

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 534 139**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.12.2005 E 10177053 (5)**

97 Fecha y número de publicación de la concesión europea: **11.02.2015 EP 2280084**

54 Título: **Métodos y ácidos nucleicos para el análisis de la expresión génica asociada con el pronóstico de trastornos proliferativos de células de próstata**

30 Prioridad:

02.12.2004 US 632426 P
02.12.2004 US 633250 P
14.03.2005 US 662220 P
03.10.2005 US 723125 P
03.10.2005 US 723054 P
30.11.2005 US 740736 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
17.04.2015

73 Titular/es:

EPIGENOMICS AG (100.0%)
Geneststrasse 5
10829 Berlin, DE

72 Inventor/es:

COTTRELL, SUSAN;
MODEL, FABIAN;
HAEFLIGER, CAROLINA;
WEISS, GUNTER;
DISTLER, JÜRGEN;
SLEDZIEWSKI, ANDREW Z.;
SONG, XIAOLING;
SKILLMAN, THOMAS L. y
THOMAS, JEFFREY G.

74 Agente/Representante:

LAZCANO GAINZA, Jesús

ES 2 534 139 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

Métodos y ácidos nucleicos para el análisis de la expresión génica asociada con el pronóstico de trastornos proliferativos de células de próstata

Campo de la invención

Los aspectos de la presente invención se relacionan con secuencias de ADN humano que exhiben patrones de expresión heterogéneos en pacientes con cáncer de próstata, y más particularmente con métodos novedosos para proporcionar un pronóstico de dichos pacientes.

Listado de secuencias

Una lista de secuencias, conforme a las Instrucciones Administrativas del PCT Parte 8, Sección 801(a), se ha proporcionado en disco compacto (1 de 1) como un archivo de texto de 3.25 MB (476_0001.txt).

Antecedentes

Cáncer de próstata. El cáncer de próstata es la neoplasia maligna más común entre los hombres en los Estados Unidos (~200,000 nuevos casos por año), y la sexta causa principal de muertes relacionadas con el cáncer entre los hombres en todo el mundo (~204,000 por año). El cáncer de próstata es principalmente una enfermedad de los ancianos, con aproximadamente 16% de los hombres entre las edades de 60 y 79 que tienen la enfermedad. De acuerdo con algunas estimaciones en la autopsia, el 80% de todos los hombres con más de 80 años de edad tiene alguna forma de enfermedad de la próstata (*por ejemplo*, cáncer, BPH, prostatitis, etc). La hipertrofia benigna de la próstata está presente en aproximadamente 50% de los hombres de 50 o más años de edad, y en 95% de los hombres de 75 o más años de edad. El cáncer de próstata, basado en estos informes, frecuentemente no es una enfermedad de la que mueren los hombres, sino más típicamente una con la que mueren. La evidencia reciente sugiere que la incidencia de cáncer de próstata de hecho puede estar disminuyendo, probablemente como resultado de un mejor tratamiento, una mejor cirugía, y una detección más temprana.

Diagnóstico del cáncer de próstata; enfoques moleculares. Las directrices actuales para la detección del cáncer de próstata han sido sugeridas por la Sociedad Americana del Cáncer y son las siguientes: A los 50 años de edad, los profesionales de la salud deben ofrecer una prueba de sangre para el antígeno prostático específico (PSA) y realizar un examen rectal digital (DRE). Se recomienda que las poblaciones de alto riesgo, tales como los afroamericanos y aquellos con una historia familiar de enfermedad de la próstata, deben comenzar la detección a los 45 años de edad. Los hombres sin patología prostática anormal generalmente tienen un nivel de PSA en sangre por debajo de 4 ng/ml. Los niveles de PSA entre 4 ng/ml y 10 ng/ml (denominados la 'Zona Gris') tienen 25% de probabilidad de tener cáncer de próstata. El resultado es que el 75% de las veces, los hombres con un DRE anormal y un PSA en esta *zona gris* tienen una biopsia negativa, o aparentemente innecesaria. Por encima de la zona gris, la probabilidad de tener cáncer de próstata es significativa (> 67%) y aumenta aun más a medida que se elevan los niveles de PSA. Existen numerosos métodos para medir PSA (porcentaje de PSA libre, velocidad del PSA, densidad del PSA, etc.), y cada uno tiene una precisión asociada para detectar la presencia del cáncer. Sin embargo, aun con las pequeñas mejoras en la detección, y las disminuciones informadas de la mortalidad asociada con la detección, la frecuencia de falsos positivos se mantiene alta. La reducción de la especificidad resulta en parte del aumento del PSA en la sangre asociado con BPH, y prostatitis. Además se ha estimado que hasta el 45% de las biopsias de próstata bajo las directrices actuales son falsas negativas, lo que resulta en la disminución de la sensibilidad aun con la biopsia.

La biopsia guiada por TRUS se considera el 'estándar de oro' para diagnosticar el cáncer de próstata. Las recomendaciones para la biopsia se basan en niveles de PSA anormales y/o un DRE anormal. Para el PSA existe una zona gris donde un alto porcentaje de las biopsias quizás no sean necesarias. Sin embargo la capacidad de detectar cáncer en esta zona gris (niveles de PSA de 4.0 a 10 ng/ml) es difícil sin la biopsia. Debido a esta falta de especificidad, el 75% de los hombres que se someten a una biopsia no tienen cáncer. Sin embargo sin biopsia, aquellos con cáncer se perderían, lo que resultaría en el aumento de la morbilidad y la mortalidad. Desafortunadamente, los riesgos asociados con una biopsia innecesaria también son altos.

Los marcadores moleculares ofrecerían la ventaja de poder usarse para analizar eficientemente incluso muestras de tejido muy pequeñas, y muestras cuya arquitectura tisular no se ha mantenido. En la última década, se han estudiado numerosos genes con respecto a la expresión diferencial entre la hiperplasia benigna de la próstata y diferentes grados de cáncer de

próstata. Sin embargo, aún no se ha demostrado que un marcador único sea suficiente para la clasificación del pronóstico de tumores de próstata en un entorno clínico.

Alternativamente, los enfoques de alta dimensión basados en el ARNm pueden, en casos particulares, proporcionar un instrumento para distinguir entre tipos de tumores diferentes y lesiones benignas y malignas. Sin embargo, la aplicación de dichos enfoques como una herramienta de diagnóstico de rutina en un ambiente clínico se ve impedida y sustancialmente limitada por la extrema inestabilidad del ARNm, los cambios en la expresión que se producen rápidamente tras ciertos desencadenantes (*por ejemplo*, recolección de las muestras), y, lo más importante, por la gran cantidad de ARNm necesaria para el análisis la que frecuentemente no se puede obtener a partir de una biopsia de rutina (ver, *por ejemplo*, Lipshutz, R. J. y otros, Nature Genetics 21:20-24, 1999; Bowtell, D. D. L. Nature Genetics Suppl. 21:25-32, 1999).

La metilación genética aberrante en el cáncer de próstata se ha observado en varios genes que incluyen GSTP1, AR, p16 (CDKN2a/INK4a), CD44, CDH1. La hipometilación en todo el genoma por ejemplo del elemento repetitivo LINE-1 también se ha asociado con la progresión del tumor (Santourlidis, S., y otros, Prostate 39:166-74, 1999).

Opciones de tratamiento del cáncer de próstata. Existen muchas estrategias de tratamiento disponibles para pacientes diagnosticados con cáncer de próstata, y la decisión para los pacientes y médicos frecuentemente no está clara. Debido a que el cáncer de próstata puede ser una enfermedad de lento desarrollo, muchos hombres eligen un enfoque de tratamiento denominado espera vigilante, o tratamiento conservador. Como el nombre lo indica, este enfoque no incluye ninguna terapia radical destinada a curar al paciente. En su lugar, la enfermedad es monitoreada cuidadosamente con el uso de pruebas de PSA y DRE. El paciente ideal para este enfoque es uno cuyo tumor es de crecimiento lento y no invasivo, y que por lo tanto es propenso a morir de otras causas antes de que el cáncer de próstata se vuelva problemático.

Para pacientes más jóvenes con la enfermedad localizada, el tratamiento curativo es más adecuado. La prostatectomía radical se usa para extirpar la próstata y con suerte todos los restos del tumor. Los márgenes quirúrgicos, las vesículas seminales, y algunas veces los ganglios linfáticos se prueban para la presencia de cáncer, y en cada caso la presencia de cáncer se correlaciona con la reducción de la supervivencia libre de la enfermedad. En general, aproximadamente 70% de los hombres se mantienen libres de la enfermedad diez años después de la cirugía (Roehl, y otros, 2004). La prostatectomía radical es una cirugía significativa, con efectos secundarios que incluyen pérdida de sangre, incontinencia, e impotencia. Se estima que la tasa de complicaciones intraoperatorias y postoperatorias es menor que 2% (Lepor, y otros, 2001).

La terapia de radiación también se usa para intentar curar a los pacientes con cáncer de próstata. Los pacientes pueden elegir entre la radioterapia externa o la braquiterapia (implantes de semillas radioactivas). Las tasas de supervivencia y los efectos secundarios son similares a los de la prostatectomía radical (D'Amico, y otros, 1998). Tanto para la prostatectomía radical como para la terapia de radiación, la probabilidad de supervivencia es muy dependiente del estadio y la diferenciación del tumor. Los tumores indolentes localizados son más propensos a curarse.

La terapia hormonal se usa frecuentemente para pacientes cuyo cáncer se ha propagado más allá de la próstata o para pacientes cuyo cáncer ha recurrido después de la prostatectomía o la terapia de radiación. En otras palabras, la terapia hormonal se usa para controlar el cáncer pero no para curarlo. La terapia hormonal se usa algunas veces junto con otras terapias tales como la radiación o como una terapia neoadyuvante antes de la cirugía. El objetivo de la terapia hormonal es reducir el efecto estimulador de los andrógenos sobre el tumor de la próstata. La reducción de las hormonas se logra a través de la orquiectomía, análogos de la hormona liberadora de la hormona luteinizante (LHRH), y antiandrógenos. Los efectos secundarios de la terapia hormonal pueden incluir impotencia, sofocos, fatiga, y reducción de la libido. A menudo, los tumores de la próstata se vuelven insensibles a los andrógenos y la terapia hormonal ya no es eficaz.

Después que el tumor se ha propagado fuera de la cápsula y la terapia hormonal ha fracasado, puede usarse quimioterapia para aliviar el dolor o retrasar la progresión de la enfermedad. La respuesta a la quimioterapia es variable, y la vida se extiende sólo para una minoría de los pacientes. Los bisfosfonatos se usan para reducir la actividad osteolítica de los tumores que han hecho metástasis en los huesos.

Estimación del pronóstico del cáncer de próstata. El DRE, la TRUS, la biopsia, y el PSA proporcionan una información inicial de estadificación del tumor, pero la MRI, las exploraciones CT, las gammagrafías ProstaScint y las gammagrafías óseas se usan para determinar la propagación del cáncer más allá de la cápsula prostática. Estas pruebas no se usan en todos los pacientes con cáncer de próstata, sino sólo en aquellos con alguna probabilidad de metástasis. Si puede confirmarse la metástasis, el paciente recibirá un tratamiento diseñado para retardar la progresión de la enfermedad. Si no se detectan metástasis, un paciente es un candidato para tratamientos potencialmente curativos tales como la prostatectomía y la terapia de radiación. Antes de extirpar la próstata, algunas veces se diseccionan los ganglios linfáticos como una prueba

final para las metástasis. Si hay metástasis presentes en los ganglios diseccionados, la cirugía se puede abortar. El análisis del tejido extirpado quirúrgicamente durante la prostatectomía es la técnica de estadificación final y el estándar de oro para aquellos pacientes que eligen someterse a cirugía. Con frecuencia, el análisis de los especímenes quirúrgicos demuestra que originalmente el paciente fue *subestadificado* por las pruebas de diagnóstico (Bostwick, 1997).

Una estimación precisa del pronóstico es crucial para la selección del tratamiento más adecuado para cada paciente. Dado que el cáncer de próstata confinado al órgano no puede conducir a la muerte, la estimación del pronóstico es también una estimación de la presencia o la probabilidad de desarrollo de metástasis. Un paciente que es propenso a desarrollar cáncer fuera de la cápsula prostática recibirá un trabajo de diagnóstico más extenso, que incluye exploraciones MRI y CT, y posiblemente tratamientos más radicales, que incluyen cirugía y radiación.

Una evaluación inicial del pronóstico se elabora a partir de los resultados de una prueba de PSA, el DRE, y el análisis de la biopsia. El tamaño, la ubicación, y el método de detección del tumor se combinan para proporcionar una puntuación de estadificación en la escala de TNM. Los pacientes con tumores en estadios superiores y valores de PSA altos son más propensos a tener un cáncer que se ha propagado o que se propagará fuera de la próstata. Un análisis histológico de la biopsia permite a un patólogo determinar la puntuación de Gleason. La puntuación de Gleason es un compuesto de los dos grados más prevalentes en la muestra de tejido, y los grados pueden estar en el intervalo entre uno y cinco. Un grado mayor indica una desdiferenciación más extrema, y puntuaciones compuestas mayores se correlacionan con mayor probabilidad de metástasis y reducción de la supervivencia libre de la enfermedad.

Los nomogramas de cáncer de próstata se han desarrollado y modificado para predecir el riesgo de recurrencia del cáncer después de la terapia primaria basado en los niveles de PSA, la calificación de Gleason, y la información de estadificación preoperatoria (Kattan y otros 2003; Kattan y otros 1998; Potter y otros 2001). El dato se deriva de las tasas reales de supervivencia de los pacientes en cohortes de miles de pacientes en múltiples instituciones. Como en todas las medidas de pronóstico en el cáncer de próstata, la estimación del riesgo de recurrencia por el nomograma es también una estimación de la probabilidad de presencia del cáncer fuera de la cápsula prostática. Debido a que las características clínicas de los tipos de cáncer que presentan los pacientes han cambiado con el amplio uso del PSA, los nomogramas son obsoletos y ya no se usan ampliamente. Sin embargo, el proceso general de integración de la información de Gleason, el estadio, y el PSA aún se usa.

Los pacientes con cáncer que se ha propagado a los ganglios linfáticos u otros sitios metastáticos se tratan con terapias sistémicas tales como las terapias hormonales. Los pacientes con enfermedad localizada (T1-T3) son candidatos para los tratamientos curativos definitivos, tales como cirugía o radiación. Aquellos pacientes con enfermedad localizada que se consideran de bajo riesgo son candidatos ideales para la espera vigilante. Aquellos con riesgo intermedio son ideales para monoterapia tal como cirugía o radiación. Aquellos con enfermedad localizada de alto riesgo deben considerarse para terapias multimodales o ensayos clínicos.

Después de la cirugía, está disponible más información del pronóstico debido a que puede analizarse directamente la propagación del tumor. Durante algunas prostatectomías, los ganglios linfáticos se diseccionan directamente y se confirma el estado del ganglio. En todas las cirugías, se comprueba la propagación del tumor a las vesículas seminales y el estado de los márgenes. Los ganglios, vesículas seminales, y márgenes positivos indican un pronóstico inferior y pueden sugerir que el paciente debe recibir tratamiento adyuvante.

Marcadores moleculares de pronóstico del cáncer de próstata; deficiencias de los enfoques de la técnica anterior. Como una alternativa a los enfoques actuales a la clasificación del pronóstico de los pacientes con carcinoma de próstata actualmente se explora una variedad de enfoques moleculares. Se prevé que el desarrollo de marcadores moleculares adecuados tendrá ventajas significativas sobre los enfoques actuales en términos de precisión, rentabilidad y/o invasividad al paciente. Se ha descubierto una variedad de marcadores moleculares que incluyen anticuerpos monoclonales. En un estudio realizado por Xu y otros (ICDB/95613763) 114 casos de cáncer de próstata mostraron que 57% de los especímenes de médula ósea tenían niveles de OVX1 elevados (mayores que 7.2 U/ml). En otros experimentos, los niveles de OVX1 eran aproximadamente 2 veces más altos en muestras de suero de pacientes con cáncer de próstata independiente de andrógenos que en los dependientes de andrógenos (p menor que 0.001), lo que sugiere que los niveles de OVX1 en suero pueden ser capaces de predecir la progresión del cáncer de próstata, ya que esta enfermedad cuando progresa típicamente se vuelve independiente de andrógenos. La expresión de la proteína PSCA y el ARNm se ha correlacionado positivamente con características adversas del tumor, tales como aumento del grado patológico (poca diferenciación celular), agravamiento del estadio clínico e independencia de andrógenos y especulativamente con la carcinogénesis de la próstata (Jpn J Clin Oncol, 4:414-9, 2004). Otros marcadores para análisis potenciales de ARNm incluyen Hepsin. La expresión de Hepsin

mostró diferencias significativas entre los pacientes de menor riesgo (puntuación pT2, G2 y Gleason menos de 7) y mayor riesgo (puntuación pT3/4, G3 y Gleason 7 o mayor) para recaída (J Urol, 171:187-91, 2004).

El gen GSTP1 es el marcador de diagnóstico de carcinoma de próstata mejor caracterizado. Zhou y otros (J Urol, 171:2195-8, 2004) recientemente correlacionado la expresión del gen GSTP1 con el grado de Gleason y el volumen del cáncer. Además, el uso del gen GSTP1 como un marcador para la detección de carcinomas de próstata ubicados en la zona periférica (*es decir*, con una alta probabilidad de metástasis) también se ha descrito en la solicitud de patente de los Estados Unidos con número de serie 10/350,763.

Otro marcador de metilación que puede ser adecuado para la clasificación del pronóstico de carcinomas de próstata es uPA. Rabbani y otros (The FASEB Journal 17:1081-1088, 2003) han demostrado que el promotor uPA está hipermetilado en células PrEC y LNCaP sensibles a hormonas e hipometilado en células PC-3 insensibles a hormonas. La desmetilación del promotor en las líneas celulares LNCaP resultó en el aumento del análisis del ARNm y resultó en un aumento de la capacidad invasiva. Singal y otros, analizaron la metilación de un panel de genes que consiste en glutatión s-transferasa P1 (GSTP1), receptor beta de ácido retinoico (RARβ), CD44, E-cadherina (ECAD) y proteína 1A de la familia de dominios de asociación a RAS (RASSF1A) en cáncer de próstata. Se calculó un índice de metilación (MI) como el número total de genes metilados, se observó un MI mayor en el estadio III en comparación con el estadio II de la enfermedad, y en las muestras con puntuación de Gleason 7 en comparación con las de puntuación de Gleason 6. Singal y otros concluyeron por lo tanto que los resultados sugieren que la metilación del panel de genes se correlaciona con las características clinicopatológicas de un mal pronóstico.

Necesidad pronunciada en la técnica. De manera significativa, sin embargo, ninguno de los marcadores mencionados hasta este momento está lo suficientemente desarrollado para proporcionar un marcador para el pronóstico de los trastornos proliferativos de células de próstata que sea lo suficientemente robusto y/o preciso para el uso eficaz en un entorno clínico.

Si bien el diagnóstico preciso del carcinoma de próstata es importante, la necesidad más apremiante en el tratamiento del cáncer de próstata está en la información para guiar la decisión de planificación del tratamiento. Los líderes en el campo coinciden en que muchos pacientes con una enfermedad no significativa clínicamente reciben tratamientos radicales innecesarios tales como prostatectomía o terapia de radiación. Sin embargo, el veinte por ciento de los pacientes que reciben estos tratamientos curativos experimenta la recurrencia de la enfermedad. Una prueba molecular podría ayudar a seleccionar pacientes para la elección de tratamiento óptimo y así reducir el exceso o la falta de tratamiento.

Actualmente la elección de la terapia se realiza sobre la base de la probabilidad de propagación de la enfermedad. Los pacientes de bajo riesgo son candidatos para la espera vigilante. Los pacientes de mediano y alto riesgo deberán recibir cirugía o radiación, y los pacientes de alto riesgo son candidatos para tratamientos adyuvantes adicionales. La estadificación, Gleason, y PSA se usan actualmente para estimar este riesgo, pero la información combinada de estas pruebas es insuficiente. Muy pocos pacientes son recomendados para la espera vigilante debido a que los clínicos no pueden estar seguros de cuáles cánceres son indolentes. Además, muchos pacientes que reciben monoterapia experimentan una recurrencia.

La calificación de Gleason actualmente juega un papel principal en la evaluación del pronóstico. Los pacientes con la enfermedad localizada y puntuaciones de Gleason altas (8-10) siempre se someten a tratamientos radicales. Los pacientes con puntuaciones de Gleason bajas (2-5) tienen la opción de diferir el tratamiento curativo y optar por la espera vigilante, sin embargo muchos eligen someterse a la terapia curativa poco después del diagnóstico. Para pacientes con puntuaciones de Gleason en el intervalo medio, que son la mayoría de los pacientes diagnosticados hoy en día, los clínicos deben usar otros indicadores de pronóstico menos confiables para información adicional.

En consecuencia existe una necesidad pronunciada en la técnica de una prueba de pronóstico eficaz, novedosa, y en particular una que pueda predecir la probabilidad de que un cáncer se haya propagado o sea propenso a propagarse fuera de la próstata basada en los patrones de metilación de muestras de biopsia. Este tipo de información es muy valiosa en los procesos de diagnóstico y planificación del tratamiento. Esta información se usaría inicialmente para llegar a la decisión acerca de si las pruebas imagenológicas son necesarias para comprobar la metástasis para un trabajo de diagnóstico completo. La cirugía es innecesaria para cualquier paciente cuyo cáncer ya se haya propagado, pero si un paciente no se selecciona para la imagenología la metástasis no se detectará hasta la cirugía o más tarde.

Además existe una necesidad pronunciada en la técnica de una prueba de clasificación molecular novedosa y eficaz de los trastornos proliferativos de células de próstata, y en particular una que sea adecuada para el análisis de muestras de biopsia para mejorar la estratificación de los pacientes en las categorías de riesgo bajo, intermedio, y alto de manera que puedan

seleccionarse planes de tratamiento óptimos para cada paciente. Con la estratificación precisa, los pacientes y los doctores pueden elegir la espera vigilante con la confianza de que existe poco riesgo de una recurrencia temprana. Por lo tanto esta prueba deberá reducir el número de cirugías y tratamientos de radiación innecesarios.

- 5 Adicionalmente, con mejores estimaciones de qué pacientes son propensos a recurrir con monoterapia, los médicos pueden hacer un mejor uso de los tratamientos adyuvantes disponibles. Si un paciente elige someterse a cirugía, la prueba puede repetirse en muestras de la prostatectomía para verificar la evaluación de su necesidad de una terapia adyuvante. Los beneficios de los diferentes enfoques de terapia adyuvante aún se están estudiando en los ensayos clínicos, y una prueba molecular podría proporcionar información valiosa para estratificar a los pacientes para este tratamiento adicional o para ensayos clínicos.
- 10 PITX2 (factor de transcripción 2 con homeodominio de tipo apareado), también conocido como PTX2, RIEG1, o ARP1, codifica un miembro de la familia de la caja homeótica RIEG/PITX, que está en la clase bicoide de proteínas de homeodominios. PITX2 codifica varios transcritos alternativos, y las mutaciones en el gen conducen al trastorno autosómico dominante síndrome de Rieger, un trastorno del desarrollo que afecta predominantemente al ojo (Semina y otros, 1996). La proteína actúa como un factor de transcripción y está involucrada en el desarrollo de varios órganos principales. Es inducida por la vía WNT, y media la proliferación específica de tipo celular mediante la inducción de genes reguladores del crecimiento (Kioussi y otros 2002).
- 15 Toyota y otros (2001) encontraron hipermetilación del gen en una gran proporción de leucemias mieloides agudas. Varios estudios realizados por el solicitante (ver WO 2005/059172) han demostrado que la hipermetilación de PITX2 está asociada con un mal pronóstico para pacientes con cáncer de mama.
- 20 El marcador GPR7 está ubicado en una isla de CpG en la región promotora de un gen sin intrones en el cromosoma 10. GPR7, o receptor acoplado a proteína G 7, es un receptor para el neuropéptido W y el neuropéptido B (Shimomura y otros 2002; Tanaka y otros 2003). La expresión de GPR7 se ha estudiado en el cerebro, y se expresa principalmente en el cerebelo y la corteza frontal (O'Dowd y otros 1995). Ishii y otros (2003) estudiaron el fenotipo de ratones que carecen de una copia funcional de GPR7. Los ratones desarrollaron obesidad del adulto y defectos metabólicos tales como disminución del gasto de energía y aumento de los niveles en sangre de glucosa e insulina. Curiosamente, estos fenotipos sólo se detectaron en ratones machos. Los ligandos de GPR7, los neuropéptidos W y B, también han sido implicados en el metabolismo y la obesidad en estudios diferentes (Samson y otros 2004; Levine y otros 2005). GPR7, que es similar en secuencia a los receptores opioides, también puede tener un papel en la señalización del dolor (Zaratin y otros 2005).
- 30 La SEQ ID NO: 63 está ubicada dentro de la región reguladora de HIST2H2BF en el cromosoma 1 en una región con varios genes de histonas. El contenido de histonas y el estado de la cromatina pueden influir en la expresión del gen codificado. La metilación y expresión alterada de un gen de histona en el cáncer de próstata pueden causar cambios en la cromatina en todo el genoma que alteran la expresión génica en formas que resultan en propiedades del tumor más agresivas. No hay artículos publicados sobre la función de esta histona particular.
- 35 El marcador referido como SEQ ID NO: 35 está ubicado en el cromosoma 3 corriente abajo del gen FOXL2 (factor de transcripción Forkhead) y dentro de o cercano a genes o EST predichos. Aunque está corriente abajo, se prevé que la metilación de este marcador tenga efecto sobre la expresión de FOXL2, que está mutado en el síndrome blefarofimosis-ptosis epicanto inverso (BPES). Este síndrome se caracteriza por anomalías oculares, craneofaciales, y ováricas. La metilación del marcador también puede afectar la expresión del EST, o el EST puede demostrar ser un exón alternativo para el gen FOXL2.
- 40 WO 02/31209 A2 describe un conjunto de genes cuya expresión está alterada en el cáncer de próstata.
- 45 WO 01/19845 A1 describe un método para detectar un trastorno proliferativo celular que se basa en la determinación del estado de metilación de al menos un gen o región reguladora asociada del gen.
- 50 EP 1 379 694 A2 describe métodos, sistemas y productos de programas informáticos para determinar el efecto y/o actividad biológica de fármacos, sustancias químicas y/o composiciones farmacéuticas que usan su efecto sobre la metilación del ADN como un marcador de su(s) efecto(s) biológico(s).
- WO 02/070742 A1 describe un método para el desarrollo de paneles de genes con fines diagnósticos y terapéuticos basados en la expresión y el estado de metilación de genes específicos.
- 55 WO 2004/035803 A2 describe el gen PITX2 tratado con bisulfito.

Resumen de la invención

El alcance de la presente invención se define en las reivindicaciones anexas.

También se describen genes, secuencias genómicas y/o regiones reguladoras de los mismos de acuerdo con la Tabla 11 (o a uno o más de ellos), la expresión de los mismos es indicativa del pronóstico de los trastornos proliferativos de células de próstata o las características de los mismos. Se prefiere particularmente que dichos genes, secuencias genómicas y/o regiones reguladoras se seleccionen del grupo que consiste en PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35. Adicionalmente se prefiere el gen PITX2. En una modalidad particularmente preferida, el estado de la metilación de las posiciones CpG de los genes, secuencias genómicas y/o regiones reguladoras de los mismos de acuerdo con la Tabla 11 (o a uno o más de ellos) es indicativo del pronóstico de los trastornos proliferativos de células de próstata o las características de los mismos. Se prefiere particularmente que dichos genes, secuencias genómicas y/o regiones reguladoras se seleccionen del grupo que consiste en PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35. Adicionalmente se prefiere el gen PITX2. Se prefiere particularmente que dicho trastorno proliferativo de células de próstata sea un carcinoma de próstata o neoplasia de próstata.

Para permitir este análisis se describe un método para el análisis de muestras biológicas para la metilación genómica asociada con el desarrollo de trastornos proliferativos de células de próstata. Dicho método se caracteriza porque al menos un ácido nucleico, o un fragmento del mismo, del grupo que consiste en SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961 (preferentemente SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961) se pone en contacto con un reactivo o una serie de reactivos capaz de distinguir entre dinucleótidos CpG metilados y no metilados dentro de la secuencia genómica, o secuencias de interés.

Se prefiere particularmente que el método y los ácidos nucleicos descritos en la presente descripción se utilicen para al menos uno de: pronóstico de; tratamiento de; monitoreo de; y tratamiento y monitoreo de trastornos proliferativos de células de próstata.

También se describe un método para determinar parámetros genéticos y/o epigenéticos del ADN genómico. El método tiene utilidad para mejorar la clasificación del pronóstico de trastornos proliferativos de células de próstata, más específicamente al permitir la mejor identificación de y diferenciación entre las formas agresivas y no agresivas de dicho trastorno. La presente descripción presenta varias mejoras sustanciales con respecto al estado de la técnica. Aunque se conocen algunos ensayos de metilación para la detección del cáncer, actualmente no existe un sistema de clasificación molecular para el *pronóstico* clasificación de tumores.

La fuente de ADN puede ser cualquier fuente adecuada. Preferentemente, la fuente de la muestra de ADN se selecciona del grupo que consiste en células o líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, eyaculado, orina, sangre, y combinaciones de los mismos. Preferentemente, la fuente es biopsias, fluidos corporales, eyaculado, orina o sangre.

Específicamente, se describe un método para proporcionar un pronóstico de los trastornos proliferativos de células de próstata, dicho método comprende: obtener una muestra biológica que comprende ácido(s) nucleico(s) genómico(s); poner en contacto el(los) ácido(s) nucleico(s), o un fragmento de estos, con un reactivo o una pluralidad de reactivos suficiente para distinguir entre secuencias de dinucleótidos CpG metilados y no metilados dentro de una secuencia objetivo del ácido nucleico sujeto, en donde la secuencia objetivo comprende, o hibrida bajo condiciones rigurosas a, una secuencia que comprende al menos 16 nucleótidos contiguos de SEQ ID NO:1 a SEQ ID NO: 64 y SEQ ID NO: 961, (preferentemente SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961) dichos nucleótidos contiguos comprenden al menos una secuencias de dinucleótidos CpG; y determinar, basado al menos en parte en dicha distinción, el estado de metilación de al menos una secuencias de dinucleótidos CpG objetivo, o un promedio, o un valor que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpGs objetivo. Preferentemente, distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas dentro de la secuencia objetivo comprende la conversión o no conversión dependiente del estado de metilación de al menos una de dichas secuencias de dinucleótidos CpG a la secuencia de dinucleótido convertida o no convertida correspondiente dentro de la secuencia seleccionada del grupo que consiste en SEQ ID NO: 65 a SEQ ID NO:320 y SEQ ID NO: 962 a SEQ ID NO: 965, y regiones contiguas de las mismas correspondientes a la secuencia objetivo. Preferentemente dicha secuencia se selecciona del grupo que consiste en las SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia dicha Secuencia se selecciona del grupo que consiste en SEQ ID Nos: 962 – 965.

También se describe un método para proporcionar un pronóstico de los trastornos proliferativos de células de próstata, que comprende: obtener una muestra biológica que tiene el ADN genómico sujeto; extraer el ADN genómico; tratar el ADN genómico, o un fragmento del mismo, con uno o más reactivos para convertir las bases citosina no metiladas en la posición

5 a uracilo u otra base que pueda detectarse como diferente de la citosina en términos de las propiedades de hibridación; poner en contacto el ADN genómico tratado, o el fragmento tratado del mismo, con una enzima de amplificación y al menos dos iniciadores que comprenden, en cada caso una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia seleccionada del grupo que consiste en SEQ ID NO:65 a SEQ ID NO:320 y SEQ ID NO: 962 a SEQ ID NO: 965 (preferentemente dicho grupo consiste en SEQ ID Nos: 133,134,261,262,189,190,317,318,101,102,229,230 y con la máxima preferencia dicho grupo consiste en SEQ ID Nos: 962 - 965) y los complementos de las mismas, en donde el ADN o el fragmento del mismo tratado se amplifica para producir un amplificado, o no se amplifica; y determinar, basado en la presencia o ausencia de, o en una propiedad de dicho amplificado, el estado de metilación de al menos una secuencia de dinucleótidos CpG seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO:64 y SEQ ID NO: 961 (preferentemente dicho grupo consiste en SEQ ID Nos: 35, 63, 19 y con la máxima preferencia es SEQ ID NO: 961), o un promedio, o un valor que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG de las mismas.

Preferentemente, al menos una de dichas moléculas de ácidos nucleicos que hibridan o moléculas de ácidos nucleicos peptídicos está unida a una fase sólida. Preferentemente, determinar comprende el uso de al menos un método seleccionado del grupo que consiste en: hibridar al menos una molécula de ácido nucleico que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia seleccionada del grupo que consiste en SEQ ID NO:65 a SEQ ID NO:320, (preferentemente dicho grupo consiste en SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia dicho grupo consiste en SEQ ID Nos: 962 - 965) y complementos de las mismas; hibridar al menos una molécula de ácido nucleico, unida a una fase sólida, que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia seleccionada del grupo que consiste en SEQ ID NO:65 a SEQ ID NO:320 y SEQ ID NO:962 a SEQ ID NO:965, (preferentemente dicho grupo consiste en SEQ ID Nos: 133,134,261,262,189,190,317,318,101,102,229,230 y con la máxima preferencia dicho grupo consiste en SEQ ID Nos: 962 - 965), y complementos de las mismas; hibridar al menos una molécula de ácido nucleico que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia seleccionada del grupo que consiste en SEQ ID NO:65 a SEQ ID NO:320 y SEQ ID NO: 962 a SEQ ID NO: 965 (preferentemente dicho grupo consiste en SEQ ID Nos: 133,134,261,262,189,190,317,318,101,102,229,230 y con la máxima preferencia dicho grupo consiste en SEQ ID Nos: 962 - 965), y complementos de las mismas, y extender al menos una de dichas moléculas de ácido nucleico hibridadas en al menos una base nucleotídica; y secuenciar el amplificado.

También se describe un método para proporcionar un pronóstico de los trastornos proliferativos de células de próstata, que comprende: obtener una muestra biológica que tiene el ADN genómico sujeto; extraer el ADN genómico; poner en contacto el ADN genómico, o un fragmento del mismo, que comprende una o más secuencias seleccionadas del grupo que consiste en SEQ ID NO:1 a SEQ ID NO: 64 y SEQ ID NO: 961 (preferentemente dicho grupo consiste en SEQ ID Nos: 35,63,19 y con la máxima preferencia dicha secuencia es SEQ ID NO:961) o una secuencia que hibrida bajo condiciones rigurosas a este, con una o más enzimas de restricción sensibles a la metilación, en donde el ADN genómico se digiere de esta manera para producir fragmentos de digestión, o no se digiere de esta manera; y determinar, basado en una presencia o ausencia de, o en una propiedad de al menos uno de dichos fragmentos, el estado de metilación de al menos una secuencia de dinucleótidos CpG de una o más secuencias seleccionadas del grupo que consiste en SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO:961, o un promedio, o un valor que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG de las mismas. Preferentemente dicho grupo consiste en SEQ ID Nos: 35,63,19 y con la máxima preferencia dicha secuencia es SEQ ID NO:961. Preferentemente, el ADN genómico digerido o sin digerir se amplifica antes de dicha determinación.

También se describen secuencias de ácidos nucleicos genómicos novedosas y químicamente modificadas, así como también oligonucleótidos y/o oligómeros PNA para el análisis de los patrones de metilación de citosina dentro de las secuencias del grupo que consiste en SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961.

Breve descripción de las figuras

La Figura 1 muestra un gráfico de ROC para un ensayo MPS de la SEQ ID NO:19 que se corre en 26 muestras congeladas de prostatectomía radical de pacientes con recurrencia de PSA temprano, y en 30 muestras de pacientes con recurrencia de PSA después de al menos 48 meses.

La Figura 2 muestra un gráfico de ROC para un ensayo MPS de la SEQ ID NO:35 que se corre en 26 muestras congeladas de prostatectomía radical de pacientes con recurrencia de PSA temprano y en 30 muestras de pacientes con recurrencia de PSA después de al menos 48 meses.

La Figura 3 muestra un gráfico de ROC para un ensayo MPS de la SEQ ID NO:37 que se corre en 26 muestras congeladas de prostatectomía radical de pacientes con recurrencia de PSA temprano y en 30 muestras de pacientes con recurrencia de PSA después de al menos 48 meses.

La Figura 4 muestra un gráfico de ROC para un ensayo MPS de la SEQ ID NO:7 que se corre en 26 muestras congeladas de prostatectomía radical de pacientes con recurrencia de PSA temprano y en 30 muestras de pacientes con recurrencia de PSA después de al menos 48 meses.

La Figura 5 muestra un gráfico de ROC para un ensayo MPS de la SEQ ID NO:63 que se corre en 26 muestras congeladas de prostatectomía radical de pacientes con recurrencia de PSA temprano y en 30 muestras de pacientes con recurrencia de PSA después de al menos 48 meses.

La Figura 6 muestra un gráfico de ROC para un ensayo MPS de la SEQ ID NO:8 que se corre en 26 muestras congeladas de prostatectomía radical de pacientes con recurrencia de PSA temprano y en 30 muestras de pacientes con recurrencia de PSA después de al menos 48 meses.

La Figura 7 muestra un gráfico de ROC para un ensayo MPS de la SEQ ID NO:64 que se corre en 26 muestras congeladas de prostatectomía radical de pacientes con recurrencia de PSA temprano y en 30 muestras de pacientes con recurrencia de PSA después de al menos 48 meses.

La Figura 8 muestra un análisis de electroforesis en gel en 12 muestras de ADN. 200 ng por ADN se aplicó a un gel de agarosa 0.8%. El gel se corrió por 4 horas a 80 Volt. El marcador de tamaño (Invitrogen, núm.: 10496-016) contiene los siguientes fragmentos: 10.000 bp, 6.000 bp.

La Figura 9 muestra los análisis ALF Express de los productos de PCR múltiple (8plex, mPCR SetD2) en comparación con los productos de PCR único (sPCR Set D2). El tamaño estándar (carriles 1,4) contenía los fragmentos de las siguientes longitudes: 50, 100, 150, 200, 250, 300, 350, 400, 450, 500 bp. Todos los fragmentos pueden amplificarse. Se observó un producto de tamaño no deseado (220 bp).

La Figura 10 muestra el desempeño del PCR múltiple. Carril 1: 100bp marcador. Carriles 2-11: desempeño del PCR múltiple de 10 muestras de prueba, carril 12: control positivo, carril 13: control de H₂O.

La Figura 11 muestra las muestras de Tumores vs. Linfocitos, clasificadas según la estadística de Wilcoxon. Los valores p corregidos por Bonferroni (superior) y AUC (inferior) se muestran a la derecha de la matriz de datos. Cada columna representa una muestra; cada fila un oligonucleótido. Los oligonucleótidos están agrupados por candidato a marcador. Los marcadores indicados están ordenados de arriba a abajo con el aumento de AUC. A la derecha de cada marcador se proporcionan el valor p de Wilcoxon corregido por Bonferroni y AUC. Más abajo se proporciona entre paréntesis la sensibilidad de AUC a una especificidad de ~ 0.75. Los datos de la metilación están centrados y normalizados con respecto a una desviación estándar para los oligonucleótidos individuales. El color representa la distancia relativa del estado de metilación del oligonucleótido desde el valor medio. El gris claro representa los CpG hipometilados dentro de un oligonucleótido mientras que el gris oscuro indica los CpG hipermetilados dentro de un oligonucleótido.

La Figura 12 muestra las clasificaciones de marcadores para Gleason alto vs. Gleason bajo. El gráfico muestra los valores p sin corregir del análisis estadístico de rangos de Wilcoxon por genes. Las líneas de puntos superior e inferior muestran los límites de Bonferroni y FDR del 5%, respectivamente.

La Figura 13 muestra la matriz de metilación de Gleason alto vs. Gleason bajo de los 10 marcadores con mejor AUC. Las puntuaciones de Gleason se muestran encima de cada grupo de muestras. Cada columna representa una muestra; cada fila un oligonucleótido (1, 2, o 3 sitios CpG cada uno). Los oligonucleótidos están agrupados por candidato a marcador. Los marcadores indicados están ordenados de arriba a abajo con el aumento de AUC. A la derecha de cada marcador se proporcionan el valor p de Wilcoxon corregido por Bonferroni y AUC. Más abajo se proporciona entre paréntesis la sensibilidad de AUC a una especificidad de ~ 0.75. Los datos de metilación están centrados y normalizados con respecto a una desviación estándar para los oligonucleótidos individuales. El color representa la distancia relativa del estado de metilación de los oligonucleótidos desde el valor medio. El gris claro representa los CpG hipometilados dentro de un oligonucleótido mientras que el gris oscuro indica los CpG hipermetilados dentro de un oligonucleótido.

La Figura 14 muestra las clasificaciones de marcadores de Recurrencia temprana vs. No recurrencia. El gráfico proporciona los valores p sin corregir del análisis de prueba de rangos de Wilcoxon por genes. Las líneas de puntos superiores e inferiores muestran los límites de Bonferroni y FDR del 5%, respectivamente.

La Figura 15 muestra la matriz de metilación de la Recurrencia temprana vs. No recurrencia de los 10 marcadores con mejor AUC. Cada columna representa una muestra; cada fila un oligonucleótido (1, 2, o 3 sitios CpG cada uno). Los oligonucleótidos están agrupados por candidato a marcador. Los marcadores indicados están ordenados de arriba a abajo con el aumento de AUC. A la derecha de cada marcador se proporcionan el valor p de Wilcoxon corregido por Bonferroni y AUC. Más abajo se proporciona entre paréntesis la sensibilidad de AUC a una especificidad de ~ 0.75. Los datos de metilación están centrados y normalizados con respecto a una desviación estándar para los oligonucleótidos individuales. El color representa la distancia relativa del estado de metilación del oligonucleótido desde el valor medio. El gris claro representa los CpG hipometilados dentro de un oligonucleótido mientras que el gris oscuro indica los CpG hipermetilados dentro de un oligonucleótido.

Para las Figuras 16-88, cada figura muestra la secuencia del amplificado analizado de cada SEQ ID NO respectiva. En cada figura, se muestra un amplificado analizado en una serie de paneles 'envueltos', donde la primera fila (fila superior) en cada panel muestra la secuencia *genómica* que se amplifica (la fila de secuencia genómica), y donde los iniciadores de amplificación directo e inverso (que definen un 'amplicón') se muestran en la fila del panel (la fila que muestra los iniciadores) inmediatamente por debajo de la primera fila. La fila más abajo de la fila que muestra los iniciadores (o, en los paneles que no muestran un iniciador, la fila más abajo de la fila que muestra la secuencia genómica) es la secuencia del amplificado convertida con bisulfito (la fila de secuencias convertidas con bisulfito; en donde las posiciones de CpG están marcadas en rojo). Las filas restantes mostradas en algunos paneles, muestran las secuencias de los oligonucleótidos de detección (oligos CG y TG) usadas para analizar el amplificado.

La Figura 16 muestra un amplificado de la SEQ ID NO:14.
 La Figura 17 muestra un amplificado de la SEQ ID NO:15.
 La Figura 18 muestra un amplificado de la SEQ ID NO:16.
 La Figura 19 muestra un amplificado de la SEQ ID NO:17.
 La Figura 20 muestra un amplificado de la SEQ ID NO:18.
 La Figura 21 muestra un amplificado de la SEQ ID NO:19.
 La Figura 22 muestra un amplificado de la SEQ ID NO:20.
 La Figura 23 muestra un amplificado de la SEQ ID NO:21.
 La Figura 24 muestra un amplificado de la SEQ ID NO:22.
 La Figura 25 muestra un amplificado de la SEQ ID NO:23.
 La Figura 26 muestra un amplificado de la SEQ ID NO:24.
 La Figura 27 muestra un amplificado de la SEQ ID NO:25.
 La Figura 28 muestra un amplificado de la SEQ ID NO:26.
 La Figura 29 muestra un amplificado de la SEQ ID NO:27.
 La Figura 30 muestra un amplificado de la SEQ ID NO:28.
 La Figura 31 muestra un amplificado de la SEQ ID NO:29.
 La Figura 32 muestra un amplificado de la SEQ ID NO:30.
 La Figura 33 muestra un amplificado de la SEQ ID NO:31.
 La Figura 34 muestra un amplificado de la SEQ ID NO:32 (amplificado A)
 La Figura 35 muestra un amplificado de la SEQ ID NO:32 (amplificado B)
 La Figura 36 muestra un amplificado de la SEQ ID NO:33.
 La Figura 37 muestra un amplificado de la SEQ ID NO:34.
 La Figura 38 muestra un amplificado de la SEQ ID NO:35.
 La Figura 39 muestra un amplificado de la SEQ ID NO:13.
 La Figura 40 muestra un amplificado de la SEQ ID NO:36.
 La Figura 41 muestra un amplificado de la SEQ ID NO:37.
 La Figura 42 muestra un amplificado de la SEQ ID NO:1.
 La Figura 43 muestra un amplificado de la SEQ ID NO:2.
 La Figura 44 muestra un amplificado de la SEQ ID NO:3.
 La Figura 45 muestra un amplificado de la SEQ ID NO:4.
 La Figura 46 muestra un amplificado de la SEQ ID NO:5.
 La Figura 47 muestra un amplificado de la SEQ ID NO:6.
 La Figura 48 muestra un amplificado de la SEQ ID NO:7.
 La Figura 49 muestra un amplificado de la SEQ ID NO:38.
 La Figura 50 muestra un amplificado de la SEQ ID NO:39.
 La Figura 60 muestra un amplificado de la SEQ ID NO:40.
 La Figura 61 muestra un amplificado de la SEQ ID NO:41.
 La Figura 62 muestra un amplificado de la SEQ ID NO:42.
 La Figura 63 muestra un amplificado de la SEQ ID NO:43.
 La Figura 64 muestra un amplificado de la SEQ ID NO:44.
 La Figura 65 muestra un amplificado de la SEQ ID NO:45.
 La Figura 66 muestra un amplificado de la SEQ ID NO:46.
 La Figura 67 muestra un amplificado de la SEQ ID NO:47.
 La Figura 68 muestra un amplificado de la SEQ ID NO:48.
 La Figura 69 muestra un amplificado de la SEQ ID NO:49.
 La Figura 70 muestra un amplificado de la SEQ ID NO:50.
 La Figura 71 muestra un amplificado de la SEQ ID NO:51.
 La Figura 72 muestra un amplificado de la SEQ ID NO:52.
 La Figura 73 muestra un amplificado de la SEQ ID NO:53.

- La Figura 74 muestra un amplificado de la SEQ ID NO:54
 La Figura 75 muestra un amplificado de la SEQ ID NO:55
 La Figura 76 muestra un amplificado de la SEQ ID NO:56
 La Figura 77 muestra un amplificado de la SEQ ID NO:57
 5 La Figura 78 muestra un amplificado de la SEQ ID NO:58
 La Figura 79 muestra un amplificado de la SEQ ID NO:59
 La Figura 80 muestra un amplificado de la SEQ ID NO:60
 La Figura 81 muestra un amplificado de la SEQ ID NO:61
 10 La Figura 82 muestra un amplificado de la SEQ ID NO:62
 La Figura 83 muestra un amplificado de la SEQ ID NO:8
 La Figura 84 muestra un amplificado de la SEQ ID NO:9
 La Figura 85 muestra un amplificado de la SEQ ID NO:10
 La Figura 86 muestra un amplificado de la SEQ ID NO:11
 15 La Figura 87 muestra un amplificado de la SEQ ID NO:12 (amplificado A)
 La Figura 88 muestra un amplificado de la SEQ ID NO:12 (amplificado B)
 La Figura 89 muestra la distribución de seguimiento de tiempos de los pacientes como se analiza en el Ejemplo 5. Las barras blancas representan la distribución de todos los pacientes excluidos (sin recaída de PSA). Las barras grises muestran la distribución del tiempo de supervivencia libre de PSA para todos los pacientes con recaída. La frecuencia se muestra en el eje Y y el tiempo (meses) se muestra en el eje X.
 20 La Figura 90 muestra el análisis de supervivencia de Kaplan-Meier del marcador PITX2 (A & B) y el análisis de la curva de ROC (C) del marcador PITX2 para diferenciar entre pacientes con cáncer de próstata de acuerdo con el Ejemplo 5. La proporción de pacientes libres de recurrencia se muestra en el eje Y, el tiempo en años se muestra en el eje X. La Figura 91 muestra el análisis de supervivencia de Kaplan-Meier del marcador GPR7 (A & B) y el análisis de la curva de ROC (C) del marcador PITX2 para diferenciar entre pacientes con cáncer de próstata de acuerdo con el Ejemplo 5. La proporción de pacientes libres de recurrencia se muestra en el eje Y, el tiempo en años se muestra en el eje X. La Figura 92 muestra el análisis de supervivencia de Kaplan-Meier del marcador SEQ ID NO:63 (A & B) y el análisis de la curva de ROC (C) del marcador PITX2 para diferenciar entre pacientes con cáncer de próstata de acuerdo con el Ejemplo 5. La proporción de pacientes libres de recurrencia se muestra en el eje Y, el tiempo en años se muestra en el eje X.
 25 La Figura 93 muestra el análisis de supervivencia de Kaplan-Meier del marcador SEQ ID NO: 35 (A & B) y el análisis de la curva de ROC (C) del marcador PITX2 para diferenciar entre pacientes con cáncer de próstata de acuerdo con el Ejemplo 5. La proporción de pacientes libres de recurrencia se muestra en el eje Y, el tiempo en años se muestra en el eje X.
 30 La Figura 94 muestra el análisis de supervivencia de Kaplan-Meier del marcador ABHD9 (A & B) y el análisis de la curva de ROC (C) del marcador PITX2 para diferenciar entre pacientes con cáncer de próstata de acuerdo con el Ejemplo 5. La proporción de pacientes libres de recurrencia se muestra en el eje Y, el tiempo en años se muestra en el eje X. La Figura 95 muestra el análisis de supervivencia de Kaplan-Meier del marcador CCND2 (A & B) y el análisis de la curva de ROC (C) del marcador PITX2 para diferenciar entre pacientes con cáncer de próstata de acuerdo con el Ejemplo 5. La proporción de pacientes libres de recurrencia se muestra en el eje Y, el tiempo en años se muestra en el eje X. La Figura 96 muestra el análisis de supervivencia de Kaplan-Meier del desempeño de PITX2 en subpoblaciones basadas en el estadio de acuerdo con el Ejemplo 5.
 40 La Figura 97 muestra el análisis de supervivencia de Kaplan-Meier del desempeño de PITX2 en subpoblaciones basadas en la puntuación de Gleason de acuerdo con el Ejemplo 5.
 La Figura 98 muestra el análisis de supervivencia de Kaplan-Meier del desempeño de PITX2 en subpoblaciones basadas en la puntuación del nomograma de acuerdo con el Ejemplo 5.
 45 La Figura 99 muestra el análisis de supervivencia de Kaplan-Meier del desempeño de la SEQ ID NO: 63 en subpoblaciones basadas en la puntuación de Gleason alta de acuerdo con el Ejemplo 5.
 La Figura 100 muestra el análisis de supervivencia de Kaplan-Meier del desempeño de la SEQ ID NO: 63 en subpoblaciones basadas en la puntuación del nomograma pobre de acuerdo con el Ejemplo 5.
 50 La Figura 101 muestra el análisis de supervivencia de Kaplan-Meier del desempeño de la SEQ ID NO: 35 en subpoblaciones T2 de acuerdo con el Ejemplo 5.
 La Figura 102 muestra el amplificado detectado tanto en muestras congeladas y PET en comparaciones de recaída bioquímica temprana vs. sin recaída bioquímica usando el ensayo de la SEQ ID NO: 19 que se muestra en la Tabla 12 según se detalla en el Ejemplo 6.
 55 La Figura 103 muestra el amplificado detectado tanto en muestras congeladas y PET en comparaciones de recaída bioquímica temprana vs. sin recaída bioquímica usando el ensayo de la SEQ ID NO: 63 que se muestra en la Tabla 12 según se detalla en el Ejemplo 6.

Como se usa en la presente descripción el término expresión se tomará en el sentido de la transcripción y traducción de un gen. El nivel de expresión de PITX2 se determina por el análisis del nivel de expresión de ARN.

5 Además, la actividad del gen transcrito puede verse afectada por variaciones genéticas tales como, pero sin limitarse a mutaciones genéticas (incluyendo pero no limitado a SNPs, mutaciones puntuales, deleciones, inserciones, longitud de repetición, reordenamientos y otros polimorfismos).

10 Como se usa en la presente descripción, el término "pronóstico" se tomará en el sentido de una predicción de la progresión de la enfermedad (por ejemplo, pero no limitado a la regresión, estasis y metástasis), en particular agresividad y potencial metastásico de un tumor de próstata.

15 Como se usa en la presente descripción, el término "marcador pronóstico" se tomará en el sentido de un indicador de una predicción de la progresión de la enfermedad, en particular agresividad y potencial metastásico de un tumor de próstata.

Como se usa en la presente descripción, el término "clasificación pronóstico" se tomará en el sentido de la clasificación de un trastorno proliferativo de células de la próstata de acuerdo con una predicción de la progresión de la enfermedad, en particular agresividad y potencial metastásico de un tumor de próstata.

20 Se usa preferentemente para definir a los pacientes con riesgos altos, bajos e intermedios de muerte o recurrencia después del tratamiento que resultan de la heterogeneidad inherente al proceso de la enfermedad. Como se usa en la presente descripción el término "agresivo" como se usa con respecto al tumor de próstata se tomará en el sentido de un trastorno proliferativo celular de próstata que tiene la capacidad biológica para propagarse rápidamente fuera de la próstata. Los indicadores de nivel de agresividad del tumor en la técnica incluyen pero sin limitarse a la etapa del tumor, grado tumoral, grado de Gleason, estado ganglionar y supervivencia. Como se usa en la presente descripción el término "sobrevida" no se limitará en el sentido de la supervivencia hasta la mortalidad (en donde dicha mortalidad puede ser ya sea independientemente de la causa o relacionada al trastorno proliferativo celular de la próstata), pero puede ser usado en combinación con otros términos para definir los términos clínicos, por ejemplo pero sin limitarse a "supervivencia libre de recurrencia" (en donde el término recurrencia incluirá tanto la recurrencia localizada como a distancia); supervivencia libre de metástasis; supervivencia libre de enfermedad (en donde el término enfermedad incluirá el cáncer de próstata y enfermedades asociadas con el mismo). La longitud de dicha supervivencia puede calcularse por referencia a un punto de inicio definido (por ejemplo, tiempo de diagnóstico o inicio del tratamiento) y un punto final definido (por ejemplo, la muerte, recurrencia o metástasis).

35 El término "relación observado/esperado" ("relación O/E") se refiere a la frecuencia de dinucleótidos CpG dentro de una secuencia de ADN particular, y corresponde al [número de sitios de CpG / (número de bases C x número de bases G)].

40 El término "isla de CpG" se refiere a una región contigua de ADN genómico que satisface los criterios de (1) tener una frecuencia de dinucleótidos CpG que corresponde a una "relación observado/esperado" >0.6, y (2) tener un "contenido de GC" >0.5. Las islas de CpG son típicamente, pero no siempre, entre aproximadamente 0.2 a aproximadamente 1 kb, o hasta aproximadamente 2 kb de longitud.

45 El término "estado de metilación" o "condición de metilación" se refiere a la presencia o ausencia de 5-metilcitosina ("5-mCyt") en uno o una pluralidad de dinucleótidos CpG dentro de una secuencia de ADN. Los estados de metilación en uno o más sitios de metilación CpG particulares (cada uno con dos secuencias de dinucleótidos CpG CpG) dentro de una secuencia de ADN incluyen "no metilado", "totalmente metilado" y "hemi-metilado."

50 El término "hemi-metilación" o "hemimetilación" se refiere al estado de metilación de un sitio de metilación CpG palindrómico, donde sólo se metila una única citosina en una de las dos secuencias de dinucleótidos de CpG del sitio de metilación CpG palindrómico (*por ejemplo*, 5'-CC^MGG-3' (cadena superior): 3'-GGCC-5' (cadena inferior)).

55 El término 'AUC' como se usa en la presente descripción es una abreviatura para el área bajo una curva. En particular, se refiere al área bajo una curva de característica operativa del receptor (ROC). La curva ROC es un gráfico de la tasa de verdaderos positivos contra la tasa de falsos positivos para los diferentes puntos de corte posibles de una prueba de diagnóstico. Se muestra la compensación entre la sensibilidad y la especificidad dependiendo del punto de corte seleccionado (cualquier aumento en la sensibilidad se acompañará por una disminución en la especificidad). El área bajo una curva ROC (AUC) es una medida de la exactitud de una prueba de diagnóstico (cuanto mayor sea el área mejor, el

óptimo es 1, una prueba aleatoria puede tener una curva ROC acostada en la diagonal con un área de 0.5; para referencia J.P. Egan. Signal Detection Theory and ROC Analysis, Academic Press, Nueva York, 1975).

5 El término "hipermetilación" se refiere al estado de metilación promedio que corresponde a un *aumento* de la presencia de 5-mCyt en un o una pluralidad de dinucleótidos CpG dentro de una secuencia de ADN de una muestra de ADN prueba, con relación a la cantidad de 5-mCyt encontrada en los dinucleótidos CpG correspondientes dentro de una muestra normal de ADN control.

10 El término "hipometilación" se refiere al estado de metilación promedio que corresponde a una disminución de la presencia de 5-mCyt en uno o una pluralidad de dinucleótidos CpG dentro de una secuencia de ADN de una muestra de ADN prueba, con relación a la cantidad de 5-mCyt encontrada en los dinucleótidos CpG correspondientes dentro de una muestra normal de ADN control.

15 El término "microarreglos" se refiere ampliamente tanto a "microarreglos de ADN," como "matrices de ADN," como se reconoce en la técnica, abarca todos los soportes sólidos reconocidos en la técnica, y abarca todos los métodos para fijar moléculas de ácido nucleico a estos o síntesis de ácidos nucleicos en estos.

20 "Parámetros Genéticos" son mutaciones y polimorfismos de genes y secuencias adicionalmente requeridos para su regulación. Se designan como mutaciones, en particular, las inserciones, deleciones, mutaciones puntuales, inversiones y polimorfismos y son particularmente preferidos los SNPs (polimorfismos de un solo nucleótido).

25 "Parámetros Epigenéticos" son, en particular, metilaciones de citosina. Parámetros epigenéticos adicionales incluyen, por ejemplo, la acetilación de histonas que, sin embargo, no puede analizarse directamente usando el método descrito pero que, a su vez, se correlaciona con la metilación del ADN.

El término "reactivo de bisulfito" se refiere a un reactivo que comprende bisulfito, disulfito, sulfito de hidrógeno o combinaciones de estos, útiles como se describe en la presente descripción para distinguir entre secuencias de dinucleótidos CpGs metiladas y no metiladas.

30 El término "ensayo de metilación" se refiere a cualquier ensayo para determinar el estado de metilación de una o más secuencias de dinucleótidos CpGs dentro de una secuencia de ADN.

35 El término "MS.AP-PCR" (Reacción en Cadena de la Polimerasa arbitrariamente iniciada sensible a la metilación) se refiere a la tecnología reconocida en la técnica que permite un tamizaje global del genoma usando iniciadores ricos en CG para enfocarse en las regiones con mayor probabilidad de contener dinucleótidos CpG, y descritas en Gonzalgo y otros, Cancer Research 57:594-599, 1997.

40 El término "MethyLight™" se refiere a la técnica de PCR basada en fluorescencia en tiempo real reconocida en la técnica descrita por Eads y otros, Cancer Res. 59:2302-2306, 1999.

45 El término ensayo "HeavyMethyl™", en la modalidad de este implementada en la presente descripción, se refiere a un ensayo, en donde las sondas *que bloquean* la metilación específica (referidas también en la presente descripción como *bloqueadores*) revisten las posiciones CpG entre, o revestidas por los iniciadores de la amplificación permitiendo la amplificación selectiva específica de la metilación de una muestra de ácido nucleico.

El término ensayo "HeavyMethyl™ MethyLight™", en la modalidad de este implementada en la presente descripción, se refiere a un ensayo HeavyMethyl™ MethyLight™, que es una variación del ensayo MethyLight™, en donde el ensayo MethyLight™ se combina con las sondas *que bloquean* la metilación específica que reviste las posiciones CpG entre los iniciadores de la amplificación

50 El término "Ms-SNuPE" (extensión del iniciador de un solo nucleótido sensible a metilación) se refiere al ensayo de reconocido en la técnica descrito por Gonzalgo & Jones, Nucleic Acids Res. 25:2529-2531, 1997.

55 El término "MSP" (PCR específica de metilación) se refiere al ensayo de metilación reconocido en la técnica descrito por Herman y otros Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996, y por patente de Estados Unidos núm. 5,786,146.

El término "COBRA" (análisis de restricción de bisulfito combinado) se refiere al ensayo de metilación reconocido en la técnica descrito por Xiong & Laird, Nucleic Acids Res. 25:2532-2534, 1997.

El término "MCA" (amplificación de la Isla de CpG metilada) se refiere al ensayo de metilación reconocido en la técnica descrito por Toyota y otros, Cancer Res. 59:2307-12, 1999, y en WO 00/26401A1.

- 5 El término "hibridación" debe entenderse como un enlace de un oligonucleótido con una secuencia complementaria a lo largo de las líneas de apareamientos de bases Watson-Crick en el ADN muestra, formando una estructura bicatenaria.

- 10 "Condiciones rigurosas de hibridación," tal como se definen en la presente descripción, implican hibridar a 68 °C en solución de SSC 5x/Denhardt's 5x/SDS 1.0%, y lavar en SSC 0.2x /SDS 0.1% a temperatura ambiente, o implican el equivalente de estas reconocidas en la técnica (por ejemplo, condiciones en las que una hibridación se lleva a cabo a 60 °C en amortiguador SSC 2.5 x, seguido por varias etapas de lavado a 37 °C en una concentración baja de amortiguador, y permanece estable). Condiciones moderadamente rigurosas, como se definen en la presente descripción, implican lavar incluir el lavado en SSC 3X a 42 °C, o lo equivalente de esto reconocido en la técnica. Los parámetros de concentración de sal y la temperatura se pueden variar para alcanzar el nivel óptimo de identidad entre la sonda y el ácido nucleico objetivo.
- 15 Orientación sobre dichas condiciones está disponible en la técnica, por ejemplo, en Sambrook y otros, 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Nueva York.; y Ausubel y otros (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, Nueva York) en la unidad 2.10.

- 20 Los términos "SEQ ID NO de matriz," "SEQ ID NO de matriz compuesta," o "secuencia de matriz compuesta" se refieren a una secuencia, hipotética o de cualquier otra forma, que consiste en un compuesto lineal de cabeza a cola (5' a 3') de todas las secuencias contiguas individuales de una matriz de sujeto (por ejemplo, un compuesto de cabeza a cola de SEQ ID NO: 1-71, en ese orden).

- 25 Los términos "SEQ ID NO de nodo de matriz," "SEQ ID NO de nodo de matriz compuesta," o "secuencia de nodo de matriz compuesta" se refiere a la *unión* entre dos secuencias contiguas individuales de la "SEQ ID NO de matriz," la "SEQ ID NO de matriz compuesta," o la "secuencia de matriz compuesta."

- 30 Con respecto a las secuencias de matriz compuesta, la frase "nucleótidos contiguos" se refiere a una región de la secuencia contigua de cualquier secuencia contigua individual de la matriz compuesta, pero no incluye una región de la secuencia de matriz compuesta que incluye un "nodo", como se definió en la presente descripción anteriormente.

Información general:

- 35 El alcance de la presente invención se define en las reivindicaciones anexas. También se describen marcadores genéticos moleculares que tienen utilidad novedosa para proporcionar un pronóstico de los trastornos proliferativos de las células de la próstata. En modalidades particulares dichos marcadores se pueden usar para clasificar el tumor de acuerdo con la agresividad y/o invasividad. Es particularmente preferido que el método y ácidos nucleicos descritos en la presente se utilicen en al menos uno del: pronóstico de; tratamiento de; seguimiento de; y tratamiento y seguimiento de los trastornos proliferativos de células de próstata.

- 40 El término 'pronóstico' se toma en el sentido de una predicción del resultado de la progresión de la enfermedad (en donde el término progresión se tomará para incluir también la recurrencia después del tratamiento). El pronóstico puede expresarse en términos de la supervivencia general del paciente, supervivencia sin enfermedad o recaída, complicaciones relacionadas con el tumor aumentadas y la tasa de progresión del tumor o metástasis, en donde una disminución de cualquiera de dichos factores (con la excepción del aumento de la tasa de progresión de complicaciones relacionadas con el tumor) con respecto
- 45 a un nivel predeterminado, es un resultado "negativo" y el aumento de estos es un resultado "positivo". Una disminución de las complicaciones relacionadas con el tumor y/o tasa de progresión del tumor o metástasis en relación a un nivel predeterminado, se considera un resultado 'positivo' y el aumento de estos es un resultado 'negativo'.

- 50 De aquí en adelante, el pronóstico también puede referirse en términos de "agresividad" en donde un cáncer agresivo se determina que tiene un alto riesgo de resultado negativo y en donde un cáncer no agresivo tiene un bajo riesgo de resultado negativo.

- 55 En un aspecto el marcador de pronóstico descrito en la presente se usa para proporcionar una estimación del riesgo de resultado negativo. La caracterización de un cáncer de próstata en términos del resultado previsto permite al médico determinar el riesgo de recurrencia y/o muerte. Esto ayuda en la selección del tratamiento ya que la reducción absoluta del riesgo de recurrencia y muerte después de los tratamientos tales como adyuvante hormonal, quimioterapia y radioterapia se puede determinar basado en el resultado negativo previsto. La reducción absoluta del riesgo atribuible al tratamiento puede

compararse después con los inconvenientes de dicho tratamiento (por ejemplo, efectos secundarios, costo) para determinar la idoneidad de dicho tratamiento para el paciente.

- 5 Por el contrario, en donde un cáncer se caracteriza como no agresivo (es decir, el resultado positivo con bajo riesgo de muerte y/o recurrencia) el paciente obtendrá bajo beneficio absoluto del adyuvante u otro tratamiento y puede tratarse adecuadamente por la espera vigilante. En la misma se encuentra una gran ventaja de la presente invención. Al proporcionar un medio para determinar qué pacientes no se beneficiarán significativamente del tratamiento de la presente invención se identifica los candidatos adecuados para la espera vigilante y se previene la sobreprescripción de terapias.
- 10 De acuerdo con el resultado previsto (es decir, el pronóstico) de la enfermedad se pueden seleccionar un tratamiento o tratamientos adecuados. En donde un cáncer se caracteriza como agresivo es particularmente preferido que el tratamiento adyuvante tal como, pero sin limitarse a, hormonal, quimioterapia o radioterapia se proporciona adicionalmente o en lugar de otros tratamientos.
- 15 Los marcadores descritos en la presente tienen más utilidad aun en la previsión del resultado de un paciente después del tratamiento quirúrgico. De aquí en adelante se referirá también como un marcador 'predictivo'. La sobreexpresión de los genes según la Tabla 11 (particularmente FOXL2, SEQ ID NO:35, SEQ ID NO:63, HIST2H2BF, GPR7 y con la máxima preferencia PITX2), están asociados con el resultado negativo de los pacientes con cáncer de próstata. Los pacientes con resultado predictivo positivo (es decir, hipometilación o sobre-expresión) tendrán como consecuencia después de dicho
- 20 tratamiento, una reducción absoluta disminuida del riesgo de recurrencia y muerte después del tratamiento con las terapias de adyuvantes post-quirúrgicas. Los pacientes con resultado negativo previsto (es decir, la hipermetilación) después de dicho tratamiento deberá proceder de una reducción absoluta relativamente mayor de riesgo de recurrencia y muerte después del tratamiento adyuvante postquirúrgica. Como consecuencia los pacientes con un resultado negativo después de dicho tratamiento se considerarán candidatos más adecuados para el tratamiento con adyuvante de pacientes con un
- 25 resultado positivo. Los pacientes con un resultado positivo como consecuencia se pueden prevenir de la sobre-prescripción del tratamiento con adyuvante.

La modificación del ADN con bisulfito es una herramienta reconocida en la técnica usada para evaluar la condición de metilación CpG. 5-metilcitosina es la modificación de base covalente más frecuente en el ADN de las células eucariotas. Esta juega un papel, por ejemplo, en la regulación de la transcripción, en la impronta genética, y en la tumorigénesis. Por lo tanto, la identificación de la 5-metilcitosina como un componente de la información genética es de gran interés. Sin embargo, las posiciones 5-metilcitosina pueden no identificarse por secuenciación, porque 5-metilcitosina tiene el mismo comportamiento de apareamiento de bases que la citosina. Además, la información epigenética realizada por la 5-metilcitosina se pierde por completo durante, *por ejemplo*, la amplificación por PCR.

- 35 El método más frecuentemente usado para analizar ADN para la presencia de 5-metilcitosina se basa en la reacción específica de bisulfito con citosina por lo que, tras la subsiguiente hidrólisis alcalina, la citosina se convierte en uracilo, que se corresponde con timina en su comportamiento de apareamiento de bases. Significativamente, sin embargo, la 5-metilcitosina permanece sin modificarse bajo estas condiciones. Por consiguiente, el ADN original se *convierte* de tal manera que la metilcitosina, que originalmente no podía distinguirse de la citosina por su comportamiento de hibridación, ahora puede detectarse ya que sólo la citosina permanece usando técnicas de biología molecular estándares, reconocidas en la técnica, por ejemplo, por amplificación e hibridación, o por secuenciación. Todas estas técnicas se basan en las propiedades diferenciales de apareamiento de bases, que pueden ahora explotarse plenamente.
- 40 La técnica anterior, en términos de sensibilidad, se define por un método que comprende encerrar el ADN que se analiza en una matriz de agarosa, evitando así la difusión y renaturalización del ADN (el bisulfito sólo reacciona con el ADN monocatenario), y sustituir todas las etapas de precipitación y purificación con diálisis rápida (Olek A, y otros, A modified and improved method for bisulfite based cytosine methylation analysis, Nucleic Acids Res. 24:5064-6, 1996). Así, es posible analizar las células individuales para la condición de metilación, ilustrando la utilidad y sensibilidad del método. Una visión general de los métodos reconocidos en la técnica para la detección de 5-metilcitosina se proporciona por Rein, T., y otros, Nucleic Acids Res., 26:2255, 1998.
- 45
- 50

La técnica de bisulfito, salvo pocas excepciones (*por ejemplo*., Zeschngk M, y otros, Eur J Hum Genet. 5:94-98, 1997), actualmente sólo se utiliza en la investigación. En todos los casos, los fragmentos cortos, específicos de un gen conocido son amplificados después de un tratamiento con bisulfito, y, o bien completamente secuenciados (Olek & Walter, Nat Genet. 1997 17:275-6, 1997), sometidos a una o más reacciones de extensión del iniciador (Gonzalzo & Jones, Nucleic Acids Res., 25:2529-31, 1997; WO 95/00669; patente de los Estados Unidos núm. 6,251,594) para analizar las posiciones de citosina individuales, o tratado por digestión enzimática (Xiong & Laird, Nucleic Acids Res., 25:2532-4, 1997). La detección por

hibridación también se ha descrito en la técnica (Olek y otros, WO 99/28498). Además, se ha descrito el uso de la técnica de bisulfito para la detección de metilación con respecto a los genes individuales (Grigg & Clark, Bioessays, 16:431-6, 1994; Zeschknig M, y otros, Hum Mol Genet., 6:387-95, 1997; Feil R, y otros, Nucleic Acids Res., 22:695-, 1994; Martin V, y otros, Gene, 157:261-4, 1995; WO 9746705 y WO 9515373).

También se describe el uso de la técnica de bisulfito, en combinación con uno o más ensayos de metilación, para la determinación del estado de metilación de las secuencias de dinucleótido CpG dentro de las secuencias del grupo que consiste en SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO:961. Preferentemente dicho grupo consiste en las SEQ ID Nos: 35, 63, 19 y con la máxima preferencia dicha secuencia es la SEQ ID NO:961. De acuerdo con la presente descripción, la determinación del estado de metilación de las secuencias de dinucleótido CpG dentro de las secuencias del grupo que consiste en SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961 y SEQ ID NO: 961 tiene utilidad pronóstica.

Procedimientos de ensayo de metilación. Varios procedimientos de ensayos de metilación se conocen en la técnica, y pueden usarse junto con la presente descripción. Estos ensayos permiten la determinación del estado de metilación de uno o una pluralidad de dinucleótidos CpG (por ejemplo, islas de CpG) dentro de una secuencia de ADN. Tales ensayos implican, entre otras técnicas, la secuenciación de ADN del ADN tratado con bisulfito, PCR (para la amplificación específica de secuencia), análisis de transferencia Southern, y uso de enzimas de restricción sensibles a la metilación.

Por ejemplo, la secuenciación genómica se ha simplificado para el análisis de los patrones de metilación del ADN y la distribución de 5-metilcitosina usando el tratamiento con bisulfito (Frommer y otros, Proc. Natl. Acad. Sci. USA 89:1827-1831, 1992). Además, se usa la digestión con enzimas de restricción de los productos de PCR amplificados a partir del ADN convertido con bisulfito, *por ejemplo*, el método descrito por Sadri & Hornsby (Nucl. Acids Res. 24:5058-5059, 1996), o COBRA (análisis de restricción de bisulfito combinado) (Xiong & Laird, Nucleic Acids Res. 25:2532-2534, 1997).

COBRA. El análisis COBRA es un ensayo de metilación cuantitativo útil para determinar los niveles de metilación de ADN en los loci de gen específico en cantidades pequeñas de ADN genómico (Xiong & Laird, Nucleic Acids Res. 25:2532-2534, 1997). En resumen, la digestión con enzimas de restricción se usa para revelar las diferencias en la secuencia dependiente de la metilación en los productos de PCR de ADN tratado con bisulfito sódico. Las diferencias en la secuencia dependiente de la metilación se introducen primero en el ADN genómico por tratamiento con bisulfito estándar de acuerdo con el procedimiento descrito por Frommer y otros (Proc. Natl. Acad. Sci. USA 89:1827-1831, 1992). La amplificación por PCR del ADN convertido con bisulfito se realiza después usando iniciadores específicos para las islas de CpG de interés, seguido de la digestión con endonucleasas de restricción, electroforesis en gel, y detección usando sondas de hibridación marcadas y específicas. Los niveles de metilación en la muestra de ADN original se representan por las cantidades relativas de producto de PCR digerido y no digerido de una manera linealmente cuantitativa a través de un amplio espectro de niveles de metilación de ADN. Además, esta técnica se puede aplicar fácilmente al ADN obtenido a partir de las muestras de tejidos microdisecadas embebidas en parafina. Los reactivos típicos (*por ejemplo*, como pueden encontrarse en un kit típico a base de COBRA) para el análisis COBRA pueden incluir, pero sin limitarse a: iniciadores de PCR para el gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); enzima de restricción y amortiguador apropiado; oligo del gen de hibridación; oligo control de hibridación; kit de marcaje de quinasa para sonda de oligo; y nucleótidos marcados. Además, los reactivos de conversión con bisulfito pueden incluir: amortiguador de desnaturalización de ADN; amortiguador sulfonación; reactivos o kits de recuperación de ADN (*por ejemplo*, precipitación, ultrafiltración, columna de afinidad); amortiguador desulfonación; y componentes de recuperación de ADN.

Preferentemente, los ensayos tales como "MethyLightTM" (una técnica de PCR en tiempo real basada en fluorescencia) (Eads y otros, Cancer Res. 59:2302-2306, 1999), reacciones de Ms-SNuPE (extensión del iniciador de un solo nucleótido sensible a metilación) (Gonzalzo & Jones, Nucleic Acids Res. 25:2529-2531, 1997), PCR específica de metilación ("MSP"; Herman y otros, Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996; patente de Estados Unidos núm. 5,786,146), y amplificación de la Isla de CpG metilada ("MCA"; Toyota y otros, Cancer Res. 59:2307-12, 1999) se usan solos o en combinación con otros de estos métodos.

MethyLightTM. El ensayo MethyLightTM es un ensayo de metilación cuantitativa de alto rendimiento que utiliza tecnología de PCR en tiempo real basada en fluorescencia (TaqManTM) que no requiere de manipulaciones adicionales después de la etapa de PCR (Eads y otros, Cancer Res. 59:2302-2306, 1999). En resumen, el proceso MethyLightTM comienza con una muestra mixta de ADN genómico que se convierte, con una reacción de bisulfito sódico, en una mezcla mixta de diferencias en la secuencia dependientes de la metilación de acuerdo con procedimientos estándar (el proceso de bisulfito convierte los residuos de citosina no metiladas a uracilo). El PCR basado en fluorescencia se realiza después ya sea en una reacción de PCR "no-sesgada" (con los iniciadores que no solapan sitios de metilación CpG conocidos), o en una reacción "sesgada"

(con iniciadores de PCR que solapan los dinucleótidos CpG conocidos). La discriminación de secuencia puede ocurrir ya sea a nivel del proceso de amplificación o a nivel del proceso de la detección de fluorescencia, o ambos.

El ensayo MethyLight™ puede usarse como una prueba cuantitativa de los patrones de metilación en la muestra de ADN genómico, en donde la discriminación de secuencia ocurre a nivel de la hibridación de la sonda. En esta versión cuantitativa, la reacción de PCR se proporciona para la amplificación no-sesgada en presencia de una sonda fluorescente que solapa un sitio potencial de metilación en particular. Se proporciona un control no-sesgado para la cantidad de ADN de entrada por una reacción en la que ni los iniciadores, ni la sonda revisten ninguno de los dinucleótidos CpG. Alternativamente, una prueba cualitativa para la metilación genómica se logra mediante el sondeo de la mezcla de PCR sesgada, ya sea con oligonucleótidos control que no "revisten" sitios de metilación conocidos (una versión basado en la fluorescencia de la técnica de "MSP"), o con oligonucleótidos que revisten los posibles sitios de metilación.

El proceso MethyLight™ puede usarse con una sonda de "TaqMan®" en el proceso de amplificación. Por ejemplo, el ADN genómico bicatenario se trata con bisulfito sódico y se sometió a uno de dos conjuntos de PCR reacciones usando las sondas TaqMan®; por ejemplo, ya sea con iniciadores sesgados y sonda TaqMan®, o iniciadores no-sesgados y sonda TaqMan®. La sonda TaqMan® se marca doble con las moléculas fluorescentes "reportero" e "inactivador", y se diseña para ser específica de una región GC de contenido relativamente alto tal que se funde a aproximadamente a 10°C de temperatura más alta en el ciclo de PCR que los iniciadores directo o reverso. Esto permite que la sonda TaqMan® permanezca completamente hibridada durante la etapa de apareamiento/extensión del PCR. Debido a que la Taq polimerasa sintetiza enzimáticamente una nueva cadena durante el PCR, eventualmente se logrará la sonda TaqMan® apareada. La actividad endonucleasa 5' a 3' de la Taq polimerasa desplazará después la sonda TaqMan® mediante la digestión para liberar la molécula reportera fluorescente para la detección cuantitativa de su señal ahora inactivada usando un sistema de detección de fluorescencia en tiempo real.

Reactivos típicos (*por ejemplo*, como se pueden encontrar en un típico kit basado en MethyLight™) para el análisis de MethyLight™ pueden incluir, pero sin limitarse a: iniciadores de PCR para el gen específico (o secuencia de ADN tratada con bisulfito o isla de CpG); sondas TaqMan®; amortiguadores y desoxinucleótidos de PCR optimizados; y Taq polimerasa.

Ms-SNuPE. La técnica Ms-SNuPE es un método cuantitativo para evaluar diferencias de metilación en sitios CpG específicos basados en el tratamiento del ADN con bisulfito, seguido por la extensión del iniciador de un solo nucleótido (Gonzalzo & Jones, Nucleic Acids Res. 25:2529-2531, 1997). En resumen, el ADN genómico se hace reaccionar con bisulfito sódico para convertir la citosina no metilada a uracilo, dejando 5-metilcitosina sin cambios. La amplificación de la secuencia objetivo deseada se realiza después usando iniciadores de PCR específicos para ADN convertido con bisulfito, y el producto resultante se aísla y se usa como molde para el análisis de metilación en el(os) sitio(s) CpG de interés. Pequeñas cantidades de ADN se pueden analizar (*por ejemplo*., secciones de patología microdisecadas), y esto evita la utilización de enzimas de restricción para determinar la condición de metilación en los sitios CpG.

Los reactivos típicos (*por ejemplo*, como pueden encontrarse en un kit típico basado en Ms-SNuPE) para el análisis de Ms-SNuPE pueden incluir, pero sin limitarse: iniciadores de PCR para el gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); amortiguadores y desoxinucleótidos de PCR optimizados; kit de extracción en gel; iniciadores control positivo; iniciadores Ms-SNuPE para gen específico; amortiguador de reacción (para la reacción Ms-SNuPE); y nucleótidos marcados. Además, los reactivos de conversión con bisulfito pueden incluir: amortiguador de desnaturalización de ADN; amortiguador sulfonación; reactivos o kit de recuperación de ADN (*por ejemplo*, precipitación, ultrafiltración, columna de afinidad); amortiguador desulfonación; y componentes de recuperación de ADN.

MSP. MSP (PCR específica de la metilación) permite evaluar virtualmente la condición de metilación de cualquier grupo de sitios CpG dentro de una isla de CpG, independiente del uso de enzimas de restricción sensibles a la metilación (Herman y otros Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996; patente de Estados Unidos núm. 5,786,146). En resumen, el ADN se modifica por bisulfito sódico convirtiendo todas las citosinas no metiladas, pero no las metiladas en uracilo, y posteriormente se amplifica con iniciadores específicos para ADN metilado frente al no metilado. MSP requiere solamente pequeñas cantidades de ADN, es sensible a 0.1% de alelos metilados de un locus determinado de la isla de CpG, y se puede realizar en el ADN extraído de muestras incluidas en parafina. Reactivos típicos (*por ejemplo*, como se pueden encontrar en un típico kit basado en MSP) para el análisis de MSP pueden incluir, pero sin limitarse a: iniciadores de PCR metilados y no metilados para el gen específico (o secuencia de ADN tratada con bisulfito o isla de CpG), amortiguadores y desoxinucleótidos de PCR optimizados, y sondas específicas.

MCA. La técnica de MCA es un método que se puede usar en el tamizaje de patrones de metilación alterados en el ADN genómico, y para aislar secuencias específicas asociadas con estos cambios (Toyota y otros, Cancer Res. 59:2307-12,

1999). En resumen, las enzimas de restricción con diferentes sensibilidades a la metilación de citosina en sus sitios de reconocimiento se usan para digerir los ADN genómicos de los tumores primarios, líneas celulares, y tejidos normales antes de la amplificación por PCR con cebado aleatorio. Los fragmentos que muestran metilación diferencial se clonan y secuencian después de resolver los productos de PCR en geles de poliacrilamida de alta resolución. Los fragmentos clonados se usan después como sondas para el análisis de Southern que confirma la metilación diferencial de estas regiones. Reactivos típicos (*por ejemplo*, como pueden encontrarse en un kit típico basado en MCA) para el análisis MCA pueden incluir, pero sin limitarse a: iniciadores de PCR para el ADN genómico de cebado aleatorio; amortiguadores de PCR y nucleótidos, enzimas de restricción y amortiguadores apropiados; oligos o sondas de hibridación de genes; oligos o sondas de hibridación control.

Se determinó que las secuencias genómicas de acuerdo con las SEQ ID NOS:1-64 y SEQ ID NO: 961 (preferentemente SEQ ID Nos:35,63,19 y con la máxima preferencia SEQ ID NO:961), y variantes tratadas de estas no naturales de estas de acuerdo con las SEQ ID NOS:65-320 y SEQ ID Nos: 962 - 965 (preferentemente SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia SEQ ID Nos: 962 - 965), tienen utilidad para proporcionar un pronóstico y/o tratamiento de los trastornos proliferativos de las células de la próstata.

También se describe un método para proporcionar un pronóstico de los trastornos proliferativos de células de próstata en un sujeto. En una modalidad particularmente preferida, se proporciona un método para la clasificación basada en la agresividad de un trastorno proliferativo de células de próstata.

Dicho método comprende las siguientes etapas:

- i) determinar los niveles de expresión de uno o más genes o secuencias de genes de acuerdo con la Tabla 11 y/o regiones reguladoras de estos; y
- ii) determinar el pronóstico de dichos trastornos proliferativos de células de próstata de acuerdo con dicho nivel de expresión.

Dicho nivel de expresión puede determinarse por cualquier medio estándar en la técnica, incluyendo pero sin limitarse al análisis de metilación, la pérdida de la heterocigosidad (de aquí en adelante también referido como LOH), los niveles de expresión del ARN y los niveles de expresión de proteínas.

Como consecuencia dicho método puede ser capaz por medio de cualquier análisis de la expresión de un ARN transcrito de este o polipéptido o proteína traducida a partir de dicho ARN, preferentemente por medio del análisis de la expresión del ARNm o el análisis de la expresión del polipéptido. Como consecuencia, la presente descripción proporciona también los ensayos y métodos de pronóstico, tanto cuantitativos como cualitativos para la detección de la expresión de los genes, secuencias genómicas y/o regiones reguladoras de acuerdo con la Tabla 11 en un sujeto con un carcinoma o neoplasia de próstata y determinación de estos tras el pronóstico y/o predicción del resultado del tratamiento en dicho sujeto. Particularmente se prefiere que dichos genes, secuencias genómicas y/o regiones reguladoras se seleccionen del grupo consistente en PITX2, SEQ ID NO:63, GPR7 y SEQ ID NO:35. Se prefiere más aun el gen PITX2.

La expresión aberrante del ARNm transcrito a partir de los genes o regiones genómicas de acuerdo con la Tabla 11, particularmente la SEQ ID NO:35, SEQ ID NO:63 y GPR7 y con la máxima preferencia PITX2 se asocian con el pronóstico y/o predicción del resultado de tratamiento del carcinoma de próstata.

Para detectar la presencia de ARNm que codifica un gen o secuencia genómica, se obtiene una muestra de un paciente. La muestra puede ser cualquier muestra adecuada que comprende la materia celular del tumor, con la máxima preferencia el tumor primario. Los tipos de muestras adecuadas incluyen las células tumorales o líneas celulares, portaobjetos histológicos, tejidos incluidos en parafina, biopsias, tejidos incluidos en parafina, fluidos corporales (tales como, pero sin limitarse a fluido de masaje prostático y orina) o cualquier otra muestra biológica adecuada y todas las combinaciones posibles de estos.

En una modalidad particularmente preferida del método dicha fuente es el tejido del tumor primario. La muestra puede ser tratada para extraer el ARN contenido en ella. Después se analiza el ácido nucleico resultante de la muestra. Muchas técnicas se conocen en el estado de la técnica para determinar niveles absolutos y relativos de la expresión génica, técnicas comúnmente usadas adecuadas para su uso en la presente invención incluyen la hibridación in situ (por ejemplo, FISH), análisis de Northern, ensayos de protección de ribonucleasa (RPA), microarreglos y técnicas basadas en PCR, tales como PCR cuantitativa y PCR de visualización diferenciada o cualquier otro método de detección de ácido nucleico.

Particularmente preferido es el uso de la técnica de reacción en cadena de polimerización/transcripción inversa (RT-PCR). El método de RT-PCR es bien conocido en la técnica (por ejemplo, ver Watson y Fleming, supra).

El método RT-PCR puede llevarse a cabo como sigue. El ARN celular total se aísla mediante, por ejemplo, el método de isotiocianato de guanidinio y el ARN total se transcribe inverso. El método de transcripción inversa implica la síntesis de ADN en un molde de ARN usando una enzima transcriptasa inversa y un iniciador oligo dT extremo 3' y/o iniciadores hexámeros aleatorios. El ADNc producido así se amplifica después por medio de PCR. (Belyavsky y otros, Nucl Acid Res 17:2919-2932, 1989; Krug y Berger, Methods in Enzymology, Academic Press, N.Y., Vol. 152, págs. 316-325, 1987). Además se prefiere la variante "tiempo real" de RT-PCR, en donde el producto de PCR se detecta por medio de sondas de hibridación (por ejemplo TaqMan, Lightcycler, Molecular Beacons & Scorpion) o SYBR verde. La señal detectada de las sondas o SYBR verde se cuantifica después ya sea por referencia a una curva estándar o mediante la comparación de los valores de Ct con los de un estándar de calibración. El análisis de los genes de mantenimiento se usa frecuentemente para normalizar los resultados.

En el análisis de transferencia Northern el ARNm total o poli(A) + se ejecuta en un gel de agarosa desnaturizante y se detecta por hibridación con una sonda marcada en el propio gel secado o en una membrana. La señal resultante es proporcional a la cantidad de ARN objetivo en la población de ARN.

Comparar las señales de dos o más poblaciones de células o tejidos revela diferencias relativas en los niveles de expresión génica. La cuantificación absoluta se puede realizar comparando la señal de una curva estándar generada usando cantidades conocidas de un transcrito in vitro que se corresponde con el ARN objetivo. El análisis de genes de mantenimiento, genes cuyos niveles de expresión se esperan que permanezcan relativamente constantes a pesar de las condiciones, se usa frecuentemente para normalizar los resultados, eliminar cualquiera de las diferencias aparentes causadas por la transferencia dispar del ARN a la membrana o carga dispar de ARN en el gel.

La primera etapa en el análisis de Northern es aislar el ARN intacto, puro a partir de las células o tejido de interés. Debido a que las membranas de Northern distinguen los ARN por tamaño, la integridad de la muestra influye el grado para el cual una señal se localiza en una única banda. Muestras de ARN parcialmente degradadas resultarán en la señal siendo dispersada o distribuida sobre diversas bandas con una pérdida total en la sensibilidad y posiblemente una interpretación errónea de los datos. En el análisis de transferencia de Northern, se pueden usar ADN, ARN y sondas de oligonucleótidos y estas sondas son preferentemente marcadas (por ejemplo marcadores radioactivos, marcadores de masa o marcadores fluorescentes). El tamaño del ARN objetivo, no la sonda, determinará el tamaño de la banda detectada, pero los métodos tal como el marcaje de cebado aleatorio, que genera sondas de longitudes variables, son adecuadas para la síntesis de sondas. La actividad específica de la sonda determinará el nivel de sensibilidad, pero se prefiere que se usen las sondas con actividades específicas altas.

En un ensayo de protección de ribonucleasa, el ARN objetivo y una sonda de ARN de una longitud definida se hibridan en solución. Después de la hibridación, el ARN se digiere con ribonucleasas específicas en ácidos nucleicos monocatenarios para eliminar cualquier sonda y ARN objetivo monocatenario no hibridado. Las ribonucleasas se inactivan, y el ARN se separa por ejemplo, por electroforesis en gel de poli(acrilamida) desnaturizante. La cantidad de sonda de ARN intacto es proporcional a la cantidad de ARN objetivo en la población de ARN. RPA puede usarse para la cuantificación absoluta y relativa de la expresión génica y también para el mapeo de la estructura del ARN, tales como los límites de intrón/exón y sitios de inicio de transcripción. Se prefiere el ensayo de protección de ribonucleasa para el análisis de transferencia Northern, ya que generalmente tiene un límite inferior de detección.

Las sondas de ARN antisentido usadas en RPA se generan por transcripción in vitro de un molde de ADN con un punto final definido y están típicamente en el intervalo de 50-600 nucleótidos. El uso de sondas de ARN que incluyen las secuencias adicionales no homólogas al ARN objetivo permite que el fragmento protegido se distinga de la sonda de longitud completa. Las sondas de ARN se usan típicamente en lugar de sondas de ADN debido a la facilidad de generar sondas de ARN monocatenario y la reproducibilidad y fiabilidad del ARN: digestión de ARN bicatenario con ribonucleasas (Ausubel y otros 2003), particularmente preferidas son las sondas con altas actividades específicas.

Particularmente preferido es el uso de microarreglos. El proceso de análisis de microarreglos se puede dividir en dos partes principales. La primera es la inmovilización de las secuencias génicas conocidas en portaobjetos de vidrio u otro soporte sólido seguido por la hibridación del ADNc marcado con fluorescencia (que comprende las secuencias que se interrogan) con los genes conocidos inmovilizados en el portaobjetos de vidrio. Después de la hibridación, los arreglos se escanearon usando un escáner para microarreglos fluorescentes. El análisis de la intensidad de fluorescencia relativa de diferentes genes proporciona una medida de las diferencias en la expresión génica.

Matrices de ADN pueden generarse inmovilizando oligonucleótidos presintetizados en portaobjetos de vidrio preparados. En este caso, las secuencias de genes representativos se fabrican y preparan usando la síntesis estándar de oligonucleótidos y métodos de purificación. Estas secuencias de genes sintetizados son complementarias a los genes de interés (con la máxima preferencia PITX2) y tienden a ser secuencias más cortas en el intervalo de 25-70 nucleótidos. Alternativamente, los oligos inmovilizados pueden sintetizarse químicamente in situ sobre la superficie del portaobjetos. La síntesis de oligonucleótidos in situ implica la adición consecutiva de los nucleótidos adecuados a las manchas en el microarreglo; las manchas que no reciben un nucleótido se protegen durante cada etapa del proceso usando máscaras físicas o virtuales.

En experimentos de perfiles de expresión de microarreglos, los moldes de ARN usados son representativos del perfil de transcripción de las células o tejidos en estudio. El ARN se aisló por primera vez de las poblaciones de células o tejidos que se comparan. Cada muestra de ARN se usa después como un molde para generar el ADNc marcado con fluorescencia a través de una reacción de transcripción inversa. El marcaje fluorescente del ADNc puede realizarse ya sea por métodos de marcaje directo o marcaje indirecto. Durante el marcaje directo, los nucleótidos modificados con fluorescencia (por ejemplo, Cy[®]3- o Cy[®]5-dCTP) se incorporan directamente en el ADNc durante la transcripción inversa. Alternativamente, el marcaje indirecto se puede lograr mediante la incorporación de nucleótidos aminoalil-modificados durante la síntesis del ADNc y conjugar después un colorante N-hidroxisuccinimida (NHS)-éster al ADNc aminoalil-modificado después que se completa la reacción de transcripción inversa. Alternativamente, la sonda puede no-marcarse, pero puede ser detectable por la unión específica con un ligando que se marca, ya sea directa o indirectamente. Marcadores adecuados y los métodos para marcar ligandos (y sondas) se conocen en la técnica, e incluyen, por ejemplo, marcadores radiactivos que se pueden incorporar por métodos conocidos (por ejemplo, desplazamiento de mella o tratamiento con quinasas). Otros marcadores adecuados incluyen, pero sin limitarse a biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetanos activados), enzimas, anticuerpos, y similares.

Para realizar el análisis de la expresión génica diferencial, el ADNc generado a partir de diferentes muestras de ARN se marcan con Cy[®]3. El ADNc resultante marcado se purifica para eliminar los nucleótidos no incorporados, colorante libre y ARN residual. Después de la purificación, las muestras de ADNc marcadas se hibridan al microarreglo. La rigurosidad de la hibridación se determina por una serie de factores durante la hibridación y durante el procedimiento de lavado, incluyendo la temperatura, fuerza iónica, longitud de tiempo y concentración de formamida. Estos factores se describen en, por ejemplo, Sambrook y otros (Molecular Cloning: A Laboratory Manual, 2da ed., 1989). El microarreglo se escanea post-hibridación usando un escáner fluorescente para microarreglos. La intensidad de fluorescencia de cada punto indica el nivel de expresión de ese gen; manchas brillantes corresponden a genes fuertemente expresados, mientras que las manchas oscuras indican expresión débil.

Una vez que se obtienen las imágenes, se deben analizar los datos sin procesar. En primer lugar, la fluorescencia de fondo se debe sustraer de la fluorescencia de cada mancha. El dato se normaliza después con una secuencia control, tal como por ejemplo un ARN exógenamente añadido, o un panel de genes de mantenimiento para responder por cualquier hibridación no específica, imperfecciones de la matriz o variabilidad en la configuración de la matriz, marcaje del ADNc, hibridación o lavado. La normalización de datos permite que se comparen los resultados de múltiples matrices.

Se describen también métodos para la detección de la presencia del polipéptido codificado por dichas secuencias génicas en una muestra obtenida a partir de un paciente.

Niveles aberrantes de la expresión del polipéptido de los polipéptidos codificados por los genes y/o regiones genómicas de acuerdo con la Tabla 11 (particularmente SEQ ID NO:35, SEQ ID NO:63, GPR7 y con la máxima preferencia PITX2) se asocian con el cáncer de próstata y pronóstico de neoplasia y/o el resultado del tratamiento.

Cualquier método conocido en la técnica para detectar polipéptidos puede usarse. Tales métodos incluyen, pero sin limitarse a espectrometría de masa, inmunodifusión, inmunolectroforesis, métodos inmunoquímicos, ensayos de aglutinante-ligando, técnicas inmunohistoquímicas, ensayos de aglutinación y complemento (por ejemplo, ver Basic and Clinical Immunology, Sites and Terr, eds., Appleton & Lange, Norwalk, Conn. págs. 217-262, 1991). Se prefieren los métodos de inmunoensayo de ligando-aglutinante que incluyen la reacción de anticuerpos con un epítipo o epítopos y competitivamente desplazar un polipéptido marcado o derivado de este.

Ciertas modalidades comprenden el uso de anticuerpos específicos para los polipéptidos codificados por los genes y/o regiones genómicas de acuerdo con la Tabla 11 (en particular SEQ ID NO: 35, SEQ ID NO: 63, GPR7 y con la máxima preferencia PITX2).

Tales anticuerpos son útiles para el pronóstico del cáncer de próstata y/o aplicaciones predictivas. En determinadas modalidades la producción de anticuerpos monoclonales o policlonales se puede inducir mediante el uso del polipéptido codificado como un antígeno. Tales anticuerpos pueden ser a su vez usados para detectar polipéptidos expresados como marcadores para el pronóstico del cáncer de próstata. Los niveles presentes de tales polipéptidos se pueden cuantificar por métodos convencionales. La unión anticuerpo-polipéptido se puede detectar y cuantificar por una variedad de medios conocidos en la técnica, tales como el marcaje con ligandos fluorescentes o radiactivos. También se describen kits para realizar los procedimientos anteriormente mencionados, en donde tales kits contienen anticuerpos específicos para los polipéptidos investigados.

Numerosos inmunoensayos de unión de polipéptidos competitivos y no competitivos son bien conocidos en la técnica. Los anticuerpos empleados en tales ensayos pueden estar sin marcar, por ejemplo como se usan en las pruebas de

aglutinación, o marcados para su uso una amplia variedad de métodos de ensayo. Los marcadores que pueden usarse incluyen radionúclidos, enzimas, fluorescentes, quimioluminiscentes, sustratos enzimáticos o cofactores, inhibidores de enzimas, partículas, colorantes y similares. Los ensayos preferidos incluyen, pero sin limitarse a radioinmunoensayo (RIA), inmunoensayos enzimáticos, por ejemplo, ensayo por inmunoabsorción ligado a enzimas (ELISA), inmunoensayos fluorescentes y similares. Los anticuerpos policlonales o monoclonales o epítomos de estos se pueden fabricar para su uso en inmunoensayos mediante cualquiera de una serie de métodos conocidos en la técnica.

En una modalidad alternativa del método de las proteínas pueden detectarse por medio de análisis de inmunoelectrotransferencia. Dicho análisis es estándar en la técnica, en resumen, las proteínas se separan por medio de electroforesis por ejemplo, SDS-PAGE. Las proteínas separadas se transfieren después a una membrana adecuada (o papel), por ejemplo nitrocelulosa, que conserva la separación espacial lograda por la electroforesis. La membrana se incuba con una proteína genérica (por ejemplo, proteína de la leche) para unirse a los restantes sitios adhesivos en la membrana. Se añade después un anticuerpo específico a la proteína de interés, dicho anticuerpo marcado de manera detectable por ejemplo con colorantes o medios enzimáticos (por ejemplo, fosfatasa alcalina o peroxidasa de rábano picante). Después se detecta la localización del anticuerpo en la membrana.

En una modalidad alternativa del método de las proteínas se pueden detectar por medio de inmunohistoquímica (el uso de anticuerpos para sondear antígenos específicos en una muestra). Dicho análisis es estándar en la técnica, en donde la detección de antígenos en los tejidos se conoce como inmunohistoquímica, mientras que la detección en células cultivadas se denomina generalmente inmunocitoquímica. En resumen, el anticuerpo primario que se detecta mediante la unión a su antígeno específico. El complejo anticuerpo-antígeno se une después mediante un anticuerpo secundario conjugado con enzima. En la presencia del sustrato y cromógeno necesario se detecta la enzima unida de acuerdo con los depósitos de color en los sitios de unión antígeno-anticuerpo. Existe un amplio intervalo de tipos de muestra adecuadas, afinidad antígeno-anticuerpo, tipos de anticuerpos, y métodos de mejoramiento de la detección. Así las condiciones óptimas para la detección inmunohistoquímica o inmunocitoquímica se deben determinar por la persona con experiencia en la técnica para cada caso individual.

Un enfoque para preparar anticuerpos de un polipéptido es la selección y preparación de una secuencia de aminoácidos de la totalidad o parte del polipéptido, sintetizar químicamente la secuencia de aminoácidos e inyectarla en un animal apropiado, normalmente un conejo o un ratón (Milstein y Kohler Nature 256:495-497, 1975; Gutfre y Milstein, Methods in Enzymology: Immunochemical Techniques 73:1-46, Langone y Banatis eds., Academic Press, 1981). Los métodos para la preparación de los polipéptidos o epítomos de estos incluyen, pero sin limitarse a la síntesis química, técnicas de ADN recombinante o aislamiento de muestras biológicas.

En la última etapa del método se determina el pronóstico del paciente, según el cual la sobreexpresión es indicativa de mal pronóstico. El término sobreexpresión se tomará en el sentido de la expresión en un nivel detectada mayor que un valor de corte predeterminado que puede seleccionarse del grupo que consiste de la media, mediana o un valor umbral optimizado. También se describe un kit para su uso en proveer un pronóstico de un sujeto con cáncer de próstata, que comprende: un medio para detectar polipéptidos de un gen o región genómica de acuerdo con la Tabla 11 (particularmente SEQ ID NO:35, SEQ ID NO:63 y GPR7 y con la máxima preferencia PITX2). Los medios para detectar los polipéptidos comprenden, preferentemente, anticuerpos, derivados de anticuerpos, o fragmentos de anticuerpos. Los polipéptidos se detectan con la máxima preferencia a través de la inmunoelectrotransferencia usando un anticuerpo marcado. En otra modalidad el kit comprende además medios para obtener una muestra biológica del paciente. Se prefiere un kit, que comprende además un recipiente adecuado para contener los medios para detectar los polipéptidos en la muestra biológica del paciente, y con la máxima preferencia comprende adicionalmente instrucciones para el uso y la interpretación de los resultados del kit. En una modalidad preferida, el kit para su uso en la determinación de la estrategia de tratamiento para un paciente con cáncer de próstata o neoplasias, comprende: (a) un medio para detectar polipéptidos de un gen o genómico región de acuerdo con la Tabla 11 (particularmente SEQ ID NO:35, SEQ ID NO:63 y GPR7 y con la máxima preferencia PITX2); (b) un recipiente adecuado para contener dichos medios y la muestra biológica del paciente, que comprende los polipéptidos en donde los medios pueden formar complejos con los polipéptidos; (c) un medio para detectar los complejos de (b); y opcionalmente (d) instrucciones de uso e interpretación de los resultados del kit.

El kit puede contener también otros componentes tales como amortiguadores o soluciones convenientes para el bloqueo, lavado o recubrimiento, empaquetados en un recipiente separado.

También se describe un kit para su uso en proveer un pronóstico de un sujeto con cáncer de próstata, dicho kit que comprende: un medio para medir el nivel de transcripción de un gen o región genómica de acuerdo con la Tabla 11 (particularmente SEQ ID NO:35, SEQ ID NO:63 y GPR7 y con la máxima preferencia PITX2). En una modalidad preferida, los medios para medir el nivel de transcripción comprenden oligonucleótidos o polinucleótidos capaces de hibridar en condiciones rigurosas o moderadamente rigurosas a los productos de transcripción de un gen o región genómica de acuerdo con la Tabla 11 (en particular, SEQ ID NO: 35, SEQ ID NO: 63 y GPR7 y con la máxima preferencia PITX2). En una

modalidad con la máxima preferencia, el nivel de transcripción se determina mediante técnicas seleccionadas del grupo de análisis de transferencia Northern, PCR transcriptasa inversa, PCR en tiempo real, protección de ARNasa, y microarreglos. En otra modalidad el kit comprende además medios para obtener una muestra biológica del paciente. Se prefiere un kit, que comprende además un recipiente adecuado para contener los medios para medir el nivel de transcripción y la muestra biológica del paciente, y con la máxima preferencia comprende adicionalmente instrucciones para el uso y la interpretación de los resultados del kit.

En una modalidad preferida, el kit para su uso en la determinación de la estrategia de tratamiento para un paciente con cáncer de próstata comprende (a) una pluralidad de oligonucleótidos o polinucleótidos capaces de hibridar en condiciones rigurosas o moderadamente rigurosas a los productos de la transcripción de un gen o región genómica de acuerdo a la Tabla 11 (particularmente, SEQ ID NO: 35, SEQ ID NO: 63 y GPR7 y con la máxima preferencia PITX2); (b) un recipiente adecuado para contener los oligonucleótidos o polinucleótidos y una muestra biológica del paciente, que comprende los productos de transcripción en donde los oligonucleótidos o polinucleótido pueden hibridarse en condiciones rigurosas o moderadamente rigurosas a los productos de la transcripción, (c) medios para detectar la hibridación de (b); y, opcionalmente, (d) las instrucciones de uso e interpretación de los resultados del kit.

El kit también puede contener otros componentes, tales como amortiguador de hibridación (donde los oligonucleótidos son para usarse como una sonda) empaquetados en un recipiente separado. Alternativamente, cuando los oligonucleótidos son para usarse en amplificar una región objetivo, el kit puede contener, empaquetados en recipientes separados, una polimerasa y un amortiguador de reacción optimizado para la extensión del iniciador mediada por la polimerasa, tal como PCR.

Con la máxima preferencia un kit como se describe en la presente se usa para la determinación de la etapa de expresión de los métodos descritos en la presente.

También se describe un método adicional para proporcionar un pronóstico de un sujeto con cáncer de próstata que comprende las siguientes etapas. En la primera etapa del método se obtiene una muestra del sujeto. Los tipos de muestras adecuadas incluyen las células tumorales o líneas celulares, portaobjetos histológicos, tejidos incluidos en parafina, biopsias, tejidos incluidos en parafina, fluidos corporales (tales como, pero sin limitarse a fluido de masaje prostático y orina) o cualquier otra muestra biológica adecuada y todas las combinaciones posibles de estos. Las técnicas adecuadas comúnmente usadas incluyen para su uso la hibridación in situ (por ejemplo FISH), análisis de Northern, ensayos de protección de ribonucleasa (RPA), microarreglos y técnicas basadas en PCR, tales como PCR cuantitativa y PCR de visualización diferenciada o cualquier otro método de detección de ácido nucleico.

Particularmente preferido es el uso de la técnica de reacción en cadena polimerización/transcripción inversa (RT-PCR). El método de RT-PCR es bien conocido en la técnica (por ejemplo, ver Watson y Fleming, supra).

El método RT-PCR puede llevarse a cabo como sigue. El ARN celular total se aísla mediante, por ejemplo, el método de isotiocianato de guanidinio y el ARN total se transcribe inverso. El método de transcripción inversa implica la síntesis de ADN en un molde de ARN usando una enzima transcriptasa inversa y un iniciador oligo dT extremo 3' y/o iniciadores hexámeros aleatorios. El ADNc producido así se amplifica después por medio de PCR. (Belyavsky y otros, Nucl Acid Res 17:2919-2932, 1989; Krug y Berger, Methods in Enzymology, Academic Press, N.Y., Vol.152, págs. 316-325, 1987). Además se prefiere la variante "tiempo real" de RT-PCR, en donde el producto de PCR se detecta por medio de sondas de hibridación (por ejemplo TaqMan, Lightcycler, Molecular Beacons & Scorpion) o SYBR verde. La señal detectada de las sondas o SYBR verde se cuantifica después ya sea por referencia a una curva estándar o mediante la comparación de los valores de Ct con los de un estándar de calibración. EL análisis de los genes de mantenimiento se usa frecuentemente para normalizar los resultados.

En el análisis de transferencia Northern el ARNm total o poli(A)+ se ejecuta en un gel de agarosa desnaturizante y se detecta por hibridación con una sonda marcada en el propio gel secado o en una membrana. La señal resultante es proporcional a la cantidad de ARN objetivo en la población de ARN.

Comparar las señales de dos o más poblaciones de células o tejidos revela diferencias relativas en los niveles de expresión génica. La cuantificación absoluta se puede realizar comparando la señal de una curva estándar generada usando cantidades conocidas de un transcripto in vitro que se corresponde con el ARN objetivo. El análisis de genes de mantenimiento, genes cuyos niveles de expresión se esperan que permanezcan relativamente constantes a pesar de las condiciones, se usa frecuentemente para normalizar los resultados, eliminar cualquiera de las diferencias aparentes causadas por la transferencia dispar del ARN a la membrana o carga dispar de ARN en el gel.

La primera etapa en el análisis de Northern es aislar el ARN intacto, puro a partir de las células o tejido de interés. Debido a que las membranas de Northern distinguen los ARN por tamaño, la integridad de la muestra influye el grado para el cual una señal se localiza en una única banda. Muestras de ARN parcialmente degradadas resultarán en la señal siendo dispersada o distribuida sobre diversas bandas con una pérdida total en la sensibilidad y posiblemente una interpretación errónea de los datos. En el análisis de transferencia de Northern, se pueden usar ADN, ARN y sondas de oligonucleótidos y estas sondas

son preferentemente marcadas (por ejemplo marcadores radioactivos, marcadores de masa o marcadores fluorescentes). El tamaño del ARN objetivo, no la sonda, determinará el tamaño de la banda detectada, pero los métodos tal como el marcaje de cebado aleatorio, que genera sondas de longitudes variables, son adecuadas para la síntesis de sondas. La actividad específica de la sonda determinará el nivel de sensibilidad, pero se prefiere que se usen las sondas con actividades específicas altas.

En un ensayo de protección de ribonucleasa, el ARN objetivo y una sonda de ARN de una longitud definida se hibridan en solución. Después de la hibridación, el ARN se digiere con ribonucleasas específicas en ácidos nucleicos monocatenarios para eliminar cualquier sonda y ARN objetivo monocatenario no hibridado. Las ribonucleasas se inactivan, y el ARN se separa por ejemplo, por electroforesis en gel de poliacrilamida desnaturizante. La cantidad de sonda de ARN intacto es proporcional a la cantidad de ARN objetivo en la población de ARN. RPA puede usarse para la cuantificación absoluta y relativa de la expresión génica y también para el mapeo de la estructura del ARN, tales como los límites de intrón/exón y sitios de inicio de transcripción. Se prefiere el ensayo de protección de ribonucleasa para el análisis de transferencia Northern, ya que generalmente tiene un límite inferior de detección.

Las sondas de ARN antisentido usadas en RPA se generan por transcripción in vitro de un molde de ADN con un punto final definido y están típicamente en el intervalo de 50-600 nucleótidos. El uso de sondas de ARN que incluyen las secuencias adicionales no homólogas al ARN objetivo permite que el fragmento protegido se distinga de la sonda de longitud completa. Las sondas de ARN se usan típicamente en lugar de sondas de ADN debido a la facilidad de generar sondas de ARN monocatenario y la reproducibilidad y fiabilidad del ARN: digestión de ARN bicatenario con ribonucleasas (Ausubel y *otros* 2003), particularmente preferidas son las sondas con altas actividades específicas.

Particularmente preferido es el uso de microarreglos. El proceso de análisis de microarreglos se puede dividir en dos partes principales. La primera es la inmovilización de las secuencias génicas conocidas en portaobjetos de vidrio u otro soporte sólido seguido por la hibridación del ADNc marcado con fluorescencia (que comprende las secuencias que se interrogan) con los genes conocidos inmovilizados en el portaobjetos de vidrio. Después de la hibridación, los arreglos se escanearon usando un escáner para microarreglos fluorescentes. El análisis de la intensidad de fluorescencia relativa de diferentes genes proporciona una medida de las diferencias en la expresión génica.

Matrices de ADN pueden generarse inmovilizando oligonucleótidos presintetizados en portaobjetos de vidrio preparados. En este caso, las secuencias de genes representativos se fabrican y preparan usando la síntesis estándar de oligonucleótidos y métodos de purificación. Las secuencias de genes sintetizadas son complementarias a los genes de interés (tales como PITX2 u otros genes de acuerdo con la Tabla 11) y tienden a ser secuencias más cortas en el intervalo de 25-70 nucleótidos. Alternativamente, los oligos inmovilizados pueden sintetizarse químicamente in situ sobre la superficie del portaobjetos. La síntesis de oligonucleótidos in situ implica la adición consecutiva de los nucleótidos adecuados a las manchas en el microarreglo; las manchas que no reciben un nucleótido se protegen durante cada etapa del proceso usando máscaras físicas o virtuales.

En experimentos de perfiles de expresión de microarreglos, los moldes de ARN usados son representativos del perfil de transcripción de las células o tejidos en estudio. El ARN se aisló por primera vez de las poblaciones de células o tejidos que se comparan. Cada muestra de ARN se usa después como un molde para generar el ADNc marcado con fluorescencia a través de una reacción de transcripción inversa. El marcaje fluorescente del ADNc puede realizarse ya sea por métodos de marcaje directo o marcaje indirecto. Durante el marcaje directo, los nucleótidos modificados con fluorescencia (por ejemplo, Cy[®]3- o Cy[®]5-dCTP) se incorporan directamente en el ADNc durante la transcripción inversa. Alternativamente, el marcaje indirecto se puede lograr mediante la incorporación de nucleótidos aminoalil-modificados durante la síntesis del ADNc y conjugar después un colorante N-hidroxisuccinimida (NHS)-éster al ADNc aminoalil-modificado después que se completa la reacción de transcripción inversa. Alternativamente, la sonda puede no-marcarse, pero puede ser detectable por la unión específica con un ligando que se marca, ya sea directa o indirectamente. Marcadores adecuados y los métodos para marcar ligandos (y sondas) se conocen en la técnica, e incluyen, por ejemplo, marcadores radiactivos que se pueden incorporar por métodos conocidos (por ejemplo, desplazamiento de mella o tratamiento con quinasa). Otros marcadores adecuados incluyen, pero sin limitarse a biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetanos activados), enzimas, anticuerpos, y similares.

Para realizar el análisis de la expresión génica diferencial, el ADNc generado a partir de diferentes muestras de ARN se marcan con Cy[®]3. El ADNc resultante marcado se purifica para eliminar los nucleótidos no incorporados, colorante libre y ARN residual. Después de la purificación, las muestras de ADNc marcadas se hibridan al microarreglo. La rigurosidad de la hibridación se determina por una serie de factores durante la hibridación y durante el procedimiento de lavado, incluyendo la temperatura, fuerza iónica, longitud de tiempo y concentración de formamida. Estos factores se describen en, por ejemplo, Sambrook y otros (Molecular Cloning: A Laboratory Manual, 2da ed., 1989). El microarreglo se escanea post-hibridación usando un escáner fluorescente para microarreglos. La intensidad de fluorescencia de cada punto indica el nivel de expresión de ese gen; manchas brillantes corresponden a genes fuertemente expresados, mientras que las manchas oscuras indican expresión débil.

Una vez que se obtienen las imágenes, se deben analizar los datos sin procesar. En primer lugar, la fluorescencia de fondo se debe sustraer de la fluorescencia de cada mancha. El dato se normaliza después con una secuencia control, tal como por ejemplo un ARN exógenamente añadido, o un panel de genes de mantenimiento para responder por cualquier hibridación no específica, imperfecciones de la matriz o variabilidad en la configuración de la matriz, marcaje del ADNc, hibridación o lavado. La normalización de datos permite que se comparen los resultados de múltiples matrices.

Se describen también métodos para la detección de la presencia del polipéptido codificado por dichas secuencias génicas en una muestra obtenida a partir de un paciente.

Niveles aberrantes de expresión del polipéptido de los polipéptidos codificados por el gen de acuerdo con la Tabla 11 (particularmente FOXL2, HIST2H2BF y GPR7 y con la máxima preferencia PITX2) se asocian con el cáncer de próstata y/o resultado del tratamiento.

Se puede usar cualquier método conocido en la técnica para detectar polipéptidos. Tales métodos incluyen, pero no sin limitarse a espectrometría de masa, inmunodifusión, inmunoelectroforesis, métodos inmunoquímicos, ensayos aglutinante-ligando, técnicas inmunohistoquímicas, ensayos de aglutinación y complemento (por ejemplo, ver Basic and Clinical Immunology, Sites and Terr, eds., Appleton & Lange, Norwalk, Conn. págs. 217-262, 1991). Son preferidos los métodos de inmunoensayo aglutinante-ligando que incluyen hacer reaccionar anticuerpos con un epítipo o epítipos y competitivamente desplazar un polipéptido marcado o derivado de este.

Ciertas modalidades comprenden el uso de anticuerpos específicos al polipéptido codificado por el gen seleccionado de la Tabla 11, preferentemente FOXL2, HIST2H2BF o GPR7 y con la máxima preferencia PITX2.

Tales anticuerpos son útiles para el pronóstico de cáncer de próstata y/o aplicaciones predictivas. En ciertas modalidades la producción de anticuerpos monoclonales o policlonales se pueden inducir por el uso del polipéptido codificado como un antígeno. Tales anticuerpos se pueden a su vez usar para detectar polipéptidos expresados como marcadores para el pronóstico del cáncer de próstata. Los niveles de tales polipéptidos presentes se pueden cuantificar por métodos convencionales. La unión de anticuerpo-polipéptido se puede detectar y cuantificar por una variedad de medios conocidos en la técnica, tales como el marcaje con ligandos fluorescentes o radiactivos. También se describen estuches para realizar los procedimientos anteriormente mencionados, en donde tales estuches contienen anticuerpos específicos para los polipéptidos investigados.

Numerosos inmunoensayos de unión de polipéptidos competitivos y no competitivos son bien conocidos en la técnica. Los anticuerpos empleados en tales ensayos se pueden marcar, por ejemplo como se usa en pruebas de aglutinación, o marcar para el uso de una amplia variedad de métodos de ensayo. Los marcadores que se pueden usar incluyen radionucleidos, enzimas, fluorescentes, quimioluminiscentes, sustratos enzimáticos o cofactores, inhibidores de enzimas, partículas, colorantes y similares. Los ensayos preferidos incluyen, pero sin limitarse a radioinmunoensayo (RIA), inmunoensayos enzimáticos, por ejemplo, ensayo inmunoabsorbente ligado a enzima (ELISA), inmunoensayos fluorescentes y similares. Los anticuerpos policlonales o monoclonales o epítipos de estos se pueden preparar para el uso en inmunoensayos mediante cualquiera de un número de métodos conocidos en la técnica.

En una modalidad alternativa del método las proteínas se pueden detectar por medio de análisis de membrana de Western. Dicho análisis es estándar en la técnica, las proteínas brevemente se separan por medio de electroforesis por ejemplo, SDS-PAGE. Las proteínas separadas se transfieren después a una membrana adecuada (o papel), por ejemplo nitrocelulosa, conservando la separación espacial lograda por la electroforesis. La membrana se incuba después con una proteína genérica (por ejemplo, proteína de la leche) para unirse quedando los lugares adhesivos en la membrana. Se añade después un anticuerpo específico a la proteína de interés, estando dicho anticuerpo marcado de manera detectable por ejemplo mediante medios colorantes o enzimáticos (por ejemplo, fosfatasa alcalina o peroxidasa de rábano picante). Se detecta después el lugar del anticuerpo en la membrana.

En una modalidad alternativa del método las proteínas se pueden detectar por medio de inmunohistoquímica (el uso de anticuerpos para sondear antígenos específicos en una muestra). Dicho análisis es estándar en la técnica, en donde la detección de antígenos en los tejidos se conoce como inmunohistoquímica, mientras que la detección en células cultivadas se denomina generalmente inmunocitoquímica. En resumen el anticuerpo primario se detecta mediante la unión a su antígeno específico. El complejo anticuerpo-antígeno se une después mediante un anticuerpo secundario conjugado a enzima. En presencia del sustrato necesario y cromógeno la enzima unida se detecta de acuerdo con los depósitos coloreados en los sitios de unión antígeno-anticuerpo. Existe una amplia variedad de tipos de muestras adecuadas, afinidad antígeno-anticuerpo, tipos de anticuerpo, y métodos mejorados de detección. Así las condiciones óptimas para la detección

inmunohistoquímica o inmunocitoquímica se deben determinar por la persona con experiencia en la técnica para cada caso individual.

Un enfoque para preparar anticuerpos para un polipéptido es la selección y preparación de una secuencia de aminoácidos de todo o parte del polipéptido, sintetizar químicamente la secuencia de aminoácidos e inyectarla en un animal adecuado, por lo general un conejo o un ratón (Milstein y Kohler Nature 256:495-497, 1975; Gutfre y Milstein, Methods in Enzymology: Immunochemical Techniques 73:1-46, Langone and Banatis eds., Academic Press, 1981). Los métodos para la preparación de los polipéptidos o epítomos de estos incluyen, pero sin limitarse a la síntesis química, técnicas de ADN recombinante o aislamiento a partir de muestras biológicas.

En una modalidad particularmente preferida, el nivel de expresión de los genes, secuencias genómicas y/o regiones reguladoras de acuerdo con la Tabla 11 se determina por análisis del nivel de metilación de dichos genes, secuencias genómicas y/o regiones reguladoras de estos se determine mediante la determinación de la condición o nivel de metilación de al menos un dinucleótido CpG de estos. Se prefiere adicionalmente que el nivel de metilación de dichos genes, secuencias genómicas y/o regiones reguladoras de estos se determine mediante la determinación de la condición o nivel de metilación de una pluralidad de dinucleótidos CpG de estos. Es particularmente preferido que dichos genes, secuencias genómicas y/o regiones reguladoras se seleccionan del grupo que consiste en PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35. Adicionalmente preferido es el gen PITX2. Dicho análisis comprende las siguientes etapas:

- i) poner en contacto el ADN genómico obtenido del sujeto con al menos un reactivo, o serie de reactivos que se distinguen entre los dinucleótidos CpG metilados y no metilados dentro de al menos una región objetivo del ADN genómico, en donde dichos nucleótidos contiguos comprenden al menos una secuencia de dinucleótido CpG; y
- ii) clasificar los trastornos proliferativos de célula de la próstata de acuerdo con su pronóstico como se determina a partir de la condición de metilación de dichas regiones objetivo analizada en i).

El ADN genómico se puede aislar mediante cualquier medio estándar en la técnica, que incluyen el uso de estuches disponibles en el comercio. En resumen, en donde el ADN de interés se encapsula mediante una membrana celular la muestra biológica se debe romper y lisar mediante medios enzimáticos, químicos o mecánicos. La solución de ADN se puede clarificar de las proteínas y otros contaminantes, por ejemplo mediante digestión con proteinasa K. El ADN genómico se recupera después de la solución. Esto puede llevarse a cabo por medio de una variedad de métodos que incluyen precipitación salina, extracción orgánica o unión de ADN a un soporte de fase sólida. La elección del método será afectada por varios factores, que incluyen el tiempo, costo y cantidad requerida de ADN. Preferentemente, se selecciona la fuente de la muestra de ADN del grupo que consiste de células o líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, eyaculado, orina, sangre, y combinaciones de estos. Preferentemente, la fuente es biopsias, fluidos corporales, eyaculado, orina o sangre. La muestra de ADN genómico se trata después de manera tal que las bases de citosina que están sin metilar en la posición 5' se convierten a uracilo, timina, u otra base que es muy diferente a la citosina en términos de comportamiento de hibridación. Esto se entenderá como "tratamiento" en la presente descripción.

El tratamiento del ADN genómico descrito anteriormente se lleva a cabo preferentemente con bisulfito (sulfito de hidrógeno, disulfito) y posterior hidrólisis alcalina que resulta en una conversión de nucleobases de citosina no metiladas a uracilo o a otra base que es diferente a la citosina en términos de comportamiento de apareamiento de bases.

Se analiza después el ADN tratado para determinar el estado de metilación de una o más secuencias de genes objetivo (antes del tratamiento) asociados con el desarrollo de carcinoma de la próstata. Se prefiere particularmente que la región objetivo comprende, o hibrida en condiciones rigurosas para al menos 16 nucleótidos contiguos de al menos un gen o secuencia genómica seleccionada a partir del grupo que consiste en los genes y secuencias genómicas que se listan en la Tabla 11. Se prefiere particularmente que dicho gen o secuencia genómica se seleccione del grupo que consiste en PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35. Adicionalmente preferido es el gen PITX2. Se prefiere adicionalmente que las secuencias de dichos genes se analizan como se describe en la lista de secuencias de acompañamiento. El método de análisis se puede seleccionar a partir de los conocidos en la técnica, que incluyen los listados en la presente descripción. Particularmente son preferidos MethyLight™, MSP™ y el uso de oligonucleótidos de bloqueo como se describe en la presente. Se prefiere adicionalmente que cualquier oligonucleótidos usados en tales análisis (que incluyen iniciadores, oligonucleótidos de bloqueo y sondas de detección) deben ser complementarios inversos, idénticos, o hibridar en condiciones rigurosas o altamente rigurosas con al menos un segmento largo de al menos 16 pares de bases de las secuencias de bases de una o más de SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965 y las secuencias complementarias a estas. Se prefiere que cualquier oligonucleótidos usados en tales análisis (que incluyen iniciadores, oligonucleótidos de bloqueo y sondas de detección) deben ser complementarios inversos, idénticos, o hibridar

en condiciones rigurosas o altamente rigurosas con un segmento largo de al menos 16 pares de bases de las secuencias bases de uno o más de SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia SEQ ID Nos: 962-965 y secuencias complementarias a estas.

5 La metilación aberrante, de uno o más genes o secuencias genómicas tomadas de los listados en la Tabla 11, y con mayor preferencia los genes o secuencias genómicas de acuerdo con PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35 se asocian con el pronóstico de trastornos proliferativos de células de la próstata. El análisis de uno o una pluralidad de secuencias permite la clasificación pronóstico de los trastornos proliferativos de células de la próstata. Con mayor preferencia, la hipermetilación de uno o más de dichos genes o secuencias genómicas se asocia con un mal pronóstico. La hipermetilación se asocia generalmente con baja expresión de ARNm y como consecuencia polipéptidos.

10 En una modalidad, el método describe el uso de uno o más genes o secuencias genómicas seleccionados a partir del grupo que consiste en los genes de acuerdo con la Tabla 11 como marcadores para proporcionar un pronóstico de los trastornos proliferativos de células de la próstata. Se prefiere particularmente que dicho gen o secuencia genómica se seleccione del grupo que consiste en PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35. Adicionalmente preferido es el gen PITX2.

15 Dicho uso de los genes y/o secuencias se puede activar por medio de cualquier análisis de la expresión del gen, por medio de análisis de la expresión de ARNm o análisis de la expresión de proteína. Sin embargo, en la modalidad más preferida, la detección de trastornos proliferativos de células de la próstata se activa por medio de análisis de la condición de metilación de dichos genes o secuencias genómicas y su promotor o elementos reguladores. Los métodos para el análisis de metilación de genes se describen en la presente.

20 En una modalidad, el método describe el uso de uno o más genes o secuencias genómicas seleccionados a partir del grupo que consiste de los genes de acuerdo con la Tabla 11 como marcadores para proporcionar un pronóstico de los trastornos proliferativos de células de la próstata. Se prefiere particularmente que dicho gen o secuencia genómica se seleccione del grupo que consiste en PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35. Adicionalmente preferido es el gen PITX2.

25 Dicho uso de los genes y/o secuencias se puede activar por medio de cualquier análisis de la expresión del gen, por medio de análisis de la expresión de ARNm o análisis de la expresión de proteína. Sin embargo, en la modalidad más preferida, la detección de trastornos proliferativos de células de la próstata se activa por medio de análisis de la condición de metilación de dichos genes o secuencias genómicas y su promotor o elementos reguladores. Los métodos para el análisis de metilación de genes se describen en la presente.

30 Niveles aberrantes de expresión de ARNm de los genes, las secuencias genómicas o genes regulados por secuencias genómicas de acuerdo con la Tabla 11 se asocian con el pronóstico de los trastornos proliferativos de células de la próstata. En consecuencia, los niveles aumentados o disminuidos de expresión de dichos genes o secuencias son asociables con factores asociados con el pronóstico de los trastornos proliferativos de células de la próstata, que incluyen, pero sin limitarse a la agresividad de la enfermedad y la progresión. Se prefiere particularmente que dicho gen o secuencia genómica se seleccione del grupo que consiste en PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35. Adicionalmente preferido es el gen PITX2.

35 Para detectar la presencia de ARNm que codifica un gen o secuencia genómica en una clasificación pronóstico de trastornos proliferativos de células de la próstata, se obtiene una muestra de un paciente. Preferentemente, se selecciona la fuente de la muestra de ADN del grupo que consiste de células o líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, eyaculado, orina, sangre, y combinaciones de estos. Preferentemente, la fuente es biopsias, fluidos corporales, eyaculado, orina o sangre. La muestra se puede tratar para extraer los ácidos nucleicos contenidos en la misma. El ácido nucleico resultante de la muestra se somete a electroforesis en gel u otras técnicas de separación. La detección implica poner en contacto los ácidos nucleicos y particularmente el ARNm de la muestra con una secuencia de ADN que sirve como sonda para formar dúplex híbridos. La rigurosidad de la hibridación se determina por un número de factores durante la hibridación y durante el procedimiento de lavado, que incluyen la temperatura, fuerza iónica, periodo de tiempo y concentración de formamida. Estos factores se describen en, por ejemplo, Sambrook y otros (Molecular Cloning: A Laboratory Manual, 2d ed., 1989). La detección del dúplex resultante se logra por lo general mediante el uso de sondas marcadas. Como alternativa, la sonda puede estar sin marcar, pero puede ser detectable por la unión específica con un ligando que se marca, ya sea directa o indirectamente. Los marcadores y métodos adecuados para marcar sondas y ligandos son conocidos en la técnica, e incluyen, por ejemplo, marcadores radiactivos que se pueden incorporar por métodos conocidos (*por ejemplo*, desplazamiento de mella o tratamiento con quinasa), biotina, grupos fluorescentes, grupos quimioluminiscentes (*por ejemplo*, dioxetanos, particularmente dioxetanos activados), enzimas, anticuerpos, y similares.

Para aumentar la sensibilidad de la detección en una muestra de ARNm transcrito a partir del gen o secuencia genómica, la técnica de reacción en cadena de polimerización/transcripción inversa se puede usar para amplificar ADNc transcrito a partir del ARNm. El método de transcripción inversa/PCR es bien conocido en la técnica (por ejemplo, ver Watson y Fleming, arriba).

El método de transcripción inversa/PCR puede llevarse a cabo como sigue. El ARN celular total se aísla mediante, por ejemplo, el método de isotiocianato de guanidinio y el ARN total se transcribe inverso. El método de transcripción inversa implica la síntesis de ADN en un molde de ARN usando una enzima transcriptasa inversa y un iniciador del extremo 3'. Típicamente, el iniciador contiene una secuencia oligo (dT). El ADNc así producido se amplifica después usando el método de PCR e iniciadores específicos EYA4. (Belyavsky y otros, Nucl Acid Res 17:2919-2932, 1989; Krug y Berger, Methods in Enzymology, Academic Press, N.Y., Vol.152, págs. 316-325, 1987).

También se describe un *estuche* para el uso al proporcionar un pronóstico de un estado de trastorno proliferativo de células de la próstata a través de pruebas de una muestra biológica. Un estuche representante puede comprender uno o más segmentos de ácido nucleico que se hibridan selectivamente al ARNm y un envase para cada uno de uno o más segmentos de ácido nucleico. En ciertas modalidades, los segmentos de ácido nucleico se pueden combinar en un solo tubo. En modalidades adicionales, los segmentos de ácido nucleico pueden incluir también un par de iniciadores para amplificar el ARNm objetivo. Tales estuches pueden incluir también cualquier amortiguadores, soluciones, disolventes, enzimas, nucleótidos, u otros componentes para las reacciones de hibridación, amplificación o detección. Los componentes del estuche preferidos incluyen reactivos para la transcripción inversa -PCR, hibridación in situ, análisis Northern y/o RPA.

También se describe una nueva aplicación del análisis de los niveles de metilación y/o patrones dentro de dichas secuencias que permite una clasificación pronóstico y de ese modo el tratamiento mejorado de los trastornos proliferativos de células de la próstata. El tratamiento de los trastornos proliferativos de células de la próstata se relaciona directamente con el pronóstico de la enfermedad particularmente, la agresividad, y el método descrito de ese modo permite al médico y paciente tomar más y mejores decisiones del tratamiento informado.

Mejoras adicionales

También se describen nuevos usos para las secuencias genómicas seleccionadas del grupo que consiste en SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961 (preferentemente SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961). Las modalidades adicionales proporcionan variantes modificadas de SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961, así como también oligonucleótidos y/o PNA-oligómeros para análisis de patrones de metilación de citosina dentro del grupo que consiste en SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961 (preferentemente SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961).

Un objetivo comprende el análisis del estado de metilación de uno o más dinucleótidos CpG dentro de al menos una de las secuencias genómicas seleccionadas del grupo que consiste en SEQ ID NO: a SEQ ID NO:64 y SEQ ID NO: 961 y secuencias complementarias a estas. Preferentemente dicho grupo consiste en las SEQ ID Nos: 35, 63, 19 y con la máxima preferencia dicha secuencia es la SEQ ID NO: 961.

También se describen ácidos nucleicos tratados, derivados de la SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961 genómicas (preferentemente las SEQ ID Nos: 35, 63, 19 y con la máxima preferencia la SEQ ID NO: 961), en donde el tratamiento es adecuado para convertir al menos una base de citosina no metilada de la secuencia de ADN genómica a uracilo u otra base que se detecta diferente a la citosina en términos de hibridación. Las secuencias genómicas en cuestión pueden comprender uno o más posiciones CpG metiladas consecutivas o aleatorias. Dicho tratamiento comprende preferentemente el uso de un reactivo seleccionado del grupo que consiste en bisulfito, sulfito de hidrógeno, disulfito, y combinaciones de estos. En una modalidad preferida, el objetivo comprende análisis de un ácido nucleico modificado que no es de origen natural que comprende una secuencia de al menos 16 bases de nucleótidos contiguos de longitud de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID Nos: 962 - 965 (preferentemente dicho grupo consiste en las SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia dicho grupo consiste en las SEQ ID Nos: 962 - 965). Particularmente se prefiere un ácido nucleico modificado que no es de origen natural que comprende una secuencia de al menos 16 bases de nucleótidos contiguos de longitud de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965 que no es idéntica a o complementaria a la SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 u otro ADN genómico humano. Adicionalmente se prefiere un ácido nucleico modificado que no es de origen natural que comprende una secuencia de al menos 16 bases de nucleótidos contiguos de longitud de una secuencia seleccionada del grupo que

consiste en SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230, 962 - 965 que es no idéntica a o complementaria SEQ ID Nos: 961, 35, 63 y 19 u otro ADN genómico humano.

Se prefiere adicionalmente que dicha secuencia comprende al menos un dinucleótido CpG, TpA o CpA y secuencias complementarias a estas. Las secuencias de SEQ ID NO:65 a SEQ ID NO:320 y SEQ ID NO: 962 a SEQ ID NO: 965 proporcionan versiones modificadas que no es de origen natural del ácido nucleico de acuerdo con la SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961, en donde la modificación de cada secuencia genómica resulta en la síntesis de un ácido nucleico que tiene una secuencia que es única y distinta de dicha secuencia genómica como sigue. Para cada ADN genómico de cadena sentido, *por ejemplo*, SEQ ID NO:1, se describen cuatro versiones convertidas. Una primera versión donde "C" se convierte a "T", pero "CpG" sigue siendo "CpG" (*es decir*, corresponde al caso donde, para la secuencia genómica, todos los residuos "C" de secuencias de dinucleótidos CpG se metilan y son así no convertidas); una segunda versión describe el complemento de la secuencia de ADN genómico descrito (*es decir* cadena antisentido), en donde "C" se convierte a "T", pero "CpG" sigue siendo "CpG" (*es decir*, corresponde al caso donde, para todos los residuos "C" de secuencias de dinucleótidos CpG se metilan y así no se convierten). Las secuencias convertida no metiladas de SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 corresponden a la SEQ ID NO:65 a SEQ ID NO:728. Se proporciona una tercera versión de cada una de las secuencias genómicas convertidas químicamente, en donde "C" se convierte a "T" para todos los residuos "C", que incluyen las secuencias de dinucleótidos "CpG" (*es decir*, corresponden al caso donde, para las secuencias genómicas, todos los residuos "C" de secuencias de dinucleótidos CpG son no metiladas); una versión químicamente convertida definitiva de cada secuencia, describe el complemento de la secuencia de ADN genómico descrito (*es decir* cadena antisentido), en donde "C" se convierte a "T" para todos los residuos "C", que incluyen las secuencias de dinucleótidos "CpG" (*es decir*, corresponden al caso donde, para el complemento (*cadena* antisentido) de cada secuencia genómica, todos los residuos de "C" de secuencias de dinucleótidos CpG son no metiladas). Las secuencias convertidas 'poco metiladas' de SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 corresponden a la SEQ ID NO:65 a SEQ:320 y SEQ ID NO: 962 a SEQ ID NO: 965.

En una modalidad alternativa preferida, tal análisis comprende el uso de un oligonucleótido u oligómero para detectar el estado de metilación de citosina dentro del ADN genómico o tratado (modificado químicamente), de acuerdo con SEQ ID NO: 1 a SEQ ID NO: 320 y SEQ ID NO: 961 a SEQ ID NO: 965. Dicho oligonucleótido u oligómero que comprende una secuencia de ácido nucleico que tiene una longitud de al menos nueve (9) nucleótidos que se hibrida, en condiciones moderadamente rigurosas o rigurosas (como se define en la presente descripción anteriormente), a una secuencia de ácido nucleico tratada de acuerdo con la SEQ ID NO:1 a SEQ ID NO:320 y SEQ ID NO: 961 a SEQ ID NO: 965 y/o secuencias complementarias a estas, o a una secuencia genómica de acuerdo con SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961 y/o secuencias complementarias a estas.

Así, la presente descripción incluye moléculas de ácidos nucleicos (*por ejemplo*, oligonucleótidos y moléculas de ácidos nucleicos peptídicos (PNA) (PNA-oligómeros)) que se hibridan en condiciones moderadamente rigurosas y/o rigurosas de hibridación a toda o una porción de las secuencias SEQ ID NO: 1 a SEQ ID NO: 320 y SEQ ID NO: 961 a 965, o a los complementos de estos. Se prefiere particularmente una molécula de ácido nucleico que hibrida en condiciones de hibridación moderadamente rigurosas y/o rigurosas a toda o una porción de las secuencias SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965 pero no es idéntica a o complementaria a SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 u otro ADN genómico humano. Se prefiere adicionalmente una molécula de ácido nucleico que hibrida en condiciones de hibridación moderadamente rigurosas y/o rigurosas a toda o una porción de las secuencias SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230, 962 - 965 pero no es idéntica a o complementaria a las SEQ ID Nos: 961, 35, 63 y 19 u otro ADN genómico humano.

La porción de hibridación de los ácidos nucleicos de hibridación es típicamente de al menos 9, 15, 20, 25, 30 o 35 nucleótidos de longitud. Sin embargo, las moléculas más largas tienen utilidad, y están así dentro del alcance de la presente descripción.

Preferentemente, la porción de hibridación de los ácidos nucleicos de hibridación es al menos 95%, o al menos 98%, o 100% idéntica a la secuencia, o a una porción de esta de la SEQ ID NO: 1 a SEQ ID NO: 320 y SEQ ID NO: 961 a 965, o a los complementos de estas.

Los ácidos nucleicos de hibridación del tipo descrito en la presente se pueden usar, por ejemplo, como un iniciador (*por ejemplo* un iniciador de PCR), o un diagnóstico y/o una sonda o iniciador pronóstico. Preferentemente, la hibridación de la sonda de oligonucleótido a una muestra de ácido nucleico se lleva a cabo en condiciones rigurosas y la sonda es 100% idéntica a la secuencia objetivo. La estabilidad del ácido nucleico dúplex o híbrido se expresa como la temperatura de fusión

o T_m , que es la temperatura a la que una sonda se disocia de un ADN objetivo. Esta temperatura de fusión se usa para definir las condiciones de rigurosidad requeridas.

Para las secuencias objetivos que se relacionan y sustancialmente idénticas a la secuencia correspondiente de la SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 (tales como variantes alélicas y SNPs), en lugar de idénticas, es útil establecer primero la temperatura más baja a la que sólo se produce la hibridación homóloga con una concentración de sal particular (por ejemplo, SSC o SSPE). Después, suponiendo que 1% de desajustes resulta en una disminución de 1 °C en la T_m , la temperatura del lavado final en la reacción de hibridación se reduce como consecuencia (por ejemplo, si se buscan las secuencias que tienen > 95% de identidad con la sonda, la temperatura de lavado final se disminuye por 5 °C). En la práctica, el cambio en la T_m puede ser entre 0.5 °C y 1.5 °C por 1% de desajuste.

Ejemplos de oligonucleótidos de longitud X (en nucleótidos), como se indicó por las posiciones de polinucleótido con referencia a, por ejemplo, SEQ ID NO:1, incluyen los correspondientes a los conjuntos (conjuntos sentido y antisentido) de oligonucleótidos de solapamiento de longitud X consecutivos, donde los oligonucleótidos dentro de cada conjunto de solapamiento consecutivo (correspondiente a un valor X dado) se definen como el conjunto finito de oligonucleótidos Z a partir de las posiciones de nucleótidos:

N para $(n + (X-1))$;

donde $n=1, 2, 3, \dots, (Y-(X-1))$;
donde Y es igual a la longitud (nucleótidos o pares de bases) de la SEQ ID NO: 1 (7572);
donde X es igual a la longitud común (en nucleótidos) de cada oligonucleótido en el conjunto (por ejemplo, $X=20$ para un conjunto de 20-mers de solapamiento consecutivo); y
donde el número (Z) de oligómeros de solapamiento consecutivos de longitud X para una determinada SEQ ID NO de longitud Y es igual a $Y-(X-1)$. Por ejemplo $Z = 7572-19 = 7553$ para cualquiera de los conjuntos sentido o antisentido de SEQ ID NO: 1, donde $X=20$.

Preferentemente, el conjunto se limita a los oligómeros que comprenden al menos un dinucleótido CpG, TpG o CpA.

Ejemplos de oligonucleótidos de 20-mer incluyen el siguiente conjunto de oligómeros (y el conjunto antisentido complementario a este), indicado por las posiciones de polinucleótido con referencia a la SEQ ID NO:1:

1-20, 2-21, 3-22, 4-23, 5-24,7553-7572.

Preferentemente, el conjunto se limita a los oligómeros que comprenden al menos un dinucleótido CpG, TpG o CpA.

Del mismo modo, ejemplos de oligonucleótidos de 25-mer incluyen el siguiente conjunto de 2,256 oligómeros (y el conjunto antisentido complementario a este), indicado por las posiciones de polinucleótido con referencia a la SEQ ID NO:1:

1-25, 2-26, 3-27, 4-28, 5-29, 7553-7572.

Preferentemente, el conjunto se limita a los oligómeros que comprenden al menos un dinucleótido CpG, TpG o CpA.

La presente descripción abarca, para cada una de SEQ ID NO: 1 a SEQ ID NO: 320 y SEQ ID NO: 961 a SEQ ID NO: 965 (sentido y antisentido), múltiples conjuntos de solapamiento consecutivos de oligonucleótidos u oligonucleótidos modificados de longitud X, donde, por ejemplo, $X = 9, 10, 17, 20, 22, 23, 25, 27, 30$ o 35 nucleótidos.

Los oligonucleótidos u oligómeros descritos en la presente constituyen herramientas efectivas útiles para comprobar los parámetros genéticos y epigenéticos de la secuencia genómica correspondiente a la SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961. Conjuntos preferidos de tales oligonucleótidos u oligonucleótidos modificados de longitud X son los conjuntos de solapamiento consecutivos de oligómeros correspondientes a SEQ ID NO: 1 a SEQ ID NO: 320 y SEQ ID NO: 961 a SEQ ID NO: 965 (y a los complementos de estos). Preferentemente, dichos oligómeros comprenden al menos un dinucleótido CpG, TpG o CpA.

Particularmente los oligonucleótidos u oligómeros preferidos son en los que la citosina de las secuencias de dinucleótido CpG (o del correspondiente dinucleótido convertido TpG o CpA) está dentro del tercio medio del oligonucleótido; es decir, donde el oligonucleótido es, por ejemplo, 13 bases de longitud, el dinucleótido CpG, TpG o CpA se coloca dentro del quinto al noveno nucleótido desde el extremo 5'.

Los oligonucleótidos descritos en la presente también se pueden modificar por el enlace químico del oligonucleótido a una o más porciones o conjugados para mejorar la actividad, estabilidad o detección del oligonucleótido. Tales porciones o conjugados incluyen cromóforos, fluoróforos, lípidos tales como colesterol, ácido cólico, tioéter, cadenas alifáticas, fosfolípidos, poliaminas, polietilenglicol (PEG), porciones de palmitilo, y otros como se escribe en, por ejemplo, las Patentes de Estados Unidos Números 5,514,758, 5,565,552, 5,567,810, 5,574,142, 5,585,481, 5,587,371, 5,597,696 y 5,958,773. Las sondas pueden existir también en forma de un PNA (ácido nucleico peptídico) que particularmente ha preferido las propiedades de apareamiento. Así, el oligonucleótido puede incluir otros grupos adjuntos tales como péptidos, y puede incluir agentes de escisión desencadenantes de la hibridación (Krol y otros, BioTechniques 6:958-976, 1988) o agentes intercalantes (Zon, Pharm. Res. 5:539-549, 1988). Para este fin, el oligonucleótido se puede conjugar a otra molécula, *por ejemplo*, un cromóforo, fluoróforo, péptido, agente de entrecruzamiento desencadenante de la hibridación, agente de transporte, agente de escisión desencadenante de la hibridación etc.

El oligonucleótido puede comprender también al menos un azúcar modificada reconocida en la técnica y/o porción de base, o puede comprender una cadena principal modificada o un enlace internucleosídico no natural.

Los oligonucleótidos u oligómeros de acuerdo con modalidades particulares de la presente descripción se usan típicamente en "conjuntos", que contienen al menos un oligómero para el análisis de al menos uno de los dinucleótidos CpG de secuencias genómicas SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 y las secuencias complementarias a estas, o al correspondiente dinucleótido CpG, TpG o CPA dentro de una secuencia de los ácidos nucleicos tratados de acuerdo con SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965 y secuencias complementarias a estas. Sin embargo, se prevé que para los factores económicos u otros puede ser preferible analizar una selección limitada de los dinucleótidos CpG dentro de dichas secuencias, y el contenido del conjunto de oligonucleótidos se altera como consecuencia.

Por lo tanto, se proporciona en modalidades particulares, un conjunto de al menos dos (2) (oligonucleótidos y/o oligómeros-PNA) útiles para detectar el estado de metilación de citosina en el ADN genómico tratado (SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965), o en el ADN genómico (SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 y secuencias complementarias a estas). Estas sondas permiten el diagnóstico y/o la clasificación de los parámetros genéticos y epigenéticos de los trastornos proliferativos de células de la próstata. El conjunto de oligómeros pueden usarse también para detectar polimorfismos de nucleótido único (SNPs) en el ADN genómico tratado (SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965), o en el ADN genómico (SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 y las secuencias complementarias a estas).

En modalidades preferidas, al menos uno, y con mayor preferencia todos los miembros de un conjunto de oligonucleótidos se unen a una fase sólida.

En modalidades adicionales, se proporciona un conjunto de al menos dos (2) oligonucleótidos que se usan como oligonucleótidos 'iniciadores' para amplificar secuencias de ADN de una de SEQ ID NO: 1 a SEQ ID NO: 320 y SEQ ID NO: 961 a SEQ ID NO: 965 y las secuencias complementarias a estas, o segmentos de estos.

Se prevé que los oligonucleótidos pueden constituir toda o parte de una "matriz" o "matriz de ADN" (*es decir*, un arreglo de diferentes oligonucleótidos y/o PNA-oligómeros unidos a una fase sólida). Se puede caracterizar por ejemplo, una matriz de ese tipo de diferentes secuencias de oligonucleótido- y/o PNA-oligómero en que se arregla sobre la fase sólida en forma de una red rectangular o hexagonal. La superficie de fase sólida puede estar compuesta de silicio, vidrio, poliestireno, aluminio, acero, hierro, cobre, níquel, plata, u oro. La nitrocelulosa así como plásticos tales como nailon, que pueden existir en forma de gránulos o también como matrices de resina, también se pueden usar. Una visión general de la técnica anterior en la fabricación de matriz de oligómero puede ser obtenida de una edición especial de Nature Genetics (Nature Genetics Supplement, volumen 21, enero 1999, y de la bibliografía citada en ella). Sondas marcadas con fluorescencia se usan frecuentemente para el escaneo de matrices de ADN inmovilizadas. La fijación simple de colorantes Cy3 y Cy5 al 5'-OH de la sonda específica es particularmente adecuada para los marcadores de fluorescencia. La detección de la fluorescencia de las sondas hibridadas se puede llevar cabo, por ejemplo, a través de un microscopio confocal. Los colorantes Cy3 y Cy5, además de muchos otros, están disponibles en el comercio.

También se prevé que los oligonucleótidos o secuencias particulares de estas, pueden constituir toda o parte de una "matriz virtual" en donde se usan los oligonucleótidos o secuencias particulares de estas, por ejemplo, como "especificadores" como parte de, o en combinación con una población diversa de sondas marcadas únicas para analizar una mezcla compleja de analitos. Tal método, por ejemplo se describe en US 2003/0013091 (Estados Unidos serie número 09/898,743, publicada el

16 de enero de 2003). En tales métodos, se generan suficientes marcadores de manera que cada ácido nucleico en la mezcla compleja (*es decir*, cada analito) se puede unir de forma única por un marcador único y así detecta (cada marcador se cuenta directamente, resultando en una lectura digital de cada especie molecular en la mezcla).

5 Se prefiere particularmente que los oligómeros descritos en la presente se utilizan para al menos uno de: pronóstico de; tratamiento de; seguimiento de; y tratamiento y seguimiento de los trastornos proliferativos de células de la próstata. Esto es posible mediante el uso de dichos conjuntos para proporcionar un pronóstico de una muestra biológica aislada de un paciente. Particularmente preferidos son los conjuntos de oligómero que comprenden al menos dos oligonucleótidos seleccionados a partir de uno de los siguientes conjuntos de oligonucleótidos.

10 En una modalidad del método, esto se logra mediante el análisis de la condición de metilación de al menos una secuencia objetivo que comprende una secuencia, o que hibrida en condiciones rigurosas con al menos 16 nucleótidos contiguos de un gen o secuencia seleccionada del grupo que consiste en los genes y las secuencias de acuerdo con la Tabla 11 y complementos de estos. Se prefiere particularmente que dicho gen o secuencia genómica se selecciona del grupo que
15 consiste en PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35 y sus secuencias se listan en la lista de secuencias de acompañamiento. Adicionalmente preferido es el gen PITX2.

También se describe un método para la determinación de parámetros genéticos y/o epigenéticos de las secuencias genómicas de acuerdo con SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 dentro de un sujeto mediante el análisis de
20 metilación de citosina y polimorfismos de nucleótido único. También se describe un método para la determinación de parámetros genéticos y/o epigenéticos de las secuencias genómicas de acuerdo con SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961 dentro de un sujeto mediante el análisis de metilación de citosina y polimorfismos de nucleótido único. Dicho método que comprende poner en contacto un ácido nucleico que comprende una o más de SEQ ID
25 NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 (preferentemente una o más de SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961) en una muestra biológica obtenida de dicho sujeto con al menos un reactivo o una serie de reactivos, en donde dicho reactivo o serie de reactivos, distingue entre dinucleótidos CpG metilados y no metilados dentro del ácido nucleico objetivo.

30 Preferentemente, dicho método comprende las siguientes etapas: En la *primera etapa*, se obtiene una muestra de tejido que se analiza. La fuente puede ser cualquier fuente adecuada. Preferentemente, se selecciona la fuente de la muestra de ADN del grupo que consiste de células o líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, eyaculado, orina, sangre, y combinaciones de estos. Preferentemente, la fuente es biopsias, fluidos corporales, eyaculado, orina o sangre.

35 El ADN genómico se aísla después de la muestra. El ADN genómico se puede aislar por cualquier medio estándar en la técnica, que incluyen el uso de estuches disponibles en el comercio. En resumen, en donde el ADN de interés se encapsula mediante una membrana celular la muestra biológica se debe romper y lisar mediante medios enzimáticos, químicos o mecánicos. La solución de ADN se puede clarificar de las proteínas y otros contaminantes, por ejemplo mediante digestión con proteinasa K. El ADN genómico se recupera después de la solución. Esto puede llevarse a cabo por medio de una
40 variedad de métodos que incluyen precipitación salina, extracción orgánica o unión de ADN a un soporte de fase sólida. La elección del método será afectada por varios factores, que incluyen el tiempo, costo y cantidad requerida de ADN.

Una vez que los ácidos nucleicos se han extraído, se usa en el análisis el ADN genómico de cadena doble.

45 En la *segunda etapa* del método la muestra de ADN genómico se trata después de manera tal que las bases de citosina que están sin metilar en la posición 5' se convierten a uracilo, timina, u otra base que es muy diferente a la citosina en términos de comportamiento de hibridación. Esto se entenderá como "pretratamiento" o "tratamiento" en la presente descripción.

50 El tratamiento descrito anteriormente del ADN genómico se lleva a cabo preferentemente con bisulfito (sulfito de hidrógeno, disulfito) y posterior hidrólisis alcalina que resulta en una conversión de nucleobases de citosina no metiladas a uracilo o a otra base que es diferente a la citosina en términos de comportamiento de apareamiento de bases.

En la *tercera etapa* del método, se amplifican fragmentos de ADN tratado, usando conjuntos de oligonucleótidos iniciadores descritos en la presente, y una enzima de amplificación. La amplificación de varios segmentos de ADN puede llevarse a
55 cabo simultáneamente en uno y el mismo recipiente de reacción. Típicamente, la amplificación se lleva a cabo usando una reacción en cadena de la polimerasa (PCR). El conjunto de oligonucleótidos iniciadores incluye al menos dos oligonucleótidos cuyas secuencias son cada una la complementaria inversa, idéntica, o se hibridan en condiciones rigurosas o altamente rigurosas a un segmento largo de al menos 16 pares de bases de las secuencias de bases de uno de SEQ ID

NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965 (preferentemente una de las SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia una de las SEQ ID Nos: 962-965) y las secuencias complementarias a estas.

- 5 En una modalidad alternativa del método, la condición de metilación de las posiciones CpG preseleccionadas dentro de las secuencias de ácido nucleico que comprende una o más de SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 (preferentemente una o más de SEQ ID Nos: 35, 63, 19 y lo más preferentemente SEQ ID NO: 961) se pueden detectar mediante el uso de oligonucleótidos iniciadores de metilación específica. Esta técnica (MSP) se ha descrito en la Patente de Estados Unidos con núm. 6,265,171 a Herman. El uso de iniciadores específicos de la condición de metilación para la
- 10 amplificación de ADN tratado con bisulfito permite la diferenciación entre ácidos nucleicos metilados y no metilados. Los pares de iniciadores MSP contienen al menos un iniciador que se hibrida a un dinucleótido CpG tratado con bisulfito. Por lo tanto, la secuencia de dichos iniciadores comprende al menos un dinucleótido CpG. Los iniciadores MSP específicos para el ADN no metilado contiene una "T" en la posición de la posición C en el CpG. Preferentemente, por lo tanto, se requiere que la secuencia base de dichos iniciadores comprenda una secuencia que tiene una longitud de al menos 9 nucleótidos que
- 15 hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de las SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 961 a 965 (preferentemente SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia SEQ ID Nos: 962-965) y las secuencias complementarias a estas, en donde la secuencia base de dichos oligómeros comprende al menos un dinucleótido CpG.
- 20 Una modalidad adicional preferida del método comprende el uso de oligonucleótidos *bloqueadores*. El uso de tales oligonucleótidos bloqueadores se ha descrito por Yu y otros, BioTechniques 23:714-720, 1997. Los oligonucleótidos de la sonda de bloqueo se hibridan con el ácido nucleico tratado con bisulfito concurrentemente con los iniciadores de PCR. La amplificación por PCR del ácido nucleico se termina en la posición 5' de la sonda de bloqueo, tal que la amplificación de un ácido nucleico se suprime donde está presente la secuencia complementaria a la sonda de bloqueo. Las sondas se pueden
- 25 diseñar para hibridar con el ácido nucleico tratado con bisulfito de manera específica a la condición de metilación. Por ejemplo, para la detección de ácidos nucleicos metilados dentro de una población de ácidos nucleicos no metilados, la supresión de la amplificación de ácidos nucleicos que no son metilados en la posición en cuestión se lleva a cabo mediante el uso de sondas de bloqueo que comprenden un 'CpA' o 'TpA' en la posición en cuestión, a diferencia de un 'CpG' si se desea la supresión de la amplificación de ácidos nucleicos metilados.
- 30 Para los métodos de PCR usando los oligonucleótidos bloqueadores, la ruptura eficaz de la amplificación mediada por polimerasa requiere que los oligonucleótidos bloqueadores no se alarguen por la polimerasa. Preferentemente, esto se logra a través del uso de bloqueadores que son 3'-desoxioligonucleótidos, u oligonucleótidos derivatizados en la posición 3' que no sean un grupo "libre" hidroxilo. Por ejemplo, los oligonucleótidos 3'-O-acetilo son representativos de una clase preferida de molécula bloqueadora.
- 35 Además, se debe impedir la descomposición mediada por polimerasa de los oligonucleótidos bloqueadores. Preferentemente, tal impedimento comprende ya sea el uso de una polimerasa que carece de actividad 5'-3' exonucleasa, o el uso de oligonucleótidos bloqueadores modificados que tienen, por ejemplo, puentes tioato en el extremo 5' de estos que hacen resistente a nucleasa la molécula bloqueadora. Las aplicaciones particulares pueden no requerir tales modificaciones en 5' del bloqueador. Por ejemplo, si los sitios de unión al bloqueador e iniciador se solapan, impidiendo de ese modo la unión del iniciador (*por ejemplo*, con exceso de bloqueador), se impedirá sustancialmente la degradación del oligonucleótido bloqueador. Esto es porque la polimerasa no extenderá el iniciador hacia, y a través del (en la dirección 5'-3') bloqueador un proceso que resulta normalmente en la degradación del oligonucleótido bloqueador hibridado.
- 40
- 45 Una modalidad de bloqueador/PCR particularmente preferida comprende el uso de oligómeros de ácido nucleico peptídico (PNA) como oligonucleótidos de bloqueo. Tales oligómeros bloqueadores de PNA son idealmente adecuados, porque no son ni descompuestos ni extendidos por la polimerasa.
- 50 Preferentemente, por lo tanto, se requiere que la secuencia base de dichos *oligonucleótidos de bloqueo* comprenda una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de las SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965 y las secuencias complementarias a estas, en donde la secuencia base de dichos oligómeros comprende al menos un dinucleótido CpG, TpG o CpA. Con mayor preferencia, se requiere que la secuencia base de dichos *oligonucleótidos de bloqueo* comprenda una
- 55 secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de preferentemente SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia SEQ ID Nos: 962-965 y las secuencias complementarias a estas, en donde la secuencia base de dichos oligómeros comprende al menos un dinucleótido CpG, TpG o CpA.

Los fragmentos obtenidos por medio de la amplificación pueden llevar un marcador directo o indirectamente detectable. Se prefieren los marcadores en la forma de marcadores de fluorescencia, radionucleidos o fragmentos de moléculas separables que tienen una masa típica que se puede detectar en un espectrómetro de masas. Donde dichos marcadores son marcadores de masa, se prefiere que los amplificados marcados tengan una sola carga neta positiva o negativa, permitiendo una mejor detectabilidad en el espectrómetro de masas. La detección puede llevarse a cabo y visualizarse por medio de, *por ejemplo*, espectrometría de masas de desorción/ionización láser asistida por matriz (MALDI) o usando la espectrometría de masas con electrospray (ESI).

La Espectrometría de Masas de Desorción/ionización Láser Asistida por Matriz (MALDI-TOF) es un desarrollo muy eficaz para el análisis de biomoléculas (Karas & Hillenkamp, Anal Chem., 60:2299-301, 1988). Un analito se embebe en una matriz absorbente a la luz. La matriz se evapora por un pulso corto de láser que transporta así la molécula de analito en la fase de vapor de manera no fragmentada. El analito se ioniza por colisiones con moléculas de la matriz. Un voltaje aplicado acelera los iones en un tubo de vuelo en campo libre. Debido a sus diferentes masas, los iones se aceleran a diferentes velocidades. Los iones más pequeños alcanzan el detector más pronto que los más grandes. La espectrometría MALDI-TOF es muy adecuada para el análisis de péptidos y proteínas. El análisis de ácidos nucleicos es algo más difícil (Gut & Beck, Current Innovations and Future Trends, 1:147-57, 1995). La sensibilidad con respecto a los análisis de ácidos nucleicos es aproximadamente 100 veces menor que el de los péptidos, y disminuye de manera desproporcionada con el aumento de tamaño del fragmento. Además, para los ácidos nucleicos que tienen una cadena principal múltiple cargada negativamente, el proceso de ionización a través de la matriz es considerablemente menos eficaz. En la espectrometría MALDI-TOF, la selección de la matriz desempeña un papel eminentemente importante. Para la desorción de los péptidos, varias matrices muy eficaces han sido encontradas que producen una cristalización muy fina. Existen ahora varias matrices de respuesta para el ADN, sin embargo, no se ha reducido la diferencia de sensibilidad entre los péptidos y ácidos nucleicos. Esta diferencia en sensibilidad se puede reducir, sin embargo, modificando químicamente el ADN de manera tal que se vuelva más similar a un péptido. Por ejemplo, los ácidos nucleicos fosforotioato, en los que los fosfatos habituales de la cadena principal se sustituyen con tiofosfatos, se pueden convertir en un ADN de carga neutra usando la química de alquilación sencilla (Gut & Beck, Nucleic Acids Res. 23: 1367-73, 1995). El acoplamiento de una etiqueta de carga a este ADN modificado resulta en un aumento de la sensibilidad de MALDI-TOF al mismo nivel que el encontrado para los péptidos. Una ventaja adicional del etiquetado de la carga es el aumento la estabilidad del análisis contra las impurezas, que hace considerablemente más difícil la detección de sustratos no modificados.

En la *cuarta etapa* del método, los amplificados obtenidos durante la tercera etapa del método se analizan para determinar la condición de metilación de los dinucleótidos CpG antes del tratamiento.

En modalidades donde los amplificados se obtuvieron por medio de la amplificación MSP, la presencia o ausencia de un amplificado es en sí mismo indicativo del estado de metilación de las posiciones de CpG cubiertas por el iniciador, de acuerdo con las secuencias de bases de dicho iniciador.

Los amplificados obtenidos por medio tanto de PCR específico a metilación como estándar se pueden analizar adicionalmente por medio de métodos basados en la hibridación, tales como, pero sin limitarse a las tecnologías de matriz como tecnologías basadas, en sonda así como por medio de técnicas tales como secuenciación y extensión dirigida por molde.

En una modalidad del método, los amplificados sintetizados en la *etapa tres* posteriormente se hibridan con una matriz o un conjunto de oligonucleótidos y/o sondas de PNA. En este contexto, la hibridación tiene lugar en de la siguiente manera: el conjunto de sondas usadas durante la hibridación se compone preferentemente de al menos 2 oligonucleótidos u oligómeros PNA; en el proceso, los amplificados sirven como sondas que se hibridan con oligonucleótidos previamente unidos a una fase sólida; los fragmentos no hibridados se eliminan posteriormente; dichos oligonucleótidos contienen al menos una secuencia base que tiene una longitud de al menos 9 nucleótidos que es complementaria inversa o idéntica a un segmento de las secuencias bases especificadas en la presente Lista de Secuencias; y el segmento comprende al menos un dinucleótido CpG, TpG o CpA.

En una modalidad preferida, dicho dinucleótido está presente en el tercio central del oligómero. Por ejemplo, en donde el oligómero comprende un dinucleótido CpG, dicho dinucleótido es preferentemente el quinto a noveno nucleótido del extremo 5' de un 13- mer. Existe un oligonucleótido para el análisis de cada dinucleótido CpG dentro de la secuencia de acuerdo con la SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961, y las posiciones equivalentes dentro de la SEQ ID NO:65 a SEQ ID NO:320 y SEQ ID NO: 962 a SEQ ID NO: 965. Dichos oligonucleótidos también pueden estar presentes en forma de ácidos nucleicos peptídicos. Se eliminan después los amplificados no hibridados. Se detectan después los amplificados hibridados.

En este contexto, se prefiere que los marcadores unidos a los amplificados sean identificables en cada posición de la fase sólida en la que se encuentra una secuencia de oligonucleótidos.

En aún una modalidad adicional del método, la condición de metilación genómica de las posiciones de CpG se pueden determinar por medio de sondas de oligonucleótidos que se hibridan con el ADN tratado con bisulfito concurrentemente con los iniciadores de amplificación de PCR (en donde dichos iniciadores pueden ser ya sea de metilación específica o estándar).

Una modalidad particularmente preferida de este método es el uso de PCR Cuantitativo en Tiempo Real basado en la fluorescencia (Heid y otros, Genome Res. 6:986-994, 1996; *ver también* la Patente de Estados Unidos núm. 6,331,393) que emplea una sonda de oligonucleótido fluorescente marcada dual (TaqMan™ PCR, usando un Sistema de Detección de Secuencia ABI Prism 7700, Perkin Elmer Applied Biosystems, Foster City, California). La reacción de PCR TaqMan™ emplea el uso de un oligonucleótido de interrogación no extendible, llamado sonda TaqMan™, que, en modalidades preferidas, se diseña para hibridar con una secuencia rica en GPC situado entre los iniciadores de amplificación directo y reverso. La sonda TaqMan™ además comprende una "porción reportera" fluorescente y una "porción inactivadora" unida covalentemente a las porciones enlazadoras (*por ejemplo*, fosforamiditas) unidas a los nucleótidos del oligonucleótido TaqMan™. Para el análisis de metilación dentro de los ácidos nucleicos posterior al tratamiento bisulfito, se requiere que la sonda sea metilación específica, como se describe en la Patente de Estados Unidos núm. 6,331,393, conocido también como ensayo MethyLight™. Variaciones sobre la metodología de detección TaqMan™ que también son adecuadas para el uso incluyen el uso de la tecnología de doble sonda (Lightcycler™) o iniciadores de amplificación fluorescente (tecnología Sunrise™). Ambas técnicas se pueden adaptar de manera adecuada para el uso con ADN tratado con bisulfito, y además para el análisis de metilación dentro de los dinucleótidos CpG.

Un método adicional adecuado para el uso de oligonucleótidos de sonda para la evaluación de metilación mediante análisis de ácidos nucleicos tratados con bisulfito En una modalidad adicional preferida del método, la *quinta etapa* del método comprende el uso de la extensión del oligonucleótido dirigido al molde, tal como MS-SNuPE como se describe por Gonzalgo & Jones, Nucleic Acids Res. 25:2529-2531, 1997.

En aún una modalidad adicional del método, la *cuarta etapa* del método comprende la secuenciación y análisis posterior de secuencia del amplificado generado en la *tercera etapa* del método (Sanger F., y otros, Proc Natl Acad Sci USA 74:5463-5467, 1977).

Mejor modo

En la modalidad más preferida el método de los ácidos nucleicos genómicos se aíslan y se tratan de acuerdo con las tres primeras etapas del método descrito anteriormente, a saber:

- a) obtener, de un sujeto, una muestra biológica que tiene el ADN genómico del sujeto;
- b) extraer o de cualquier otra forma aislar el ADN genómico;
- c) tratar el ADN genómico de b), o un fragmento de este, con uno o más reactivos para convertir las bases de citosina que no son metiladas en la posición 5 de esta a uracilo o a otra base que se detecta diferente a la citosina en términos de propiedades de hibridación; y en donde
- d) amplificar después que el tratamiento en c) se lleva a cabo en una manera específica de metilación, específicamente mediante el uso de iniciadores específicos de metilación u *oligonucleótidos de bloqueo*, y adicionalmente en donde
- e) detectar los amplificados llevados a cabo por medio de una sonda en tiempo real, como se describe anteriormente.

Preferentemente, donde la amplificación posterior de d) se lleva a cabo por medio de iniciadores específicos de metilación, como se describió anteriormente, dichos iniciadores específicos de metilación comprenden una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de las SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 962 a SEQ ID NO: 965 y secuencias complementarias a estas, en donde la secuencia de dichos oligómeros comprende al menos un dinucleótido CpG. Con mayor preferencia, donde la amplificación posterior de d) se lleva a cabo por medio de iniciadores específicos de metilación, como se describió anteriormente, dichos iniciadores específicos de metilación comprenden una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia SEQ ID Nos: 962 - 965 y secuencias complementarias a estas, en donde la secuencia base de dichos oligómeros comprende al menos un dinucleótido CpG.

En una alternativa la modalidad más preferida del método, la amplificación posterior de d) se lleva a cabo en presencia de

oligonucleótidos de bloqueo, como se describió anteriormente. Dichos *oligonucleótidos de bloqueo*, que comprenden una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de las SEQ ID NO: 65 a SEQ ID NO: 320 y SEQ ID NO: 961 a 965 y las secuencias complementarias a estas, en donde la secuencia base de dichos oligómeros comprende al menos un dinucleótido CpG, TpG o CpA . Preferentemente dichos *oligonucleótidos de bloqueo* que comprenden una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia SEQ ID Nos.: 962-965 y secuencias complementarias a estas, en donde la secuencia base de dichos oligómeros comprende al menos un dinucleótido CpG, TpG o CpA .

Etapa e) del método, es decir, la detección de los amplificados específicos indicativos de la condición de metilación de una o más posiciones de CpG de acuerdo con SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961 (preferentemente SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961) se lleva a cabo por medio de métodos de detección en tiempo real como se describió anteriormente.

También se describe un método para el análisis de la condición de metilación del ADN genómico descrito en la presente SEQ ID NO: 1 a SEQ ID NO: 64 y SEQ ID NO: 961, (preferentemente SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO.: 961) y complementos de estas sin necesidad de pretratamiento.

En la *primera etapa* de tales modalidades adicionales, la muestra de ADN genómico se aísla de tejido o de fuentes celulares. Preferentemente, tales fuentes incluyen líneas celulares, cortes histológicos, fluidos corporales, o tejidos embebidos en parafina. En la *segunda etapa*, se extrae el ADN genómico. La extracción puede ser por medios que son estándar para un experto en la técnica, que incluyen pero sin limitarse a el uso de lisados de detergente, sonicación y agitación con vórtex con perlas de vidrio. Una vez que los ácidos nucleicos se han extraído, el ADN genómico de cadena doble se usa en el análisis.

En una modalidad preferida, el ADN se puede escindir antes del tratamiento, y esto puede ser mediante cualquier medio estándar en el estado de la técnica, particularmente, con endonucleasas de restricción sensibles a la metilación.

En la *tercera etapa*, el ADN se digiere después con una o más enzimas de restricción sensibles a la metilación. La digestión se lleva a cabo tal que la hidrólisis del ADN en el sitio de restricción es informativo de la condición de metilación de un dinucleótido CpG específico.

En la *cuarta etapa*, que es opcional, pero una modalidad preferida, se amplifican los fragmentos de restricción. Esto se lleva a cabo preferentemente usando una reacción en cadena de la polimerasa, y dichos amplificados pueden llevar marcadores detectables adecuados como se discutió anteriormente, es decir marcadores fluoróforo, radionucleidos y marcadores de masa.

En la *quinta etapa* se detectan los amplificados. La detección puede ser por cualquier medio estándar en la técnica, por ejemplo, pero sin limitarse a, análisis de electroforesis en gel, análisis de hibridación, incorporación de etiquetas detectables dentro de los productos de PCR, análisis de matriz de ADN, análisis MALDI o ESI.

Posterior a la determinación del estado de metilación de los ácidos nucleicos genómicos el pronóstico del cáncer de próstata se deduce basado en el estado de metilación de al menos una secuencia de dinucleótido CpG de la SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961, o un promedio, o un valor que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG de la SEQ ID NO:1 a SEQ ID NO:64 y SEQ ID NO: 961. Preferentemente dicho pronóstico se basa en el estado de metilación de al menos una secuencia de dinucleótido CpG de los genes PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35 (SEQ ID NOS: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961), o un promedio, o un valor que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG de SEQ ID Nos: 35, 63, 19 y con la máxima preferencia SEQ ID NO: 961. La hipometilación de dichas posiciones de CpG se asocia con buen pronóstico, y la hipermetilación se asocia con mal pronóstico. La hipermetilación se asocia generalmente con baja expresión de ARNm y como consecuencia polipéptidos. El punto de corte para determinar hipo e hiper metilación puede ser la mediana del nivel de metilación para una población dada, o es preferentemente un nivel de corte optimizado. Para el análisis de la SEQ ID NO: 35 se prefiere que la línea de corte sea de entre 30% y 40% de metilación, y con la máxima preferencia 36.42%. Para el análisis de la SEQ ID NO: 63 se prefiere que la línea de corte sea de entre 2% y 10% de metilación, y con la máxima preferencia 5.96%. Para el análisis de GPR7 se prefiere que la línea de corte sea de entre 20% y 10% de metilación, y con la máxima preferencia 16.65%. Para el análisis de PITX2 se prefiere que la línea de corte sea de entre 20% y 10% de metilación, y con la máxima preferencia 14.27%.

En donde los métodos de análisis de la expresión (con la máxima preferencia por medio de análisis de metilación), de los marcadores descritos en la presente (preferentemente PITX2, SEQ ID NO: 63, GPR7 y SEQ ID NO: 35) se usan para determinar el pronóstico de un cáncer de próstata dichos métodos se usan preferentemente en combinación con otras variables de pronóstico clínicos usados para determinar el pronóstico con la máxima preferencia la puntuación Gleason, pero que incluyen también la puntuación nomograma y el nivel de PSA (es decir, que dichas variables se tienen en cuenta o toman en cuenta). En una modalidad preferida, el análisis de la expresión pronóstica de cada uno de los marcadores PITX2 y SEQ ID NO: 35 como se describe en la presente se lleva a cabo en pacientes quienes se presentan con cáncer de próstata confinado al órgano (T2). PITX2 y SEQ ID NO: 35 tienen una clara ventaja al determinar el pronóstico del cáncer de próstata T2, mientras que es la práctica clínica actual que todos los pacientes que se presentan con cáncer de próstata T2 se consideran que tienen un buen pronóstico. De acuerdo con la presente invención dichos pacientes T2 con mal pronóstico (hipermetilación) podrían tener un pronóstico más típico que los pacientes que se presentan con cáncer de próstata T3 (no confinado al órgano) y se pueden tratar adecuadamente. Además el análisis de la expresión pronóstica de cada uno de los marcadores PITX2 y SEQ ID NO: 63 tiene utilidad adicional en el análisis de los pacientes que presentan alta puntuación de Gleason (8 o superior). Dichos pacientes se consideran actualmente que tienen un mal pronóstico, sin embargo por medio del método de análisis de la expresión de PITX2 y SEQ ID NO: 63 descrito en la presente por primera vez es posible diferenciar los pacientes con alta puntuación Gleason con un mal pronóstico (hipermetilación) de los que tienen un buen pronóstico. En la actualidad todos los pacientes con alta puntuación Gleason se consideran candidatos para el tratamiento adyuvante post-quirúrgico, como consecuencia el método descrito en la presente permite la prevención del sobretratamiento de dichos pacientes. Además el análisis de la expresión pronóstica del marcador SEQ ID NO: 63 tiene utilidad adicional en el análisis de los pacientes con mal pronóstico basado en la puntuación nomograma

Estuches

Además, se describe un estuche, dicho estuche comprende, por ejemplo: un reactivo que contiene bisulfito; un conjunto de oligonucleótidos iniciadores que contienen al menos dos oligonucleótidos cuyas secuencias en cada caso corresponden, son complementarias, o hibridan en condiciones rigurosas o muy rigurosas a un segmento de 16 bases de longitud de las secuencias SEQ ID NO: 1 a SEQ ID NO: 320 y SEQ ID NO: 961 a 965; oligonucleótidos y/o oligómeros de PNA; así como también instrucciones para llevar a cabo y evaluar el método descrito. En una modalidad preferida adicional, dicho estuche puede comprender además reactivos estándar para llevar a cabo un análisis de metilación específica de posición CpG, en donde dicho análisis comprende una o más de las siguientes técnicas: MS-SNuPE, MSP, MethyLight™, HeavyMethyl™, COBRA, y secuenciación de ácido nucleico. Sin embargo, un estuche a lo largo de las líneas de la presente descripción puede contener además sólo parte de los componentes antes mencionados.

Preferentemente dicho estuche comprende un reactivo que contiene bisulfito; un conjunto de oligonucleótidos iniciadores que contienen al menos dos oligonucleótidos cuyas secuencias en cada caso corresponden, son complementarias, o hibridan en condiciones rigurosas o muy rigurosas a un segmento de 16 bases de longitud de las secuencias SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230 y con la máxima preferencia las SEQ ID Nos: 962 - 965; oligonucleótidos y/o oligómeros de PNA ; así como también instrucciones para llevar a cabo y evaluar el método descrito. En una modalidad preferida adicional, dicho estuche puede comprender además reactivos estándar para llevar a cabo un análisis de metilación específica de posición CpG, en donde dicho análisis comprende una o más de las siguientes técnicas: MS-SNuPE, MSP, MethyLight™, HeavyMethyl™, COBRA, y secuenciación de ácido nucleico. Sin embargo, un estuche a lo largo de las líneas de la presente descripción puede contener además sólo parte de los componentes antes mencionados.

Se escribe además una composición de materia útil para proporcionar un pronóstico de pacientes con cáncer de próstata. Dicha composición comprende al menos un ácido nucleico de 18 pares de base de longitud de un segmento de un secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230, 962 - 965, y una o más sustancias que tomadas del grupo que comprende : cloruro magnésico, dNTP, polimerasa taq , albúmina de suero bovino, un oligómero en particular un oligonucleótido u oligómero-ácido nucleico péptido (PNA), dicho oligómero comprende en cada caso al menos una secuencia base que tiene una longitud de al menos 9 nucleótidos que es complementaria a, o hibrida bajo condiciones rigurosas o moderadamente rigurosas a un ADN genómico pretratado de acuerdo con una de las SEQ ID Nos: 133,134,261,262, 189,190,317,318, 101,102,229,230, 962 - 965 y las secuencias complementarias a esta. Se prefiere que dicha composición de materia comprenda una solución amortiguadora adecuada para la estabilización de dicho ácido nucleico en una solución acuosa y permitir reacciones basadas en la polimerasa dentro de dicha solución. Los amortiguadores adecuados se conocen en la técnica y se encuentran comercialmente disponibles.

Se describe además un método para proporcionar un diagnóstico de neoplasia o carcinoma de próstata en un sujeto. Dicho método comprende las siguientes etapas:

- i) determinar los niveles de expresión de uno o más genes o secuencias de genes de o de acuerdo con CDRN2A, ELK1, GSTP1, RARB, PTGS2, RASSF1, ESR2, ONECUT2, BTG4, SLC35F2, HOXB5, LIMK1, HIST1H4J, SEQ ID NO: 35, EPAS1, NOTCH1, SEQ ID NO: 55, PTPRN2, Q9NP73, MX1, DOCK10, CCND2, ISL1, SNAPC2, GRN, H2AFY2, WDFY3, FOS, FAT, Q86SP6, SLC38A1, SNRPN, GPRK5, FBN2, ARHGEF18, RHOC, KBTBD6, NR2E1, PSD, DRG1, Q8N365, SEQ ID NO: 44, Q96S01, CD37, CMYA3, SEQ ID NO: 61, Q8NCX8 y ZNF566 y/o regiones reguladoras de estos; y
- ii) determinar la presencia o ausencia de dicha neoplasia o carcinoma de próstata de acuerdo con dicho nivel de expresión.

Dicho nivel de expresión puede determinarse por cualquier medio estándar en la técnica que incluyen pero sin limitarse a análisis de metilación, pérdida de heterocigosidad (de aquí en adelante mencionado además como LOH), niveles de expresión de ARN y niveles de expresión de proteína.

En consecuencia dicho método se puede activar por medio de cualquier análisis de la expresión de un ARN transcrito de la misma o polipéptido o proteína traducida a partir de dicho ARN, preferentemente por medio del análisis de expresión ARNm o análisis de expresión de polipéptido.

En consecuencia, los ensayos y métodos de pronóstico, tanto cuantitativo y cualitativo para detectar la expresión de los genes, secuencia genómicas y/o regiones reguladoras de acuerdo con la Tabla 11 aplicados a CDRN2A, ELK1, GSTP1, RARB, PTGS2, RASSF1, ESR2, ONECUT2, BTG4, SLC35F2, HOXB5, LIMK1, HIST1H4J, SEQ ID NO: 35, EPAS1, NOTCH1, SEQ ID NO: 55, PTPRN2, Q9NP73, MX1, DOCK10, CCND2, ISL1, SNAPC2, GRN, H2AFY2, WDFY3, FOS, FAT, Q86SP6, SLC38A1, SNRPN, GPRK5, FBN2, ARHGEF18, RHOC, KBTBD6, NR2E1, PSD, DRG1, Q8N365, SEQ ID NO: 44, Q96S01, CD37, CMYA3, SEQ ID NO: 61, Q8NCX8 y ZNF566 son adecuados para dichos propósitos diagnósticos. Además, las composiciones de la materia, kits y ácidos nucleicos como se describen anteriormente para el análisis de los genes y secuencias de acuerdo con la Tabla 11 anteriormente se usan además en el análisis de CDRN2A, ELK1, GSTP1, RARB, PTGS2, RASSF1, ESR2, ONECUT2, BTG4, SLC35F2, HOXB5, LIMK1, HIST1H4J, SEQ ID NO: 35, EPAS1, NOTCH1, SEQ ID NO: 55, PTPRN2, Q9NP73, MX1, DOCK10, CCND2, ISL1, SNAPC2, GRN, H2AFY2, WDFY3, FOS, FAT, Q86SP6, SLC38A1, SNRPN, GPRK5, FBN2, ARHGEF18, RHOC, KBTBD6, NR2E1, PSD, DRG1, Q8N365, SEQ ID NO: 44, Q96S01, CD37, CMYA3, SEQ ID NO: 61, Q8NCX8 y ZNF566 y por ello son aplicables también en la detección de carcinoma o neoplasias de próstata.

Aunque la presente invención se ha descritos con especificidad de acuerdo con algunas de sus modalidades preferidas, los siguientes EJEMPLOS y TABLAS sirven solamente para ilustrar la invención y no pretenden limitar la invención dentro de los principios y alcance de las interpretaciones más amplias y configuraciones equivalentes de estas.

Tablas 1-12

TABLA 1: Montaje experimental de tamizaje MeST con número de muestras en los grupos.

Experimento	Número de comparaciones	Grupo "Agresivo"	Grupo "no Agresivo"
recurrencia de PSA vs. no recurrencia	3	5 tumores de pacientes que tenían recurrencia de PSA <2 años después de la cirugía	5 tumores de pacientes sin recurrencia de PSA después de al menos 4 años de seguimiento
Etapas tardía vs. etapa temprana	1	5 tumores en etapa III y IV	5 tumores en etapa I y II
Gleason alto vs. Gleason bajo	1	5 tumores de grado 4 o 5	5 tumores de grado 1, 2, o 3
NAT (PSA recurrencia)	1	4 muestras normales de	4 muestras normales de

Experimento	Número de comparaciones	Grupo "Agresivo"	Grupo "no Agresivo"
vs. NAT (no recurrencia)		tejido adyacente a los tumores de pacientes que tenían recurrencias de PSA <2 años después de la cirugía	tejido adyacente a los tumores de pacientes sin recurrencia de PSA después de al menos 4 años de seguimiento
Zona periférica vs. Zona de transición	1	5 tumores de zona periférica	5 tumores de zona de transición

TABLA 2: Resultados resumidos de los experimentos de tamizaje MeST

Experimento:	
Número total de MeSTs	441
puntuación > 0	278
MeSTs de MCA	242
puntuación > 0	126
MeSTs de AP-PCR	199
puntuación > 0	152

TABLA 3: Número total de MeSTs y número de MeSTs positivos a partir de cada comparación de tamizaje.

Experimento	Número de MeSTs	Número de MeSTs positivos	Porcentaje de MeSTs positivos
PSA recurrencia vs. no recurrencia I	41	28	68%
PSA recurrencia vs. no recurrencia II	113	74	65%
PSA recurrencia vs. no recurrencia II	69	43	62%
Etapa tardía vs. etapa temprana	64	39	61%
Gleason alto vs. Gleason bajo	63	48	76%
NAT (PSA recurrencia) vs. NAT (no recurrencia)	103	66	64%

Experimento	Número de MeSTs	Número de MeSTs positivos	Porcentaje de MeSTs positivos
recurrencia)			
Zona periférica vs. Zon transición	76	48	63%

TABLA 4 Resumen de los datos de PCR en tiempo real para siete candidatos. Los valores P son de una prueba Wilcoxon y sin corregir para comparaciones múltiples. La sensibilidad se calcula cuando la especificidad se fija a 0.85 o mayor.

Candidato	Nombre del Gen	AUC	Sensibilidad	Especificidad	Valor P (Wilcoxon)
SEQ ID NO: 19	GPR7	0.76	0.50	0.87	0.0008
SEQ ID NO: 35	Región genómica corriente abajo de FOXL2	0.75	0.54	0.87	0.0016
SEQ ID NO: 37	ABHD9	0.7	0.38	0.86	0.0115
SEQ ID NO: 7	GSTP1	0.69	0.38	0.87	0.0176
SEQ ID NO: 63	Región reguladora HIST2H2BF	0.68	0.35	0.87	0.0241
SEQ ID NO: 8	RARB	0.68	0.50	0.85	0.0214
SEQ ID NO: 64	PMF1	0.47	0.15	0.87	0.7138

TABLA 5 A Resumen de las muestras usadas en el estudio chip. Para las categorías de resultados, "Sin recurrencia" significa que el paciente no tuvo una recaída y se dispone de un mínimo de información de 48 meses de seguimiento. "Recurrencia temprana" indica que el paciente experimentó una recaída del PSA en menos de 2 años después de la cirugía. "Otros" incluye los pacientes sin información de seguimiento y pacientes que no se ajustaron al criterio para las categorías de recurrencia.

Categorías de Gleason	Número de la muestra	Categorías de resultados
A) Gleason bajo(1+2, 2+1, 2+2, 2+3, 3+2, y 3+3)	135	1. Sin recurrencia (n=42)
		2. Recurrencia temprana (n=6)
		3. Otro (n=87)
B) Gleason intermedio (2+4, 4+2, 3+4, 4+3, 2+5, 5+2)	73	1. Sin recurrencia (n=34)
		2. Recurrencia temprana (n=33)
		3. Otro (n=6)
C) Gleason alto (3+5, 5+3, 4+4, 4+5, 5+4, y 5+5)	99	1. Sin recurrencia (n=9)
		2. Recurrencia temprana (n=26)
		3. Otro (n=64)
D) Sin información de Gleason	4	1. Sin recurrencia (n=3)
		2. Recurrencia temprana (n=0)
		3. Otro (n=1)

TABLA 6 Valor p de Wilcoxon corregido / AUC / sensibilidad / especificidad de los amplificadores. Se reporta sensibilidad a una especificidad fijada de ~0.75.

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 19	2.05E-06	0.72	0.58	0.75
SEQ ID NO: 41	2.17E-06	0.71	0.51	0.75
SEQ ID NO: 37	4.15E-06	0.71	0.55	0.75
SEQ ID NO: 51	8.68E-06	0.71	0.51	0.75
SEQ ID NO: 35	2.60E-05	0.70	0.48	0.75
SEQ ID NO: 4	5.52E-05	0.69	0.51	0.75
SEQ ID NO: 17	5.89E-05	0.69	0.51	0.75
SEQ ID NO: 16	6.82E-05	0.69	0.47	0.75
SEQ ID NO: 9	9.30E-05	0.69	0.57	0.75
SEQ ID NO: 47	1.80E-04	0.68	0.46	0.75
SEQ ID NO: 49	2.05E-04	0.68	0.49	0.75
SEQ ID NO: 42	3.84E-04	0.68	0.45	0.75
SEQ ID NO: 57	4.34E-04	0.68	0.49	0.75
SEQ ID NO: 1	7.44E-04	0.67	0.51	0.75
SEQ ID NO: 7	8.06E-04	0.67	0.43	0.75

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 62	8.06E-04	0.67	0.49	0.75
SEQ ID NO: 8	9.30E-04	0.67	0.47	0.75
SEQ ID NO: 58	1.24E-03	0.67	0.46	0.75
SEQ ID NO: 3	2.79E-03	0.66	0.48	0.75
SEQ ID NO: 13	3.84E-03	0.66	0.45	0.75
SEQ ID NO: 23	8.68E-03	0.65	0.43	0.75
SEQ ID NO: 29	9.30E-03	0.65	0.41	0.75
SEQ ID NO: 39	9.30E-03	0.65	0.41	0.75
SEQ ID NO: 5	1.05E-02	0.65	0.47	0.75
SEQ ID NO: 56	0.03	0.64	0.46	0.75
SEQ ID NO: 10	0.10	0.62	0.41	0.75
SEQ ID NO: 2	0.11	0.62	0.40	0.75
SEQ ID NO: 6	0.13	0.62	0.40	0.75
SEQ ID NO: 50	0.25	0.61	0.42	0.75
SEQ ID NO: 33	0.36	0.61	0.47	0.75
SEQ ID NO: 15	0.37	0.61	0.35	0.75
SEQ ID NO: 60	0.39	0.61	0.43	0.75
SEQ ID NO: 55	0.46	0.60	0.31	0.75
SEQ ID NO: 52	0.47	0.60	0.38	0.75
SEQ ID NO: 20	0.87	0.60	0.33	0.75
SEQ ID NO: 21	0.87	0.60	0.42	0.75
SEQ ID NO: 24	1.00	0.59	0.35	0.75

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 54	1.00	0.59	0.36	0.75
SEQ ID NO: 30	1.00	0.59	0.36	0.75
SEQ ID NO: 43	1.00	0.42	0.22	0.75
SEQ ID NO: 48	1.00	0.58	0.35	0.75
SEQ ID NO: 45	1.00	0.58	0.32	0.75
SEQ ID NO: 26	1.00	0.58	0.38	0.75
SEQ ID NO: 31	1.00	0.57	0.42	0.75
SEQ ID NO: 38	1.00	0.56	0.34	0.75
SEQ ID NO: 28	1.00	0.56	0.36	0.75
SEQ ID NO: 32	1.00	0.56	0.33	0.75
SEQ ID NO: 40	1.00	0.56	0.30	0.75
SEQ ID NO: 27	1.00	0.55	0.34	0.75
SEQ ID NO: 61	1.00	0.45	0.21	0.75
SEQ ID NO: 11	1.00	0.54	0.34	0.75
SEQ ID NO: 44	1.00	0.54	0.28	0.75
SEQ ID NO: 53	1.00	0.54	0.21	0.75
SEQ ID NO: 22	1.00	0.54	0.28	0.75
SEQ ID NO: 18	1.00	0.47	0.20	0.75
SEQ ID NO: 59	1.00	0.47	0.25	0.75
SEQ ID NO: 36	1.00	0.47	0.22	0.75
SEQ ID NO: 12	1.00	0.53	0.24	0.75
SEQ ID NO: 34	1.00	0.52	0.27	0.75

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 46	1.00	0.48	0.19	0.75
SEQ ID NO: 25	1.00	0.51	0.25	0.75
SEQ ID NO: 14	1.00	0.51	0.24	0.75

TABLA 7: Valor p de Wilcoxon corregido / AUC / sensibilidad / especificidad de los amplificadores. Se reporta sensibilidad a una especificidad fijada de - 0.75.

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 19	2.42E-04	0.72	0.59	0.76
SEQ ID NO: 35	2.48E-04	0.72	0.59	0.76
SEQ ID NO: 37	0.05	0.66	0.37	0.76
SEQ ID NO: 20	0.07	0.66	0.44	0.76
SEQ ID NO: 13	0.08	0.65	0.46	0.76
SEQ ID NO: 9	0.11	0.65	0.37	0.76
SEQ ID NO: 41	0.11	0.65	0.37	0.76
SEQ ID NO: 57	0.11	0.65	0.41	0.76
SEQ ID NO: 51	0.12	0.65	0.49	0.76
SEQ ID NO: 4	0.22	0.64	0.40	0.76
SEQ ID NO: 1	0.24	0.64	0.41	0.76
SEQ ID NO: 10	0.57	0.62	0.38	0.76
SEQ ID NO: 15	0.61	0.62	0.38	0.76
SEQ ID NO: 17	0.74	0.62	0.32	0.76
SEQ ID NO: 58	0.93	0.62	0.40	0.76
SEQ ID NO: 62	1.00	0.61	0.49	0.76
SEQ ID NO: 43	1.00	0.39	0.11	0.76

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 29	1.00	0.61	0.37	0.76
SEQ ID NO: 16	1.00	0.60	0.35	0.76
SEQ ID NO: 23	1.00	0.60	0.40	0.76
SEQ ID NO: 5	1.00	0.60	0.44	0.76
SEQ ID NO: 42	1.00	0.60	0.32	0.76
SEQ ID NO: 49	1.00	0.59	0.37	0.76
SEQ ID NO: 3	1.00	0.59	0.40	0.76
SEQ ID NO: 47	1.00	0.59	0.35	0.76
SEQ ID NO: 7	1.00	0.59	0.27	0.76
SEQ ID NO: 28	1.00	0.59	0.37	0.76
SEQ ID NO: 2	1.00	0.59	0.29	0.76
SEQ ID NO: 18	1.00	0.42	0.14	0.76
SEQ ID NO: 45	1.00	0.58	0.38	0.76
SEQ ID NO: 8	1.00	0.58	0.25	0.76
SEQ ID NO: 46	1.00	0.42	0.13	0.76
SEQ ID NO: 30	1.00	0.58	0.32	0.76
SEQ ID NO: 39	1.00	0.58	0.35	0.76
SEQ ID NO: 54	1.00	0.57	0.32	0.76
SEQ ID NO: 25	1.00	0.43	0.17	0.76
SEQ ID NO: 6	1.00	0.57	0.30	0.76
SEQ ID NO: 11	1.00	0.57	0.41	0.76
SEQ ID NO: 36	1.00	0.57	0.33	0.76

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 60	1.00	0.57	0.32	0.76
SEQ ID NO: 61	1.00	0.43	0.13	0.76
SEQ ID NO: 55	1.00	0.56	0.27	0.76
SEQ ID NO: 33	1.00	0.56	0.27	0.76
SEQ ID NO: 21	1.00	0.56	0.21	0.76
SEQ ID NO: 56	1.00	0.56	0.32	0.76
SEQ ID NO: 24	1.00	0.56	0.27	0.76
SEQ ID NO: 38	1.00	0.55	0.37	0.76
SEQ ID NO: 48	1.00	0.55	0.33	0.76
SEQ ID NO: 26	1.00	0.45	0.22	0.76
SEQ ID NO: 59	1.00	0.54	0.38	0.76
SEQ ID NO: 34	1.00	0.47	0.22	0.76
SEQ ID NO: 52	1.00	0.53	0.22	0.76
SEQ ID NO: 50	1.00	0.53	0.27	0.76
SEQ ID NO: 14	1.00	0.47	0.21	0.76
SEQ ID NO: 22	1.00	0.47	0.27	0.76
SEQ ID NO: 32	1.00	0.52	0.29	0.76
SEQ ID NO: 53	1.00	0.52	0.25	0.76
SEQ ID NO: 44	1.00	0.48	0.27	0.76
SEQ ID NO: 40	1.00	0.49	0.13	0.76
SEQ ID NO: 27	1.00	0.50	0.33	0.76
SEQ ID NO: 12	1.00	0.50	0.21	0.76

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 31	1.00	0.50	0.16	0.76

TABLA 8: Valor p de Wilcoxon corregido / AUC / sensibilidad / especificidad de los amplificadores. Se reporta sensibilidad a una especificidad fijada de ~0.75.

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 19	2.42E-04	0.72	0.59	0.76
SEQ ID NO: 35	2.48E-04	0.72	0.59	0.76
SEQ ID NO: 37	0.05	0.66	0.37	0.76
SEQ ID NO: 20	0.07	0.66	0.44	0.76
SEQ ID NO: 13	0.08	0.65	0.46	0.76
SEQ ID NO: 9	0.11	0.65	0.37	0.76
SEQ ID NO: 41	0.11	0.65	0.37	0.76
SEQ ID NO: 57	0.11	0.65	0.41	0.76
SEQ ID NO: 51	0.12	0.65	0.49	0.76
SEQ ID NO: 4	0.22	0.64	0.40	0.76
SEQ ID NO: 1	0.24	0.64	0.41	0.76
SEQ ID NO: 10	0.57	0.62	0.38	0.76
SEQ ID NO: 15	0.61	0.62	0.38	0.76
SEQ ID NO: 17	0.74	0.62	0.32	0.76
SEQ ID NO: 58	0.93	0.62	0.40	0.76
SEQ ID NO: 62	1.00	0.61	0.49	0.76
SEQ ID NO: 43	1.00	0.39	0.11	0.76
SEQ ID NO: 29	1.00	0.61	0.37	0.76
SEQ ID NO: 16	1.00	0.60	0.35	0.76

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 23	1.00	0.60	0.40	0.76
SEQ ID NO: 5	1.00	0.60	0.44	0.76
SEQ ID NO: 42	1.00	0.60	0.32	0.76
SEQ ID NO: 49	1.00	0.59	0.37	0.76
SEQ ID NO: 3	1.00	0.59	0.40	0.76
SEQ ID NO: 47	1.00	0.59	0.35	0.76
SEQ ID NO: 7	1.00	0.59	0.27	0.76
SEQ ID NO: 28	1.00	0.59	0.37	0.76
SEQ ID NO: 2	1.00	0.59	0.29	0.76
SEQ ID NO: 18	1.00	0.42	0.14	0.76
SEQ ID NO: 45	1.00	0.58	0.38	0.76
SEQ ID NO: 8	1.00	0.58	0.25	0.76
SEQ ID NO: 46	1.00	0.42	0.13	0.76
SEQ ID NO: 30	1.00	0.58	0.32	0.76
SEQ ID NO: 39	1.00	0.58	0.35	0.76
SEQ ID NO: 54	1.00	0.57	0.32	0.76
SEQ ID NO: 25	1.00	0.43	0.17	0.76
SEQ ID NO: 6	1.00	0.57	0.30	0.76
SEQ ID NO: 11	1.00	0.57	0.41	0.76
SEQ ID NO: 36	1.00	0.57	0.33	0.76
SEQ ID NO: 60	1.00	0.57	0.32	0.76
SEQ ID NO: 61	1.00	0.43	0.13	0.76

Marcador	Valor-P	AUC	Sensibilidad	Especificidad
SEQ ID NO: 55	1.00	0.56	0.27	0.76
SEQ ID NO: 33	1.00	0.56	0.27	0.76
SEQ ID NO: 21	1.00	0.56	0.21	0.76
SEQ ID NO: 56	1.00	0.56	0.32	0.76
SEQ ID NO: 24	1.00	0.56	0.27	0.76
SEQ ID NO: 38	1.00	0.55	0.37	0.76
SEQ ID NO: 48	1.00	0.55	0.33	0.76
SEQ ID NO: 26	1.00	0.45	0.22	0.76
SEQ ID NO: 59	1.00	0.54	0.38	0.76
SEQ ID NO: 34	1.00	0.47	0.22	0.76
SEQ ID NO: 52	1.00	0.53	0.22	0.76
SEQ ID NO: 50	1.00	0.53	0.27	0.76
SEQ ID NO: 14	1.00	0.47	0.21	0.76
SEQ ID NO: 22	1.00	0.47	0.27	0.76
SEQ ID NO: 32	1.00	0.52	0.29	0.76
SEQ ID NO: 53	1.00	0.52	0.25	0.76
SEQ ID NO: 44	1.00	0.48	0.27	0.76
SEQ ID NO: 40	1.00	0.49	0.13	0.76
SEQ ID NO: 27	1.00	0.50	0.33	0.76
SEQ ID NO: 12	1.00	0.50	0.21	0.76
SEQ ID NO: 31	1.00	0.50	0.16	0.76

TABLA 9: Iniciadores de acuerdo con el EJEMPLO 3.

Gen	Iniciadores:	Longitud Amplificada:
1• CCND2 (SEQ ID NO: 1)	AAAAACAACCTTAACTC	504
CCND2 (SEQ ID NO.: 1)	AAACAT (SEQ ID NO: 322) TTTGGAGGGATAGAAT GTGA (SEQ ID NO: 321)	
2• CDKN2A (SEQ ID NO: 2)	AGATTATTAGTTTTATT	
CDKN2A (SEQ ID NO.: 2)	TGAGGGATT (SEQ ID NO: 323) CCACCCTAACTCTAACC ATTG (SEQ ID NO: 324)	
3• CD44 (SEQ ID NO: 3)	TTTTTGTTGGGTGTGT TTT	403
CD44 (SEQ ID NO.: 3)	(SEQ ID NO: 325) CACTTAACTCCAATCCC CC (SEQ ID NO: 326)	
4• EDNRB1 (SEQ ID NO: 4)	GA AAAACTTCTCAAATC	446
EDNRB1 (SEQ ID NO.: 4)	AACAA (SEQ ID NO: 328) GAAGGGATGAATGAAT AAAAGT (SEQ ID NO: 327)	
5• ELK1 (SEQ ID NO: 5)		

Gen	Iniciadores:	Longitud Amplificada:
ELK1 (SEQ ID NO.: 5)	TCCAATAAACACAAACC TAAATC (SEQ ID NO: 330) ATATGGGATTGATGGA AGATAG (SEQ ID NO: 329)	
6• FOS (SEQ ID NO: 6)	TTTTTATTAGGATGAG GGATATT (SEQ ID NO: 331) ACTACTTCCCACCCAAC C (SEQ ID NO: 332)	
7• GSTP1 (SEQ ID NO: 7)	CTACTATCTATTACTC CCTAAAC (SEQ ID NO: 334) TTGGTTTATGTTGGGA GTTTTG (SEQ ID NO: 333)	347
8• RARB (SEQ ID NO: 8)	GGGAGTTTTTAAGTTTT GTGAG (SEQ ID NO: 335) TTAATCTTTTCCCAAC CC (SEQ ID NO: 336)	488
9• PTGS2 (SEQ ID NO: 9)	AACATATCAACCTTTCT TAACCTT (SEQ ID NO: 338) AGAGGGGGTAGTTTTT ATTTTT	344

Gen	Iniciadores:	Longitud Amplificada:
	(SEQ ID NO: 337)	
10• RASSF1 (SEQ ID NO: 10)	AGTGGGTAGGTTAAGT GTGTTG (SEQ ID NO: 339) CCCCAAAATCCAACTA AA (SEQ ID NO: 340)	319
RASSF1 (SEQ ID NO.: 10)		
11• ESR2 (SEQ ID NO: 11)	AGTTGGAGAAATTGAAA AGATTA (SEQ ID NO: 341) TAACAAACCCAAAACCT CTCTA (SEQ ID NO: 342)	441
ESR2 (SEQ ID NO.: 11)		
12• DRG1 (SEQ ID NO: 12)	TTTAGTTGTGAAAAAGG GATTT (SEQ ID NO: 343) CCCTATAACCTCCACAC TATCTC (SEQ ID NO: 344)	436
DRG1 (SEQ ID NO.: 12)		
13• DRG1 (SEQ ID NO: 12)	CCCATCCACAATTAAA A (SEQ ID NO: 346) GTTTTGGAGGGAGTAG AGATT (SEQ ID NO: 345)	
DRG1 (SEQ ID NO.: 12)		
14• CMYA3 (SEQ ID NO: 13)		488
CMYA3 (SEQ ID NO.: 13)		

Gen	Iniciadores:	Longitud Amplificada:
	ACTCCCCAAAATCCCA CT (SEQ ID NO: 348) TGTTTATAGGTTTGATGG ATTAGA (SEQ ID NO: 347)	
15• ONECUT2 (SEQ ID NO: 14)		462
ONECUT2 (SEQ ID NO.: 14)	GAAGAGGTGTTGAGAA ATTAAAA (SEQ ID NO: 349) CCCACCCTAAGTTACCA TAAA (SEQ ID NO: 350)	
16• MX1 (SEQ ID NO: 15)	TGTAGGAGAGGTTGGG AAG	341
MX1 (SEQ ID NO.: 15)	(SEQ ID NO: 351) CCAAACATAACATCCAC TAAAA (SEQ ID NO: 352)	
17• DOCK10 (SEQ ID NO: 16)		459
DOCK10 (SEQ ID NO.: 16)	TACCTCTCCCTCTACC AAAC (SEQ ID NO: 354) GTTTTTAAGTGTGGGT GATTT (SEQ ID NO: 353)	
18• BTG4 (SEQ ID NO: 17)		199
BTG4 (SEQ ID NO.: 17)		

Gen	Iniciadores:	Longitud Amplificada:
	AAGAGTTTTAGGAAATG TGTTTT (SEQ ID NO: 355) TTCTACTCAGCAAACCC TCTAC (SEQ ID NO: 356)	
19• DMRTC2 (SEQ ID NO: 18)	TAAGGTAAGGGAAGGT TAGAAA (SEQ ID NO: 357) ATTACCACAACCTCCAA TAAAA (SEQ ID NO: 358)	477
DMRTC2 (SEQ ID NO.: 18)		
20• GPR7 (SEQ ID NO: 19)	TTATTATTTAGATGGA GTGAGGTT (SEQ ID NO: 359) ACCCAAATTACCCACAA A (SEQ ID NO: 360)	442
GPR7 (SEQ ID NO.: 19)		
21• FAT (SEQ ID NO: 20)	TACTCACCCCAATCTTC ACTA (SEQ ID NO: 362) AGATGTTTTATTTGT TTGGGA (SEQ ID NO: 361)	444
FAT (SEQ ID NO: 20)		
22• ISL1 (SEQ ID NO: 21)	ATCTCCCAAAAACAATC ACA (SEQ ID NO: 364) TAAAAATGGAAGGGAA GATAGA	412
ISL1 (SEQ ID NO.: 21)		

Gen	Iniciadores:	Longitud Amplificada:
	(SEQ ID NO: 363)	
23• GPRK5 (SEQ ID NO: 22)	TTGTGGTTATTTGAGA	416
GPRK5 (SEQ ID NO.: 22)	TGGTA (SEQ ID NO: 365) CCCTCCCCCTAACTAAA A (SEQ ID NO: 366)	
24• SLC35F2 (SEQ ID NO: 23)	TTTATTTATTAGGTGAA	366
SLC35F2 (SEQ ID NO.: 23)	GAGTTTGTT (SEQ ID NO: 367) TCCTCCTACCACCTAA AA (SEQ ID NO: 368)	
25• C14orf59 (SEQ ID NO: 24)	AAAACCTCCTCCCTCTA	492
C14orf59 (SEQ ID NO.: 24)	TAAAT (SEQ ID NO: 370) TTGGAGAGATGTGTTG GTTAG (SEQ ID NO: 369)	
26• SNRPN (SEQ ID NO: 25)	CCCCTCTCATTACAACA	407
SNRPN (SEQ ID NO: 25)	ATACT (SEQ ID NO: 372) TTTTTAGAATAAAGGAT TTTAGGG (SEQ ID NO: 371)	
27• ARHGEF18 (SEQ ID NO: 26)		454
ARHGEF18 (SEQ ID NO.: 26)		

Gen	Iniciadores:	Longitud Amplificada:
	TTTTAGGAATGTAGGAT ATAAGGG (SEQ ID NO: 373) CCCCACATAAAACCTA TCC (SEQ ID NO: 374)	
28• SNX8 (SEQ ID NO: 27)		273
SNX8 (SEQ ID NO.: 27)	TCCCTCCTAACTCAACA CTAA (SEQ ID NO: 376) TAGGGTTTGATGATGT GATTTT (SEQ ID NO: 375)	
29• FBN2 (SEQ ID NO: 28)	GAGAGGGAGGGTTAAG GTT	193
FBN2 (SEQ ID NO.: 28)	(SEQ ID NO: 377) ACTACTACCTCCTTTCC CAAAT (SEQ ID NO: 378)	
30• HOXB5 (SEQ ID NO: 29)		356
HOXB5 (SEQ ID NO.: 29)	CTCCTCAATTCTCACCA AAA (SEQ ID NO: 380) GTGGAAAAAGGAGAGT AAATTG (SEQ ID NO: 379)	
31• LIMK1 (SEQ ID NO: 30)		
LIMK1 (SEQ ID NO.: 30)		

Gen	Iniciadores:	Longitud Amplificada:
	AAACCCTACTTCCTACA AACAA (SEQ ID NO: 382) AGGGAGGTTTGGTGTA TTTT (SEQ ID NO: 381)	
32• PSD/Q9H469 (SEQ ID NO: 31)		333
PSD/Q9H469 (SEQ ID NO: 31)	AAGGTATTATTTTGGG GTTTT (SEQ ID NO: 383) AACTATCCAACCTCTTC CACTT (SEQ ID NO: 384)	
33• SLC38A1 (SEQ ID NO: 32)		334
SLC38A1 (SEQ ID NO.: 32)	GGGTGTTGGGGAGTTT TA (SEQ ID NO: 385) CTTACAATAACTCACTA TCCTTTCC (SEQ ID NO: 386)	
34• SLC38A1 (SEQ ID NO: 32)		317
SLC38A1 (SEQ ID NO.: 32)	ACAAACATCCTTTAATA ATTCTCC (SEQ ID NO: 388) AGAGTGTTGGTATTAGAT TTGGTTTT (SEQ ID NO: 387)	
35• HIST1H4J (SEQ ID NO: 33)		421
HIST1H4J (SEQ ID NO.: 33)	TTAGTTGAGAAAGTGG GGGT (SEQ ID NO: 389) CTACCTCAAACCAAAT CCTC	

Gen	Iniciadores:	Longitud Amplificada:
	(SEQ ID NO: 390)	
36• Q96S01 (SEQ ID NO: 34)	ACCCCAATCAACTACAT	498
Q96S01 (SEQ ID NO.: 34)	AACTAA (SEQ ID NO: 392) GTGAGAGTGGGTGTTG AAAT (SEQ ID NO: 391)	
37• Región genómica corriente abajo de FOXL2 (SEQ ID NO: 35)	ATTTTAGGTGTAAGTTT	473
Región genómica corriente abajo de FOXL2 (SEQ ID NO: 35)	AAGGTTGT (SEQ ID NO: 393) ATCTACCTTTCCCACC C (SEQ ID NO: 394)	
38• ORC4L (SEQ ID NO: 36)	GTTGAGAGGTAAGGTA	477
ORC4L (SEQ ID NO.: 36)	TGAAGG (SEQ ID NO: 395) TTAATTCCCCTCTTTAA CCTAATAA (SEQ ID NO: 396)	
39• ABHD9 (SEQ ID NO: 37)	GTGTTAGGGTTTAGGG	209
ABHD9 (SEQ ID NO.: 37)	GTTT (SEQ ID NO: 397) CCTTTCCAACCTCTTCC T (SEQ ID NO: 398)	
40• CD37 (SEQ ID NO: 38)		466
CD37 (SEQ ID NO.: 38)		

Gen	Iniciadores:	Longitud Amplificada:
	CCTCATCAACCAACCGT A (SEQ ID NO: 400) TAGATGGGGATAGGAA GTTGT (SEQ ID NO: 399)	
41• GRN (SEQ ID NO: 39)		466
GRN (SEQ ID NO: 39)	CATTCCAACTAACCCC A (SEQ ID NO: 402) TTATTAGGAGAGGGGA AGAAGT	
	(SEQ ID NO: 401)	
42• EPAS1 (SEQ ID NO: 40)		479
EPAS1 (SEQ ID NO.: 40)	TATTGGATGTTTTGGT AGGTT (SEQ ID NO: 403) AATCACCTCCCTCCCTT A (SEQ ID NO: 404)	
43• NOTCH1 (SEQ ID NO: 41)		424
NOTCH1 (SEQ ID NO.: 41)	CCCACCCTAAAACTCA CTA (SEQ ID NO: 406) GGAGGGGTTTGAGTAA TTG (SEQ ID NO: 405)	
44• MLLT3 (SEQ ID NO: 42)		
MLLT3 (SEQ ID NO.: 42)		

Gen	Iniciadores:	Longitud Amplificada:
	GATTAGGTTTGGAGTT GTTTT (SEQ ID NO: 407) AATCAGATCCATCTTTC ACTTT (SEQ ID NO: 408)	
45• SOLH (SEQ ID NO: 43)		389
SOLH (SEQ ID NO: 43)	CCCAACTCCCCAAATAA A (SEQ ID NO: 410) GGGGATAAGTGGTTAA TGAGT (SEQ ID NO: 409)	
46• SEQ ID NO: 44 (SEQ ID NO: 44)		448
SEQ ID NO: 44 (SEQ ID NO: 44)	ATAGTGTGGATTTTATG GGATATT (SEQ ID NO: 411) CAAAACCTATTCCTTA CCT (SEQ ID NO: 412)	
47• Q8N365 (SEQ ID NO: 45)		475
Q8N365 (SEQ ID NO.: 45)	GTAGTAAATGGGTTATG GATTTTAG (SEQ ID NO: 413) CCATACCACCCACAAA CA (SEQ ID NO: 414)	
48• Q9NWV0 (SEQ ID NO: 46)		330
	TTTGTAGTAGTGGATTTG GTTAGTATT	
Q9NWV0 (SEQ ID NO.: 46)		

Gen	Iniciadores:	Longitud Amplificada:
	(SEQ ID NO: 415) CATCTCTTAACCCCACT TTCA (SEQ ID NO: 416)	
49• SEQ ID NO: 47 (SEQ ID NO: 47)	TCCCTCTAAAAACCAAA AATC (SEQ ID NO: 418) GAGGAAAGAAGGAGGA TATAGG (SEQ ID NO: 417)	335
SEQ ID NO: 47 (SEQ ID NO: 47)		
50• H2AFY2 (SEQ ID NO: 48)	GTTAAAGGGGTATTGG TTTTT (SEQ ID NO: 419) TTTCTTTTCTCTCACC TATTAAAC (SEQ ID NO: 420)	490
H2AFY2 (SEQ ID NO.: 48)		
51• RHOC (SEQ ID NO: 49)	GTAAGAGGGATAGGGA ATTGG (SEQ ID NO: 421) AACACCCAAACCAAAAT AAAA (SEQ ID NO: 422)	440
RHOC (SEQ ID NO: 49)		
52• NR2E1 (SEQ ID NO: 50)	GGAGTTTGTGAAAAGT GGG (SEQ ID NO: 423) ACTCAACAAATACAATA ATCTAAACC (SEQ ID NO: 424)	429
NR2E1 (SEQ ID NO.: 50)		

Gen	Iniciadores:	Longitud Amplificada:
53• KBTBD6 (SEQ ID NO: 51)	ACTATACCAACAAACT ACAAAATAAA (SEQ ID NO: 426) GAAGGTTGAGGAGGAG TTAGA (SEQ ID NO: 425)	228
KBTBD6 (SEQ ID NO.: 51)		
54• TRPM4 (SEQ ID NO: 52)	GGTTGGAAGTGGAGG ATT (SEQ ID NO: 427) CCAACTCTAAAAACAAA AACAA	438
TRPM4 (SEQ ID NO.: 52)		
	(SEQ ID NO: 428)	
55• TCEB3BP1 (SEQ ID NO: 53)	GAGTTGGTTTTGTTGAG GTGT (SEQ ID NO: 429) AAAATACCTTCCCACTA ACCTT (SEQ ID NO: 430)	325
TCEB3BP1 (SEQ ID NO.: 53)		
56• Q8NCX8 (SEQ ID NO: 54)	TAGTGGAATTATGAGG GGG (SEQ ID NO: 431) CCAAATACACCTCTACC AAAA (SEQ ID NO: 432)	300
Q8NCX8 (SEQ ID NO.: 54)		
57• SEQ ID NO: 55 (SEQ ID NO: 55)		384
SEQ ID NO: 55 (SEQ ID NO: 55)		

5

10

15

20

25

30

35

40

45

50

55

Gen	Iniciadores:	Longitud Amplificada:
	CCTTACCCCTCCTCCT AAA (SEQ ID NO: 434) TAGGATTGTGGTTGGT GTT (SEQ ID NO: 433)	
58• SNAPC2 (SEQ ID NO: 56)		362
SNAPC2 (SEQ ID NO.: 56)	GGTTTAGGGTATTTTAA GGGG (SEQ ID NO: 435) AAATAAATCCAATCC CAAA (SEQ ID NO: 436)	
59• PTPRN2 (SEQ ID NO: 57)		378
PTPRN2 (SEQ ID NO.: 57)	CCCTCTACTCACTTTAC CAAAA (SEQ ID NO: 438) GGGGAGGTGTTTAGTG GTT (SEQ ID NO: 437)	
60• WDFY3 (SEQ ID NO: 58)		435
WDFY3 (SEQ ID NO.: 58)	TGTTGGTGGTTATTTTT AATTTTT (SEQ ID NO: 439) ACCCAATTATCCTTTCT CAAC (SEQ ID NO: 440)	
61• ZNF566 (SEQ ID NO: 59)	GAGGATTGTGTTAAG GTTTTT	457
ZNF566 (SEQ ID NO.: 59)		

Gen	Iniciadores:	Longitud Amplificada:
	(SEQ ID NO: 441) CCAACTCCACTATCTAC CACAT (SEQ ID NO: 442)	
62• Q9NP73 (SEQ ID NO: 60)	GGGGTTTTAAGAGTTG GTTTT (SEQ ID NO: 443) CCTCCCTCACTCACTTA CAA (SEQ ID NO: 444)	
63• SEQ ID NO: 61 (SEQ ID NO: 61)	TGGGGATATTGGATGT TTTT (SEQ ID NO: 445) CCTTCCAACCTAACCTC C (SEQ ID NO: 446)	497
64• Q86SP6 (SEQ ID NO: 62)	CCCAACACTCATTAC CTATCT (SEQ ID NO: 448) GGAGTTTTAATTTTGG GATTT (SEQ ID NO: 447)	342

TABLA 10: Detección de oligonucleótidos de acuerdo con el Ejemplo 3.

No:	Gen	Oligo:
1	ONECUT2	TAGAGGCGCGGGTTAT
	(SEQ ID NO.: 14)	(SEQ ID NO.: 449)
2	ONECUT2	TAGAGGTGTGGGTTAT
	(SEQ ID NO.: 14)	(SEQ ID NO.: 450)

No:	Gen	Oligo:
3	ONECUT2	TTGCGATTGGTACGTA
	(SEQ ID NO.: 14)	(SEQ ID NO.: 451)
4	ONECUT2	TGTGATTGGTATGTAGT
	(SEQ ID NO.: 14)	(SEQ ID NO.: 452)
5	ONECUT2	TTTTGTGCGTACGGAT
	(SEQ ID NO.: 14)	(SEQ ID NO.: 453)
6	ONECUT2	TTTTGTGTGTATGGAT
	(SEQ ID NO.: 14)	(SEQ ID NO.: 454)
7	ONECUT2	TTAAGCGGGCGTTGAT
	(SEQ ID NO.: 14)	(SEQ ID NO.: 455)
8	ONECUT2	TTAAGTGGGTGTTGAT
	(SEQ ID NO.: 14)	(SEQ ID NO.: 456)
9	MX1	GTTACGAGTTTATTCGA
	(SEQ ID NO.: 15)	(SEQ ID NO.: 457)
10	MX1	GGTTATGAGTTTATTTGAA
	(SEQ ID NO.: 15)	(SEQ ID NO.: 458)
11	MX1	AACGCGCGAAAGTAA
	(SEQ ID NO.: 15)	(SEQ ID NO.: 459)
12	MX1	TTGGGAATGTGTGAAA
	(SEQ ID NO.: 15)	(SEQ ID NO.: 460)
13	MX1	GAGTTTTCGTCGATTT
	(SEQ ID NO.: 15)	(SEQ ID NO.: 461)
14	MX1	AGGAGTTTTTGTTGATT
	(SEQ ID NO.: 15)	(SEQ ID NO.: 462)
15	MX1	TATGCGCGGGAAGATT
	(SEQ ID NO.: 15)	(SEQ ID NO.: 463)
16	MX1	GTATGTGTGGGAAGAT
	(SEQ ID NO.: 15)	(SEQ ID NO.: 464)
17	DOCK10	GATCGGAATTCGGGTT
	(SEQ ID NO.: 16)	(SEQ ID NO.: 465)
18	DOCK10	ATTGGAATTGGGTTG
	(SEQ ID NO.: 16)	(SEQ ID NO.: 466)
19	DOCK10	TAGTAGTCGCGTTTTT
	(SEQ ID NO.: 16)	(SEQ ID NO.: 467)
20	DOCK10	AGTAGTTGTGTTTTTG

No:	Gen	Oligo:
	(SEQ ID NO.: 16)	(SEQ ID NO.: 468)
21	DOCK10	ATTTTCGCGGGAAGTT
	(SEQ ID NO.: 16)	(SEQ ID NO.: 469)
22	DOCK10	GTGATTTTGTGGGAA
	(SEQ ID NO.: 16)	(SEQ ID NO.: 470)
23	BTG4	AGTTAGCGTTTGTCCG
	(SEQ ID NO.: 17)	(SEQ ID NO.: 471)
24	BTG4	TAGTTAGTGTTTGTGG
	(SEQ ID NO.: 17)	(SEQ ID NO.: 472)
25	BTG4	TTCGGCGTCGATGTAT
	(SEQ ID NO.: 17)	(SEQ ID NO.: 473)
26	BTG4	TTTGGTGTTGATGATT
	(SEQ ID NO.: 17)	(SEQ ID NO.: 474)
27	DMRTC2	GGGTATTCGACGTTT
	(SEQ ID NO.: 18)	(SEQ ID NO.: 475)
28	DMRTC2	AGGGGTATTTGATGTT
	(SEQ ID NO.: 18)	(SEQ ID NO.: 476)
29	DMRTC2	ATTCGAAGCGTTATTG
	(SEQ ID NO.: 18)	(SEQ ID NO.: 477)
30	DMRTC2	TTTGAAGTGTTATTGGT
	(SEQ ID NO.: 18)	(SEQ ID NO.: 478)
31	DMRTC2	AAATCGTATTGGTCGT
	(SEQ ID NO.: 18)	(SEQ ID NO.: 479)
32	DMRTC2	AAATTGTATTGGTTGTT
	(SEQ ID NO.: 18)	(SEQ ID NO.: 480)
33	DMRTC2	AATTCGTGTATAGATCGG
	(SEQ ID NO.: 18)	(SEQ ID NO.: 481)
34	DMRTC2	ATTTGTGTATAGATTGGG
	(SEQ ID NO.: 18)	(SEQ ID NO.: 482)
35	DMRTC2	AAAAGTAGCGCGAGTT
	(SEQ ID NO.: 18)	(SEQ ID NO.: 483)
36	DMRTC2	AGTAGTGTGAGTTTGG
	(SEQ ID NO.: 18)	(SEQ ID NO.: 484)
37	GPR7	GAACGTAGTCGCGGTT
	(SEQ ID NO.: 19)	(SEQ ID NO.: 485)

No:	Gen	Oligo:
38	GPR7	GGAATGTAGTTGTGGT
	(SEQ ID NO.: 19)	(SEQ ID NO.: 486)
39	GPR7	ATTTTGGCGAATTCGG
	(SEQ ID NO.: 19)	(SEQ ID NO.: 487)
40	GPR7	TGGTGAATTTGGGGGA
	(SEQ ID NO.: 19)	(SEQ ID NO.: 488)
41	GPR7	TTAGTCGGTAGGCGTT
	(SEQ ID NO.: 19)	(SEQ ID NO.: 489)
42	GPR7	ATTTAGTTGGTAGGTGT
	(SEQ ID NO.: 19)	(SEQ ID NO.: 490)
43	GPR7	TTTTCGTAGTCGGCGG
	(SEQ ID NO.: 19)	(SEQ ID NO.: 491)
44	GPR7	TTTTTTGTAGTTGGTGG
	(SEQ ID NO.: 19)	(SEQ ID NO.: 492)
45	Grasa	TAAGCGTTAATAGAACGA
	(SEQ ID NO.: 20)	(SEQ ID NO.: 493)
46	Grasa	AGTGTTAATAGAATGAAAT
	(SEQ ID NO.: 20)	(SEQ ID NO.: 494)
47	Grasa	TTGTATTCGTCGTTAT
	(SEQ ID NO.: 20)	(SEQ ID NO.: 495)
48	Grasa	TTTTGTTGTTATAGGAGT
	(SEQ ID NO.: 20)	(SEQ ID NO.: 496)
49	Grasa	ATAGGAGTCGTCGAGA
	(SEQ ID NO.: 20)	(SEQ ID NO.: 497)
50	Grasa	ATAGGAGTTGTTGAGAG
	(SEQ ID NO.: 20)	(SEQ ID NO.: 498)
51	ISL1	TTAATGCGGTCGGTTA
	(SEQ ID NO.: 21)	(SEQ ID NO.: 499)
52	ISL1	AGTTAATGTGGTTGGT
	(SEQ ID NO.: 21)	(SEQ ID NO.: 500)
53	GPRK5	TTTGATTCGCGGTCGG
	(SEQ ID NO.: 22)	(SEQ ID NO.: 501)
54	GPRK5	ATTTGATTTGTGGTTGG
	(SEQ ID NO.: 22)	(SEQ ID NO.: 502)
55	GPRK5	TATCGTCGGTCGAGTT

No:	Gen	Oligo:
	(SEQ ID NO.: 22)	(SEQ ID NO.: 503)
56	GPRK5	TTTATTGTTGGTTGAGT
	(SEQ ID NO.: 22)	(SEQ ID NO.: 504)
57	GPRK5	ATGTCGAGAGTTCGTA
	(SEQ ID NO.: 22)	(SEQ ID NO.: 505)
58	GPRK5	GTTGAGAGTTTGTATGT
	(SEQ ID NO.: 22)	(SEQ ID NO.: 506)
59	SLC35F2	TTTCGGCGTTTAAAT
	(SEQ ID NO.: 23)	(SEQ ID NO.: 507)
60	SLC35F2	TTTTTGGTGTAAAAATT
	(SEQ ID NO.: 23)	(SEQ ID NO.: 508)
61	SLC35F2	ATTTTCGAAGTGTGGG
	(SEQ ID NO.: 23)	(SEQ ID NO.: 509)
62	SLC35F2	TTTGAAGTGTGGGTT
	(SEQ ID NO.: 23)	(SEQ ID NO.: 510)
63	SLC35F2	TTTCGGAAGACGGGAG
	(SEQ ID NO.: 23)	(SEQ ID NO.: 511)
64	SLC35F2	TTTTGGAAGATGGGAG
	(SEQ ID NO.: 23)	(SEQ ID NO.: 512)
65	SLC35F2	AATTCGGTCGTCGTTT
	(SEQ ID NO.: 23)	(SEQ ID NO.: 513)
66	SLC35F2	AGAATTTGGTTGTTGTT
	(SEQ ID NO.: 23)	(SEQ ID NO.: 514)
67	C14orf59	GACGTAGGGACGGAGA
	(SEQ ID NO.: 24)	(SEQ ID NO.: 515)
68	C14orf59	GATGTAGGGATGGAGA
	(SEQ ID NO.: 24)	(SEQ ID NO.: 516)
69	C14orf59	TATCGTGG TTTTTTACGTAT
	(SEQ ID NO.: 24)	(SEQ ID NO.: 517)
70	C14orf59	ATTGTGGTTTTTATGTATA
	(SEQ ID NO.: 24)	(SEQ ID NO.: 518)
71	C14orf59	GTGTTTCGAGAGCGAGT
	(SEQ ID NO.: 24)	(SEQ ID NO.: 519)
72	C14orf59	TGTTTGAGAGTGAGTGT
	(SEQ ID NO.: 24)	(SEQ ID NO.: 520)

No:	Gen	Oligo:
73	C14orf59	TTTATTCGGTGTTCTGA
	(SEQ ID NO.: 24).	(SEQ ID NO.: 521)
74	C14orf59	TATTTGGTGTGAGAG
	(SEQ ID NO.: 24)	(SEQ ID NO.: 522)
75	SNRPN	TTTTTGCGGTCGCGTA
	(SEQ ID NO.: 25)	(SEQ ID NO.: 523)
76	SNRPN	TTTTGTGGTTGTAGG
	(SEQ ID NO.: 25)	(SEQ ID NO.: 524)
77	SNRPN	AGTATGCGCGTTAGTT
	(SEQ ID NO.: 25)	(SEQ ID NO.: 525)
78	SNRPN	TGAGTATGTGTGTAGT
	(SEQ ID NO.: 25)	(SEQ ID NO.: 526)
79	SNRPN	TAGCGGTAGGTTTCGTA
	(SEQ ID NO.: 25)	(SEQ ID NO.: 527)
80	SNRPN	TAGTGGTAGGTTTGTGA
	(SEQ ID NO.: 25)	(SEQ ID NO.: 528)
81	SNRPN	TTTGCGTTAGATTCGT
	(SEQ ID NO.: 25)	(SEQ ID NO.: 529)
82	SNRPN	TGTGTTAGATTTGTTGT
	(SEQ ID NO.: 25)	(SEQ ID NO.: 530)
83	ARHGEF18	GTCGATTCGGTTGATT
	(SEQ ID NO.: 26)	(SEQ ID NO.: 531)
84	ARHGEF18	AAGGTTGATTTGGTTG
	(SEQ ID NO.: 26)	(SEQ ID NO.: 532)
85	ARHGEF18	GGCGGTTTCGAAGATT
	(SEQ ID NO.: 26)	(SEQ ID NO.: 533)
86	ARHGEF18	AGATGGGTGGTTTGA
	(SEQ ID NO.: 26)	(SEQ ID NO.: 534)
87	ARHGEF18	AAGTCGTTATGAGCGA
	(SEQ ID NO.: 26)	(SEQ ID NO.: 535)
88	ARHGEF18	AAGTTGGTTATGAGTGA
	(SEQ ID NO.: 26)	(SEQ ID NO.: 536)
89	ARHGEF18	TGGTCGATACGGTATT
	(SEQ ID NO.: 26)	(SEQ ID NO.: 537)
90	ARHGEF18	TTGTGGTTGATATGGT

No:	Gen	Oligo:
	(SEQ ID NO.: 26)	(SEQ ID NO.: 538)
91	SNX8	AGGACGCGATAGGGAT
	(SEQ ID NO.: 27)	(SEQ ID NO.: 539)
92	SNX8	AGGATGTGATAGGGAT
	(SEQ ID NO.: 27)	(SEQ ID NO.: 540)
93	SNX8	ATTCGTCGTATGTGA
	(SEQ ID NO.: 27)	(SEQ ID NO.: 541)
94	SNX8	ATTTGTTGTATGTGAAG
	(SEQ ID NO.: 27)	(SEQ ID NO.: 542)
95	SNX8	GTCGTTTGCGTATTTA
	(SEQ ID NO.: 27)	(SEQ ID NO.: 543)
96	SNX8	GGTTGTTGTGTATTTAA
	(SEQ ID NO.: 27)	(SEQ ID NO.: 544)
97	SNX8	ATTGTATACGCGCGTT
	(SEQ ID NO.: 27)	(SEQ ID NO.: 545)
98	SNX8	TTGTATATGTGTGTTGG
	(SEQ ID NO.: 27)	(SEQ ID NO.: 546)
99	FBN2	TAAAGCGAGTAGACGG
	(SEQ ID NO.: 28)	(SEQ ID NO.: 547)
100	FBN2	TATAAAGTGAGTAGATGG
	(SEQ ID NO.: 28)	(SEQ ID NO.: 548)
101	HOXB5	ATAGTTTTCGGCGGGT
	(SEQ ID NO.: 29)	(SEQ ID NO.: 549)
102	HOXB5	TATAGTTTTGGTGGGT
	(SEQ ID NO.: 29)	(SEQ ID NO.: 550)
103	HOXB5	TTTTTCGGCGTAGATA
	(SEQ ID NO.: 29)	(SEQ ID NO.: 551)
104	HOXB5	TGTTTTTGGTGTAGAT
	(SEQ ID NO.: 29)	(SEQ ID NO.: 552)
105	HOXB5	AGTCGAGGGCGTTAGA
	(SEQ ID NO.: 29)	(SEQ ID NO.: 553)
106	HOXB5	AGTTGAGGGTGTAGA
	(SEQ ID NO.: 29)	(SEQ ID NO.: 554)
107	HOXB5	TTTTCGAGGAATTCGT
	(SEQ ID NO.: 29)	(SEQ ID NO.: 555)

No:	Gen	Oligo:
108	HOXB5	TTTTTTGAGGAATTTGTT
	(SEQ ID NO.: 29)	(SEQ ID NO.: 556)
109	LIMK1	TATCGGATTATCGCGG
	(SEQ ID NO.: 30)	(SEQ ID NO.: 557)
110	LIMK1	ATTGGATTATTGTGGGG
	(SEQ ID NO.: 30)	(SEQ ID NO.: 558)
111	LIMK1	GTCGGTAGTTTATCGGAT
	(SEQ ID NO.: 30)	(SEQ ID NO.: 559)
112	LIMK1	GTTGGTAGTTTATTGGAT
	(SEQ ID NO.: 30)	(SEQ ID NO.: 560)
113	LIMK1	TAGGAGACGTTACGTT
	(SEQ ID NO.: 30)	(SEQ ID NO.: 561)
114	LIMK1	AGATGTTATGTTAGGGT
	(SEQ ID NO.: 30)	(SEQ ID NO.: 562)
115	PSD/Q9H469	TTTTTCGAAGCGGATT
	(SEQ ID NO.: 31)	(SEQ ID NO.: 563)
116	PSD/Q9H469	TTTGAAGTGGATTTTGG
	(SEQ ID NO.: 31)	(SEQ ID NO.: 564)
117	PSD/Q9H469	GAAACGCGGTTTAAAT
	(SEQ ID NO.: 31)	(SEQ ID NO.: 565)
118	PSD/Q9H469	GGAAATGTGGTTTAAATT
	(SEQ ID NO.: 31)	(SEQ ID NO.: 566)
119	SLC38A1	TTTGCGGTAACGTTTA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 567)
120	SLC38A1	TTGTGGTAATGTTTAGG
	(SEQ ID NO.: 32)	(SEQ ID NO.: 568)
121	SLC38A1	TAGCGGTCGCGGATTA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 569)
122	SLC38A1	GTAGTGGTTGTGGATT
	(SEQ ID NO.: 32)	(SEQ ID NO.: 570)
123	SLC38A1	TTAGGGACGCGAATTA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 571)
124	SLC38A1	AGGGATGTGAATTAGG
	(SEQ ID NO.: 32)	(SEQ ID NO.: 572)
125	SLC38A1	TGCGTTTAAGATCGCGT

No:	Gen	Oligo:
	(SEQ ID NO.: 32)	(SEQ ID NO.: 573)
126	SLC38A1	TGTGTTTAAGATTGTGT
	(SEQ ID NO.: 32)	(SEQ ID NO.: 574)
127	SLC38A1	ATTTCGGTTTTCGAAA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 575)
128	SLC38A1	ATATTTTGGTTTTGAAAA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 576)
129	SLC38A1	AGAGCGTAGTTGATTCGA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 577)
130	SLC38A1	AGAGTGTAGTTGATTGA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 578)
131	SLC38A1	AGGAATTACGTACGTT
	(SEQ ID NO.: 32)	(SEQ ID NO.: 579)
132	SLC38A1	TATGTATGTTTGGAGGG
	(SEQ ID NO.: 32)	(SEQ ID NO.: 580)
133	SLC38A1	TTTGTAAACGCGGGGAA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 581)
134	SLC38A1	TTTTGTAATGTGGGGA
	(SEQ ID NO.: 32)	(SEQ ID NO.: 582)
135	HIST1H4J	TATGGCGGTGATCGTT
	(SEQ ID NO.: 33)	(SEQ ID NO.: 583)
136	HIST1H4J	TTTATGGTGGTGATTGT
	(SEQ ID NO.: 33)	(SEQ ID NO.: 584)
137	HIST1H4J	TTACGGCGTTTCGGAT
	(SEQ ID NO.: 33)	(SEQ ID NO.: 585)
138	HIST1H4J	TTATGGTGTGTTTGGATT
	(SEQ ID NO.: 33)	(SEQ ID NO.: 586)
139	HIST1 H4J	ATGCGTTTACGTCGT
	(SEQ ID NO.: 33)	(SEQ ID NO.: 587)
140	HIST1 H4J	AGATGTGTTTATGTTGT
	(SEQ ID NO.: 33)	(SEQ ID NO.: 588)
141	HIST1H4J	TATTGTCGCGTAGTAT
	(SEQ ID NO.: 33)	(SEQ ID NO.: 589)
142	HIST1H4J	GGATATTGTTGTGTAGT
	(SEQ ID NO.: 33)	(SEQ ID NO.: 590)

No:	Gen	Oligo:
143	HIST1H4J	ATCGAAATCGTAGAGG
	(SEQ ID NO.: 33)	(SEQ ID NO.: 591)
144	HIST1H4J	ATTGAAATTGTAGAGGG
	(SEQ ID NO.: 33)	(SEQ ID NO.: 592)
145	Q96S01	ATCGGTTTTTCGAGGT
	(SEQ ID NO.: 34)	(SEQ ID NO.: 593)
146	Q96S01	ATTGGTTTTTGAGAGTT
	(SEQ ID NO.: 34)	(SEQ ID NO.: 594)
147	Q96S01	GGTCGATTTTCGCGTA
	(SEQ ID NO.: 34)	(SEQ ID NO.: 595)
148	Q96S01	TGGTTGATTTTGTGTA
	(SEQ ID NO.: 34)	(SEQ ID NO.: 596)
149	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	AATCGTGCGGTTGATA
		(SEQ ID NO.: 597)
150	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	TGTAGAATTGTGTGGT
		(SEQ ID NO.: 598)
151	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	AAAATTCGAGGTCGGG
		(SEQ ID NO.: 599)
152	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	AAAATTGAGGTTGGG
		(SEQ ID NO.: 600)
153	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	TTTCGCGGTTTCGAG
		(SEQ ID NO.: 601)
154	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	TTTGTGGTTTGAGAA
		(SEQ ID NO.: 602)
155	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	AATAGGCGATGTACGG
		(SEQ ID NO.: 603)
156	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	TAGGTGATGTATGGGT
		(SEQ ID NO.: 604)

No:	Gen	Oligo:
157	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	TTGGTCGGTTAATCGA (SEQ ID NO.: 605)
158	GENOMIC REGION DOWNSTREAM FROM FOXL2 (SEQ ID NO: 35)	TTTGGTTGGTTAATTGA (SEQ ID NO.: 606)
159	REGIÓN GENÓMICA CORRIENTE ABAJO DE FOXL2 (SEQ ID NO: 35)	TAGCGGTCGCGAAAAT (SEQ ID NO.: 607)
160	REGIÓN GENÓMICA CORRIENTE ABAJO DE FOXL2 (SEQ ID NO: 35)	AGTTTAGTGGTTGTGA (SEQ ID NO.: 608)
161	CMYA3 (SEQ ID NO.: 13)	TTTACGCGGGGTTTGA (SEQ ID NO.: 609).
162	CMYA3 (SEQ ID NO.: 13)	TTTATGTGGGGTTTGTAG (SEQ ID NO.: 610)
163	CMYA3 (SEQ ID NO.: 13)	TTACGTCGTTATTAGGT (SEQ ID NO.: 611)
164	CMYA3 (SEQ ID NO.: 13)	TTTTATGTTGTTATTAGGT (SEQ ID NO.: 612)
165	CMYA3 (SEQ ID NO.: 13)	TATTTGGACGTCGGGT (SEQ ID NO.: 613)
166	CMYA3 (SEQ ID NO.: 13)	TATTTGGATGTTGGGT (SEQ ID NO.: 614)
167	CMYA3 (SEQ ID NO.: 13)	TTTGTCGGAAGCGGA (SEQ ID NO.: 615)
168	CMYA3 (SEQ ID NO.: 13)	TTTGTTGGAAAGTGGA (SEQ ID NO.: 616)
169	ORC4L (SEQ ID NO.: 36)	ATTCGATCGTTACGT (SEQ ID NO.: 617)
170	ORC4L (SEQ ID NO.: 36)	ATTGGATTGTTATGTTT (SEQ ID NO.: 618)
171	ORC4L (SEQ ID NO.: 36)	ATAAGACGGAGTTCGT (SEQ ID NO.: 619)
172	ORC4L (SEQ ID NO.: 36)	AAGATGGAGTTTGTGTTG (SEQ ID NO.: 620)

No:	Gen	Oligo:
173	ORC4L	TTGCGTTATCGACGTT
	(SEQ ID NO.: 36)	(SEQ ID NO.: 621)
174	ORC4L	ATTTTGTGTTATTGATGT
	(SEQ ID NO.: 36)	(SEQ ID NO.: 622)
175	ORC4L	GAGTAACGCGTGTGAT
	(SEQ ID NO.: 36)	(SEQ ID NO.: 623)
176	ORC4L	TATGAGTAATGTGTGTG
	(SEQ ID NO.: 36)	(SEQ ID NO.: 624)
177	ABHD9	TATTTGGGCGCGATAG
	(SEQ ID NO.: 37)	(SEQ ID NO.: 625)
178	ABHD9	ATTTGGGTGTGATAGG
	(SEQ ID NO.: 37)	(SEQ ID NO.: 626)
179	ABHD9	GTGGGACGCGTTGAAG
	(SEQ ID NO.: 37)	(SEQ ID NO.: 627)
180	ABHD9	TGTGGGATGTGTTGAA
	(SEQ ID NO.: 37)	(SEQ ID NO.: 628)
181	ABHD9	GGCGGTTTCGATAGAA
	(SEQ ID NO.: 37)	(SEQ ID NO.: 629)
182	ABHD9	GGGTGGTTTGTATAGA
	(SEQ ID NO.: 37)	(SEQ ID NO.: 630)
183	CCND2	ATAAGTCGTTTCGAGGT
	(SEQ ID NO.: 1)	(SEQ ID NO.: 631)
184	CCND2	AAGTTGTTTGAGGTGT
	(SEQ ID NO.: 1)	(SEQ ID NO.: 632)
185	CCND2	TAGCGGTTACGTAGGA
	(SEQ ID NO.: 1)	(SEQ ID NO.: 633)
186	CCND2	AGTGGTTATGTAGGAAA
	(SEQ ID NO.: 1)	(SEQ ID NO.: 634)
187	CCND2	TACGTGTTTTAACGTAT
	(SEQ ID NO.: 1)	(SEQ ID NO.: 635)
188	CCND2	TGATATGTGTTTTAATGTA
	(SEQ ID NO.: 1)	(SEQ ID NO.: 636)
189	CCND2	TTAGGGTCGTCGTAGGT
	(SEQ ID NO.: 1)	(SEQ ID NO.: 637)
190	CCND2	TTAGGGTTGTTGTAGGT

No:	Gen	Oligo:
	(SEQ ID NO.: 1)	(SEQ ID NO.: 638)
191	CCND2	TTAGTACGGTCGGTTT
	(SEQ ID NO.: 1)	(SEQ ID NO.: 639)
192	CCND2	GTTAGTATGGTTGGTTT
	(SEQ ID NO.: 1)	(SEQ ID NO.: 640)
193	CDKN2A	TAGGTATCGCGTACGT
	(SEQ ID NO.: 2)	(SEQ ID NO.: 641)
194	CDKN2A	TTAGGTATTGTGTATGTT
	(SEQ ID NO.: 2)	(SEQ ID NO.: 642)
195	CDKN2A	TTCGCGTCGTGGAGTA
	(SEQ ID NO.: 2)	(SEQ ID NO.: 643)
196	CDKN2A	TTTTGTGTTGTGGAGTA
	(SEQ ID NO.: 2)	(SEQ ID NO.: 644)
197	CDKN2A	TAGTCGCGCGTAGGTA
	(SEQ ID NO.: 2)	(SEQ ID NO.: 645)
198	CDKN2A	TAGTTGTGTGTAGGTAT
	(SEQ ID NO.: 2)	(SEQ ID NO.: 646)
199	CDKN2A	TAGGTATCGTGCGATA
	(SEQ ID NO.: 2)	(SEQ ID NO.: 647)
200	CDKN2A	GTAGGTATTGTGTGATAT
	(SEQ ID NO.: 2)	(SEQ ID NO.: 648)
201	CD44	TAGGTTCGGTTCGTTAT
	(SEQ ID NO.: 3)	(SEQ ID NO.: 649)
202	CD44	TAGGTTTGTTTGTATT
	(SEQ ID NO.: 3)	(SEQ ID NO.: 650)
203	CD44	GTTTCGCGTTTAGGGA
	(SEQ ID NO.: 3)	(SEQ ID NO.: 651)
204	CD44	GTTTTGTGTTTAGGGAT
	(SEQ ID NO.: 3)	(SEQ ID NO.: 652)
205	CD44	GTTTCGTTTCGGATATTA
	(SEQ ID NO.: 3)	(SEQ ID NO.: 653)
206	CD44	TTTGTGTTGGATATTATGG
	(SEQ ID NO.: 3)	(SEQ ID NO.: 654)
207	CD44	TTTGCGTAGATCGGT
	(SEQ ID NO.: 3)	(SEQ ID NO.: 655)

No:	Gen	Oligo:
208	CD44	TTTGGTGTAGATTGGT
	(SEQ ID NO.: 3)	(SEQ ID NO.: 656)
209	EDNRB1	TTCGTTTTTCGGAAG
	(SEQ ID NO.: 4)	(SEQ ID NO.: 657)
210	EDNRB1	TTTGTTTTTGGAAGG
	(SEQ ID NO.: 4)	(SEQ ID NO.: 658)
211	EDNRB1	TAGAGTCGGATTCGTT
	(SEQ ID NO.: 4)	(SEQ ID NO.: 659)
212	EDNRB1	AGTTGGATTGTTTGT
	(SEQ ID NO.: 4)	(SEQ ID NO.: 660)
213	EDNRB1	GTATTTTCGTAGCGTT
	(SEQ ID NO.: 4)	(SEQ ID NO.: 661)
214	EDNRB1	TGGTATTTTGTAGTGT
	(SEQ ID NO.: 4)	(SEQ ID NO.: 662)
215	EDNRB1	ATTTCGAGTAAACGGT
	(SEQ ID NO.: 4)	(SEQ ID NO.: 663)
216	EDNRB1	TTTGAGTAAATGGTGGA
	(SEQ ID NO.: 4)	(SEQ ID NO.: 664)
217	ELK1	TGTCGTACGTTATGTT
	(SEQ ID NO.: 5)	(SEQ ID NO.: 665)
218	ELK1	GGTGTTGTATGTTATGT
	(SEQ ID NO.: 5)	(SEQ ID NO.: 666)
219	ELK1	TGGGCGTAGTAGTCGG
	(SEQ ID NO.: 5)	(SEQ ID NO.: 667)
220	ELK1	ATGGGTGTAGTAGTTGG
	(SEQ ID NO.: 5)	(SEQ ID NO.: 668)
221	ELK1	ATTGGGTTTCGCGTAGG
	(SEQ ID NO.: 5)	(SEQ ID NO.: 669)
222	ELK1	ATTGGGTTTGTGTAGG
	(SEQ ID NO.: 5)	(SEQ ID NO.: 670)
223	ELK1	TTGATTGGCGGACGAG
	(SEQ ID NO.: 5)	(SEQ ID NO.: 671)
224	ELK1	TTGATTGGTGGATGAG
	(SEQ ID NO.: 5)	(SEQ ID NO.: 672)
225	FOS	TACGGATTGGTCGTTT

No:	Gen	Oligo:
	(SEQ ID NO.: 6)	(SEQ ID NO.: 673)
226	FOS	TATGGATTGGTTGTTT
	(SEQ ID NO.: 6)	(SEQ ID NO.: 674)
227	FOS	TTCGATTAGTTCGGAT
	(SEQ ID NO.: 6)	(SEQ ID NO.: 675)
228	FOS	TATTTTGATTAGTTTGGAT
	(SEQ ID NO.: 6)	(SEQ ID NO.: 676)
229	FOS	TTTCGTGGTTTTATCGTA
	(SEQ ID NO.: 6)	(SEQ ID NO.: 677)
230	FOS	TTTTGTGGTTTTATTGTA
	(SEQ ID NO.: 6)	(SEQ ID NO.: 678)
231	FOS	GTCGAGCGTAGAGTAT
	(SEQ ID NO.: 6)	(SEQ ID NO.: 679)
232	FOS	TAGGAGGTTGAGTGTA
	(SEQ ID NO.: 6)	(SEQ ID NO.: 680)
233	GSTP1	GTCGGTCGTAGAGGGG
	(SEQ ID NO.: 7)	(SEQ ID NO.: 681)
234	GSTP1	GTTGGTTGTAGAGGGG
	(SEQ ID NO.: 7)	(SEQ ID NO.: 682)
235	GSTP1	TTCGCGGTTTTCGAGT
	(SEQ ID NO.: 7)	(SEQ ID NO.: 683)
236	GSTP1	TTTGTGGTTTTTGAGTT
	(SEQ ID NO.: 7)	(SEQ ID NO.: 684)
237	GSTP1	GTCGCGCGTATTTATT
	(SEQ ID NO.: 7)	(SEQ ID NO.: 685)
238	GSTP1	GGGTTGTGTGTTTAT
	(SEQ ID NO.: 7)	(SEQ ID NO.: 686)
239	GSTP1	GAGTCGTCGCGTAGTT
	(SEQ ID NO.: 7)	(SEQ ID NO.: 687)
240	GSTP1	GGAGTTGTTGTGTAGTT
	(SEQ ID NO.: 7)	(SEQ ID NO.: 688)
241	CD37	ATCGAGAGCGTTATGA
	(SEQ ID NO.: 38)	(SEQ ID NO.: 689)
242	CD37	AGGAATATTGAGAGTGT
	(SEQ ID NO.: 38)	(SEQ ID NO.: 690)

No:	Gen	Oligo:
243	CD37	TATCGAGCGAGTCGGT
	(SEQ ID NO.: 38)	(SEQ ID NO.: 691)
244	CD37	TATTGAGTGAGTTGGTT
	(SEQ ID NO.: 38)	(SEQ ID NO.: 692)
245	CD37	TTAGCGTACGTGACGG
	(SEQ ID NO.: 38)	(SEQ ID NO.: 693)
246	CD37	AGTGTATGTGATGGGG
	(SEQ ID NO.: 38)	(SEQ ID NO.: 694)
247	CD37	GGGTACGTAGTTACGT
	(SEQ ID NO.: 38)	(SEQ ID NO.: 695)
248	CD37	TATGTAGTTATGTGGGT
	(SEQ ID NO.: 38)	(SEQ ID NO.: 696)
249	GRN	TAGCGCGATGATTCGT
	(SEQ ID NO.: 39)	(SEQ ID NO.: 697)
250	GRN	AGTGTGATGATTTGTTT
	(SEQ ID NO.: 39)	(SEQ ID NO.: 698)
251	GRN	TAATCGGGTAGCGTTT
	(SEQ ID NO.: 39)	(SEQ ID NO.: 699)
252	GRN	TAGTAATTGGGTAGTGT
	(SEQ ID NO.: 39)	(SEQ ID NO.: 700)
253	GRN	TTCGATTTTCGCGGTTT
	(SEQ ID NO.: 39)	(SEQ ID NO.: 701)
254	GRN	GTTTGATTTTGTGGTTT
	(SEQ ID NO.: 39)	(SEQ ID NO.: 702)
255	GRN	TTTAACGGGGCGTTAT
	(SEQ ID NO.: 39)	(SEQ ID NO.: 703)
256	GRN	AATGGGGTGTTATTGT
	(SEQ ID NO.: 39)	(SEQ ID NO.: 704)
257	EPAS1	TTCGGCGTTATTCGAG
	(SEQ ID NO.: 40)	(SEQ ID NO.: 705)
258	EPAS1	GGTTTGGTGTTATTTGA
	(SEQ ID NO.: 40)	(SEQ ID NO.: 706)
259	EPAS1	TATTCGTGCGGTTTTA
	(SEQ ID NO.: 40)	(SEQ ID NO.: 707)
260	EPAS1	ATTTGTGTGGTTTATGT

No:	Gen	Oligo:
	(SEQ ID NO.: 40)	(SEQ ID NO.: 708)
261	EPAS1	GAATTTAACGCGCGGT
	(SEQ ID NO.: 40)	(SEQ ID NO.: 709)
262	EPAS1	GGAATTTAATGTGTGGT
	(SEQ ID NO.: 40)	(SEQ ID NO.: 710)
263	EPAS1	TTCGCGAGTTTTTCGG
	(SEQ ID NO.: 40)	(SEQ ID NO.: 711)
264	EPAS1	TTTGTGAGTTTTTTGGTA
	(SEQ ID NO.: 40)	(SEQ ID NO.: 712)
265	NOTCH1	TTTACGGGCGGGAGTT
	(SEQ ID NO.: 41)	(SEQ ID NO.: 713)
266	NOTCH1	TTTTATGGGTGGGAGT
	(SEQ ID NO.: 41)	(SEQ ID NO.: 714)
267	NOTCH1	TAGCGGGCGAGTAGTT
	(SEQ ID NO.: 41)	(SEQ ID NO.: 715)
268	NOTCH1	TAGTGGGTGAGTAGTT
	(SEQ ID NO.: 41)	(SEQ ID NO.: 716)
269	NOTCH1	AGGCGGTTTCGATTTT
	(SEQ ID NO.: 41)	(SEQ ID NO.: 717)
270	NOTCH1	TGGAGGTGTTTTGAT
	(SEQ ID NO.: 41)	(SEQ ID NO.: 718)
271	NOTCH1	ATTTCGGCGGTTGGAT
	(SEQ ID NO.: 41)	(SEQ ID NO.: 719)
272	NOTCH1	AGGTAATTTTGGTGGTT
	(SEQ ID NO.: 41)	(SEQ ID NO.: 720)
273	MLLT3	TAATTCGGTTAGATTTTCGG
	(SEQ ID NO.: 42)	(SEQ ID NO.: 721)
274	MLLT3	TAATTTGGTTAGATTTTGG
	(SEQ ID NO.: 42)	(SEQ ID NO.: 722)
275	MLLT3	TTTAGAGTCGCGTTTT
	(SEQ ID NO.: 42)	(SEQ ID NO.: 723)
276	MLLT3	TAGAGTTGTGTTTTGT
	(SEQ ID NO.: 42)	(SEQ ID NO.: 724)
277	MLLT3	TAGAATAGCGCGGTTA
	(SEQ ID NO.: 42)	(SEQ ID NO.: 725)

No:	Gen	Oligo:
278	MLLT3	GATAGAATAGTGTGGTTA
	(SEQ ID NO.: 42)	(SEQ ID NO.: 726)
279	SOLH	TAGGGGACGTGTACGA
	(SEQ ID NO.: 43)	(SEQ ID NO.: 727)
280	SOLH	TAGGGGATGTGTATGA
	(SEQ ID NO.: 43)	(SEQ ID NO.: 728)
281	SOLH	GACGGAACGTATGTTT
	(SEQ ID NO.: 43)	(SEQ ID NO.: 729)
282	SOLH	TGGGGATGGAATGTAT
	(SEQ ID NO.: 43)	(SEQ ID NO.: 730)
283	SOLH	ATTCGTGGGGACGGAA
	(SEQ ID NO.: 43)	(SEQ ID NO.: 731)
284	SOLH	ATTTGTGGGGATGGAA
	(SEQ ID NO.: 43)	(SEQ ID NO.: 732)
285	SEQ ID NO: 44	TAGGGTTCGTTCTGATT
	(SEQ ID NO.: 44)	(SEQ ID NO.: 733)
286	SEQ ID NO: 44	TTTAGGGTTTGTGTAT
	(SEQ ID NO.: 44)	(SEQ ID NO.: 734)
287	SEQ ID NO: 44	TTACGATTTTCGGGGT
	(SEQ ID NO.: 44)	(SEQ ID NO.: 735)
288	SEQ ID NO: 44	TTTATGATTTTGGGGT
	(SEQ ID NO.: 44)	(SEQ ID NO.: 736)
289	SEQ ID NO: 44	AAGTAGACGTCGAGA
	(SEQ ID NO.: 44)	(SEQ ID NO.: 737)
290	SEQ ID NO: 44	TAGATGTTGGAGAGGG
	(SEQ ID NO.: 44)	(SEQ ID NO.: 738)
291	SEQ ID NO: 44	TGTCGTTACGTTTAGG
	(SEQ ID NO.: 44)	(SEQ ID NO.: 739)
292	SEQ ID NO: 44	GTTGTTATGTTTAGGGT
	(SEQ ID NO.: 44)	(SEQ ID NO.: 740)
293	Q8N365	AGATTCGAGAGACGG
	(SEQ ID NO.: 45)	(SEQ ID NO.: 741)
294	Q8N365	TAGATTTGGAGAGATGG
	(SEQ ID NO.: 45)	(SEQ ID NO.: 742)
295	Q8N365	AGATCGGCGTAGGGAT

No:	Gen	Oligo:
	(SEQ ID NO.: 45)	(SEQ ID NO.: 743)
296	Q8N365	AGATTGGTGTAGGGAT
	(SEQ ID NO.: 45)	(SEQ ID NO.: 744)
297	Q8N365	TACGTGTTATCGGCGA
	(SEQ ID NO.: 45)	(SEQ ID NO.: 745)
298	Q8N365	TATGTGTTATTGGTGATA
	(SEQ ID NO.: 45)	(SEQ ID NO.: 746)
299	Q8N365	TACGTGTCGGGTCGTA
	(SEQ ID NO.: 45)	(SEQ ID NO.: 747)
300	Q8N365	GTATGTGTTGGGTTGTA
	(SEQ ID NO.: 45)	(SEQ ID NO.: 748)
301	Q9NWV0	GTCGCGTGGATACGTG
	(SEQ ID NO.: 46)	(SEQ ID NO.: 749)
302	Q9NWV0	GGTTGTGTGGATATGT
	(SEQ ID NO.: 46)	(SEQ ID NO.: 750)
303	Q9NWV0	TTGGCGATTTTACGA
	(SEQ ID NO.: 46)	(SEQ ID NO.: 751)
304	Q9NWV0	TTGGTGATTTTATGAGA
	(SEQ ID NO.: 46)	(SEQ ID NO.: 752)
305	Q9NWV0	TGTCGTAGCGTTATGT
	(SEQ ID NO.: 46)	(SEQ ID NO.: 753)
306	Q9NWV0	AGGTGTTGTAGTGTTAT
	(SEQ ID NO.: 46)	(SEQ ID NO.: 754)
307	SEQ ID NO: 47	GTAACGCGTTTGTTTT
	(SEQ ID NO.: 47)	(SEQ ID NO.: 755)
308	SEQ ID NO: 47	TGGTAATGTGTTTGGT
	(SEQ ID NO.: 47)	(SEQ ID NO.: 756)
309	SEQ ID NO: 47	TTCGAGCGTTTACGT
	(SEQ ID NO.: 47)	(SEQ ID NO.: 757)
310	SEQ ID NO: 47	TTTGAGTGTTTATGTT
	(SEQ ID NO.: 47)	(SEQ ID NO.: 758)
311	SEQ ID NO: 47	TTACGTGGGAGCGTTT
	(SEQ ID NO.: 47)	(SEQ ID NO.: 759)
312	SEQ ID NO: 47	TTATGTGGGAGTGTTT
	(SEQ ID NO.: 47)	(SEQ ID NO.: 760)

No:	Gen	Oligo:
313	SEQ ID NO: 47	TAGTTTACGCGGGTA
	(SEQ ID NO.: 47)	(SEQ ID NO.: 761)
314	SEQ ID NO: 47	AGTTTATGTGGGTATG
	(SEQ ID NO.: 47)	(SEQ ID NO.: 762)
315	H2AFY2	ATGAAAGGCGCGAGAA
	(SEQ ID NO.: 48)	(SEQ ID NO.: 763)
316	H2AFY2	ATGAAAGGTGTGAGAA
	(SEQ ID NO.: 48)	(SEQ ID NO.: 764)
317	H2AFY2	GAATCGTGGTTTCGTT
	(SEQ ID NO.: 48)	(SEQ ID NO.: 765)
318	H2AFY2	GGAATTGTGGTTTTGT
	(SEQ ID NO.: 48)	(SEQ ID NO.: 766)
319	H2AFY2	TAGTTTATCGCGGTAA
	(SEQ ID NO.: 48)	(SEQ ID NO.: 767)
320	H2AFY2	TTTATTGTGGTAAATGGT
	(SEQ ID NO.: 48)	(SEQ ID NO.: 768)
321	RHOC	TGCGGTTCTGAAGATTA
	(SEQ ID NO.: 49)	(SEQ ID NO.: 769)
322	RHOC	GGTGTGGTTTGAAGAT
	(SEQ ID NO.: 49)	(SEQ ID NO.: 770)
323	RHOC	GAACGCGTTTATAGCGT
	(SEQ ID NO.: 49)	(SEQ ID NO.: 771)
324	RHOC	GGAATGTGTTTTAGTGT
	(SEQ ID NO.: 49)	(SEQ ID NO.: 772)
325	RHOC	TTTAGCGTCGGGGAT
	(SEQ ID NO.: 49)	(SEQ ID NO.: 773)
326	RHOC	TTTAGTGTTGGGGAT
	(SEQ ID NO.: 49)	(SEQ ID NO.: 774)
327	NR2E1	TAGTCGTATTAGCGGT
	(SEQ ID NO.: 50)	(SEQ ID NO.: 775)
328	NR2E1	TAGTTGTATTAGTGTTT
	(SEQ ID NO.: 50)	(SEQ ID NO.: 776)
329	NR2E1	TGCGTTTTATTGCGGG
	(SEQ ID NO.: 50)	(SEQ ID NO.: 777)
330	NR2E1	TGTGTTTTATTGTGGT

No:	Gen	Oligo:
	(SEQ ID NO.: 50)	(SEQ ID NO.: 778)
331	NR2E1	TTTCGAAGTTTCGCGG
	(SEQ ID NO.: 50)	(SEQ ID NO.: 779)
332	NR2E1	TTTGAAGTTTGTGGG
	(SEQ ID NO.: 50)	(SEQ ID NO.: 780)
333	NR2E1	TAGCGCGAATCGTTTA
	(SEQ ID NO.: 50)	(SEQ ID NO.: 781)
334	NR2E1	AGTGTGAATTGTTAGT
	(SEQ ID NO.: 50)	(SEQ ID NO.: 782)
335	KBTBD6	AGACGTTTCGCGTTTT
	(SEQ ID NO.: 51)	(SEQ ID NO.: 783)
336	KBTBD6	GAAGATGTTTGTGTTT
	(SEQ ID NO.: 51)	(SEQ ID NO.: 784)
337	KBTBD6	TTTCGGAAGACGTTT
	(SEQ ID NO.: 51)	(SEQ ID NO.: 785)
338	KBTBD6	AGTTTGGGAAGATGT
	(SEQ ID NO.: 51)	(SEQ ID NO.: 786)
339	KBTBD6	TATTAGCGTTCGTCGT
	(SEQ ID NO.: 51)	(SEQ ID NO.: 787)
340	KBTBD6	GTTATTAGTGTGTTGT
	(SEQ ID NO.: 51)	(SEQ ID NO.: 788)
341	TRPM4	TAGGTGAGCGTCGGAT
	(SEQ ID NO.: 52)	(SEQ ID NO.: 789)
342	TRPM4	TAGGTGAGTGTGGAT
	(SEQ ID NO.: 52)	(SEQ ID NO.: 790)
343	TRPM4	AGTAGAGTCGGCGGAG
	(SEQ ID NO.: 52)	(SEQ ID NO.: 791)
344	TRPM4	AGTAGAGTTGGTGGAG
	(SEQ ID NO.: 52)	(SEQ ID NO.: 792)
345	TCEB3BP1	GACGTTAGCGAGTATT
	(SEQ ID NO.: 53)	(SEQ ID NO.: 793)
346	TCEB3BP1	AGGGATGTTAGTGAGT
	(SEQ ID NO.: 53)	(SEQ ID NO.: 794)
347	TCEB3BP1	TGTCGGGTCGAGTAGT
	(SEQ ID NO.: 53)	(SEQ ID NO.: 795)

5
10
15
20
25
30
35
40
45
50
55

No:	Gen	Oligo:
348	TCEB3BP1	TGTTGGGTTGAGTAGT
	(SEQ ID NO.: 53)	(SEQ ID NO.: 796)
349	Q8NCX8	TTACGGCGGG TTTTAA
	(SEQ ID NO.: 54)	(SEQ ID NO.: 797)
350	Q8NCX8	TTATGGTGGGTTTTATT
	(SEQ ID NO.: 54)	(SEQ ID NO.: 798)
351	Q8NCX8	TTCGGTTTATACGATT
	(SEQ ID NO.: 54)	(SEQ ID NO.: 799)
352	Q8NCX8	ATTTGGTTTATATGATTTT
	(SEQ ID NO.: 54)	(SEQ ID NO.: 800)
353	Q8NCX8	TTACGTCGATTATCGG
	(SEQ ID NO.: 54)	(SEQ ID NO.: 801)
354	Q8NCX8	TATGTTGATTATTGGGGT
	(SEQ ID NO.: 54)	(SEQ ID NO.: 802)
355	SEQ ID NO: 55	AATTACGCGGGTGTT
	(SEQ ID NO.: 55)	(SEQ ID NO.: 803)
356	SEQ ID NO: 55	GGAATTTATGTGGGTG
	(SEQ ID NO.: 55)	(SEQ ID NO.: 804)
357	SEQ ID NO: 55	TAGTTCGTTTCGGTTA
	(SEQ ID NO.: 55)	(SEQ ID NO.: 805)
358	SEQ ID NO: 55	AGTTTGTTTTGGTTAGT
	(SEQ ID NO.: 55)	(SEQ ID NO: 806).
359	SEQ ID NO: 55	GTCGTGTACGAATTCGT
	(SEQ ID NO.: 55)	(SEQ ID NO.: 807)
360	SEQ ID NO: 55	GTTGTGTATGAATTTGTT
	(SEQ ID NO.: 55)	(SEQ ID NO.: 808)
361	SEQ ID NO: 55	AGCGTTTAAGTCGCGG
	(SEQ ID NO.: 55)	(SEQ ID NO.: 809)
362	SEQ ID NO: 55	TAGTGTTTAAGTTGTGG
	(SEQ ID NO.: 55)	(SEQ ID NO.: 810)
363	SNAPC2	TTACGATTTTCGGTGTT
	(SEQ ID NO.: 56)	(SEQ ID NO.: 811)
364	SNAPC2	TAGAGTTATGATTTTGGT
	(SEQ ID NO.: 56)	(SEQ ID NO.: 812)
365	SNAPC2	AGGCGTGCGTCGATAT

No:	Gen	Oligo:
	(SEQ ID NO.: 56)	(SEQ ID NO.: 813)
366	SNAPC2	AGGTGTGTGTTGATATT
	(SEQ ID NO.: 56)	(SEQ ID NO.: 814)
367	SNAPC2	GTAGCGTCGAGGGTTT
	(SEQ ID NO.: 56)	(SEQ ID NO.: 815)
368	SNAPC2	GTAGTGTGAGGGTTT
	(SEQ ID NO.: 56)	(SEQ ID NO.: 816)
369	SNAPC2	ATACGTTGACGTAGTT
	(SEQ ID NO.: 56)	(SEQ ID NO.: 817)
370	SNAPC2	TTGTATATGTTGATAGT
	(SEQ ID NO.: 56)	(SEQ ID NO.: 818)
371	PTPRN2	GACGTAGTTCGTACGT
	(SEQ ID NO.: 57)	(SEQ ID NO.: 819)
372	PTPRN2	GGATGTAGTTTGTATGT
	(SEQ ID NO.: 57)	(SEQ ID NO.: 820)
373	PTPRN2	TTAGTCGTTTCGAGAT
	(SEQ ID NO.: 57)	(SEQ ID NO.: 821)
374	PTPRN2	GTTTTAGTTGTTTGAGA
	(SEQ ID NO.: 57)	(SEQ ID NO.: 822)
375	PTPRN2	TTACGCGTATCGGGAT
	(SEQ ID NO.: 57)	(SEQ ID NO.: 823)
376	PTPRN2	TATGTGATTGGGATTTA
	(SEQ ID NO.: 57)	(SEQ ID NO.: 824)
377	PTPRN2	TTCGCGTTTCGTAAGA
	(SEQ ID NO.: 57)	(SEQ ID NO.: 825)
378	PTPRN2	TTTGTGTTTGTAGATAT
	(SEQ ID NO.: 57)	(SEQ ID NO.: 826)
379	WDFY3	TATTGAGTCGGTCGTA
	(SEQ ID NO.: 58)	(SEQ ID NO.: 827)
380	WDFY3	ATTGAGTTGGTTGTAGA
	(SEQ ID NO.: 58)	(SEQ ID NO.: 828)
381	WDFY3	TAGATAGCGTCGTTGG
	(SEQ ID NO.: 58)	(SEQ ID NO.: 829)
382	WDFY3	TAGATAGTGTGTTGGA
	(SEQ ID NO.: 58)	(SEQ ID NO.: 830)

No:	Gen	Oligo:
383	ZNF566	ATAAAATACGACGTTGA
	(SEQ ID NO.: 59)	(SEQ ID NO.: 831)
384	ZNF566	ATAAAATATGATGTTGAATT
	(SEQ ID NO.: 59)	(SEQ ID NO.: 832)
385	ZNF566	AGCGTTTTTTACGAAT
	(SEQ ID NO.: 59)	(SEQ ID NO.: 833)
386	ZNF566	TAGTGTTTTTTATGAATGA
	(SEQ ID NO.: 59)	(SEQ ID NO.: 834)
387	Q9NP73	TATTACGGATTTTCGTT
	(SEQ ID NO.: 60)	(SEQ ID NO.: 835)
388	Q9NP73	GTTATTATGGATTTTTGTT
	(SEQ ID NO.: 60)	(SEQ ID NO.: 836)
389	Q9NP73	TGCGTTATATTCGGTT
	(SEQ ID NO.: 60)	(SEQ ID NO.: 837)
390	Q9NP73	AGTTGTGTTATATTTGGT
	(SEQ ID NO.: 60)	(SEQ ID NO.: 838)
391	Q9NP73	AATTCGCGTTATGAAG
	(SEQ ID NO.: 60)	(SEQ ID NO.: 839)
392	Q9NP73	TTGAGGAATTTGTGTTAT
	(SEQ ID NO.: 60)	(SEQ ID NO.: 840)
393	Q9NP73	GTCGGCGTTCGATAGT
	(SEQ ID NO.: 60)	(SEQ ID NO.: 841)
394	Q9NP73	TGTTGGTGTTCGATAGT
	(SEQ ID NO.: 60)	(SEQ ID NO.: 842)
395	SEQ ID NO: 61	AGAATTCGTAAACGGG
	(SEQ ID NO.: 61)	(SEQ ID NO.: 843)
396	SEQ ID NO: 61	ATTTGTAAATGGGGGT
	(SEQ ID NO.: 61)	(SEQ ID NO.: 844)
397	SEQ ID NO: 61	TTCGGTTATCGACGGT
	(SEQ ID NO.: 61)	(SEQ ID NO.: 845)
398	SEQ ID NO: 61	TTTGGTTATTGATGGTT
	(SEQ ID NO.: 61)	(SEQ ID NO.: 846)
399	Q86SP6	ATCGTGAGCGTAGCGT
	(SEQ ID NO.: 62)	(SEQ ID NO.: 847)
400	Q86SP6	ATTGTGAGTGTAGTGTA

No:	Gen	Oligo:
	(SEQ ID NO.: 62)	(SEQ ID NO.: 848)
401	Q86SP6	TAGGCGTGAGAATCGG
	(SEQ ID NO.: 62)	(SEQ ID NO.: 849)
402	Q86SP6	TAGGTGTGAGAATTGG
	(SEQ ID NO.: 62)	(SEQ ID NO.: 850)
403	Q86SP6	ATCGTGTTTCGGATCGG
	(SEQ ID NO.: 62)	(SEQ ID NO.: 851)
404	Q86SP6	TATTGTGTTTGGATTGG
	(SEQ ID NO.: 62)	(SEQ ID NO.: 852)
405	Q86SP6	AGGTTGAGTTTGCGG
	(SEQ ID NO.: 62)	(SEQ ID NO.: 853)
406	Q86SP6	TAGGTTTGAGTTTGTGG
	(SEQ ID NO.: 62)	(SEQ ID NO.: 854)
407	Q86SP6	TATAAAGCGCGAGTGT
	(SEQ ID NO.: 62)	(SEQ ID NO.: 855)
408	Q86SP6	TATAAAGTGTGAGTGTG
	(SEQ ID NO.: 62)	(SEQ ID NO.: 856)
409	RARB	ATGTCGAGAACGCGAG
	(SEQ ID NO.: 8)	(SEQ ID NO.: 857)
410	RARB	GGATGTTGAGAATGTGA
	(SEQ ID NO.: 8)	(SEQ ID NO.: 858)
411	RARB	AGCGATTCGAGTAGGG
	(SEQ ID NO.: 8)	(SEQ ID NO.: 859)
412	RARB	AGTGATTTGAGTAGGG
	(SEQ ID NO.: 8)	(SEQ ID NO.: 860)
413	RARB	TATCGTCGGGGTAGGA
	(SEQ ID NO.: 8)	(SEQ ID NO.: 861)
414	RARB	TATTGTTGGGGTAGGA
	(SEQ ID NO.: 8)	(SEQ ID NO.: 862)
415	RARB	AACGTATTCGGAAGGT
	(SEQ ID NO.: 8)	(SEQ ID NO.: 863)
416	RARB	GAATGTATTTGGAAGGT
	(SEQ ID NO.: 8)	(SEQ ID NO.: 864)
417	PTGS2	AGCGTTTTTCGAGAGTT
	(SEQ ID NO.: 9)	(SEQ ID NO.: 865)

No:	Gen	Oligo:
418	PTGS2	GAGTGTTTTGAGAGTT
	(SEQ ID NO.: 9)	(SEQ ID NO.: 866)
419	PTGS2	TTACGTCGGGATAGAT
	(SEQ ID NO.: 9)	(SEQ ID NO.: 867)
420	PTGS2	GGAAGTTATGTTGGGA
	(SEQ ID NO.: 9)	(SEQ ID NO.: 868)
421	PTGS2	TATCGTTTTAGGCGTA
	(SEQ ID NO.: 9)	(SEQ ID NO.: 869)
422	PTGS2	TTTATTGTTTAGGTGTAT
	(SEQ ID NO.: 9)	(SEQ ID NO.: 870)
423	RASSF1	TACGGGTATTTTCGCGT
	(SEQ ID NO.: 10)	(SEQ ID NO.: 871)
424	RASSF1	ATATGGGTATTTTGTGT
	(SEQ ID NO.: 10)	(SEQ ID NO.: 872)
425	RASSF1	AGAGCGCGTTTAGTTT
	(SEQ ID NO.: 10)	(SEQ ID NO.: 873)
426	RASSF1	GAGAGTGTGTTTAGTTT
	(SEQ ID NO.: 10)	(SEQ ID NO.: 874)
427	ESR2	TAGGAACGTTTCGACGT
	(SEQ ID NO.: 11)	(SEQ ID NO.: 875)
428	ESR2	TTTAGGAATGTTTGATGT
	(SEQ ID NO.: 11)	(SEQ ID NO.: 876)
429	ESR2	ATGGCGTTTTTCGTAG
	(SEQ ID NO.: 11)	(SEQ ID NO.: 877)
430	ESR2	TAGATGGTGTTTTTGT
	(SEQ ID NO.: 11)	(SEQ ID NO.: 878)
431	ESR2	ATAAGCGATTTAACGAT
	(SEQ ID NO.: 11)	(SEQ ID NO.: 879)
432	ESR2	AGTGATTTAATGATAAGTT
	(SEQ ID NO.: 11)	(SEQ ID NO.: 880)
433	ESR2	ATTTTCGAGGATTACGT
	(SEQ ID NO.: 11)	(SEQ ID NO.: 881)
434	ESR2	ATATTTTGAGGATTATGTT
	(SEQ ID NO.: 11)	(SEQ ID NO.: 882)
435	DRG1	TAGTCGTAAAACGTAG

No:	Gen	Oligo:
	(SEQ ID NO.: 12)	(SEQ ID NO.: 883)
436	DRG1	AGTTGTAAAATGTAGATT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 884)
437	DRG1	ATCGTATTGGTTCGCGG
	(SEQ ID NO.: 12)	(SEQ ID NO.: 885)
438	DRG1	ATTGTATTGGTTTGTGG
	(SEQ ID NO.: 12)	(SEQ ID NO.: 886)
439	DRG1	TTTACGTTTTGCGATT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 887)
440	DRG1	AGTTTTATGTTTTGTGAT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 888)
441	DRG1	ATTTTACGCGGAATT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 889)
442	DRG1	GTGATTTTATGTGGAAT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 890)
443	DRG1	ATTCGAAGTATCGCGT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 891)
444	DRG1	ATTTGAAGTATTGTGTTT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 892)
445	DRG1	ATTCGAATTCGAGTA
	(SEQ ID NO.: 12)	(SEQ ID NO.: 893)
446	DRG1	TTTGAATTTGAGTAGAG
	(SEQ ID NO.: 12)	(SEQ ID NO.: 894)
447	DOCK10	TATATTGTCGGGGAGA
	(SEQ ID NO.: 16)	(SEQ ID NO.: 895)
448	DOCK10	ATATTGTTGGGGAGAG
	(SEQ ID NO.: 16)	(SEQ ID NO.: 896)
449	BTG4	GTGTTGCGTAGAGATA
	(SEQ ID NO.: 17)	(SEQ ID NO.: 897)
450	BTG4	GGTGTGTGTAGAGAT
	(SEQ ID NO.: 17)	(SEQ ID NO.: 898)
451	BTG4	TTGGTTTTTCGGAATAA
	(SEQ ID NO.: 17)	(SEQ ID NO.: 899)
452	BTG4	GGTTTTTTGGAATAAGAT
	(SEQ ID NO.: 17)	(SEQ ID NO.: 900)

No:	Gen	Oligo:
453	GPR7	ATTGAGGGCGTATAGA
	(SEQ ID NO.: 19)	(SEQ ID NO.: 901)
454	GPR7	TGAGGGTGTATAGATTT
	(SEQ ID NO.: 19)	(SEQ ID NO.: 902)
455	GPR7	GGAGATCGAAGTTTGT
	(SEQ ID NO.: 19)	(SEQ ID NO.: 903)
456	GPR7	GGGGAGATTGAAGTTT
	(SEQ ID NO.: 19)	(SEQ ID NO.: 904)
457	ISL1	AGATTTTGCAGAAAGATA
	(SEQ ID NO.: 21)	(SEQ ID NO.: 905)
458	ISL1	TTAGATTTTGTGAAAGATA
	(SEQ ID NO.: 21)	(SEQ ID NO.: 906)
459	ISL1	AAGATATCGAAATTAAGTT
	(SEQ ID NO.: 21)	(SEQ ID NO.: 907)
460	ISL1	AAGATATTGAAATTAAGTTT
	(SEQ ID NO.: 21)	(SEQ ID NO.: 908)
461	GPRK5	AGATTTTCGAGGGAGA
	(SEQ ID NO.: 22)	(SEQ ID NO.: 909)
462	GPRK5	AGATTTTTCGAGGGAGAT
	(SEQ ID NO.: 22)	(SEQ ID NO.: 910)
463	FBN2	AATTGGTCGTTAGTTTT
	(SEQ ID NO.: 28)	(SEQ ID NO.: 911)
464	FBN2	GTTAATTGGTTGTTAGTT
	(SEQ ID NO.: 28)	(SEQ ID NO.: 912)
465	FBN2	TTAGGGATCGGATTTG
	(SEQ ID NO.: 28)	(SEQ ID NO.: 913)
466	FBN2	ATTAGGGATTGGATTTG

	(SEQ ID NO.: 28)	(SEQ ID NO.: 914)
467	LIMK1	TTTAATTCGAAAGGGAA
	(SEQ ID NO.: 30)	(SEQ ID NO.: 915)
468	LIMK1	TTAATTTGAAAGGGAAAG
	(SEQ ID NO.: 30)	(SEQ ID NO.: 916)
469	PSD/Q9H469	AGGTTAGTCGAGAAGT
	(SEQ ID NO.: 31)	(SEQ ID NO.: 917)

470	PSD/Q9H469	AGGTTAGTTGAGAAGTA
	(SEQ ID NO.: 31)	(SEQ ID NO.: 918)
471	PSD/Q9H469	ATTGTTACGAAGTGGA
	(SEQ ID NO.: 31)	(SEQ ID NO.: 919)
472	PSD/Q9H469	TTGTTATGAAGTGGAAG
	(SEQ ID NO.: 31)	(SEQ ID NO.: 920)
473	Q96S01	TATTGATCGGTTGGAG
	(SEQ ID NO.: 34)	(SEQ ID NO.: 921)
474	Q96S01	TATTGATTGGTTGGAGG
	(SEQ ID NO.: 34)	(SEQ ID NO.: 922)
475	ABHD9	AGTTAGAGCGTATTATTT
	(SEQ ID NO.: 37)	(SEQ ID NO.: 923)
476	ABHD9	AATAGTTAGAGTGATTATT
	(SEQ ID NO.: 37)	(SEQ ID NO.: 924)
477	MLLT3	TTAGTGCGTCGGAGATA
	(SEQ ID NO.: 42)	(SEQ ID NO.: 925)
478	MLLT3	AGTGGTTGGAGATAGA
	(SEQ ID NO.: 42)	(SEQ ID NO.: 926)
479	SOLH	TATATTCGTGAGGGTA
	(SEQ ID NO.: 43)	(SEQ ID NO.: 927)
480	SOLH	TATATTTGTGAGGGTAG
	(SEQ ID NO.: 43)	(SEQ ID NO.: 928)
481	Q9NWV0	ATTTTACGAGAAGGTT
	(SEQ ID NO.: 46)	(SEQ ID NO.: 929)
482	Q9NWV0	GATTTTATGAGAAGGTT
	(SEQ ID NO.: 46)	(SEQ ID NO.: 930)
483	H2AFY2	ATGGGAATCGTGTTT
	(SEQ ID NO.: 48)	(SEQ ID NO.: 931)
484	H2AFY2	ATGGGAATTGTGTTT
	(SEQ ID NO.: 48)	(SEQ ID NO.: 932)
485	RHOC	AGGTTTACGAAAAGG
	(SEQ ID NO.: 49)	(SEQ ID NO.: 933)
486	RHOC	AAGGTTTATGAAAAGG
	(SEQ ID NO.: 49)	(SEQ ID NO.: 934)
487	KBTBD6	TAGAGTCGGGTTTGTGA
	(SEQ ID NO.: 51)	(SEQ ID NO.: 935)

5

10

15

20

25

30

35

40

45

50

55

488	KBTBD6	TAGAGTTGGGTTTTGTA
	(SEQ ID NO.: 51)	(SEQ ID NO.: 936)
489	TRPM4	ATGGGGGTCGAGATTT
	(SEQ ID NO.: 52)	(SEQ ID NO.: 937)
490	TRPM4	ATGGGGGTTGAGATTT
	(SEQ ID NO.: 52)	(SEQ ID NO.: 938)
491	TCEB3BP1	GGTAGTCGGTATTAGG
	(SEQ ID NO.: 53)	(SEQ ID NO.: 939)
492	TCEB3BP1	TGGTAGTTGGTATTAGG
	(SEQ ID NO.: 53)	(SEQ ID NO.: 940)
493	TCEB3BP1	TGTAGTCGAAGGTTAG
	(SEQ ID NO.: 53)	(SEQ ID NO.: 941)
494	TCEB3BP1	TGTAGTTGAAGGTTAGT
	(SEQ ID NO.: 53)	(SEQ ID NO.: 942)
495	Q8NCX8	GGAGTTTATTCGGTTTAT
	(SEQ ID NO.: 54)	(SEQ ID NO.: 943)
496	Q8NCX8	GGAGTTTATTTGGTTTATA
	(SEQ ID NO.: 54)	(SEQ ID NO.: 944)
497	WDFY3	AATTGTAGTCGTTTAGTT
	(SEQ ID NO.: 58)	(SEQ ID NO.: 945)
498	WDFY3	AAATTGTAGTTGTTTAGTT
	(SEQ ID NO.: 58)	(SEQ ID NO.: 946)
499	ZNF566	TATTTTTTCGGTAGAGAT
	(SEQ ID NO.: 59)	(SEQ ID NO.: 947)
500	ZNF566	TTTTTTTGGTAGAGATTG
	(SEQ ID NO.: 59)	(SEQ ID NO.: 948)
501	ZNF566	ATGGGTTGCGTTTATA
	(SEQ ID NO.: 59)	(SEQ ID NO.: 949)
502	ZNF566	TGGGTTGTGTTTATAAG
	(SEQ ID NO.: 59)	(SEQ ID NO.: 950)
503	SEQ ID NO: 61	TTTAGGGCGAGGTTAT
	(SEQ ID NO.: 61)	(SEQ ID NO.: 951)
504	SEQ ID NO: 61	TTTAGGGTGAGGTTATT
	(SEQ ID NO.: 61)	(SEQ ID NO.: 952)
505	SEQ ID NO: 61	TGTATATTCGTAGGGT
	(SEQ ID NO.: 61)	(SEQ ID NO.: 953)

506	SEQ ID NO: 61	TTGTATATTTTGTAGGGT
	(SEQ ID NO.: 61)	(SEQ ID NO.: 954)
507	PTGS2	ATTTGAGCGGTTTTGA
	(SEQ ID NO.: 9)	(SEQ ID NO.: 955)
508	PTGS2	AATTTGAGTGGTTTTGA
	(SEQ ID NO.: 9)	(SEQ ID NO.: 956)
509	RASSF1	AGTAAATCGGATTAGGA
	(SEQ ID NO.: 10)	(SEQ ID NO.: 957)
510	RASSF1	AGTAAATTGGATTAGGAG
	(SEQ ID NO.: 10)	(SEQ ID NO.: 958)
511	DRG1	TGTATGAACGTGTAGTT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 959)
512	DRG1	GTGTATGAATGTGTAGT
	(SEQ ID NO.: 12)	(SEQ ID NO.: 960)

TABLA 11: Genes y secuencias de acuerdo con la lista de secuencias.

Gen (HUGO o SPTREMBL ID o EST Gene ID)	Ref.-Sec	SEQ ID NO genómica:	Secuencia codificante metilada convertida SEQ ID NO:	Secuencia antisentido metilada convertida SEQ ID NO:	Secuencia codificante no metilada convertida SEQ ID NO:	Secuencia antisentido no metilada convertida SEQ ID NO:
CCND2	NM_001759	1	65	66	193	194
CDRN2A	NM_000077	2	67	68	195	196
CD44	NM_000610	3	69	70	197	198
EDNRB1	NM_000115	4	71	72	199	200
ELK1	NM_005229	5	73	74	201	202
FOS	NM_005252	6	75	76	203	204
GSTP1	NM_000852	7	77	78	205	206
RARB	NM_000965	8	79	80	207	208
PTGS2	NM_000963	9	81	82	209	210
RASSF1	NM_170715	10	83	84	211	212
ESR2	NM_001437	11	85	86	213	214
DRG1	NM_004147	12	87	88	215	216
CMYA3		13	89	90	217	218
ONECUT2	NM_004852	14	91	92	219	220
MX1	NM_002462	15	93	94	221	222
DOCK10		16	95	96	223	224
BTG4	NM_017589	17	97	98	225	226

Gen (HUGO o SPTREMBL ID o EST Gene ID)	Ref.-Sec	SEQ ID NO genómica:	Secuencia codificante metilada convertida SEQ ID NO:	Secuencia antisentido metilada convertida SEQ ID NO:	Secuencia codificante no metilada convertida SEQ ID NO:	Secuencia antisentido no metilada convertida SEQ ID NO:
DMRTC2	NM_033052	18	99	100	227	228
GPR7	NM_005285	19	101	102	229	230
Grasa	NM_005245	20	103	104	231	232
ISL1	NM_002202	21	105	106	233	234
GPRK5	NM_005308	22	107	108	235	236
SLC35F2	NM_017515	23	109	110	237	238
C14orf59	NM_174976	24	111	112	239	240
SNRPN	NM_003097	25	113	114	241	242
ARHGEF18	NM_015318	26	115	116	243	244
SNX8	NM_013321	27	117	118	245	246
FBN2	NM_001999	28	119	120	247	248
HOXB5	NM_002147	29	121	122	249	250
LIMK1	NM_002314	30	123	124	251	252
PSD; Q9H469	NM_002779; NM_024326	31	125	126	253	254
SLC38A1	NM_030674	32	127	128	255	256
HIST1H4J	NM_003495	33	129	130	257	258
Q96S01	No aplicable	34	131	132	259	260
Región genómica corriente abajo de FOXL2	NM_023067	35	133	134	261	262
ORC4L	NM_002552	36	135	136	263	264
ABHD9	NM_024794	37	137	138	265	266
CD37	NM_001774	38	139	140	267	268
GRN	NM_002087	39	141	142	269	270
EPAS1	NM_001430	40	143	144	271	272
NOTCH1	NM_017617	41	145	146	273	274
MLLT3	NM_004529	42	147	148	275	276
SOLH	NM_005632	43	149	150	277	278
ENSESTG00002636932	No aplicable	44	151	152	279	280
Q8N365	NM_144697	45	153	154	281	282
Q9NWV0	NM_017891	46	155	156	283	284
ENST00000339569	No aplicable	47	157	158	285	286

Gen (HUGO o SPTREMBL ID o EST Gene ID)	Ref.-Sec	SEQ ID NO genómica:	Secuencia codificante metilada convertida SEQ ID NO:	Secuencia antisentido metilada convertida SEQ ID NO:	Secuencia codificante no metilada convertida SEQ ID NO:	Secuencia antisentido no metilada convertida SEQ ID NO:
H2AFY2	NM_018649	48	159	160	287	288
RHOC	NM_005167	49	161	162	289	290
NR2E1	NM_003269	50	163	164	291	292
KBTBD6	NM_152903	51	165	166	293	294
TRPM4	NM_017636	52	167	168	295	296
TCEB3BP1	NM_020695	53	169	170	297	298
Q8NCX8	NM_178564	54	171	172	299	300
secuencia genómica	secuencia genómica	55	173	174	301	302
SNAPC2	NM_003083	56	175	176	303	304
PTPRN2	NM_002847	57	177	178	305	306
WDFY3	NM_014991	58	179	180	307	308
ZNF566	NM_032838	59	181	182	309	310
Q9NP73	NM_018466	60	183	184	311	312
	NM_173660	61	185	186	313	314
Q86SP6	No aplicable	62	187	188	315	316
Región reguladora HIST2H2BF	NM_003529	63	189	190	317	318
PMF1	NM_000711	64	191	192	319	320
PITX2	NM_000325	961	962	963	964	965

TABLA 12: Iniciadores y sondas para los ensayos de MSP-MethyLight™

	Directo	Inverso	Sonda
SEQ ID NO: 19	tgcgattttggcgaattc	aaaaaccgcgcgactacgaa	aactaaacgcctaccgactaaatccgcct
SEQ ID NO: 35	gagttttcggttcgga	cgcgaccgctaaactcg	cgaccaaaccgaaccggtacatcg
SEQ ID NO: 37	cggtatgtctgcggttc	ctaacaccgcttcgccg	cctaaccgacaacgcgcgtaat
SEQ ID NO:	agttgcgcggcgatttc	gccccaaactaaatcacgacg	cggtcgacgttcggggtgtagcg

	Directo	Inverso	Sonda
7			
SEQ ID NO: 63	atggcgtataggttcgtgtttc	ctccaacgactaatacgcgaa	cgcttccaaaactcgacgtaataacgc
SEQ ID NO: 8	atataaactaaaaacgaaacgataaacga	tttacgttttttatttcgggc	cgaacgaacgcaaacgaaacaccg
SEQ ID NO: 64	ccactaactcgtaccgtacgtat	ggtagcgagtcgatcggtt	acgttcgtctccgctaaattatccgc
SEQ ID NO: 43	tcgttttttagtcgtttgggtc	tatcgaaaccccgaaaccg	caccgtcgctcccacgaca
SEQ ID NO: 40	tttcgttttttcggtcggtt	gaccgcgcaaaaaactcg	cgctcgaataacgcggaacccg
SEQ ID NO: 32	ctcgcaaaaaacgaaaatacgc	agttcgcggttttaaacgtgtc	ccgacaccgacccacgcgt
SEQ ID NO: 37	cggatgtcgtcgcgtttc	ctaacaccgcttcgccg	cctaaccgacaacgcgcgtaat
SEQ ID NO: 9	aaacgaaaactctacccgaatacgc	cgcggttttgcggtt	ccgaaatcccgcatacgcgcgcg
SEQ ID NO: 19	tcgggattttggcgaattc	aaaaacccgcgcgactacgaa	aactaaaacgcctaccgactaatatccgct
SEQ ID NO: 34	gtatttatttggtatttcgtattataattcgag	gactaaaaacgcgaaatccga	acgatccgatctaaaaaccgactcttcgaa
SEQ ID NO: 35	igagttttcgcggttcgga	cgcgaccgcctaaactcg	cgaccaaatacgaacccgtacatcg

Tabla 13: Componentes para todos los ensayos QM de acuerdo con el Ejemplo 5

Componente	compañía	Conc. de solución de reserva
Amortiguador de reacción ROX	Eurogentec	10x
MgCl ₂	Eurogentec	50mM
dNTPs	MBI	25mM cada uno
Cebador Director	TIB Molbiol	6,25µM
Cebador Inverso	TIB Molbiol	6,25µM
Sonda cg	Eurogentec	4µM
Sonda. tg	Eurogentec	4µM
HotGoldStar-Taq	Eurogentec	5U/µl
Agua	Fluka	

Tabla 14: Condiciones de reacción optimizadas para todos los ensayos QM de acuerdo con el Ejemplo 5

Gen	dNTPs	Amortiguador	MgCl ₂	Iniciadores	Sondas	Taq	Valores de la línea de base	Umbral	Hibridación
PITX2	250µM	1x	3mM	625 nM	200 nm	1U	3/23	0,05	62 °C
Chr3-EST	200µM	1x	3.5mM	625 nM	200 nm	1U	6/22	0,05	60°C
ABHD9	200µM	1x	2.5mM	625 nM	200 nm	1U	6/25	0,08	60°C
GPR7	250µM	1x	3mM	625 nM	150 nM	1U	6/24	0,05	60°C
HIST 2H2BF	250µM	1x	3mM	625 nM	250	1U	3/22	0,05	60°C
OCND2	250µM	1x	3mM	625 nM	250 nM	1U	3/22	0,08	60°C

Tabla 15: Programa de ciclos para los ensayos de QM de acuerdo con el Ejemplo 5. Para las temperaturas de hibridación ver la Tabla 14.

	T [°C]	T	Ciclos
Desnat. Inicial	95.0	10min	
Desnaturalización.	95.0	15seg	45x (PITX2 SOx)
Hibridación	Variable	60seg	

Tabla 16: Características clínicas de la población de pacientes de acuerdo con el Ejemplo 5. La edad se expresa como la media, y todas las demás variables se expresan como el número de pacientes. No toda la información estuvo disponible para todos los pacientes.

Variable Clínica		Baylor	Stanford	VMMC	Total
Edad (medio)		61.1	61.7	61.1	61.3
PSA	0-4	25	33	18	76
	4-10	120	139	99	358
	>10	60	72	30	162
Puntuación de Gleason	5-6	137	164	118	419
	7	37	44	19	100
	8-10	26	31	25	82
Etapa	Confinado a órgano	110	211	113	434
	No confinado a órgano	94	33	35	162
Recurrencia basada en el PSA		22	10	13	45
Decisión para tratar basada en la recurrencia		3	14	4	21
Total de muestras		206	244	162	612

Tabla 17: Desempeño de los seis marcadores de acuerdo con el Ejemplo 5 mediante el uso de la mediana del nivel de metilación como valor de corte

	Valor de P	Eventos en grupo hipometilado	Eventos en grupo hipermetilado	AUC (5 años)
PITX2	0.000017	15	49	0.64
GPR7	0.0016	20	45	0.64
Región reguladora HIST2H2BF	0.018	22	43	0.60
SEQ ID NO: 35	0.0059	21	44	0.61
ABHD9	0.018	22	43	0.58
CCND2	0.22	27	38	0.61

Tabla 18: Resultados del análisis de regresión de Cox para PITX2 de acuerdo con el Ejemplo 5. Con el uso de la regresión por pasos el marcador se mantiene en el modelo. Los valores de p se refieren a la hipótesis nula "el índice de riesgos es igual a cero".

Variable	Valor P	Índice de Riesgos	Intervalo de confianza inferior	Intervalo de confianza superior
PITX2	0.0043	2.222	1.284	3.845
Etapas de la enfermedad	0.0692	1.713	0.965	3.061
Categoría de Gleason	0.0107	1.798	1.146	2.821
PSA	0.075	1.254	0.977	1.609
Categoría Nomograma	0.0866	2.187	0.894	5.353

Tabla 19: Resultados del análisis de regresión de Cox para la SEQ ID NO: 63 de acuerdo con el Ejemplo 5. El marcador permanece en el modelo.

Variable	Valor P	Índice de Riesgos	Intervalo de confianza inferior	Intervalo de confianza superior
SEQ ID NO: 63	0.0239	2.918	1.152	7.393
Etapas de la enfermedad	0.0599	1.735	0.977	3.081
Categoría de Gleason	0.0106	1.799	1.146	2.822
PSA	0.0732	1.25	0.979	1.596
Categoría Nomograma	0.0384	2.526	1.051	6.071

Tabla 20: Secuencias de iniciadores y sondas de los ensayos de acuerdo con el Ejemplo 5.

CCND2			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Cebador Director	Tttttgtaaagatagtttgatttaagtat (SEQ ID NO: 983)		
Cebador Inverso	caaacctttctccctaaaaacc (SEQ ID NO: 984)		
Sonda-CG	Cgcccgaacacgatcg (SEQ ID NO: 985)	FAM	BHQ1
Sonda-TG	Caccaccaacacaatcaaccctaacac (SEQ ID NO: 986)	Hex	BHQ1

SEQ ID NO: 63			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Cebador Director	tgattattatgtttaaggatatttagttg (SEQ ID NO: 987)		
Cebador Inverso	caataactctaaaaaaacctttaatc (SEQ ID NO: 988)		
Sonda-CG	Cgctccccgcgaatacgcg (SEQ ID NO: 989)	FAM	BHQ1
Sonda-TG	Taaaccactccccacaaatacaacaac (SEQ ID NO: 990)	Hex	BHQ1

GRP7			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Cebador Director	Catccctacacttccaaac (SEQ ID NO: 991)		
Cebador Inverso	Ggagttgttaggagaaaagt (SEQ ID NO: 992)		

GRP7			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Sonda-CG	Cgaacacccaaccgacaaacg (SEQ ID NO: 993)	FAM	BHQ1
Sonda-TG	Caaacacccaaccaacaacatctca (SEQ ID NO: 994)	Hex	BHQ1
Chr3_EST			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Cebador Director	ttgtagggttttttgggtt (SEQ ID NO: 995)		
Cebador Inverso	Ctcaaaacccttaaaacataaa (SEQ ID NO: 996)		
Sonda-CG	Ataaccacactacgcctcc (SEQ ID NO: 997)	FAM	BHQ1
Sonda-TG	Ataaccacactacacctcccaca (SEQ ID NO: 998)	Yakima amarillo	BHQ1
ABHD9			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Cebador Director	Ggtgttaggggttaggggtt (SEQ ID NO: 999)		
Cebador Inverso	Ccaaataattacctaactcaata (SEQ ID NO: 1000)		
Sonda-CG	Aactattttctatcgaaaccgccc (SEQ ID NO: 1001)	FAM	BHQ1
Sonda-TG	Aactattttctatcaaaaccacccctct (SEQ ID NO: 1002)	Yakima amarillo	BHQ1
PITX2			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Cebador Director	Gtaggggagggagtagatgtt (SEQ ID NO: 1003)		
Cebador Inverso	Ttctaactcctcttcacaataa (SEQ ID NO: 1004)		
Sonda-CG	Agtcggagtcgggagagcga (SEQ ID NO: 1005)	FAM	TAMRA
Sonda-TG	Agttgagtgaggagtgaaaggaga (SEQ ID NO: 1006)	VIC	TAMRA

GRP7			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Sonda-TG	ataaccacactacacacctcccaca	Yakima amarillo	BHQ1
ABHD9			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Cebador Director	ggtgttaggttaggggt		
Cebador Inverso	ccaaatatttacctaactcaata		
Sonda-CG	aactatttctatcgaaacgcgcg	FAM	BHQ1
Sonda-TG	aactatttctatcaaaaccaccacctct	Yakima amarillo	BHQ1
PITX2			
Iniciadores + sondas	Secuencia	Marcador 5'	Marcador 3'
Cebador Director	gtaggggagggagtagatgt		
Cebador Inverso	ttctaactcctctccacaataa		
Sonda-CG	agtcggagtcggagagcga	FAM	TAMRA
Sonda-TG	agttggagttggagagtgaaggaga	VIC	TAMRA

TABLA 21

Sec. con núm. de ident.	Clasificación	Tipo de tejido	AUC	Sensibilidad	Especificidad
19	Recurrencia	PET & congelado	0.62	0.35	0.86
63	Recurrencia	PET & congelado	0.68	0.34	0.85
35	Recurrencia	PET & congelado	0.6	0.27	0.85
37	Recurrencia	PET & congelado	0.61	0.18	0.85
19	Recurrencia	PET	0.58	0.5	0.88
63	Recurrencia	PET	0.63	0.17	0.89

Sec. con núm. de ident.	Clasificación	Tipo de tejido	AUC	Sensibilidad	Especificidad
35	Recurrencia	PET	0.69	0.33	0.89
37	Recurrencia	PET	0.69	0.17	0.89
19	Recurrencia	Congelado	0.62	0.34	0.86
63	Recurrencia	Congelado	0.68	0.31	0.86
35	Recurrencia	Congelado	0.6	0.28	0.85
37	Recurrencia	Congelado	0.6	0.2	0.85
19	Gleason	PET & congelado	0.72	0.49	0.86
63	Gleason	PET & congelado	0.73	0.39	0.86
35	Gleason	PET & congelado	0.62	0.21	0.86
37	Gleason	PET & congelado	0.76	0.48	0.86
19	Gleason	PET	0.72	0.36	0.89
63	Gleason	PET	0.7	0.38	0.86
35	Gleason	PET	0.56	0.12	0.86
37	Gleason	PET	0.77	0.5	0.86
19	Gleason	Congelado	0.74	0.56	0.87
63	Gleason	Congelado	0.76	0.58	0.86
35	Gleason	Congelado	0.65	0.28	0.87
37	Gleason	Congelado	0.77	0.47	0.87

Ejemplos

Resumen de la investigación

El objetivo de la presente investigación es desarrollar marcadores genéticos que puedan identificar a pacientes con cáncer de próstata que tienen tumores agresivos con potencial metastásico. Se decidió usar el análisis de metilación para identificar marcadores genéticos expresados diferencialmente.

La investigación comenzó con una etapa de tamizaje del genoma completo para descubrir nuevos marcadores. Este enfoque utiliza métodos de biología molecular para la determinación de la metilación diferencial entre grupos predefinidos de muestras de pacientes. Las secuencias metiladas diferencialmente identificadas mediante el uso de dichos métodos de tamizaje genómico se denominan en la presente como Etiquetas de Secuencias de Metilación (conocidas además como MeST). La etapa de tamizaje del genoma completo identifica sitios CpG metilados diferencialmente. Información adicional relacionada con el área que rodea al sitio CpG identificado se obtiene mediante análisis BLAST de la secuencia que se encuentra en la etapa de tamizaje (MeST o Etiquetas de las Secuencias de Metilación) y el mapeo al genoma humano.

Después de la identificación de candidatos mediante el tamizaje del genoma completo, se desarrollaron los ensayos de MSP MethyLight™ para un subconjunto de los candidatos prometedores. Estos ensayos se usaron para analizar la metilación en

56 muestras de prostatectomías con información de resultado clínico. Los ensayos de MSP MethyLight™ son sensibles y cuantitativos y se basan en la co-metilación de CpG; por lo tanto, este formato de ensayo proporciona datos complementarios a la tecnología de matriz de ADN.

5 Todas las candidatas a MeST prometedoras y algunas candidatas adicionales se analizaron después por una matriz de oligonucleótidos de metilación mediante el uso de la tecnología de chip propiedad del solicitante, como se describe en detalle más abajo. Este proceso proporciona información relativa al desempeño preliminar de las MeST o genes candidatos si se usa un conjunto adecuado de muestras. Los marcadores candidatos se seleccionarán en base a los datos obtenidos del análisis de datos del chip. La selección se basa principalmente en el AUC del marcador en la discriminación entre las clases deseadas.

Para completar el proceso de desarrollo, los marcadores candidatos seleccionados para el desarrollo de ensayos a partir del estudio de chips se probaron en ensayos de PCR en tiempo real para su posterior validación en la población objetivo y en el material de muestras usadas para una prueba de diagnóstico o pronóstico potencial (tejidos embebidos en parafina).

Ejemplo 1: Tamizaje de MeST

Diseño experimental

20 Se usó ADN genómico mezclado de muestras de cáncer de próstata para el tamizaje en el genoma completo de marcadores asociados con la agresividad tumoral. Se aplicaron dos métodos diferentes: Reacción en cadena de la polimerasa con el uso de manera arbitraria de iniciadores con metilación específica (MS-APPCR) (Liang y otros, 1998) y amplificación de islas de CpG metilados (MCA) (Toyota y otros, 1999). Estas tecnologías distinguen entre sitios CpG metilados y no metilados mediante el uso de enzimas sensibles a la metilación; en general, el ADN genómico se corta con una enzima de restricción sensible a la metilación. Los fragmentos metilados se amplifican preferentemente porque la escisión en sitios no metilados impide la amplificación de los objetivos no metilados. Los fragmentos de las etiquetas de secuencias metiladas (MeST) obtenidos mediante el uso de estas técnicas se secuencian y se mapean con el genoma humano mediante el uso de la utilidad BLAST en la base de datos Ensembl (www.ensembl.org).

30 La definición primaria de la agresividad tumoral para la fase de tamizaje se basó en la recurrencia del PSA después de una prostatectomía radical. Un tumor agresivo se definió como el que tuvo una recurrencia en menos de 24 meses. Un tumor no agresivo se definió como el que no recurrió después de al menos 48 meses de seguimiento con pruebas regulares del PSA. Se mezclaron cinco muestras en cada categoría, y había tres mezclas para cada categoría. La mediana del tiempo hasta la recurrencia del PSA para los pacientes con tumores agresivos fue de 5.1 meses. La mediana del tiempo de seguimiento para los pacientes sin recurrencia fue de 60.3 meses. Ninguno de estos pacientes recibió terapias neoadyuvantes o adyuvantes antes de la recaída del PSA.

Además se incluyeron otras cuatro comparaciones con indicadores alternativos de agresividad. Las muestras con grados de Gleason cuatro y cinco se compararon con muestras con grados de Gleason uno, dos y tres. Los tumores en etapa tardía (III y IV) se compararon con los tumores en etapa temprana (I y II). Los tumores de la zona periférica se compararon con los tumores de la zona de transición. Por último, los tejidos normales adyacentes a los tumores en pacientes con recurrencia temprana del PSA (<2 años) se compararon con los tejidos normales adyacentes a los tumores en pacientes sin recurrencia del PSA (>4 años de seguimiento).

Tamizaje

El proceso de tamizaje de MeST usualmente resulta en un gran número de secuencias que representan marcadores potenciales. Algunas de ellas son redundantes o no pueden aparearse con el genoma. Las secuencias restantes se seleccionan mediante el uso de un procedimiento de puntuación que evalúa:

- Aparición mediante el uso de múltiples métodos
- Aparición en múltiples comparaciones de mezclas del mismo tipo
- Lugar en la Isla de CpG.
- Lugar en la región promotora
- Lugar cerca o dentro de de la secuencia del gen.
- Asociación de gen cercano con cáncer
- Clase de gen (factor de transcripción, factor de crecimiento, etc.)
- Elemento Repetitivo (puntuación negativa)

En este sistema de puntuación, una secuencia MeST recibe un punto por cada uno de los criterios positivos (primeros siete criterios), y recibe una puntuación de menos 8 por tener un contenido de secuencia repetitiva superior al 50% (puntuación negativa). En este último caso la MeST siempre tiene un resultado global negativo. Las Tablas 2 y 3 resumen los resultados de los experimentos de tamizaje de MeST.

Con el uso de los criterios de puntuación y la investigación basada en la literatura de la función génica potencial, se seleccionaron 80 genes para el diseño de amplicones en chip. Las MeST de las comparaciones basadas en la recurrencia del PSA se priorizaron, pero muchas de las MeST de las otras comparaciones se incluyeron en la lista de diseño de amplicones.

Ejemplo 2: Estudio PCR en tiempo real

Diseño experimental

Los ensayos de MSP se desarrollaron en el Taqman™ 7900 para los candidatos fuertes a partir del tamizaje de MeST con el objetivo de obtener los primeros datos sobre el desempeño de las MeST como marcadores de la agresividad del cáncer de próstata.

Como se usa en la presente descripción el término MSP-MethyLight se tomará para significar un ensayo que comprende la amplificación de una secuencia tratada con bisulfito mediante iniciadores específicos de metilación y la detección de los amplificados resultantes mediante la detección con oligonucleótidos MethyLight (conocidos además como 'sondas').

Los ensayos de MSP-MethyLight se desarrollaron para un número de MeST (ver la Tabla 12). Los ensayos se probaron en ADN metilado artificialmente y en diluciones de ADN metilado en ADN no metilado para garantizar el desempeño del ensayo. Todos los ensayos fueron capaces de amplificar una cantidad tan pequeña como 100 picogramos de ADN metilado en presencia o ausencia de 20 a 100 nanogramos de ADN no metilado. La mayoría de los ensayos se cuantificaron entre 0.1 y 100% de metilación.

De los 36 ensayos, 21 estuvieron metilados en una mezcla de ADN de tumor de próstata. Estos 21 ensayos se probaron primero en 46 muestras del proceso de tamizaje. Estas 46 muestras incluyeron 14 prostatectomías de pacientes que presentaron recurrencias en menos de 24 meses y 18 prostatectomías de pacientes que no tuvieron recurrencias después de al menos 48 meses. Adicionalmente, existieron catorce pacientes sin información de seguimiento; nueve tenían alto grado de Gleason (Puntuación 8-10) y 5 tenían bajo grado de Gleason (Puntuación 2-6). Los datos de este experimento se usaron para seleccionar siete ensayos para un conjunto de muestras independientes.

El segundo conjunto de muestras consistió en 26 muestras congeladas de prostatectomías radicales a partir de pacientes con recurrencia temprana del PSA, con una mediana de tiempo hasta la recurrencia del PSA de 6 meses, y 30 muestras de pacientes sin recurrencia del PSA después de al menos 48 meses (la mediana del tiempo de seguimiento fue de 60 meses). Los ensayos MethyLight™ se usaron para medir la cantidad de ADN metilado en cada locus al comparar el ciclo umbral (Ct) con respecto a una curva estándar de ADN metilado. Un ensayo de control se usó para medir la cantidad total de ADN. Todas las muestras se corrieron por triplicado para todos los ensayos. Las secuencias de iniciadores y sondas se enumeran en la Tabla 12. La relación de ADN metilado con respecto a la cantidad total de ADN se usó para indicar el estado de metilación de cada candidato en cada muestra.

Resultados

Los valores de metilación para cada ensayo se usaron para construir las curvas ROC (Figuras 2 a 8) y calcular la sensibilidad, la especificidad y los valores de p. Los datos se resumen en la Tabla 4. Los valores del AUC para algunos de los candidatos sugieren que la metilación del marcador podría tener valor pronóstico. Seis candidatos tenían una AUC de 0.68 o mayor. Los candidatos más fuertes, SEQ ID NO: 19 (GPR7) y SEQ ID NO: 35 (región genómica corriente abajo de FOXL2), fueron significativos por una prueba de Wilcoxon después de la corrección de Bonferroni. Cuando la especificidad se establece en 87% para estos dos ensayos, la sensibilidad es alrededor de 50% para cada uno. (SEQ ID NOS correlacionadas con los nombres de los genes en la tabla 11).

Ejemplo 3: Estudio de chips

En el estudio de chips, un panel de genes compuesto de genes candidatos y MEST seleccionados se analizaron en 329 muestras mediante el uso de la tecnología de microarreglos del solicitante.

Conjunto de muestras

5

El conjunto de muestras incluyó 329 muestras congeladas obtenidas a partir de prostatectomías radicales. Sólo se usaron muestras con un por ciento tumoral estimado de al menos 70%, y la mediana del por ciento tumoral estimado (por volumen) fue del 90%. Algunos proveedores de muestras lograron alto por ciento tumoral ya sea por extracción de una sección de la próstata congelada de la que se sabe que contiene un tumor o por disección de tejido normal lejos del tumor. Los pacientes que recibieron terapia neoadyuvante no se excluyeron del estudio. La información clínica sobre la supervivencia libre de enfermedad no se usó para ningún paciente que recibió terapia adyuvante antes de la recurrencia de la enfermedad.

10

15

La puntuación de Gleason estuvo disponible para casi todas las prostatectomías. Para algunas muestras, estuvo disponible además la puntuación de Gleason de la porción del tumor proporcionada al solicitante. El conjunto de muestras consistió en muestras que se clasificaron en una de las dos categorías extremas de Gleason (alta o baja), muestras de pacientes que se clasificaron para las dos categorías de recaída (temprana o sin recaída), o muestras que cayeron en múltiples categorías (por ejemplo, alto grado de Gleason y recurrencia temprana). Dentro del conjunto de muestras, hubo 135 muestras con bajas puntuaciones de Gleason (1+2, 2+1, 2+2, 2+3, 3+2, y 3+3). Habían 99 muestras con altas puntuaciones de Gleason (3+5, 5+3, 4+4, 4+5, 5+4, y 5+5). Para algunas de las muestras, estaba disponible la información del seguimiento clínico. Sesenta y cinco pacientes experimentaron recurrencia del PSA en menos de dos años, y 88 pacientes no tuvieron recurrencias después de al menos 4 años de seguimiento.

20

Para propósitos de control se incluyeron muestras adicionales. Con el objetivo de controlar la calidad y la funcionalidad de los oligos, se usó ADN no metilado (ADN de phi-29) y ADN metilado artificialmente (Promega). Adicionalmente, 16 muestras de ADN de linfocitos se procesaron en paralelo a las muestras del ensayo.

25

Matriz

30

La matriz contenía los oligos que representan 62 candidatos diferentes. Cincuenta y uno de los candidatos eran MeST. Una MeST, SEQ ID NO:32, se representó por dos amplicones no solapados, ambos cerca del exón uno del gen. Para todas las MeST, el amplicón se diseñó lo más cercano posible a la secuencia de la MeST en una región rica en CpG. Se incluyó un amplicón para el gen ELK1 del cromosoma X para el análisis de las muestras masculinas y femeninas de linfocitos controles.

35

Otros genes analizados incluyeron CCND2 (Ciclina D2 Padar y otros 2003), CD44 (Woodson y otros 2004), EDNRB1 (receptor B de endotelina; Woodson y otros 2004; Nelson y otros 1997), GSTP1 (glutación S-transferasa pi; Maruyama y otros 2002), RARB (receptor del ácido retinoico, beta; Singal y otros 2004), PTGS2 (prostaglandina-endoperoxido sintasa 2; Yegnasubramanian y otros 2004), RASSF1 (familia 1 del dominio de asociación a Ras; Liu y otros 2002), ESR2 (receptor 2 de estrógeno; Zhu y otros 2004), DRG1 (proteína 1 de unión a GTP regulada por el desarrollo; Bandyopadhyay y otros 2003), y CDKN2A (p16; Halvorsen y otros 2000). DRG1 se representó con dos amplicones. En todos los casos, las áreas ricas en CpG cercanas al promotor o exón 1 fueron objetivos. Para p16, se usó el área rica en CpG que abarca el exón 2 porque en la literatura se han observado tasas de metilación más altas (Nguyen y otros 2000).

40

Un resumen completo de todos los genes analizados puede encontrarse en la Tabla 11.

45

Métodos estadísticos: Análisis de los Datos de Chip

Desde intensidades de hibridación en bruto a índices de metilación

50

El logaritmo del índice de metilación ($\log(\text{CG}/\text{TG})$) en cada posición CpG se determina de acuerdo con un enfoque de procesamiento previo estandarizado que incluye las siguientes etapas:

- Para cada punto la mediana de la intensidad de los píxeles del fondo se resta de la mediana de la intensidad de los píxeles del objeto. Esto da un buen valor estimado de las intensidades de hibridación con corrección del fondo;
- Tanto para los oligonucleótidos de detección de GC como TG de cada posición de CpG, se toma la mediana con corrección de fondo de las 4 intensidades de puntos redundantes;
- Para cada chip y cada par de oligos de CG/TG, se calcula el índice $\log(\text{CGfTG})$; y

55

- Para cada muestra, se toma la mediana de las intensidades del $\log(CG/TG)$ sobre las repeticiones redundantes del chip. Este índice $\log$ tiene la propiedad de que el ruido de hibridación tiene una varianza aproximadamente constante en todo el intervalo de posibles tasas de metilación (Huber y otros, 2002).

5 Análisis de componentes principales

El análisis de componentes principales (PCA) proyecta vectores de medición (por ejemplo, datos de chips, perfiles de metilación en varios sitios CpG, etc.) en un nuevo sistema de coordenadas. Los nuevos ejes de coordenadas se conocen como componentes principales. El primer componente principal se extiende en la dirección de la mayor varianza de los datos. Los componentes posteriores se ordenan por la disminución de la varianza y son ortogonales y no correlacionados entre sí. Diferentes posiciones de CpG contribuyen con diferentes pesos a la extensión de la nube de datos a lo largo de los diferentes componentes. El PCA es una técnica no supervisada, es decir no tiene en cuenta ninguna información de grupo o marcador de los puntos de datos (para más detalles, ver, por ejemplo, Ripley, 1996).

El PCA se usa típicamente para proyectar datos de gran número de dimensiones (en nuestro caso los datos de matriz de metilación) sobre subespacios de menores dimensiones con el objetivo de visualizar o extraer características con gran varianza a partir de los datos. En el presente informe usamos 2 proyecciones dimensionales para el control estadístico de la calidad de los datos. Investigamos el efecto de diferentes parámetros del proceso sobre los datos de los chips y se excluyó que los parámetros cambiantes del proceso causaran grandes alteraciones en los valores de medición.

Una versión robusta del PCA se usó para detectar chips extremos individuales y excluirlos del análisis adicional (Model y otros, 2002).

Gráficos de control T^2

Para controlar la estabilidad general del proceso de producción de chips usamos métodos del campo de control de procesos con estadística multivariable (MVSPC). Nuestra herramienta principal es el gráfico de control T^2 , que se usa para detectar desviaciones significativas del proceso de chips a partir de condiciones de trabajo normales (Model y otros, 2002).

El gráfico T^2 se construye de la siguiente manera:

1. Ordenar los datos del chip con respecto a un parámetro del proceso (por ejemplo fecha de la hibridación o robot de imprimir);
2. Definir un conjunto de datos históricos, que describe el proceso del chip bajo condiciones de trabajo normales (por ejemplo los primeros 75 chips hibridados). En el gráfico, los datos del conjunto de datos históricos se indican mediante un símbolo especial del gráfico; y
3. Calcular la distancia de cada nuevo chip con respecto al conjunto de datos históricos. Si la distancia de varios chips consecutivos excede un límite control dado el proceso tiene que considerarse como fuera de control.

El uso de gráficos T^2 para monitorear el proceso de producción de chips nos permite detectar y eliminar eficazmente la mayoría de las fuentes de errores sistemáticos.

40 Pruebas de Hipótesis

Nuestra tarea principal es identificar marcadores que puedan hacer una contribución significativa a la predicción de la clase de las muestras. Una contribución significativa se detecta cuando la hipótesis nula, que plantea que un modelo de predicción que incluye el marcador no mejora el desempeño de la clasificación con respecto a un modelo sin el marcador, puede rechazarse con $p < 0.05$. Debido a que aplicamos esta prueba a todo un conjunto de marcadores potenciales, tenemos que corregir los valores de p para múltiples pruebas. Hacemos esto al aplicar la corrección conservativa de Bonferroni, que simplemente multiplica los valores de p de los marcadores individuales con el número de posibles marcadores probados. Además damos resultados con el método menos conservativo de la Tasa de Tamizaje Falso (FDR) (Dudoit y otros 2002).

A lo largo de este informe un marcador (a veces referido además simplemente como gen o amplicón) es una región genómica de interés (ROI). Usualmente consiste en varias posiciones de CpG en el área respectiva. Para probar la hipótesis nula de que un marcador no tiene poder predictivo utilizamos las pruebas de suma de rangos de Wilcoxon para comparar los grupos. Un resultado significativo de la prueba ($p < 0.05$) indica un cambio entre las distribuciones de los respectivos $\log$ del índice de metilación, es decir $\ln(CG/IG)$. La media de todos los oligos para cada marcador se usó para combinar los CpG antes de que se generaran las estadísticas de Wilcoxon. Este enfoque tiene la ventaja de que favorece los marcadores que muestran co-metilación.

Un valor de p significativo para un marcador significa que la metilación de esta ROI tiene cierta correlación sistemática a la pregunta de interés como se da por las dos clases. Generalmente un valor de p significativo mediante el uso de la prueba de suma de rangos de Wilcoxon implica además un buen desempeño de la clasificación.

Predicción de clases mediante el análisis ROC

- 5 Se usó el análisis de Característica Operativa del Receptor (ROC) para estimar qué tan bien el conjunto de CpG de un marcador seleccionado puede diferenciar entre diferentes clases de tejidos. Una curva ROC es un gráfico de la tasa de verdaderos positivos (sensibilidad) contra la tasa de falsos positivos (1 - especificidad) para un marcador de todos los posibles umbrales de prueba. El Área Bajo la Curva (AUC) de una curva ROC da la probabilidad de clasificar correctamente una muestra aleatoria y por lo tanto puede usarse para evaluar el desempeño global del marcador. El AUC se relaciona con la prueba estadística de Wilcoxon y los intervalos comparativos de los marcadores mediante las AUC o los valores de p de Wilcoxon son equivalentes. La media de todos los oligos para cada marcador se usó para combinar los CpG antes del análisis de ROC. Este enfoque tiene la ventaja de que favorece los marcadores que muestran co-metilación.

Desempeño Experimental

15 Extracción de ADN

- Las muestras se recibieron de colaboradores externos ya sea como tejidos congelados o como ADN genómico extraído. El ADN de las muestras de tejidos se aisló en Epigenomics Berlín mediante el uso del kit DNA Mini Kit de Qiagen.
- 20 La calidad del ADN de todas las muestras suministradas y extraídas se evaluó primero mediante mediciones fotométricas. Se determinaron los coeficientes de extinción a 260 nm y 280 nm así como los índices A260/280 y se calcularon las concentraciones resultantes. Para la mayoría de los ADN usados, se determinaron los índices de A260/280 entre 1.6 y 1.9 lo que indica una pureza suficiente. Para algunas muestras se calcularon índices en el intervalo de 1.2-1.5. Sin embargo, estos ADN también se procesaron.
- 25 Después de las mediciones fotométricas se aplicaron 200 ng de los ADN genómicos a un gel de agarosa al 0.8% y se realizó una electroforesis en gel. La Figura 8 muestra una imagen de un gel típico. Se observaron sólo signos menores o ningún signo de degradación, lo que indica una buena calidad general de los ADN usados.

Tratamiento con bisulfito y PCR Múltiple

- 30 Los ADN genómicos totales a partir de las muestras seleccionadas así como los ADN controles se trataron con bisulfito que convierte las citosinas no metiladas en uracilo. Las citosinas metiladas se conservan. El tratamiento con bisulfito se realizó mediante el uso del proceso de tratamiento con bisulfito de dioxano de Epigenomics. Con el objetivo de evitar un procesamiento conjunto de todas las muestras con el mismo fondo biológico que resulte en un potencial sesgo del proceso en los datos más tarde, las muestras se agruparon aleatoriamente en lotes de procesamiento. Los lotes de 50 muestras se aleatorizaron en cuanto a la puntuación de Gleason y al resultado del PSA. Se realizaron dos reacciones independientes de bisulfito por muestra de ADN. Después del tratamiento con bisulfito, se usaron 11.25 ng de cada muestra en 8 reacciones de PCR múltiple (mPCR) posteriores que contenían 8 pares de iniciadores cada una.
- 35 Para el seguimiento de los resultados de mPCR, se realizó una electroforesis en gel para todos los productos de PCR. Para encontrar la mejor composición de ocho pares de iniciadores en un conjunto de mPCR, se usó un análisis de ALF, que compara una mezcla de productos de PCR individuales con diferentes variantes de conjuntos de mPCR.
- 40

Análisis de Expresión de ALF:

- 45 Para la evaluación de los fragmentos amplificados, los productos de mPCR de ADN de Promega se analizaron mediante el uso de la tecnología ALF Express. Los resultados para esos mPCR se compararon con la mezcla de productos de PCR individuales. La Figura 9 ilustra el resultado para una PCR 8-plex. Los 64 fragmentos (ocho PCR 8-plex) seleccionados para el estudio pudieron amplificarse en los experimentos de mPCR realizados. En algunos casos se obtuvieron productos secundarios no deseados.

50 Electroforesis en gel de agarosa:

- Como se mencionó anteriormente se realizaron dos reacciones independientes de bisulfito y PCR por muestra de ADN y los productos de PCR obtenidos se aplicaron a un gel de agarosa al 2%. En la Figura 10 se muestra una imagen de un gel típico que ilustra el desempeño de mPCR para 10 muestras. Ningún producto visible de la PCR implica el fracaso del tratamiento con bisulfito o la amplificación por PCR. Después la PCR se repitió con el doble de la cantidad de ADN (22.5 ng). Los ADN tratados con bisulfito que fallaron nuevamente se excluyeron del estudio. Si obtuvimos sólo una sonda de hibridación (64 productos de PCR mezclados) a partir de una muestra, se hibridaron 4 chips mediante el uso de esta única

sonda. Si las sondas a partir de dos tratamientos independientes de bisulfito de una muestra se amplificaron con éxito, ambas sondas se hibridaron a dos chips cada una.

Cuatro (4) de un total de 331 muestras procesadas (que incluyeron muestras controles) no pudieron amplificarse, a pesar de varios intentos. Estas muestras no se procesaron más aún.

Resultados del Estudio de Chips

El análisis primario fue una comparación de las muestras con altas puntuaciones de Gleason (8-10) y bajas puntuaciones de Gleason (2-6 sin el componente de los grados 4 o 5). Estas clases son las categorías A y C de la Tabla 5. Para nuestra segunda comparación, usamos un grupo de muestras de pacientes con recurrencia temprana del PSA después de la cirugía (<2 años) y un grupo sin recurrencia (>4 años de seguimiento). Estas clases son A1, B1, C1, y D1 para el grupo sin recurrencia y A2, B2, y C2 para el grupo con recurrencia (ver Tabla 5). Estos dos conjuntos de muestras (Gleason y resultado clínico) se superpusieron un poco. Nuestra tercera comparación analizó sólo los pacientes con grado intermedio de Gleason (3+4, 4+3, 2+5, 5+2, 2+4, 4+2), para determinar si la metilación de nuestras secuencias candidatas se correlaciona con la recurrencia temprana en estos pacientes. Por lo tanto, sólo se incluyeron las categorías B1 y B2.

Tejido tumoral vs. Linfocitos

Con el objetivo de evaluar el valor diagnóstico del chip, dieciséis muestras de linfocitos se incluyeron en el estudio. Los tejidos de cáncer de próstata y los linfocitos se compararon mediante pruebas estadísticas de suma de rangos de Wilcoxon. Una vista clasificada para los diez mejores amplificados se muestra en la Figura 12. Mientras que el grupo de linfocitos es algo homogéneo, la figura muestra mayor variabilidad para las muestras de cáncer de próstata. Las diferencias entre grupos son significativas en el nivel 0.05 (después de 5% corrección de tasa de falso descubrimiento) por los amplificados de CDRN2A, ELK1, GSTP1, RARB, PTGS2, RASSF1, ESR2, ONECUT2, BTG4, SLC35F2, HOXB5, LIMK1, HIST1H4J, SEQ ID NO: 35, EPAS1, NOTCH1, SEQ ID NO: 55, PTPRN2, Q9NP73, MX1, DOCK10, CCND2, ISL1, SNAPC2, GRN, H2AFY2, WDFY3, FOS, FAT, Q86SP6, SLC38A1, SNRPN, GPRK5, FBN2, ARHGEF18, RHOC, KBTBD6, NR2E1, PSD, DRG1, Q8N365, SEQ ID NO: 44, Q96S01, CD37, CMYA3, SEQ ID NO: 61, Q8NCX8 y ZNF566

Marcadores candidatos para Gleason

Comparación del grado de Gleason alto vs grado de Gleason bajo

Se usaron las pruebas estadísticas de rangos de Wilcoxon para analizar las diferencias en los perfiles de metilación de los pacientes clasificados como de alto grado de Gleason (Puntuación 8-10) y bajo grado de Gleason (Puntuación 2-6, sin componentes de grado 4 o 5). La clase de alto grado de Gleason consiste de 98 muestras y la clase de bajo grado de Gleason consiste de 135 muestras. La Figura 12 muestra los resultados de este análisis.

Para 25 amplificados, el valor de p corregido por Bonferroni de la prueba de Wilcoxon está por debajo de 0.05. Para una discusión de la relevancia biológica ver la sección más abajo. La Figura 13 muestra la matriz de metilación de los 10 mejores marcadores.

Las AUC/sensibilidad/especificidad de los amplificados de marcadores candidatos se muestran en la Tabla 6.

La Figura 13 muestra la matriz de metilación de alto grado de Gleason vs bajo grado de Gleason de los 10 marcadores con mejor AUC. Cada columna representa una muestra; cada fila un oligonucleótido (1, 2, o 3 sitios CpG cada uno). Los oligonucleótidos están agrupados por candidato a marcador. Los marcadores indicados están ordenados de arriba a abajo con el aumento del AUC. En el lado derecho de cada marcador se indica el valor de p de la prueba de Wilcoxon corregido por Bonferroni y el AUC. Debajo se indica la sensibilidad del AUC a una especificidad de ~ 0.75 encerrada entre corchetes. Los datos de metilación están centrados y normalizados a una desviación estándar para los oligonucleótidos individuales. El color representa la distancia relativa del estado de metilación del oligonucleótido a partir del valor medio. El gris claro representa los CpG hipometilados dentro de un oligonucleótido mientras que el gris oscuro indica los CpG hipermetilados dentro de un oligonucleótido.

Marcadores candidatos para la recurrencia del PSA

Comparación de la Recurrencia temprana vs. No Recurrencia

A continuación analizamos las diferencias en los perfiles de metilación de los pacientes clasificados como recurrencia temprana (recaída del PSA en menos de 24 meses) y sin recurrencia (sin recaída del PSA después de al menos 4 años).

Para tres (3) amplificados el valor de p corregido por Bonferroni de la prueba del índice de probabilidad (LR) está por debajo de 0.05. Para una discusión de la relevancia biológica ver la sección más abajo.

5 Las AUC/sensibilidad/especificidad de los amplificados de los marcadores superiores se muestran en la Tabla 7.

10 La Figura 15 muestra la matriz de metilación de la Recurrencia temprana vs No recurrencia de los 10 marcadores con mejor AUC. Cada columna representa una muestra; cada fila un oligonucleótido (1, 2, o 3 sitios CpG cada uno). Los oligonucleótidos están agrupados por candidato a marcador. Los marcadores indicados están ordenados de arriba a abajo con el aumento del AUC. En el lado derecho de cada marcador se indica el valor de p de la prueba de Wilcoxon corregido por Bonferroni y el AUC. Más abajo se indica la sensibilidad del AUC a una especificidad de - 0.75 encerrada entre corchetes. Los datos de metilación están centrados y normalizados a una desviación estándar para los oligonucleótidos individuales. El color representa la distancia relativa del estado de metilación del oligonucleótido a partir del valor medio. El gris claro representa los CpG hipometilados dentro de un oligonucleótido mientras que el gris oscuro indica los CpG hipermetilados dentro de un oligonucleótido.

Marcadores candidatos para la recurrencia del PSA en pacientes con puntuaciones de Gleason intermedias

20 Comparación de la Recurrencia temprana vs. No Recurrencia

Por último, analizamos las diferencias en los perfiles de metilación de muestras con grado de Gleason intermedio de pacientes clasificados como recurrencia temprana (recaída del PSA en menos de 24 meses) y sin recurrencia (sin recaída del PSA después de al menos 4 años). El grado de Gleason intermedio incluyó a todos los pacientes con puntuaciones de 2+5, 5+2, 3+4, 4+3, 2+4, 4+2, 1+5, 5+1, y estos pacientes son un subconjunto del grupo usado en la sección 6.6.2. La mayoría fue grado de Gleason 3+4 o 4+3. Aunque ningún amplificado mostró un valor de p corregido por Bonferroni por debajo de 0.05, varios marcadores mostraron AUC prometedoras (ver la Tabla 8). Es probable que esta comparación tuviera baja potencia debido al pequeño conjunto de muestras para esta comparación

30 Para una discusión de la relevancia biológica ver la sección más abajo.

Cometilación revelada por análisis de microarreglos

35 Debido al diseño del actual estudio de chips, fuimos capaces de determinar áreas dentro de fragmentos de marcadores que estaban co-metilados. En este diseño, al menos dos pares de oligos, que contienen cada uno 1, 2 o 3 sitios CpG, se incluyeron para cada uno de los fragmentos de marcadores analizados. Los detalles de los sitios CpG dirigidos en el análisis de matriz pueden encontrarse en las figuras 16 en adelante. La evidencia de co-metilación es evidente en las figuras de matriz clasificada. En las figuras de matriz clasificada a partir del análisis de micromatrices cada fragmento de marcador se agrupa horizontalmente. Dado que cada fragmento de marcador representa de uno a tres amplicones y un mínimo de cuatro y hasta treinta sitios CpG individuales, puede determinarse una amplia información relacionada con el estado de metilación del fragmento. Las cajas grises oscuras consecutivas dentro de la agrupación de un fragmento en una dirección vertical indican co-metilación de los oligonucleótidos (y CpG dentro de ese oligo). Estos datos se analizarán más aún para las áreas más discriminatorias dentro de un fragmento y esta información se utilizará para el diseño del ensayo de PCR en tiempo real.

45 Resumen de Resultados

En el análisis primario, una comparación de muestras con altos y bajos grados de Gleason, 25 marcadores cumplen los criterios de significación estadística mediante el uso de una metodología estadística muy conservativa. Esta comparación se basa en el grado de Gleason como un indicador sustituto de la agresividad, pero se usó como el análisis primario porque la información del grado de Gleason estuvo disponible para casi todos los tumores. Para el análisis adicional estuvieron disponibles menos muestras, en base al tiempo hasta la recaída del PSA, pero aún así dos marcadores alcanzaron la significación estadística.

50 Discusión

Aspectos biológicos

Tamizaje de MeST

El proceso de tamizaje de MeST fue muy exitoso, al producir más de 400 candidatos. En las PCR en tiempo real y los estudios de chips, los candidatos a MeST se comportaron bien. En el estudio de PCR en tiempo real, tres MeST superaron a GSTP1. En el estudio de chips, los cinco candidatos superiores en la comparación del grado de Gleason todos son MeST. Por lo tanto, el proceso de tamizaje contribuyó con valiosos marcadores candidatos para distinguir los tumores agresivos y los no agresivos.

Además, las MeST de todas las comparaciones de tamizaje estuvieron representadas en la lista de los principales candidatos para puntuación. El principal candidato a marcador en la comparación del grado de Gleason, GPR7, se descubrió en dos comparaciones de resultados, la comparación del grado de Gleason y la comparación basada en la etapa. Otro participante principal, DOCK10, se descubrió en la comparación basada en la zona prostática. De los marcadores principales, sólo ABHD9 se descubrió en la comparación de tejido normal adyacente a tumores a partir de pacientes en las dos categorías de resultados. Concluimos que todas las comparaciones de tamizaje del genoma completo produjeron marcadores candidatos importantes.

Evaluación de candidatos por MethyLight

Muchas MeST candidatas se escogieron para el desarrollo de ensayos de PCR en tiempo real mientras se recolectaban las muestras para el estudio de chips. Los ensayos se tamizaron previamente en una muestra mezclada de ADN a partir de muchas muestras de tumores de próstata. Esta etapa permitió seleccionar previamente sólo aquellos ensayos que podrían ser potencialmente informativos. Más de un tercio de los ensayos se eliminó en esta etapa. A continuación, los ensayos restantes se probaron en el ADN a partir de las muestras del tamizaje con el objetivo de priorizar los ensayos. Después, en el último conjunto de 56 muestras independientes, seis de los siete ensayos priorizados tuvieron un buen desempeño.

El éxito del experimento de PCR en tiempo real sugiere que existe una co-metilación significativa en estos marcadores. Por lo tanto, los ensayos de PCR en tiempo real, todos los cuales requieren cierto grado de co-metilación, serán candidatos adecuados para una opción de ensayo final. El experimento de PCR en tiempo real se basó en gran medida en la cuantificación de las diferencias de metilación: Para casi todos los ensayos, la diferencia entre el grupo de recurrencia temprana y el grupo sin recurrencia fue una diferencia cuantitativa de metilación. El tipo de ensayo final necesitará fuertes capacidades cuantitativas.

Evaluación de candidatos por matriz de metilación

El experimento de chips fue un gran éxito, lo que demuestra el potencial como marcadores para muchas secuencias candidatas. En la comparación de muestras con grados de Gleason altos y bajos, 25 amplicones fueron significativamente diferentes. La gran mayoría de estos están hipermetilados en muestras con alto grado de Gleason. Los tres candidatos analizados por PCR en tiempo real se encontraban entre los seis marcadores principales en esta comparación del grado de Gleason. Por lo tanto hay consistencia entre los dos métodos de medición de la metilación.

La comparación basada en las características de recaída del PSA de los pacientes tuvo menores números de muestras. A pesar de estos números pequeños, todavía pudimos demostrar que al menos tres (3) de nuestros marcadores candidatos pueden distinguir de manera significativa los pacientes que experimentan una recaída temprana de los pacientes que no experimentaron recaída. Generalmente, la metilación es mayor en los pacientes que experimentan recurrencia temprana. En esta comparación, los tres candidatos principales en el estudio de PCR en tiempo real fueron los tres marcadores principales más significativos en el análisis de chips. Los valores de AUC para GPR7, SEQ ID NO: 35 (corriente abajo de FOXL2), y ABHD9 fueron 0.72, 0.72, y 0.66 respectivamente en los datos del chip de resultados clínicos y 0.76, 0.75, y 0.70 respectivamente en los datos de PCR en tiempo real para los resultados clínicos.

El tratamiento para pacientes con grados de Gleason altos y bajos frecuentemente es claro. A cualquier persona con alto grado de Gleason se le recomendará un tratamiento agresivo, que incluye un tratamiento definitivo (cirugía o radiación) y posiblemente terapia adyuvante. Los pacientes con bajo grado de Gleason tienen la opción de diferir el tratamiento definitivo. Aunque todavía hay algunas incertidumbres para estos pacientes, las mejores opciones son aún menos claras para los pacientes con niveles de Gleason intermedios. Además, la mayoría de los pacientes diagnosticados con cáncer de próstata hoy en día tienen puntuaciones de Gleason intermedias de 6 o 7. Estos son los pacientes que pueden recibir más ayuda mediante una prueba de clasificación molecular. Los amplicones con las mayores AUC en la comparación basada en el resultado clínico fueron GPR7 (AUC=0.72) y la SEQ ID NO: 35 (AUC=0.72). Cuando esta comparación se restringió a pacientes con puntuaciones de Gleason intermedias (1+5, 5+1, 2+4, 4+2, 3+4, 4+3, 2+5, 5+2), el AUC para estos dos

marcadores todavía fue de 0.72 o mayor. Estos resultados sugieren que un ensayo basado en la metilación proporcionará información incluso para pacientes con puntuaciones de Gleason en el intervalo medio.

Biología de genes marcadores

Varios marcadores interesantes se identificaron por los estudios de PCR en tiempo real y de chips. Uno de los marcadores por PCR en tiempo real es una proteína G acoplada a un receptor (GPR7; SEQ ID NO:19), sin embargo se sabe muy poco sobre el producto del gen. Un segundo marcador a partir del estudio en tiempo real se localiza en una isla de CpG en el promotor del gen que contiene dominio de Abidrolasa 9 (ABHD9; SEQ ID NO:37). El gen más cercano a la SEQ ID NO:35 es FOXL2. La MeST está en una isla de CpG a varias kilobases corriente abajo de este gen. La SEQ ID NO:63 está en un área con varios genes de histonas.

En el estudio de chips surgieron marcadores adicionales. NOTCH1 (SEQ ID NO:41) controla una vía de señalización que regula las interacciones entre células adyacentes. Muchos laboratorios han estudiado el papel de este gen en la carcinogénesis y las metástasis. Poco se sabe sobre muchos de los candidatos, que incluyen DOCK10 (SEQ ID NO:16), SEQ ID NO:51, que está en el promotor de un gen denominado repetición de Kelch y que contiene el dominio BTB (POZ) 6, y la SEQ ID NO :17, que se localiza entre una EST y el gen 4 de Translocación de células B. Se ha demostrado que BTG4 tiene propiedades inhibitoras del crecimiento (Buanne y otros 2000).

PTGS2 (SEQ ID NO:9) es el único gen que se encuentra previamente en la literatura por estar más metilado en tumores de próstata de pacientes que presentaron recurrencia temprana después de la prostatectomía (Yegnasubramanian y otros 2004). PTGS2, conocido además como ciclooxigenasa (COX2), es un candidato prometedor adicional tanto en los análisis del grado de Gleason como del resultado clínico, con AUC de 0.69 y 0.65 respectivamente. GSTP1 es el marcador de metilación más altamente estudiado en el cáncer de próstata, y si bien no hay datos publicados que demuestran directamente su valor pronóstico, existe alguna evidencia de que su metilación se correlaciona con el grado de Gleason (Maruyama y otros 2002). Sin embargo, esta correlación no se confirmó en otro estudio (Woodson y otros 2004). En los datos instantáneos, la metilación de GSTP1 se correlaciona significativamente con el grado de Gleason, pero el AUC en la comparación del resultado clínico es de sólo 0.58.

Aspectos Médicos

Los candidatos de metilación que han surgido a partir de nuestra comparación del grado de Gleason son marcadores pronósticos informativos. Hemos demostrado que algunos de estos candidatos que se correlacionan con las categorías del grado de Gleason además pueden predecir la recaída del PSA, incluso en pacientes con puntuaciones intermedias del grado de Gleason. Por lo tanto es probable que nuestro análisis basado en el grado de Gleason proporcione marcadores que brinden información adicional al grado de Gleason.

Como marcadores individuales, los candidatos de los chips alcanzan 40-60% de sensibilidad cuando la especificidad se establece en 75%. En el estudio en tiempo real, la sensibilidad de tres de los marcadores fue mayor, al alcanzar 50-60% a una especificidad de 85%. El mejor desempeño en el tiempo real podría deberse a las capacidades cuantitativas del MSP-MethyLight. Aproximadamente el 20% de los pacientes experimentan recaídas dentro de 5-10 años después de la cirugía. Si esto se estableció como la prevalencia de tumores agresivos en la población con prostatectomía radical, entonces un marcador como el nuestro con 50% de sensibilidad y 85% de especificidad tendría un valor predictivo negativo de 0.87 y un valor predictivo positivo de 0.45. Por lo tanto, un marcador con este desempeño definiría un grupo de pacientes con sólo una probabilidad del 13% de recurrencia después de la cirugía y un grupo de pacientes con una probabilidad de 45% de recurrencia. El primer grupo podría seguirse solamente para determinar un aumento del PSA, y el segundo grupo serían candidatos para terapias adyuvantes.

Mientras que estos candidatos se han estudiado en muestras de prostatectomías, serían útiles además para el análisis de biopsias. Un marcador que predice un resultado después de una prostatectomía se correlaciona con la agresividad y el potencial metastásico del tumor, y estas propiedades también estarán presentes en la biopsia. Después de las pruebas de biopsias y de determinación de la etapa, los pacientes optan por una espera vigilante, terapia curativa definitiva, o una combinación de tratamientos (tal como cirugía más radioterapia o ablación de andrógenos). Una prueba molecular con alto valor predictivo negativo permitiría que más pacientes escogieran una espera vigilante. Una prueba molecular con un valor predictivo positivo suficientemente alto seleccionaría un subconjunto de pacientes que no deberían recibir radiación o cirugía solamente. Por lo tanto, estos candidatos a marcadores de metilación tienen el potencial de reducir tanto los tratamientos insuficientes como los excesivos para el cáncer de próstata.

Ejemplo 4: Análisis cuantitativo de la metilación en tiempo real

El ADN genómico se analizó mediante el uso de la técnica de PCR en tiempo real después de la conversión con bisulfito.

El ensayo QM (= ensayo cuantitativo de metilación) es un método basado en la PCR en tiempo real para la detección cuantitativa de la metilación del ADN. El principio del ensayo se basa en la amplificación específica sin metilación de la región objetivo y una detección específica de metilación mediante hibridación competitiva de dos sondas diferentes específicas para el estado CG o el TG, respectivamente. Para el presente estudio, se usaron sondas TaqMan que se marcaron con dos colorantes fluorescentes diferentes ("FAM" para las sondas específicas de CG, "VIC" para las sondas específicas de TG) y se modificaron adicionalmente por una molécula desactivadora ("TAMRA" o "desactivador no fluorescente/ ligador de surco menor").

La evaluación de los datos en bruto del ensayo QM es posible con dos métodos diferentes:

1. Medir las intensidades de fluorescencia absolutas (FI) en la fase logarítmica de la amplificación. Diferenciar en ciclos umbrales (Ct) de la sonda específica de CG y de TG.

En la siguiente serie de ensayos cuantitativos de metilación se cuantifica la cantidad de ADN amplificado de la muestra con referencia al gen GSTP1 para normalizar con respecto al ADN de entrada. Para la estandarización, los iniciadores y la sonda para el análisis del gen GSTP1 carecen de dinucleótidos CpG de manera que la amplificación es posible, independientemente de los niveles de metilación. Como no hay posiciones variables sin metilación, sólo se requiere un oligonucleótido de sonda.

Las reacciones se calibran con referencia a patrones de ADN de niveles de metilación conocidos con el objetivo de cuantificar los niveles de metilación dentro de la muestra. Los patrones de ADN estaban compuestos de ADN genómico amplificado de phi29 tratado con bisulfito (es decir no metilado), y/o ADN genómico amplificado de phi29 tratado con la enzima metilasa Sss1 (de esta manera se metila cada posición CpG en la muestra), que después se trata con solución de bisulfito. Se usaron siete estándares de referencia diferentes con 0%, (es decir solamente ADN genómico amplificado de phi29), 5%, 10%, 25%, 50%, 75% y 100% (es decir solamente genómico de phi29 tratado con Sss1).

Tratamiento con bisulfito

El tratamiento con bisulfito se llevó a cabo basado en el método descrito por Olek y otros Nucleic Acids Res. 1996 Dec 15;24(24):5064-6, y optimizado para el flujo de trabajo del laboratorio del solicitante.

Patrones de cuantificación

Las reacciones se calibran con referencia a patrones de ADN de niveles de metilación conocidos con el objetivo de cuantificar los niveles de metilación dentro de la muestra. Los patrones de ADN estaban compuestos de ADN genómico humano amplificado con phi29 tratado con bisulfito (Promega) (es decir no metilado), y/o ADN genómico amplificado de phi29 tratado con la enzima metilasa Sss1 (de esta manera se metila cada posición CpG en la muestra), que después se trata con solución de bisulfito. Se usaron siete estándares de referencia diferentes con 0%, (es decir solamente ADN genómico amplificado de phi29), 5%, 10%, 25%, 50%, 75% y 100% (es decir solamente genómico de phi29 tratado con Sss1). Los lotes de 2000 ng de ADN genómico humano (Promega) se trataron con bisulfito. Para generar ADN MDA metilado, 13 tubos de 4.5 µg de ADN-MDA (700 ng/µl) se trataron con Sss1.

Ensayo de control

El diseño del ensayo de GSTP1-C3 lo hace adecuado para cuantificar ADN a partir de diferentes fuentes, que incluyen muestras frescas/congeladas, muestras remotas tales como plasma o suero, y ADN obtenido a partir de un espécimen en archivos tal como material embebido en parafina. Los siguientes oligonucleótidos se usaron en la reacción para amplificar el amplificado control:

Iniciador control1: GGAGTGGAGGAAATTGAGAT (SEQ ID NO:966)

Iniciador control2: CCACACAACAAATACTCAAAAC (SEQ ID NO:967)

Sonda Control: FAM-
TGGGTGTTTGTAAATTTTGTGTTAGGTT-TAMRA
(SEQ ID NO:968)

Programa del Ciclo (40 ciclos): 95°C, 10 min
 95°C, 15 seg
 58°C, 1 min

Diseño del ensayo y condiciones de reacción

Se desarrollaron dos ensayos para el análisis del gen PITX2 (SEQ ID NO:961) Ensayo 1:

Iniciadores:	GTAGGGGAGGGAAGTAGATGTT (SEQ ID NO:969)
	TTCTAATCCTCCTTTCCACAATAA (SEQ ID NO:970)
Sondas:	FAM-AGTCGGAGTCGGGAGAGCGA-TAMRA (SEQ ID NO:971)
	VIC-AGTTGGAGTTGGGAGAGTGAAAGGAGA -TAMRA (SEQ ID NO:972)
Amplicon (SEQ ID NO:973):	
GAGGGGAGGGAAGTAGATGAGCGGCGGAAGAGTCGGGAGTCGGAGTCGGGAGAGCGAAGGAG AGGGGAHTGGCGGGGATTAGGAGHAAICGAGGAGIAGGAGIACGGATHTAATGTGGAAAGGAGGAHA GAA	
Longitud del fragmento: 143 bp	

Las posiciones de los iniciadores, sondas y dinucleótidos CpG se resaltan.

Componentes PCR (suministrados por Eurogentec) : 3 mM amortiguador de MgCl₂, 10x amortiguador, Hotstart TAQ, 200 µM dNTP, 625 nM cada iniciador, 200 nM cada sonda.

Programa del Ciclo (45 ciclos):	95°C, 10 min
	95°C, 15 seg
	62°C, 1 min
Ensayo 2 :	
cebadores:	AACATCTACTTCCCTCCCCTAC (SEQ ID NO:974)
	GTTAGTAGAGATTTTATTAATTTTATTGTAT (SEQ ID NO:975)
Sondas:	FAM-TTCGTTGCGCGGT-MGBNQF (SEQ ID NO:976)
	VIC-TTTGGTTGTGTGGTTG- MGBNQF (SEQ ID NO:977)
Amplicon (SEQ ID NO:978):	
GTAGIAGAGATTtAATAAATTtAATGtAAGTGGCGCGCGGGCGGtCGGtCGAGtCGGtTCCGCGGtTGGC CAttAGGAGCGAGtAAGCGtCGGGCGAGCGtCGGGGGGAGCGAGIAGGGGCGACGAGAAACGAGG tAGGGGAGGGAAGTAGATGt	
Longitud del fragmento: 164 bp	

Las posiciones de las sondas, iniciadores y posiciones CpG se resaltan.

Las sondas cubren tres posiciones CpG co-metiladas.

Componentes PCR (suministrados por Eurogentec) : 2,5 mM amortiguador de MgCl₂, 10x amortiguador, Hotstart TAQ, 200 µM dNTP, 625 nM cada iniciador, 200 nM cada sonda.

Programa (45 ciclos):	95°C, 10 min
	95°C, 15 seg
	60°C, 1 min

El grado de metilación en un locus específico se determinó mediante las siguientes fórmulas:

Usando intensidad de fluorescencia absoluta: tasa de metilación = $100 + I(CG) / (I(CG) + I(TG))$

(I = intensidad de fluorescencia de la sonda CG o la sonda TG)

Usando el ciclo de umbral Ct: tasa de metilación = $100 * CG / (CG + TG) = 100 / (1 + TG/CG) = 100 / (1 + 2^{\Delta Ct})$

(suponiendo eficacia PCR E=2; $\Delta Ct = Ct(\text{metilado}) - Ct(\text{no metilado})$)

Ejemplo 5: Validación

El objetivo principal de esta fase de la investigación fue confirmar la importancia de los candidatos a marcadores previamente identificados y optimizar los valores de corte de la metilación. Los marcadores deben ser adecuados para dividir a los pacientes que experimentan prostatectomía en dos grupos: uno con una alta probabilidad de recurrencia del PSA y uno con baja probabilidad de recurrencia del PSA. Adicionalmente, los marcadores deben proporcionar información adicional al análisis del grado de Gleason. Los marcadores que cumplen estos criterios tendrán un papel clínico importante en la selección de pacientes sometidos a prostatectomía para la terapia adyuvante.

El solicitante había identificado previamente varios marcadores con niveles de metilación significativamente mayores en pacientes que experimentaron recurrencia del PSA dentro de los 24 meses posteriores a la cirugía en comparación con los pacientes que no experimentaron recurrencia del PSA (ver anteriormente los Ejemplos 1 a 4). Seis de estos marcadores se transfirieron a una plataforma en tiempo real (Ensayo QM). Estos ensayos se usaron para analizar los niveles de metilación de 612 muestras de prostatectomías embebidas en parafina a partir de una cohorte de pacientes con ganglios negativos de tres instituciones.

El objetivo principal de la presente descripción fue proporcionar marcadores que puedan diferenciar entre pacientes con baja probabilidad de recurrencia del PSA después de la cirugía y aquellos con una alta probabilidad de recurrencia del PSA. Se proporciona además el desempeño de estos marcadores en comparación con indicadores pronósticos tradicionales tales como el grado de Gleason y la información de la etapa.

Es un objetivo adicional de la presente descripción determinar dónde los marcadores son más informativos en relación con la evaluación de pronóstico clínico actual y en consecuencia proporcionar modalidades de uso particularmente preferidas. Particularmente se prefiere que una prueba molecular descrita en la presente se combine, ya sea formal o informalmente, con información a partir de otras fuentes pronósticas, en particular el grado de Gleason.

Métodos: Descripción del ensayo QM

Cada ensayo QM se desarrolló para mejorar el desempeño sin alterar drásticamente las condiciones estándar con el objetivo de permitir una multiplexación futura. Las concentraciones de iniciadores y sondas, la concentración de MgCl₂ y la temperatura de hibridación se optimizaron bajo condiciones fijas del amortiguador y la polimerasa. Los ensayos se diseñaron y optimizaron para asegurar el análisis cuantitativo de la metilación de cada marcador entre 10 y 100 por ciento de metilación. Los productos del ensayo se comprobaron en un gel de agarosa y no se detectaron productos no deseados. Los resultados del procedimiento de optimización se muestran en las siguientes tablas.

Conjunto de muestras

- 5 Se analizaron muestras de tejidos de prostatectomías embebidos en parafina de 605 pacientes. Las muestras fueron proporcionadas por el Colegio de Medicina Baylor SPORE, Departamento de Urología de la Universidad de Stanford, y el Hospital Virginia Mason en Seattle. Las muestras de Stanford y Virginia Mason se prepararon al encontrar primero el bloque quirúrgico con el por ciento más alto de tumor, y después cortar el bloque. Se prepararon tres tubos, cada uno con tres secciones de 10 micrómetros de grosor. El procedimiento fue ligeramente diferente en Baylor. Se extrajo un núcleo de tejido del tumor dentro del bloque de la prostatectomía, y después este núcleo se cortó en secciones de 10 micras. Se incluyeron diez secciones en cada uno de los tres tubos.
- 10 Una sección adyacente se montó en un portaobjetos y se tiñó con H&E para el análisis histológico. Un patólogo revisó estos portaobjetos para una determinación independiente del grado de Gleason y el porcentaje tumoral. Los resultados del grado de Gleason se usaron para todos los análisis en este informe. Los valores del grado de Gleason originales del proveedor están disponibles, pero no se usaron para el análisis debido a los sesgos conocidos e hipotéticos entre los proveedores. Stanford, por ejemplo, usa un porcentaje de Gleason 4/5 para informar el grado, mientras que los otros dos proveedores usan el sistema tradicional. Los valores medidos del grado de Gleason proporcionaron una medida independiente y uniforme.
- 15 Se encontraron pocas muestras sin células tumorales en el portaobjetos de H&E, y estos pacientes se omitieron del análisis. Adicionalmente, encontramos unos pocos pacientes que no tenían un PSA nadir después de la cirugía. Estos pacientes también se excluyeron del estudio. En total, se incluyeron 612 pacientes en el análisis de los datos.
- 20 Debido a su técnica de extracción, el porcentaje tumoral de las muestras proporcionadas por Baylor fue mayor que los otros proveedores.
- 25 Todos los pacientes, de 40-80 años de edad, sometidos a cirugía en las tres instituciones durante ciertos años se incluyeron en el estudio, con la excepción de los pacientes que recibieron terapia neoadyuvante o adyuvante (antes del aumento del PSA) y pacientes con ganglios positivos en el momento de la cirugía. Para Baylor, el periodo de tiempo fue 1993-1998, para Virginia Mason fue de 1996-2000, y para Stanford fue de 1996-1999.
- 30 La cohorte general es similar a otras cohortes de prostatectomías descritas en la literatura, tales como la cohorte recogida por William Catalona y descrita en 2004 (Roehl *et al.*). Las cohortes de pacientes de cada proveedor son similares para casi todos los parámetros clínicos. Una excepción es el tipo de recurrencia. Mientras que otras instituciones típicamente esperan hasta que el PSA del paciente aumente a 0.2 ng/ml o más después de la cirugía, el Departamento de Urología de Stanford trata a muchos pacientes cuando su PSA aumenta a 0.05. Por lo tanto, Stanford tiene una mayor tasa de recurrencia basada en los criterios para la decisión de tratar y una menor tasa de recurrencia basada en los criterios del nivel del PSA (0.2 ng/ml). Ver la sección 6.1 para un resumen de los criterios de definición de eventos. La Figura 89 proporciona un histograma de los periodos de seguimiento para la cohorte de pacientes (se incluyen los tres proveedores). Las barras blancas consisten en los pacientes que no tuvieron una recurrencia antes de ser censurados, y las barras sombreadas consisten en los pacientes que experimentaron recurrencia. Al seleccionar los pacientes que recibieron cirugía en el período 1993-2000, hemos asegurado que la mediana del tiempo de seguimiento de la cohorte (66 meses) sea lo suficientemente larga para tener un número significativo de pacientes que han sufrido una recaída.
- 35 Para la desparafinación, las 627 muestras PET proporcionadas se procesaron directamente en el tubo en el que se entregaron por los proveedores. Se añadió 1 ml (Virginia Mason y Baylor) o 1.8 ml (Stanford) de limoneno a cada tubo y se incubaron a temperatura ambiente durante 10 minutos en un termomezclador con agitación en vórtex ocasional. Las muestras se centrifugaron a 16,000 x g por 5 minutos. El sobrenadante de limoneno se eliminó, y si no se detectó un sedimento, se repitió la centrifugación a mayor velocidad y se eliminó el limoneno restante. Para las muestras de Stanford, el proceso de desparafinación se repitió una vez con 1.6 ml de limoneno para deshacerse de la parafina residual.
- 40 Para la desparafinación, las 627 muestras PET proporcionadas se procesaron directamente en el tubo en el que se entregaron por los proveedores. Se añadió 1 ml (Virginia Mason y Baylor) o 1.8 ml (Stanford) de limoneno a cada tubo y se incubaron a temperatura ambiente durante 10 minutos en un termomezclador con agitación en vórtex ocasional. Las muestras se centrifugaron a 16,000 x g por 5 minutos. El sobrenadante de limoneno se eliminó, y si no se detectó un sedimento, se repitió la centrifugación a mayor velocidad y se eliminó el limoneno restante. Para las muestras de Stanford, el proceso de desparafinación se repitió una vez con 1.6 ml de limoneno para deshacerse de la parafina residual.
- 45 Para la lisis del tejido, se añadieron 190 µl de amortiguador de lisis y 20 µl de proteinasa K a cada muestra desparafinada. Para las muestras de Stanford, se usaron 570 µl de amortiguador de lisis y 60 µl de proteinasa K. Después de agitación en vórtex, las muestras se centrifugaron brevemente y se incubaron en un termoagitador a 60 °C durante 40 horas. Después de la incubación, las muestras se comprobaron para asegurar que la lisis fue completa, y después la proteinasa se inactivó a 95°C durante 10 minutos. Si las muestras lisadas no se usaron directamente para la extracción de ADN, se almacenaron a -20 °C.
- 50 Para la lisis del tejido, se añadieron 190 µl de amortiguador de lisis y 20 µl de proteinasa K a cada muestra desparafinada. Para las muestras de Stanford, se usaron 570 µl de amortiguador de lisis y 60 µl de proteinasa K. Después de agitación en vórtex, las muestras se centrifugaron brevemente y se incubaron en un termoagitador a 60 °C durante 40 horas. Después de la incubación, las muestras se comprobaron para asegurar que la lisis fue completa, y después la proteinasa se inactivó a 95°C durante 10 minutos. Si las muestras lisadas no se usaron directamente para la extracción de ADN, se almacenaron a -20 °C.
- 55 Para la lisis del tejido, se añadieron 190 µl de amortiguador de lisis y 20 µl de proteinasa K a cada muestra desparafinada. Para las muestras de Stanford, se usaron 570 µl de amortiguador de lisis y 60 µl de proteinasa K. Después de agitación en vórtex, las muestras se centrifugaron brevemente y se incubaron en un termoagitador a 60 °C durante 40 horas. Después de la incubación, las muestras se comprobaron para asegurar que la lisis fue completa, y después la proteinasa se inactivó a 95°C durante 10 minutos. Si las muestras lisadas no se usaron directamente para la extracción de ADN, se almacenaron a -20 °C.

Los lisados se aleatorizaron en base al proveedor de las muestras y la recurrencia del PSA. El ADN se aisló mediante el uso de un kit DNeasy Tissue de QIAGEN con unas pocas modificaciones. Se distribuyeron 400 µl de amortiguador AL/E a los tubos de recogida y se añadieron 200 µl de lisado. Las muestras se mezclaron por agitación durante 15 segundos. Las mezclas de lisado/amortiguador se aplicaron a columnas de las placas de DNeasy de 96 pocillos. La placa se selló y se centrifugó a 5790xg durante 10 minutos. Las columnas se lavaron una vez con 500 µl de AW1 y después 500 µl de AW2. El ADN se eluyó con 120 µl de amortiguador AE. Por lo tanto, el volumen final de ADN extraído fue de aproximadamente 120 µl. El ADN se almacenó a -20°C.

Tratamiento con bisulfito

El ensayo de PCR en tiempo real CFF se usó para cuantificar la concentración de ADN de las muestras después de la extracción.

Secuencia CFF:

TAAGAGTAATAATGGATGGATGATGGATAGATGAATGGATGAAGAAAGAAAGGATGAGTGAGAGAA AGGAAGGGAGATGGGAGG (84bp) (SEQ ID NO: 979)	
CFF- iniciador directo	TAAGAGTAATAATGGATGGATGATG (SEQ ID NO: 980)
CFF-iniciador inverso	CCTCCCATCTCCCTTCC (SEQ ID NO: 981)
Sonda TaqMan CFF	ATGGATGAAGAAAGAAAGGATGAGT (SEQ ID NO: 982)

Ajustamos la concentración de cada muestra de ADN genómico de manera que estuviera presente 1 µg de ADN medido con CFF1 en 44 µl. El tratamiento con bisulfito del ADN genómico derivado del tejido embebido en parafina se realizó mediante el uso de un protocolo de 96 pocillos. Cuarenta y cuatro µl de ADN genómico (con aproximadamente 1 µg de ADN amplificable), 83 µl de solución de bisulfito 4.9M (pH 5.45-5.5), y 13 µl de solución DME se pipetearon en los pocillos de la placa. Las muestras se mezclaron completamente después se colocaron en un termociclador con el siguiente programa:

- 5:00 min desnaturalización de ADN a 99°C
- 22:00 min incubación a 60°C
- 3:00 min desnaturalización de ADN a 99°C
- 1:27:00 horas de incubación a 60°C
- 3:00 min desnaturalización de ADN a 99°C
- 2:57:00 horas de incubación a 60°C
- Enfriamiento a 20°C

Después de las incubaciones, cada muestra se dividió en dos alícuotas de 70 µL. Cada alícuota se combinó con 280 µL de amortiguador preparado AVL/portador ARN y 280 µL etanol. Los pocillos se sellaron y las muestras se mezclaron vigorosamente durante 15 segundos. La placa se incubó a temperatura ambiente durante 10 minutos. La primera alícuota se aplicó a la placa QIAamp 96 y la placa se centrifugó durante cuatro minutos a 5790 x g. El proceso se repitió con la segunda alícuota de modo que ambas alícuotas se aplicaron a la misma columna de unión. Las columnas se lavaron con 500 µL de amortiguador AW1, y después 500 µL 0.2 M de NaOH, y después dos veces con 500 µL amortiguador AW2. El ADN se eluyó con 100 µl de amortiguador de elución (Qiagen) previamente calentado a 70 grados C. Los bisADN se almacenaron a -20 °C.

Las muestras de ADN tratadas con bisulfito se almacenaron en 8 placas de 96 pocillos (placas 01-08). Las muestras y los controles se combinaron en dos placas de reacción de PCR de 384 pocillos para cada ensayo QM. Cada placa de ensayo QM contenía las muestras de 4 placas de 96 pocillos (realmente se usaron 85 pocillos por placa) y 1 placa de 96 pocillos

con el ADN patrón (7 mezclas del ADN de calibración y agua para la reacción de PCR control sin molde). Las placas de ensayos QM se corrieron tres veces.

Las placas de PCR de 384 pocillos se pipetearon con la estación de trabajo TECAN. El programa de pipeteo transfirió primero 10 µl de la mezcla maestra y después 10 µl del ADN respectivo al pocillo designado. La mezcla maestra se pipeteó en un tubo Falcon y se distribuyó en 8 frascos con tapa de rosca de 500 µl para el pipeteo automático con la estación de trabajo TECAN.

Todos los ensayos QM se corrieron en un dispositivo en tiempo real ABI TAQMAN 7900HT (software SDS 2.2.) con un volumen de reacción de 20 µl. Los ensayos de PITX2 y CCND2 se corrieron con emulación 9600, y los otros ensayos no. Se usó una configuración automática de la muestra para transferir los nombres correctos de las muestras y los colorantes detectores/reporteros al software de TAQMAN. Las condiciones de los ciclos se ajustaron manualmente y se usó ROX como colorante de referencia pasivo. Todas las placas de PCR de 384 pocillos se analizaron con el software SDS2.2 mediante el uso de las configuraciones del análisis manuales (configuración de la línea de base con valores de inicio y parada y umbral manual) para producir archivos de resultados para cada corrida individualmente.

Métodos: Evaluación del desempeño del Marcador

Definición de Eventos

Después de una prostatectomía exitosa en un paciente con enfermedad no metastásica, no deben quedar células de la próstata en su cuerpo y por lo tanto sus niveles de PSA deben caer a cero. Los niveles de PSA de un paciente se miden típicamente cada 6-12 meses después de la cirugía para asegurarse de que el paciente permanece libre de cáncer de próstata. Si el PSA se vuelve detectable y aumenta a un nivel determinado, el médico y el paciente pueden decidir sobre una terapia adicional. Por lo tanto, el retorno y el aumento de los niveles de PSA son la principal indicación de recurrencia de la enfermedad.

Una recaída del PSA posterior a la cirugía se muestra típicamente por un aumento gradual o rápido de sus niveles en una serie de pruebas secuenciales. En dependencia de las características clínicas del paciente o el enfoque de la institución, los pacientes pueden tratarse tan pronto como se detecta el PSA, cuando alcanza un cierto umbral, o cuando los síntomas clínicos acompañan el aumento del PSA. La mayoría de las instituciones consideran que un nivel de PSA de 0.2 ng/ml es significativo, y si el PSA de un paciente alcanza este nivel y se confirma que está en aumento en las pruebas posteriores, se le ofrecerá una terapia adicional. El Departamento de Urología de Stanford, uno de los proveedores de muestras, considera 0.05 ng/ml como una recurrencia del PSA, y considera un tratamiento para los pacientes cuando su PSA alcanza este nivel.

Un evento en este estudio incluye todas las recurrencias basadas en el PSA. Un nivel del PSA de 0.2 ng/ml, confirmado en pruebas posteriores, se ha demostrado que proporciona la mejor sensibilidad y especificidad para la detección de recurrencia (Freedland y otros 2003). El aumento del PSA a este nivel normalmente precede cualquier desarrollo de recurrencia clínica; por lo tanto, casi todos los pacientes de este estudio están libres de recurrencia clínica en el momento de la recurrencia del PSA. Debido a que Stanford frecuentemente trata a los pacientes con una recurrencia del PSA antes de alcanzar este valor de corte de 0.2ng/ml, muchos de sus pacientes recurrentes serían censurados en el presente estudio si el nivel de PSA de 0.2ng/ml fue el único evento considerado. Por lo tanto, los pacientes de cualquiera de las tres instituciones que reciben terapia debido a los niveles de PSA también se consideran un evento en este estudio.

En resumen, un evento se define en el presente estudio como cualquier aumento del PSA a 0.2 ng/ml (confirmado en prueba posterior) O una decisión de tratar al paciente sobre la base de criterios del PSA.

Procesamiento de datos en bruto de QM

Todos los análisis en este informe se basan en la evaluación de CT. Suponiendo condiciones óptimas de la PCR en tiempo real en la fase de amplificación exponencial, la concentración del ADN metilado (C_{met}) puede determinarse mediante

$$C_{met} = \frac{100}{1 + 2^{(CT_{CG} - CT_{TG})}} [\%],$$

donde

CT_{CG} denota el ciclo umbral del reportero CG (canal FAM) y CT_{TG} denota el ciclo umbral del reportero TG (canal VIC).

Los umbrales para los ciclos se determinaron mediante inspección visual de las gráficas de amplificación (Guía del usuario del sistema de detección de secuencias ABI PRISM 7900 HT). Los valores para los ciclos (CT_{CG} y CT_{TG}) se calcularon con estos umbrales por el software ABI 7900. Siempre que la curva de amplificación no excedió el umbral, el valor del ciclo se estableció en el ciclo máximo por ejemplo 50.

El paquete de software R, versión 2.2. (Gentleman and Ihaka 1997), se usó para el análisis estadístico.

Adicionalmente, usamos el paquete de "supervivencia", versión 2.11-5 (<http://cran.at.r-project.org/src/contrib/Descriptions/survival.html>), para el análisis de supervivencia.

Se utilizó el código de propietario para la validación cruzada k veces, el análisis ROC y las funciones gráficas.

Cada conjunto de datos se representa en un objeto de datos del propietario, denominado "Matriz de Datos Anotados" (ADM). Este objeto de datos contiene las mediciones después del control de la calidad y la promediación, así como todas las anotaciones necesarias para las muestras y los ensayos.

Curvas de calibración del ensayo QM

Una serie de mezclas de ADN-MDA metilado y ADN-MDA no metilado, que estaban en el intervalo de 0 a 100 por ciento metiladas, se incluyeron por triplicado en cada placa de PCR de QM. Estos ADN se usaron para asegurar el desempeño uniforme del ensayo QM en todas las placas de PCR. Todos los ensayos mostraron fuertes capacidades cuantitativas entre 10 y 100%, y algunos ensayos fueron capaces de distinguir consistentemente 5% de ADN metilado de ADN no metilado.

Métodos Estadísticos

Después de control de calidad, cada ensayo se analizó estadísticamente.

Regresión Cox

La relación entre los tiempos de supervivencia libres de recurrencia (RFS) y las covariables se analizaron mediante el uso de modelos de riesgo proporcional de Cox (Cox y Oates 1984; Harrel 2001).

El peligro, es decir, el riesgo inmediato de una recaída, se modela como

$$h(t | x) = h_0(t) \cdot \exp(\beta x) \quad (3)$$

y

$$h(t | x_1, \dots, x_k) = h_0(t) \cdot \exp(\beta_1 x_1 + \dots + \beta_k x_k) \quad (4)$$

para análisis de regresión univariada y multivariada, respectivamente, donde k es 10, m es 100 t es el tiempo medido en meses después de la cirugía, $h_0(t)$ es el riesgo de la línea de base (no especificado), x_i son las covariables (por ejemplo las mediciones de los ensayos) y β_i son los coeficientes de regresión (parámetros del modelo). Los β_i se estimarán al maximizar la probabilidad parcial del modelo de riesgo proporcional de Cox Las pruebas de los índices de Probabilidad se realizan para probar si la metilación está relacionada con el riesgo. La diferencia entre $-2\log(\text{Probabilidad})$ del modelo completo y el modelo nulo es aproximadamente χ^2 -distribuido con k grados de libertad bajo la hipótesis nula $\beta_1 = \dots = \beta_k = 0$.

El hecho de asumir riesgos proporcionales se evaluó mediante los residuos de Schoenfeld escalados (Therneau y Grambsch 2000). Para el cálculo, el análisis y el diagnóstico del modelo de Riesgo proporcional de Cox R se usan las funciones R "coxph" y "coxph.zph" del paquete de "supervivencia".

Análisis de regresión por pasos

Para múltiples modelos de regresión de Cox se usó un procedimiento por pasos (Venables y Ripley 1999; Harrel 2001) con el objetivo de encontrar sub-modelos que incluyan solo las variables relevantes. Usualmente se logran dos efectos mediante estos procedimientos:

- 5
 - Las variables (tasas de metilación) que básicamente no se relacionan con la variable dependiente (DFS/MFS) se excluyen ya que no añaden información relevante al modelo.
 - De un conjunto de variables altamente correlacionadas, sólo se retiene la que tiene la mejor relación con la variable dependiente.

- 10 La inclusión de ambos tipos de variables puede conducir a inestabilidades numéricas y a una pérdida de potencia. Por otra parte, el desempeño del predictor puede ser bajo debido a un exceso de ajuste.

El algoritmo aplicado pretende minimizar el criterio de información de Akaike (AIC).

- 15 El AIC se relaciona con el desempeño de un modelo, los valores más pequeños prometen un mejor desempeño. Mientras la inclusión de variables adicionales siempre mejora el ajuste del modelo y por lo tanto aumenta la probabilidad, el segundo término penaliza la estimación de parámetros adicionales. El mejor modelo presentará un modelo de compromiso con buen ajuste y usualmente un número pequeño o moderado de variables. El cálculo de la regresión por pasos con AIC se realiza con la "etapa" de la función R.

- 20 Curvas de supervivencia de Kaplan-Meier y pruebas Log-rank

- 25 Las curvas de supervivencia se estimaron a partir de datos de RFS mediante el uso del estimador de Kaplan-Meier para la supervivencia (Kaplan y Meier, 1958). Las pruebas de log-rank (Cox y Oates, 1984) se usan para probar las diferencias de dos curvas de supervivencia, por ejemplo la supervivencia en grupos hipermetilados vs. hipometilados. Adicionalmente, se aplicó una variante de la prueba de log-rank usualmente referida como la prueba de Wilcoxon Generalizada (para la descripción ver Hosmer y Lemeshow 1999). Para el análisis de Kaplan-Meier se usan las funciones "survfit" y "survdiff" del paquete de "supervivencia".

- 30 Independencia de marcadores individuales y paneles de marcadores con respecto a otras covariables

- 35 Para comprobar si los presentes marcadores proporcionan información adicional e independiente, se incluyeron otros factores clínicos relevantes en el modelo de Riesgo Proporcional de Cox y se calcularon los valores de p para los pesos para cada factor (Prueba Wald) (Therneau y otros 2000). Para el análisis de factores adicionales en el modelo de Riesgo Proporcional de Cox, se usa la función R "coxph".

Correcciones de prueba múltiple

- 40 No se realizó ninguna corrección de pruebas múltiples.

Estimación de la densidad

- 45 Para las variables numéricas, se realizó la estimación de la densidad de núcleo con una función gaussiana y ancho de banda variable. El ancho de banda se determina mediante el uso de la "regla del pulgar" de Silverman (Silverman 1986). Para el cálculo de las densidades se usa la función R "densidad".

Análisis de la Sensibilidad y la Especificidad

- 50 El método de cálculo de la sensibilidad y la especificidad mediante el uso de la fórmula de Bayes se basó en los estimados de Kaplan-Meier (Heagerty y otros 2000) para las probabilidades de supervivencia en los grupos de marcador positivo y de marcador negativo para un tiempo dado T_{Umbral} . Los ROC se calcularon para diferentes tiempos de referencia T_{umbral} (3 años, 4 años, 5 años, 6 años).

Validación cruzada en k veces

- 55 Para el análisis de la selección del modelo y la robustez del modelo se usó la validación cruzada en k veces (Hastie y otros 2001). El conjunto de observaciones se divide aleatoriamente en k partes. A su vez, cada parte se usó como un conjunto de

prueba, mientras que las restantes partes k-1 constituyen el conjunto de entrenamiento. Este procedimiento se repite m veces.

Resultados

Las 605 muestras se procesaron como se describió anteriormente. Todas las muestras se analizaron con ensayos QM de seis marcadores con tres réplicas. Los datos se filtraron para el control de la calidad, y se analizaron como se describe en la sección de métodos. El desempeño clínico de cada marcador se resume más abajo y las curvas de supervivencia de Kaplan-Meier y las curvas ROC de acuerdo con las figuras 90 a 95. Los valores de p para la comparación de las curvas de supervivencia informadas en los gráficos se basan en la prueba común de log-rank. Los resultados del uso de la prueba de Wilcoxon Generalizada son esencialmente los mismos (datos no mostrados).

El desempeño de los marcadores se examinó primero mediante el uso de la mediana del nivel de metilación como valor de corte. Dado que este valor de corte se fijó antes de examinar los datos, los valores de p pueden usarse para juzgar el desempeño de los marcadores. Cualquier marcador con un valor de p significativo al usar la mediana de la metilación como valor de corte se considera como validado. La mediana del nivel de metilación podría no ser el mejor valor de corte para todos los marcadores, y para estos marcadores la separación del pronóstico puede optimizarse más aún al seleccionar el valor de corte de metilación que resulte en el menor valor de p. Dado que el valor de corte se optimiza específicamente para el valor de p, el valor de p ya no puede usarse para indicar la significación estadística.

Para juzgar la significación del desempeño de un marcador mediante el uso de la mediana de la metilación como valor de corte, usamos un valor de p de 0.005 (suponiendo la corrección para 6 comparaciones). Sobre la base del valor de p (menos de 0.008) y la separación de eventos, PITX2 es el candidato más fuerte. GPR7 junto con la SEQ ID NO:35. Por lo tanto, estos dos marcadores se consideran marcadores validados del pronóstico postquirúrgico del cáncer de próstata, la SEQ ID NO: 63 no fue significativa con el uso de la mediana del nivel de metilación como valor de corte (valores de p 0.018 y 0.0059), pero se desempeña bien cuando el valor de corte de la metilación se optimiza. Ver la Tabla 17 para los resultados.

La Figura 90 A muestra el análisis de supervivencia de Kaplan-Meier del marcador PITX2 de las 585 muestras de pacientes que pasaron el filtro del control de la calidad mediante el uso del valor de corte optimizado de la metilación (13.5%). La figura 90B muestra el análisis de supervivencia de Kaplan-Meier del marcador PITX2 con el uso del valor predefinido de la mediana de metilación como valor de corte, el valor de p fue 0.000017. La Figura 90C muestra el análisis de la curva ROC del marcador PITX2 después de 5 años de seguimiento. El valor de corte de la mediana de metilación se marca como un triángulo, y el valor de corte optimizado de la metilación se muestra como un diamante. El AUC fue de 0.64.

La Figura 91 A muestra el análisis de supervivencia de Kaplan-Meier del marcador GPR7 de las 596 muestras de pacientes que pasaron el filtro del control de la calidad mediante el uso del valor de corte optimizado de la metilación (18.06%). La Figura 91 B muestra el análisis de supervivencia de Kaplan-Meier de dicho marcador mediante el uso del valor predefinido de la mediana de metilación como valor de corte, el valor de p fue 0.0016. La Figura 91C muestra el análisis de la curva ROC de dicho marcador después de 5 años de seguimiento. El valor de corte de la mediana de metilación se marca como un triángulo, y el valor de corte optimizado de la metilación se muestra como un diamante. El AUC fue de 0.64.

La Figura 92 A muestra el análisis de supervivencia de Kaplan-Meier del marcador SEQ ID NO:63 de las 599 muestras de pacientes que pasaron el filtro de control de la calidad mediante el uso del valor de corte optimizado de la metilación (5.79%). La figura 92B muestra el análisis de supervivencia de Kaplan-Meier de dicho marcador mediante el uso del valor predefinido de la mediana de la metilación como valor de corte, el valor de p fue 0.018. La Figura 92C muestra el análisis de la curva ROC de dicho marcador después de 5 años de seguimiento. El valor de corte de la mediana de metilación se marca como un triángulo, y el valor de corte optimizado de la metilación se muestra como un diamante. El AUC fue de 0.60.

La Figura 93 A muestra el análisis de supervivencia de Kaplan-Meier del marcador SEQ ID NO:35 de las 598 muestras de pacientes que pasaron el filtro de control de la calidad mediante el uso del valor de corte optimizado de la metilación (36.77%). La figura 93B muestra el análisis de supervivencia de Kaplan-Meier de dicho marcador mediante el uso del valor predefinido de la mediana de la metilación como valor de corte, el valor de p fue 0.059. La Figura 93C muestra el análisis de la curva ROC de dicho marcador después de 5 años de seguimiento. El valor de corte de la mediana de metilación se marca como un triángulo, y el valor de corte optimizado de la metilación se muestra como un diamante. El AUC fue de 0.61.

La Figura 94 A muestra el análisis de supervivencia de Kaplan-Meier del marcador ABHD9 de las 592 muestras de pacientes que pasaron el filtro de control de la calidad mediante el uso del valor de corte optimizado de la metilación (28.41%). La figura 94B muestra el análisis de supervivencia de Kaplan-Meier de dicho marcador mediante el uso del valor

predefinido de la mediana de la metilación como valor de corte, el valor de p fue 0.018. La Figura 94C muestra el análisis de la curva ROC de dicho marcador después de 5 años de seguimiento. El valor de corte de la mediana de metilación se marca como un triángulo, y el valor de corte optimizado de la metilación se muestra como un diamante. El AUC fue de 0.58.

La Figura 95 A muestra el análisis de supervivencia de Kaplan-Meier del marcador CCND2 de las 604 muestras de pacientes que pasaron el filtro de control de la calidad mediante el uso del valor de corte optimizado de la metilación (2.22%). La figura 95B muestra el análisis de supervivencia de Kaplan-Meier de dicho marcador mediante el uso del valor predefinido de la mediana de la metilación como valor de corte, el valor de p fue 0.22. La Figura 95C muestra el análisis de la curva ROC de dicho marcador después de 5 años de seguimiento. El valor de corte de la mediana de metilación se marca como un triángulo, y el valor de corte optimizado de la metilación se muestra como un diamante. El AUC fue de 0.61.

Evaluación de marcadores en subconjuntos clínicos de los pacientes

Varios factores pronósticos clínicos se utilizan comúnmente para evaluar el cáncer de próstata. El análisis histológico del tumor con la cuantificación del estado de diferenciación tumoral mediante el uso del sistema de grados de Gleason es un indicador pronóstico de particular importancia en la práctica clínica actual. El análisis se continuó al determinar si los marcadores podrían mejorar el análisis del grado de Gleason al subdividir los pacientes dentro de una categoría de Gleason. Además investigamos si los marcadores podrían añadir información a otros indicadores pronósticos, tales como la estimación del riesgo con nomograma (Han y otros 2003) y la etapa de la enfermedad.

Para estos análisis, usamos el análisis de Kaplan-Meier para determinar si nuestros marcadores todavía son informativos en subgrupos de la población, y el análisis de regresión de Cox para determinar si los marcadores proporcionan información independiente de las variables clínicas pronósticas. La puntuación de Gleason (mediante el uso de llamadas Charite Gleason) se dividió en tres grupos (6 o menos, 7, y 8 a 10), la etapa se dividió en dos grupos (T2/confinado al órgano y T3/no confinado al órgano), el PSA se dividió en cuatro grupos (de 0 a 4 ng/ml, de 4 a 10 ng/ml, de 10 a 20 ng/ml, y mayor que 20 ng/ml), y la estimación por nomograma de la supervivencia libre de PSA durante 5 años se dividió en dos grupos (de 90 a 100% y de 0 a 89%).

PITX2

Con el modelo de regresión de Cox, PITX2 es un valioso marcador pronóstico independiente de otra información de pronóstico clínico (Tabla 18). En otras palabras, la metilación de PITX2 añade más información al grado de Gleason que el PSA o la etapa de la enfermedad. El índice de riesgo para PITX2 es de 2.2. En el análisis de supervivencia de los subgrupos, PITX2 tiene el potencial de ser un marcador significativo para todos los pacientes de cáncer de próstata.

Es particularmente interesante ver una fuerte separación dentro del subgrupo de pacientes con la enfermedad confinada al órgano (Figura 96). Los pacientes con la enfermedad confinada al órgano (T2) deben curarse con cirugía. Aquellos que no se curan mediante cirugía deben haber tenido algunas células que dejaron la próstata antes de la cirugía, y por lo tanto tenían células tumorales con características agresivas temprano en el desarrollo. PITX2 puede separar el grupo T2 en un grupo hipometilado con una probabilidad muy pequeña de recurrencia (~5%) y un grupo hipermetilado con un pronóstico más similar a los pacientes T3.

La Figura 96 muestra el análisis de supervivencia del desempeño de PITX2 en subpoblaciones en base a la etapa. El gráfico izquierdo superior muestra el desempeño de la etapa de la enfermedad como marcador pronóstico. El gráfico derecho superior muestra el desempeño de PITX2 en pacientes pT2. El gráfico izquierdo inferior muestra el desempeño de PITX2 en pacientes pT3.

PITX2 es capaz además de estratificar a los pacientes dentro de subcategorías de Gleason. La Figura 97 muestra que el análisis de supervivencia en pacientes con bajo grado de Gleason (Puntuación de 5 o 6) y pacientes con alto grado de Gleason (Puntuación de 8, 9 o 10) resulta en bajos valores de p. Los pacientes con altas puntuaciones de Gleason son actualmente candidatos para ensayos clínicos de terapias adyuvantes postquirúrgicas. Pero los valores de PITX2 sugieren que este no es un grupo uniforme. Los pacientes con PITX2 hipometilado, con alto grado de Gleason tienen 85% de probabilidad de supervivencia libre de la enfermedad a los diez años, mientras que los pacientes hipermetilados con alto grado de Gleason tienen una probabilidad muy baja (~35%). Estos pacientes con alta probabilidad de recurrencia de la enfermedad son los pacientes que deben seleccionarse para una terapia adyuvante o ensayos clínicos.

La Figura 97 muestra el análisis de supervivencia del desempeño de PITX2 en subpoblaciones en base a las categorías de la puntuación de Gleason. El gráfico izquierdo superior muestra el desempeño de la puntuación de Gleason como marcador pronóstico. Los pacientes con puntuación de Gleason de 5 y 6 están marcados como A, los pacientes con puntuación de

Gleason 7 están marcados como B, y los pacientes con puntuación de Gleason 8, 9, y 10 están marcados como C. El gráfico derecho superior muestra el desempeño de PITX2 en los pacientes con puntuación de Gleason 5 y 6. El gráfico izquierdo inferior muestra el desempeño de PITX2 en los pacientes con puntuación de Gleason 7. El gráfico derecho inferior muestra el desempeño de PITX2 en pacientes con puntuación de Gleason 8, 9, y 10.

- 5 Los nomogramas de cáncer de próstata se crean en base a grandes cohortes de pacientes. Ellos combinan matemáticamente la información de la etapa, el grado de Gleason, y los niveles preoperatorios del PSA en un solo indicador pronóstico. Como muestra la Figura 98, el nomograma por sí mismo es muy fuerte. Pero PITX2 es capaz de subdividir más aún a los pacientes.
- 10 La Figura 98 muestra el análisis de supervivencia del desempeño de PITX2 en subpoblaciones en base a la estimación del riesgo por nomograma. El gráfico izquierdo superior muestra el desempeño del nomograma como marcador pronóstico. El gráfico derecho superior muestra el desempeño de PITX2 en pacientes con una probabilidad del 90% de una supervivencia libre del PSA durante 5 años de acuerdo con el nomograma. El gráfico izquierdo inferior muestra el desempeño de PITX2 en pacientes con menos del 90% de probabilidad de supervivencia libre de PSA durante 5 años de acuerdo con el nomograma.

15 SEQ ID NO.: 63

- 20 Con el análisis de regresión de Cox, la SEQ ID NO:63 es un valioso marcador pronóstico independiente de otra información de pronóstico clínico (Tabla 19). El índice de riesgo es de 2.9. En el análisis de supervivencia de los subgrupos, la SEQ ID NO:63 parece tener el potencial de ser un marcador significativo para algunos subgrupos, tales como los pacientes de alto grado de Gleason (Figura 99) y pacientes con mal pronóstico en base al nomograma (Figura 100).

25 La Figura 99 muestra el análisis de supervivencia del desempeño de la SEQ ID NO:63 en pacientes con puntuaciones de Gleason 8, 9, y 10.

La Figura 100 muestra el análisis de supervivencia del desempeño de la SEQ ID NO:63 en pacientes con menos del 90% de probabilidad de una supervivencia libre de PSA durante 5 años de acuerdo con el nomograma.

La SEQ ID NO:35 es un marcador para algunos subgrupos, tales como los pacientes pT2 (Figura 101).

30 Discusión

PITX2, la SEQ ID NO:35, y GPR7 todos muestran información pronóstica significativa cuando la mediana del nivel de metilación se usa como valor de corte. El establecimiento del valor de corte de la metilación incluso superior a la mediana mejora el desempeño de estos tres marcadores. Esto tiene el efecto de disminuir el grupo de marcadores positivos y aumentar la especificidad de la prueba. La mediana del nivel de metilación no es óptimo para la SEQ ID NO: 63. En cambio, un menor valor de corte separa más claramente los grupos de buen y mal pronóstico para este marcador. Los valores de corte de la metilación optimizados para estos cuatro marcadores caen todos en el intervalo para el cual sus respectivos ensayos son técnicamente muy adecuados.

- 40 Los pacientes cuyas muestras se analizaron en este estudio son representativos de la población a la que se dirigiría una prueba de prostatectomía. Por lo tanto, es posible especular sobre la información que estos marcadores podrían proporcionar para futuros pacientes. PITX2, por ejemplo, tiene una sensibilidad de alrededor del 60% y una especificidad del 70%. En el análisis de Kaplan-Meier en la Figura 90, el grupo de marcadores positivos tiene aproximadamente tres veces el riesgo de recurrencia después de diez años que el que tiene el grupo de marcadores negativos. En la Figura 97, los
- 45 pacientes con grado de Gleason 8-10 que son positivos para PITX2 tienen un 65% de probabilidad de recurrencia del PSA en 10 años. Por el contrario, los pacientes con grado de Gleason 8-10 que estaban negativos en los marcadores sólo tenían un 15% de probabilidad de recaída del PSA. La adición de la información de marcadores de metilación a la estratificación del grado de Gleason permitirá a los médicos identificar un subgrupo de mal pronóstico que más se pueden beneficiar de la terapia adyuvante. Si estos marcadores de metilación se incorporan en el procedimiento de selección de pacientes para los
- 50 ensayos clínicos de terapia adyuvante, los médicos pueden comenzar a ver un claro beneficio a la adición de tratamientos adyuvantes tempranos para pacientes con mal pronóstico.

Además de añadir información al grado de Gleason, PITX2 y algunos de los otros marcadores pueden además estratificar los pacientes con la enfermedad confinada al órgano. Los pacientes con la enfermedad que está verdaderamente confinada al órgano se curarán mediante la eliminación completa del órgano. Los pacientes con la enfermedad que parece estar confinada al órgano, pero tienen micrometástasis no detectadas, no se curarán mediante cirugía. Estos dos grupos de pacientes, ambos con lesiones operables pequeñas, tienen tumores con capacidades muy diferentes para las metástasis. PITX2 y algunos de los otros marcadores parecen detectar estas diferencias subyacentes en la agresividad tumoral básica.

La capacidad de estos marcadores para añadir información a los marcadores usados actualmente es esencial. El grado de Gleason y la etapa ya proporcionan información pronóstica importante, una nueva prueba que no sustituiría sino complementarías estas fuentes tradicionales de información es más valiosa y tendría mayor probabilidad de ser adoptada fácilmente en la práctica clínica.

En el análisis de los marcadores en los subgrupos de pacientes, los marcadores frecuentemente parecían más fuertes en pacientes con mal pronóstico en base a las variables clínicas tradicionales. Los pacientes con grado de Gleason de 8-10 y los pacientes con baja probabilidad por el nomograma para una supervivencia libre del PSA están bien estratificados por los presentes marcadores en grupos de buen y mal pronósticos. Para una prueba de prostatectomía, estos son los pacientes ideales a los cuales dirigirse, ya que la prueba se usaría para seleccionar un grupo de pacientes con mal pronóstico que más pueden beneficiarse de la terapia adyuvante. Para los pacientes T3 (enfermedad no confinada al órgano), se prefiere la SEQ ID NO: 63 como marcador. En general, este análisis demuestra que los presentes marcadores son especialmente adecuados para identificar pacientes de mal pronóstico.

Ejemplo 6: Prueba de ensayos en tejido embebido en parafina.

En el siguiente análisis, se analizó la metilación dentro de muestras de tejido de próstata embebido en parafina por medio de los ensayos que se muestran en la Tabla 12 para el análisis de las SEQ ID NO: 19, 35, 37 y 63. Esto se comparó después con la misma medición llevada a cabo en las muestras congeladas descritas en el Ejemplo 2.

Muestras

Las muestras fueron muestras de prostatectomías embebidas en parafina o tejidos frescos congelados como se describió en el Ejemplo 3. Las muestras se cortaron, el tejido se lisó, después se extrajo el ADN y se convirtió con bisulfito.

Estaban disponibles 309 muestras embebidas en parafina, de éstas se incluyeron en el análisis todas las muestras con al menos 1 ng de ADN por PCR, donde se usó entre 1 y 10 ng de ADN por PCR.

Reactivos:

1x Taqman amortiguador A de PCR
0.25 mM dNTPs
3.5 mM MgCl₂
900 nM cada iniciador
300 nM sonda.
1 unidad de AmpliTaq Gold

Perfil de ciclos térmicos:

Etapa 1. 95°C ->10min (Taq Activación)
Repeticiones:1
Etapa 2. 95°C ->15s (Desnaturalización.)
63,0°C->1 min (Hibridación./Extensión)
Repeticiones:50

Se compararon las siguientes categorías

1. Recaída bioquímica temprana (recaída del PSA después de la prostatectomía en menos de 24 meses) vs. sin recaída bioquímica (sin aumento significativo del PSA durante al menos 4 años de monitoreo del PSA después de la cirugía)
En el grupo sin recaída estaban disponibles 132 muestras y en el grupo de recaída temprana estaban disponibles 59 muestras.

2. Alto grado de Gleason (puntuación=8-10) vs Bajo grado de Gleason (puntuación = 6-2 sin componentes de grados 4 o 5)
59 muestras estaban disponibles en el grupo de alto grado de Gleason y 64 muestras estaban disponibles en el grupo de bajo grado de Gleason.

Resultados

Reivindicaciones

- 5 1. Un método para proporcionar un pronóstico de un sujeto con un trastorno proliferativo de células de la próstata, que comprende las etapas de:
 - 10 a) determinar el estado de expresión del gen PITX2 en una muestra biológica obtenida de dicho sujeto; y
 - b) determinar el pronóstico de dicho sujeto basado en dicha expresión, en donde la subexpresión es indicativo de un pronóstico negativo, en donde la expresión se determina mediante la medición del nivel de ARNm.
- 15 2. El método de acuerdo con la reivindicación 1, en donde al menos una variable pronóstica adicional se tiene en cuenta al determinar el pronóstico de c).
- 20 3. El método de acuerdo con la reivindicación 2, en donde dicha variable pronóstica se selecciona del grupo que consiste en nomograma, nivel de PSA y puntuación de Gleason.
- 25 4. El método de acuerdo con cualquiera de las reivindicaciones 1 a 3, en donde dicha muestra se selecciona del grupo que consiste en células o líneas celulares, cortes histológicos, biopsias, tejido embebido en parafina, fluidos corporales, eyaculado, orina, sangre, y combinaciones de los mismos.
5. El método de acuerdo con cualquiera de las reivindicaciones 1 a 4, en donde la expresión se determina por el uso de al menos una técnica seleccionada del grupo que consiste en análisis de membrana de Northern, PCR con transcripción inversa, PCR en tiempo real, protección RNAsa, y análisis de microarreglos.
6. Uso de un oligonucleótido contra PITX2 en el método de acuerdo con cualquiera de las reivindicaciones 1 a 5.
- 30 7. Uso de un método de acuerdo con cualquiera de las reivindicaciones 1 a 5 para proporcionar el pronóstico de un sujeto con un trastorno proliferativo de células de la próstata.

Gráfico de ROC

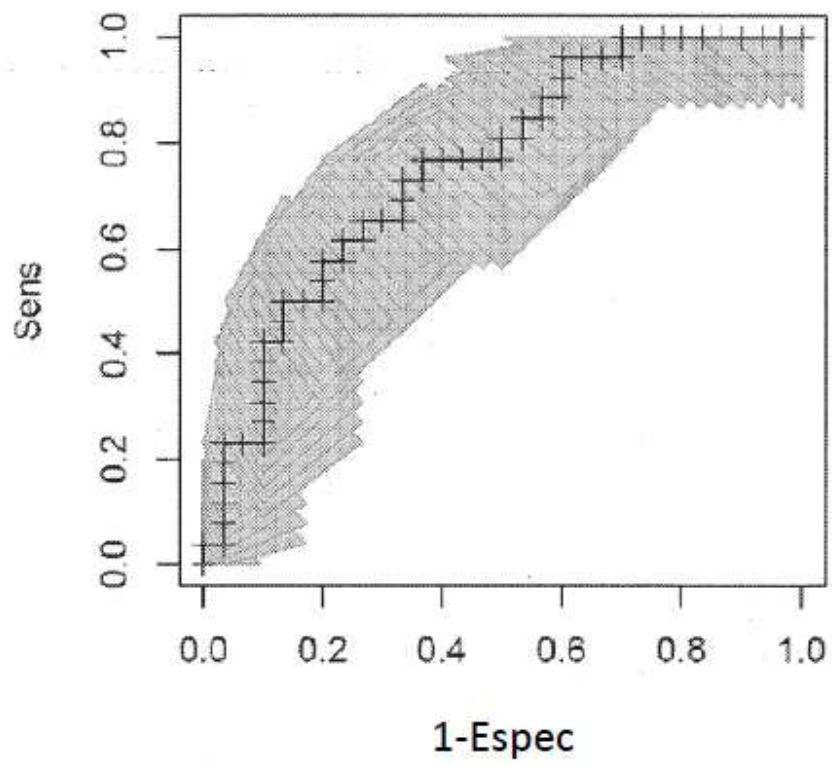


Figura 1

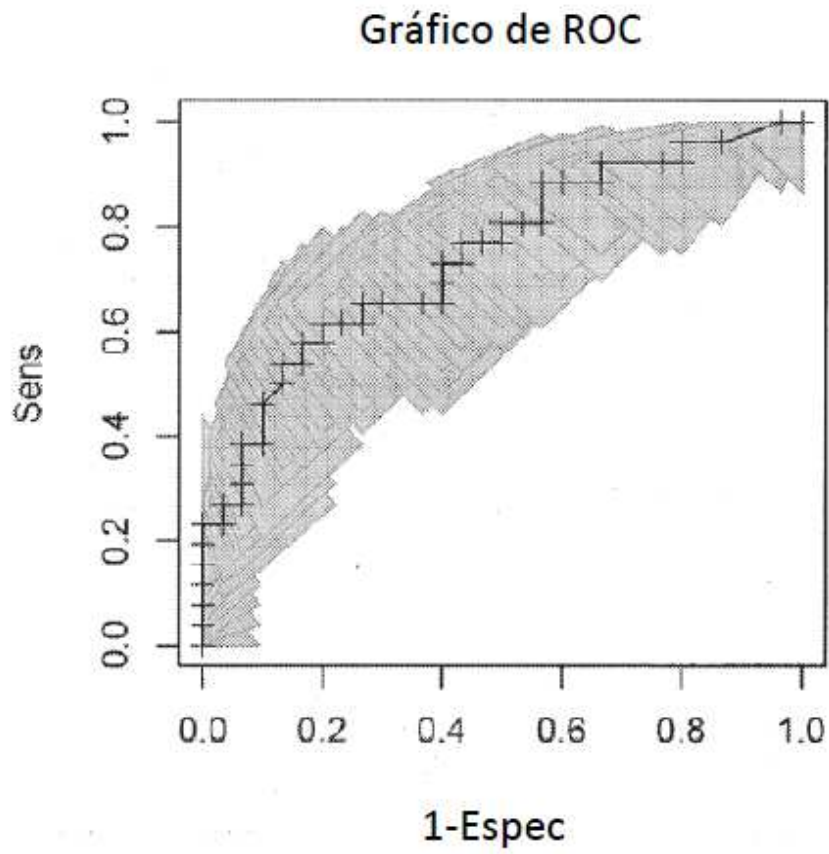


Figura 2

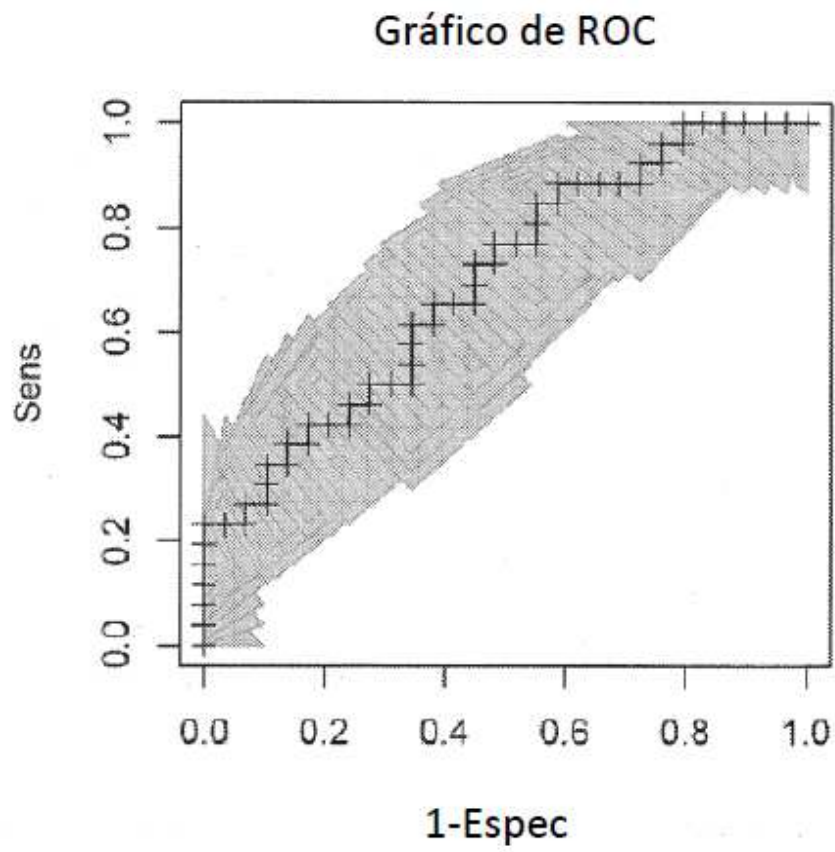


Figura 3

Gráfico de ROC

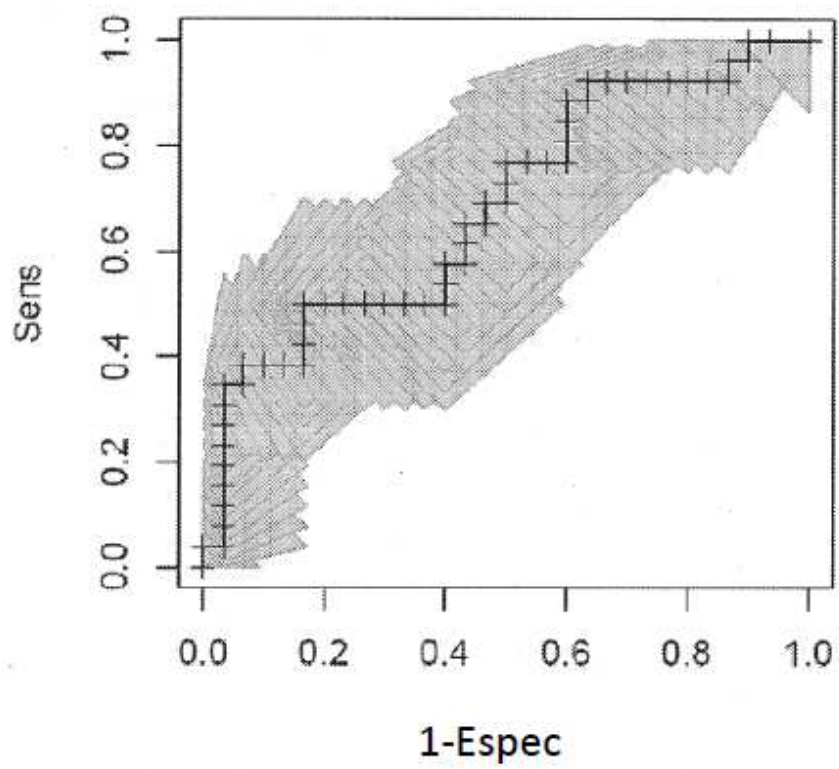


Figura 4

Gráfico de ROC

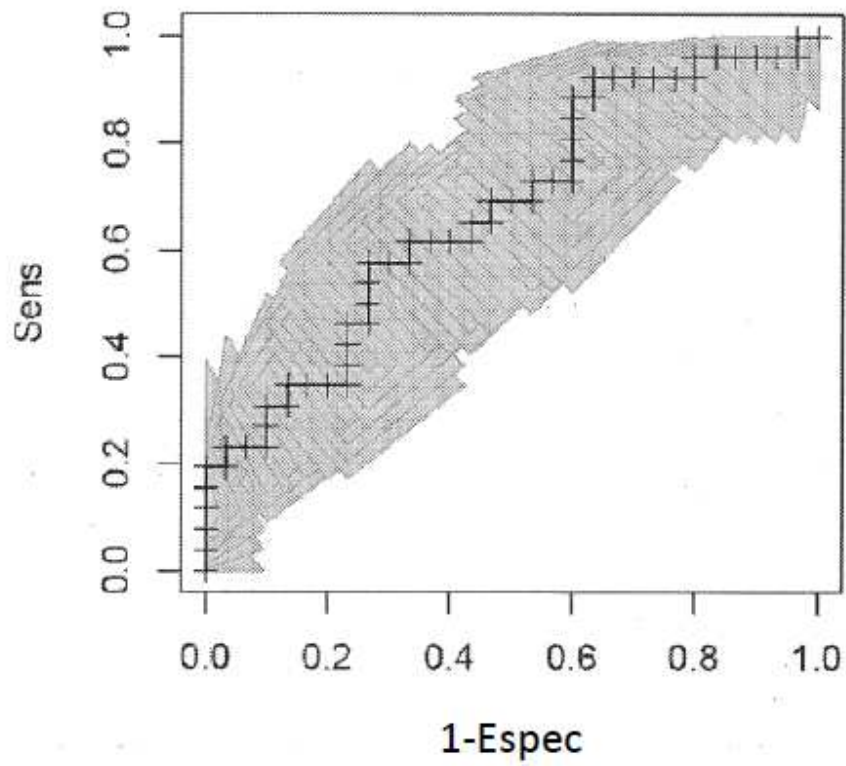


Figura 5

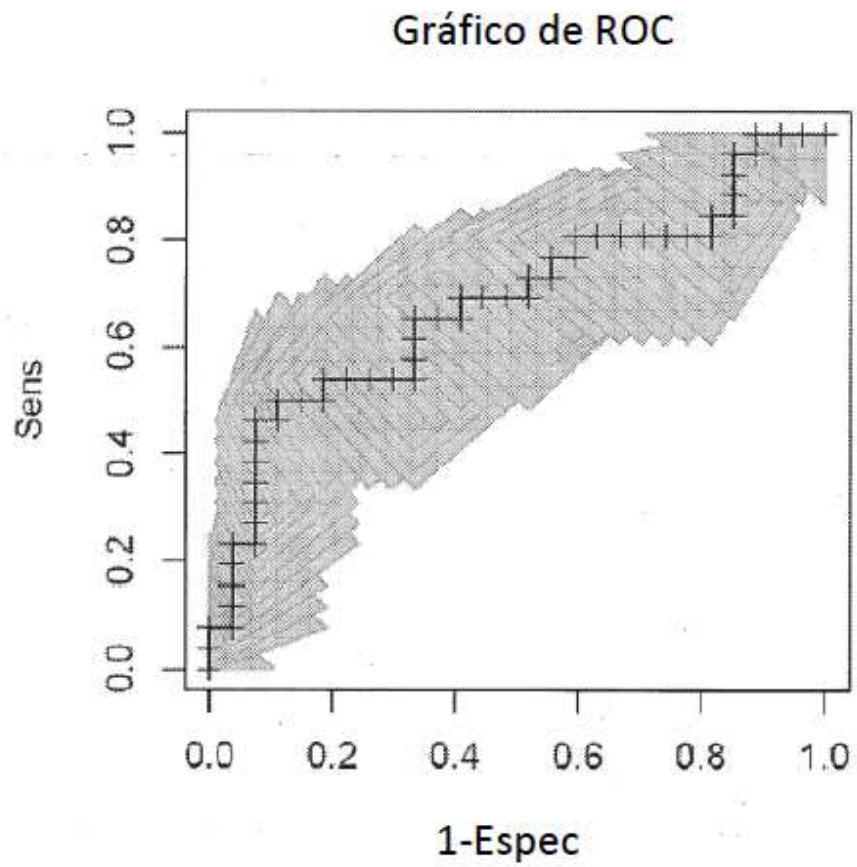


Figura 6

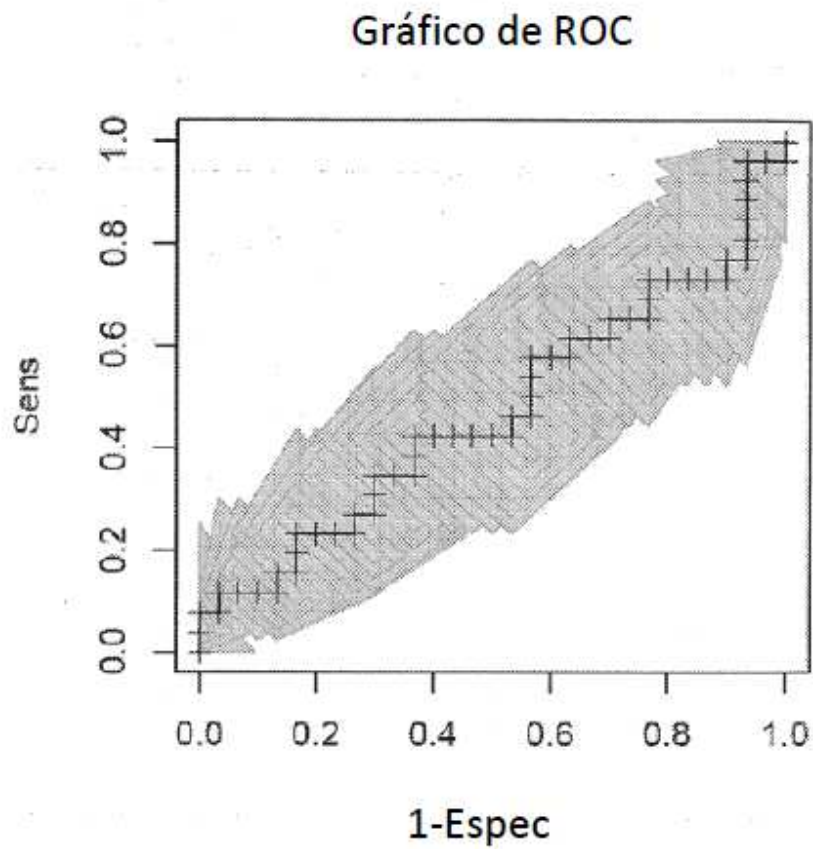


Figura 7

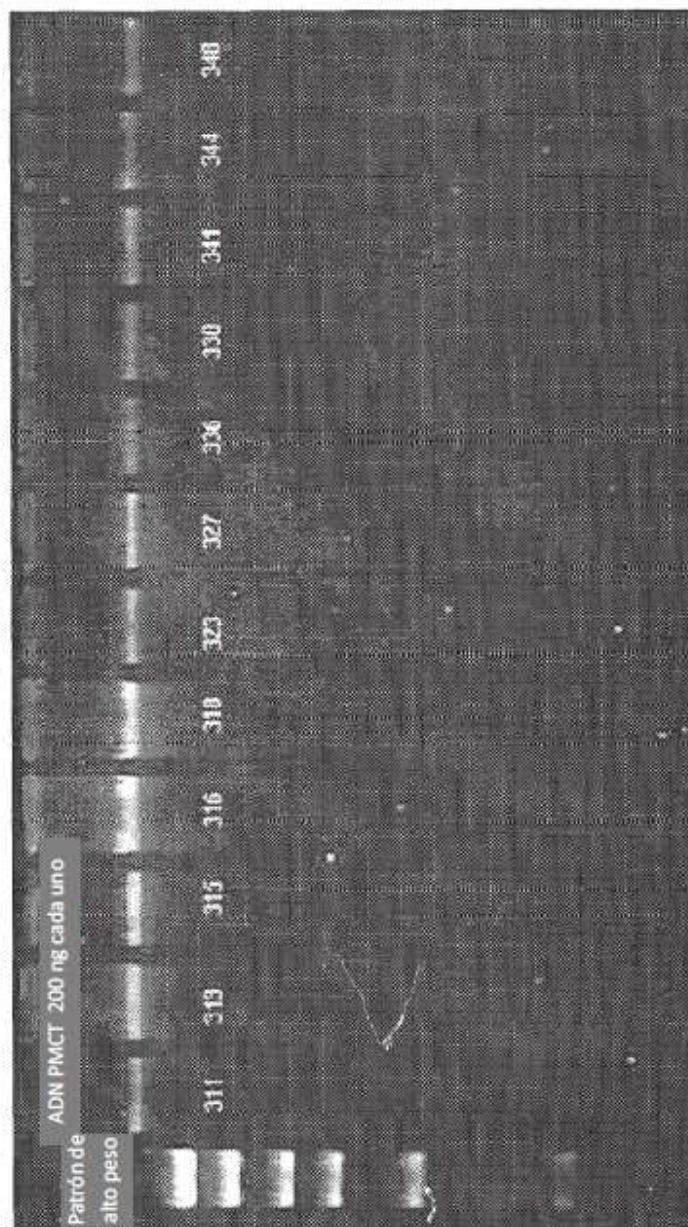


Figura 8

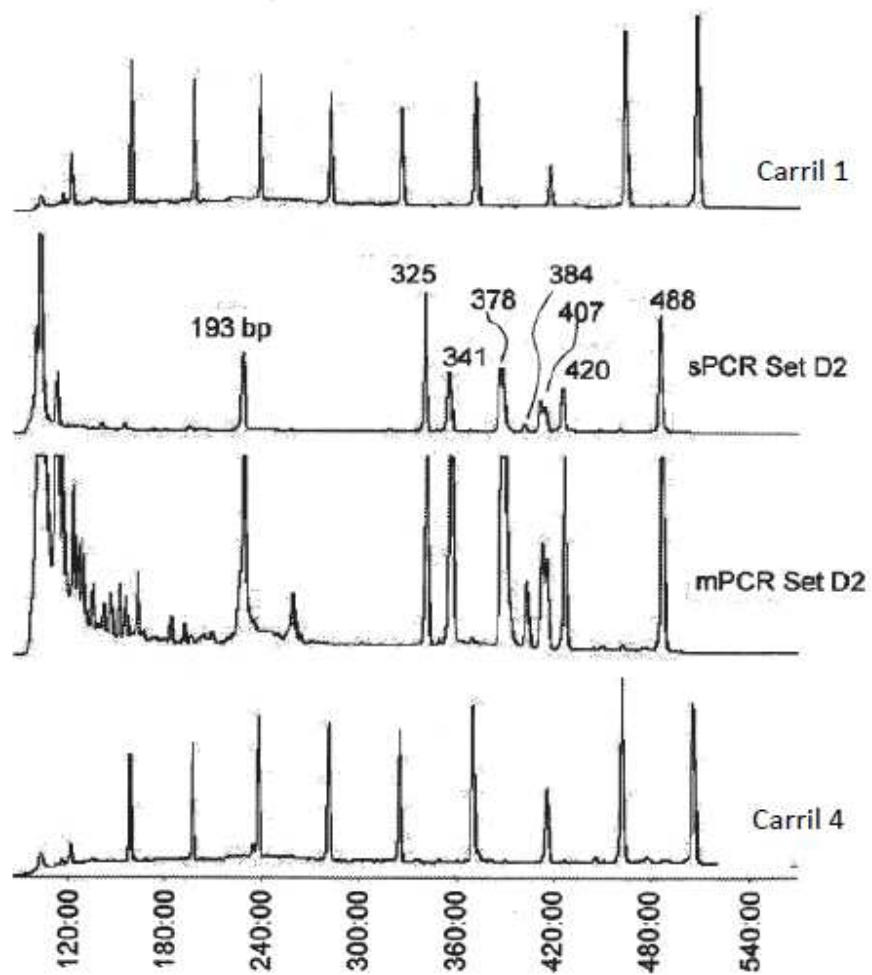


Figura 9

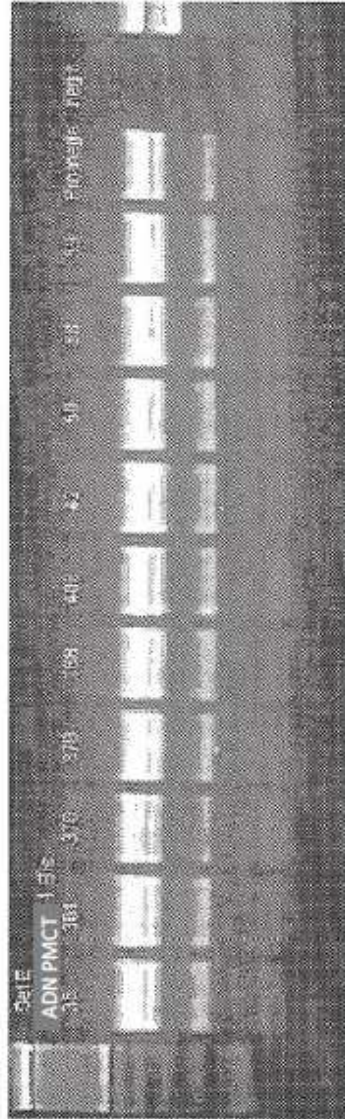


Figura 10

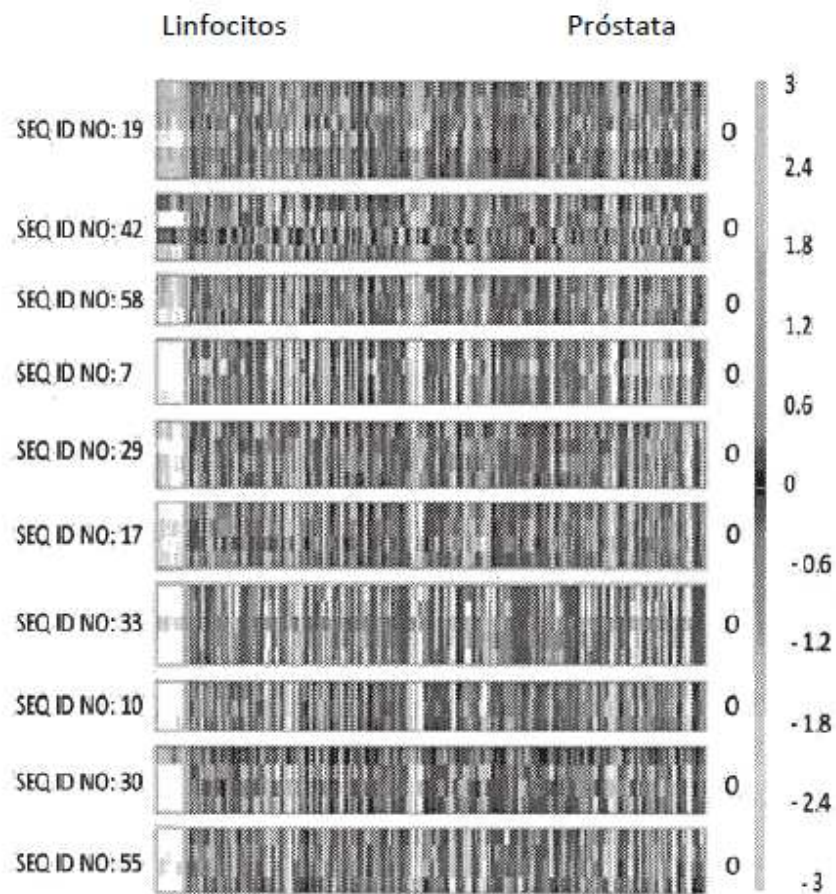


Figura 11

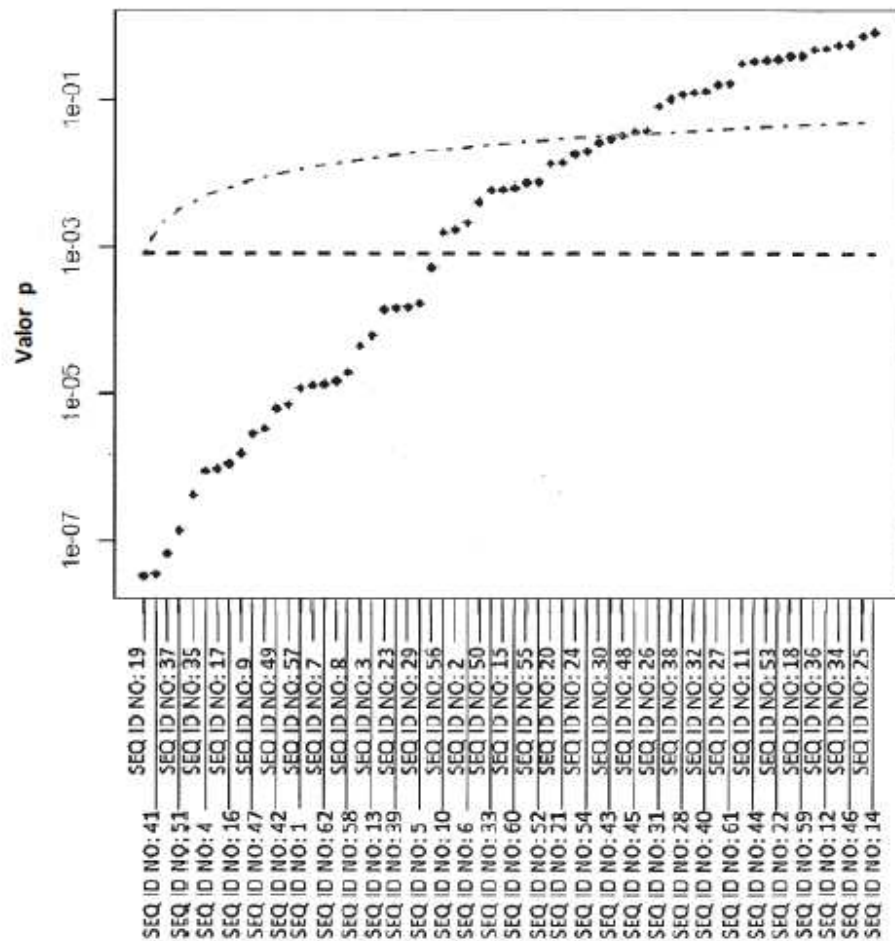


Figura 12

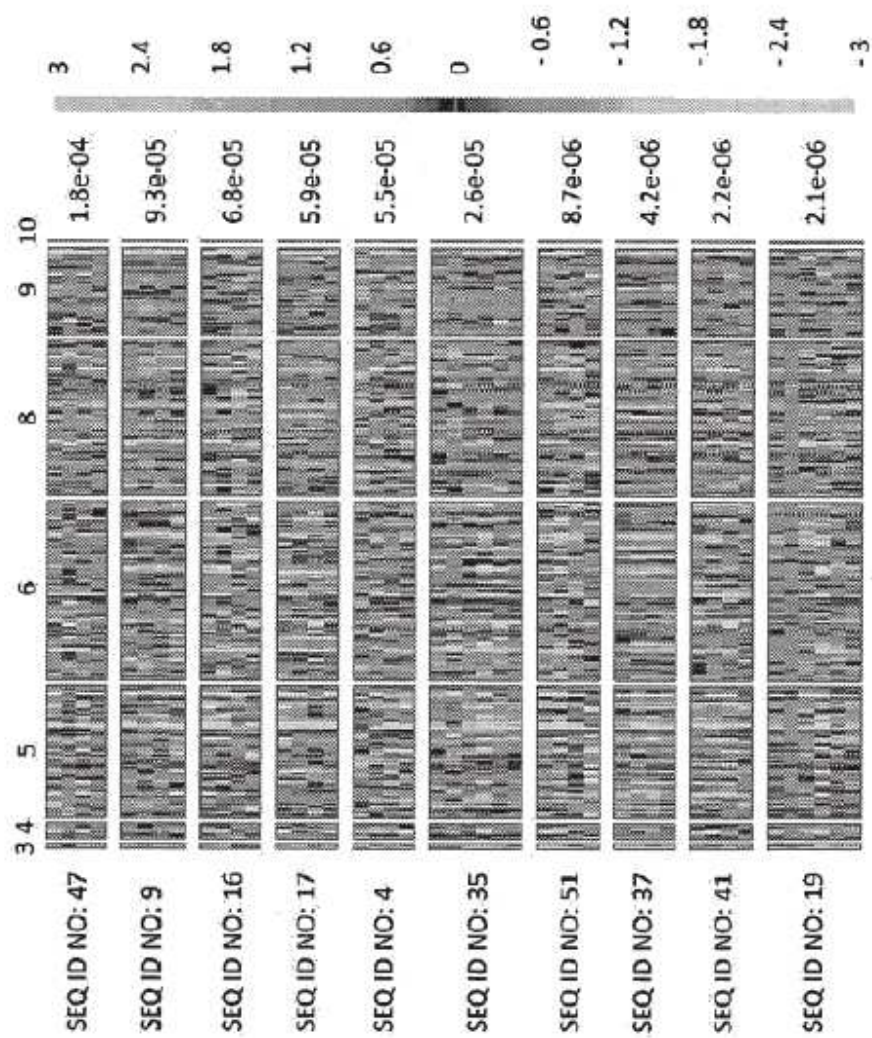


Figura 13

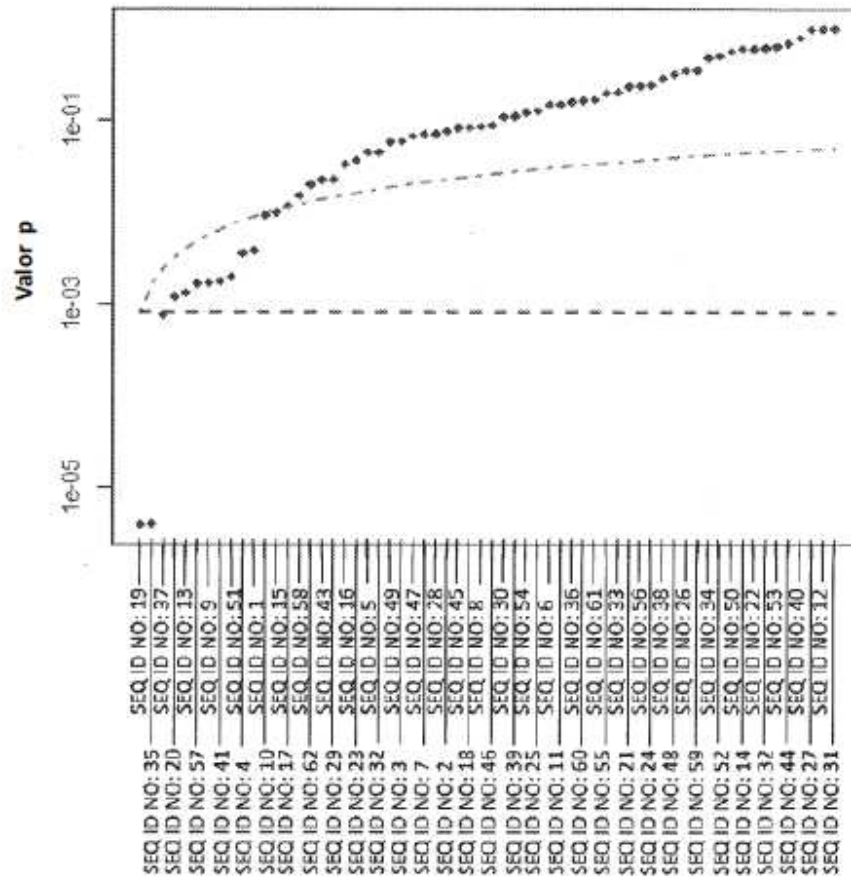


Figura 14

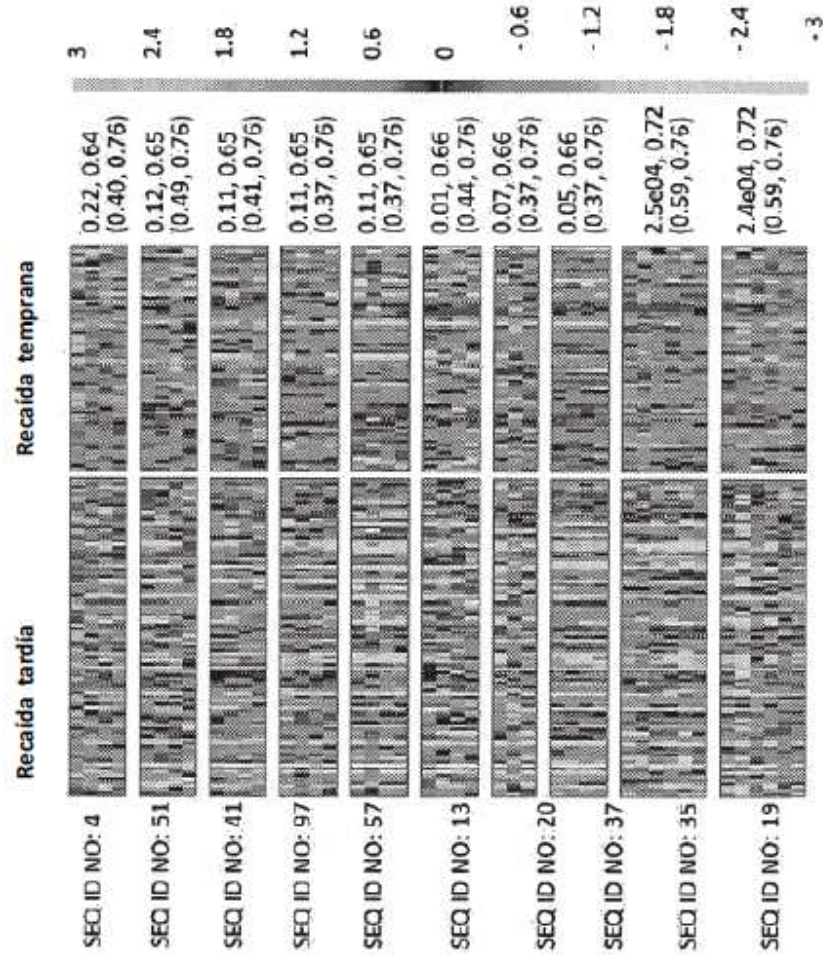


Figura 15

Figura 16: Amplificado de la SEQ ID NO: 14

1	GAAGAGGTGCTGACAAATTAAAAATTACAGGTTAGTTAAATGCATCCCTGCCGCCGCTGCAGGCTCCGCCCTTTGCATTAAAG	80
1	GAAGAGGTGTTGAGAAATTAAAA	80
1	GAAGAGGTGCTGACAAATTAAAAATTACAGGTTAGTTAAATGCATCCCTGCCGCCGCTGCAGGCTCCGCCCTTTGCATTAAAG	80
1	TTAAG	80
1	TTAAG	80
81	CGGGCGCTGATTGTGCGCGCCCTGCGCGACCGCGGGGAGGACTGGCGGCCCGCGGAGCGGACCGGCTAGAGCGCGGGTTAC	160
81	CGGGCGCTGATTGTGCGCGCCCTGCGCGACCGCGGGGAGGACTGGCGGCCCGCGGAGCGGACCGGCTAGAGCGCGGGTTAC	160
81	<u>CGGGCGCTGATT</u>	160
81	<u>CGGGCGCTGATT</u>	160
161	ATTGTTCTGCAGCCGCTCCCTCTTTGTCCTCCCTAGCGGCGCAAGCTGCGAGGTACAGCCCTCTATTGTTCTAGGAG	240
161	ATTGTTCTGCAGCCGCTCCCTCTTTGTCCTCCCTAGCGGCGCAAGCTGCGAGGTACAGCCCTCTATTGTTCTAGGAG	240
241	CACAGAAACCTCTGTTGGCGCGGCTGCGGAGGTAGAGCGGCAACATCCAGTACTTACTGCCACTGCCAGCGCAGTTG	320
241	CACAGAAACCTCTGTTGGCGCGGCTGCGGAGGTAGAGCGGCAACATCCAGTACTTACTGCCACTGCCAGCGCAGTTG	320
241	TTGGCATTGCTACGTA	320
241	TTGGCATTGCTACGTA	320

321	CCGCGTTTGTGCGCACGGACCCCGCGCGGTGTCGCTGCGACTGCGCTGCCCCCTAGGAGCAAGCCACGGGCCACAGGG	400
321	CGCGTTTGTGCGGACGGATTCGCGCGGTGTGCGTGGCGATTCGCTGCTTTTAGGAGTAAAGTTACGGGTTAGAGGG	400
321	<u>TTTGTGTCGTATCGAT</u>	
321	<u>TTTGTGCGTACGGAT</u>	
401	GCARATGTCCAGGTCCCCCGCTGGGAAGGACACACTATACCCCTATGGCAAGCCAGGGTGGG	480
401	TTTATGGTAAGTTAGGGTGGG	480
401	GLAAATGTTTAGGTTTTTCGTGGGAAGGATATATATATTTATGCTAAGTLAGGGTGGG	480

Figura 17: Amplificado de la SEQ ID NO: 15

1	TGCAGGAGAGGTTGGGAAGGGTGGGGGACGGGGCTCGGGGGAGGTCTCCAGGGACTCTAGTAAGCGGGCAAGGCCGCC	80
1	TCTAGGAGAGGTTGGCAAG	80
1	TGACGAGAGGTTGGGARGGGTGGGGGACGGGTTCGGGGAGGTTCGAGGGATTTAGTAAGCGGGCAAGGGGGTC	80
81	GGGAAAGTTTCAGATCCACGGCTGCGGGGCCACGAGCCACCCGACCCGACCCGACCTTCCGTCGACTTCTATTTC	160
81	CGGAAAGTTTACATTTACGATTCGCGGGGLACGAGTLLALLCGAACGCGGATATGTTTTTCGTCGATTTTATTTC	160
81	<u>GGTTATGAGTTTATTGAA</u>	160
81	<u>GTTACGAGTTTATTTCGA</u>	160
161	CTGGGAACGCGCGAAAGCAACCCAAAGTCACACTGCGGAGGTGGCTGGGAGGGAGGTTCAAGGAGTTCGCCGATCC	240
161	CTGGGAACGCGCGAAAGTAAATTAAGTLAGATGCGGAGTCTGCTGGGAGGGAGGTTAAGGAGTTCGTCGATTC	240
161	<u>TTGGGAATGCTGAAA</u>	240
161	<u>AACCCCGCAAGCTAA</u>	240
161	<u>AGGAGTTTGTGTTGATT</u>	240
161	<u>GAGTTTTCGTCGATTT</u>	240
241	TGCTCAATTAACCGGCTTCCGAGCTGGGCGGAGATGGGGCATGCGGGAAGACCCCTGCCCGTGTTCCTCCCGACCGC	320
241	T	320
241	TGCTCAATTAACCGGCTTCCGAGTGGGTCGACATGGGGATGCGGGAAGATTTTGTGCTGCTGTTTTTTTATCGT	320
241	<u>GTTATGCTGTGGGAAGAT</u>	320
241	<u>TATGCGCGGGAAGATT</u>	320

321	CCCAGTGGATGCCATGCCCTGG	400
321	TTTACTCGATCTTATCTTTGG	400
321	TTTAGTGGATGCTTATGCTTGG	400

459	GCTTCTAAGTGTGGGTGATTTCCGGGGGAAGTCCCTGGGGCAGCAGCCAGCATCTGATGCTAGCATCCTATCCATACCA	380
459	<u>GTTTTAAAGTGTGGGTGATTT</u>	380
459	GTTTCTAAGTGTGGGTGATTTCCGGGGGAAGTCTTGGGGTAGTAGTATLTGATGCTAGCATCTATTTATATATTA	380
459	<u>ATTTTCGGCGCAAGTT</u>	380
459	<u>GTGATTTTGTGGGA</u>	380

379	ACCCAGATGTTCTCGGTGCAGTAGGTCACACTGCCGGGGAGAGGTGGCAAGAGACAAATCAACTCTAGTGGAAACGG	300
379	ATLLAGATGTTCTCGGTGAGTAGGTTATATGTCGGGGAGAGGTGGTAAGAGATAAATLAAATCTTAGTGGAAACGG	300
379	<u>ATATTGTTGGGGAGAG</u>	300
379	<u>TATATTGTCGGGGAGA</u>	300

299	GAACCGAGTGGGAACAACAGCCTCACAGCAAGACCGGGGCTCAGACGCCGGCTCTTGGCCTGCTTTGGTCGACCCGAA	220
299	GAACCGAGTGGGAACAATAGCTTCTAAGCAAGATCCCGCTTAGTAGTCCGCTTCTTGGTTCGTCGATCGAA	220
299	<u>AGTAGTTGTGTTTTTGG</u>	220
299	<u>TAGTAGTCGGGTTTTT</u>	220

219	TCCGGGTGTGTGTTTCAATTTCTTTTTCGCCAGAAATATCATTAGGAACTAAACGGCTCTTTTAAAGAGTGTCAAAGT	140
219	TCCGGGTGTGTGTTTCAATTTCTTTTTCGCCAGAAATATCATTAGGAACTAAACGGCTCTTTTAAAGAGTGTCAAAGT	140
219	<u>TTTGGGTTG</u>	140
219	<u>TTCCGGCTT</u>	140

139	GTTACTCTTGCCAGTCACCAATCAGGAGGGCGCCAAAGCTTCCAGAGCCAGTTCGCCACAGCCAGGACIATTCCTAPR	60
139	GTTATLTCTTGLUAGTLCATLAATGAGGAGGGCGTTAAAGCTTTTtAGAGttAGTTtttAtAGttAGGAtTATTtTtTAA	60
59	GCCAGAGGTCGCAAACTTGAACACACACGGAAGCAAAATTGCCCTGGCAGAGGGGAAGAGGCA	-20
59	GTTTGGTAGAGGGGAACAGCTA	-20
59	GtLAGAGGTCGGAAACTTGAATAGAACCGCAAGtAAATTGttTGGtAGAGGGGAAGAGGLA	-20

Figura 19: Amplificado de la SEQ ID NO: 17

1 AAGAGTCCGAGGAAATGTGCCCGGGATTACTGGTATTTTGTGGCTCCGCGAACAGATGCCAACTTGGCTAAGCAG 80
1 AAGAGTTTTAGGAAATCTGTTTTT
1 AACAGTttttAGGAAATGTtttttCCGGATTAtTGGTATTtTGTGGTtttTCGGAATtAAGATgttAAtTTGGtTAAAGtAg 80
1 TTGGTTTTTCGGAATAA
1 GGTTTTTTGGATTAGAT
81 TTCTCGATCTCCGCGCTCGATGTATCCCCCTAGCGAATCTCAGCTGGTGTGCGAGACACAGTCAGCGTCTGCCGGT 160
81 TTTtTGGATtTCGGCGTcGATGTAttttTTAGCGAAttTtAGTtGGTgtTGGCTAGAGATAGtAGTtAGCGTtTGTCCGT 160
81 TTCGGCGCTCGATGTAT
81 TTTGGTGTtTcATGTAtT
161 GCGCGCGCCAGGAGGAGCAGAGGGTCTGGTGAGTAGAA 240
161 GTAGAGGGTTTGGCTGACTAGAA 240
161 GCGCGGtTtAGCAGcAGtAGAGGGtTtGGTgAGTAGAA 240

321	CTGCTTCGGCTGAGAGAGGGATGCTTGAAGTTCAATTGGTCTTTCTGACACACGGTAAGAAAGCAGCGCCAGCTTGGCG	400
321	TTGCTTCGGCTGAGAGAGGGATGCTTGAAGTTCAATTGGTCTTTCTGAGALAGGTAAGAAAGTAGCGCGAGCTTGGCG	400
321	AAAAGTAGCGCGAGCTT	400
321	TT	AGTAGTAGGACTTTGG
401	GTTCTCTTGGCCACITTTCTCACAGTGCTTTTGGGCGTCTTCTTGACTGAATCTGACTCCATTTGGAGGCTGTGTAAAC	480
401	TTTTATGGAGCTTCTCTAAT	480
401	GTTCTCTTGGCTAATTTTtAlACTGTtTTTGGGCGTtTTtTTGAtTGAATLIGALTTtATTGGAGGtTGTGTAAAC	480

1	CTATTATTACAGATGGACTGAGGTTCGACGACTGGGATGGAAGAAGGAATCCCTTAAATTTGGGGGAATTTCTGTTCTC	80
1	TTATTATTTAGATGGAGTGAGGT	80
1	TATTATTTACATGGACTCAGGTTGACCATTCGGATGGAAGAAGGAATLLTTAAATTTGGGGGAATTTCTGTTCTT	80
81	TGTTCTAAGACCAATTTTACTTTGGGGTCTGGGGTGGGGCGGGCGGTTCAGGGCAGTGGAAACGCACTCGCGCTCGCCCATC	160
81	TGTTCTAAGCAATATTTTAATTGGGGTCTGGGGTGGGGCGGGCGGTTCAGGGCAGTGGAAACGCACTCGCGCTCGCCCATC	160
81	GGAAATGTAGTTGTGGT	160
81	<u>GAACGPAGTCGCGGT</u>	160
161	CCTGCACCTTCAGGGCGCTGGAGGACCGCGGGGACCGAGCTCCGACCTGCGGAACTCGGGGAGGACAGACAGG	240
161	ttTGtAtTtTAGGGCGCGGAGGCAATCGCGGGACCGAGTTCGGGALTTTGGCGAALTCGGGGAGGTAGATAGG	240
161	ATTTTGGCCAAATTCG	240
161	<u>TGGTGAATTTGGGGGA</u>	240
241	GGAGGGGACACCCAGCGGCAGCGCTCTCAGCCCTCCCGACCGCGGGCGGCTTTCTCTGACAGTCCAGGAAAGCA	320
241	GGAGGGGCAATCTTAGTCGGTAGGGCTTTAGTtTtTTCGTAGTCGGCGGCTTTTtTtTGAtAGTtTAGCAAAAGGtA	320
241	<u>ATTTAGTTGGTAGTCT</u>	320
241	<u>TTTTCTGATCTCGCG</u>	320
241	<u>TTAGTCGGTAGGCGT</u>	320

321	GACCCCTCCCAGGCCAGGCAGGTAAGACTGCTGTGGAGCTTGCTTTACTGAGSGCGCACAGACCCTGGGGG	400
321	GATTTTATTAAAGTAGTACGTAAGGTAAGAATGCTGTTGAGTTTGCTGAGTTTGGGCGCATAGATLTTGGGGG	400
321		ATTGAGGGCGGTATAGA GCGGAG 400
321		TGAGGGGTATACATTT GGAG 400

401	ACCGAAGCTTCACATCGCGGATTCTGTCGGCTAACCTGGGT	480
401	TTGTGGGGTAATTTGGGT	480
401	ATCGAAGTTTGtAtATCGCGGATTCTGTCGGCTAAttTCGGT	480
401	ATTGAAGTTT	480
401	ATCGAAGTTGT	480

Figura 22: Amplificado de la SEQ ID NO: 20

1	AGATGTTTACATCTGCTGGGAATCCCGAATTTCCAGGCTCCCTGAGAAATCTATCTGCTGGAAGAAGTTGCACCTTCG	80
1	AGATGTTTATATTGCTTCGGA	80
1	AGATGTTTACATCTGCTGGGAATCCCGAATTTCCAGGCTCCCTGAGAAATCTATCTGCTGGAAGAAGTTGCACCTTCG	80
1	<u>TTGTATTTCG</u>	80
1	<u>TTTTC</u>	80
81	TCGCCACAGGAGGCCGCCGAGAGCCAGGAGTGGAGATTTCAAGACTTGGCAGCATCTCTACTGATCTTTAAGCGC	160
81	TCGTTATAGGAGTCTGAGAGATTAGGATGGAGATTTTACATTTAGGATTTGTACGATTTTATCTGATCTTTAAGCGC	160
81	<u>TCGTTAT</u>	160
81	<u>TTGTATAGGAGT</u>	160
81	<u>ATAGGAGTCTGCGAGA</u>	160
81	<u>ATAGGAGTCTGCGAG</u>	160
161	CAACAGAACGAAATGGACATGCCCTATTACCTACCAAAATCGAATTTATTTATATGAGACACAGAAATCTGCTTCAAGTT	240
161	TAACAGAACGAAATGGATATGTTTATTAGTATATCAAAATCGAATTTATTTATATGAGAGAGAGAAATCTGCTTCAAGTT	240
161	<u>TAATAGAACGA</u>	240
161	<u>TAATAGAACGAAT</u>	240

241 CCACATTGTCCTTTAAATATCGACCCGCCCTAACCCCTGTCGCCAAAGGCTGTGGAAATTAGCCAGGGGCTGGGGAG 320
241 TtAtATTGTtTtTAAATAATCGAtTtTtTtAAAtTtTGtTtTtAAAGGtTCGtGGAAATTAGGGAGGGGCTGGGGAG 320

321 ACGGTTCCAGAAACAAACGGCTTTCCTCCAGCAGGGCCTCTACATCCAGTCCACCCCAAGCTGATTTTGTAACTCATATA 400
321 ACGGTTtLlAGAAALAAAACGtTtTtTtTAGtACGgtTtTtAtAtTtTAGTtLlAtttAAGtGtGtTtTtGTAACTtAtATA 400

401 AAGAATGGGCCCCCGCTGGTCAGTGAAGACTGGGGTGAGCA 480
401 TAGTGAAGATTGGGGTCAGTA 480
401 AAGAATGGCtTtCGGtGGTGGTtLlAGTGAAGALTGGGGTGAGLA 480

Figura 23: Amplificado de la SEQ ID NO: 21

1	TAAATAATGGACGGGAAGACAGATCCCAATTTGNAATCCCTCCTTGAGAAAAACAATCAAACTAGTTCTCTGGGGAAGAAA	80
1	TAAATAATGGACGGGAAGATAGA	80
1	TAAATAATGGACGGGAAGACATAGATTTAAATTTGNAATTTTCTTGAGAAAAATAAATLAAAAATAGTTTLLTGGGGAAGAAA	80
81	GCCTCAGCTAGGTCAGGAGGAAACTTACGCATCTTTGATTCCTTTCCGCTTTGGAGAGGCATTACCCGTTCAAGCCCCAG	160
81	GTCTTAGTTAGTTAGGTTAGGAGGAAATTTACGTAATTTTGATTTCTTTTCGCTTTGGAGAGGATTTATTCGTTTAAAGTTTAA	160
81		AG 160
81		160
161	CCAAATGCGGCCCGCCAGGATTTACAGCTCTTATCAAGGGTACATTTTCGGAAGATATAAAGAAAGGAGTTTCCCTAACCA	240
161	TTAAATGCGGTCGGTACGATTATAGTTTATTAATCAAGGGTTAGATTTTCGGAAGATATAAAGAAAGGAGTTTCTTAAACA	240
161	TTAATGTTGGTTGGT	240
161	TTAATGCGGTCGGTTA	240
241	CACCGGGTCTTCTCACACCAACACAGACACCGAAATTAAGCCTTCTATGGTTTGTCGTAGGACTTTTGTAGATAAAACT	320
241	TTATCGGGTCTTCTTATAGTATCAGATATCGAAATTAAGCTTTCTATCGTTTGTCGTAGGACTTTTGTAGATAAAACT	320
241		320
241		320

321	GGCAGCCCCATTAGATAAGAAATGGCTTTTACAGACCTTTGGGGGAAGCCGGGGGTTCTTCGIGGCCCTTCCTGTGACTG	400
321		
321	GGTAGCTTTATTACATAAGAAATGGCTTTTACAGAGCTTTGGGGGAAGCCGGGGGTTCTTCGIGGCCCTTCCTGTGACTG	400
401	TCCTTTGGGAGAC	480
401	TTTTTGGGAGAT	480
401	TTTTTGGGAGAT	480

Figura 24: Amplificado de la SEQ ID NO: 22

1 TTCTGCTCACTCTGAGATGGCAGATTGGGGTCCCTGTGTGTCCTTGGACAGATGAGAATGCCGAGAGCTCGTATGCCTTGC 80
1 TTGTGGTTATTTTGAGATGGTA 80
1 TTCTGCGTTATTTTGAGATGGTATTTGAGATGTCGAGAGTTCGATGTCTTGT 80
1 ATCTCCACACGTTCCGA 80
1 GTTCGACGTTTGTATGT 80

81 CCAGGGCCACACAGCAAGGCCGGTGCCCATGCGGTGTCCGAGGACCCACHTGACCCCTGCTGTCCCTCAGGCCACAT 160
81 ttAPGGGLALALAGtAAGStCGGTGtLlATGCGGTGTtLlTAGGAtttAtTGAAtttTGTGTtLlTlAGGttAtAt 160

161 TAGGATCTCAGACCTCGGCTTGGCTGTGAAGATCCCCAGSAGGAGACCTGATCCCGGCCGGGTGGGCACtGTGTGGCTACA 240
161 TAGGATtLlLAGAtttTGGGtTTGGtTGTGAAGATtLlLcAGSAGAGAtttGATtCCGGtCCGGTCGGtAtTGTGGtTAtA 240
161 AGATTTTCGAGGGAGA 240
161 AGATTTTCGAGGGAGAT 240
161 ATTGATTTCGGTTGG 240
161 TTGATTCCGGTCCG 240

241	TGGGTGACTGCTCGGCTGCTGCTCAATGCACCTTGAGACCCACACACCTGCTCACCGCGCGCCGAGCCCCAGAGCC	TG	320
241	TGGGTGAGTGTGTGGGTGCTTGGTGTGCTTAAATGTAATGATTTGAGATTTATTTATTCGTCGGTCCGACGTTTTACAGT	TTG	320
241		<u>TTTATTGTTGGTTGAGT</u>	320
241		<u>TATCGTCGCTCGACCTT</u>	
321	CCGCGAGAGCCTCTCGGCTCTTCTGCACTGCTGATCACACAGAGTCACAGTCTCTGCGAGGCTGCTGGGGGGGGGG	CCCC	400
321		TT	400
321	TCGCGAGAGCTTCTCGTCTTTTATGCTATGCTGATTAATATAGAGTTATAGTTTCTGTGAGGGTGTGTGGGGGGGGG	GGTT	400
401	TTAGCCAGGGGGAGGG		480
401	TTAGTTAGGGGGAGGG		480
401	TTAGCCAGGGGGAGGG		480

Figura 25: Amplificado de la SEQ ID NO: 23

1	TCGACTCACCAGGTGAAGAGTTTGCCTTTTATCCTGCGCAGCAGGCTGGAGACTCGGCCGCGCTCCCTCCGCGAGGGG	80
1	<u>TTTTTTATTAGGTGAAGAGTTTGT</u>	80
1	TttAtTlAtttAGCTGAACAGTTTGTtTtTTTATtTtTGCGLAGtAGGtTGGAGAtTCGGtCGtCGtTtTtTtTtCGGAGGGG	80
1	<u>AGATTTTGGTTCTTTT</u>	80
1	<u>AAATCGGTCGTCGTTT</u>	80
81	CTCTGGGGCGCCGGCCCGCTGGCGAGTCTGCCCTCCATCGGGGCCCTTGCGCGTCTCCGGCGCCCAAAACCCCGAAGT	160
81	tTtTGGGGCGtCGGGGtTCGtTGGCGAGtTtGtTtTtLAtCGGGtTtTTCGGCGTtTtCGCGtTtAAAtTtCGAAGT	160
81	<u>TTTTTGGTGTATAAAT</u>	160
81	<u>TTTCGGCCTTAAAT</u>	160
81	<u>ATTTTGAAGT</u>	160
81	<u>TTTGAAGT</u>	160
161	QCCCGCTCCCGAAGACGGAGGGCGTGCCCTGTGCGGCGGCTGGGCTCCGCTCTGGCCCTGCGCTCCCTCCCTCGCTC	240
161	GtCGGtTtCGGAAGACGGAGGGCGTGtTtGTGCGGtCGStTCGGGtTtCGTtTtTGGtTtTGGGtTtTtTtTtTtTtTtTt	240
161	<u>TTTCGGAGACGGGAG</u>	240
161	<u>TTTTGGAGATCGGAG</u>	240
161	<u>GTCCG</u>	240
161	<u>GTGGGT</u>	240

241	TCCTGGCTGGCGTCCCTCTCGCTCCCTCCGCTCGGCTCCCTTCAGCCACCAATCATTAACCTGTCAAGCCCTAAGGGCCAGACA	320
241	TTTGTCTGGCGTllllTTCGTttttTCGltttTCGltTCGgtllttTAGCGATTAAITATTAAATGTAAAGtttTAAGGGtTAGATA	320
321	ATACATTTCAGAGATTACAACCTGGCACCTCCAGGGTGGCAGGAGGA	400
321	TTTAGGGTGGTAGGAGGA	400
321	ATATATTGTAGATTATAAATTGGGAtttTtTAGGGTGGTAGGAGGA	400

Figura 26: Amplificado de la SEQ ID NO: 24

492	TTGGAGAGATGTGCTGGCCAGGACCGCTGACACAGGTAATTCGGTTGGCCCCCACTCACCCGGTGCCCGAGAGCGAG	413
492	TTGGAGAGATGTGTTGGTTAG	413
492	TTGGAGAGATGTGTTGGTTAGGATLCGCTGATCAGGTAATTCGGTTGGCTTTTATCTTATTCGGTGTTCGAGAGCGAG	413
492		TGTTTCAGAGTGA 413
492		GTGTTTCAGAGCGAG 413
492		TATTTCCCTGTTTCAGAG 413
492		TTTATTCGGTGTTCGA 413
412	TGTGACTGTATCGATTGTTTCAATTGATTTTTTTAAAGCTAGATPAGCAGAGGCGTGGTGTCTGTACACCGAGATATATAC	333
412	TGTGACTGTATCGATTGTTTCAATTGATTTTTTTAAAGCTAGATPAGCAGAGGCGTGGTGTCTGTACACCGAGATATATAT	333
412	TGT	333
412	T	333
332	ATACACAAATACAGATACTCAAGCGCACACAGATTACACGTACAGAGTTCACACACAAGGACACACACACACATCGTGG	253
332	ATATATATAATATAGATATTTAAGCGTATATAGATTATACGTACAGAGTTCACAGALAAGGALATATATATATATCGTGG	253
332		ATTGTCG 253
332		TATCGTGG 253

252	TCCCTCAGGCACACATACCTCCAAAGAGCCCTTGGTAACCTGGTGAAGTCCCTGAGTGTGCCCAAGATGTGTGTGGG	173
252	TttttACGTAtAtAtAtttttAAAGAGcttTGGTAActTGGCAGGttttTGAGTgtTgtttAAAGATGTGTGTGGG	173
252	<u>TttttTATGTATA</u>	173
252	<u>TTTTTACGTAT</u>	173

172	TGGGACGCAGGGACGGAGAAAGGGAGTGTGAAGCAGGATTTCCAGCAGCTCAAGTCCCGCTGATTCCTCTGCCCCCGA	93
172	TGCCCCACctAGGGACGGAGAAAGGGAGTGTGAAGTACGGATTTtLAGLAGGtLAAGTgtCGTGTGATTTtttTGcttttCGA	93
172	<u>GACGTAGGGACGGAGA</u>	93
172	<u>GAATGTAGGGATCGAAG</u>	93

92	CCCCAAATCCTTCCCTTTCTCCTCTAGGGAGGTCGTGAACCAAGAAATCCCAAPCCTGGGCTTGTttCCACATCCACAGAG	13
92	ATTATAGAG	13
92	ttttAAATttTttttTttTttttTAGGGAGGTTGAATLAGAAATtttLAPAtttTGGGtTTGTTtttAtAtttAtAGAG	13

12	GGGAGGAGTCCT	-67
12	GGGAGGAGTTTT	-67
12	GGGAGGAGTttt	-67

Figura 27: Amplificado de la SEQ ID NO: 25

407	TC	TCCAGAACAAAGGACTTTAGGGCCCAATTCGGTTTATTCAGTACTCCAGTCCTA	AAAACITGGAATATCTGATGAA	328
407	TTTTAGAAATAAAGGATTTAGGG			328
407	Tt	TtTtAGAAATAAAGGATTTAGGGCtttAAATTCGTTATTtAGTALTtAAAGTtTt	AAAAALTtGGAATATtTGATGAA	328
327	TAAAAGTGGCGGCTCCCGAGGCTGTCTCTTGAGAGAGGCCACGGSCACAGCTGACCTTGCCCGCTCCCATCGCGTCACTGA	248		
327	TAAAAGTGGGTCGtTtTtLAGGtTGTtTtTTCAGAGAGAGtTtACGGtTtAGTtTGAltTtTGtTCGtTt	ATCGCGTtAtTGA	248	
247	CCGCTCCTCAGACAGATGCGTCAGGCA	TCTCCGGCGGCGGCTCCTTGGCGCAGACTCGCTGCAGCAGCGGCAGGCTT	168	
247	TCGtTtTtTAGATAGATCGGtTAGGtATtTtCGGCGGtCGtTtTtATtTtTGCGLAGATTCGtTGLAGtAGCGGLAGGtTt	168		
247		TGGTGTAGATTTGTCT	TACCCGCTACGCTT	168
247		TTTGGCTTAGATTCTGT	TAGTGTAGGTTT	168
167	CGCACACATCCCGCTGAGCATGCGGCGCCAGCCTCCCTCTCGCGCCCGGACGCGTGCTTGT	TTTGGCGCAGTGCAGGGG	88	
167	CGtAtAtATtTtCGtTtCGtTtTGAGtATGCGGtTAGtTtGtTtTtCGGtCGGtAGGCGtNGtTTGT	TTGTTCGtAGtStAGGGG	88	
167	CGTA	AGTATCGCGGCTTAGTT	TTTTTCGCTTGTAGG	88
167	TCTA	TGAGTATGTGTAGT	TTTTTGGGTCGGTA	

Figura 28: Amplificado de la SEQ ID NO: 26

1	TCCAGGAATGCAGGACACAAAGGGGCGAGCCTACCTCAGGGCAGGGCCAGCGGGGTTCCCTCATCCTTAGGGCCAGTCCCGG	80
1	TTTTAGCAATCTAGCATATAACGG	80
1	TtLtAGGAATGLAGGALATAAAGGGGTAGTtTAtTtTAGGGtAGGGtTAGCGGGCTTtTtLtAtTtTTAGGtAGTtLtCG	80
81	GGGCTCAGATGATTCAGGGAGCGCCACCGCGCTGACTGCCACCTCCACATCACTCTGCAGAGCGGTTGAGCATGAAA	160
81	GGGCTLAGATGATTLAGGGAAGGTCGATTCGGTGTGATtGtLtAtTtAtLtAtTtTtTGCAGAGCGGTTGAGtATGAAA	160
81	<u>AAGTTTCATTTGCTTG</u>	160
81	<u>GTGGATTTCGGTTGATT</u>	160
161	GACCACTGTATCGCACAGAGCCTCCTAGAGAAACAGCAGATCTACCTGGAGATGGCCGAGATGGCGGGCTCGAAGACCT	240
161	GATtAGTtGATtGATtGATtAGAGtTtTtTAGAGAAAtAGtAGATtTAtTtTGGACATGGtCGAGATGGCGGGtTCGAAGAtTt	240
161	<u>AGATGGGTGGTTTGA</u>	240
161	<u>GGCGGTTTCGAAGATT</u>	240
241	GCCCCAGCCCCGAGGCTTATTCCTGGAGGGACCCATCCGACACCCCTCCAGGGGAGGTAATTCTCAAGTCGGCCATGA	320
241	GtLtLAGtLtCGAGGtTtTtTtCCTGGAGGGGAtTtTtTCGAGAtTtTtGAGGGGAGtTtTtTtAAGTCGGtTtATGA	320
241	<u>AAGTCGGTTATGA</u>	320
241	<u>AAGTTGGTTATGA</u>	320

321	GCGAGAGTAAGTTGGCTGCCACACCTCAAGGTCAGTCTTGCCGGGGTGGGCTCCTCAGGGGAACCCAGGCCAGGCTC	400
321	GCGAGASTAGTTGGTIGTtttAtAttttAAGGTCtAGTtTTGtCGGGGTGGGtTtTtTAGGGAActtttACGttAGGtTt	400
321	<u>GCGA</u>	400
321	<u>GTGA</u>	400
401	ACGGCTCATTGTGGCCGACACGGGACCCCTGTCCAGGACAGGCTTCTATGTGGGG	480
401	GGATAGGTTTTTATGTGGGG	480
401	ACGGTtTATTGGGtCCGAtACGGtAttttTGTtLAGGAtAGGtTtTtTATtCTCGGG	480
401	<u>TTGTGGTTCATATCGT</u>	480
401	<u>TGGTCCGATACGGTATT</u>	480

Figura 29: Amplificado de la SEQ ID NO: 27

273	CAGGGTCTGATGATGATGATTTTACACCCCGCCGATGTGAAGCCACAGACAGATCATTCAGCTCACGCTGAAGGGCACA	194
273	TACCGTTTGATGATGATGATTTT	194
273	TAGGGTCTGATGATGATGATTTTATTTTCGTCGATGCTGAAGGttatAGAGACAGATTATTLAGttTtACGTTGAAGGGtATA	194
273	<u>ATTTGTTGTATGTGAAG</u>	194
273	<u>ATTTCGTCGATGTGA</u>	194
193	CGGGCGTCAAGGCGACAGACAGAGAGGAGCGGCACAGGATGCTGTGATGCTGGCTCAGGCTGGGCATCCGAGTTGATG	114
193	CGGGCGTtTAGGGtAGtAGAGAGtAGAGAGGACCCGATACCGATCTGTGATGCTGGGtTtAGGtTGGGtATtCGAGTTGATG	114
193	<u>AGGACGCGATAGGGAT</u>	114
193	<u>AGGATGTCATACCGAT</u>	114
113	AGCGCTGGCTGGCGTTTCAAGGCGCTGGAAACTGACACGCGCGCTGGCGCGCTTCGCAATTAAAGGAAGTGCAGCCTCATC	34
113	AGCGtTGGtTGGCGTtTAGCGtTtAGCGtTtTGAAAtTGLALACGCGCGTGGGtCGTtGGGtATtTAAGGAAGTtAGtTtCATC	34
113	<u>GGTtGTTTCTGTATTTAA</u>	34
113	<u>ATTGTATACCGCGCTT</u>	34
113	<u>GTGTTTGGGTATTTA</u>	34
113	<u>TTGATATATGtGtGtTtG</u>	34
33	GGAAAGAGCCCTGCCAGTGTGCTGAGCCAGGAGGA	-46
33	TTAGTGTTCAGTTAGCAGCGA	-46
33	GGAAAGAGGtTtGtTtAGTGTtTGAGtTAGGAGGA	-46

Figura 30: Amplificado de la SEQ ID NO: 28

1	GAGAGGGAGGGCCCAAGGTCGGGCGCCGAGAGGAGGGGCGCGGGGGCTCGAGGGGGCGAGCCGCTAGCGCGCTTT	80
1	GAGAGGGAGGGCTTACGTT	80
1	GAGAGGGAGGGGCTLAAGGTCGGGCGTCGAGAGGAGGGGCTCGGGGGGCTGGAGGGGGCGAGGGCTTAGCGCGCTTT	80
81	GCTTCGGTCCCGTCAATGCGCGCTAGCTCTGTTTGACTGCCTGTACAAAGCGAGTAGACGGGGGGGATGGGGCGA	160
81	GtTtTtCgTtTtCgTtAAtTgGtCgTAgTtTgTtTtGAtTgTtTgTtAaAGtGAGTAGACGGCGGGGGATGGGGCGA	160
81	<u>GTAAATTTGGTTGTTAGTT</u>	160
81	<u>ATAAATTTGAGTAGATGG</u>	160
81	<u>ANTGGTCGTTAGTTTT</u>	160
161	TTACGGACCGGATTTGGGAAGGAGGGCGACAGC	240
161	ATTGGGAAGGAGGTTAGTAGT	240
161	TTACGGACCGGATTTGGGAAGGAGGAGGtAGtAGt	240
161	<u>TTAGGGATTCGATTTC</u>	240
161	<u>TTACGGATCGGATTTC</u>	240

Figura 31: Amplificado de la SEQ ID NO: 29

356	GTGGAAAAAGGAGGCAAACTTCCTCAGCCCGTCTCAGCTCAGTCCCGAGGAGGAAAGAAACAGGCCGAGTGAAGGT	277
356	CTGCAAAAGCAGAGTAAATTG	277
356	GTGGAAAAAGGAGAGGCAAACTTCCTCAGCCCGTCTCAGCTCAGTCCCGAGGAGGAAAGAAACAGGCCGAGTGAAGGT	277

276	GCTGGAAAGGAGGAGGACGCGAGGGGAAAGCCCTGTGGGAGCCGAGGGCGTCTAGAGAGACCCGGGAAGGAGGCTCT	197
276	GLTGGAAAGGAGGAGGAGGACGCGAGGGGAAAGGCTGTGGGAGCCGAGGGCGTCTAGAGAGATCCGGGAAGGAGGCTCT	197
276		197
276	<u>AGTCGAGGGCGTTAGA</u>	197
276	<u>ACTTCAGCGCTCTTACA</u>	197

196	CGGGTGGGGAGCCAGGAGACCTGCTCTCGGGGAGAGAGGCGGGGCCAGCCGCTCTCCTGGACGCCGCCCGCCGACCA	117
196	CGGGTGGGGAGCTAGGACATCTGTCTCGGGGAGAGAGAGGCGGGGCTTAGCGCTTCTTTGGACGCTTTCGCTCGTACA	117
196		ATA 117
196	<u>TTTTTGGGCTAGATA</u>	TATA 117
196	<u>FGTTTTTTGGTGTAGAT</u>	

116	GCTCCCGGCGGGTGCTCTCAGGGCTCAGTCTCAGAGCCCAACCCAGCATCCCGGCCCTTCCTTCCCGAGGAACTCGCC	37
116	GtTtTCGGCGGGTGCTCTCAGGGCTTCTACTATCTGAGCTTATTTAGATATTCGGCCGCTTCTTTTTCGAGGAATTCGCT	37
116	GTTCCTTCGGCGGGT	TTTTTTGAGGAATTTGPT 37
116	<u>CTTTTTCGGTGGGT</u>	<u>TTTTTCGAGGAATTCGT</u>

36 TCAGCCGTCATCAGGCTTCCTGGTGAGAACTCAGGAG	-43
36 TTTGGTGAGAATTGAGGAG	-43
36 TTAGCTTGATTAGGCTTTCTGGTGAGAATTGAGGAG	-43

20 GCCTGCAGGAAGCAGGCC
20 GTTTGTAGGAAGTAGGTTT
20 GtTTGTAGGAAGtAGGgttt

Figura 33: Amplificado de la SEQ ID NO: 31

```

1 AAGGCACTACTCTGCGGTCCCGAAGCGGATCCCTGGCACTGGGACCTGGGACGCGGTGAGAGCGAGCCCTGAAGCTG 80
1 AAGCTATTATTTTTGGCGTTT 80
1 AAGGCACTACTCTGCGGTCCCGAAGCGGATCCCTGGCACTGGGACCTGGGACGCGGTGAGAGCGAGCCCTGAAGCTG 80
1 TTTTTCGARAGCGGATT 80
1 TTTGAACTGGGATTTCG 80

81 ASTCCAGAGGCTGGGGCGTCATGTCACGGGCGGACCCAGCCGACAGCCCTACAACTAAACCCAGGCCAGCCGAGAA 160
81 AGTttAGAGGCTtGGGGCGCTtATTGtttAGGGCGGAtttAGtttLAGAtttTALAAtTAAAtttAAGGtttAGtCGAGAA 160
81 AGGTTAGTCGAGAA 160
81 AGGTTAGTCGAGAA 160

161 GCACCTACCACTGGCGGCTGAGCTGAGCTCTCATCTTCTTAGGACTCTGGACCTTTGGAGGAAATGAAGGGAUCCAGT 240
161 GEAATATtAAtTGCGGGGtGAGtTGAGtTtTtTtTtTAGGAtTtTGGAAtTTTGGAGGAAATGAAGGGAUCCAGT 240
161 GT 240
161 GTA 240

```

241	TGGAATATAGTC	TTTCTTCGGCCCCCTACATCTCTCTCCCGCGGAAACGCGGCCCCAACTGTTACGAACTGGAG	320
241			AACTGGAAG 320
241	TGGAATATAGT	TTTTTTCGGCGTTTTTATATTTTtttttCGGCGGAAACGCGGtttAAATCTTTACGAACTGGAG	320
241			<u>GGAAATGIGGTTTTAAAT</u> 320
241			<u>GAAACGGCGTTTAAAT</u> 320
241			<u>ATTGTTACGAACTGGA</u> 320
241			<u>TTGTTATCGAACTGGAG</u>
321	AGGCTGGATAGCT		400
321	AGGTTGGATAGTT		400
321	AGGTTGGATAGTT		400

Figura 34: Amplificado de la SEQ ID NO: 32a

1	GGGTGCTGGGAGCCCCATCCGCCACGAGGTGGATAATTTAGGGGGCAGGGGCTCGCACAAAGGGCGAAGATGCGCTCTC	80
1	GGGTGTTGGGGAGTTT	80
1	GGGTGCTGGGCGAGCTTCTATCGGCTAGGACGTGGATAATTAGGGGGTAAAGGGGCTCGTATAAAGGGCGAAGATGCGGTCTT	80
81	CTCCCGGCACCGACCCACCGCTTTAAAGGCTCTGCGGCAACGTTTAGGACGCGAACTAGGTGCCGGGGTTGGGACG	160
81	tTtCGGtAECGAtttACGGGTTTAAAGCtttTCGGGtAACGTTTAGGGACGCGAACTAGGTGtCGGGGGTTGGGAGG	160
81	<u>TTTGGGTAACGTTT</u>	160
81	<u>TTCTGCTAATGTTTACG</u>	160
81	<u>TTAGGGACGCGAATTA</u>	160
81	<u>ACGGATGTGAATTAGG</u>	160
161	AGGTGGGACTGGAAAGCGGAGAGTCTGCGCCCAAGACCGCGCAAGAGGTGGCCTTTTGTTCGCTGGCTCCAGCACATCC	240
161	AGGTGGGAtTGGAAAGCGACAGtAGTtTGCgttTAAGAtCGCGtAAGAGGTGGtTTTGTTCGtTGGtTtCAGtAtTt	240
161	<u>TGCGTTTAAAGATCGCGT</u>	ATATTT 240
161	<u>TCGTGTTTAAAGATTGTGT</u>	ATTT 240

[illegible]

Figura 35: Amplificado de la SEQ ID NO: 32b

317	AGACTGTGGTACCAGATCTGGTTCCATTTTACGAGAGCCGAGAACCCACGACGTCCTGGAGGGAGAGCGTAGCTGACCCG	238
317	TCCTTTCC	238
317	AGACTCTCGGTATACAGATCTGGTTTtATTTTACGAGAGtLAGGAAtLACGLACGTTTGGAGGGAGAGCGTAGCTGATTCG	238
317		
317	<u>AGGCGTAGTTGATTCC</u>	238
317	<u>AGGATTTACGTACGTT</u>	238
317	<u>AGGATTTACGTACGTT</u>	238
317	<u>TACGTATCTTTTCGACCC</u>	238
237	AGGGCGCGACCGTGTTCTGGGAGTGTGAAATCCCGCCCTTTGCTGAAATTTTCTCTCTCCGCAAGCCCGAGAGGAGA	158
237	AGGGCGCGATCGTCTTCTGGGAGTGTGAAATCCCGCCCTTTGCTGAAATTTTCTCTCTCCGCAAGCCCGAGAGGAGA	158
237	A	158
237	A	158
157	TCGTGTGATGTCGGGGGGTCTCTCGTGGCGCGGCTGGGTAGCGCGCGGACCACCCTAAACCGGGGCACTGTTGCC	78
157	TCGTGTGATGTCGGGGGGTCTCTCGTGGCGCGGCTGGGTAGCGCGCGGACCACCCTAAACCGGGGCACTGTTGCC	78
157	<u>TACCGGTCCCGGATTA</u>	78
157	<u>GTAGTGGTTGTGGATT</u>	78

77	GTCCCTTCCTGGTCTCTGCCCCCTCTGTAAACGGCGGGAAGTAGGAGGGCGGGGAGNAATTACTAAAGGATGCTTGC	-2
77		
77	CTttttTtTtTGCTGCTTCTGTGTTTTTGTGTACCGCGGGGAGTAGGAGGGCGGGGAGNAATTACTAAAGGATGCTTGT	-2
77		
77	<u>TTTGTACCGCGGGAA</u>	-2
77	<u>TTTTGTATGTGGGA</u>	-2

Figura 36: Amplificado de la SEQ ID NO: 33

```

1 TCAGCTGAGAAAGTGGCGGCCCTAAAGAGGCCCTTTGTTGATAGAAAGGACGCTCAACCCAGAACCTAGAGGGTG 80
1 TTAGTTGAGAAAGTGGGGGT 80
1 TTAGTTGAGAAAGTGGGGGTTTAAAGAGGCTTTTGTGTTGATAGAAAGGACGCTTAATTAACGAAACCTAGAGGGTG 80
1 ATCGAAATCGTAGACG 80
1 ATTGAAATTTGAGAGGG 80

81 CGGCCCTGGCGCTTGAGCGGTGACACCATCCATGGCGGTGACCGCTCTTGGCGTCTCTGTATAGGTCACGGC 160
81 CGGCTTTGGCGCTTGACCGCGGTAGATTATATATGGCGGTGATCGTCTTGGCGTCTCTGTATAGGTCACGGC 160
81 TTTATCGTGGTCAATTGT
81 TATGGCGGTGATCGTT
81 TTATGGT 160

161 GTCCCGGATCACGTTCTCCAGGACACCTTCACGACCCCGGAGTCTCCCTGTAGATGAGCGCGGAGATGCGCTTCACGC 240
161 GTTTCGGATTACGTTTCTTAGGAAATCTTLAGTATTCGCGAGTTTCTGTAGATGAGGTCGAGATCCGCTTTACGT 240
161 GTTTCGGAT
161 GTTTTCGATT
161 ATGCGTTTTCCT 240

241 CGCCCGCCCGAGCAAGCGCGCGATGGCGGCTTGGTATGCCCTGGATATTGTGCGGAGTACTTTACGGTGGCGCTTA 320
241 CGTGGCGGAGTAAGGGGTGGATGGTCCGTTCGCTGATGCTTTCGATATTCTCGCGAGTATTTTACGGTGGCGCTTA 320
241 TGT
241 CGT
241 GGATATTGTTGTGATG
241 TATTGTCGCGTAT

```

321	GGCCGGCCCTTGGCCAAAGACCCCTTCCGGCCCTTGGCCGGCCAGACATGACGAGCAGCAGCAGCTCACCACACGCTTGT	400
321	GGGtCGgttTtTgttAAGAtttTtTtCGtTtTtTtCGGGtLAGAtTtGACGAGtAAGAGGAGTtTtAtttAACGtTtTGT	400
401	GAGGACTCTGGCCCTGAGGCAG	480
401	GAGGATTTTGGTTTGACCTAC	480
401	GAGCAAtTtTGgttTGAGGtAG	480

Figura 37: Amplificado de la SEQ ID NO: 34

[illegible]

18	CATGCAGCTGACTGGGGC	-61
18	TATCTAGCTTGATCGCCCT	-61
18	TATGTAGTTCATTCGGGt	-61

Figura 38: Amplificado de la SEQ ID NO: 35

1	ACCTCAGCTGTAAAGCCCAAGGCTGCTGCCTGCAAAAGGCAATGCCGTGGAGACTGGGTTCCACAGCGACCTGGGTTTTTAA	80
1	ATT TTAGGTGTAAAGTTTAAGGTGTGT	80
1	AttTtAGGTGTAAAGTtTtAAAGGtTtGtTtTGLAAAGtAAATGtCGTGGAGAtTGGGTTtLAtAGCGAtTtTGGGTTTTTAA	80
81	GTCACCGAAAGCTACAGGSCAGGGTCACAAACACTCTCCCGCCTCTCTGCAGAACCGTCCGGCTGACAGGAGCGTTG	160
81	GtTAtCGAAAGtTAtAGCGStAGGGTtAtAAAtAtTtTtTCGtTtTtTGLAGAtCGTGGGtTGAAtAGGAGAGCGTTG	160
81		160
81		160
	<u>TGTAGAtTtTCTGTGCT</u>	
	<u>AATCGTGGGTTGATA</u>	
161	CGCAGAAATTCGAGCGCGCGCTCGAGGAGTCTCCCGCGCCCGGAGAGGACAGAGCGCCCTGAGAGGCGCAGCT	240
161	CGtAGAAAtTCGAGGtCGGGCGtTGGAGGAGTtTtCGCGGtTCGAGAGAGGAtAGAGGCGtTtTtTGAGAGGCGtAGTt	240
161	AAAATTCGAGGTCGGG	240
161	AAAATTCGAGGTGGG	240
	<u>TTTGTGGTTTCGAGAA</u>	
241	GCAACAGCCGATGCACGSGTTCGGATCTGGCCGGCCAAACCGAGCCAGCGTCCGAAACACAGAGTCGCCACAGAGGT	320
241	GGAAtAGGCGAtGtACGGGtTCGGAtTGGtCGCtAtAAtCGAGtTtAGCGGTCGCGAAAtLAGAGTCGtLAtAGAGGT	320
241	AATAGCCGATGTACGG	320
241	TAGGTGATGTATGGT	320
241		320
	<u>TTTGGTTGGTTAATTGA</u>	
	<u>TGGTCGGTTAATCGA</u>	
	<u>AGTTTAGTGGTTGTGA</u>	

321	GAGCACGATTCCATTTCCTGGGGATCGGGTCCGGTGGGGAGCCACTGTCCCTAACGCCCTCCAGAGACTTTTAAGTAGACAA	400
321	GAGTACGATTtTATTTTCTGGGGATCGGGTtCGGTGGGGAGcttAtTGTtTTAAAGctTtTtAGAGAtTTTNAAGTAGCATAA	400
401	TATAATGCTGCTAGGAGCAAGGIGCAGGAATTTAGCGGCAGAAAGCCTTTCGGGGGGGTGGGGAAGGCAGAT	480
401		480
401	TATAATGLTGGTAGGAGLAPGGIGLAAAGAAATTTAGCGGTAGAAGctTTTCGGGGGGGTGGGGAAGGCtAGAT	480

Figura 39: Amplificado de la SEQ ID NO: 13

488	TGTCCTAGGTC TGATGGACCAAGACCCAGCGGGCCCTGAACCTTTGTCGGAAAGCGGAGCCGCCAGGCCATTTGGCAGCTGCC	409
488	TGTTTACGTTTGA TGGATTAGA	409
488	TGTTTTAGCTT TGATGGATTAGATTTTACGGGGLTGAALTTTGTGGAAAGCGGAGTCGTTAGGTATTTGGTAGTGTG	409
488	<u>TTTGTGGAAACCGGA</u>	409
488	<u>TTTGTGGAAACCGGA</u>	409
408	AGAGCGCCACCAAGCCCGGTCTCAACCTGGGAGGTCAGCGGGCCCAAGCGCCCTGGGAGGCAGTGGCCCCACCTGGACG	329
408	AGAGCGGtAtTAGGTCGGGTtTAACTTGGGAGGtTAGCGGtTAAGCGGtTTCGCCAGGtAGTGGtTtAtTGGACG	329
408	<u>TATTTGGACG</u>	329
408	<u>TATTTGGATG</u>	329
328	CCCGGCCGTCGGGGCTCGGAGCCCGCTCTGCTCCACCAAGCTGTCCGCTGGCCCTACGCCGCCACCGCAGGAGG	249
328	TCGGGtCGTTCGGGGCTCGGAGtTtCGTtTtCGtTtTtAtTAAGtTGTtTtCGTTCGGtTtTACGtCGtTtAtTAGGtAGGAGG	249
328	<u>TCCGGT</u>	249
328	<u>TTTATGTTGTATTAGGT</u>	249
248	AGCGGCTACCAAGTACCTCAGAGTGGGACTCTCCATCCTCAAGTCTGGCTCCAGAAAGTGTAGCAAGCTCCACCGCGG	169
248	AGCCCTATtAGTAtTtAGAGTGGGAtTtTtTtATtTtLAAGTtTGGGtTtAGAAAGTGTAGtAGtTtTtACCGCGG	169
248	<u>TTTATGTTGG</u>	169
248	<u>TTTACCGCGG</u>	169

8 TCGCGAGC	-71
8 TGGGGAGT	-71
8 TCGCGAGt	-71

[illegible]

321 CCCCTTCCCAATCCCAATCTCCACAGCCTTCATCCCTCCACATTTGGATTCACCTTGGCCACCGACGCCCTGGGCT 400

321 TTTTCTTLLAATTTTAAATTTTACAGCTTTCTATATTtTtTtAATTTTCGCTtAtCGACGtTtTtTGGtT 400

321 ATTTCTCTTATTGATGT 400

321 TTGGCTTATCGACGTT

401 TCGTTTCAGCAAGTTTACCCCAACCATACCTCTCGCATAAAGCCCTGCATTTACCGGTCAAGAGGGGAACCAA 480

401 TTATTAGCTTAAAGCGGGAATTAA 480

401 TCGTTTCtAGTtAGTtTtAttttAtAtTtAtTtTtTCGtATAAAGtTtTGLtTTtAttAGGtTtAAGAGCGGGAAtTAA 480

Figura 41: Amplificado de la SEQ ID NO: 37

1	GTGTCAGGGCTCAGGGGCCCATGCCCCCTTGGGCTGGGGCCCCCTCCGTGCGGTCGGGATTGTGGGACGGCGTGAAGGAAG	80
1	CTCTTAGCGTTTAGGGGTTT	80
1	GTGTTAGGGGTTTAGGGGTTTATGCTTTTGGGTTTGGGTTTCTGTCGCTCGCGATTGTGGGACGGGTTGAAGGAAG	80
1	<u>TCTGGGATGTGTTGAA</u>	80
1	<u>CTGGGACGGCGTTGAAG</u>	80
81	AGGGGGGGGCTCCGACAGAAACAGCCAGAGCCGACCCACTCACCTGACTGCCAGCTAACACCTGGGGCGGACAGGGAC	160
81	AGGGGGGGGCTCCGATAGAAATAGCTAGAGCGTATTATLALTGAGTGTAGGTAAATATTGGGGCGGATAGGGAT	160
81	<u>GGTGGTGTTCATAGT</u> <u>AGTTAGAGCTATTATT</u>	160
81	<u>GGCGGTTTCGATAGAA</u>	160
81	<u>AATAGTTAGAGTATTATT</u>	160
161	ACCAAAACAAGGTAGGGTGGGAGGCTGGGGAGGGAAGAGGTTGGAAGG	240
161	AGGACACCTTGGAAAGG	240
161	ACCAAAACAAGGTAGGGTGGGAGGCTGGGGAGGGAAGAGGTTGGAAGG	240

Figura 42: Amplificado de la SEQ ID NO: 1

```

504 TCTGGAGGGATAGAATGTGATCCTCGAGCAGAAACCATCGTGTTTCTGGGGGCTCTTTGGGGCCCGTGGGGCGACCGCGT 425
504 TTTGGAGGGATAGAATGTGA 425
504 TtTGGAGGGATAGAATGTGATtLTCGAGLAGAAALATCGTGTTTtTGGGGGtTtTTCGGGGCtCtGTGGGGCGACCGCGT 425

424 GCAGCGTCTAGGGCCGCCGACGGCCGGGGGAGGGCTCCAGCGGTTCGCCCGCGCGGACCGGCCACCGAGGAAACCC 345
424 GlAGCGTtTAGGtCCGtCGtAGGtCGGGGtAGGGtTtLTCGGGTtLTCGGGLAGCGGLACGLAGGAAAtTC 345
424 TTAGGGTCGTCTAGGT AGTGGTTATGTAGGAAA 345
424 TTACGGTGTCTTAGGT TAGCGTTACGTAGGA 345

344 GCTTCCTCGCCCTGCATCTGCTGACAAAGCCGCCGAGGTGTCTGCGTGAGCGTGCCACACCGATGCAGCTTTTCFAGG 265
344 GlTtLTCGtLTCGtLTCGtATtTGTATtTGTGATAAAGtCGtTCGAGGTtTTCGtTGAGCGTCCtAtAtCGATGtAGtTTTTCFAGG 265
344 AAGTTGTTTGAGGtGT 265
344 ATAAGTCTGTTTCGAGGT 265

264 AAATGGCCCGGGAGGGGGAGGGGGCGAGTGAGGGATTAGTTCGGGTCCCGCTCGTCCCGCTGGGCAAAACACCCACCC 185
264 AAATGGtLCGGAGGGGGAGGGGGCGAGTCAGGGATTAGTtCGGtTtTtTTCGtTtTtTTCGtLAAAtACtAtTtTt 185

```

184	CCTTCCTTTCCCGGAGGCATGAGGCTCCACGGTCGGTCCTTCGCCCCCATCTCGCCGACCCCTCCGGTTTCGCGCCT	105
184	ttTTtttttCGGAGGtATGAGGttTttACGtTCGGTtLLTTCGtTCGtAittTCGtTCGtAittTtCGGTTTCGCGGttT	105
104	CATCGCCAGGCTTTTCCTGGGAGCCGACTGCGGTGAAGTCGCGCCGCCACGACCGTCGGCCCTGACACGTCGCTTAACG	25
104	tAtTCGttAGGtTTTTttttTGGGAGLCGAtTGGGTGAAGTCGCGGTCGGttTAGtACGGTCGGtttTGAtACGTGtTtTAACG	25
104		
104		
104		
104		
24	CATGCTTCAGTCAAGGCTGTCCTT	-55
24	ATGTTGAGTTAAGGTTGTTTTT	-55
24	tATGTTTGAGTTAAGGTTGTTTTT	-55
24	<u>TAT</u>	-55
24	<u>TA</u>	-55

Figura 43: Amplificado de la SEQ ID NO: 2

1	AGATCATCAGTCCTCACCTCAGGACCTTCGGGGCATCTATGGGGCATGGTTACTGCTCTGGTGTCTCCCGCAGCCG	80
1	AGATTATTAGTTTTTTATTGGGGATT	80
1	AGATLAPTAGTTTCTTATTTCAGGGATTTTTCGGGGATTTATGCGGGATATGGTTATGTTTCTGGTGTCTCTCGTAgTcG	80
1		<u>TAGTCG</u> 80
1		<u>TAGTTG</u> 80
81	CGCGCAGGTACCGTGCACATCGGCATGGCCAGCTCCTCAGCCAGGTCCACGGGCAGCGGCCCCAGGCATCGCGCAGG	160
81	CGCGLAGGTATCGTCCGATATCGCGATGGTTLAGTTTCTTACTAGTTTACGGGTACACGGGTTTAgGtATCGCGGtAGG	160
81	<u>CGCGTAGGTA</u>	<u>TTAGGTATTGGTATG</u> 160
81	<u>TGTGTAGGTAT</u>	<u>TAGGTATCGCGTACG</u> 160
81	<u>GTAGGTATTGCTGCATAT</u>	160
81	<u>TAGGTATCGTGGATA</u>	160
161	TCCAGCCGCGCCCGCCCGGTGCAGCACCCAGCGTGTCCAGGAAGCCCTCCCGGGCAGCGTCTGTCCACCGGTCCGGT	240
161	TTTAgTcGCGGtTTCGGTTCGGTGTAGTATTATTACCGTGTCTTAGGAAGTTTTCGGGtAGCGTCTGTGLACGGGTCCGGT	240
161	<u>TT</u>	240
161	<u>T</u>	240

241	GAGAGTGGGGGTTCGGCGCAGTTGGGCTCGCGGCCGTGGAGCAGCAGAGCTCCGCCACTCGGGCGCTGCCCATCATCA	320
241	GAGAGTGGCGGGTTCGGCGCAGTTGGGCTTCGGCGGCCGTGGAGCAGCAGAGCTCCGCCACTCGGGCGCTGCCCATCATCA	320
241	TTTTGTGTTGTGGACTA	320
241	TTCCGCTCGTGGAGTA	320
321	TGACCTGCCAGGAGAGACAGATGGTCAGAGCCAGGGTGG	400
321	GAATCCCTTAGACTTAGGGTGG	400
321	TGALLTGTLTAGAGAGAGATAGATGGTLAGAGTTAGGGTGG	400

1	CCCTTCCTTGGGTGTGTCTTGGCTGGCTCCCTCCCTCGCTTAGGTGACATGTTTTCAACCTCGAATAAACAACTCGACC	80
1	<u>TTTTTGGTTGGGTGCTGTTTT</u>	80
1	TttTTGtTTGGGTGTGTtTtTTCGtTCGtTtTtTtTtTCGtTTAGGtTAATGTTTTtAAAtTCGAATAAAAtTGTAGt	80
81	CAACTTCGGAGGCGAGCTCAATTGCGCAGCGGAGCCAGCCCTCGCCAGCTTCGGTCCGCCATCCTCGTCCGCTCCCTCCGC	160
81	tAAATTCGAGGtAGctTtAATTGtTtAGGGAtTtTAGtTtTGtTAGGTTCGtTAATtTCGtTtCGTtTtTCGtTtTCGt	160
81	<u>TAGCTCCGCTTCGTTAT</u>	160
81	<u>TAGGTTTGGTTTGTTATTT</u>	160
161	CGGCTCTGCCCGCGCCAGGATTCCTCCAGCTCCTTTGCGCGCGCCCTCGCTTCGGTTCGGACACCATGGACAAGTT	240
161	CGGtTtTGtTtCGGtTtTAGGAtTtTtAGTtTtTTTCGtTCGCTtCGGtTtCGGAtAtTAATGGAtAAGTT	240
161	<u>GTTTCGCGCTTTAGGCA</u>	240
161	<u>GTtTTGTGTTAGGGAT</u>	240
161	<u>GTTTCGTTTCGGATATTA</u>	240
161	<u>TTTCGTTTCGATATTAATGG</u>	240
241	TTGGTCGCACGCAGCTGGGGACTCTGCTCGTTCGGCTGAGCTTGGCGCAGATCGGTCACTGCCCGCGCAGCCTGGGC	320
241	TTGGTGGtACGTAGtTTCGGGAtTtTcttTCGTGtCGtTGAGtTtGGCGTAGATCGGTGAGTgtTCGTCGtAGtTGGGt	320
241	<u>TTTGGCGGTAGATCGGT</u>	320
241	<u>TTTGGCTGAGATTGGT</u>	320

321	AGCAAGATGGGTGGGGTGTCTCAGGGCGGACCCGGGGCGCAGCCCTCCGCTGACTGGCCCTGGGGGACTGGAGTCAA	400
321		
321	AGTAAGATGGGTGCGGGGTGTATTAGGGCGGATTCGGGGGtAGtTTtTCGGtTGACTCGGtTTTGGGGGAtTGGAGTLAA	400
401	GTG	480
401	GTG	480
401	GTG	480

Figura 45: Amplificado de la SEQ ID NO: 4

446	GAAGCGATGAATGAATATAAGTACTTGCTCTGATGGCAGCAGACACCCGAGCAAAACGGTGGAGGCTACACTGTCTGGCAT	367
446	GAAGGGATGAATGRATAAAGT	367
446	GAAGCGATGAATGAATATAAGTACTTGCTCTGATGGCAGCAGACACCCGAGCAAAACGGTGGAGGCTATATGTCTGGCAT	367
446	<u>TTTGACTAATATCGTGGCA</u>	GTAT 367
446	<u>ATTTCGAGTAAACGGT</u>	TGGTAT 367
366	TCCTCCAGCGTTTCCTCAGAGCCGACCGCCCTGCAGCTCAAGGGAGGGCTGCTCTCTCCAGAGCAGGCTGGAAACCA	287
366	TTCTCGTAGCGTTTCCTCAGAGCTGAGTCTGAGTCTGAGTCTAAGGGAGGGCTGCTCTCTCTCCAGAGCAGGCTGGAAACCA	287
366	<u>TTTCTGTAGCGTT</u> <u>ACTTCCGATTTGTTTCTA</u>	287
366	<u>TTTCTGTAGCGTT</u> <u>TAGATCGGATTCGTT</u>	287
286	GCTGGGTTCCGCCPCCCGGGAAGTGGTCTCCATTCGTGCTCTGCAATCTGCTTTGTCAGATCCGAGAGCTAAACATTCG	207
286	GCTGGGTTTCGTTTTCGGGAAGTGGTCTCCATTCGTGCTCTGCAATCTGCTTTGTCAGATCCGAGAGCTAAACATTCG	207
286	<u>TTTGTCTTTTGGGAAG</u>	207
286	<u>TTGTCTTTTGGGAAG</u>	207
206	GGCTTGGTGTGTAATTAATCATTTGATTTGACCTTATTCCTGGGGCTTCGGTTTGGCTTACTAGTTTGGGATTTTAAAAA	127
206	GGCTTGGTGTGTAATTAATCATTTGATTTGACCTTATTCCTGGGGCTTCGGTTTGGCTTACTAGTTTGGGATTTTAAAAA	127

126 AATTAAATAATTAGCCTATAGAGAGGCAAAATTAATAATTAGTTGGTAAAGCAAGGAGCGCGAGTGTTTCAAGCCGTTTG 47
 126 AATTAAATAATTAGCTTATAGAGAGCGGCAAAATTAATAATTAGTTGGTAAAGCAAGGAGCGCGAGTGTTTGAAGTCGTTTG 47

46 GAGGGAACAGCGGTTTCCAGTTCCCTGCTGACTTGAGAAGCTCTG -33
 46 TTGTTGATTTGAGAGGTTTGTG -33
 46 GAGGGAALAGCGGTTTLLAAGTTTCTGCTGACTTGAGAAGTCTTG -33

Figura 46: Amplificado de la SEQ ID NO: 5

1	ACATGGGATTGATGGAGACAGCCCCCAAGGACGGGGGTGGTGCCCTGTTTTCTCTGATTGSCGGACGGGGACTTC	80
1	ATATGGGATTGATGGAGATAG	80
1	ATATGGGATTGATGGAGATAGtttttAAGGACGGGGCTGGTGGtttTGTttttTttTtTGATTGGCGGACGAGGGGAtTt	80
1	<u>TTGATTGGCGGACGAG</u>	80
1	<u>TTGATTGGTGGATGAG</u>	80
81	TAGACCCCTGCCAAACACCAAGGGGCGGTCTCGCACGCCATGCTCTCAGCGGAAGCAGGGCTTGTATAGAAATGGGC	160
81	TAGAttttTGttAAAtAtttAAGGGGCGGTGTGtACGtAtGtTgtTtAGCGGAAGtAGGGtTtGTATAGAAATGGGC	160
81	<u>GGTtTGTATGtTtATCT</u>	160
81	<u>TGTCTACGTTATGTT</u>	160
161	GTAGACCGCGCGGAGCTGGGTCTGGACTGGGCTCCGGTAGCGCGGACGGGAGCTCTACGGTTAGACAAGACCATAGAG	240
161	GTAGLAGLCGGCGGGAGTGGGTGGGAtTGGGtTtCGCGTAGCGCGGtAGGGtAGtTtTACGGTTAGAtAAGAttATAGAG	240
161	<u>GTAGTACGTTGG</u>	240
161	<u>GTAGTAGTCGG</u>	240
241	CTGGGTAAGACGAATTTAGAACACAGCGGAGGTAGGACGGCAGGATGGCTGTCAAGGCACACCAAGAGCATTGAGTGGC	320
241	tTGGGTAAGACGAATTTAGAAAtAtAGCGGAGGTAGGACGGGLAGGAtGGtTGTtAGGtAtACGAAGAGtATTGAGTGGT	320

321	AGAAACGAAATGCTTT	CAGAGGGCAGGGTT	GAACTGCGGCAGAGTT	AAAGCCAGCAACT	GGCTCTGCATTGATG	400
321	AGAAACGAAATGCTTT	AGAGGGLAGGTTGA	ALTGCGGTLAGAGTT	AAAAAGGGGRGA	ATTGGATTGATG	400
401	GCGGTGGCTGGACT	CAACCTAGNATAG	GGSCCGACTAGG	AGACCCGATCC	AGTCTGTGCCCA	480
401					GATTTAGGTTTGTGTTTATTGGA	480
401	GCGGTGGCTGGACT	TAATTTACAATAG	GGGGLCCGATAG	CGAGAGCGCGAT	TTAGGTLTGTGTL	480
					LATTGGA	

Figura 47: Amplificado de la SEQ ID NO: 6

1	CTTTCATCCAGCATGAGGACATTTAAGATGAAATGTCGGTGGCAGGATCGTTTCCTTCACCTGCTGCATGCGGCACTGG	80
1	<u>TTTTTATTAGGATGAGGGATATT</u>	80
1	TTTTTATTGACCATGACGGATATTTAAGATGAAATGTCGGTGGCAGGATCGTTTCCTTCACCTGCTGCATGCGGCACTGG	80
81	GAACTCGGCCACCTGTGTCCGGAACCTGCTCGCTCAGTCGGCTTTCCTTCACCTGCTGCATGCGGCACTTTCTCCACGG	160
81	GAAATCGTGTGTTTAAATGCTGCTCGGAAATTTCTGCTTACGTCGGCTTTCCTTCACCTGCTGCATGCGGCACTTTCTCCACGG	160
81		TACGG 160
81		<u>TATCG 160</u>
161	ACCTGGCGGCTCCAGTCCCACTTCATTCACCGGTCACCTGCCATCTCGACCACTCCGACCTGCGAGTGGCTGGTGCAG	240
161	ATTGCGTGGTGTGTTAGTGTAAATTTTATTTTACGGTAAATGTTATTCGATTAAGTCCGATTTGAGTGGTGGTGTAG	240
161	<u>ATTGGTGGT</u>	240
161	<u>ATTGGTGGT</u>	240
	<u>TATTTGATTAAGTTTGGAT</u>	
	<u>TTCGATTAGTTCGGAT</u>	
241	CCGGCCCTCGTCTCCGTCGCCGCCATCCAGACCAAGACCCCTTCGGAGTCCCGGCCCTCCGCTGGGGC	320
241	TCGGTTCGTGTTTTCGTTGGTTCATCGTACATTAAGATTTTATTTTCGGAGTTCGGTTCCTTCGGTGGGT	320
241	<u>TTTCCTGCTTTTATCGTA</u>	320
241	<u>TTTGTGGTTCATTAAGTA</u>	320

321	TTACTCCAGGGCTGGCGTTGTGAAGACCATTGACAGGAGGGCCGAGCCAGACCATTTGGCAGGAGGGGCAAGGTGGACAGG	400
321	TTATTTTAGGGTTGGCGTTGTGAAGATTATGATAGGAGGCGAGCGTAGAGTATTGGTAGGAGGGGLAAGGTGGAAALAGG	400
321		400
321	<u>TAGGAGGTTGAGTGTGTA</u> DTGGAGCGTAGAGTAT	
401	TGAGGAACTCTACCGTACTCTTCTCTGGGAATGTGGGGGCTGGGTGGGAAGCAGC	480
401	GGTGGGTGGGAAGTACT	480
401	TGACCAATTCTACCGTATTCTTTTGGGAATGTGGGGGCTGGGTGGGAAGTACT	480

27	GGGGCCTAGGGAGCTAAACAGACAGCAG	-52
27	GTTTACCGGACTAATATAGATAGTAG	-52
27	GGGGELTAGGGGAGCTAAATAGATAGTAG	-52

Figura 49: Amplificado de la SEQ ID NO: 38

466	CRGATGGGGACAGGAAGCTGTGGACGAAAGCCCCAGGTC	387
466	TAGATGGCGATACCAAGTTCT	387
466	LAGATGGGGALAGGAAGTGTGGADGAAAGTtttAGGTTtCGTGGGACAGCTGAtAGTAGtAGGGtACGtAGtACGtG	387
466		<u>TATGTAGTTATGTC</u> 387
466		<u>GGTACCTAGTTACGT</u> 387
386	GGTCCCCAGGGGAATGTGAAGCGGAGGGtCCAGGGGAACTGGGGATTAAACAATATTTACACCCACAGGGGAAGTGC	307
386	GGTttttLAGGGGAATGTGAAGCGCGCAGCGCtttTAGCCGAATGGGGATTAAATAATATTTALAGGLAGtAGGGGAAGTGT	307
386	<u>GGT</u>	307
386		307
306	CCAGCGCACGTGACGGGGCGGGGCGGGACtTTCCGGACGCCCGCCGCTAACGCTATCGAGCGAGCCGGTTGTAGACGTG	227
306	ttACCGtACGTACCGGGGGCGGGGCGGGALTTTGGGGAGGGCGGGGtTTAACGGTATCGAGCGAGtCGGTTCTAGACGTG	227
306	<u>AGTGATATGTGATGGG</u>	227
306	<u>TAGCGTACGTACCG</u>	<u>TATTGAGTGAGTTGGTT</u> 227
		<u>TATCGAGCGAGTCGGT</u> 227
226	GTCCAGGTTTCTGCACAGGAATATCGAGAGCGTCAATGAACCCGAGCTATAGAGAAAGGAGATGAGTTCAGAGAATCGAA	147
226	GTttTAGGTTttTGtAtAGGAATATCGAGAGCGTtATGAAttCGAGtATAGAGAAACGACATCAGGtTAGAGAATCGAA	147
226		<u>ATCGAGACCGTTAIGA</u>
226		<u>AGGAATATTGAGAGTGT</u> 147

146	GGGTCCTCGGCGACCGGGGTGATGCCACCCGGGCGGAGAGGTCAGGGGTGGGGCGTGAAATTGGCGGGAGGGCC	67
146	GGGTTGTTGGCGAGACGGGGGTGATGTTTATCCGGCGGAGAGGTTAGGGGTGGGGTCGTGAATTGGCGGGAGGGGT	67

66	TGGAGAGGGGATAGGTAGAAAAGGGGGGGCTCGACGGGAATGCCCGCAGGGCTGGCTGATGAGG	-13
66		-13
66	TGGAGAGGGGATAGGTAGAAAAGGGGGGGCTCGACGGGAATGGCTAGGGTGGGTGATGAGG	-13

Figura 50: Amplificado de la SEQ ID NO: 39

```

466 CCACCAGGAGAGGGGAAGAAGCCAGCACCTACCGACAGGGGTGGAGCTCCCTCAAGATGGTGTGGTCCCTGCTTTGGGG 387
466 TTATTAGCAGACCCCAAGACT 387
466 TTTTLAGGAGAGGGGAAGAAGTLAGTATTTATCTGATAGGGGTGGAGTGGGTTAAGAAATGCTGTGCTTTTGGGG 387

386 GAATGCTGGGGAGCTAGAAAGCCCTTCTAAGGGGGCGTCACGTGCAATTACTGCTTCCCTCTTTCCCATATAAACTCCCTCT 307
386 GAATGCTGGGGAGGTAGAAAGTCTTTCTAAGGGGGGTGATGTAATTATGCTTTTCTTTTtataaaaattctctctt 307
386 AATGGGGGTGTTATTTG 307
386 TTTAAGGGGGCGTTAT 307

306 AGTGTATCAGATACCCCAAGGAGTTTCAGTAAGCGGTTCTTCTGTTGTCTCCGCTGAGACTCCAGGGGAACCTCAAGCT 227
306 AGTGTATTAGAAATTTTAAGCAGTTTtAGTAAGCGGTTCTTCTGTTGTCTCCGCTGAGACTCCAGGGGAATTTCAAGCT 227

226 CACATGGCCCTGGCGGGCCCTGGGCAGGAGCAGGGCAGAGGCTCTGGCTGGCCGCTCTCTACCTGCGTCCGACTCCCGG 147
226 TATATGGTtTGGCGGGtTTTGGGTAGGACTAGCCGACAGAGGTTGGCGGGtCGTtTTtTAttTGCGTtCGAtTtCGG 147
226 TCGATTTCGG 147
226 GTTTGATTTTGTG 147

```

146 GTCCTTGGGACGACGACCGGGTAGTGGCTCAGACTACAGACCCGCGGATGACCCGCTACCTCAGTTTCCATTGGC 67
146 GTTtTTGGGtAGtAGtAAACCGGGTAGCGGtTAGAATATAGAtttTAGCGCGATGAtCGtTtPAtTtTAGTtttATTGGt 67
146 GIIT TAAATGGGTAGCGTTT ACTGTGACGATTGGTTT 67
146 GTTT TAGAAATGGGTAGTGT TAGCGCGATGATTGGT

66 CAGGGATCAGGGCTACCGGCTCCCAATGGCTACTTATCATATAGAGGTGGGCCAGCCCTGGGAATG -13
66 TGGCGTTACTTTGGCAATG -13
66 TAGGGATACGGGtAttCGGtTtTtAtTTGGtTAUttATtAAtATAGAGGTGGGSttAGtLLTGGAAATG -13

Figura 60: Amplificado de la SEQ ID NO: 40

1	CAC	TGG	ATG	TCC	CTGG	CAG	TC	CGG	TG	TCT	CG	AAG	CAC	AG	AGG	CAG	CC	ACG	AT	CTG	CC	CCG	CTG	CC	CT	80
1	<u>TAT</u>	<u>TGC</u>	<u>ATG</u>	<u>TTC</u>	<u>TTC</u>	<u>GAT</u>	<u>GTT</u>	<u>TTC</u>	<u>GAT</u>	<u>GTT</u>	<u>TTC</u>	<u>GAT</u>	<u>GTT</u>	<u>TTC</u>	<u>GAT</u>	<u>GTT</u>	<u>TTC</u>	<u>GAT</u>	<u>GTT</u>	<u>TTC</u>	<u>GAT</u>	<u>GTT</u>	<u>TTC</u>	<u>GAT</u>	<u>GTT</u>	80
1	LAC	TGG	ATG	TTC	TGG	TAG	TC	CGG	TG	TCT	CG	AAG	CAC	AG	AGG	CAG	CC	ACG	AT	TG	TC	CG	TC	TG	TC	80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										80
1																										

321	CCCCGGCTTGGCCGGCTTGGCGAGGCTATGCTCCGGGAAGCTGGAA	TCCAA	CGCGCGCGGCTGAACGCTGAGC	400
321	tttccgcctttggtcccggtttggcgagggtatgtgtcgggaaggtggaatt	aaacg	cggtcggttgaaatcgttgcagt	400
321			<u>ggaatttaattgtgtgt</u>	400
321			<u>gaaatttaacggcggt</u>	
401	CGCGGGAAPCCAGCGCGGAGCGCGGTATAGAGTGCAGGATGAGGAGGACCGACT	TCGCTAAGGGAGCGAGCTGACT	480	
401		TAAGGAGGAGGTGATT	480	
401	CGCGGGAALTAGCGCGGAGCGCGGTATAGAGTGCAGGATGAGGAGGATCGAT	TCGTAAAGGAGGAGCTGATT	480	

104	GTCAGCTCTGCCCCCTGCAGAACGGTTTGACAGAGGCAATCTCGCGCGCTGGATCGGGGGGGCCCTGCGCCTGCCAGAC	25
104	GTtAgtTcTgctttcttTctAgAcCGGTTTGNAtACAGGGAATtTCGGCGgtTgATGGGGGGgtttTgtttTgtTAgaT	25
104	A''''TCGGCGCTTGGAT	25
104	ACCTAAATTTTGGTGGTT	25
24	TCTCAGTGAGCCTCCAGGGTGGG	-55
24	TAGTGAGTTTTAGGGTGGG	-55
24	TttTAGTGAGttTttAGGGTGGG	-55

1	GACCAAGCGTTGGAGTTGTTTCTTTCGGGGGGCTTCGCCCTCCTGACGCCCTACTGAGTTGCCATCAGAACCTGGAGGATTGT	80	
1	GATTAGGTTTGGAGTTGTTTTT	80	
1	GATTAGGTTTGGAGTTGTTTTTTCGGGGGGGTTTTTCTTTCGGGGGGGTTTTTTCGATAGTGTATAGAACGTCGACGATTGT	80	
81	CCTCTGAGGACGCGCGGGGCAACCGGAGTCTTAAATAGCAGATAAAGTTAATACTGACTGCTAATGTGCGTAAATTGGA	160	
81	ttttTGAGCACCGCGGGGTAATCGGAGTTTAAATAAGTAAAGTTAATACTGATGTAAATGTGCGTAAATTGGA	160	
161	TTATAATCTTTAGGCGTATGGTGACAACGAGCTAATCCATCTCTGAGTATTAATCCGGTTAGACTCCCGGTTGACAGC	240	
161	TTATAATCTTTAGGCGTATGGTGATAAATAGCGTAATtATtCTGAGTATTAATCCGGTTAGATtTCGGGTGATAGT	240	
161		TAATTCGGTTAGATTTTCGG	240
161		TAATTTGGTTAGATTTTCGG	240
241	CAGTCCGAGGAGTGTTCACAGAGTCGCGCCCTCTCCAAATGTTCTCTCACTCGCGTTGTATTTGATCTCAGTCGCGCGGAGATA	320	
241	tAGTCGACGAGTGTtttTAGAGTCGCGGtttttGtAAATGTTtttTtATCGCGGTTGTATTTTCATtTtAGTCCtCCGAGATA	320	
241	TTTAGAGTCGCGGTTTT	TTTAGTCGTCGGAGATA	320
241	TAGAGTTGTGTTTTTGT	ACTGCTTCGAGATA	320
241		GATA	320
241		TA	320

321	GAACAGCGCGGCCACGGGTGGGGCGGGAGAGGAGTATCAACTACTGGCTTCCTTTACAGACGCAGATAACCTTTAAAAA	400
321	GAALAGCGGGGLACGGGTGGGGCGGGAGAGGAGTATTAATTAATGGCTTCTTTATAGACCCACATAATCTTTAAAAA	400
321	GA	400
321	<u>GAATAGTGTGGTGA</u>	400
321	<u>GAATAGCGCGGTTA</u>	400
401	GGAAACGGATAGATCTAGACCATTCATTCAAAGTGAAAGATGGATGTGACT	480
401	AAAGTGAAGATGGATGTGATT	480
401	GGAAACGGATAAGATTTAGACCATTCATTCAAAGTGAAAGATGGATGTGALT	480

Figura 63: Amplificado de la SEQ ID NO: 43

389	GGGGACAAGTGGTTAATGAGCAGGCCCAACTGGCCGGTCCCTCCTGTATGCCACACACACACCCCGTGAGGGCAGCAA	310
389	GGCGATAAGTGGTTAATGAGT	310
389	GGGGALAAAGTGGTTAATGAGLAGGGTAAATGGELCGGTtTtTGTTtTGTATGtAtAtAtAttCGTGAGGGTAGLAA	310
389	<u>TATATTTTGTGAGGGTAG</u>	310
389	<u>TATATTTCGTGAGGGTA</u>	310
309	GGAGCATCCAGCCAGGGGACGTGCACGAGAGAGAGGCCAGGATGATCACACTGCCCGGGCTCAGCACAGGGCTGAGTCCCG	230
309	GGAGTATtTAGtTAGGGGACCTtACGAGAGAGGGtTAGGATGATtAtAtTGTtLGGGGtTAGLAtAGGGtTGAGTtLGG	230
309	<u>TAGGGGACGTGTACGA</u>	230
309	<u>TAGGGGATCTGTATGA</u>	230
229	GAGGTGACCATtGGCCCGCCCGCTCCATCTCCGGCCCACTGAGGTCTGCTGGGGGTCTCTGGTGTGGCCtTGACAGCCAC	150
229	GAGGTGATtATTGtCCGtTtCGtTtAtTtTGCGGtAtTGAGGtTtGTtGGGGTtTtTGGTGTGGtTtTGATAGtTtAt	150
149	CGTGGCAGGACAGGGTCTCGAGGTGCCTCTCGGTAGCTCTGGGTCTAGTGGGGTGGACCCGTGGGACGGACCGCATG	70
149	tCGTGGtAGGALAGGCTtTCGAGGTGtTtTCGGTAGGTGGTGGGtTtAGTGGGGTGTALtCGTGGGACGGACCGCATG	70
149	<u>GACGGACCGTATG</u>	70
149	<u>TGGCATGCAATATAT</u>	70
149	<u>ATTCTGGGGACGGAA</u>	70
149	<u>ATTGTGGCCATGCCAA</u>	70

69	CTCCTGAGGCTCCAGTGGATGCAGGCCAAGGCACTCCCGGGTCAGCGTCCCTACCTGGGAGCTGGG	-10
69		-10
69	TTTATTGGGGAGTTGGG	-10
69	TTTATTGGGGAGTTGGG	-10

321	CGGAGAGGGCCGGTGGGTGGTGGTGGAGCGGAACGTGGGGTGGTGGTGGAGGGTCACTGGGATTTATCTCTGGG	400
321	CGGAGAGGGTCCGTCGCTGCTGGTGGAGCGGACGTGGGGTGGTGGTGGAGGGTCACTGGGATTTATCTCTGGG	400
321	CGGAGA	400
321	TGGAGAGCG	
401	GCCCGGGATGAGCCCTCCGCAAACTCCAGGCAGGGGACAGGTCCTG	480
401	AGGTAGGGGAATAGGTTTG	480
401	GtCCGGGATGAGTtTCCGtAAAGtTtLAGGtAGGGGAAtAGGtTtTG	480

Figura 65: Amplificado de la SEQ ID NO: 45

```

1 GTAGCAATGGGCTATGGACTTCAGTAACTAGAGTCACGTTTACCGGGGACACAAACCAATGACGGCGCCAGATCTGTGT 80
1 GTAGTAAATGGGTTATGGATTTTAC 80
1 GTAGCAATGGGCTATGGGTTATGGTAACTAGAGTACGTTTACCGGGGACACAAACCAATGACGGCGCCAGATCTGTGT 80
1 TACGTCTTTATCGGCGA 80
1 TATGTGTTATTTGGTGATA 80

81 AAGCAGGGGGAGCCACGTTGCGGGCCGACGTGTCTGGGAGGGGGGGAAGGGGGGGGGCTTACTGACAGAGCGCGAGGTG 160
81 AAGCAGGGGGGGTACGTTGCGGGTGGTACGTGTCTGGGAGGGGGGGAAGGGGGGGGGCTTACTGACAGAGCGCGAGGTG 160
81 GATCTCTTTGGGTTGTA 160
81 TACGTTGCGGTCGTA 160

161 GGTAGATAGACCGGGAGAGAGCGGCGAGGAGCTGGAAACCGAGCCAGGCTCAGCCGCTCACAACACAGGTAGGAGATTC 240
161 GGTAGATAGATTCGGAGAGAGCGGCGAGGAGCTGGAAACCGAGCTAGGCTGAGCTGGCTGACAAAGGTAGGAGATTC 240
161 TGGATTTGGAGAGATGG 240
161 AGATTGGAGAGACCG 240

241 GGAGGACAAAGCGAGCTTTAGGGAGTATTTGGCAGTGCATTTGGGGCTGAATGCAGATCTTTGGTATTTGGAGGTTTG 320
241 GGAGGACAAAGCGAGCTTTAGGGAGTATTTGGGAGTGCATTTGGGGCTGAATGCAGATCTTTGGTATTTGGAGGTTTG 320

```

321	CATCACGTGGCAGGTGCGAGCCCAAGAGACCGCGCCAGGGATCCTGATCTTGAAAGATTGGGAGAGGCCAGGAGGCCAG 400
321	TATLACGTGGTLAGGTGCGAGTtLACACACATCCCGCctAGGGATtLTGATtTTGAAAGATTGGGAGAGGGtAGGAGGtTAG 400
321	AGATCGGCGGTAGGGAT
321	AGATTGCTCTACGGAT
401	TCCTATGGGGCGGTTCCCGTAGGATCACCAAGGTGGTTCGGCTGCCACAGATTGACTGCTTGTGGTGGCATGG 480
401	TCCTTGTGCGGTGGTATGG
401	TtTtTATCGCGCGGTTtTtCGTAGGATtAtLAGGTGGTTCGGTGTtLACAGTTGAtTgtTTGTGGGTGCTAtGG 480

Figura 66: Amplificado de la SEQ ID NO: 46

1	CCCCAGCAGTGGACCTGGTCAGCACCTCTGGGCTCGGACATCAGTGTCTCTAAAGGTGGAAGGGTTAGACCCCTCTAGGGG	80
1	TTTACTAGTGGATTGGTTAGTATT	80
1	ttttAGtAGTGGAttTGGTtAGtAttTtTGGGgtTGGcAtAttAGTGTtTAAAGGTGGAAGGGTTAGAttttTtTAGGGG	80
81	AAGGGCCACCGCCCTCCACAGAGGCTCTGGCTTAGCCCGCGTGGACACGTGAGGGGGGCACCTACCGTGTtTtTCCATGG	160
81	AAGGGtttAtCGttttTtLAtAGAGgtTtTGGtTtTAGgtCGCGTGGAtACGTGAGGGGGgtAttTtAtCGTGTtTtTtATGG	160
81	<u>GCTGTGTGGATATGT</u>	160
81	<u>GTGCGGTGGATACGTG</u>	160
161	ACTTGCTGGCGACTCCACGAGAGAAGGDCAGCCAGGAGGGCGGAGGTGCCCGCAGCGCCATGCCAGGACAGATGCCGACAGC	240
161	AtTtGtTGGGAtTtLtACGAGAGAgttAGtTAGGAGGGCGACCTtCGtAGCGttAtGttAGGAGtAGAtGGGLAGAtL	240
161	<u>TTCGGATTTTACGA</u>	240
161	<u>TTCGTGATTTTAAAGGA</u>	240
161	<u>GATTTTATCAGAAAGGT</u>	240
161	<u>ATTTTACGAGAAGGT</u>	240
241	CTGCCACAGGGAGGAGCATGCGGAGGCCAAAGAGATGGAGTGGGGCTGAGGCAGGTGCGTGGCGCCAAATGAAAGTGGGG	320
241	TGAAAGTGGGG	320
241	tTgttAtAGCCAGGAGtATGCCGAGttAAAGAGATGGAGTGGGGtTAGgtAGGTGGGTGGGGgtAAATGAAAGTGGGG	320

321 TCAAGAGATG	400
321 TTAAGAGATG	400
321 TTAAGAGATG	400

Figura 67: Amplificado de la SEQ ID NO: 47

335	GAGGAAGAGGAGGATACAGCGCGCCAAATGTGTGGCCTACGTGGGAGCGCCCGGAGCGCTCCACGTCCCGGCGCTGCGC	256
335	GAGGAAGAGGAGGATATAGG	256
335	GAGCAAGAGGAGGATACAGCGCGTAAATGTGTGGTACGTGGGAGCGLLCGAGCGGTTLLACGTTTLCGGTTTGGGT	256
335	<u>TTTTGAGTGTATTATGTT</u>	256
335	<u>TTGAGCGGTTTACGT</u>	256
335	<u>TTACGTGGGAGCGTTT</u>	256
335	<u>TTATGTGGGAGTGTTT</u>	256
255	CGGGCGGCGCTTTCCCGGCGCATCTTGCCAAACGCCCTGGCCCAACCCGGAGGGCCCCCGGGCCCCCGCACGCCCTCGTG	176
255	CGGGCGGCTTTTCTTCCGSGATCTTGGTAACCGGTTTGGTLLAAGTCGGAGGGGTTTCCGGGTTTCGTACGTTTtTGG	176
255	<u>GTAACCGGTTTCCTTT</u>	176
255	<u>TGGTAATGCTTTGGT</u>	176
175	ACCAGTCAGCATCCAGGCTCCTCTTTTCAGCCCGCGGGGTGTTTCATGGGGGTAGAAGTCCTAACCCCTCCACCCCTCTCCT	96
175	AtTAGTLAGGATLLAGGTTLLTTTTCAGATTCCGGGTGTTTtATGGGGGTACAAGTCTAAAtttTtAtttttTtTtT	96

95	CCCCAGCAGTCCACCGCGGGTATGGAGAGAAATGAAGTTCTTTGTCTCTAAGGGNTTCAAAACCAGAAACCGAGGGACCTC	16
95		
95	ttttAGTAGTtLACGCGGGTATGGAGAGAAATGAAGTTTTCGTCTCTAAGGGATTtLAAAttACAAACCGAGGGAttTt	16
95	AGTTTtATGTGGGTATG	16
95	<u>TAGTTTtACGCGGGTtA</u>	16
15	TGGTTCCAGAGGGA	-64
15	TGGTTTtTAGAGGGA	-64
15	TGGTTtTtAGAGGGA	-64

Figura 68: Amplificado de la SEQ ID NO: 48

```

1 GCCAAAGGGGCATTGGTTCCCCAAAGGGCCAAACGAGCTTCCTCCAGGAGTTTCCCGTCGGGTCAGTGGACACCGTGGG 80
1 GTTAAAGGGGTATTGGTTTTT 80
1 Gttaaaggggtattggtttttttaagggtttaaaccagattttllttaggagtttllcgtcggcggtgcagtggattatcgtggg 80

81 CCTGACTGAGGCCCTGGTCTTGGCAGGGTTTTTTTGGCAACGGCCCTGGATGAGTGAAGCGGATGAAAGCCGCCACAAGAA 160
81 ttgcattcaggcttgcgttttttgcctttaggcttttttggtaacgggtttaggatgagtgaagggtgaaggcgccgagagagaa 160
81 ATGAAAGCCGCCACAA 160
81 ATGAAGGGTGTGAGAA 160

161 TGGGCTGACAAACCAAACTCGGGGCCCTGAGTGTTTTGAGGGGACCGCATGGGAACCGTGGCTCCGCTGCCCTGATTTA 240
161 TGGGTTGAGAAATLAAAGTGGGGGtllTGAGTGTTTTGAGGGGACCGCATGGGAATCGTGGTtTCGtTgtttTGATTTTA 240
161 ATGGGAATCGTGGTT 240
161 ATGGGAATTTGTTGTTT 240
161 GGAATTGTGTTTGT 240
161 GAATCGTGGTTTCGTT 240

241 AGGCCAGTTACCCGTGACAGCCACGCTGGAACGTAGGTGTCTTCCAGCTTCTGTCCACCTCAGAGGAATTCAAAACAA 320
241 AGGGTAGTTA ttttTgttatagttacgttggaaatgcacctgttttttagtttTGTttatltTAGAGGAATTUAAAALAA 320

```

321	GGGCCCCAGTTC	TAGCCCCACAGACCCACGATCCCTAGGGCCATTAGTCCACCGGGCAAAATGGTTCACATTG	400
321	GGGtttAGTtT	tTAAGtlltAlAGAlullLAGGAtttttTAGGGgttATTAGtttAtCGGGtAAANTGGTTtTtTtAtTTG	400
321		TAGTTATCGCGGTAA	400
321		TTTATTGTGCTAAATGGT	
401	CTCATTGTGTTTT	GTGAAATCCAAATGTGCACCTCAGTGGCTGCACAGTTGCTACTGAAACGGTTCATATAGGTGACA	480
401		GTTTAATAGGTGAGA	480
401	tTtATTGTGTTTT	GTGAAATtLAAATTGTLAULTtAtTGGgtTGTAtAGTTGtTAtTGAAACGGTtLAATAGGTGACA	480
481	GAAGAAAGAAA		560
481	GAAGAAAGAAA		560
481	GAAGAAAGAAA		560

Figura 69: Amplificado de la SEQ ID NO: 49

1	GTAACAGGGACAGCAACTGGGGCCCTGGAGCCGGCAAGCGGTGAGCGNGTCTTCCACACCGCTATGTCCCTATTAAA	80
1	CTAAGACGGATACGGAAATGG	80
1	GTAAGAGGGATAGGGAACTGGGGCGTTGGAGTCCGGAAACGGGTGAGCGAGTCTTTTAAATCGTATGCTTTTATTATAA	80

81	AAGAAACCCCGCTCCTTTCGTAACCCCTTTTCCCTTAGCTCTTAAGCTSCCGGGCCGCGACGGGGAACTGSCCAACTCAT	160
81	AAGAAATTCGTCTTTCGTGAAATCTTTTCTTTTAGTCTTTAAGTCTGCGGTGCGACGGGGAACTGCTAAATCTAT	160

161	CCTTCGTCGGGCTCTCCGCTCCTCGGGCTCTTGGGCGCAGAACAGCTCCTGAGCGCGCGTGGGCGAGGCGAGTCAATGAGC	240
161	TTTCGTCCGCTTCTCGCTTCTCGGCTTCTGGGCTAGAAATAGGCTTCTGAGCGCGCGTGGGCGAGGCGAGTCAATGAGC	240

241	GGGAGGACGTGAGAGAGTGAACACAGGCTGCCGCCACTGCAAAATGGTTCGGAGGGGGTCCGCCCGAACACACAG	320
241	GGGAGGACGTGAGAGAGTGAATAGGCTGCTTTTCTTCTGCAAAATGGTTCGGAGGGGGTCCGCCCGAACAG	320
241		GGTGGGTTGAGAA
241		TCGGGTTCGAAGATTA

321	GATAGCAAGCTCCACGGGAAGGGAGCTCTGSCCCGGACGGGGGTGGGGACGGGTCTCTAGCGCGGGGATGCGGAG	400
321	GATAGTAAAGGTTLACGGGAAGGGAGTCTTGGTLOGGACGGGGGTGGGGACGGGTCTCTAGCGCGGGGATGCGGAG	400
321	<u>GAACCTTTATGGGAAAGG</u>	400
321	<u>AGCTTTACGGGAAGG</u>	400
321	<u>TTTATAGCGTCGGGCAT</u>	400
321	<u>TTTATAGTCTTGGGCAT</u>	400
401	CTGGGGAGAGGGGAGGGGACCCCATCTCTGGCTTGGGTGCC	480
401	TTTTATTTCCGTTTGGGCTTT	480
401	TTGGGACACGGGTCGGGACCTTCTTCTGGTTGGGTGTT	480

1	GGAGTTTTC	GAAAAGT	CGGCGCT	CGGGGCC	CAGTC	AATGGGG	CGCCCC	CGCGCGG	CGGCT	GAGTGG	AGCT	TAGCG	GAACCG	80
1	CGACTTTCTGAAAGTGGG													
1	CGAGTTTTC	GAAAAGT	GGGGT	TCGGG	GLLAG	TAAATGGGG	CGTLL	CGCGCGG	CGGGT	TGATGG	AGCT	TAGCG	GAACCG	80
1													TAGCGCGGAATCG	80
1													AGTGTGAATG	80
81	CTCACCG	CGCGCG	CGCCCA	ATTA	TACCG	CCCTT	TGCGGG	CGCGCG	CGCGCG	CGCGCG	CGCGCG	CGCGCG	CGCGCG	160
81	TTACG	CGCGG	TTTTAA	TAAATCG	GLTT	TGTGGG	GLCG	TTCGG	GLCG	AT	TAGCG	CT	TTATG	160
81													TACTCGTATTNGCCT	160
81													TACTTGTATTAGTGCCT	160
161	TTGCGG	CTCGCG	CGCGCG	CGCGG	ATTAC	CTTTT	CAAC	AGCGG	TTTTG	TCAGG	CGG	AGTTT	TCAC	240
161	TTGCG	CTCG	CGCG	CGCGG	ATTAT	CTTTT	CAAT	AGCGG	TTTTG	TTAGG	CGG	AGTTT	TTAT	240
241	TGACA	ATTG	CGCCCA	AT	TAGAT	AG	TTTGT	CAAG	CGAC	CTT	CGGAA	CGG	CAAC	320
241	TGACA	ATTG	GTTTT	AAAT	TAGAT	AG	TTTGT	TAGA	CGG	AT	TTGCG	GAAG	AGT	320

[illegible]

Figura 71: Amplificado de la SEQ ID NO: 51

228	GAAGGTTGAGGAGGAGCCAGAGCGGGTCCGTGACGGTTTCTGCGCATCAGCGCGCGTGGCCATCTCCACCATGCAGTCC	149
228	GAAGGTTGAGGAGGAGTTAGA	149
228	GAAGGTTGAGGAGGAGCAGTACAGTCGGGTTATGACGGTTTTCGCTATAGCGTTCGTCGTTATTTLLATATGTAGTTT	149
228	TACAGTCGGGTTTGTA	149
228	TACAGTCGGGTTTGTA	149
228	TACAGTCGGGTTTGTA	149
228	TTT 149	
228	AGTTT 149	
148	CGCGAAGACGCCCGCGGTCTCGCCGCTAGCCAGTCCCGGTGGTGGGAGCGGCCUAGAAGCATTCACAAACCCACAGT	69
148	CGCGAAGACGCTTCGGGTTTCGCTGCTTAGCTAGCTTTCGCTGGGAGCGGCTTAAGAAGATTTCAAATTTAAGT	69
148	ACACGTTTCGGGTTT	69
148	GAAGTGTGTTTGTTT	69
148	CGGAGACGTTT	69
148	TGGAGCATCT	69
68	TTCCGCCCTTTTTCACGGGTCCAGAGGAATTAAAGGACACGGCCCAATTCTGCAGCCCTGCTGGCACAGC	-11
68	TTTATTTTCTAGCTTTGCTTGGTATAGT	-11
68	TTCCGCTTTTTCACGGGTTCAGAGGAATTAAAGCATACCGGCTTCTTGTAGCTTCTGCTGGTALAGT	-11

321 CGGACTCCTGTGTGCGGTTGGGAGGCGACGGGGTGGGTGGCTCTCTGTGTCATGCCACACAACCTCGCTGCCCTCTCACCCC 400
 321 CGGATTTCTGTGTGCTGGGCGAGCGTACGGGGTGGGTGGCTCTCTGTGTCATGCCACACAACCTCGCTGCCCTCTCACCCC 400

401 ACTTCCACCCCTACATTTGTCTGTGTCACAGACTGG 480
 401 TCTTTTGTGTTTACACTGG 480
 401 ACTTCTAATTTTATATTTGTTTGTCTTCTAGAGTTGG 480

Figura 73: Amplificado de la SEQ ID NO: 53

1	GAGCTGGCCCTGCTGAGGTGCGGGCGGAGTAGTTGGAGAGAGGGTCATACTCCAGGCTCTGTGGGTGGCCCTGGGAGTTGTC	80
1	GAGTTGGTTTGTGAGGCT	80
1	GAGTTGGCTTTTGAGAGCTGTCGGGTCAGTAGTTGGAGAGAGGGTATATTTAGGTGTTGGGTGGTCTGGAGTTGTC	80
1	TGTCGGGTCCAGTAGT	80
1	TGTTGGGTTCAGTAGT	80
81	CACGACGTACTTSCACTTGGGACGGGGCGGCTGTGCGCGCGGGCTGGCTACAGCCTTCGCGGACCTATTCCAGGGCAC	160
81	TATTACGTATTTGTTATTTGGGACGGGGCGGCTGTGCTGCGCGGGCTGGCTTATAGCTTTGGGGACGTATTTTAGGGTAT	160
161	CGCCAGCCCTCCACCTGTGCTGACCGCTGTCCAGGAGCGCCAGCAGTACTTGTGCTGGGCTCGGCGAGGGCGGCC	240
161	CGTATATTTTATTTGTTGTTGATTTTGTAGGACGTTAGCAGTATTTGCTGTCTGGTTCGGCTAGCGGGCGGCT	240
161	ACGGATGCTAGTCAGT	240
161	GACGTTAGCCAGTATT	240
241	ACTCGCGCTCGGCTGTACCGGCAATCAGGGCTTAATAGGCGGTGGCTGCGGGGCTGTAGTGGAGGCCAGTGGGAAGGC	320
241	AGGTTAGTGGGAAGCT	320
241	AGTGGGGTGGGTGGTAGTGGTATTAGGGTTAATAGGGTCTGTCTGGGGTGTACTCGGAAGCTTACTGGGAAGCT	320
241	TGTTAGTGGTATTTAGG	320
241	GCTAGTGGGTATTTAGG	320
241	TGTAGTGGAGGTTAG	320
241	TGTAGTGAAGGTTAGT	320

321 ATCCT	400
321 ATTTT	400
321 ATTTT	400

Figura 74: Amplificado de la SEQ ID NO: 54

[illegible]

Figura 75: Amplificado de la SEQ ID NO: 55

384	TAGGACTTGCGGCTGGTGGTTCTCGGCCATGCCCTGGCCUCAAAGAGTCGTTGGTGAGATGCCCTCAGTGGAGCGGGGGGCC	305
384	TAGGA TTGTGGTTGGTGGTT	305
384	TAGGATTTGTGGGTTGGTGGTTTCGGCTATGCTTTGGCTTAAAGAACTCCTTGGTGCAAGATGTTTACGGAGCGGGGGCGCT	305
304	CGCGGGCGCCACGCTTGTTGGGTTGTTACAGGTGGCCACACGTCGTCGTGCACGAACCTGCCCCGACGGGAGCGGGCGCTAG	225
304	CGLCGGGTGCGTACGGTTGGGTTGGTTACGGTCGTGGTAAACGTCGTGGTGACGAAATCGCTTTGCTAGGGAGCGGCTTTAG	225
304		<u>CTTGCGTATGCAATTGGTT</u>
304		<u>GTCGTGTACGAATTCTGT</u>
224	CACGGCAGCTCGCCCGGCCGACGACGACCTCTTCTTCCCACCTGCGAATGCCCGCGATCATGCTCCCCACAGACCCACAC	145
224	TATAGGTAAGTCGCTTCGGTACAGTGAATTTTCTTTCTTTATTCGGAATGTTCCGATATGTTTTTAAAGATTAAGAT	145
224		<u>AGTTTGTTTTCCTTAGT</u>
224		<u>TAGTTCGTTTCGGTTA</u>
144	ACAAACACCTTGGGAACCAACGCGGTGCTCTCTGTTTACCGCAGCGCTCAACCCCGCGCCCTTGAGTCCACAGATGGC	65
144	ATAAATAATTTGGCAATTAACCGGCTGTTTCTGTTATCCGAGCGCTAAAGTCGCGGGCGCTTGAGTTTAAATGGC	65
144		<u>AAATTAACCGGCTGTT</u>
144		<u>GGAAATTAAGGGTG</u>
		AGGCTTTAAGTCGCGG
		TAGTGTTTAAGTTGTCG

64	GAAGTTCCGGGAGGTGGAGTGGGATCCCGGCGTGCCTGCCAGGACAGGAGGGGACG	-15
64	TTTACGACAGGAGGTTACG	-15
64	GAAGTTCCCGCCAGCTCGAGTGGGATTCGGGCTGCGGTGCTTAGGACAGGAGGTTAGG	-15

321	GCAGAGCCCTCGCAAGCAGCCCTGGAGCTGGACCCAGTCT	400
321	TTTGGAGTTGGATTAGTTT	400
321	GLAGAGCGGCTGGGAGGACGTTTCGAGCTGGATTACGTT	400

Figura 77: Amplificado de la SEQ ID NO: 57

378	GGGAGGTGCTCAGTGGTCTCAGTGGTCCCGACACCTTAGCTCTCTCCGCCCCCTGCGGGCCGACTGCGCCAGGGGCC	299
378	GGGAGGTGTTTAGTGGT	299
378	GGGAGGTGCTCAGTGGTCTCAGTGGTCCCGACACCTTAGCTCTCTCCGCCCCCTGCGGGCCGACTGCGCCAGGGGCC	299
378	TTAGTGGTTCGAGAT	299
378	GTITTAGITGTTTCAGA	299
298	GGGGTTATGTCGGCAGGCTCGAACGACTGGCAGCGTGGCTGTGATTCCTCCACTCCTTTCGCCCGCAGAGCGGTGGT	219
298	GGGGTTATGTCGGCAGGCTCGAACGACTGGCAGCGTGGCTGTGATTCCTCCACTCCTTTCGCCCGCAGAGCGGTGGT	219
218	CTAGGCGCAGACCATTGACCTGGGTAGGGGGAAGGCAAGGACGCACTCCGACGCTTGGTTTTTCACCGCCGCTCTCCCTT	139
218	CTAGGCGCAGACCATTGACCTGGGTAGGGGGAAGGCAAGGACGCACTCCGACGCTTGGTTTTTCACCGCCGCTCTCCCTT	139
218	CACGTAGTTCCCTACGT	139
218	GGATGTAGTTTGTATGT	139
138	CGGCGTTTGGGCTGGCGGCGGACATTTGGCTGGACCAATAGTGGCAAGTTCACCTTCCTCTCTCGCAAGCAC	59
138	CGGCGTTTGGGCTGGCGGCGGACATTTGGCTGGACCAATAGTGGCAAGTTCACCTTCCTCTCTCGCAAGCAC	59
138	TTTGTGTTTGTAAAGATAT	59
138	TTCCGCTTTCGTAAAGA	59

58	TTTCTACGGCCATCGGGACCTAAACACTCGGGAGGACCTGGTAAGTGAAGCAGAGGG	-21
58	TTTTGGTAAAGTGAAGTGAAGCAGAGGG	-21
58	TTTCTACGGCCATCGGGACCTAAACACTCGGGAGGACCTGGTAAGTGAAGCAGAGGG	-21
58	TATCTGTATCGGGATTTA	-21
58	TTACGGCTATCGGGAT	-21

Figura 78: Amplificado de la SEQ ID NO: 58

[illegible]

321 GGCCTCTCCCGGCGCGCGCCCTCTCTCTCCGACCCCTGCCCCCTCCGCGGGCTGACCTCTCTCAGTCCGAGTT 400
321 GGCCTCTCCCGGCGCGCGCCCTCTCTCTCCGACCCCTGCCCCCTCCGCGGGCTGACCTCTCTCAGTCCGAGTT 400

401 TCTGAAGGACTCAGCTGAGAAAGGACACTGGST 480
401 GTTGAAGGAGGATTAATTGGST 480
401 TCTGAAGGALTCAGTGGAGAAAGGATTAATTGGST 480

Figura 79: Amplificado de la SEQ ID NO: 59

```

1 GAGGACCTGCTAAGGCTTTCTCATCCACGAGGCCACCAATGGGCTTGCATTACAAAGGAATGCTCCCTGCTCATCTGCT 80
1 GAGGATTTGGTTAAGGTTTT 80
1 GAGGAULTGCTGTAGGCTTTCTTATTTAATTAGGTTATTATGGGTTGCGTTTAAAGCAATGTTTtTGTTTtTtTGTT 80
1 ATGGGTTCCGTTTATA 80
1 TGGGTTCTGTGTTTATAG 80

81 TTATAGTAAAGTCATTACAGAGG7GGTGGTTATGTTCAATTGATTCAGTGCACGAGACACAGTCACCTCTGATACAGC 160
81 TTATAGTAAAGTATTACAGAGG7GGTGGTTATGTTCAATTGATTCAGTGCACGAGACACAGTCACCTCTGATACAGC 160

161 AAACCCCTGCAGGCTGTATCAGGTCGCCCTCTGCAATCCCTGGGACCTCTCTATCTGTCTACAAAGTGGGGTCAATAAA 240
161 AAAttttGtACGGTGTATtAGGGTtttttTtGtAtGtlttTGGGAltTtTtTATtTGTtTtTAAAGTGGGTLATAAAA 240
161 ATARRA 240
161 ATAAATA 240

241 TACGACGCTGAATCCAAAGSCCTTGGCTCAACTGCACCCGCTCCCTCAATGCAACCGAAGCCCATGAGGCCCTAGCGTCT 320
241 TACGACGCTGAATtCAAAGStt7TGGCTtCAAtTtCAACGSttTGGtTtATGtAAUCGAAGlttTATAGSttTAGCGTtT 320
241 TACGACGCTTGA TAGTCTTT 320
241 TATGATCTTTGAAT ACCGTTT 320

```

321	TCCACGAATGAGGGCCACTCCCATGGCTACCTCGAGAAATCTGCTCCACGCTCTGACCCGCGCGCTTCCGACAGTGTCT	400
321	TtTACGAATGAGGGCTATtTtATCGGtLAlTCCAGAAATtTGtTtACGtTtTGRAAttCGCGGtTtTtLAtAGTGTtT	400
321	<u>TTTATCAATGA</u>	400
321	<u>TTTACGAT</u>	400
401	GGTACAGCCACATATTCCTCCCGCCAGAGATTGACGATGTGGCAGATATGTGSAATTGG	480
401	ATGTGCTAGATATGGAGTTGG	480
401	GGTALAGLLALAIATtTtTtCGGTAGAGATTGAGGATGTGTGATATGTGCACTTGG	480
401	<u>TAATTTTTCGGTACAGAT</u>	480
401	<u>TTTTTTTCTAGAGATTG</u>	480

Figura 80: Amplificado de la SEQ ID NO: 60

1	GGGGTCTCAAGAGCTGGCTCCTACACATTAAGGGCTTAAACAGTGGCTGGTACTAGTCCACCAATAGTCCTTTTGCCCC	80
1	GGGGTTTTAAGAGTTGGTTTT	80
1	GGGGTTTTAAGAGTTGGTTTTATATATAAAGGGCTTACAATAGTGGTGGTAACTAGCTATTATATAGTTTTTGTGTTTT	80
91	ADCCAAACGGGTACACAGCCGGGGCCCCCGGCCCTTGGGGCCATTACGGATTTCCGGCTCCCTCACAGAGGGAGT	160
91	ATCTAATCGGTGALALAGTTCGGGGTTCGGGGCTTGGGGCTATTACGGATTTCCGGCTTTCTTATAGAAGGAGT	160
91	<u>GTATATATCGAATTTTGT</u>	160
91	<u>TATTACCGAATTTGGT</u>	160
161	CACGCAACGGTGGCTGCAGTTGGTCATATCCGGCCCTTSCGATCAGGGTTGAGGACCCGGCCCATGAAGTGG	240
161	TAETGTAACGGTGGCTGTTAGTTGGTATATTCGGTTCGGATAGGGTTGAGGAACTCGGCTATGAAGTGG	240
161	<u>AGTTGTGTAAATTTGGT</u>	240
161	<u>TGGTTATATTCGGT</u>	240
161	<u>ATTCCGCTTATCAAG</u>	240
241	TSTTTGTTACCGTAGGCAACCCAGCTTTCAGGACCTCATTCGTGTGTGGGCGCCGACAGTCTGCAGTCAAGTGA	320
241	TTGTAAGTGAAGTGA	320
241	TSTTTGTTACCGTAGGCAATLALAGTTTCACGATTTATTGCTGTGTGTGGCTTCGATAGTCTGCAGTGA	320
241	<u>TGTTGCTGTTCATAGT</u>	320
241	<u>CTCCCGCTTCATAGT</u>	320

321 GGAGG	400
321 CGAGG	400
321 GGAGG	400

Figura 81: Amplificado de la SEQ ID NO: 61

1	TGGGGACACTGCATGTTTTTCTCTGAGGCGGTGGAGGTTTCACATCCCTGCCGAGGTGTGGGTGGCCCTTTTTTCCGCTGG	80
1	TGGGGATATTGGATGTTTTT	80
1	TGGGGATATTGGATGTTTTTATTTCAGGGTCGTGAGGTTTATATTtTTGtCGAGGTGTGGCTCCTtTtTTTTTtCGtTGG	80
81	TTTAAACATCCACACCTGGGGCTGTCTCTCCCGAGGGGAGGCTACTGTGCCCTTTTCTGGGCTGTCAAAAGCTGG	160
81	TAAATATTTATATALLGGGGTGTGTLTTLTLLTAGGGGAGGTTATGTGTtCGTTTTtT7GGGtTGtT7AAAAGtTGG	160
81	TTTAGGGCTAGCGTTAT	160
81	TTTAGGCTGAGCTTATT	160
161	CCCCACTCTTTCTGGGGCCGTCTGTCTTGTCTGTCTGCTCTCGGGGCAAGCTGTIAGAGCAGCCCAACCTGCA	240
161	GGCTATtT7GTGGGtCGT7TGTLITGtTGTIGtTCTTGGGGTAAAGTGT7IAGAGTAGtTtTATtTtAAAtT7GA	240
241	AACCAAGCTGGCCCCACAGACCTGMAAACCTGCMAACCGGGCTGGCCGACACAACCCACAACCTGAGAGCTGGGCTG	320
241	AATTAGGtTGGtTtTtTtACAtT7CAAAAtT7GtAAACGGGtTGGtTtTtATAGAACTtAAAtT7GtAGAtTGGGtTtG	320

321	GCCCCACAGAAUCCGCAACCGGGGUTGGCCCAAGAACCCGCCCTTCGAGCTCAGCTGCAATCTTCGTCGCCCGGT	400
321	GTTTATAGAAATCGAAACGGGGG-TGGTATTAGAAATCGCTTTTGAGCTTAGTGTGAATGTTCTCGTLLCGGT	400
321	ACAAATTCGTAAACGGG	400
321	ATTGTAAATGGGGT	400
401	GGTACCTGGCTGAGAGAGGAAGGGCTCTGGGAATTCGGTCAACGACGGCTCCACACCCCGTAGGCTAGGTGGGTTCTG	480
401		G 480
401	GCTATTTTCCTTTGAGAGACCAAGGGCTTTGGGAATTCGGTCAACGACGGCTGTAATATTCGTAAGGTAGGTGGGTTCTG	480
401	TTCCGGTATTCGACGGT	480
401	TTTGGCTTATTGATGGTT	480
401	TTCTATATATTTCTGACGGT	480
401	TGTATATTTCTGTAAGGT	480
481	CAGGCCAGGCTGGAAG	560
481	GAGGTAGGTGGAAG	560
481	GAGGTLAGGTGGAAG	560

Figura 82: Amplificado de la SEQ ID NO: 62

[illegible]

102	CTAGCGGTTACAAAAGTGTGCTTTTGGCTGGCCATGATGCTAAAGACGAGTTCACAAAGCGCGAGTGTGTGTGCCA	23
102		A 23
102	ATACCGGTTAATAAAGTCGTTATTTGGTGTATATGATGGTAAAGAGTGAAGTTGTTAATAAAGCGCGAGTGTGTGTGTTA	23
102		TATTAAGTGTGTGAGTGTG
102		<u>TATTAAGCGCGAGTGT</u>
22	GACAGTGTAAATCAGTGTTCG	-57
22	CATAGTGTAAATCAGTGTTCGG	-57
22	GATAGTGTAAATCAGTGTTCGG	-57

Figura 83: Amplificado de la SEQ ID NO: 8

1	GGGAGTTTTTAAAGCTCTGTGAGAAATCCCTGGGAGTTGGTGATGTCAGACTACTTGGGTCAATTGAAGGTTAGCAGCCCGGG	80
1	GGGAGTTTTTAAAGTTTTGTCAG	80
1	GGGAGTTTTTAAAGCTCTGTGAGAAATCTGGGAGTTGGTGATGTLAGALTAGTTGGGTLATTTGAAGGTTAGLAGTTCGGG	80
81	TAGGGTTCACCGAAGTTCACCTGSCATATATTAGSCAATTCATCTCTGTGTGACACAGTAGTAGGAACTCAG	160
81	TAGGGTTTATCGAAAGTTLATTCGATATATTAGGTAATTTAATTTTATTTCTGTGTGATAGAGTAGTAGGAACTCAG	160
161	CTGCTCAGAGCCAGGAGGCTCTATTCTTCCCAAGGGGGACAGAAATCCCAATCGAGCTGTTGAGACTGGAT	240
161	TTGTTTAGGGLAGGAGGGTTTATTTTCTTCTTAAAGGGGGGATAGAAATTTTATCCGAGCTGTTGAGACTTCCGAT	240
161	GGAT	240
161	AT	240
241	GCCGACACCGGAGCGATCCGACAGGGTTTGTCTGGGCACCGTCGGGGTAGGATCCGGAACCGCATTCGGAAGGCTTTT	320
241	GtCGACACCGGAGCGATCCGAGLAGGGTTTGTCTGGGATCCGTCGGGGTAGGATCCGGAACCGCATTCGGAAGGCTTTT	320
241	GTGAGAAATGTA	320
241	GTGAGAAATGTA	320
241	ATCGATTCGAGTAGGG	320
241	ATCGATTCGAGTAGGG	320

321	GAAGCAATTACCTTGGAGGAGAACTTGGATCTTTC	TGGCAACCCCGCCCTGGCTCCATTGCGCGACCAACCCCTGGA	400
321	GAAGTATTTATTTGGAGGAGAACTTGGCAATTTT	TGGCAATTTTCCTTCCGTTGGATTGGLCGAGLAAAGCTTGGAA	400
401	AAATGGTAAATGATCATTTGGATCAATTACAGGCTTT	TAGCTGGCTTGCTGTCATATTCATGATTCGGCGCTGGGAAA	480
401		GGCTTGGGAAA	480
401	AAATGGTAAATGATCATTTGGATCAATTACAGGCTTT	TAGCTGGCTTGCTGTCATATTCATGATTCGGCGCTGGGAAA	480
481	AAGACCAA		560
481	AAGATTAA		560
481	AAGATTAA		560

Figura 84: Amplificado de la SEQ ID NO: 9

1	ACAGCGGGTAGTCCCCACACTCTCCGTGTCTGATCCCTCCCTCTCCTCCCGAGTCCACCGCGCCAGGCCACAGGTTTTCGG	80
1	ACAGCGCGTAGTCTTTATTTT	80
1	AGAGGGGGTAGTtttttLALtTttTGTTTCGAttttttttttttttECAGTtttACGtttttAGGCCTAtACGTTTTCG	80
1		80
1		80
	<u>TTTATGTCTTAGCTCTAT</u>	
	<u>TATCGTTTTAGGCGTA</u>	
81	CCAGATGTCTTTTCTTCTTCGCGACTCTTTGCGCGAGCGCTTCGAGAGCCAGTTCGCACTGATCGCCTTGGCATCGGATA	160
81	ttAGATGLTtttttATLttTCGLAGTtttTGTtTCGAGCGTttTCGAGAGCTtAGTttTCGATTCATCGctttttTCGATGGGATA	160
81		160
81	<u>GAAGTttttTCACACTT</u>	
	<u>AGCGTTTTTCGAGAGTT</u>	
161	CCGCGCGAGCGCCAGAGGACACTTGGCTTCTCTCCAGSAACTGAGCGGTCCTGAGGTCGCGGGCGCGACGGCAATCCCC	240
161	tCGGGGGAGGGtACAGGGAAtTTGGTttttTtTtTAGSAACTGAGCGGtttTCAGGTTtCGCGCGCGtAGGGAAtttt	240
161		240
161	<u>AAATGCACTGGTTTGTGA</u>	
	<u>ATTGAGCGGTTTTTGA</u>	
241	TCtCCGCGCGCGCGCGCGTGTCTGGTGTACGtCtTTAGAGGTCGAGGAAGTCACGTCGGGACAGACATGGGGCCAGT	320
241	TtTtTCGtCGtCGtCGtCGtTtGGTtTGtACGtTtTAGAGGTCGAGGAAGTtACGTCGGGAtAGAtTGGGGCCAGT	320
241		320
241	<u>GGACATTATGTTGGGA</u>	
	<u>TTACGTCGGGATAGAT</u>	

321	AGGTTAAGAAAGGCTGACATCTT	400
321	AGGTTAAGAAAGGTTGATATCTT	400
321	AGGTTAAGAAAGGTTGATATCTT	400

Figura 85: Amplificado de la SEQ ID NO: 10

1 AGTGGGTAGGCCAAGTGCTTGCTTCAGCAACCGGACGAGGAGGCGACGGCCGGATGTGGGACCTCTTCCTCTAGC 80
1 **AGTGGGTAGGTTAAGTGTGTTG** 80
1 AGTGGGTAGGTTAAGTGTGTTGTTTACTAATCAACGGATLACGAGGGCTAGGGTGGATGTGGGATLTTTCTTTTAGT 80
1 80
1 AGTAAACCCCAATTACCA 80
1 AGTAAATTGGATTAGGAG 80

81 ACAGTAAAGCTGGCTCCAGAAACACGGGATCTCCGGTGGTGTCTTTGGGTCGGCGGTGTGTGCGCGCTCGCGGGTGG 160
81 ALAGTAAAGTTGGCTTTLAGAAATACGGCTATLTCGGGTGGGTCTTTCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 160
81 ATATGGCTATTTTCTCT 160
81 TACGGCTATTTTCGGCT 160

161 GGTGTGAGGAGGGGACGAAGAGGCTGAAGGAGGGCAAGCGCGGGGGGCTCTGCGAGAGCGGCGCCDAGCCCCCTTGG 240
161 GGTGTGAGGAGGGGACGAAGGCGGAGGAGGGAAGGGCGGGGGCTTTGCGAGAGCGCGCTtAGtttCGttttCGG 240
161 GAGAGTGTGTTTTAGTTTT 240
161 AGAGCGCGTTTTAGTTTT 240

241 GCGCCACAGTCCCTGCACCCAGGTTTCGATTGGCGGGTCTCTCTCAGTCTCTTCGCGCGCCAGCTGTGGATCTCGGG 320
241 **TTTACTTTTCGATTTTTGGGG** 320
241 GATTTAAGTttttTGtttttACGTTTTtttATTGGCGGGtttttATTAGTtttttttTtAGTtttttttTtCGTtAGTtGGATtttTGGGG 320

321 AGTCTGGGAGCACATTCGAGGATACGCTCTCAGCCCTGCACATTAAGTGTCATTCAGAACAGATCTCTGATTC 400
321 AGTTTGCGGATCAtAtATTCGAGGATACGCTCTCAGTCTCTGATTCATTAAGTGCTATTAGAAAGATGCTLTGATTTT 400
321 ATATTTTCAGGATTATGTT 400
321 ATTTGAGGATTACGI

401 AAGGAGCCAGTCGCTGAACACAGAGGCTTTGGGTTGTCA 480
401 **TAGAGAGGTTTTGGGTTTGTTA** 480
401 AAGGAGLLAGTGCTGAACACAGAGGCTTTGGGTTGTGA 480

Figura 87: Amplificado de la SEQ ID NO: 12a

1	CCGACCTCTGAAAAGGGACCCAGGACCCGGCAGATGCTGGCGAGGAACAGGGGUCGAGCATGAGTCGGCTTTGCACCA	80
1	TTTACTTCTCAAAAAGGGATT	80
1	tttAGtTGTGAAAAGGGATLLAGGAtTCGLAGATGtTGGCGAGGAAtAGGGGTCTCGAGtATGAGTCGGtTTTGTAtTA	80
81	CGCCCTGGGCTCCACGACTGANGCCTTCAGTACCATGCAGACAGCTGGGAGGGSGTCCCGCTGGGCGGGGCCGAGCT	160
81	CGtTtTCCGCTTtTtAGtAtTGAAAGGCTTLAGTAtTATGLAGAGLAGTGGGAGCGGGTtTCGTTGGGCGGGGTCGAGT	160
161	GGGCGGACCCCTCGGCGCGCTCGGGGAGGCCAGACAGGGTGCGCTCTCTGGCTCCGCTCCGCGGGCTCTATGACAC	240
161	GGGCGGAtTtTCGGCGCGTTCGGGACGgttTAGAtAGCGCTGGtTtTtTGGtTtCGtTtCGCGGGtTtTATGAtAt	240
161		AT 240
161		AT 240
241	CGCActGGCTCCGCGGACGGGCGGGGTGCGGTGAGGGGGAGGAGACAGCCCAAGCTTTTCGACTCCGCGCTGACCCCTA	320
241	CGtAtTGGTTCGCGGACGGGCGGGGTGGGTTCAGCGGACGACAtAGtTtACGtTtTTCGCAAtTCGGGTGAtTtTtTA	320
241	CGTATTGGTTCCCGG	AGTATTATGTATTGTGAT
241	TGTATTGCTTTCTCG	TTTACGTTTTCGGATT
		ATTTTA 320

321	CCCGAACTCTCTCGGCTAATTCGAAGTACCGCGCTGGCTGCTGAGTACGCTCTGGTGGCGGTGGCAGTTTCCCGGC	400
321	CCCGGAAUUCCTCGGCTAATTCGAAGTATCGGCTTGGCTGTGTG-AGTACCGCTTGGTGGCGGTGGTACGTTTCTCCGC	400
321	TTGTTGAAAT	ATTGGAAGTATTCGGT
321	CCCGGAAT	ATTGGAAGTATTCGGT

401	GGGTGTGTGAAGGGAGACAGTGTGGAGGCCACAGG	480
401	GAGTAGTCTCGGACGTTATAGGC	480
401	GGTCTGTCAACGGAGATACGCTGGAGGTATAGGG	480

Figura 88: Amplificado de la SEQ ID NO: 12b

497	GTCTTCGAGGGAGTAGAGACCCCATCTTCAGGCTGCTCCCTGGATTAAGAAATGCCCTTTCATCATTCAAA	418
497		CCCTATAACCTCCACACTA
497	GTCTTCGAGGGAGTAGAGACCTCATCTTCAGGCTGCTCCCTGGATTAAGAAATGCCCTTTCATCATTCAAA	418
417	ATTATTTTACCTCCGCTTTTGTACTCCGAAACCCGAGTAGAGCAGCATGGGAGCCUAAAATAAGTGCATPACTTGCANTG	338
417	PCR	338
417	ATTATTTTACCTCCGCTTTTGTACTCCGAAACCCGAGTAGAGCAGCATGGGAGCCUAAAATAAGTGCATPACTTGCANTG	338
417	TTTGAATTTTGCAGTACAG	338
417	ATTTCGAATTTCCAGTA	338
337	GATAATTCGAGAGACCTGCTTGCACATAGTTTGTATATACCPAAAAATTGAAATCTATATTACACCCCATCAGTAG	258
337	GATAATTTTLAGAGAACTGCTGCTTGCACATAGTTTGTATATACCPAAAAATTGAAATCTATATTACACCCCATCAGTAG	258
257	GAAATGGATPAATGAAATAACCTCTATCCCTGTTTGAATATGATTACAGCCGTTAAACCGTAGATCTGTGCTCTGA	178
257	GAAATGGATPAATGAAATAACCTCTATCCCTGTTTGAATATGATTACAGCCGTTAAACCGTAGATCTGTGCTCTGA	178
257	AGTTTATAAATGTAGAT	178
257	TAGCTCTTAAACCTAG	178

177	CTGGCTACCCAGGACATGTTAAATGCCAAACAACTAGCAGAAAGGCCCTACGTAGIGATACCATTTATGTTATTTTA	98
177	CTGGCTATCTTAGGALATGTTAAATGGGAAALAACTAGLAGAAGGGCATCTACGTAGIGATATCTTTATGTTATTTTA	98
97	AAATTAAATGTPACATTTATATAGGTGTACCTGTGTATGAACTGCCACTTTAAAGCCCTAAACAGAAAAATACATCAGAAGCT	18
97		AAATC 18
97	AAATTAAATGCTATATTATATAGGTGTACCTGTGTATGAACTGTGTATGTTAAAGCTTTAAATAGAAAAATATATACAAAGCT	18
97		TGTATGAACTGTGTATGTT 18
97		GTGTAATGAAATGCTGTCT
17	TTTAATCTGTGGGATCGG	-62
17	CGTTTTTCACAACTAAA	-62
17	TTTAATCTGTGGGATCGG	-62

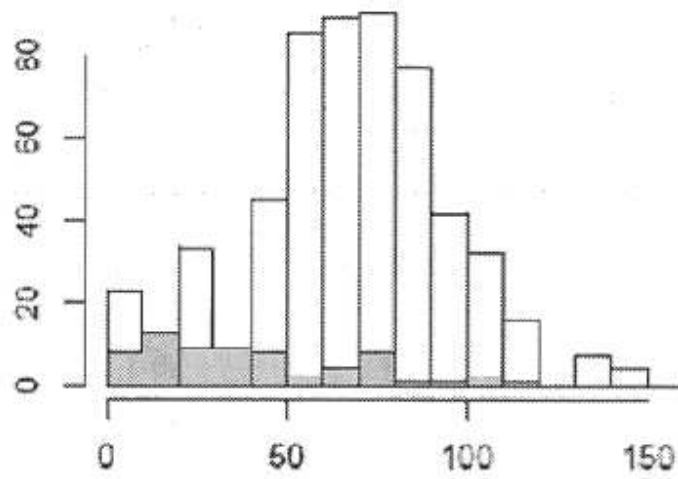


Figura 89

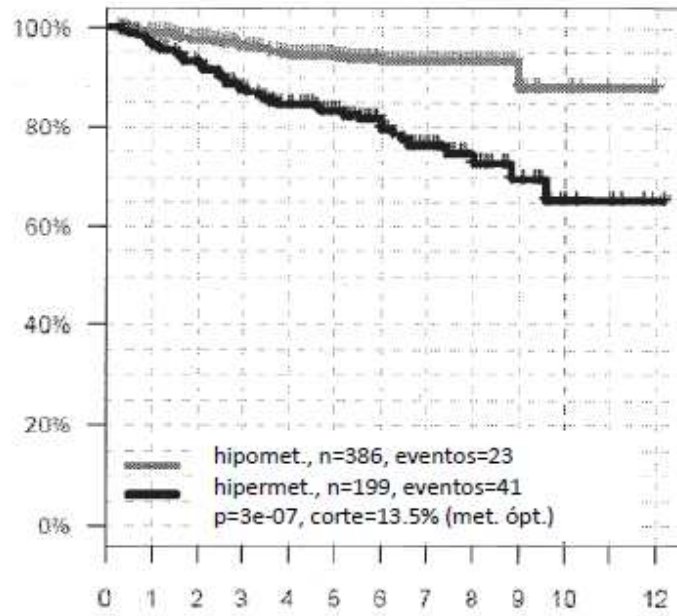


Figura 90 A

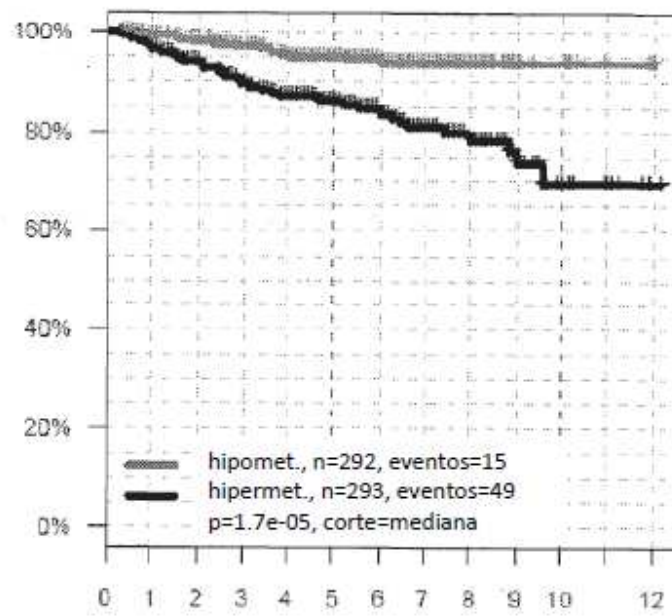


Figura 90 B

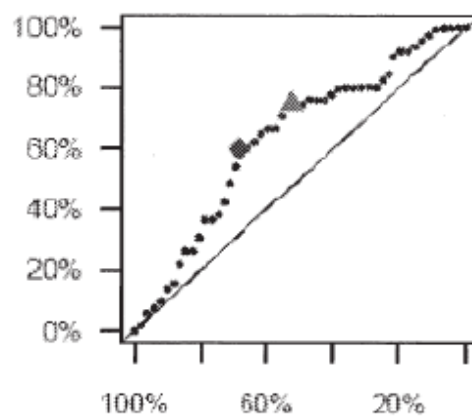


Figura 90 C

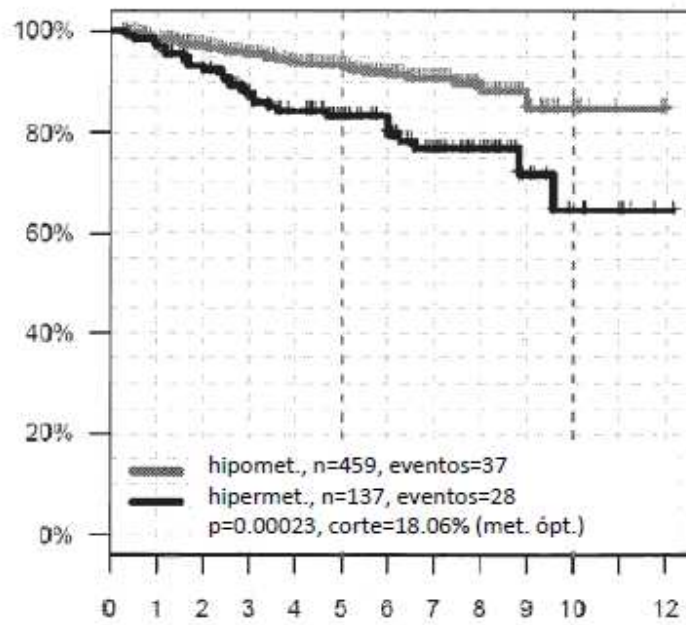


Figura 91 A

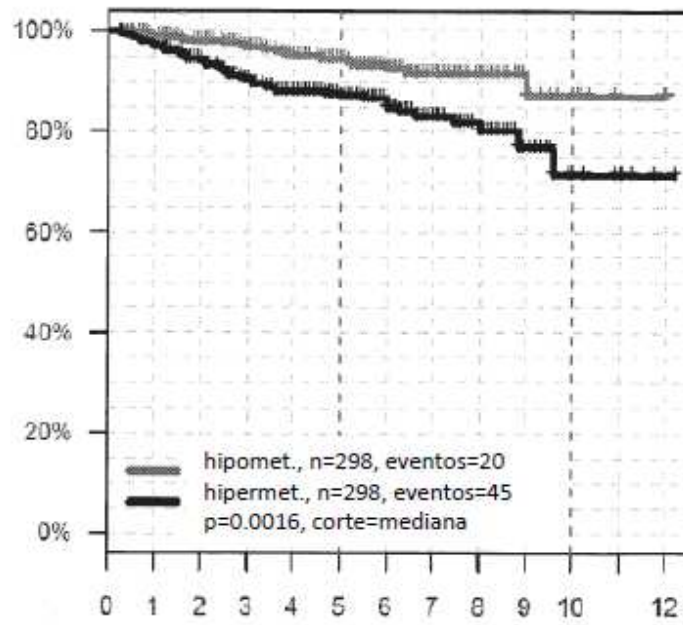


Figura 91 B

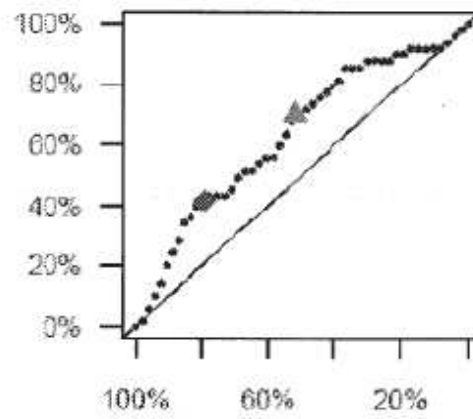


Figura 91 C

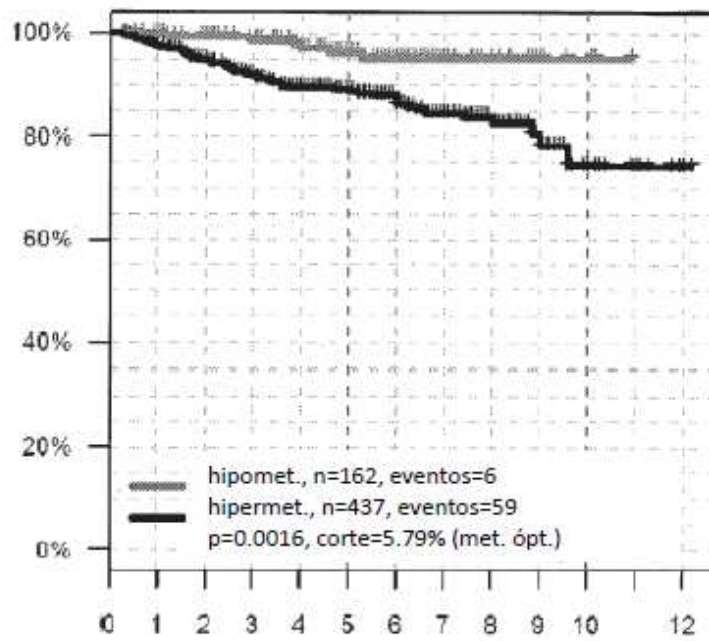


Figura 92 A

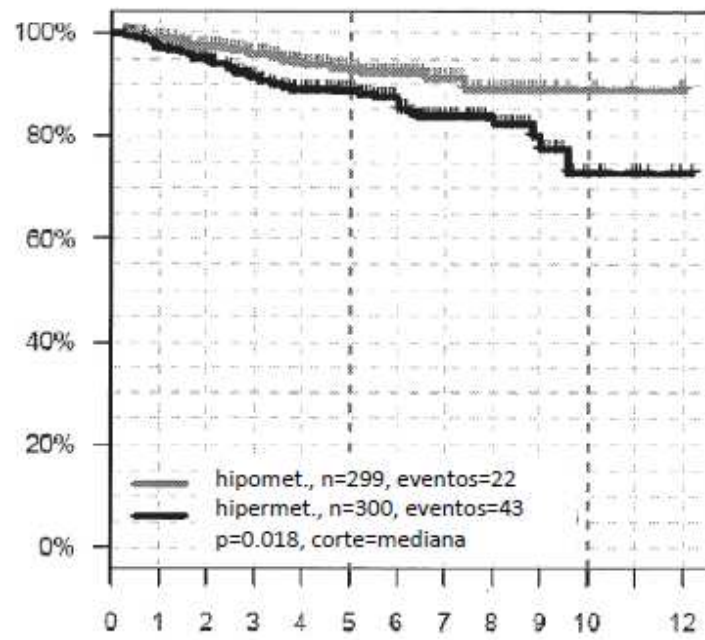


Figura 92 B

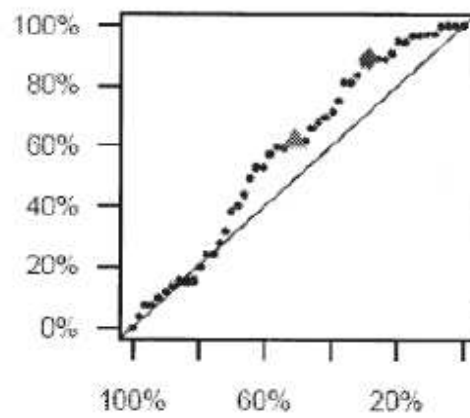


Figura 92 C

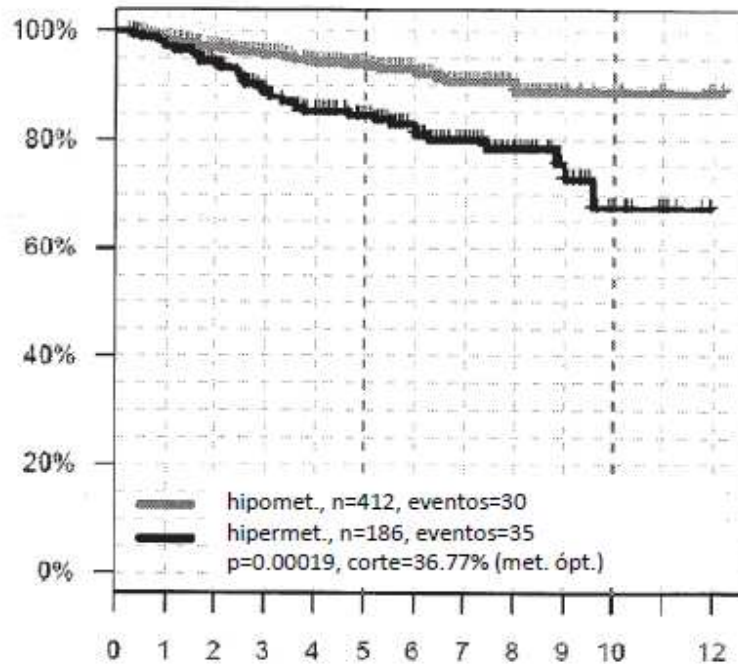


Figura 93 A

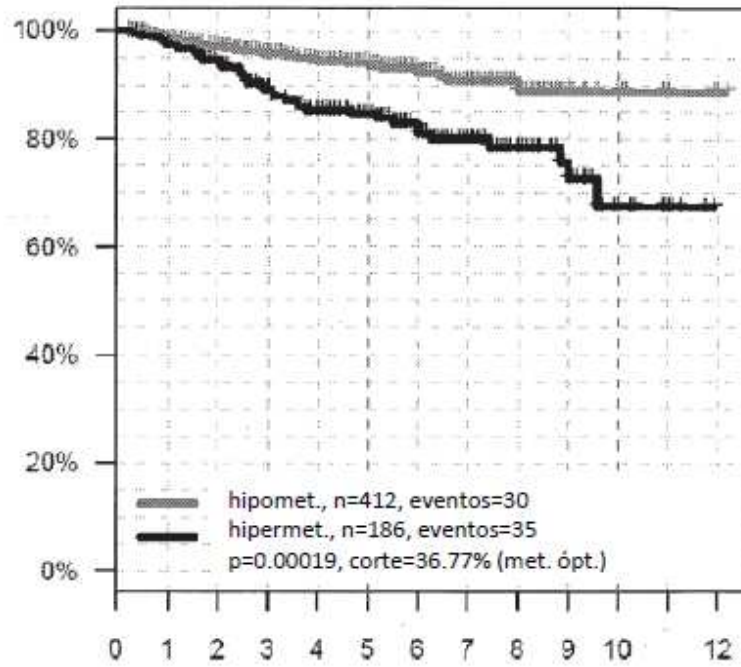


Figura 93 B

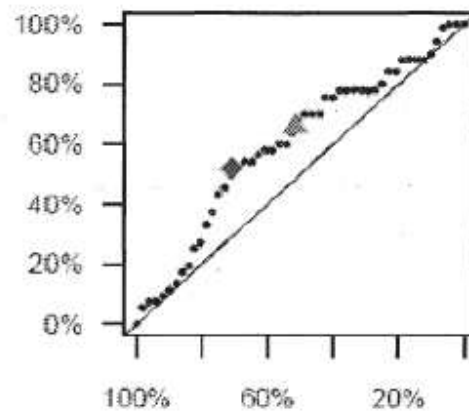


Figura 93 C

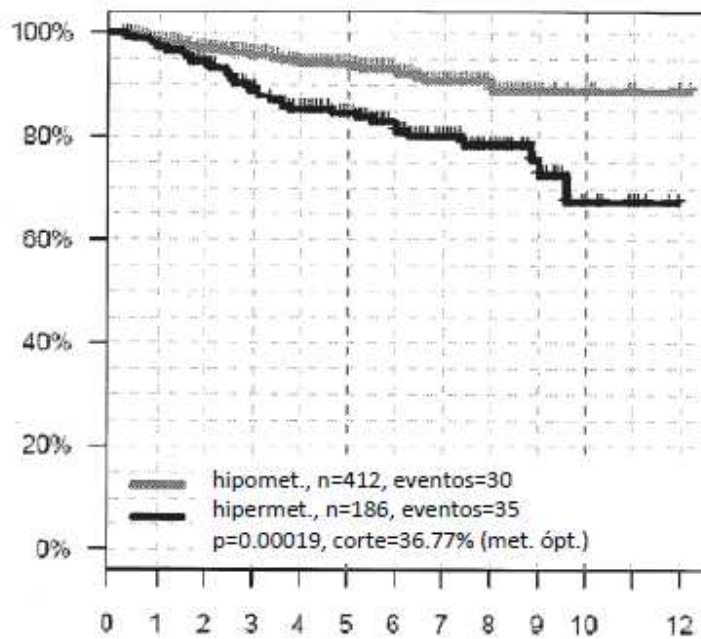


Figura 94 A

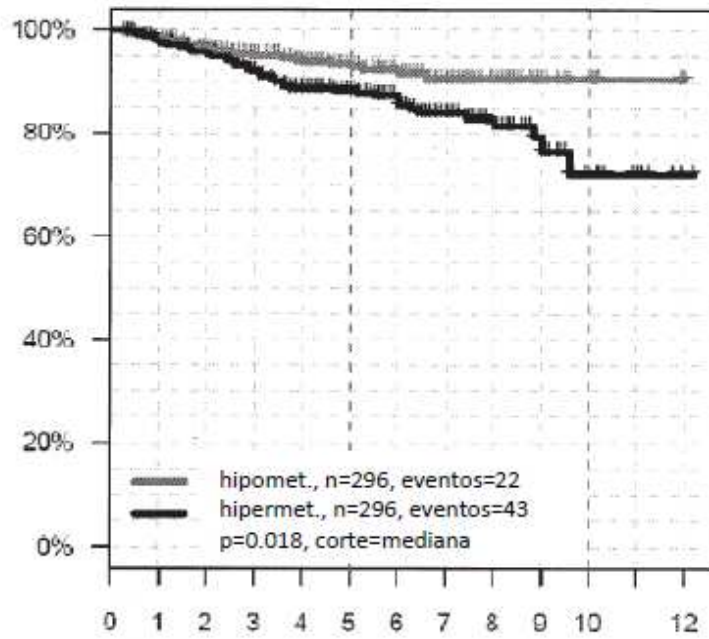


Figura 94 B

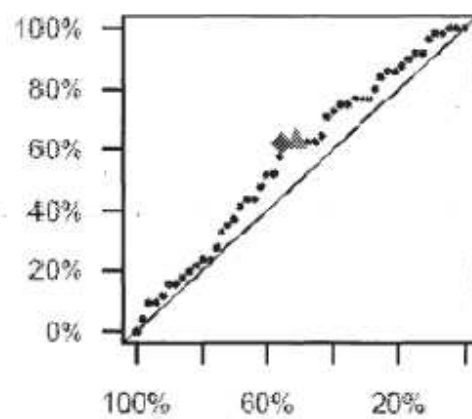


Figura 94 C

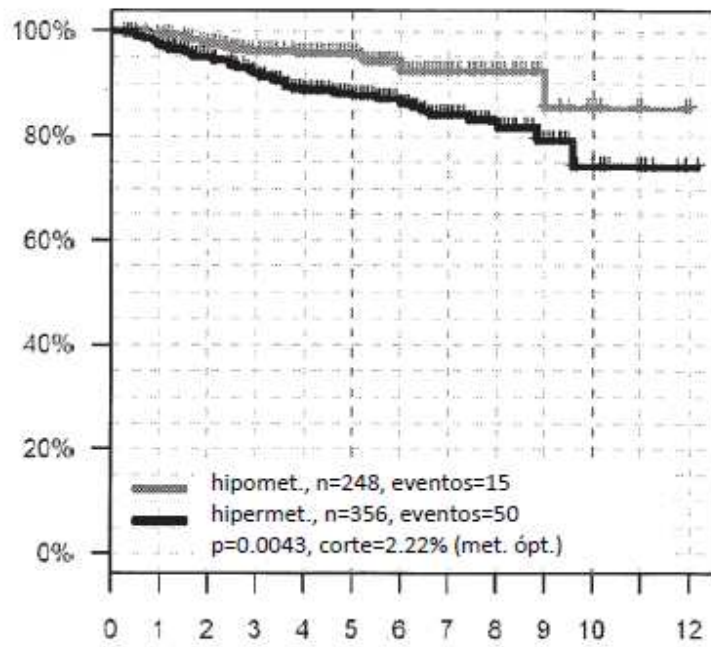


Figura 95 A

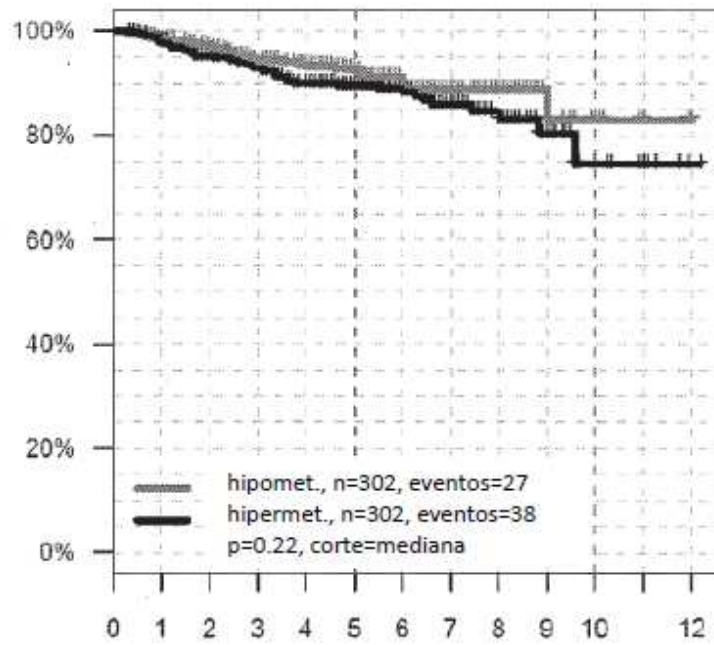


Figura 95 B

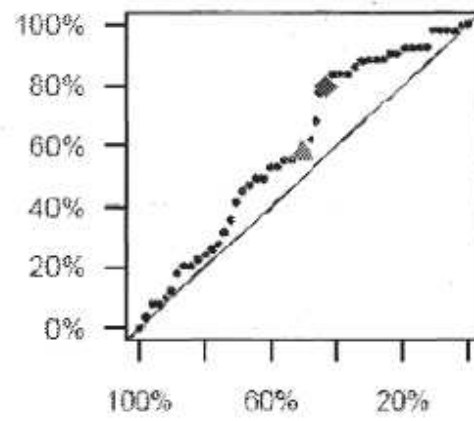


Figura 95 C

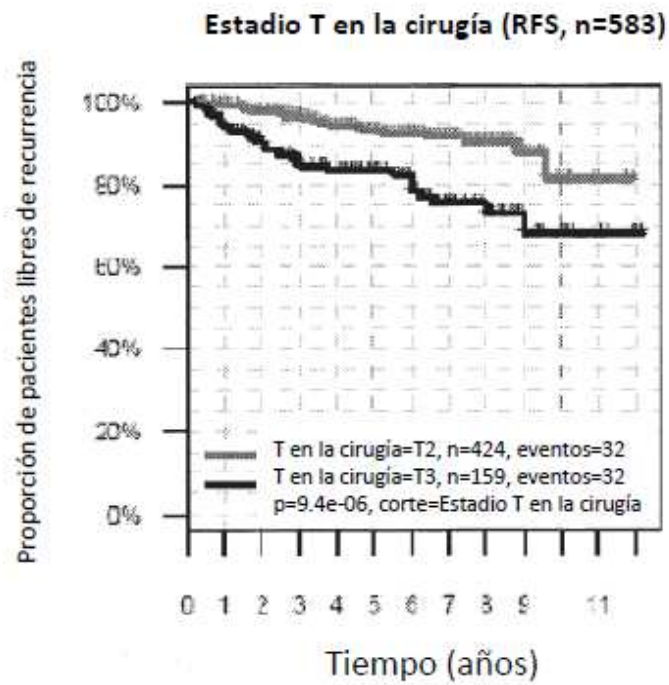


Figura 96/1

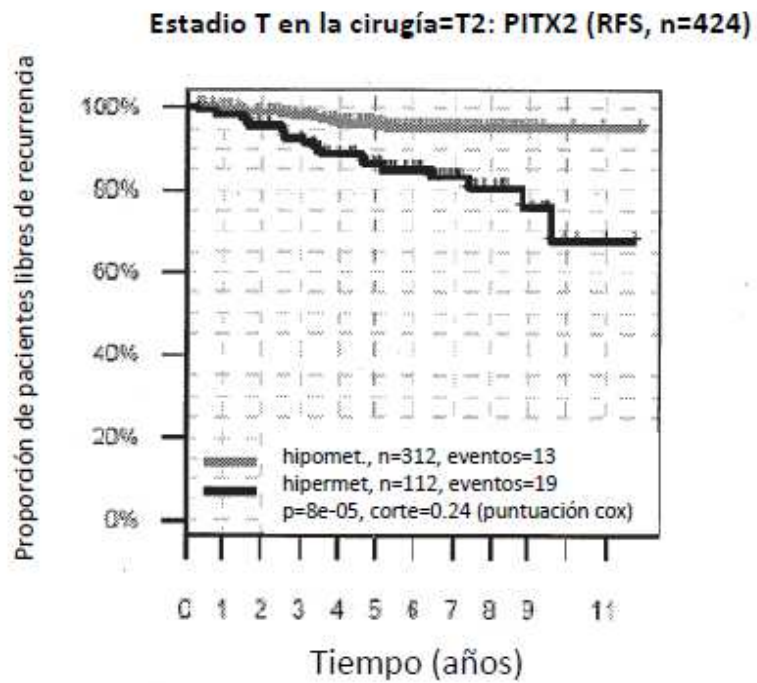


Figura 96/2

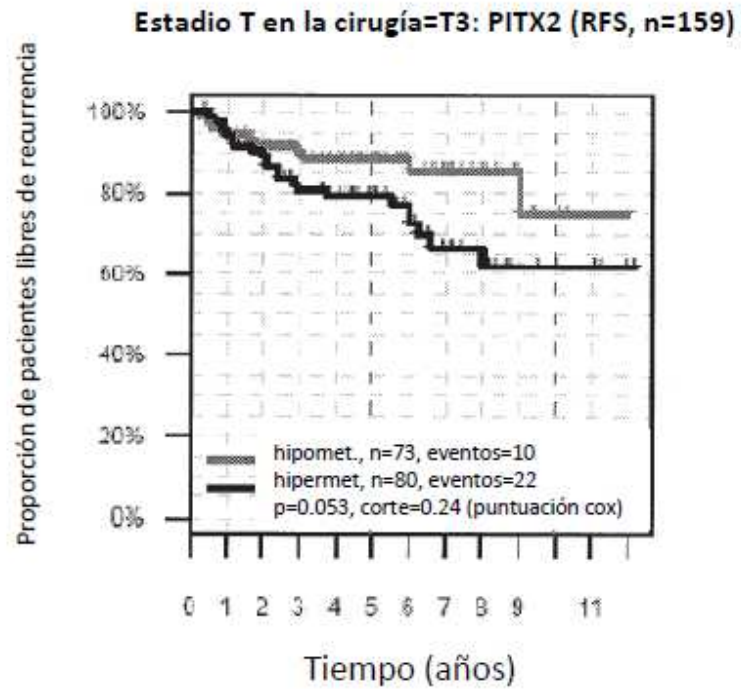


Figura 96/3

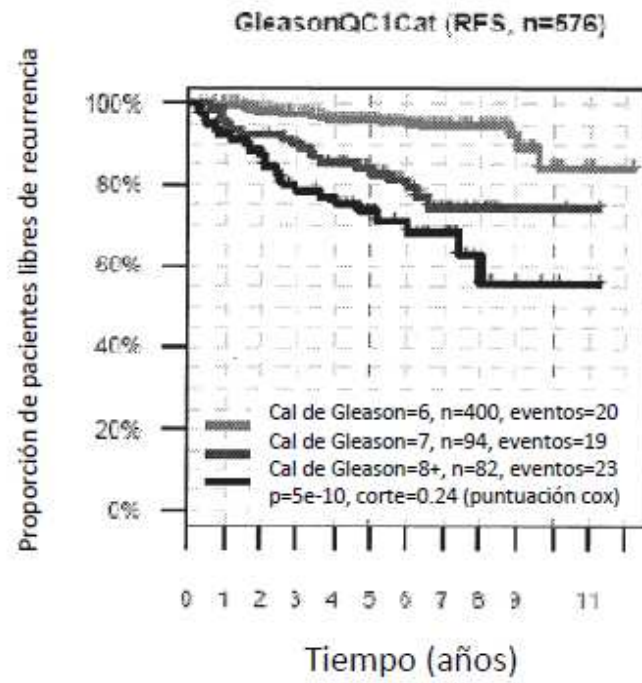


Figura 97/1

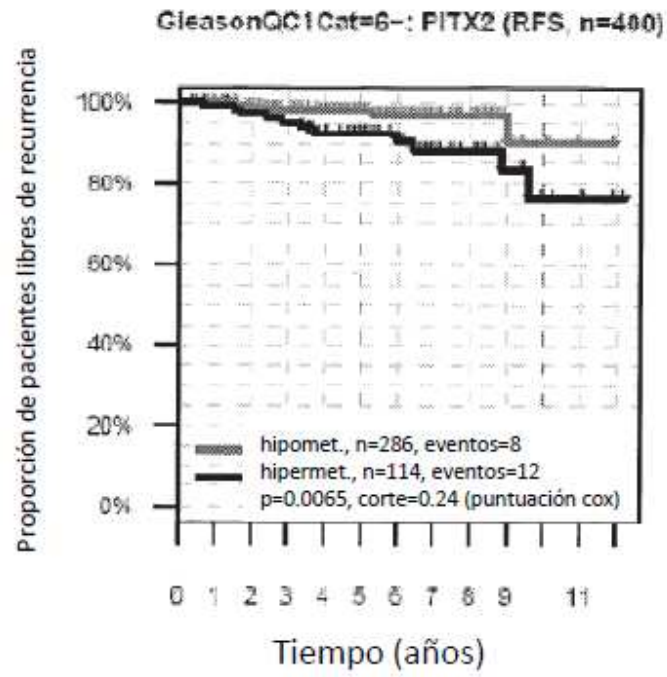


Figura 97/2

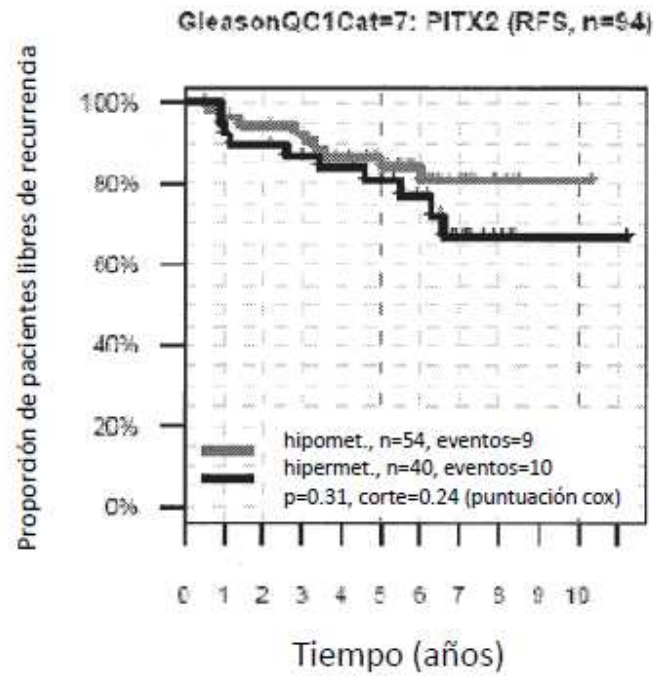


Figura 97/3

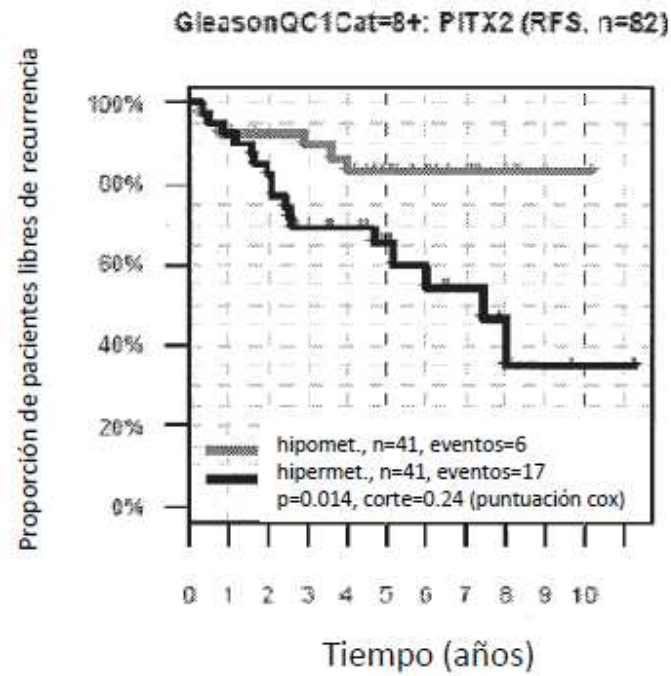


Figura 97/4

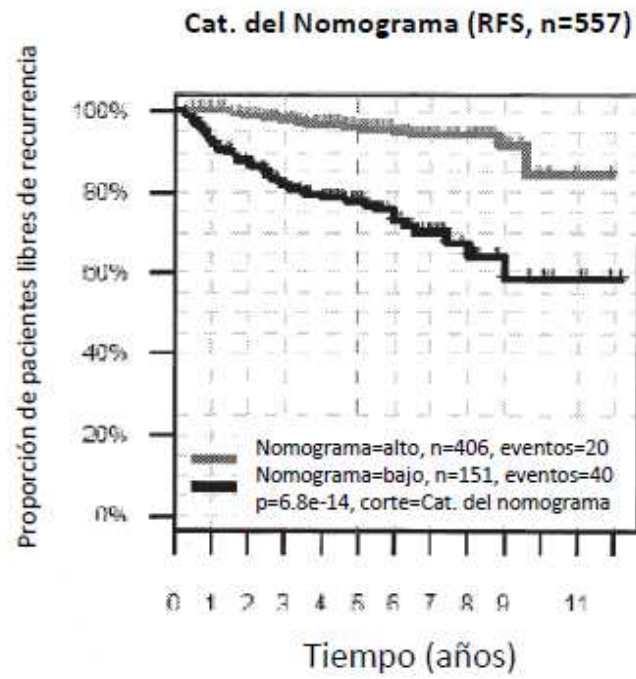


Figura 98/1

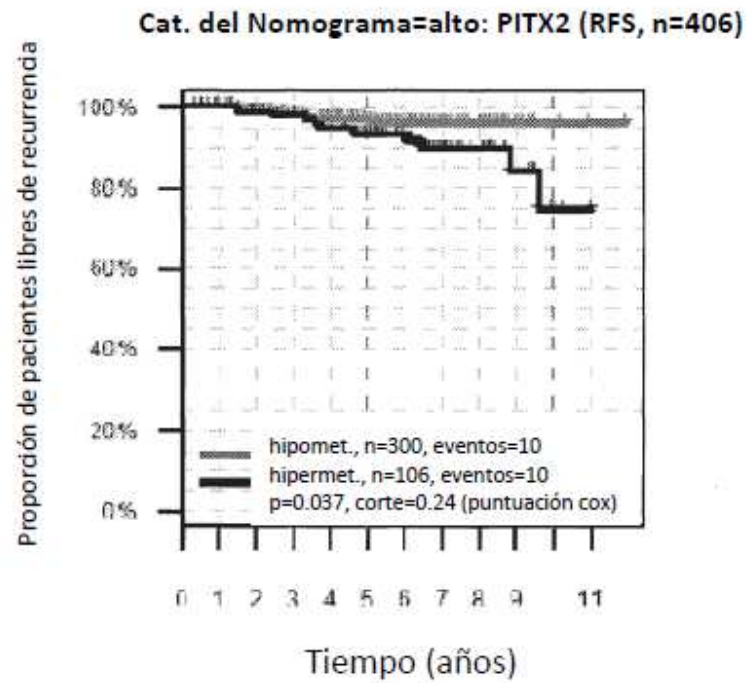


Figura 98/2

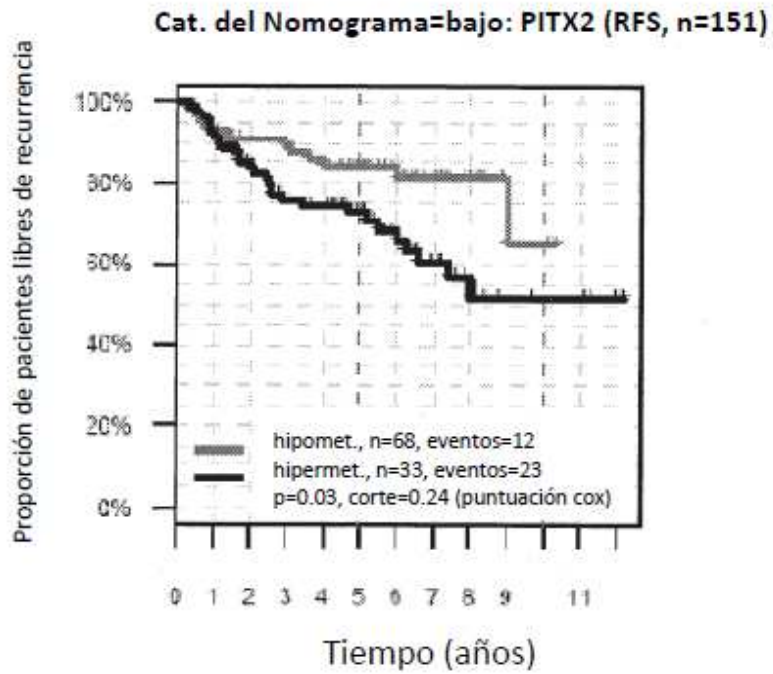


Figura 98/3

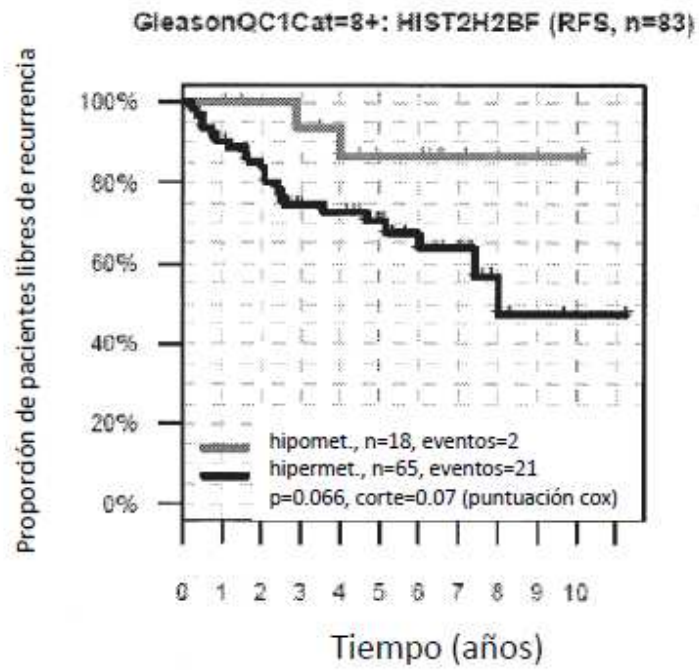


Figura 99

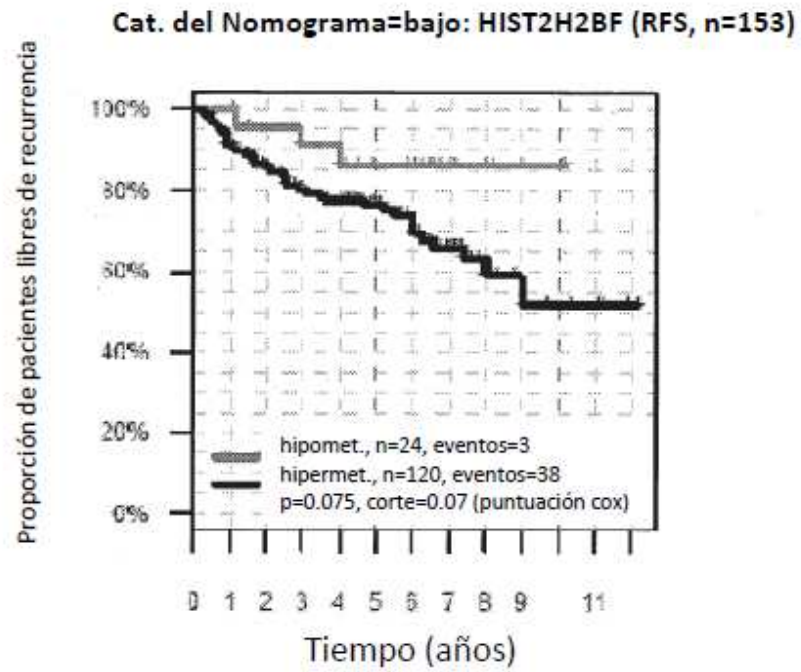


Figura 100

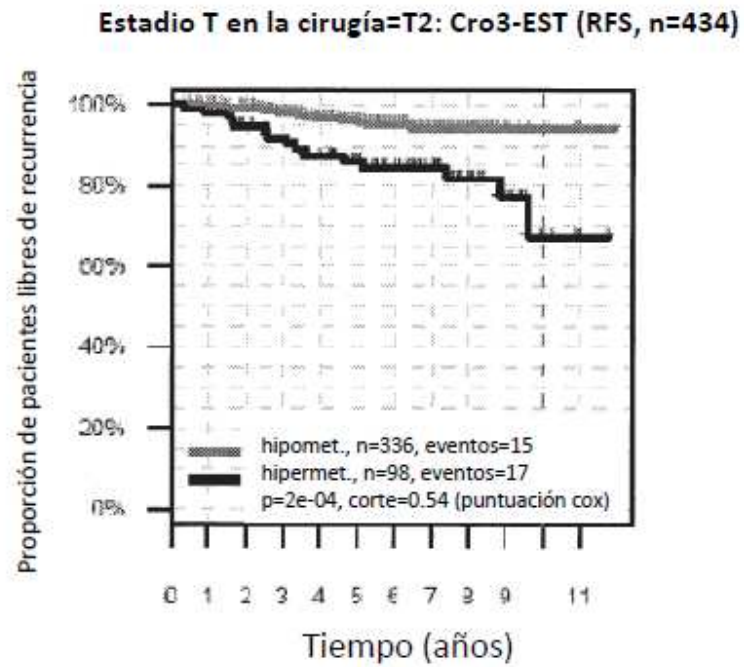


Figura 101

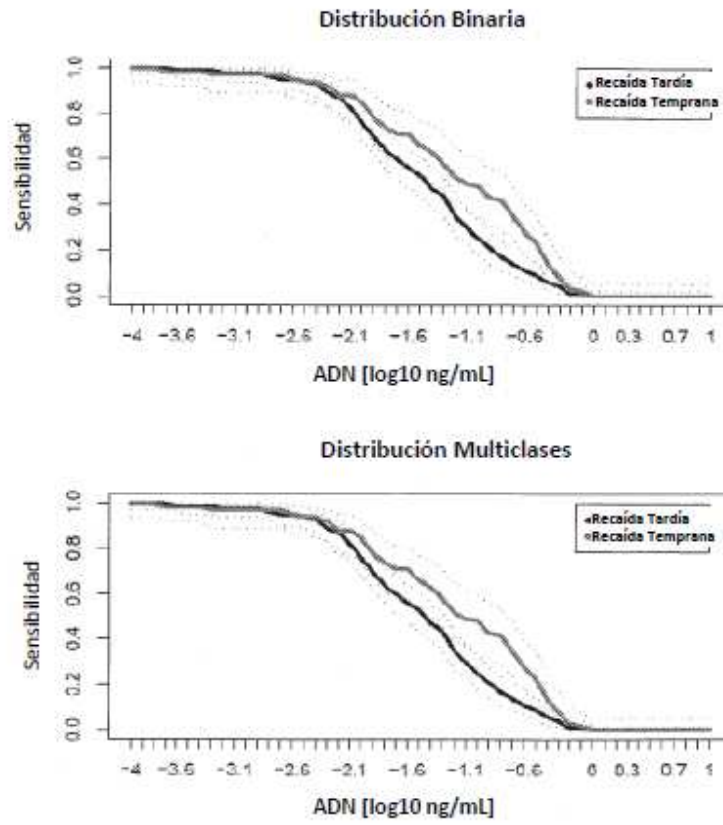
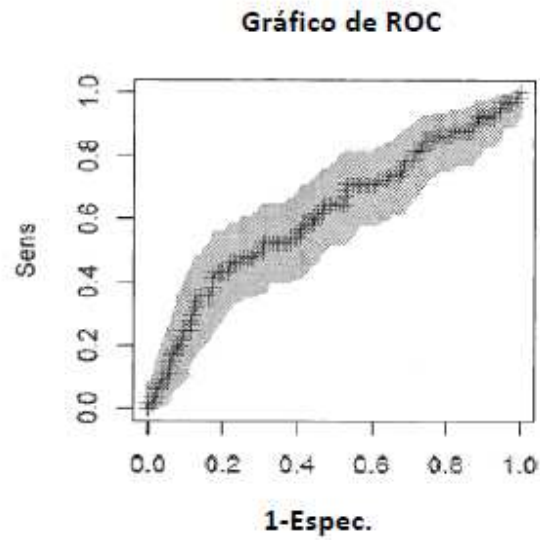


Figura 102/1



AUC	0.62 [0.54,0.68]
Sens/Espec	0.35 / 0.86
P de Wilcoxon	0.0072
Recaída Temprana (pos)	65
Recaída Tardía (neg)	140

Figura 102/2

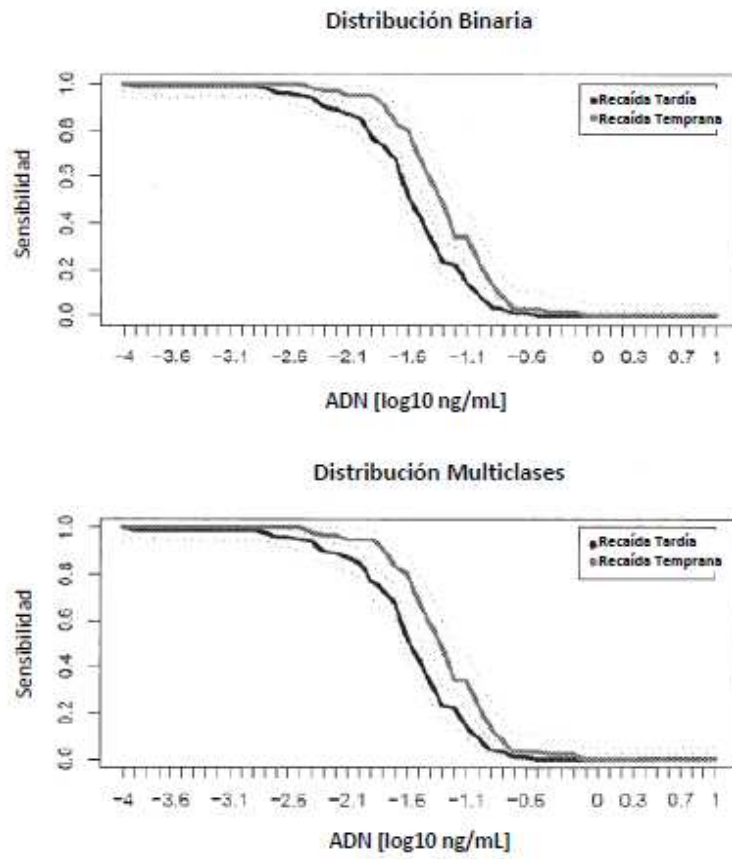
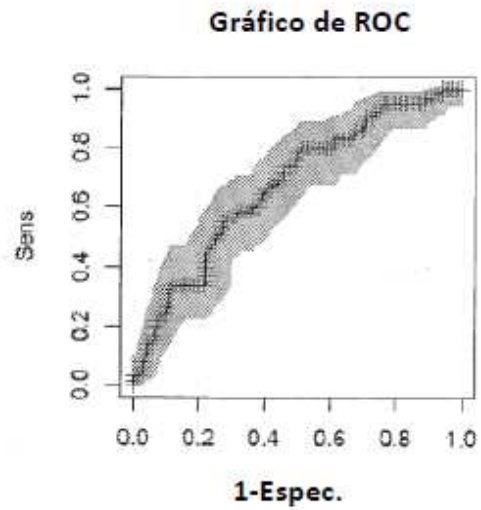


Figura 103/1



AUC	0.68 [0.61,0.74]
Sens/Espec	0.34 / 0.85
P de Wilcoxon	1e-04
Recaída Temprana (pos)	65
Recaída Tardía (neg)	141

Figura 103/2

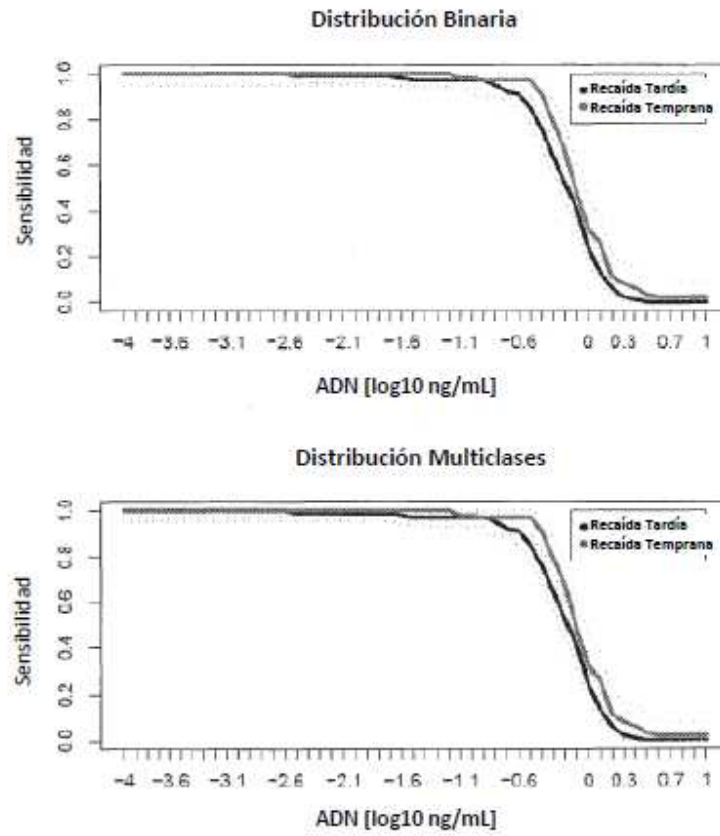
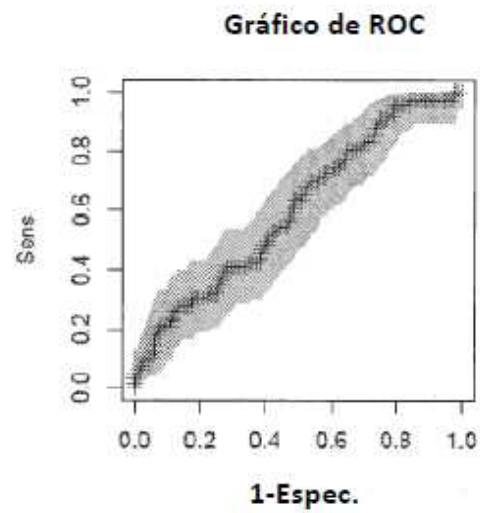


Figura 104/1



AUC	0.6 [0.53,0.67]
Sens/Espec	0.27 / 0.85
P de Wilcoxon	0.0178
Recaída Temprana (pos)	66
Recaída Tardía (neg)	143

Figura 104/2

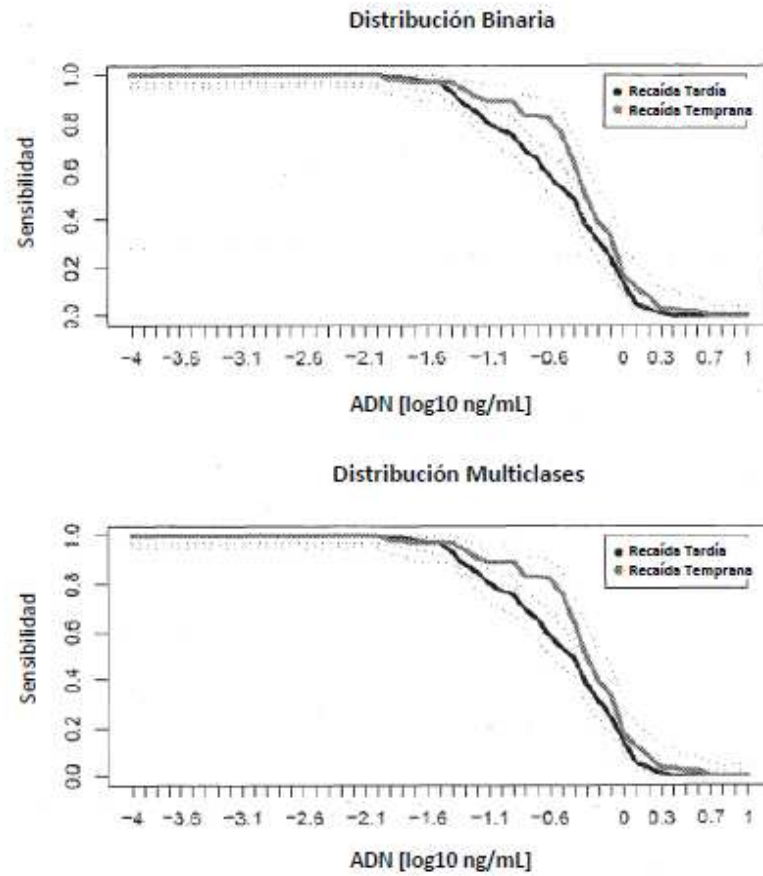
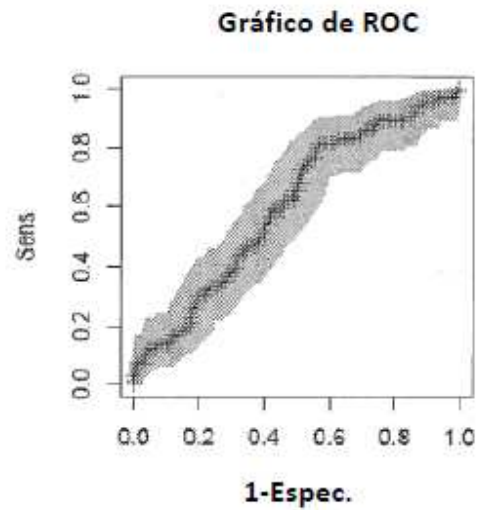


Figura 105/1



AUC	0.61 [0.53,0.67]
Sens/Espec	0.18 / 0.85
P de Wilcoxon	0.0146
Recaída Temprana (pos)	66
Recaída Tardía (neg)	143

Figura 105/2

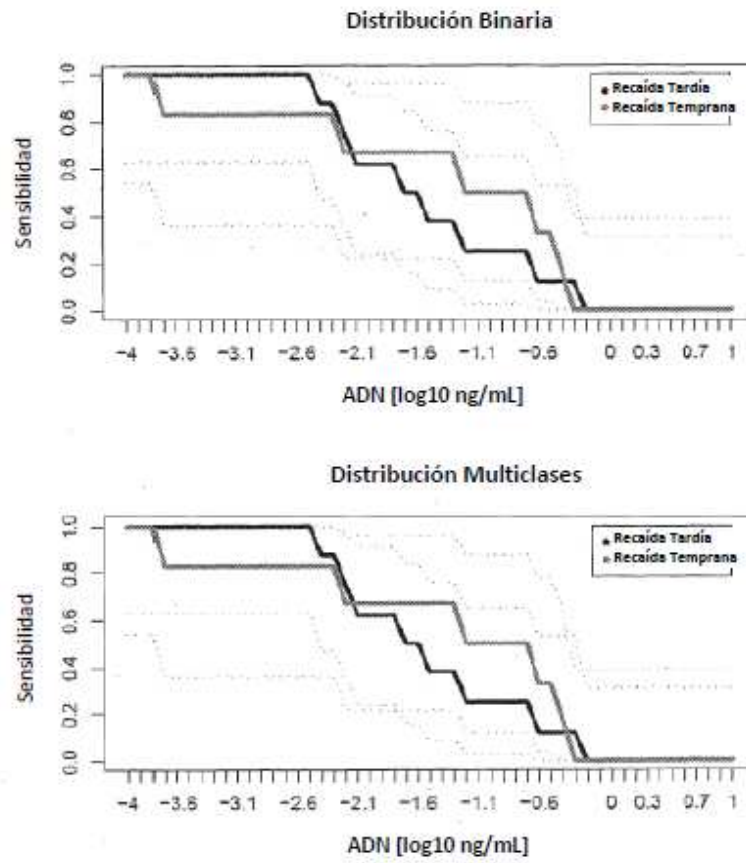
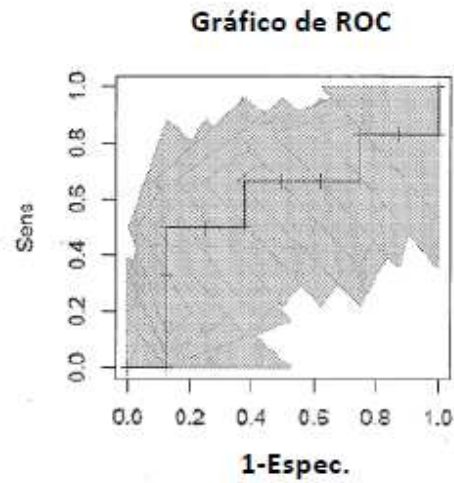


Figura 106/1



AUC	0.58 [0.29,0.82]
Sens/Espec	0.5 / 0.88
P de Wilcoxon	0.662
Recaída Temprana (pos)	6
Recaída Tardía (neg)	8

Figura 106/2

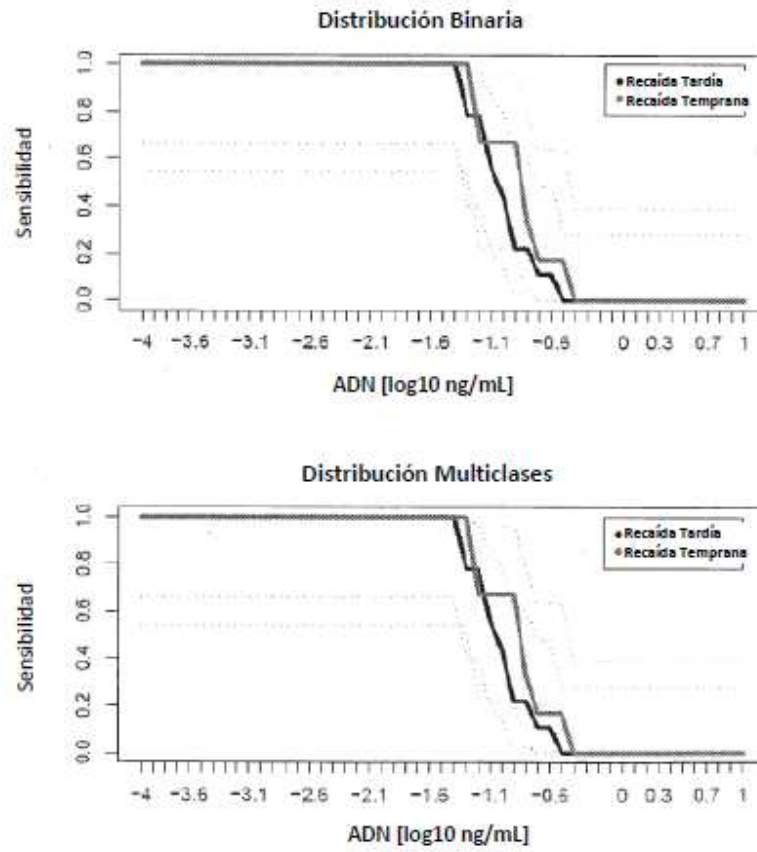
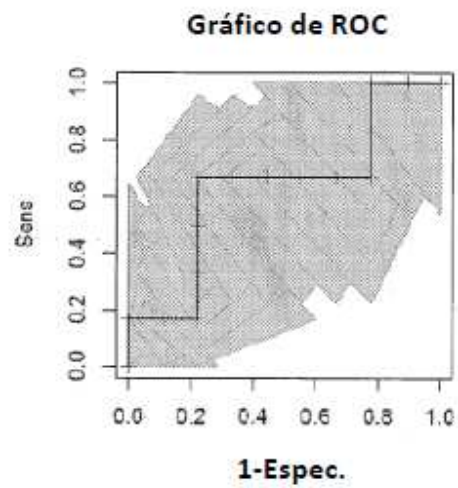


Figura 107/1



AUC	0.63 [0.32,0.84]
Sens/Espec	0.17 / 0.89
P de Wilcoxon	0.4559
Recaída Temprana (pos)	6
Recaída Tardía (neg)	9

Figura 107/2

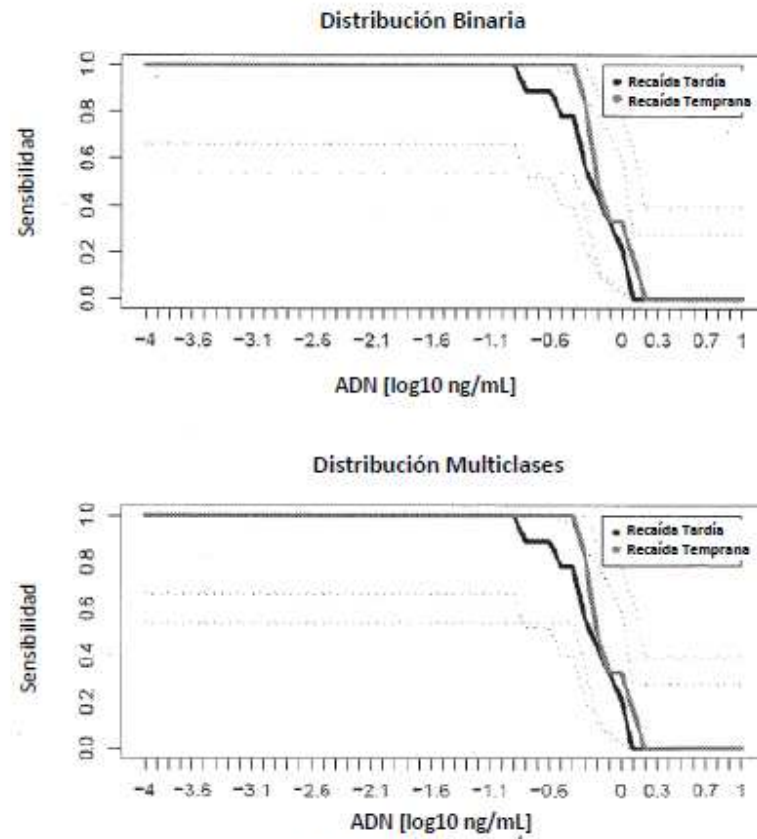
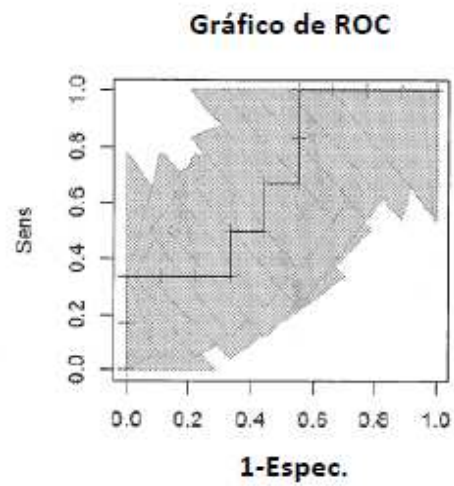


Figura 108/1



AUC	0.69 [0.38,0.88]
Sens/Espec	0.33 / 0.89
P de Wilcoxon	0.2721
Recaída Temprana (pos)	6
Recaída Tardía (neg)	9

Figura 108/2

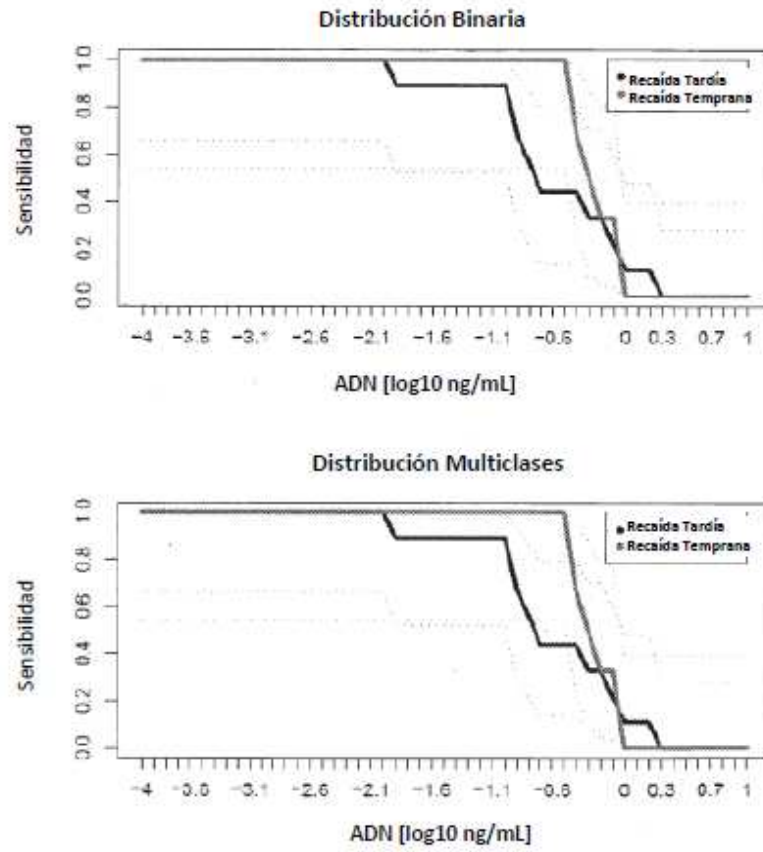
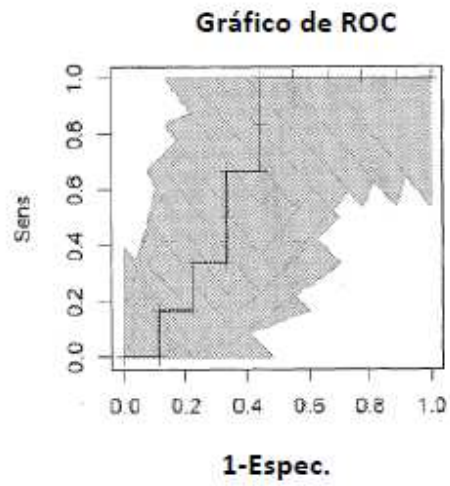


Figura 109/1



AUC	0.69 [0.38,0.88]
Sens/Espec	0.17 / 0.89
P de Wilcoxon	0.2721
Recaída Temprana (pos)	6
Recaída Tardía (neg)	9

Figura 109/2

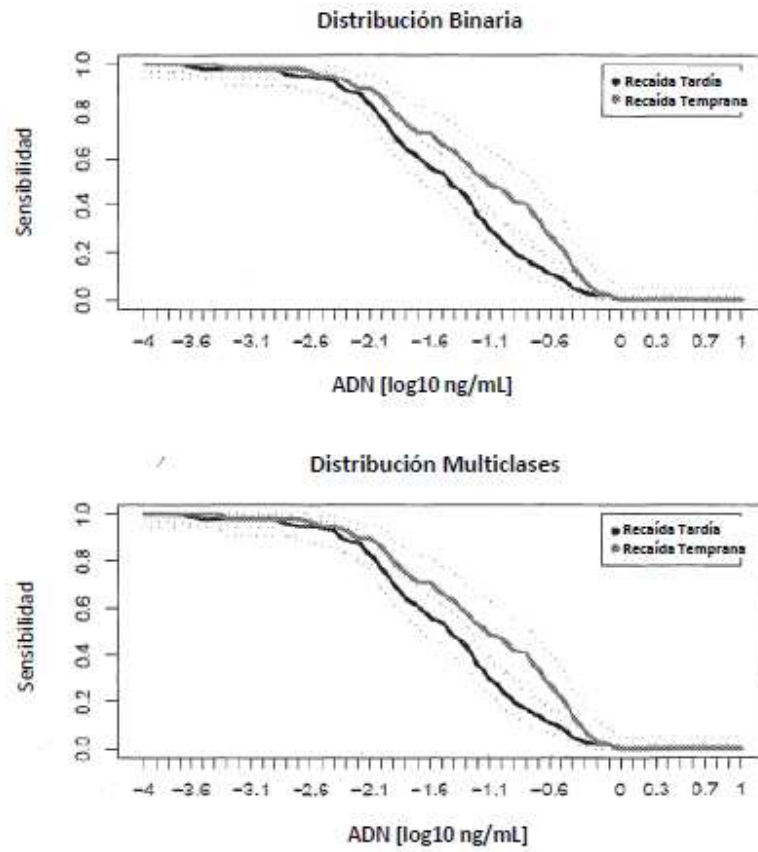


Figura 110/1

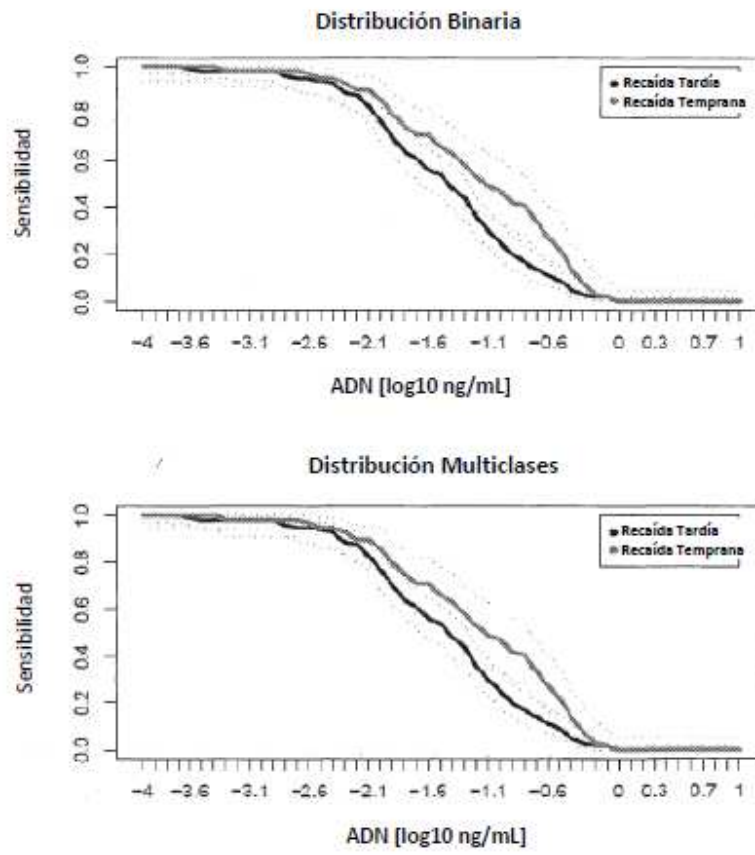
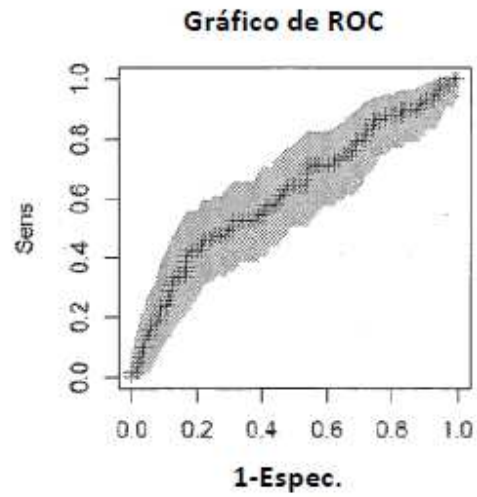


Figura 110/1



AUC	0.62 [0.54,0.69]
Sens/Espec	0.34 / 0.86
P de Wilcoxon	0.0085
Recaída Temprana (pos)	59
Recaída Tardía (neg)	132

Figura 110/2

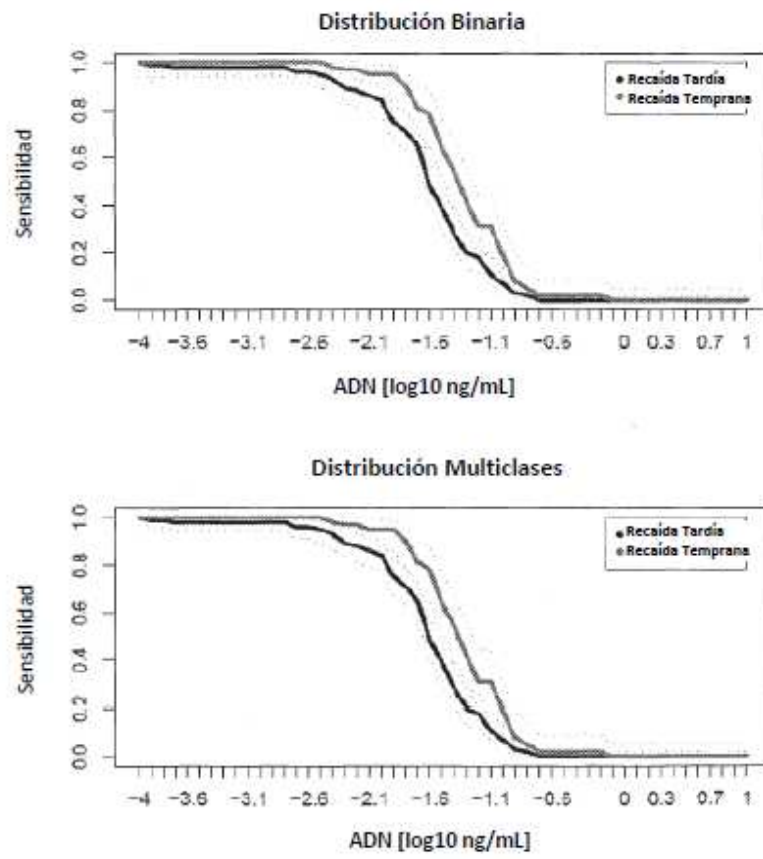
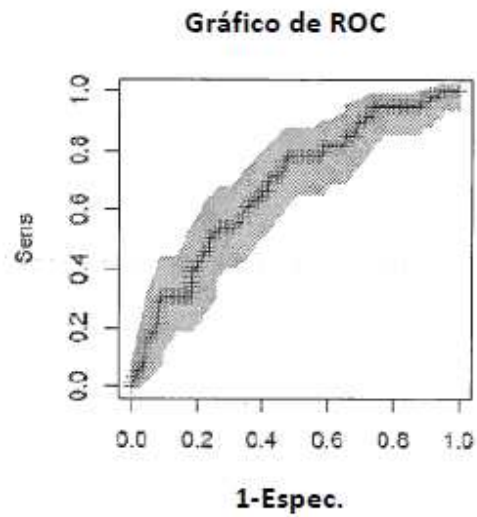


Figura 111/1



AUC	0.68 [0.61,0.75]
Sens/Espec	0.31 / 0.86
P de Wilcoxon	1e-04
Recaída Temprana (pos)	59
Recaída Tardía (neg)	132

Figura 111/2

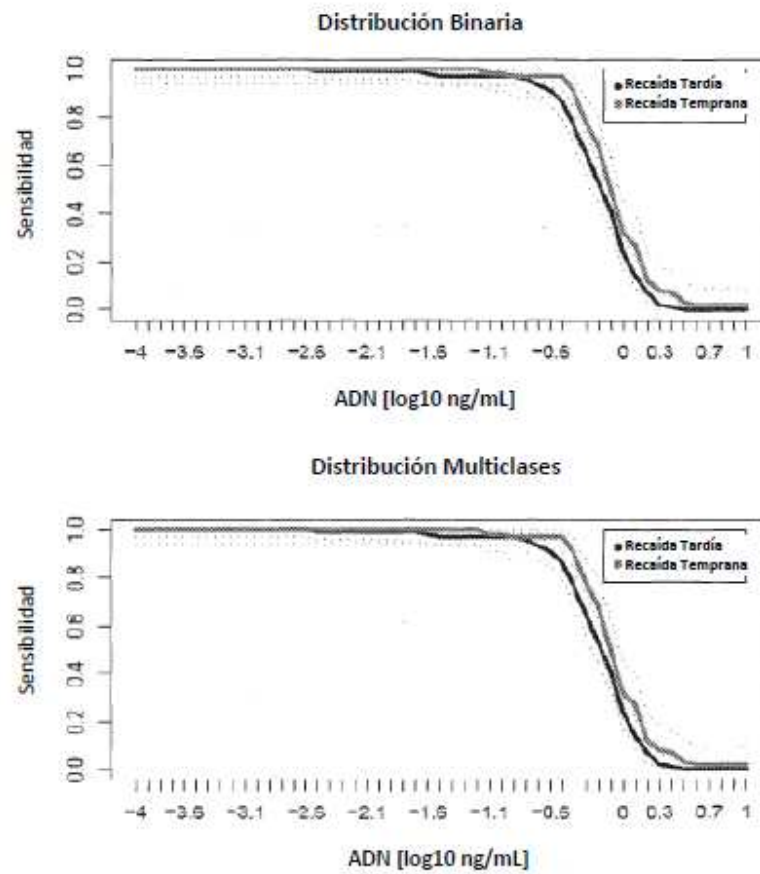
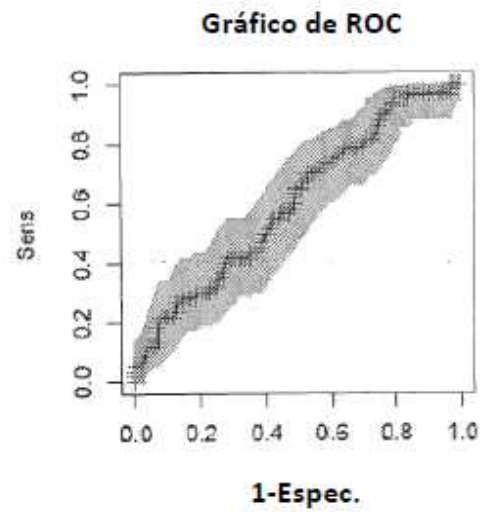


Figura 112/1



AUC	0.6 [0.53,0.67]
Sens/Espec	0.28 / 0.85
P de Wilcoxon	0.0264
Recaída Temprana (pos)	60
Recaída Tardía (neg)	134

Figura 112/2

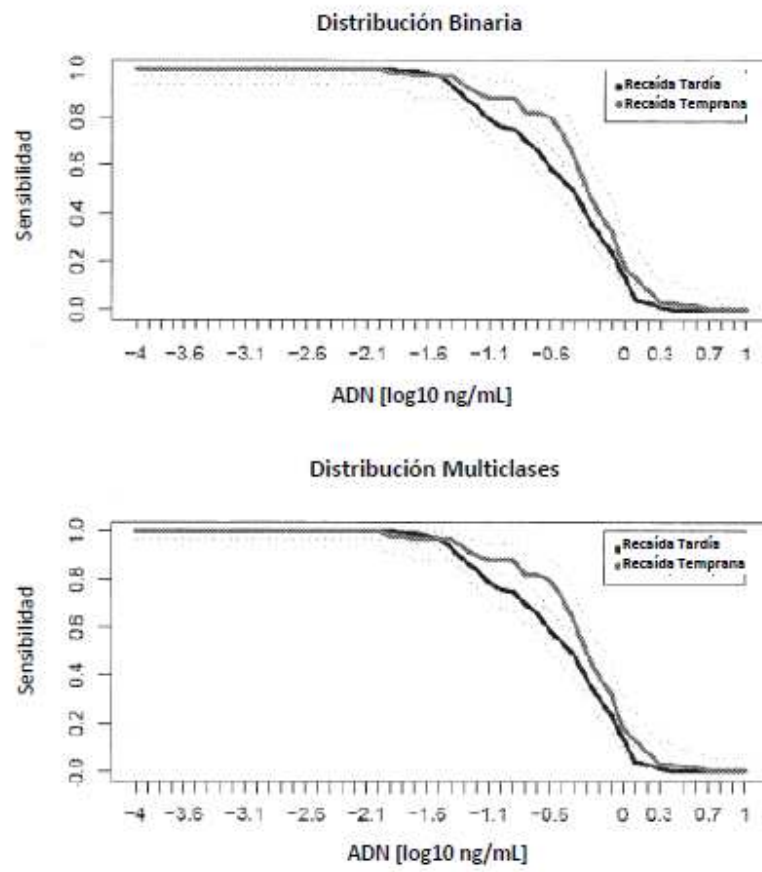
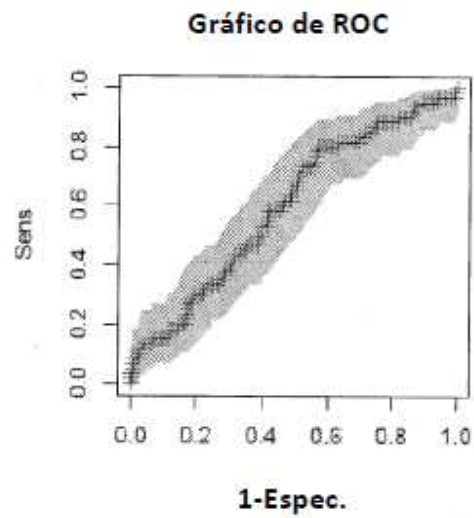


Figura 113/1



AUC	0.6 [0.53,0.67]
Sens/Espec	0.2 / 0.85
P de Wilcoxon	0.0272
Recaída Temprana (pos)	60
Recaída Tardía (neg)	134

Figura 113/2

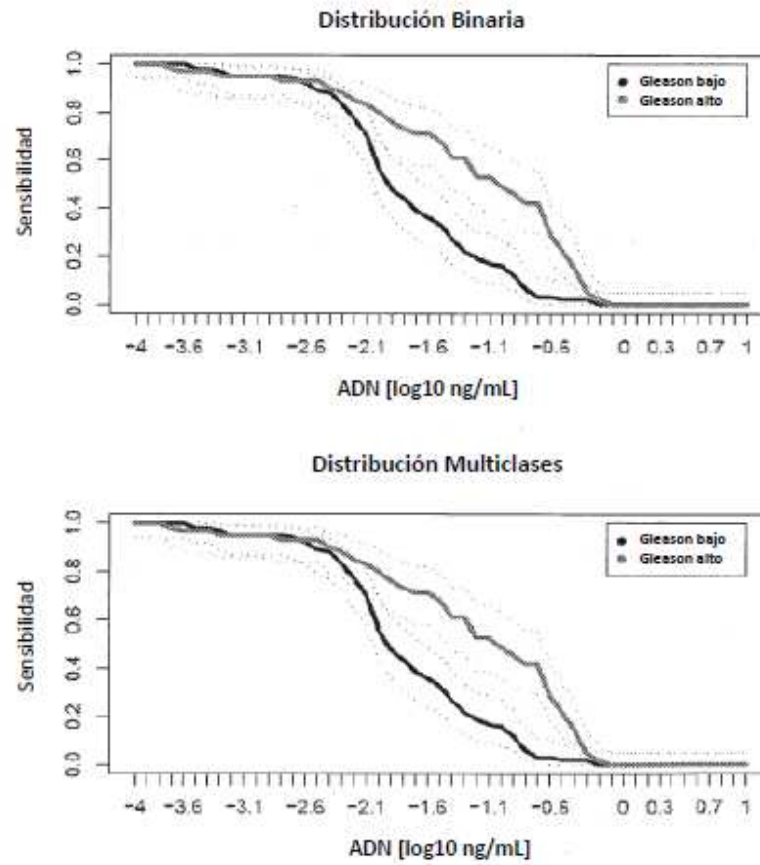
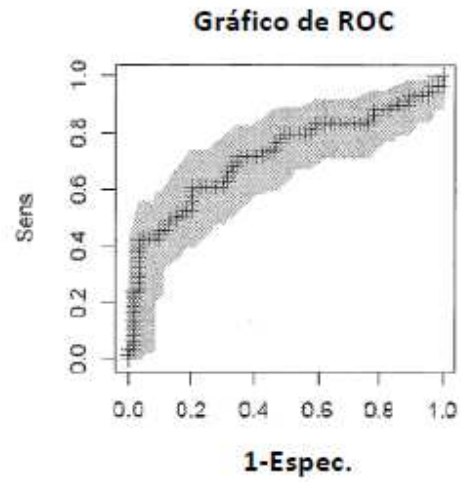


Figura 114/1



AUC	0.72 [0.63,0.79]
Sens/Espec	0.49 / 0.86
P de Wilcoxon	0
Gleason alto (pos)	59
Gleason bajo (neg)	64

Figura 114/2

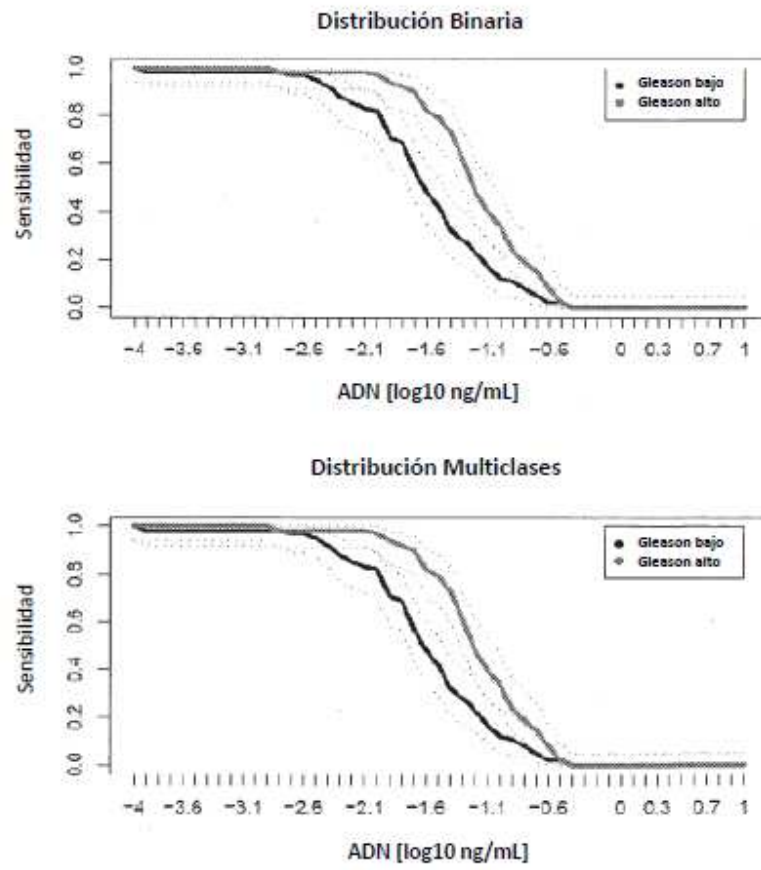
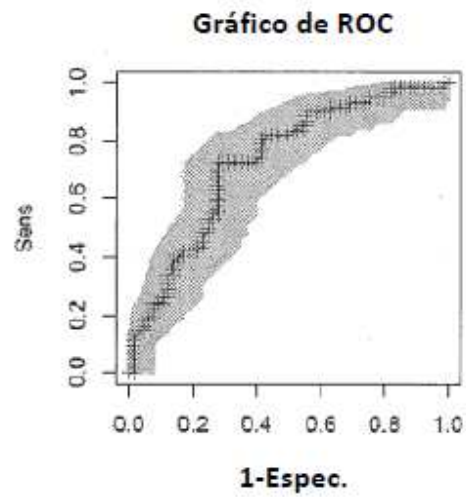


Figura 115/1



AUC	0.73 [0.64,0.8]
Sens/Espec	0.39 / 0.86
P de Wilcoxon	0
Gleason alto (pos)	62
Gleason bajo (neg)	65

Figura 115/2

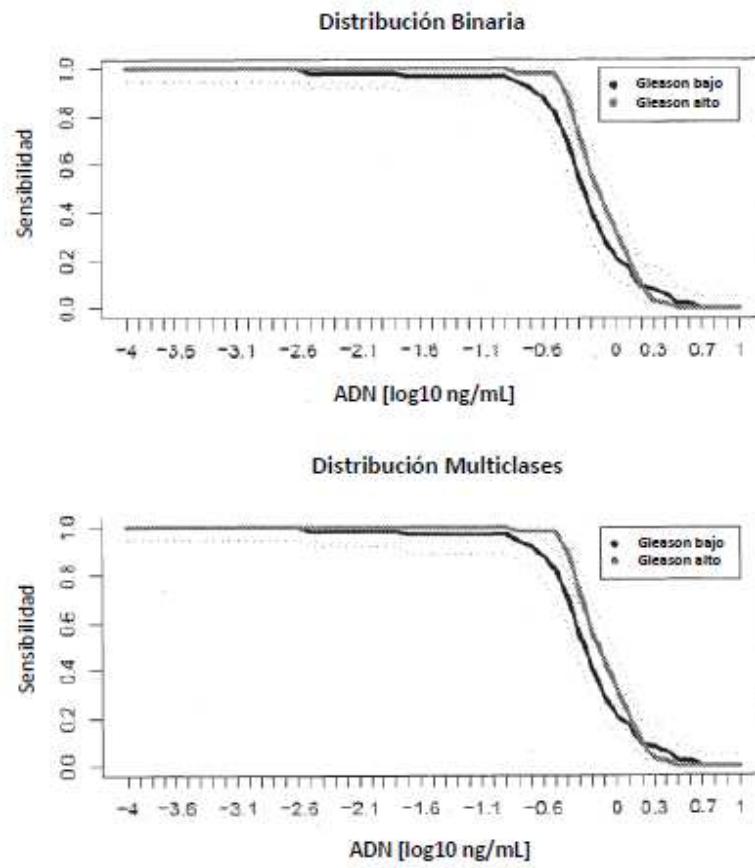
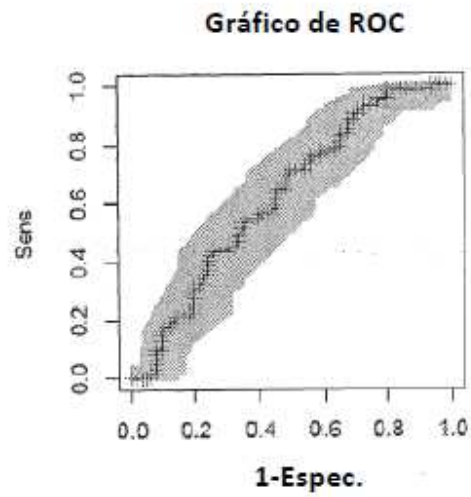


Figura 116/1



AUC	0.62 [0.53,0.7]
Sens/Espec	0.21 / 0.86
P de Wilcoxon	0.0191
Gleason alto (pos)	62
Gleason bajo (neg)	66

Figura 116/2

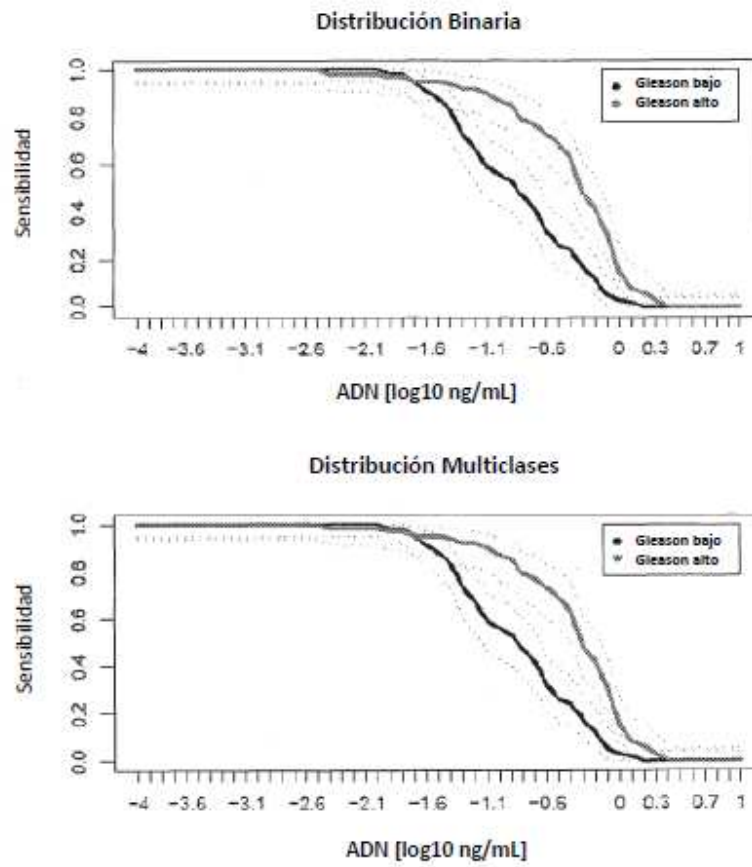
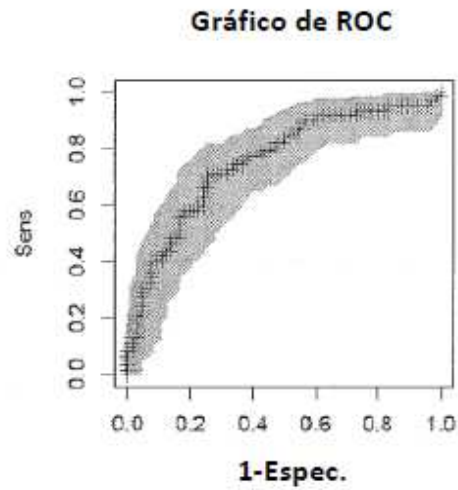


Figura 117/1



AUC	0.76 [0.67,0.83]
Sens/Espec	0.48 / 0.86
P de Wilcoxon	0
Gleason alto (pos)	62
Gleason bajo (neg)	66

Figura 117/2

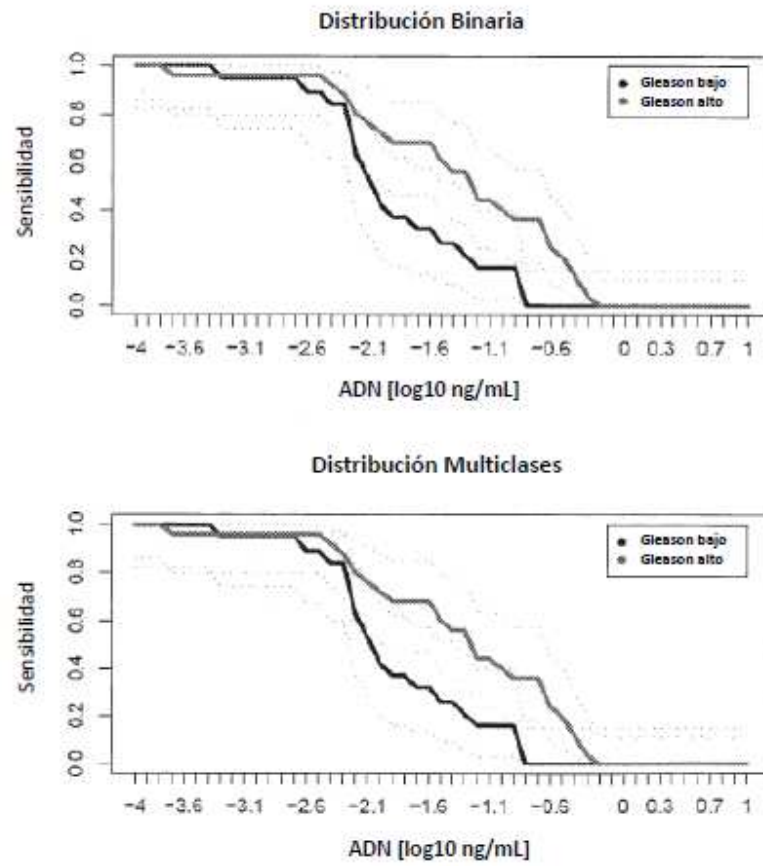
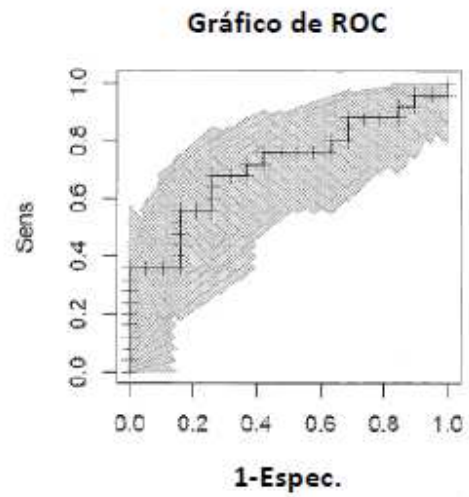


Figura 118/1



AUC	0.72 [0.55,0.83]
Sens/Espec	0.36 / 0.89
P de Wilcoxon	0.0146
Gleason alto (pos)	25
Gleason bajo (neg)	19

Figura 118/2

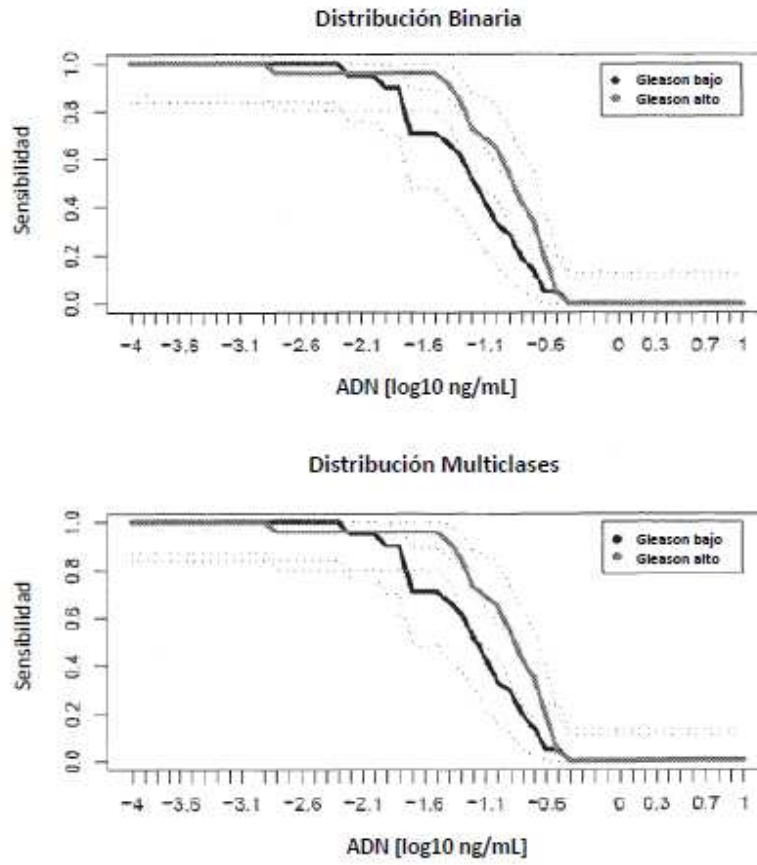
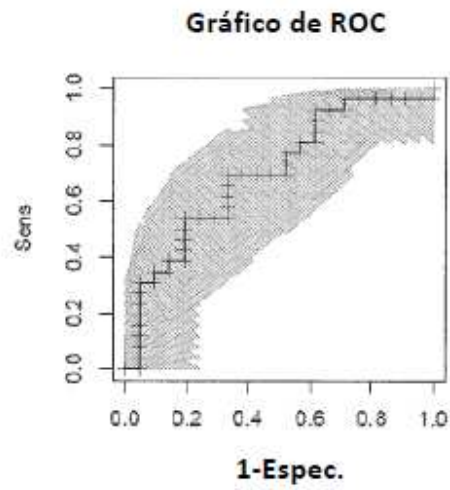


Figura 119/1



AUC	0.7 [0.55,0.83]
Sens/Espec	0.38 / 0.86
P de Wilcoxon	0.0217
Gleason alto (pos)	26
Gleason bajo (neg)	21

Figura 119/2

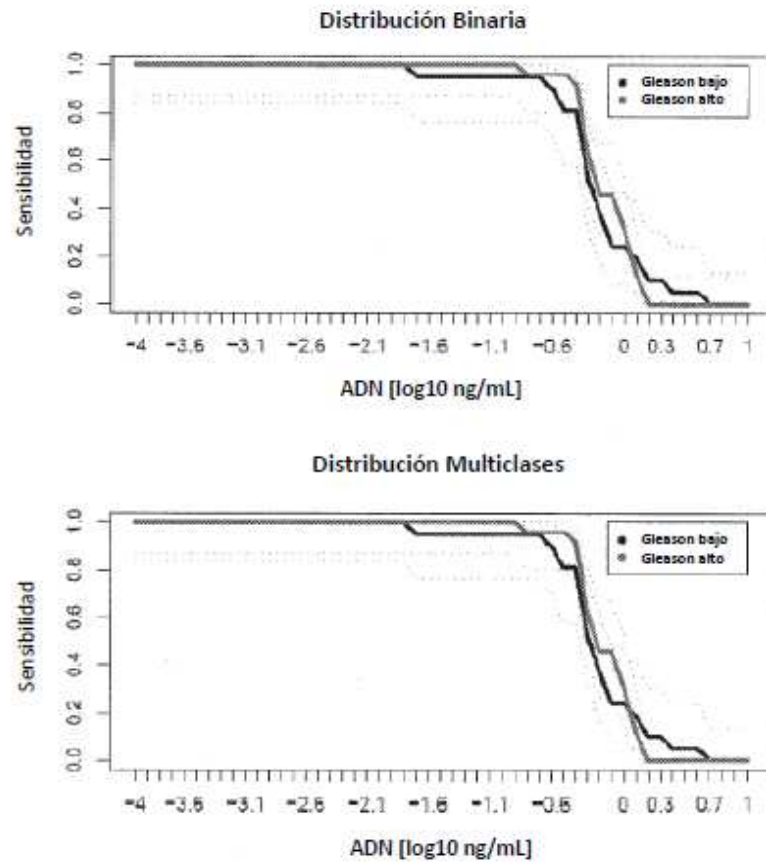
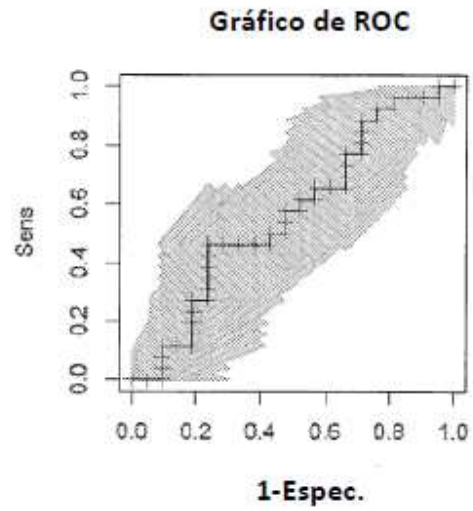


Figura 120/1



AUC	0.56 [0.4,0.7]
Sens/Espec	0.12 / 0.86
P de Wilcoxon	0.4771
Gleason alto (pos)	26
Gleason bajo (neg)	21

Figura 120/2

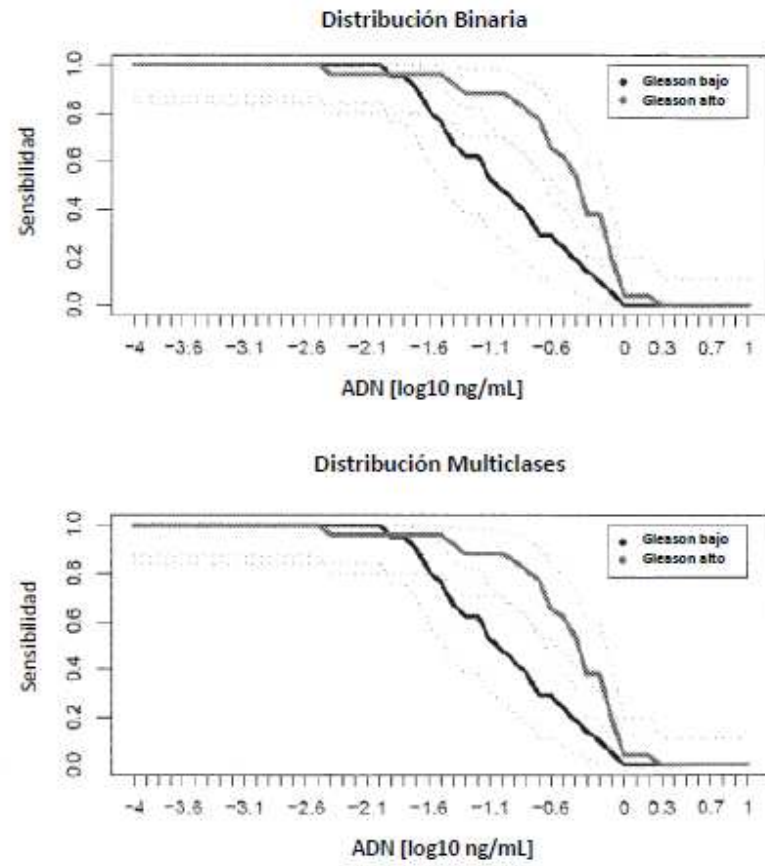
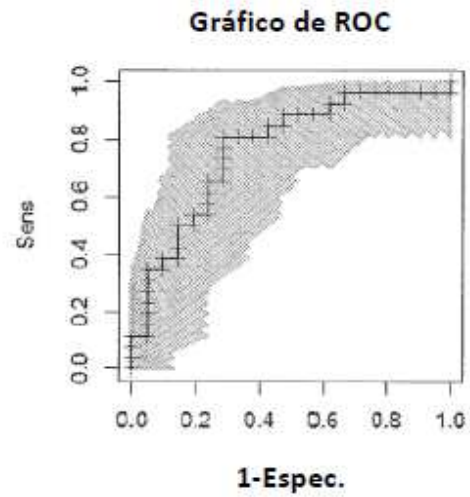


Figura 121/1



AUC	0.77 [0.62,0.88]
Sens/Espec	0.5 / 0.86
P de Wilcoxon	0.0014
Gleason alto (pos)	26
Gleason bajo (neg)	21

Figura 121/2

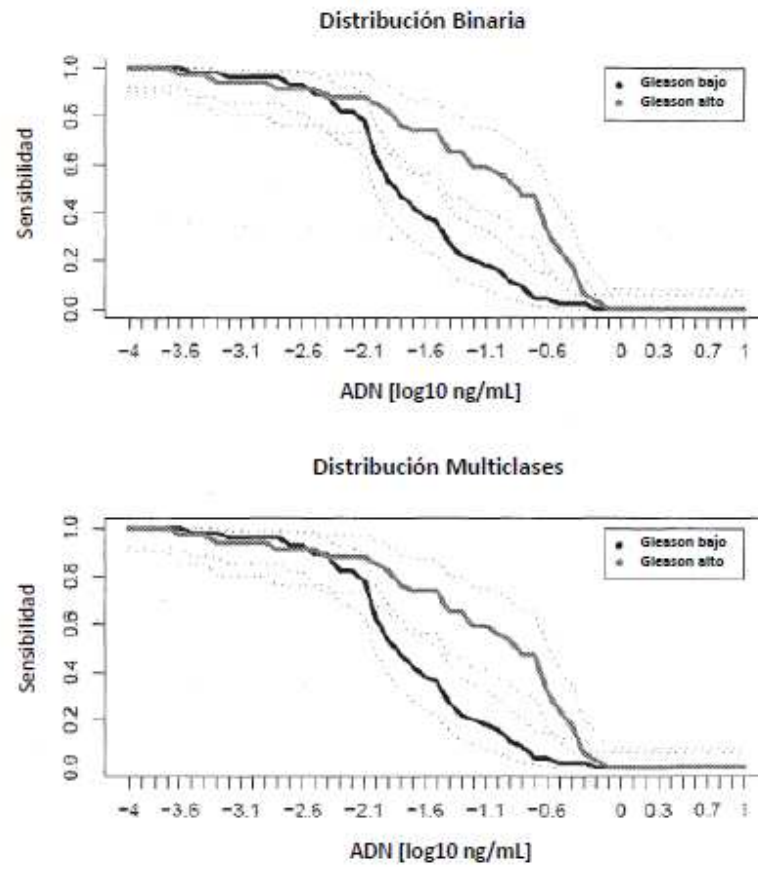
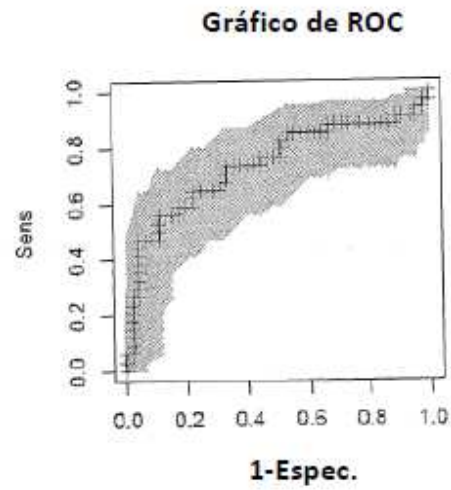


Figura 122/1



AUC	0.74 [0.62,0.83]
Sens/Espec	0.56 / 0.87
P de Wilcoxon	2e-04
Gleason alto (pos)	34
Gleason bajo (neg)	45

Figura 122/2

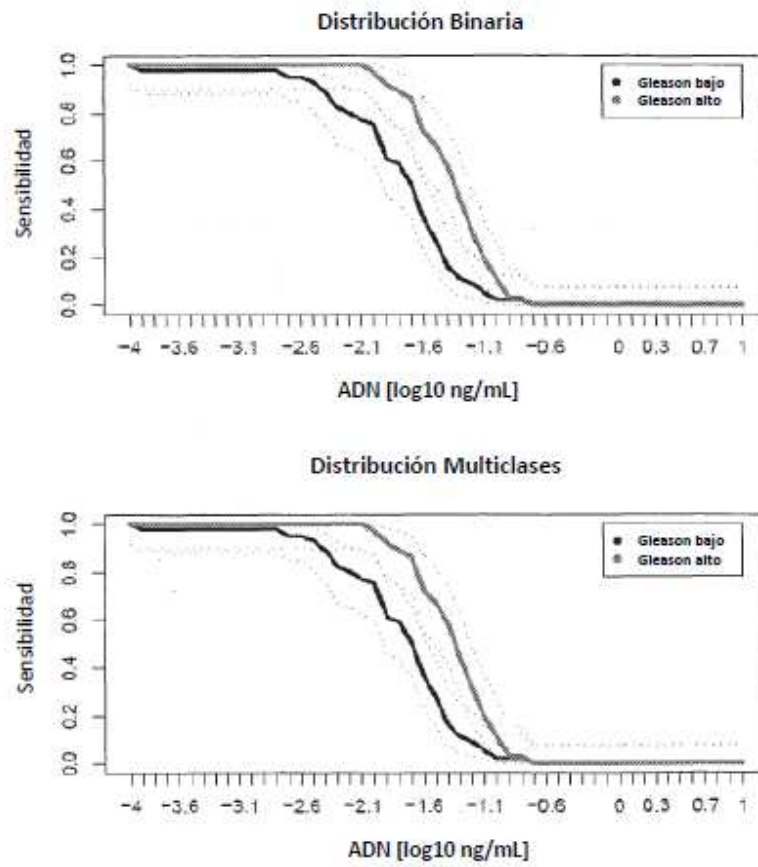
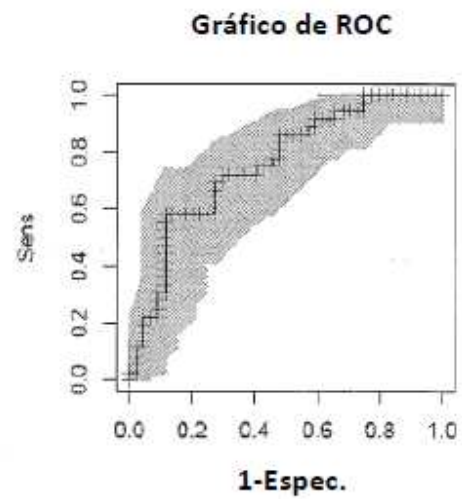


Figura 123/1



AUC	0.76 [0.65,0.85]
Sens/Espec	0.58 / 0.86
P de Wilcoxon	0
Gleason alto (pos)	36
Gleason bajo (neg)	44

Figura 123/2

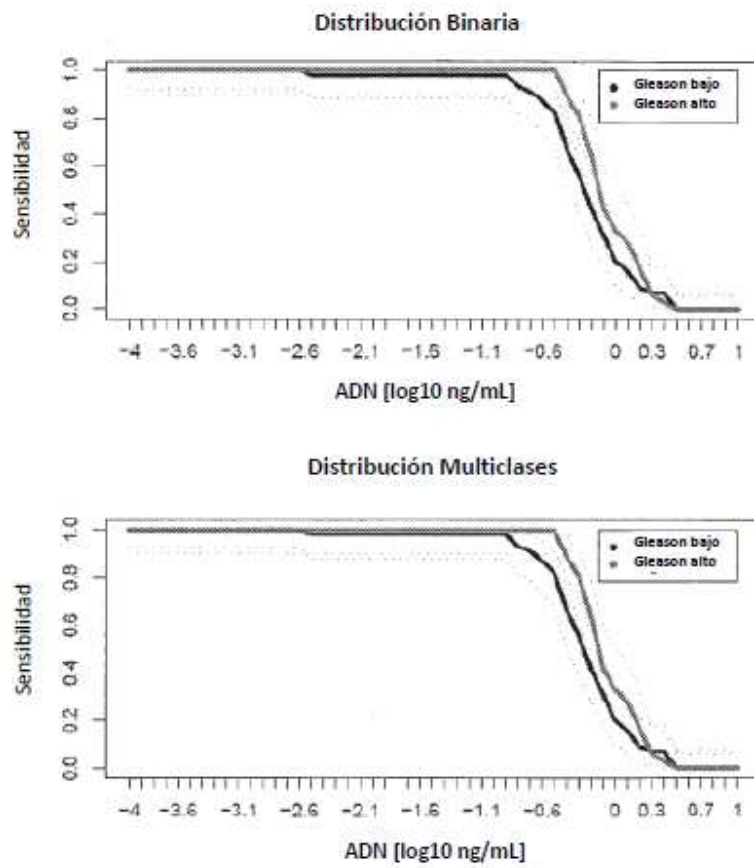
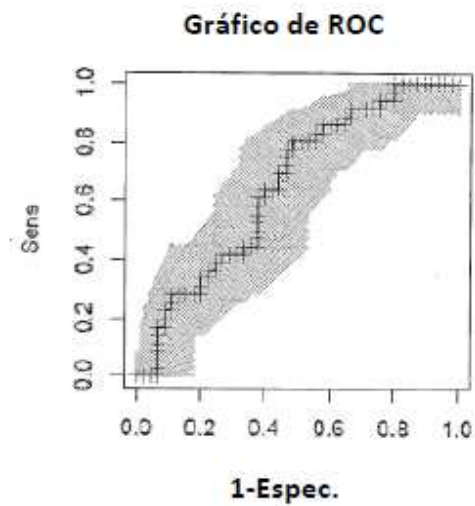


Figura 124/1



AUC	0.65 [0.54,0.76]
Sens/Espec	0.28 / 0.87
P de Wilcoxon	0.0186
Gleason alto (pos)	36
Gleason bajo (neg)	45

Figura 124/2

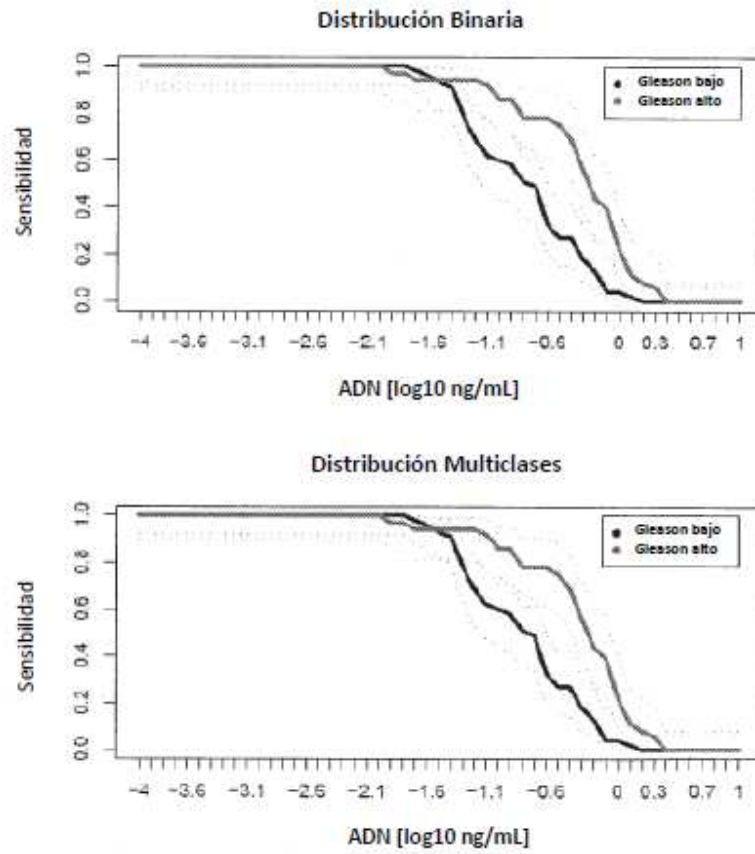
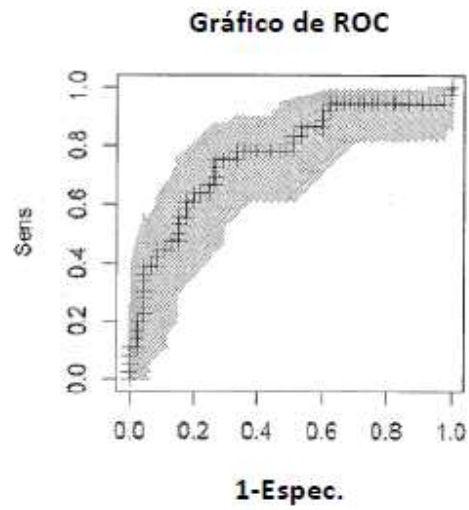


Figura 125/1



AUC	0.77 [0.66,0.85]
Sens/Espec	0.47 / 0.87
P de Wilcoxon	0
Gleason alto (pos)	36
Gleason bajo (neg)	45

Figura 125/2