 800
CCCGCAGTCC TCGCCCCTTC AATGAGATCC CCTCTCCTGG TGACAATGGC 850
TGGCTAAACC TGTACCATTT CTGGAGGGAG ACGGGCACAC ACAAAGTCCA 900

CCTTCACCAT GTCCAGAATT TCCAGAAGTA TGGCCCGATT TACAGGTAAG 950
CCTGGCAGAG GGTGGGAGCC GAAGGACAGG GAGGAGGAGG GGAAGTGGTA 1000
GCCCCGCTGT A 1011
```

30

- (2) INFORMACIÓN PARA SEQ ID NO:29

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 1011 pb  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Doble  
 (D) TOPOLOGÍA: Circular

35

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

5 (iv) FUENTE INMEDIATA: ■ (B) CLON: S071

(viii) POSICIÓN EN GENOMA:

(A) CROMOSOMA/SEGMENTO: 6q26-27

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 29

10

```

CTGTACTGAA TTACAGCCCC AAATCTGGGT CAACTGGGGA GAGACGACGA 50
GGATTAGGGT TCCAAGGTGA AACTGTGCCA TTGCGCTCCA GCCTGGGCAA 100
CAAGAATGAA ACTCTCTTAA AATAAAATAA AATAAAATAA AATAAAATAG 150
CCTAAGGATG CATTCTCAG AACTTATCCC TGTTGTTCAA TGATGTGTGT 200
CTATACAGTG GGGCCATAAC TAAGACGTAT GTTGCCCAAG CTGGCAAGAT 250
AGCTCTGACC TTCTCTTGGG CCCCTCATT CCCTCAAACA CAGGTTGTCT 300
GCAGTCTTGA CCAATGGCTG CCAGGGCATG GACTCCGCTG CAGGGGCCAG 350
TGGGAGGCCC CAGCTCAGGC AAAAGCACAG GCAGATATTT CAGGAGTCTG 400
CTAGGGCTGG CACTGAGGGC AGAGACAGAG GGGTCTCCCT GTCCTTTGGA 450
GAACCTCACG CTGCAGAAAT TCCAGACTGA ACCTTGATAC CGAGTAGGGG 500
AGGAGCTGTC TGCGGGTTTG AGCCTGCAGC AGGAGGAAGG ACGTGAACAT 550
TTTATCAGCT TCTGGTATGG CCTTGAGCTG GTAGTTATAA TCTTGCCCT 600
GGTGGCCCAG GGCTACAGTC ATCCTAGCAG FCCCCGCTGA AGTGGAGCAG 650
GTACAGTCAC AGCTGTGGGG ACAGCAATGC TGGCCAAGGG TCTTCCCCCA 700
CGCTCAGTCC TGGTCAAAGG CTGCCAGACC TTTCTGAGTG CCCCAGGGA 750
GGGGCTGGGG CGTCTCAGGG TGCCCACTGG CGAGGGAGCT GGCATCTCCA 800
CCCGCAGTCC TCGCCCCTTC AATGAGATCC CCTCTCCTGG TGACAATGGC 850
TGGCTAAACC TGTACCATT CTGGAGGGAG ACGGGCACAC ACAAAGTCCA 900
CCTTCACCAT GTCCAGAATT TCCAGAAGTA TGGCCCGATT TACAGGTAAG 950
CCTGGCAGAG GGTGGGAGCC GAAGGACAGG GAGGAGGAGG GGA CTGGGTA 1000
GCCCTGCTGT A 1011
    
```

(2) INFORMACIÓN PARA SEQ ID NO:30

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1000 pb

(B) TIPO: Ácido Nucleico

(C) TIPO DE CADENA: Doble

(D) TOPOLOGÍA: Circular

20

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

25 (iv) FUENTE INMEDIATA:

(B) CLON: S085

(viii) POSICIÓN EN GENOMA:

(A) CROMOSOMA/SEGMENTO: 22q11

30

35

40

45

5

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:30

```

AGTCAAATGA CGGTCATAGT TTGGGTGATG GTCACGGCTC AGGTTCTTTT 50
TTACACGTGC TTGCTTTGGT TGTTGTTGTT GTTCTTTGTT TTCTTGAGGC 100
AGTATCTGGC TGTGTCTCCC AGGCTGGCGT GCAATGGCAG GATCATAGCT 150
CACTGCAACC TCAAACCTCT GGCTGAAGCA ATCTTTGTGC CCTAGCCTCC 200
CAAGTTGTTG GGATTACAGT CGTGCCCCAC CATGCCTGGC TAAGTTGTTT 250
TTTGTTTTTT GTTTTTTTTT TTTTTTTCGA GACAGAGTTT TGCTCTTGTT 300
GCCCAGGCCG GAGTGCAGTG GTGTGATCTT GGCTCTCTGC AACCTCCCGG 350
GTTCAAGCGA TTCTCCTGCC TCAGCCTCCC AAAGTGATGG GATTACAGGC 400
CTGAGCCACT GTGCCTGGCC ACATGTGCTT TCCCATTCGG TCCTTGACGC 450
AGATCTTTGA GAGAGCTCAT TTGACACTCA GGAGATGCTT CTCTAACCTG 500
CTCAGAATCA GGGCCCTGGG TATTCAGGGA GGTAGAGGGA GCAGACTGCA 550
AAGCCAGTCG TGCTCCCATC GCTCCCACTT CTCTCTCCCT CTCCATGTTT 600
TCTGTCTCCC CCACCCAGCC TAGGGCATT C'TCCCCACA GTCCAGCCTG 650
CATCTGGCAC AGTGTCACTG CTCAGCCCAG GGATACTCAC AGCCTGGGTG 700
CCTGGCTCCT TTTTTCAGCT CATCAAACCA GGTAAGGGA GGTTGAGATT 750
CTGCCAACCA TTGACTCAAT TCATCCAAAT CTTCATCAC TGGAACTCTG 800
GGAGTGGCTG GATTTGAACC AGGACCTCTG AGTACTATTG CTAAGTAACT 850
GGGGGTCTCA GTGAAAGAGA GAAAAGAGCT GATAGGCCTC TTCTGTGTT 900
ATCATGTCAG GCCATCTTTT GAAACTCTTT TCTGCAATGC TACTGAAGTA 950
TTTATGCACG TGACCTGTGC TCTTCTGTCA GTCTAGGGGT GCTGGCTGAG 1000
    
```

10

(2) INFORMACIÓN PARA SEQ ID NO:31

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1000 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA:

(ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA:
- (B) CLON: S125

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 22q11.2-qter

30 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:31

```

AGTCAAATGA CGGTCATAGT TTGGGTGATG GTCACGGCTC AGGTTCTTTT 50
TTACACGTGC TTGCTTTGGT TGTTGTTGTT GTTCTTTGTT TTCTTGAGGC 100
AGTATCTGGC TGTGTCTCCC AGGCTGGCGT GCAATGGCAG GATCATAGCT 150
CACTGCAACC TCAAACCTCT GGCTGAAGCA ATCTTTGTGC CCTAGCCTCC 200
CAAGTTGTTG GGATTACAGT CGTGCCCCAC CATGCCTGGC TAAGTTGTTT 250
TTTGTTTTTT GTTTTTTTTT TTTTTTTCGA GACAGAGTTT TGCTCTTGTT 300
GCCCAGGCCG GAGTGCAGTG GTGTGATCTT GGCTCTCTGC AACCTCCCGG 350
    
```

GTTCAAGCGA	TTCTCCTGCC	TCAGCCTCCC	AAAGTGATGG	GATTACAGGC	400
CTGAGCCACT	GTGCCCTGGCC	ACATGTGCTT	TCCCATTCGG	TCCTTGACAGC	450
AGATCTTTGA	GAGAGCTCAT	TTGACACTCA	GGAGATGCTT	CTCTAACCTG	500
CTCAGAATCA	GGGCCCTGGG	TATTCAGGGA	GGTAGAGGGA	GCAGACTGCA	550
AAGCCAGTCG	TGCTCCCATC	GCTCCCCTT	CTCTCTCCCT	CTCCATGTTT	600
TCTGTCTCCC	CCACCCAGCC	TAGGGCATT	CTCCCCACA	GTCCAGCCTG	650
CATCTGGCAC	AGTGTCACCTG	CTCAGCCCAG	GGATACTCAC	AGCCTGGGTG	700
CCTGGCTCCT	TTTTTCAGCT	CATCAAACCA	GGTAAAGGGA	GGTTCAGATT	750
CTGCCAACCA	TTGACTCAAT	TCATCCAAAT	CTTCAATCAC	TGGAATCCTG	800
GGAGTGCTG	GATTTGAACC	AGGACCTCTG	AGTACTATTG	CTAAGTAACT	850
GGGGGTCTCA	GTGAAAGAGA	GAAAAGAGCT	GATAGGCCTC	TTCCTGTGTT	900
ATCATGTCAG	GCCATCTTTT	GAAACTCTTT	TCTGCAATGC	TACTGAAGTA	950
TTTATGCACG	TGACCTGTGC	TCTTCTGTCA	GTCTAGGGGT	GCTGGCTGAG	1000

(2) INFORMACIÓN PARA SEQ ID NO:32

- 5 (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 1000 pb
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Doble
  - (D) TOPOLOGÍA: Circular
- 10 (ii) TIPO DE MOLÉCULA: ADN Genómico
- (iii) HIPOTÉTICO: no
- 15 (iv) FUENTE INMEDIATA:
  - (B) CLON: S132
- (viii) POSICIÓN EN GENOMA:
  - (A) CROMOSOMA/SEGMENTO: 22

20 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:32

GGTGTGACCT	TATCCTCTCT	GAACCTCAGT	TTCCTCATCC	GTAAAATGAA	50
AAGCTGCTAG	ATTGTTGTAA	AAAAATTAAA	TGGAATAGGC	TAGGCGCGGT	100
GGCTCACGCC	TGTAATCCCA	GCACCTTAGA	AGGTGGAAGA	GGGTGGATCA	150
CTTGAGGTCA	GGAGTTTTGA	GACCAGCCTG	GCCAACACGG	TGAAACCCCA	200
TCTCTACTAA	AAATAAAAAA	TTAGCTNNGG	TGCGGTGGCT	CACACCTGTA	250
ATCCAGCAC	TTTGGGAGGC	TGAGACGGGT	GGATCACCTG	AAGTCAGGAG	300
TTCAAGGCCA	GCCTGGGCAA	CATGCTGAAA	CCACGTCTCT	ACTAAAAATA	350
CAAAAATTAG	CCAGGTGTGG	TGGCACACGC	CTGTAGTCCC	AGCTACTTGG	400
GAGGCTGAGG	CGGAAGAATC	GCTTGAACCC	AGTAGGCAGA	GGTTGCAGTG	450
AGCCGAGATA	AGAGTCACTG	CACTCCAGCC	TGGGTGACAG	AGCAAGACTC	500
CCTCTCAGAA	AATAAAATAA	AATAAAATAA	AATAAAATAA	AATAAAATAA	550
AATAAAATTC	TAAAAGGGCT	GGCATTTGCC	TAGCACTTAT	ATGCCCAATA	600
AGTAATAGCT	ATCAATATCC	CCACCCCTAC	CACTGTGCTG	AAATTTAGTT	650
TCTTTTGTG	ACCCCCATT	AGACTTAAGG	CAGAATTCCT	ACCGTACTCC	700
TCTGTAAATT	TCTGTTTCT	GGCACATAGT	TGGGTCTCAG	TGAAACATGG	750
TGAGTGAATG	AGCAAATGCA	AGGAATCTCC	AGGCCATCTG	GGAGCCCTCC	800
CAGGCGGGTG	AGTTCGGGAA	ACTCATAGTC	TGTCCTCAAT	GGCCCACTGA	850
AAGGTAGAGA	GTCTGGGTC	CCACCTCCGC	ACCCCATCT	CCTGACTCAC	900
TGCTGAAAAA	TAAATAAATA	AATAAAATAA	CACTTATCCG	GAGCCTCCCA	950
CATGCCCTGC	CAGGACTGCA	AGGAGCCAG	CAGAATGATG	ACCGGCGTGC	1000

25 (2) INFORMACIÓN PARA SEQ ID NO:33

- 25 (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 1000 pb
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Doble
  - 30 (D) TOPOLOGÍA: Circular
- (ii) TIPO DE MOLÉCULA: ADN Genómico

- (iii) HIPOTÉTICO: no
- (iv) FUENTE INMEDIATA: (B) CLON: S136
- (viii) POSICIÓN EN GENOMA:  
(A) CROMOSOMA/SEGMENTO: 22q12-qter
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:33

5

```

CCACTACATA TCCCATACAG GCTAATCAAC ATGTCAAAGT TCACACAGTT 50
ATTGTGTACC CCTGGGCTCA ATCTCAAGTG TTCTGGTTGG TCGTCCAAGG 100
TTACTTTTTT TTTTTTTTTT TTTTTTTTGA GATGGAGTCT TGCTCTGTTG 150
CCCAAGCTGG AGTGCAATGG CATGATCTTG GCTCACTGCA ACCTCCGCCT 200
CCTGGGTTC AAGGATTCTC CTGCCTCAGC CTCCTGAGTA GTTGGGATTA 250
CAGGCATGCA CTACCATGCC TGGCTAATTT TTGTATTTTT AGTAGAGGTG 300
GAGTTTCTCC ATGTTGTTCA GGCTGGTCTT GAACTCCCAA CCTCAGGCAA 350
TCCACCTCGG CCTCCCAAAG TACTGGGGTT ACAGGCATGA GCCACTGCGC 400
CTGGCCCAAG GTTACTTTTC ACTACATCTT CCTACCTGTA TCACTTACTG 450
CCGTGTGTAT AACTTCCACA TTTTCTTTCT TTTCTTTTCT TTTCTTTTCT 500
TTTCTTTTCT TTCTTTTCTT TCTTTCCTTC TTTCTCTCTC TTTCTCTCTC 550
TCTTTCTCTC TGTCCCCTCC TTCCTTCTCC TTCCTTCTTC CTTCCCTCCT 600
TCCTTTCCTT CCTTCCTTCC TTCTTTCAAC ACAGAGTCTC ACTCTGTCAC 650
CTAGGCAGGA GTGCAGTGGC CCAGTCTCAG CTCACTGCAA CCTCCGCCTC 700
CTGGGCTCAA GCAATTCTCT CACCTCAGCC TCCCGAGTAG CTGGGATTAC 750
AGGCATGTGC CACCATACCC AGCTAATTTT TGTATTTTTA GTAGAGACGG 800
GATTTACCA TATTTTCCAA GCTGGTCTCG AACTCCTGAC CTCAAGGGAT 850
CTGCCCGACT CAGCCTCCCA AACTGCTGGG ATCATAGGTG TGAGCCATCA 900
TGCTTGGCCC ACACTTTCTA TGTTAATCTA ATTTAGATGA TTTAATCTAT 950
ATACAGTTTC TATATTAATC TAATTTAGAT GACTTAATCT ATATACAAC 1000
    
```

10

(2) INFORMACIÓN PARA SEQ ID NO:34

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 1000  
(B) TIPO: Ácido Nucleico  
(C) TIPO DE CADENA: Doble  
(D) TOPOLOGÍA: Circular
- (ii) TIPO DE MOLÉCULA: ADN Genómico
- (iii) HIPOTÉTICO: no
- (iv) FUENTE INMEDIATA:  
(B) CLON: S159
- (viii) OSICIÓN EN GENOMA:  
(A) CROMOSOMA/SEGMENTO: 21q22-qter

15

20

25

30

35

40

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:34

```

AAACCCCGTC TCTACTAAAA ATACAAAAGT TAGTTGAGCA TGGTGGCAGC 50
GGCCTGTAAT CCCACCTATA ATCCCACCTA CTCGGGAGGC TGAGGCAGGA 100
GAATCGCTTG AACCCAGGAT GGGGCGATTG CAGTGAGCCG AGATCGTGCC 150
ACTGCACTCC AGCCTGGGTG ACAGAGCGAG ACTCCATCTC AAAAAAAAAA 200
AAAAAAAAACA GAATCATAGG CCAGGCACAG TGGCTAATTG TACCTTGGGA 250
GGCTGAGACG GGAGGATCGA GACCATCCTG GGCACCATAG TGAGACCCCA 300
TCTCTACAAA AAAAAAAAAA AATTTTTTTT AAATAGCCAG GCATGGTGAG 350
GCTGAAGTAG GATCACTTGA GCCTGGAAGG TCGAAGCTGA AGTGAGCCAT 400
GATCACACCA CTACACTCCA GCCTAGGTGA CAGAGCAAGA CACCATCTCA 450
AGAAAGAAAA AAAAGAAAAGA AAAGAAAAGA AAAGAAAAGA AAAGAAAAGA 500
AAAGAAAAGA AAAAACGAAG GGGAAAAAAAA GAGAATCATA AACATAAATG 550
TAAAATTTCT CAAAAAATC GTFATGACCA TAGGTTAGGC AAATATTTCT 600
TAGATATCAC AAAATCATGA CCTATTAAAA AATAATAATA AAGTAAGTTT 650
CATCAAACCT TAAAAGTTCT ACTCTTCAA AGATACCTTA TAAAGAAAGT 700
AAAAAGACAC GCCACAGGCT AAGAGAAAGT ACTTCTAATC ACATATCTAA 750
AAAAGGACTT GTGTCCAGAT TAAAGAATTC TTACACATCA ATAAGACAAC 800
CCAAATAAAA ATCGGCAAAA GATTTGAAGA GATATTTAAC CAAAGAAAAC 850
ATATAAATGT GTCCGGGCGC GATGGTAATC CCAGCACTTT GAGAGGCCGA 900
GGCAGGCGGA TCACTTGAGG TCAGGAGTTT AGGACCAGTC TGGCCAACAT 950
GGTGAAACCC TGTCTCTAAT AAAAATACAA AAATTAGCTG GGTGTGGTGG 1000
    
```

5

(2) INFORMACIÓN PARA SEQ ID NO:35

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 14 00 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

10

(ii) TIPO DE MOLÉCULA: ADN Genómico

15

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (B) CLON: S176

20

(viii) POSICIÓN EN GENOMA:

- (A) CROMOSOMA/SEGMENTO: 7q21-7q22

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:35

25

```

CCATATGTTT GTTTCCTCTA CTA CTGCTCC TCCCTGACCC TTAAGAAACA 50
CTGCCATAGA GCCCTACAGC TTGATGGGAG AAGTCCTATC CCTTAGGCAT 100
GGAAAGCTAT TAAGAATGTG AGAACTGTGT ATGAGGAAAC TAATTTAATA 150
ATTCCTTAGA ATGGAACCAG TTGAAAATTT CCAGCTCCAC AAACCTGAAGT 200
GAAATCATTT TTTCTCCACT CCTTACTAGT AAATTTACTG TTCTATGTTA 250
AAAGAAAAAA AAAATCAACC AGCATTTAAA TTATGGCAAC CTAAAATGTG 300
TCCAGTATCT TAGAATAATT TCCCCACTGA CCTATTCCTC TGTAATAGTA 350
AAACATATAC ACAAATGTTT ATAGCTACAT TAGTCATAAT AGCCGAAAGG 400
TAAAAACAAC CCAAATGCCC ATCAACTAGA TAAATGTATT TAAAAAATAT 450
GACCCAGGCG AGGTGGCTCA GGCTGTAAAT CCCAGCACTT TAGGAGGCTG 500
AGGTGGGTGG ATGACCCAGG AGTTCAAGGC CAACCTGGTG AACATAGTGA 550
GACCCCATCT CTACAAAAC TAAAAATAAAA AATTAGCCAG ATGTTGTGGT 600
GTACACCTGT AGTCCAAGCT ACTCAGGACG GTGAGGAAGG AAGATCACTT 650
GAGCCCGGGA GTTTGAGGCT GCAGTGAGCT ATGATCACAC CATGGCACTC 700
CAGCCTGGGC AAGAAAGTGA GACCAAATTA TAAAAAAA AAAAAAAAAA 750
AAAAAAAAAA AAAAAACAGA AGAAGAAGCA CTGATGCATA GGCCATGAAT 800
    
```

AAACTTTGTA	AATATTATGC	TAAGTAAAAG	AAGCCAGAGA	TGAAAATCAC	850
ATATTGTAAT	TGTATGACTC	CATGTGTTTT	TTTAAAAAGG	TCCACACAGA	900
AAAGCTATTA	GTAGTTGCTC	ACAGCTGGAA	GGCAAGGAGG	GCACGTAAGT	950
GGGTGATAGC	TATAGGACAC	AAGGATTATT	TCTGAAATGA	TGAAAATGTT	1000
CTAAAACCGT	GGTAATGGTT	TTACAACCCT	GTGAATATAC	TAAAAACTAC	1050
TGAATTGTAT	ACTTAAAATG	GGTGAATTAG	ACGGCATATG	AATTATATAT	1100
CAATAAAGGT	ATTACCCAAG	AAAAAGAATA	CAGTATCTTC	ATATTCTATA	1150
TTCTCCTCTC	TTAGCTTTAC	TCAGATTTCA	CCTCTGTCCA	GTCACCTTTC	1200
CACATTAACT	CCAGGCAACT	CCAAAAGTTA	TTCTTCCTGC	TTCATTCATC	1250
CCCCAAATAA	ATTACATTCA	CTACTGCGAA	GATAACTGGC	CAGAAACTCA	1300
ATTCTGAAG	TTCTGGCAA	TGGTTCTTAG	ACTCCAAATG	GAGCAGAATA	1350
ATTTGCAACT	GGGCTTAAAC	ACGATTGTCT	TTTTTAAGGC	ATCCTCAGTT	1400

(2) NFORMACIÓN PARA SEQ ID NO:36

- 5 (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 1250 pb
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Doble
  - (D) TOPOLOGÍA: Circular
- 10 (ii) TIPO DE MOLÉCULA: ADN Genómico
- (iii) HIPOTÉTICO: no
- 15 (iv) FUENTE INMEDIATA:
  - (B) CLON: S189
- (viii) POSICIÓN EN GENOMA:
  - (A) CROMOSOMA/SEGMENTO: 22q11.2-qter
- 20 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:36

GTTGCTCTGG	CGATTTCGCA	CTCGAAAATG	ACACTTACTA	TTCAGCTAGA	50
GATTAGAATC	TCAAGCAGTA	GGGCATTTTT	TAATAAAAAA	TTAAATTAAA	100
AATAGATTTG	CCATTGTCTG	CTTAATAAAA	CTAGTAGCTC	TGCTGGCTTA	150
GAGGGGAAAT	AACATATTTT	TTCGGATTTT	TATATATTTA	TCTGAGCAGT	200
GCTAAAAAAT	AAAACAAAGT	TACTAATATT	CATATCTTGA	GCAATTGTAC	250
ATTGCTTCTA	ACTATACATT	CAATCTCTCT	GGCACATCCA	CTGTGGCCCT	300
GAGCAGCCAG	TACAGGCTCT	TCTACCAAAA	CGAAGCAAGC	CACTCCAAAA	350
CCTGACGCGT	GCAGGTGTCA	CGAAACACCA	GGTGCAGCTT	GACAGATGTG	400
AGCCAAATAA	GGAAACATTC	AGCCCAGCAC	TGCCCAACAG	TCATGATGTA	450
TATTTTCTAC	ATCTGTGCTC	TAAAATATGG	TGGCCACTAG	CTGCAGGTGG	500
CTATTGAGAC	TAAGGAACTG	TATTTTAAAT	TTTATTTTCAT	TTCAACTCAT	550
TTAAAGTAGC	CACATGCCGC	TAATGGCTAC	TGATCTAGAG	GGCAGCTGGG	600
ATGTTACTCT	TGAGAATGTC	TCCAGCATT	TACCTGTTGC	TCTCTCTCAC	650
TCACATTTCC	CATTCTAGCA	CAAACAAAAC	AAAACAAAAC	AAAACAAAAC	700
AAAACAAAAC	AAAACAAAAA	AACCACAACA	CCTACAGTTC	TCCAAACAGG	750
GCATCTGTTT	TGTTCCCTCG	GGGGGGTCTT	GTCTATGTTG	TTACAGTGGC	800
CCTGGATTTT	CATACTCCTA	GCCTTCCTGG	AAGACATCCT	TTTCATCCTC	850
ACAACCCAAC	CCAGGCTTTA	TCTCTTCTGT	GAAGCTGTCC	TTGATTTTCC	900
GTTCTATCTT	CCCTGCTTGT	GAATGGGTCA	GCTCTCCTTC	CCCACCGCCC	950
TGTGCGTGTG	AACATCTTTG	TTCAGTATAC	TGCAGTGGGT	CGGGAGTATG	1000
TCCCTTCCAG	ACTGGAAGGC	AGAGAGGGTG	GCTGTAAGGA	TTGGCACTTT	1050
GGGCCAGGCA	CAGTGCTCAT	GCCTGTAATC	CCAGCACTTT	GGGAGGCTGA	1100
GGCAGGAGAA	TGCCCCGACC	CCAGGAGACA	GAGTTTGCAG	TGAGACGAGA	1150
TTGCACCACT	GCACTCTAGC	CTGAGGGATA	GAGCAAGACT	CCCTCTCAAA	1200
AAAATAAATA	AATAAATAAA	TAAATAAATA	AAAATAAAAA	ATTAAAGAGG	1250

25 (2) INFORMACIÓN PARA SEQ ID NO:37

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 1200 pb

- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

- 5 (ii) TIPO DE MOLÉCULA: ADN Genómico
- (iii) HIPOTÉTICO: no
- (iv) FUENTE INMEDIATA:  
10 (B) CLON: S199
- (viii) POSICIÓN EN GENOMA:  
(A) CROMOSOMA/SEGMENTO: 6q21
- 15 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:37

```

TTTCATGTTT ACAGATTGGA TTAATATTGT TAAACTGTCC ATACTACCCA 50
AAGCAATCCA TAGATTC AAT GCAACCCCTGA TATAGTTTGA ATGTATGTAG 100
GCACCAAAAT CTCATGTTGA ACTTTAATCC CCAGTGTGG AGGTGGAGCC 150
TGGTGTAAAG TGTTTAGATT ATGAAGGTGA ATCCCTCATG AACGGCTTGG 200
GCCATCTGCT TGGTGATAAG TGAGCTCTTG TTCTGAGTTC ACATGAGATA 250
CAGTCATTTA AAAGCCTGTG GTACCCAAAC TCTCTCTTGC TCTTGCTTCT 300
GTTACAGCCA TGTGATATAC CTGCTATCCT TTGCCTTTGC CTTCTGCCAT 350
GATTGGAAGC TTCCTGAGTC CTCCCCAGAA ACAGATGTAA CTATGCTTCC 400
TGTACAGCCT GCAGAACCAA GAACAAACTG AAACCTTTTT GTTATAAATT 450
GCCCAGGATT AGGTGGGTGT TTTGTTTTGT TTTGTTTTGT TTTGTTTTGT 500
TTTTTGAGAT GGAGTCTCGC TCTGTCTCCC AGGCTGGAGT GCAATGATAC 550
AATCTCGGCT CACTGCAACC FCCACCTCCC CGTTCAGCA ATTCTCCTGC 600
CTCAGCCTCC TGAGTAGCTG GGATTACAGG CGCACGCCAT CATGCCCGGC 650
TAATTTTTGT ATTTTTAGTA GAGACGGGT TTCACCACAT TGGTCAGGCT 700
GGTCTCGAAC TCCTGACCTC ATGATCCACC CGCCTTGGCC TCCCAAAGTG 750
CTGGGATFAC AGGCCTAAGC CACCATGCCC AGCCAGGTGG TTTTTTATAG 800
TAGTGCAAGA ATGGCCGAAT ACAAACCCCT ATCAAATAC CAATGACATT 850
TGTCAGGGAC ATTTTTAAA ATTCTGAAAT TTATATGGAA CCACAAAAGA 900
CCCAGAATAG CCAAACATA CCTGAGCAA AAGAACAAC CTGGAAGAAT 950
CACATTACCT GACTTCAAAG TGTACTACAG AGCTCTTATA ATCAAACAT 1000
CATGGTACTA GCATAACAAC AGACACATAG ACCAATGGAA CACAATAGAG 1050
AACCAGAAA CAAATCCATA CACCTACTGT GAACTCATTT TTGACAAAGG 1100
TGCCAAGAAC ATACATGGGA GAAAGGACAG TATCTCCAA TAAATGGTGC 1150
TGAGAAAAGT GGATATACAT ATGCATAAGA ATGTAAC TAG ACCCCTATCT 1200
    
```

- 20 (2) INFORMACIÓN PARA SEQ ID NO:38
- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
25 (A) LONGITUD: 1000 pb  
(B) TIPO: Ácido Nucleico  
(C) TIPO DE CADENA: Doble  
(D) TOPOLOGÍA: Circular
- (ii) TIPO DE MOLÉCULA: ADN Genómico
- (iii) HIPOTÉTICO: no
- (iv) FUENTE INMEDIATA:  
30 (B) LON: S040
- 35
- 40

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:38

```
GCTGCAATAA ACACGGGAGT GTAGGTATCT TTAAAAGAAG GTGGTGATTT 1900
CATCTCTTCT GGGTATGTAT CCAAAATAGG GTCACCTGTTG GGTTATAAGG 1950
TGGTTAGGTT TTGAATTTCT TTAGGAACCT CCATACTGTT TTCCATAATG 2000
GGTGCACCAA TCATCATTCC CACCAACAAT GTACAAGTGT TTTATTTTCT 2050
TCACACCCTC ATCAATATTT ATCTCTTGTC TTTTATAA TAGCCATCCT 2100
AAAGACTGTA AGGCGTTTTA TTTCTAATCT CAGATTTTAC TGTAGAAACA 2150
GTGATGACAC AGTCTCCAGC TTCCCTGTCT TTGTCTCTGG AGAAAAAAGC 2200
CACCCCTGACT TGCAGGGCCA GTCAGTGTTA GCAGCTACTA AGCCTGGTAC 2250
CAGAAGAAAC CTGAGCGGGT TCCCAGGCTC CTCATCTATG GTACAGCCCT 2300
GATTTGTGAT AGTGGGTCGG GGACAGGGCT TACTCTCACC ATCGGCAGCC 2350
TGGAGCCTGG AGCCTGGAGA TTTGCACFTC ATCACTGTTA TCAGCATAGT 2400
AGTTGGTGTG CCAATATGAT TCGACATGCA ACAAAAACCT CCAGGAGACC 2450
TAAGGTGTTT ATTTGATTAT ACTACCTGCT TCCTTTTTAG TCATCTGATG 2500
TGGTGCTGCT CAGTTTTAGC ATCTCTGCTT TGATTGGAAA TTCTGAGGTT 2550
CTCAAAGTA ATTCTTATA ATATTTATAG TTTCACTCAT GGATTTTTTT 2600
CTCAGACCCA AATGTACAGC CAGGTTGAGG CACAATTTCA TGGTCAAGGC 2650
CATTGGATCA GACTCACATG AGTGGACGCC TCTAAAGGTC CTGGCCAGTG 2700
CGATAAAGTA GCAGCGACAA TGATAAAGAA GAAGAATTAG AAAGGCAGAA 2750
TTAAAGGTAT AACCAATTCAC TGATGAAAGG ACTGTGTGGG GGAGAAATTT 2800
CTAATTGTCT ACACAGAAAT TATTAGAATT AATGAGATAC ATAGCAAATT 2850
```

5

(2) INFORMACIÓN PARA SEQ ID NO:39

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1050 pb
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Doble
- (D) TOPOLOGÍA: Circular

10

(ii) TIPO DE MOLÉCULA: ADN Genómico

15

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (A) BIBLIOTECA:
- (B) CLON: S066

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39

```
GGGGCCTAGC CCAGTTGGAG GGACAAGAGC TGGAAACTGG GTTCCTTAGG 50
GTGGTGCCAG AGTGGGCAGA GACCTCTGGG CAGCCACGTT CCAAGTCCAG 100
AGCAAGGGGA GGCTCATCCT AGAAAAGAGG CCAGAGGAGC CATAACCACC 150
ATTGTTTCCTT GGGTTAAGGA GTCCTTTTTT AAAACCATCA AACTAAGAA 200
TCCAGTGCAT TATGAATCCA AGGGGTGAGG CTCAGTGTGC CAATGCCCCA 250
GAACAGTCTA AGAAAGCTCC TTTTCCCTTT CCAGGCAGCT CGAGCTTTAC 300
CTTCCCAAAT TCTCCATTGA GGGCTCCTAT CAGCTGGAGA AAGTCCTCCC 350
CAGTCTGGGG ATCAGTAACG TCTTACCTC CCATGCTGAT CTGTCCGGCA 400
TCAGCAACCA CTCAAATATC CAGGTGTCTG AGGTGGGTTT AGAAGCTCCT 450
ATGCATCTGC TTCCCAAGAT CTATTCTGTT CTATTCTTTC TATTCTACTC 500
TACCCCATTT CATTCATTTC CATTCCACTC AACTCCACTC CACTCCACTC 550
CACTCCAGTT CACTCTATT CATTCCACTC CACTCCAGTT CACTCTATT 600
AATTCCACTC CACTCCACTC CAGTTCACTC TATTCCAGTT CACTCCACTC 650

CACTCCACTC CACTCCAGTT CACTCTATT CATTCCACTC CATTCCACTC 700
CTCCACTCCT CTCATCCACT CCACTCTACT CCTCCACTCC ACATCTCCAC 750
TCCACTCCTC CACTCCACTC CTCCACTCCA CTCATCCACT CCACTCCTCC 800
ACTCCACTCC TCCACTCCAC TCCTCCACTC CACTCCACTC ATCCACTCCA 850
CTCTTCCATT CCACTCCATT CCACTCCTCC ACTCCACTCT TCCACTCCAC 900
TCCATTCCAC TCCCTCCACT CACTCCACTC TATTCTATT CATTCCATT 950
CACTCTACTC TATTCTATT CATTCCATT CAGTCAACT CACTCCACTC 1000
TCTACTATT CATTCCACTC CTCTCCCTC CACTCCATT CATTGCAGTC 1050
```

25

(2) INFORMACIÓN PARA SEQ ID NO:40

- 5 (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 4580 pb
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Doble
  - (D) TOPOLOGÍA: Circular
- 10 (ii) TIPO DE MOLÉCULA: ADN Genómico
- (iii) HIPOTÉTICO: no
- 15 (iv) FUENTE INMEDIATA:
  - (A) BIBLIOTECA:
  - (B) CLON: S077
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:40

```

GGATCCCAAT TCATTCCGGG CTGACACGCT CACTGGCAGG CGTCGGGCAT 50
CACCTAGCGG TCACTGTTAC TCTGAAAACG GAGGCCTCAC AGAGGAAGGG 100
AGCACCAGGC CGCCTGCGCA CAGCCTGGGG CAACTGTGTC TTCTCCACCG 150
CCCCCGCCCC CACCTCCAAG TTCCTCCCTC CTTGTTGCC TAGGAAATCG 200
CCACTTTGAC GACCGGGTCT GATTGACCTT TGATCAGGCA AAAACGAACA 250
AACAGATAAA TAAATAAAAT AACACAAAAG TAACTAACTA AATAAAATAA 300
GTCAATACAA CCCATTACAA TACAATAAGA TACGATACGA TAGGATGCGA 350
TAGGATACGA TAGGATACAA TACAATAGGA TACGATACAA TACAATACAA 400
TACAATACAA TACAATACAA TACAATACAA TACAATACAA TACAATACGC 450
CGGGCGCGGT GGCTCATGCC TGTCATCCCG TCACTTTGGG ATGCCGAGGT 500
GGACGCATCA CCTGAAGTCG GGAGTTGGAG ACAAGCCCGA CCAACATGGA 550
GAAATCCCCT CTCAATTGAA AATACAAAAC TAGCCGGGCG CGGTGGCACA 600
TGCCCTATAAT CCCAGCTGCT AGGAAGGCTG AGGCAGGAGA ATCGCTTGAA 650
CCTGGGAAGC GGAGGTTGCA GTGAGCCGAG ATTGCGCCAT CGCACTCCAG 700
TCTGAGCAAC AAGAGCGAAA CTCCGTCTCA AAAATAAATA CATAAATAAA 750
TACATACATA CATAACATA TACATACATA CATAACATA TAAATTAATA 800
TAAATAAATA AAATAAATA AATAAATGGG CCCTGCGCGG TGGCTCAAGC 850
CTGTCTATCC CTCACCTTGG GAGGCCAAGG CCGGTGGATC AAGAGGCGGT 900
CAGACCAACA GGGCCAGTAT GGTGAAACCC CGTCTCTACT CACAATACAC 950
AACATTAGCC GGGCGCTGTG CTGTGCTGTA CTGTCTGTAA TCCAGCTAC 1000

```

20

(2) INFORMACIÓN PARA SEQ ID NO:41

- 25 (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 1000 pb
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Doble
  - (D) TOPOLOGÍA: Circular
- 30 (ii) TIPO DE MOLÉCULA: ADN Genómico
- (iii) HIPOTÉTICO: no
- 35 (iv) FUENTE INMEDIATA:
  - (A) BIBLIOTECA:
  - (B) CLON: S097

40

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:41

	GGACATGAGG	CTTCCCAGCC	AACTGCAGGT	GCACAACATA	AATGTATCTG	50
	CAAACAGACT	GAGAGTAAAG	CTGGGGGCAC	AAACCTCAGC	ACTGCCAGGA	100
	CACACACCCCT	TCTCGTGGAT	TCTGACTTTA	TCTGACCCGG	CCCACTGTCC	150
	AGATCTTGTT	GTGGGATTGG	GACAAGGGAG	GTCATAAAGC	CTGTCCCCAG	200
	GGCACTCTGT	GTGAGCACAC	GAGACCTCCC	CACCCCCCA	CCGTTAGGTC	250
	TCCACACATA	GATCTGACCA	TTAGGCATTG	TGAGGAGGAC	TCTAGCGCGG	300
	GCTCAGGGAT	CACACCAGAG	AATCAGGTAC	AGAGAGGAAG	ACGGGGCTCG	350
	AGGAGCTGAT	GGATGACACA	GAGCAGGGTT	CCTGCAGTCC	ACAGGTCCAG	400
	CTCACCCCTGG	TGTAGGTGCC	CCATCCCCCT	GATCCAGGCA	TCCCTGACAC	450
	AGCTCCCTCC	CGGAGCCTCC	TCCCAGGTGA	CACATCAGGG	TCCCTCACTC	500
	AAGCTGTCCA	GAGAGGGCAG	CACCTTGGAC	AGCGCCCACC	CCACTTCACT	550
	CTTCTCCCT	CACAGGGCTC	AGGGCTCAGG	GCTCAAGTCT	CAGAACAAAT	600
	GGCAGAGGCC	AGTGAGCCCA	GAGATGGTGA	CAGGGCAATG	ATCCAGGGGC	650
	AGCTGCCTGA	AACGGGAGCA	GGTGAAGCCA	CAGATGGGAG	AAGATGGTTC	700
	AGGAAGAAAA	ATCCAGGAAT	GGGCAGGAGA	GGAGAGGAGG	ACACAGGCTC	750
	TGTGGGGCTG	CAGCCCAGGA	TGGGACTAAG	TGTGAAGACA	TCTCAGCAGG	800
	TGAGGCCAGG	TCCCATGAAC	AGAGAAGCAG	CTCCCACCTC	CCCTGATGCA	850
	CGGACACACA	GAGTGTGTGG	TGCTGTGCCC	CCAGAGTCGG	GCTCTCCTGT	900
	TCTGGTCCCC	AGGGAGTGAG	AAGTGAGGTT	GACTTGTCCC	TGCTCCTCTC	950
5	TGCTACCCCA	ACATTCACCT	TCTCCTCATG	CCCCTCTCTC	TCAAATATGA	1000

(2) INFORMACIÓN PARA SEQ ID NO:42

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD : 994 pb  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Doble  
 (D) TOPOLOGÍA: Circular

- (ii) TIPO DE MOLÉCULA: ADN Genómico

- (iii) HIPOTÉTICO: no

- (iv) FUENTE INMEDIATA:  
 (B) CLON: S103

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:42

	CTCTGACTCT	CCGCGGTGGT	TGTTGGGGCT	TCTTGGCTTT	GTTTTGTTGT	50
	TTGTTTGAT	TTTATTTTTT	TCTCTCTGAC	ACCTATTTTA	GACAAATCTA	100
	AGGGAAAAAG	CCTTGACAAT	AGAACATTGA	TTGCTGTGTC	CAACTCCAGT	150
	ACCTGGAGCT	TCTCTTTAAC	TCAGGACTCC	AGCCCATTTG	TAGACGTGTG	200
	TTTCTAGAGC	CTGCTGGATC	TCCCAGGGCT	ACTCACTCAA	GTTCAAGGAC	250
	CAACAAGGGC	AGTGGAGGTG	CTGCATTGCC	TGCGGTCAAG	GCCAGCAAGG	300
	TGGAGTGGAT	GCCTCAGAAC	GGACGAGATA	ATGTGAACTA	GCTGGAATTT	350
	TTTATTCTTG	TGAATATGTA	CATAGGCAGC	ACTAGCGACA	TTGCAGTCTG	400
	CTTCTGCACC	TTATCTTAAA	GCACCTACAG	ATAGGCCTTC	TTGTGATCTT	450
	GCTCTATCTC	ACAGCACACT	CAGCACCCCC	TTCTCTGCCC	ATTCCCCAGC	500
	CTCTCTTCT	ATCCCATCCC	ATCCCATCCC	ATCCCATCCC	ATCCCATCCC	550
	GCTCTTTTCC	TACTTTTCCCT	TCCCTCAAAG	CTTCCATTCC	ACATCCGGAG	600
	GAGAAGAAGG	AAATGAATTT	CTCTACAGAT	GTCCCATTTT	CAGACTGCTT	650
	TAAAAAAAAT	CCTTCTAATC	TGCTATGCTT	GAATGCCACG	CGGTACAAAG	700
	GAAAAAGTAT	CATGGAAATA	TTATGCAAAT	TCCCAGATTT	GAAGACAAAA	750
	ATACTCTAAT	TCTAACCAGA	GCAAGCTTTT	TTATTTTTTA	TACAGGGGAA	800
	TATTTTATTC	AAGGTAAAA	TCTAAATAAA	ATATAATTGT	TTTTTATCTT	850
	TTCTACAGCA	AATTTATAAT	TTTAAGATTTC	CTTTTCTTGT	TTATCAGCAG	900
	TTGTTATTAC	ATCCTTGTGG	CACATTTTTT	TTTAATTTTG	TAAAGGTGAA	950
25	AAAAGCTTTT	ATGAGCTCAT	CTAGCAATCA	GATTTTCTTG	TGGA	994

(2) INFORMACIÓN PARA SEQ ID NO:43

- 5 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 13 66 pb  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Doble  
 (D) TOPOLOGÍA: Circular

10 (ii) TIPO DE MOLÉCULA: ADN Genómico

(iii) HIPOTÉTICO: no

15 (iv) FUENTE INMEDIATA: (B) CLON: S110

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:43

```
GGAATTC AAT GGAATATAAC GAAATGGATA GGATCAGAAC GGAACAGAGC 50
GGAGTGGAGT TGAGTGGAGT GGATCGGAGT GCAGTGGAAA GGAATGGAAT 100
AGAATGGAAT GGAATGCAGT GGAGTGGAAAT GGAATGAAGT GGAATGGAAT 150
TGAGTGGAGT GGATCGGAGT GCAGTGGAAA GGAATGGAAG AGAATGGAAT 200
GGAATGGAAT GCAGTGGACT GGAATGGAAT GGAGTGGAGT GGAGTGCAGT 250
GGGAATCGAG TGGAGTGGAG TGGGAATGGAC TGGGAATGGAA TGGATTGGAG 300
TGGAGTGCAG TGGGAATCGAG TGGAGTGGAG TGGGAATGGAG TAGAATGGAA 350
TGGAGTGGAG TGTAGTGGAA TGGGAATGGAA TGGTGAATGA ATGTCAGCTA 400
AGATTGTGCA ACTGCATTCC AGTCTGGGTG ACAAAGTGAG ATCCAGTCGA 450
AGTAAAGGAA TGGGAATGGAA TAGAGTAAAA TGGGAATGGAA TGGTGTGGAG 500
TGGGAATGGAA TGGAGAGGAA TGGAGTGGAG TGGAGTGGAG TGGAGTGGAA 550
TGGAGTGGAG TGGGAATGGAG AGTGATGGAG AGGAATGGAA TGGGAATGGAA 600
TGGGAATGGAG TGGGAATGGAA TGGGAATGGAG TGGGAATGGAA TGGGAATGTAG 650
AGGAGTGGAG TGGATTGGAG TGGAGTGGAA TGGAGTGGAA TAGAGTGAAA 700
TTTAGTGGAG TGTAAATGGAG TGGAGTGGAG TGGCAGTTGA GTGGCATGGA 750
TCAGGTGCAG TGGGAATGGAA TGGGAATGGAG TGGAGTGGAG AGGAGTGGAG 800
TGGGAATCGAA TGGGAATGGCA TGGAGTGGAG TGGGAATGGAG TGGATTGGAA 850
TTGAATGCAG TGGGAATGGAA TGGGAATGGAG TGGAGTGGAG TGGAGTGGAG 900
TGGAGTGGAG GGGGAATGGAA TGGAGTGGAG TAAAATGGTT TGGGAATGGAG 950
TGGGGTGGAA TGGAGTGGGT TGGGAATGGAG TGGAGTGGAG TAGAACGGAG 1000
TGATTGGGGT GGAATGGAAT AGAGTGGAAAT GGAATGGAGT GGAGTGGAGT 1050
AGAACGGAGT GATTGGAGTG GAATGGAATA CAGTAGAGTG GAATGCAGTG 1100
GAGTGGAAATG GAATGGAGTG GAGTGGCATG GAAAGGAATG GAGAGGAATG 1150
GAATGGAATG GAATGGAAATG GAATGGAATG GAATGGAATG GAACGGTGAA 1200
ATAAAATGTG AGTTAAGATT GTGCCACTGC ATTGCAGTCT GGGGGACAGA 1250
GTGAGATACA GTCGAAATA AGGAATGGAA GGGACTGGAG TAGAATGGAA 1300
TGGGAATTGAG TGGAGTGGAA TGGGAATGAAG TGGAGAGGAA TGGGAATGGAG 1350
TGGGAATGCAA TGGAGG 1366
```

20 (2) INFORMACIÓN PARA SEQ ID NO:44

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 19  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

30 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:44

**TGGCTCAGAC ACCTCATTG 19**

(2) INFORMACIÓN PARA SEQ ID NO:45

- 35 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 21  
 (B) TIPO: Ácido Nucleico

- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:45

5

**CACCACTGTA TTCCCAGTTT G**

**21**

(2) INFORMACIÓN PARA SEQ ID NO:46

10

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 22
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

15

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:46

**CACTTGCCAT CCCTGCCACA CA**

**22**

20 (2) INFORMACIÓN PARA SEQ ID NO:47

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 23
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:47

30

**AGCGCACCCC CAATTTCCGG TAT**

**23**

(2) INFORMACIÓN PARA SEQ ID NO:48

35

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 22
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

40

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:48

**TGGGGACATG AACACACTTT GC**

**22**

(2) INFORMACIÓN PARA SEQ ID NO:49

45

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 22
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:49

**GAGGCCAGG ACCAGATGAA AT**

**22**

55

(2) INFORMACIÓN PARA SEQ ID NO:50

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 23
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla

60

(D) TOPOLOGÍA: Lineal

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:50

**CACCTGTCAG GCAAGGCTTA AAC**

**23**

10 (2) INFORMACIÓN PARA SEQ ID NO:51

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24

(B) TIPO: Ácido Nucleico

(C) TIPO DE CADENA: Sencilla

(D) TOPOLOGÍA: Lineal

15

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:51

**CAACACTGAG CGCTTTTAGG GACT**

**24**

20

(2) INFORMACIÓN PARA SEQ ID NO:52

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24

(B) TIPO: Ácido Nucleico

(C) TIPO DE CADENA: Sencilla

(D) TOPOLOGÍA: Lineal

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:52

30

**TCAGGCAAGG CTTAAACAGG GATA**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:53

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24

(B) TIPO: Ácido Nucleico

(C) TIPO DE CADENA: Sencilla

(D) TOPOLOGÍA: Lineal

35

40

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:53

**ACACTGAGCG CTTCTAGGGA CTTC**

**24**

45 (2) INFORMACIÓN PARA SEQ ID NO:54

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24

(B) TIPO: Ácido Nucleico

(C) TIPO DE CADENA: Sencilla

(D) TOPOLOGÍA: Lineal

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:54

55 **TGAGCGCTTC TAGGGACTTC TTCA**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:55

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 18

(B) TIPO: Ácido Nucleico

(C) TIPO DE CADENA: Sencilla

60

(D) TOPOLOGÍA: Lineal

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:55

**CCCTGCCCTA CCCACTTG**

**18**

10 (2) INFORMACIÓN PARA SEQ ID NO:56

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 18
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

15

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:56

**AGGCCAGGA CCAGATGA**

**18**

20

(2) INFORMACIÓN PARA SEQ ID NO:57

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:57

30

**GCACCTGTCA GGCAAGGCTT AAAC**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:58

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 22
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

35

40

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:58

**CCAGCCATGA AGTGGCTGTG AG**

**22**

45 (2) INFORMACIÓN PARA SEQ ID NO:59

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 22
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:59

**CCCGCTTCAA AGTTCCCAGT TC**

**22**

55

(2) INFORMACIÓN PARA SEQ ID NO:60

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 21
  - (B) TIPO: Ácido Nucleico

60

- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:60

5 **CCTCCCATTT CAGCCTCCTG A** **21**

(2) INFORMACIÓN PARA SEQ ID NO:61

- 10 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

15 **GTCTGCCACA GTGCTGGAAA CTAA** **24**

(2) INFORMACIÓN PARA SEQ ID NO:62

- 20 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 21
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:62

**GCACCCGAGC CTAAGGCAAT A** **21**

30 (2) INFORMACIÓN PARA SEQ ID NO:63

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 18
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

35 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:63

40 **GCATGGCGGA AGAAACAA** **18**

(2) INFORMACIÓN PARA SEQ ID NO:64

- 45 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 19
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

50 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:64

**TGGCAACAGA GCGAGACTC** **19**

(2) INFORMACIÓN PARA SEQ ID NO:65

- 55 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 21
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:65

**CCTGGGTGAC AGCGAGAATC T**

**21**

5 (2) INFORMACIÓN PARA SEQ ID NO:66

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 24  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:66

**TGTCCCTTGC CTTGTCTCAC TAAA**

**24**

15

(2) INFORMACIÓN PARA SEQ ID NO:67

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 21  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

20

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:67

**CAGCCTTGGT GACAGAGCAA A**

**21**

30

(2) INFORMACIÓN PARA SEQ ID NO:68

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 21  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:68

**TGTGTTGAGG GTGGGGTACA T**

**21**

40

(2) INFORMACIÓN PARA SEQ ID NO:69

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 18  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 69

50

**CCTGGGCAAG AGAGCAAG**

**18**

(2) INFORMACIÓN PARA SEQ ID NO:70

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 21  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

55

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 70

**CACATCCCAA AACCACCCTA C**

**21**

5 (2) INFORMACIÓN PARA SEQ ID NO:71

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 20  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO.-71

**GCATTTCCCC TGCTTGIACT**

**20**

15

(2) INFORMACIÓN PARA SEQ ID NO:72

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 24  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

20

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 72

**GATCACATTT GCTAACCCT TCTC**

**24**

30

(2) INFORMACIÓN PARA SEQ ID NO:73

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 26  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 73

**GGCAACATAT CAAGACCCCC ATCTCT**

**26**

40

(2) INFORMACIÓN PARA SEQ ID NO:74

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 26  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

45

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 74

**GAAGCTGCCC CTCACCACTA CATTTT**

**26**

55

(2) INFORMACIÓN PARA SEQ ID NO:75

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 24  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:75

**GATCACATTT GCTAACCACT TCTC**

**24**

5 (2) INFORMACIÓN PARA SEQ ID NO:76

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:76

**TATAAATTAC CCAGTCTCAG GAAG**

**24**

15

(2) INFORMACIÓN PARA SEQ ID NO:77

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 20
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

20

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:77

**GTGATACAGC AAGCCTCATC**

**20**

30

(2) INFORMACIÓN PARA SEQ ID NO:78

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:78

**AGAGACTCCT GGAAAGATAA AAGT**

**24**

40

(2) INFORMACIÓN PARA SEQ ID NO:79

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

45

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:79

**GTCTGGAGAA CAGTGGCCCT TGT**

**23**

(2) INFORMACIÓN PARA SEQ ID NO:80

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

55

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:80

**CAGGAAGCTG AGGCAGGAGA ATCT**

**24**

5 (2) INFORMACIÓN PARA SEQ ID NO:81

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 18
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:81

**AAGGCTCCAG TGGGGTAT**

**18**

15

(2) INFORMACIÓN PARA SEQ ID NO:82

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

20

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:82

**AAAACAAGGC AGTAGTCAAT AAAG**

**24**

30

(2) INFORMACIÓN PARA SEQ ID NO:83

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD:23
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:83

**GGCATGAGAA TCGCTTGAAC CTG**

**23**

40

(2) INFORMACIÓN PARA SEQ ID NO:84

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:84

50

**GGCCTCCATG ATGTTTCCAA TGAT**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:85

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

55

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:85

**TCAGGAGGCA TGAGAATCGC TTGA**

**24**

5 (2) INFORMACIÓN PARA SEQ ID NO:86

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:86

**GGCCTCCATG ATGTTTCCCA ATGA**

**24**

15

(2) INFORMACIÓN PARA SEQ ID NO:87

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

20

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:87

25

**CTCGCCCTCT CCTATAAGCA GTTT**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:88

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

30

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:88

**GCAGAGATAA TTTGGAGTGG GATG**

**24**

40

(2) INFORMACIÓN PARA SEQ ID NO:89

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 18
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:89

50

**CTTGGGTGCC TGTAATCC**

**18**

(2) INFORMACIÓN PARA SEQ ID NO:90

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 18
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

55

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:90

**GGTAGAGCTC CCCCATCT**

**18**

5 (2) INFORMACIÓN PARA SEQ ID NO:91

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 22
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:91

**GCAGAATATT GGGGCTCATC AC**

**22**

15

(2) INFORMACIÓN PARA SEQ ID NO:92

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

20

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:92

**AAACAAGGAA AGGAGAGGAG AGGA**

**24**

30

(2) INFORMACIÓN PARA SEQ ID NO:93

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:93

**AAGGTTGTGG GATGACTACT ACA**

**23**

40

(2) INFORMACIÓN PARA SEQ ID NO:94

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 21
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:94

50

**TGGTCAACAC AGCAAGACAT T**

**21**

(2) INFORMACIÓN PARA SEQ ID NO:95

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 22
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

55

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:95

**TCCTGCCACC TGCTTGCTTT CT**

**22**

5 (2) INFORMACIÓN PARA SEQ ID NO:96

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:96

**ATTGCACTCC AGCCTGGGTGA TAC**

**23**

15

(2) INFORMACIÓN PARA SEQ ID NO:97

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 20
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

20

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:97

**CGCTTGAGCC TTGGAGATTG**

**20**

30

(2) INFORMACIÓN PARA SEQ ID NO:98

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:98

**GAGCAGTCAG AATTCAGGAG TTGT**

**24**

40

(2) INFORMACIÓN PARA SEQ ID NO:99

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

45

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 99

**TGGGCAACAA GAGCAAACT CCAT**

**24**

55

(2) INFORMACIÓN PARA SEQ ID NO:100

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:100

**GGGACTTGGG CTGAGGGCTT TAC**

**23**

5

(2) INFORMACIÓN PARA SEQ ID NO:101

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:101

15

**ATATCAATAT CAGGCAGCCA CAGG**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:102

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

20

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:102

**CCGTTTCAGA GCAGAGGTTT AGC**

**23**

30

(2) INFORMACIÓN PARA SEQ ID NO:103

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ.ID NO:103

**TCTCATTGGT TTCAAAGAAC TTA**

**23**

40

(2) INFORMACIÓN PARA SEQ ID NO:104

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

45

(A)

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:104

**AGACTCCATC TCAAACAAAA GA**

**23**

55

(2) INFORMACIÓN PARA SEQ ID NO:105

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:105

**TCATGTGCAT GGAGCCTGGT TCAT**

**24**

5

(2) INFORMACIÓN PARA SEQ ID NO:106

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:106

15

**CCCAGCCTTG GCAAGAGTGA GGT**

**23**

(2) INFORMACIÓN PARA SEQ ID NO:107

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 18
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

20

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ IDNO:107

25

**GGCGACTGAG CAAGACTC**

**18**

(2) INFORMACIÓN PARA SEQ ID NO:108

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 22
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

30

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:108

**TTAAGCAAAG TAGCCTCAAA CA**

**22**

40

(2) INFORMACIÓN PARA SEQ ID NO:109

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 19
- (B) TIPO: Nucleic Acid
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109

50

**GGGCGACTGA GCAAGACTC**

**19**

(2) INFORMACIÓN PARA SEQ ID NO:110

(i) SEQUENCE CHARACTERISTICS:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

55

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:110

**ACTCATTACC TTGCATGCAT GATA**

**24**

5

(2) INFORMACIÓN PARA SEQ ID NO:111

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 21
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

10

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:111

15

**CATTACCTTG CATGCATGAT A**

**21**

(2) INFORMACIÓN PARA SEQ ID NO:112

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 21
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

20

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:112

**TGGGCAACAG AGTAAGACTC A**

**21**

(2) INFORMACIÓN PARA SEQ ID NO:113

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

30

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:113

**GTTCAGTACC GTTCACCTCT TTA**

**23**

40

(2) INFORMACIÓN PARA SEQ ID NO:114

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 30
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

45

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:114

**GTAAGACTCA GTCTCCAAA AAAAAAAAAAG**

**30**

55

(2) INFORMACIÓN PARA SEQ ID NO:115

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 38
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:115

**AGGAATGGTT TCTCTGTTAG TAAATGGT**

**38**

5

(2) INFORMACIÓN PARA SEQ ID NO:116

(i) CARACTERÍSTICAS DE LA SECUENCIA:

10

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

15

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:116

**CAGCCTGGGC AACAGAATG AAAC**

**24**

20

(2) INFORMACIÓN PARA SEQ ID NO:117

(i) CARACTERÍSTICAS DE LA SECUENCIA:

25

- (A) LONGITUD: 19
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

25

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:117

**TGGCCCCTGC AGCGGAGTC**

**19**

30

(2) INFORMACIÓN PARA SEQ ID NO:118

(i) CARACTERÍSTICAS DE LA SECUENCIA:

35

- (A) LONGITUD: 21
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

40

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:118

**GAATTCATTT GCGGAAAGAT T**

**21**

45

(2) INFORMACIÓN PARA SEQ ID NO:119

(i) CARACTERÍSTICAS DE LA SECUENCIA:

50

- (A) LONGITUD: 21
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

50

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:119

**CTAGGGAGGC TGGAGTATTC A**

**21**

55

(2) INFORMACIÓN PARA SEQ ID NO:120

(i) CARACTERÍSTICAS DE LA SECUENCIA:

60

- (A) LONGITUD: 20
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

60

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:120

**AGAGCAAGAC CCCGTCTCAT**

**20**

5

(2) INFORMACIÓN PARA SEQ ID NO:121

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 20
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

10

15

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:121

**AGTCCATGGG CCTTTTAAACA**

**20**

20 (2) INFORMACIÓN PARA SEQ ID NO:122

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

25

(ii) TIPO DE MOLÉCULA: Oligonucleotide ssDNA

30

(iii) HIPOTÉTICO: no

(iv) FUENTE INMEDIATA:

- (B) CLON AMPLIFIED: S125

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:122

**GAGAATCACT TGAACCCAGG AAG**

**23**

40 (2) INFORMACIÓN PARA SEQ ID NO:123

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

40

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:123

**AGAACCAGCT GTTAGTTTCG TTGA**

**24**

50

(2) INFORMACIÓN PARA SEQ ID NO:124

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 25
- (B) TIPO: Ácido Nucleico
- (C) TIPO DE CADENA: Sencilla
- (D) TOPOLOGÍA: Lineal

55

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:124

60

**GGTTGCAGTG AGCCGAGATA AGAGT**

**25**

5	(2) INFORMACIÓN PARA SEQ ID NO:125	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 25	
	(B) TIPO: Ácido Nucleico	
	(C) TIPO DE CADENA: Sencilla	
	(D) TOPOLOGÍA: Lineal	
10	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:125	
	<b>TGTGCCAGGA ACCAGAAATT TACAG</b>	<b>25</b>
15	(2) INFORMACIÓN PARA SEQ ID NO:126	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	a. LONGITUD: 20	
	b. TIPO: Ácido Nucleico	
	(C) TIPO DE CADENA: Sencilla	
	(D) TOPOLOGÍA: Lineal	
20	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:126	
	<b>GGCCCAAGGT TACTTTTCAC</b>	<b>20</b>
25	(2) INFORMACIÓN PARA SEQ ID NO:127	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 16	
	(B) TIPO: Ácido Nucleico	
30	(C) TIPO DE CADENA: Sencilla	
	(D) TOPOLOGÍA: Lineal	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:127	
35	<b>GGGCCACTGC ACTCCT</b>	<b>16</b>
40	(2) INFORMACIÓN PARA SEQ ID NO:128	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 22	
	(B) TIPO: Ácido Nucleico	
	(C) TIPO DE CADENA: Sencilla	
	(D) TOPOLOGÍA: Lineal	
45	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:128	
	<b>CATGGTGAGG CTGAAGTAGG AT</b>	<b>22</b>
50	(2) INFORMACIÓN PARA SEQ ID NO:129	
	(i) SEQUENCE CHARACTERISTICS:	
	(A) LONGITUD: 25	
	(B) TIPO: Ácido Nucleico	
55	(C) TIPO DE CADENA: Sencilla	
	(D) TOPOLOGÍA: Lineal	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:129	
60	<b>GTGGCGTGTC TTTTACTTT CTTTA</b>	<b>25</b>

(2) INFORMACIÓN PARA SEQ ID NO:130

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 18
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

5

10 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:130

**AGGCAGCCCA GGAACAAT**

**18**

(2) INFORMACIÓN PARA SEQ ID NO: 131

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 22
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

15

20

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:131

**CCAAGATAGC GGCCAAGATA GT**

**22**

25

(2) INFORMACIÓN PARA SEQ ID NO:132

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

30

35

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:132

**GAGGGCAGCT GGGATGTTAC TCTT**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:133

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 24
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

40

45

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:133

**TGCCCTGTTT GGAGAACTGT AGGT**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:134

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 19
  - (B) TIPO: Ácido Nucleico
  - (C) TIPO DE CADENA: Sencilla
  - (D) TOPOLOGÍA: Lineal

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134

**CTCCCCAGAA ACAGATGTA**

**19**

60

- (2) INFORMATION FOR SEQ ID NO:135
- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 20  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:135
- GTGAGCCGAG ATTGTATCAT** **20**
- (2) INFORMACIÓN PARA SEQ ID NO:136
- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 19  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:136
- TCGGGGACAG GGCTTACTC** **19**
- (2) INFORMACIÓN PARA SEQ ID NO:137
- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 24  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:137
- ATCATTGTGCG CTGCTACTTT ATCG** **24**
- (2) INFORMACIÓN PARA SEQ ID NO:138
- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 21  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:138
- CTACTCTACC CCATTTTCATT C** **21**
- (2) INFORMACIÓN PARA SEQ ID NO:139
- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 19  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:139
- GTAGAGTGGAG TGGATGAGA** **19**
- (2) INFORMACIÓN PARA SEQ ID NO:140

5	(i) CARACTERÍSTICAS DE LA SECUENCIA: (A) LONGITUD: 21 (B) TIPO: Ácido Nucleico (C) TIPO DE CADENA: Sencilla (D) TOPOLOGÍA: Lineal	
10	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:140  <b>ATCAGGCAAA AACGAACAAA C</b>	<b>21</b>
15	(2) INFORMACIÓN PARA SEQ ID NO:141  (i) CARACTERÍSTICAS DE LA SECUENCIA: (A) LONGITUD: 17 (B) TIPO: Ácido Nucleico (C) TIPO DE CADENA: Sencilla (D) TOPOLOGÍA: Lineal	
20	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:141  <b>CGGCATCCCA AAGTGAC</b>	<b>17</b>
25	(2) INFORMACIÓN PARA SEQ ID NO:142  (i) CARACTERÍSTICAS DE LA SECUENCIA: (A) LONGITUD: 23 (B) TIPO: Ácido Nucleico (C) TIPO DE CADENA: Sencilla (D) TOPOLOGÍA: Lineal	
30	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:142  <b>CAGAGAGGGCA GCACCTTGGA CAG</b>	<b>23</b>
35	(2) INFORMACIÓN PARA SEQ ID NO:143  (i) CARACTERÍSTICAS DE LA SECUENCIA: (A) LONGITUD: 23 (B) TIPO: Ácido Nucleico (C) TIPO DE CADENA: Sencilla (D) TOPOLOGÍA: Lineal	
40	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:143  <b>GGCTTCACCT GCTCCCGTTT CAG</b>	<b>23</b>
45	(2) INFORMACIÓN PARA SEQ ID NO:144  (i) CARACTERÍSTICAS DE LA SECUENCIA: (A) LONGITUD: 22 (B) TIPO: Ácido Nucleico (C) TIPO DE CADENA: Sencilla (D) TOPOLOGÍA: Lineal	
50	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:144  <b>TCTGCCCATT CCCCAGCCTC TC</b>	<b>22</b>
55		
60		

(2) INFORMACIÓN PARA SEQ ID NO:145

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 24  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Linea

5

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:145

10

**TACCGCGTGG CATTCAAGCA TAGC**

**24**

(2) INFORMACIÓN PARA SEQ ID NO:146

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 18  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

15

20

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:146

**TCCAGTCTGG GTGACAAA**

**18**

25

(2) INFORMACIÓN PARA SEQ ID NO:147

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 20  
 (B) TIPO: Ácido Nucleico  
 (C) TIPO DE CADENA: Sencilla  
 (D) TOPOLOGÍA: Lineal

30

35

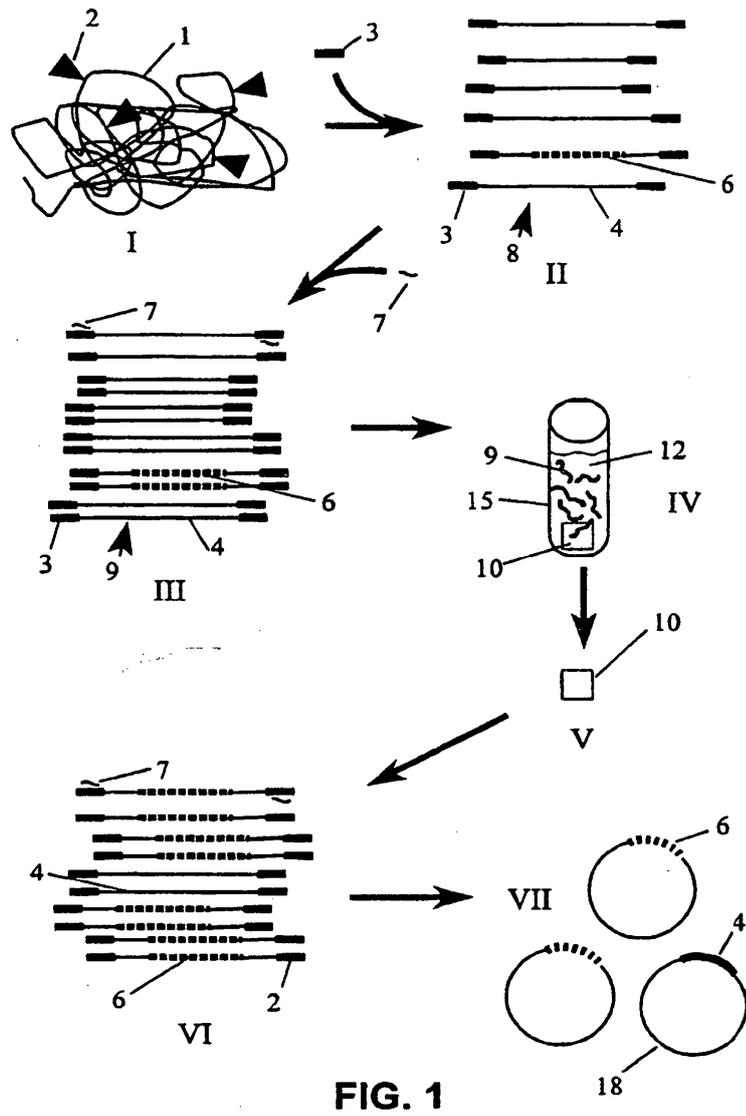
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:147

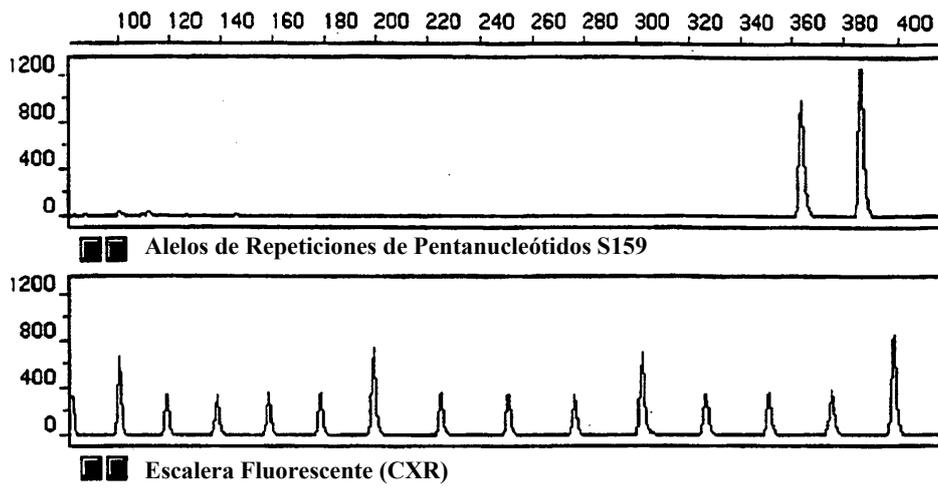
**CAATCCACTC CACTCCTCTA**

**20**

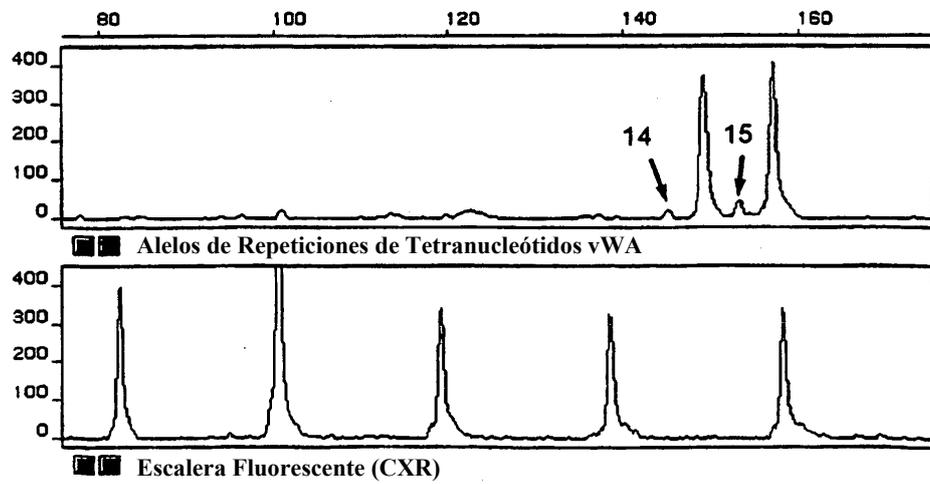
## REIVINDICACIONES

1. Un método para detectar una secuencia de ADN objetivo de repeticiones intermedias en tándem que tiene una incidencia baja de artefactos por tartamudeo, que comprende las etapas de:
- 5 (a) proporcionar una muestra de ADN que tiene al menos una secuencia objetivo de repeticiones intermedias en tándem, en la que la secuencia objetivo de repeticiones intermedias en tándem es una región del ADN que contiene al menos una unidad de repetición que consiste en una secuencia de cinco (5), seis (6), o siete (7) pares de bases repetidas en tándem al menos dos (2) veces; y
- 10 (b) detectar la secuencia objetivo de repeticiones intermedias en tándem en la muestra de ADN, en la que la secuencia objetivo de repeticiones intermedias en tándem de la muestra de ADN proporcionada en la etapa (a) se amplifica antes de la etapa (b), y en la que la secuencia objetivo de repeticiones intermedias en tándem se amplifica mediante el uso de al menos un cebador oligonucleotídico, que comprende una secuencia que es complementaria y que flanquea a una región de un marcador de ADN bicatenario que contiene una secuencia molde de repeticiones intermedias en tándem, en la que la secuencia molde de repeticiones intermedias en tándem es una región del marcador de ADN que contiene la secuencia de unidades de repetición repetidas en tándem al menos dos (2) veces,
- 15 con tal de que el marcador de ADN tenga la secuencia SEQ ID N°:26.
2. El método de la reivindicación 1, en el que la muestra de ADN proporcionada en la etapa (a) es ADN genómico humano.
3. El método de la reivindicación 2, en el que el cebador oligonucleotídico usado en la amplificación de la secuencia objetivo de repeticiones intermedias en tándem tiene un marcador fluorescente unido de manera covalente a él.
- 20 4. El método de la reivindicación 1, que comprende además detectar los polimorfismos en la secuencia objetivo de repeticiones intermedias en tándem amplificada.
5. El método de la reivindicación 4, en el que el cebador oligonucleotídico proporcionado en la etapa (b) comprende una secuencia seleccionada de uno de los grupos de secuencias que consisten en: SEQ ID N°:107, SEQ ID N°:108, SEQ ID N°:109, SEQ ID N°:110 y SEQ ID N°: 111.
- 25 6. Un equipo para la detección de al menos una secuencia objetivo de repeticiones intermedias en tándem en una muestra de ADN, en el que la secuencia objetivo de repeticiones intermedias en tándem es una región de la muestra de ADN que contiene al menos una unidad de repetición que consiste en una secuencia de cinco (5), seis (6), o siete (7) pares de bases repetidas en tándem al menos dos (2) veces, y el equipo comprende:
- 30 un recipiente que tiene un par de cebadores oligonucleotídicos para amplificar la al menos una secuencia objetivo de repeticiones intermedias en tándem, en el que cada cebador oligonucleotídico comprende una secuencia de ácidos nucleicos de una longitud mayor de 15 nucleótidos, cuyas secuencias son complementarias a las cadenas opuestas de un marcador de ADN bicatenario que flanquea a la secuencia molde de repeticiones intermedias en tándem,
- 35 y en el que el marcador de ADN tiene la secuencia SEQ ID N°:26.
7. El equipo de la reivindicación 6, que comprende además el marcador de ADN.
8. Un cebador oligonucleotídico que comprende una secuencia de una longitud mayor de 15 nucleótidos, cuya secuencia es complementaria a una cadena de un marcador de ADN bicatenario adyacente a una secuencia molde de repeticiones intermedias en tándem, en el que la secuencia molde de repeticiones intermedias en tándem es una región del marcador de ADN bicatenario que contiene al menos una unidad de repetición que consiste en una secuencia de cinco (5), seis (6), o siete (7) pares de bases repetidas en tándem al menos dos (2) veces, en el que la secuencia del marcador de ADN bicatenario es SEQ ID N°:26.
- 40 9. El cebador oligonucleotídico de la reivindicación 8, en el que el cebador oligonucleotídico comprende una secuencia seleccionada del grupo que consiste en: SEQ ID N°:107, SEQ ID N°:108, SEQ ID N°:109, SEQ ID N°:110 y SEQ ID N°: 111.
- 45 10. Un equipo según la reivindicación 6, en el que al menos uno del par de cebadores oligonucleotídicos es un cebador según la reivindicación 8 o la reivindicación 9.





**FIG. 2**



**FIG. 3**

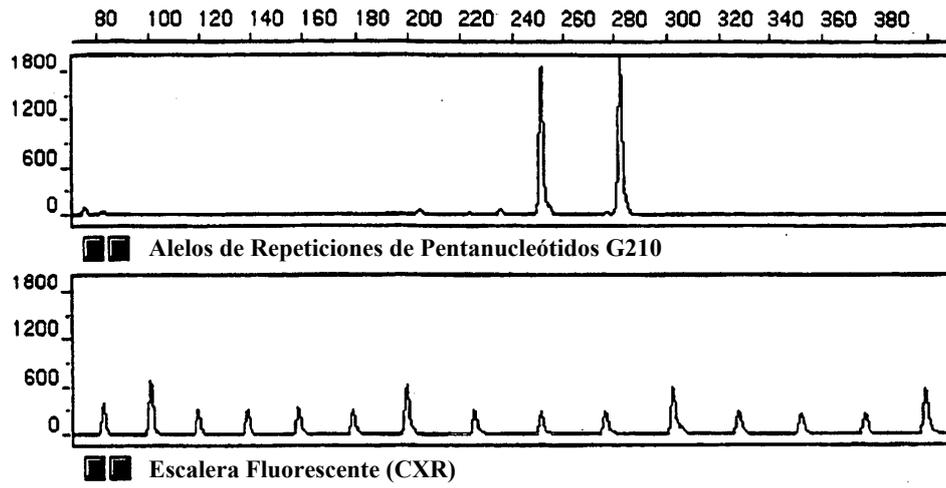


FIG. 4

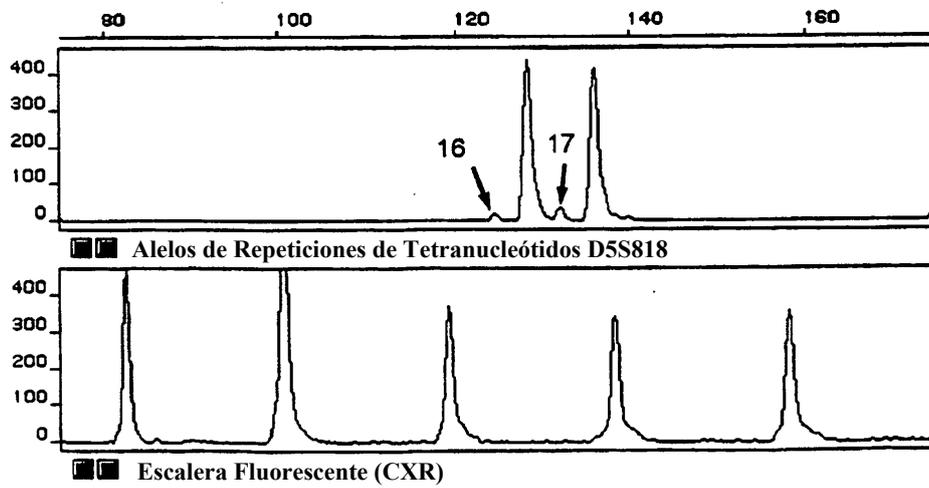
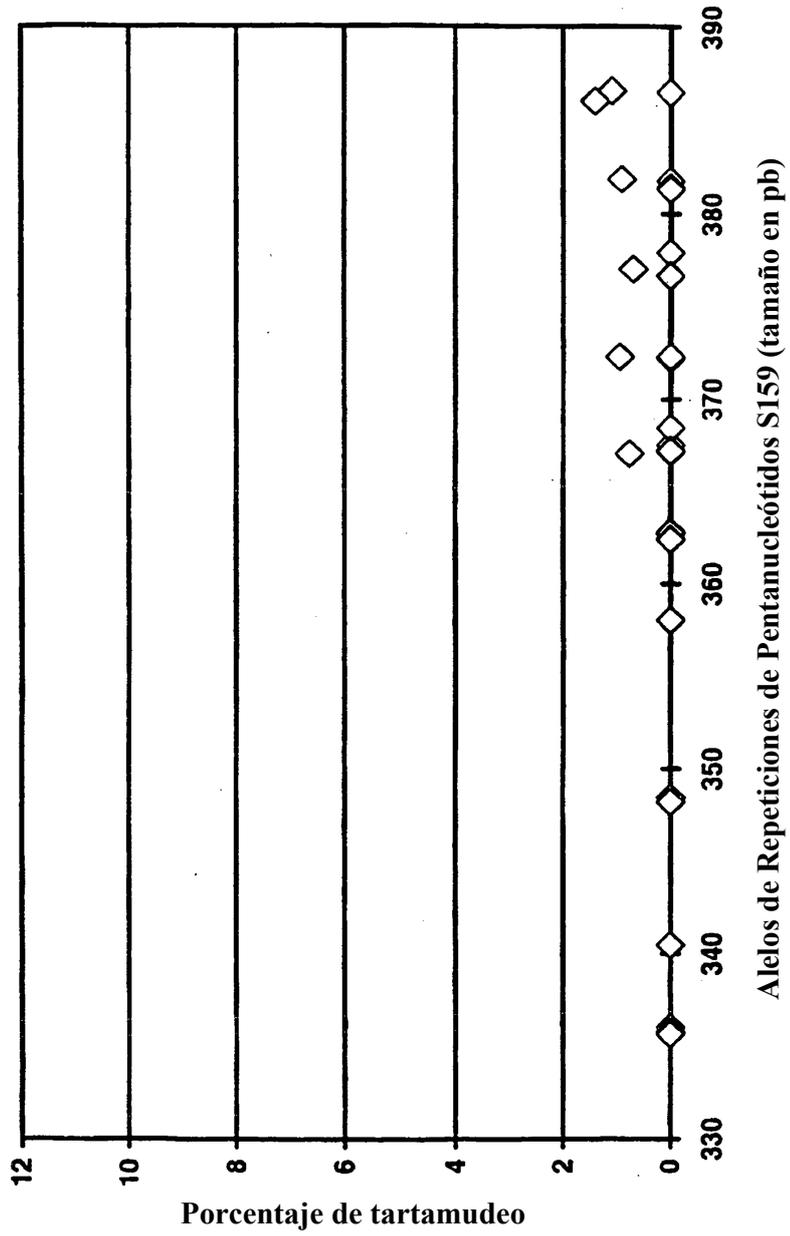


FIG. 5



**FIG. 6**

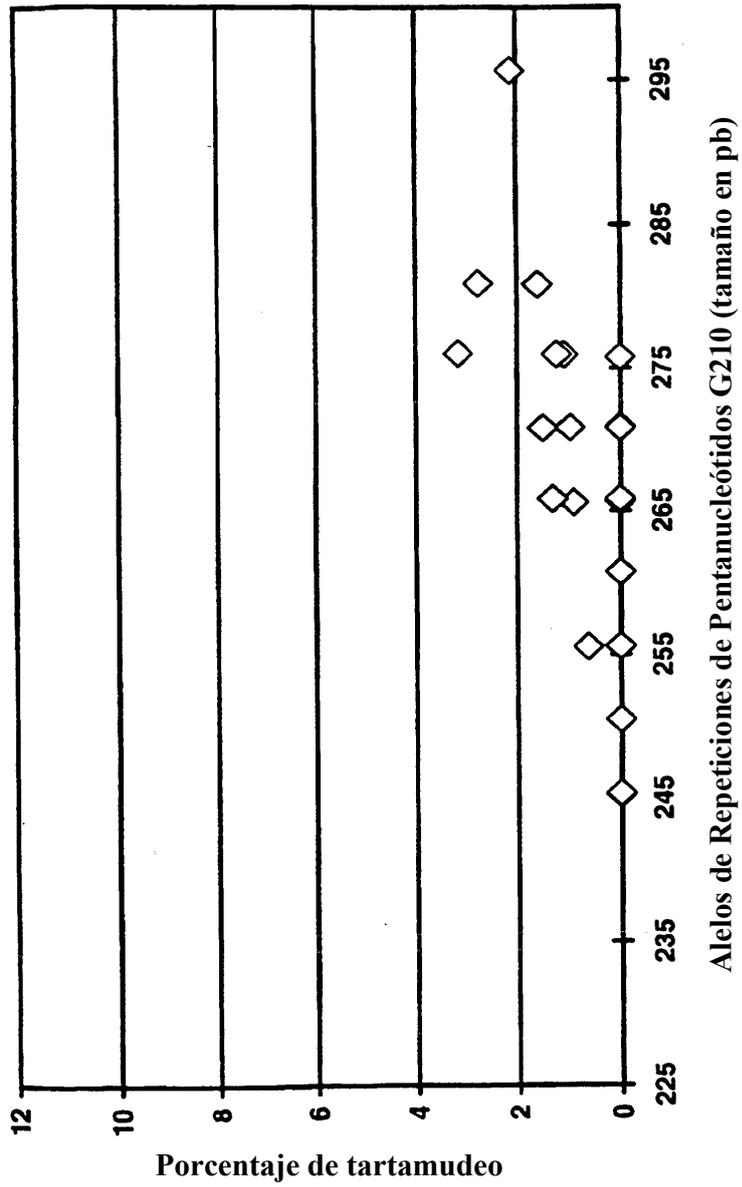
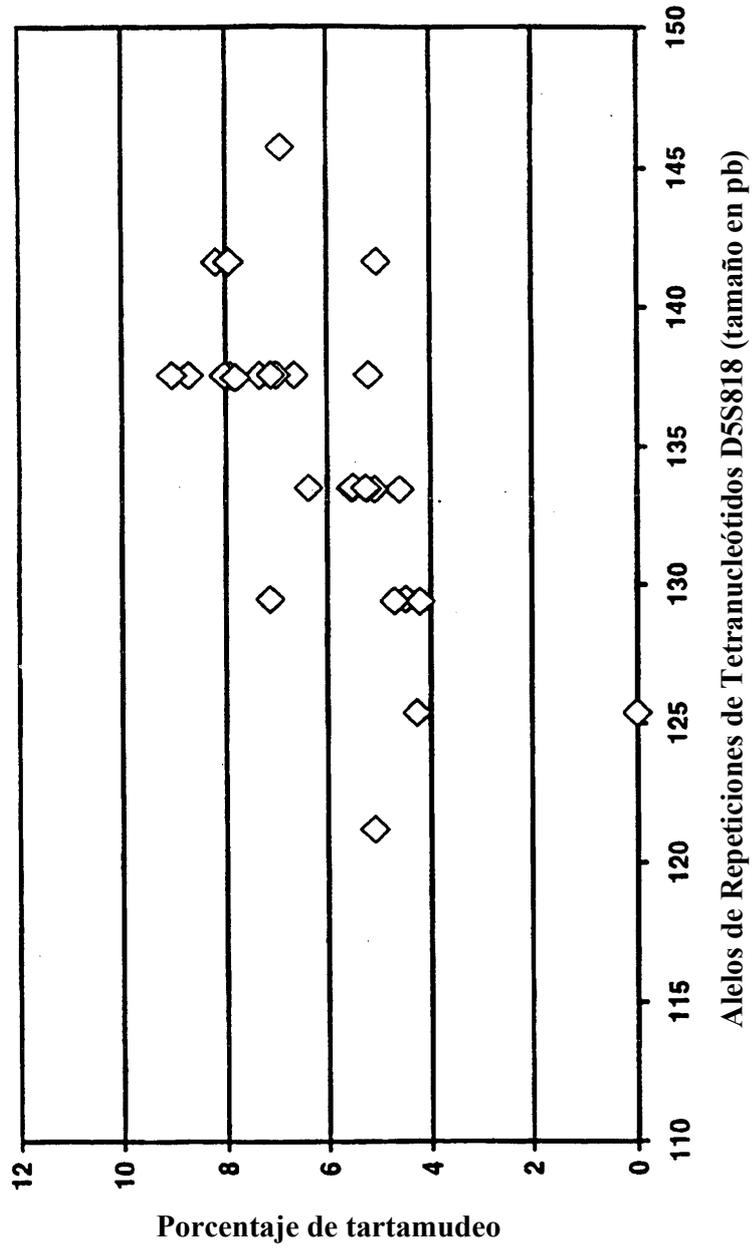
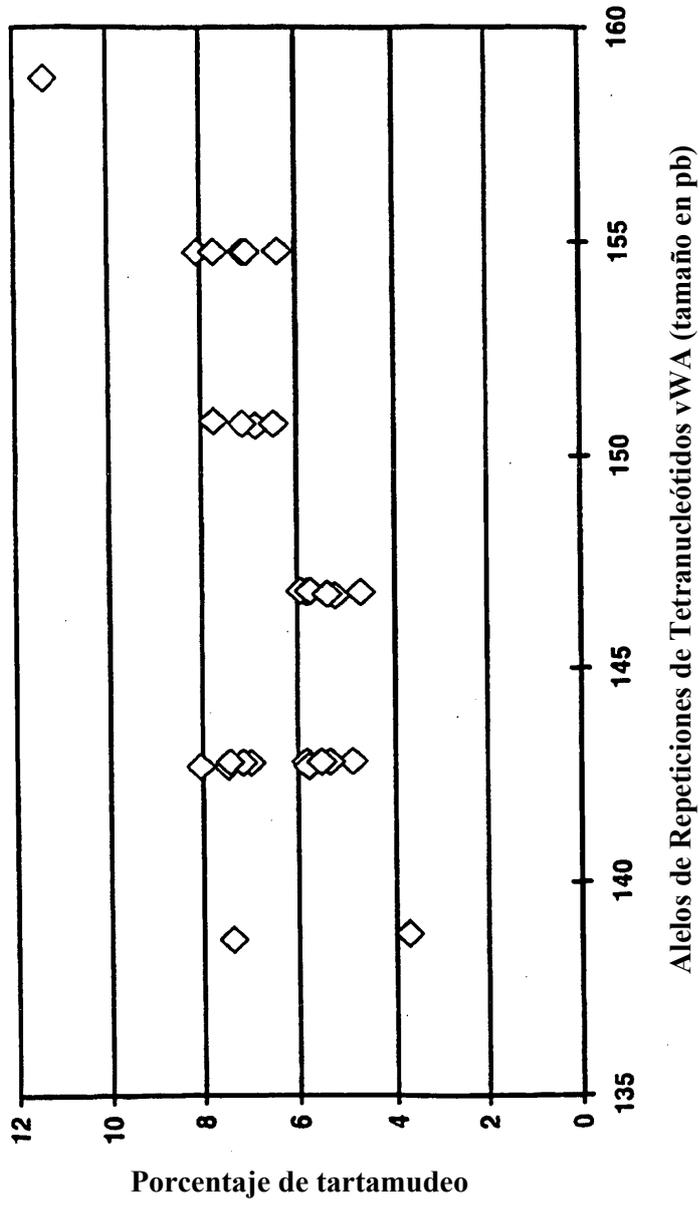


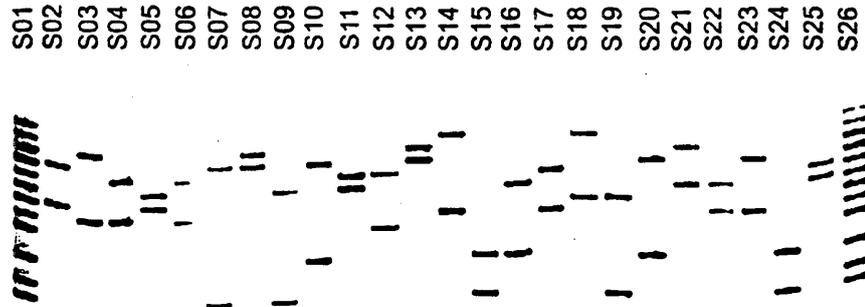
FIG. 7



**FIG. 8**

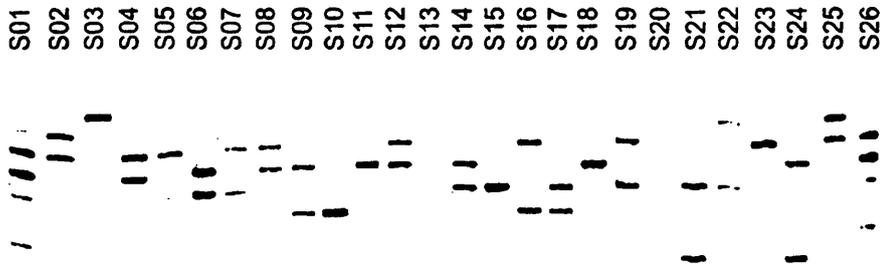


**FIG. 9**



**FIG. 10**

REPETICIÓN DE PENTANUCLEÓTIDOS POLIMÓRFICA S159



**FIG. 11**

REPETICIÓN DE PENTANUCLEÓTIDOS POLIMÓRFICA G210