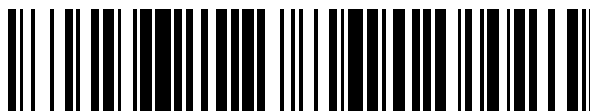


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 375 351**

51 Int. Cl.:
A61K 39/395 (2006.01)
A61K 38/00 (2006.01)
A61K 48/00 (2006.01)
C07K 16/00 (2006.01)
C12N 5/12 (2006.01)
C12P 21/00 (2006.01)
C07H 21/04 (2006.01)
C12Q 1/68 (2006.01)
G01N 33/543 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **02728645 .9**
96 Fecha de presentación: **01.04.2002**
97 Número de publicación de la solicitud: **1553980**
97 Fecha de publicación de la solicitud: **20.07.2005**

54 Título: **ÁCIDO NUCLEICO Y PROTEÍNA CORRESPONDIENTE DENOMINADA 238P1B2 ÚTIL EN EL TRATAMIENTO Y LA DETECCIÓN DE CÁNCER.**

45 Fecha de publicación de la mención BOPI:
29.02.2012

45 Fecha de la publicación del folleto de la patente:
29.02.2012

73 Titular/es:
AGENSYS, INC.
2225 COLORADO AVENUE
SANTA MÓNICA CA 90404, US

72 Inventor/es:
Raitano, Arthur B.;
Challita-Eid, Pia M.;
Faris, Mary;
Hubert, Rene S.;
Morrison, Robert Kendall;
Ge, Wangmao y
Jakobovits, Aya

74 Agente: **Carpintero López, Mario**

ES 2 375 351 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

Ácido nucleico y proteína correspondiente denominada 238P1B2 útil en el tratamiento y la detección de cáncer

Campo de la invención

5 La invención descrita en el presente documento se refiere a un gen y su proteína codificada, denominada 238P1B2, expresada en ciertos cánceres, y a procedimientos de diagnóstico y terapéuticos y composiciones útiles en el tratamiento de cánceres que expresan 238P1B2.

Antecedentes de la invención

10 El cáncer es la segunda causa principal de muerte de seres humanos junto a enfermedad coronaria. A nivel mundial, millones de personas mueren de cáncer cada año. Solamente en los Estados Unidos, como se indica por la Sociedad Americana del Cáncer, el cáncer provoca la muerte de más de medio millón de personas anualmente, con más de 1,2 millones de casos nuevos diagnosticados por año. Aunque las muertes de enfermedad coronaria han estado reduciéndose significativamente, las resultantes de cáncer generalmente están aumentando. Se predice que en la primera parte del próximo siglo el cáncer se convertirá en la causa principal de muerte.

15 A nivel mundial, varios cánceres destacan como los más letales. En particular, los carcinomas de pulmón, próstata, mama, colon, páncreas y ovario representan las causas principales de muerte por cáncer. Estos y virtualmente todos los demás carcinomas comparten una característica letal común. Con muy pocas excepciones, la enfermedad metastásica de un carcinoma es letal. Además, incluso para los pacientes de cáncer que inicialmente sobreviven sus cánceres primarios, la experiencia habitual ha mostrado que sus vidas se alteran de forma dramática. Muchos pacientes de cáncer experimentan fuertes ansiedades provocadas por la consciencia del potencial de reaparición o fracaso del tratamiento. Muchos pacientes de cáncer experimentan debilitamientos físicos después del tratamiento. Además, muchos pacientes de cáncer experimentan una recidiva.

20 A nivel mundial, el cáncer de próstata es el cuarto cáncer más prevalente en hombres. En Norte América y el Norte de Europa, es con diferencia el cáncer más habitual en hombres y es la segunda causa de muerte por cáncer en hombres. Solamente en los Estados Unidos, más de 30.000 hombres mueren anualmente de esta enfermedad, superado solamente por el cáncer de pulmón. A pesar de la magnitud de estos números, aún no existe tratamiento eficaz para cáncer de próstata metastásico. La prostatectomía quirúrgica, radioterapia, terapia de ablación hormonal, castración quirúrgica y quimioterapia continúan siendo las principales modalidades de tratamiento. Desafortunadamente, estos tratamientos son ineficaces para muchos y con frecuencia están asociados con consecuencias no deseables.

25 Con respecto al diagnóstico, la falta de un marcador de tumor de próstata que pueda detectar de forma precisa tumores de etapa temprana localizados, sigue siendo una limitación significativa en el diagnóstico y tratamiento de esta enfermedad. Aunque el ensayo de antígeno específico de próstata en suero (PSA) ha sido una herramienta muy útil, sin embargo su especificidad y utilidad general con frecuencia se considera insuficiente en varios aspectos importantes.

30 El progreso en la identificación de marcadores específicos adicionales para cáncer de próstata ha mejorado por la generación de xenotrasplantes de cáncer de próstata que pueden recapitular diferentes etapas de la enfermedad en ratones. Los xenotrasplantes de LAPC (Cáncer de Próstata Los Ángeles) son xenotrasplantes de cáncer de próstata que han sobrevivido el pase en ratones inmunodeficientes graves combinados (SCID) y han mostrado la capacidad de imitar la transición de dependencia de andrógenos a independencia de andrógenos (Klein y col., 1997, Nat. Med. 3:402). Marcadores de cáncer de próstata más recientemente identificados incluyen PCTA-1 (Su y col., 1996, Proc. Natl. Acad. Sci. USA 93: 7252), antígeno de membrana específica de próstata (PSM) (Pinto y col., Clin Cancer Res 1996 Sep 2 (9): 1445-51), STEAP (Hubert, y col., Proc Natl Acad Sci USA. 1999 Dec 7; 96(25): 14523-8) y antígeno de célula madre de próstata (PSCA) (Reiter y col., 1998, Proc. Natl. Acad. Sci. USA 95: 1735).

35 Aunque marcadores previamente identificados tales como PSA, PSM, PCTA y PSCA han facilitado los esfuerzos para diagnosticar y tratar cáncer de próstata, existe la necesidad de identificación de marcadores adicionales y dianas terapéuticas para cánceres de próstata y relacionados para mejorar adicionalmente el diagnóstico y la terapia.

40 El carcinoma de células renales (RCC) supone aproximadamente 3 % de los tumores malignos de adultos. Una vez que los adenomas alcanzan un diámetro de 2 a 3 cm, existe potencial tumoral. En el adulto, los dos tumores renales malignos principales son adenocarcinoma de células renales y carcinoma de células transicionales del uréter o pelvis renal. La incidencia de adenocarcinoma de células renales se estima en más de 29.000 casos en los Estados Unidos y más de 11.600 pacientes murieron de esta enfermedad en 1998. El carcinoma de células transicionales es menos frecuente, con una incidencia de aproximadamente 500 casos por año en los Estados Unidos.

45 La cirugía ha sido la terapia principal para adenocarcinoma de células renales durante muchas décadas. Hasta hace poco, la enfermedad metastásica ha sido refractaria para cualquier terapia sistémica. Con los desarrollos recientes en las terapias sistémicas, particularmente inmunoterapias, el carcinoma de células renales metastásicas puede

enfocarse de forma agresiva en pacientes apropiados con posibilidad de respuestas duraderas. Sin embargo, sigue existiendo una necesidad de terapias eficaces para estos pacientes.

5 De todos los nuevos casos de cáncer en los Estados Unidos, el cáncer de vejiga representa aproximadamente 5 % en hombres (quinto neoplasma más habitual) y 3 % en mujeres (octavo neoplasma más habitual). La incidencia está aumentando lentamente, simultáneamente con una población crecientemente más vieja. En 1998, hubo una estimación de 54.500 casos, incluyendo 39.500 en hombres y 15.000 en mujeres. La incidencia ajustada por edad en los Estados Unidos es 32 por 100.000 para hombres y 8 por 100.000 en mujeres. La relación hombre/mujer histórica de 3:1 se puede estar reduciendo en relación con los patrones de tabaquismo en mujeres. Hubo una estimación de 11.000 muertes de cáncer de vejiga en 1998 (7.800 en hombres y 3.900 en mujeres). La incidencia del cáncer de vejiga y la mortalidad aumentan fuertemente con la edad y serán un problema creciente a medida que la población se hace más vieja.

15 La mayoría de los cánceres de vejiga recurren en la vejiga. El cáncer de vejiga se trata con una combinación de resección transuretral de la vejiga (TUR) y quimioterapia intravesical o inmunoterapia. La naturaleza multifocal y recurrente de cáncer de vejiga apunta a las limitaciones de TUR. La mayoría de los cánceres invasivos de músculo no se curan solamente por TUR. La cistectomía radical y la desviación urinaria es el medio más eficaz para eliminar el cáncer pero conlleva un impacto innegable en la función urinaria y sexual. Continúa habiendo una necesidad significativa de modalidades de tratamiento que sean beneficiosas para pacientes con cáncer de vejiga.

20 Se estima que se produjeron 130.200 casos de cáncer colorrectal en 2000 en los Estados Unidos, incluyendo 93.800 casos de cáncer de colon y 36.400 de cáncer rectal. Los cánceres colorrectales son los terceros cánceres más habituales en hombres y mujeres. Las tasas de incidencia se redujeron significativamente durante 1992-1996 (-2,1 % por año). La investigación sugiere que estas reducciones se han debido a exploración aumentada y retirada de pólipos, evitando la progresión de pólipos a cánceres invasivos. Hubo una estimación de 56.300 muertes (47.700 de cáncer de colon, 8.600 de cáncer rectal) en 2000, que suponen aproximadamente 11 % de todas las muertes por cáncer en Estados Unidos.

25 En la actualidad, la cirugía es la forma más habitual de terapia para cáncer colorrectal y, para cánceres que no se han propagado, es frecuentemente curativa. La quimioterapia, o quimioterapia más radiación, se proporciona antes o después de cirugía a la mayoría de los pacientes cuyo cáncer ha perforado en profundidad la pared del intestino o se ha propagado a los ganglios linfáticos. En ocasiones se necesita una colostomía permanente (creación de una apertura abdominal para eliminación de residuos corporales) para cáncer de colon e infrecuentemente se requiere para cáncer rectal. Continúa habiendo una necesidad de modalidades de diagnóstico y tratamiento eficaces para cáncer colorrectal.

30 Se estima que se produjeron 164.100 nuevos casos de cáncer de pulmón y bronquial en 2000, que suponen el 14 % de todos los diagnósticos de cáncer de Estados Unidos. La tasa de incidencia de cáncer de pulmón y bronquial está reduciéndose significativamente en hombres, de un máximo de 86,5 por 100.000 en 1984 a 70,0 en 1996. En los años 90, la tasa de aumento entre mujeres comenzó a ralentizarse. En 1996, la tasa de incidencia en mujeres fue de 42,3 por 100.000.

35 El cáncer de pulmón y bronquial provocó 156.900 muertes estimadas en 2000, lo que suponen el 28 % de todas las muertes por cáncer. Durante 1992-1996, la mortalidad de cáncer de pulmón se redujo significativamente entre hombres (-1,7 % por año) mientras que las tasas para mujeres aún aumentaban significativamente (0,9 % por año). Desde 1987, han muerto más mujeres cada año de cáncer de pulmón que de cáncer de mama, que, durante más de 40 años, fue la causa principal de muerte por cáncer en mujeres. La reducción de incidencia de cáncer de pulmón y tasas de mortalidad resultó con mayor probabilidad de la reducción de las tasas de tabaquismo durante los 30 años anteriores; sin embargo, los patrones de tabaquismo en reducción entre mujeres están por detrás de los de los hombres. Es preocupante que aunque las reducciones del uso de tabaco en adultos se han ralentizado, el uso de tabaco en jóvenes está aumentando de nuevo.

40 Las opciones de tratamiento para cáncer de pulmón y bronquial se determinan por el tipo y etapa de cáncer e incluyen cirugía, radioterapia y quimioterapia. Para muchos cánceres localizados, la cirugía es habitualmente el tratamiento de elección. Debido a que la enfermedad habitualmente se ha propagado en el momento en que se descubre, con frecuencia se requieren radioterapia y quimioterapia en combinación con cirugía. La quimioterapia sola o combinada con radiación es el tratamiento de elección para cáncer de pulmón de células pequeñas; en este régimen, un gran porcentaje de pacientes experimentan remisión, que en algunos casos es de larga duración. Existe sin embargo, una necesidad continua de enfoques de tratamiento y diagnóstico eficaces para cánceres de pulmón y bronquiales.

45 Se estima que se producirán 182.800 nuevos casos invasivos de cáncer de mama entre mujeres en los Estados Unidos durante 2000. Adicionalmente, se esperaba que se diagnosticaran aproximadamente 1.400 nuevos casos de cáncer de mama en hombres en 2000. Después de aumentar aproximadamente 4 % por año en los años 80, las tasas de incidencia de cáncer de mama en mujeres se han nivelado en los años 90 a aproximadamente 110,6 casos por 100.000.

Solamente en los Estados Unidos, se estima que se produjeron 41.200 muertes (40.800 mujeres, 400 hombres) en 2000 debido a cáncer de mama. El cáncer de mama se clasifica en segundo lugar entre las muertes por cáncer en mujeres. De acuerdo con los datos más recientes, las tasas de mortalidad se redujeron significativamente durante 1992-1996 con las reducciones mayores en mujeres jóvenes, tanto blancas como negras. Estas reducciones fueron probablemente el resultado de detección más temprana y tratamiento mejorado.

Teniendo en cuenta las circunstancias médicas y las preferencias del paciente, el tratamiento de cáncer de mama puede implicar lumpectomía (retirada local del tumor) y retirada de los ganglios linfáticos bajo el brazo; mastectomía (retirada quirúrgica de la mama) y retirada de los ganglios linfáticos bajo el brazo; radioterapia; quimioterapia; o terapia hormonal. Con frecuencia, se usan dos o más procedimientos en combinación. Numerosos estudios han mostrado que, para enfermedad de etapa temprana, las tasas de supervivencia a largo plazo después de lumpectomía más radioterapia son similares a las tasas de supervivencia después de mastectomía radical modificada. Avances significativos en las técnicas de reconstrucción proporcionan varias opciones para reconstrucción del pecho después de mastectomía. Recientemente, dicha reconstrucción se ha realizado a la vez que la mastectomía.

La escisión local de carcinoma ductal *in situ* (DCIS) con cantidades adecuadas de tejido mamario normal circundante pueden evitar la reaparición local del DCIS. La radiación en la mama y/o tamoxifeno pueden reducir las probabilidades de que aparezca DCIS en el tejido mamario restante. Esto es importante porque DCIS, si se deja sin tratar, puede desarrollarse a cáncer de mama invasivo. Sin embargo, existen efectos secundarios o secuelas graves para estos tratamientos. Existe por lo tanto, la necesidad de tratamientos de cáncer de mama eficaces.

Se estima que hubo 23.000 nuevos casos de cáncer de ovario en los Estados Unidos en 2000. Supone el 4 % de todos los cánceres entre mujeres y se clasifica el segundo entre cánceres ginecológicos. Durante 1992-1996, las tasas de incidencia de cáncer de ovario se redujeron significativamente. Como consecuencia de cáncer de ovario, hubo una estimación de 14.000 muertes en 2000. El cáncer de ovario provoca más muertes que cualquier otro cáncer del sistema reproductor femenino.

La cirugía, la radioterapia y la quimioterapia son opciones de tratamiento para cáncer de ovario. La cirugía habitualmente incluye la retirada de uno o ambos ovarios, las trompas de Falopio (salpingo-ooforectomía) y el útero (histerectomía). En algunos tumores muy tempranos, sólo se retirará el ovario implicado, especialmente en mujeres jóvenes que desean tener hijos. En enfermedad avanzada, se realiza un intento de retirar toda la enfermedad intra-abdominal para potenciar el efecto de la quimioterapia. Continúa existiendo una necesidad importante de opciones de tratamiento eficaces para cáncer de ovario.

Se estima que hubo 28.300 nuevos casos de cáncer pancreático en los Estados Unidos en 2000. Durante los últimos 20 años, las tasas de cáncer pancreático se han reducido en hombres. Las tasas entre mujeres han permanecido aproximadamente constantes pero pueden estar empezando a reducirse. El cáncer pancreático provocó una estimación de 28.200 muertes en 2000 en los Estados Unidos. Durante los últimos 20 años, hubo una ligera pero significativa reducción en las tasas de mortalidad entre hombres (aproximadamente -0,9 % por año) mientras que las tasas han aumentado ligeramente entre mujeres.

La cirugía, la radioterapia y la quimioterapia son opciones de tratamiento para cáncer pancreático. Estas opciones de tratamiento pueden prolongar la supervivencia y/o aliviar síntomas en muchos pacientes pero no es probable que produzcan una cura para la mayoría. Existe una necesidad significativa de opciones de diagnóstico y terapéuticas adicionales para cáncer pancreático.

El documento WO01/27158 desvela varios polipéptidos receptores olfatorios humanos y polinucleótidos que los codifican. Los documentos WO02/16548 y EP-A-1270724, publicados el 2 de enero de 2003, desvelan proteínas receptoras acopladas a proteína G y polinucleótidos que las codifican. No se establece ninguna relación entre ninguna de estas proteínas y su expresión en tejido tumoral.

Sumario de la invención

La presente invención se refiere a un gen, designado 238P1B2, que se ha descubierto ahora que se sobre-expresa en el cáncer o cánceres enumerados en la Tabla I. El análisis de expresión por transferencia de Northern de la expresión del gen 238P1B2 en tejidos normales muestra un patrón de expresión restringido en tejidos adultos. Se proporcionan las secuencias de nucleótidos (Figura 2) y aminoácidos (Figura 2 y Figura 3) de 238P1B2. El perfil relacionado con tejido de 238P1B2 en tejidos adultos normales, en combinación con la sobre-expresión observada en los tumores enumerados en la Tabla I, muestra que 238P1B2 se sobre-expresa de forma aberrante en al menos algunos cánceres y por lo tanto actúa como una diana de diagnóstico, profiláctica, de pronóstico y/o terapéutica útil para cánceres del tejido o tejidos tales como los enumerados en la Tabla I.

De acuerdo con la presente invención, se proporciona un procedimiento para detectar la presencia de cáncer de próstata en una muestra de ensayo, comprendiendo el procedimiento poner en contacto la muestra con una sonda que se une específicamente a un polinucleótido de la Figura 2B (SEC IN N° 8872); y detectar la unión del polinucleótido en la muestra a la misma.

Los inventores describen polinucleótidos correspondientes o complementarios a todos o parte de los genes de 238P1B2, ARNm y/o secuencias codificantes, preferentemente en forma de aislada, incluyendo polinucleótidos que codifican proteínas relacionadas con 238P1B2 y fragmentos de 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más de 25 aminoácidos contiguos; al menos 30, 35, 40, 45, 50, 55, 60, 65, 70, 80, 85, 90, 95, 100 o más de 100 aminoácidos contiguos de una proteína relacionada con 238P1B2, así como los péptidos/proteínas en sí mismos; ADN, ARN, híbridos de ADN/ARN y moléculas relacionadas, polinucleótidos u oligonucleótidos complementarios o que tienen al menos un 90 % de homología con los genes 238P1B2 o secuencias de ARNm o partes de las mismas y polinucleótidos u oligonucleótidos que hibridan con los genes de 238P1B2, ARNm o con polinucleótidos que codifican 238P1B2. Los inventores también describen medios para aislar ADNc y los genes que codifican 238P1B2. También se describen moléculas de ADN recombinante que contienen polinucleótidos de 238P1B2, células transformadas o transducidas con tales moléculas y sistemas de huésped-vector para la expresión de productos génicos de 238P1B2. Puede existir una condición de que la secuencia de ácido nucleico completa de la Figura 2 no esté codificada y/o la secuencia de aminoácidos completa de la Figura 2 no esté preparada. La secuencia de ácido nucleico completa de la Figura 2 puede prepararse y/o la secuencia de aminoácidos completa de la Figura 2 puede prepararse, estando cualquiera de las dos en formas de dosis unitaria humana respectivas.

La invención proporciona procedimientos para detectar la presencia y el estado de polinucleótidos de 238P1B2 en diversas muestras biológicas, así como procedimientos para identificar células que expresan 238P1B2. Una realización típica de la presente invención proporciona procedimientos para controlar los productos génicos de 238P1B2 en una muestra tisular o hematológica que tiene o se sospecha que tiene alguna forma de desregulación del crecimiento tal como cáncer.

Breve descripción de las figuras

Figura 1. La secuencia SSH de 238P1B2 de 210 nucleótidos

Figura 2. La secuencia de ADNc y aminoácidos de la variante 1A de 238P1B2 se muestra en la Figura 2A. La metionina de partida está subrayada. El marco de lectura abierto se extiende desde los ácidos nucleicos 2133 a 2897 incluyendo el codón de parada. La secuencia de ácido nucleico y aminoácidos de 238P1B2 variante 1B se muestra en la Figura 2B, el codón para la metionina de inicio está subrayado. El marco de lectura abierto se extiende desde el ácido nucleico 1947 a 2897 incluyendo el codón de parada. La secuencia de ácido nucleico y aminoácidos de 238P1B2 variante 2 se muestra en la Figura 2C, el codón para la metionina de inicio está subrayado. El marco de lectura abierto se extiende desde el ácido nucleico 2133 a 2897 incluyendo el codón de parada. La secuencia de ácido nucleico y aminoácidos de 238P1B2 variante 3 se muestra en la Figura 2D, el codón para la metionina de inicio está subrayado. El marco de lectura abierto se extiende desde el ácido nucleico 2133 a 2897 incluyendo el codón de parada. La secuencia de ácido nucleico y aminoácidos de 238P1B2 variante 4 se muestra en la Figura 2E, el codón para la metionina de inicio está subrayado. El marco de lectura abierto se extiende desde el ácido nucleico 2133 a 2897 incluyendo el codón de parada. La secuencia de ácido nucleico y aminoácidos de 238P1B2 variante 5 se muestra en la Figura 2F, el codón para la metionina de inicio está subrayado. El marco de lectura abierto se extiende desde el ácido nucleico 2133 a 2897 incluyendo el codón de parada. La secuencia de ácido nucleico y aminoácidos de 238P1B2 variante 6 se muestra en la Figura 2G, el codón para la metionina de inicio está subrayado. El marco de lectura abierto se extiende desde el ácido nucleico 2133 a 2897 incluyendo el codón de parada.

Figura 3. La secuencia de aminoácidos de 238P1B2 variante 1A se muestra en la Figura 3A; tiene 254 aminoácidos. La secuencia de aminoácidos de 238P1B2 variante 1B se muestra en la Figura 3B; tiene 316 aminoácidos. La secuencia de aminoácidos de 238P1B2 variante 2 se muestra en la Figura 3C; tiene 254 aminoácidos.

Figura 4. A. Alineamiento de secuencia de ácido nucleico de 238P1B2 variante 1 con receptor olfatorio de ratón MOR14-1. B. Alineamiento de secuencia de ácido nucleico de 238P1B2 variante 1 con el receptor olfatorio de ratón MOR14-10. C. Alineamiento de secuencia de aminoácidos de 238P1B2 v.1A con el receptor olfatorio de ratón MOR14-1. D. Alineamiento de secuencia de aminoácidos de 238P1B2 v.1A con OPCRPHOR-1 específico de próstata. E. Alineamiento de secuencia de aminoácidos de 238P1B2 variante 1 con el receptor olfatorio humano 5112. F. Alineamiento de Clustal de secuencia de aminoácidos de las tres variantes de 238P1B2, que representa que 238P1B2 V1B contiene 62 aa adicionales en su extremo N-terminal en relación con V1A y que 238P1B2 V2 porta una mutación puntual I a T en el aa 225 en relación con V1A.

Figura 5. Perfil de hidrofiliya de aminoácidos de A) 238P1B2 y B) 238P1B2 variante 1A determinado por análisis de secuencia de algoritmo informático usando el procedimiento de Hopp y Woods (Hopp T.P., Woods K.R., 1981. Proc. Natl. Acad. Sci. U.S.A. 78:3824-3828) al que se accede en el sitio web ProtScale (www.expasy.ch/cgi-bin/protscale.pl) a través del servidor de biología molecular ExPasy.

Figura 6. Perfil hidropático de aminoácidos de A) 238P1B2 y B) 238P1B2 variante 1A determinado por análisis de secuencia de algoritmo informático usando el procedimiento de Kyte y Doolittle (Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132) al que se accede en el sitio web ProtScale (www.expasy.ch/cgi-bin/protscale.pl) a través del servidor de biología molecular ExPasy.

Figura 7. Perfil aminoacídico de porcentaje de restos accesibles de A) 238P1B2 y B) 238P1B2 variante 1A determinado por análisis de secuencia de algoritmo informático usando el procedimiento de Janin (Janin J., 1979 Nature 277:491-492) al que se accede en el sitio web ProtScale (www.expasy.ch/cgi-bin/protscale.pl) a través del servidor de biología molecular ExPasy.

Figura 8. Perfil de aminoácidos de flexibilidad media de A) 238P1B2 y B) 238P1B2 variante 1A determinado por análisis de secuencia de algoritmo informático usando el procedimiento de Bhaskaran y Ponnuswamy (Bhaskaran R., y Ponnuswamy P.K., 1988. Int. J. Pept. Protein Res. 32:242-255) al que se accede en el sitio web Protscale (www.expasy.ch/cgi-bin/protscale.pl) a través del servidor de biología molecular ExPasy.

Figura 9. Perfil de aminoácidos de giro beta de A) 238P1B2 y B) 238P1B2 variante 1A determinado por análisis de secuencia de algoritmo informático usando el procedimiento de Deleage y Roux (Deleage, G., Roux B. 1987 Protein Engineering 1:289-294) al que se accede en el sitio web Protscale (www.expasy.ch/cgi-bin/protscale.pl) a través del servidor de biología molecular ExPasy.

Figura 10. Presentación esquemática de variantes de nucleótidos de 238P1B2. Las variantes 238P1B2 v.2 a 238P1B2 v.6 son variantes con variaciones de nucleótido sencillo. La caja negra muestra la misma secuencia que 238P1B2 v.1. Los números corresponden a los de 238P1B2 v.1. Los SNP se indican encima de la caja.

Figura 11. Presentación esquemática de variantes proteicas de 238P1B2. La variante de nucleótidos 238P1B2 v.1 en la Figura 10 codifica variantes de proteína 238P1B2 v.1A y v.1B. La variante de nucleótidos 238P1B2 v.2 en la Figura 10 codifica la variante proteica 238P1B2 v.2. La caja negra muestra la misma secuencia que 238P1B2 v.1. Los números en "()" debajo de la caja corresponden a los de 238P1B2 v.1A. Las diferencias de aminoácidos sencillos se indican encima de la caja.

Figura 12. A, B. Predicción de estructura secundaria para 238P1B2 variante 1a y variante 1b. Las estructuras secundarias de las proteínas 238P1B2 variante 1a (A) y de variante 1b (B) se predijeron usando el procedimiento de Red Neural Jerárquica HNN (Guermeur, 1997, en la Dirección Web pbil.ibcp.fr/cgibin/npsa_automat.pl?page=npsa_nn.html), a la que se accede desde el servidor de biología molecular ExPasy (en la Dirección Web www.expasy.ch/tools/). Este procedimiento predice la presencia y localización de hélices alfa, tramos extendidos y bucles aleatorios de la secuencia proteica primaria. El porcentaje de la proteína en una estructura secundaria dada también se enumera para cada variante.

C, D, E, F. Predicción transmembrana para 238P1B2 variante 1a y 1b. C, E. Representaciones esquemáticas de la probabilidad de existencia de regiones transmembrana y orientación de 238P1B2 variante 1a (C) y variante 1b (E) basadas en el algoritmo TMpred de Hofmann y Stoffel que usa TMBASE (K. Hofmann, W. Stoffel. TMBASE - A database of membrane spanning protein segments Biol. Chem. Hoppe-Seyler 374:166, 1993). D, F. Representación esquemática de la probabilidad de la existencia de regiones transmembrana y la orientación extracelular e intracelular de 238P1B2 variante 1a (D) y variante 1b (F) basándose en el algoritmo TMHMM de Sonnhammer, von Heijne, y Krogh (Erik L.L. Sonnhammer, Gunnar von Heijne y Anders Krogh: A hidden Markov model for predicting transmembrane helices in protein sequences. en Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology, p 175-182 Ed J. Glasgow, T. Littlejohn, F. Major, R. Lathrop, D. Sankoff, y C. Sensen Menlo Park, CA: AAAI Press, 1998). Se accede a los algoritmos TMpred y TMHMM desde el servidor de biología molecular ExPasy (en la Dirección Web www.expasy.ch/tools/). Los resultados de los programas de predicción de transmembrana presentados representan 238P1B2 variante 1a como más probablemente con 6 dominios transmembrana y la variante 1 b 7 dominios transmembrana.

Figura 13: Topología probable y estructura de 238P1B2 variantes 1 y 1B.

Figura 14. Expresión de 238P1B2 por RT-PCR. Se preparó ADNc de primera cadena de grupo vital 1 (hígado, pulmón y riñón), grupo vital 2 (páncreas, colon y estómago) y grupo de cáncer de próstata. Se realizó normalización por PCR usando cebadores para actina y GAPDH. Se realizó una PCR semi-cuantitativa, usando cebadores para 238P1B2 a 26 y 30 ciclos de amplificación. Los resultados muestran expresión fuerte de 238P1B2 en el grupo de cáncer de próstata pero no en el grupo vital 1 y grupo vital 2.

Figura 15. Expresión de 238P1B2 en tejidos normales. Se exploraron dos transferencias de northern de tejido múltiple (Clontech) ambas con 2 µg de ARNm/carril con el fragmento SSH de 238P1B2. Los patrones de tamaño en kilobases (kb) se indican en el lateral. Los resultados muestran ausencia de expresión de 238P1B2 en los 16 tejidos normales ensayados.

Figura 16. Expresión de 238P1B2 en Tejidos Normales Múltiples. Se analizó una transferencia puntual de ARNm que contenía 76 muestras diferentes de tejidos humanos usando una sonda SSH de 238P1B2. La expresión solamente se detectó en placenta.

Figura 17. Expresión de 238P1B2 en Muestras de Ensayo de Cáncer de Paciente Humano y Tejidos Normales. Se extrajo ARN de un grupo de tres muestras de ensayo de paciente con cáncer de próstata, así como de próstata normal (PN), vejiga normal (VN), riñón normal (RN), colon normal (CN), pulmón normal (PuN), mama normal (MN), y ovario normal (ON). Se exploró transferencia de Northern con 10 µg de ARN total/carril con secuencia SSH de 238P1B2. Los patrones de tamaño en kilobases (kb) están indicados en el lateral. Los resultados muestran expresión de 238P1B2 en el grupo de cáncer de próstata y ovario normal.

Figura 18. Expresión de 238P1B2 en tejidos de pacientes con cáncer de próstata. Se extrajo ARN de próstata normal (N), xenotransplantes de cáncer de próstata (LAPC-4AD, LAPC-4AI, LAPC-9AD, LAPC-9A1), líneas celulares de cáncer de próstata (LNCaP y PC3) y tumores de pacientes con cáncer de próstata (T). Se exploraron transferencias de Northern con 10 µg de ARN total con el fragmento SSH de 238P1B2. Los patrones de tamaño en kilobases están en el lateral. Los resultados muestran expresión fuerte de 238P1B2 en tejidos de tumor de próstata. El panel inferior representa la tinción con bromuro de etidio del gel.

Descripción detallada de la invención

Descripción de las secciones

I) Definiciones

II) Polinucleótidos de 238P1B2

II.A.) Usos de polinucleótidos de 238P1B2

II.A.1.) Control de anomalías genéticas

5 **II. A.2.) Antisentido**

II. A.3.) Cebadores y pares de cebadores

II. A.4.) Aislamiento de moléculas de ácido nucleico que codifican 238P1B2

II. A.5.) Moléculas de ácido nucleico recombinante y sistemas de huésped-vector

III.) Proteínas relacionadas con 238P1B2

10 **III.A.) Proteínas que portan motivo**

III.B.) Expresión de proteínas relacionadas con 238P1B2

III.C.) Modificaciones de proteínas relacionadas con 238P1B2

III.D.) Usos de proteínas relacionadas con 238P1B2

IV.) Procedimientos para la detección de 238P1B2

15 **V.) Procedimientos para controlar el estado de genes relacionados con 238PIB2 y sus productos**

VI.) Identificación de moléculas que interaccionan con 238P1B2

VII.) Procedimientos de diagnóstico y pronóstico.

VIII.) KITS

I.) Definiciones:

20 A no ser que se defina de otro modo, todos los términos de la técnica, notaciones y otros términos científicos o terminología usada en el presente documento pretenden tener los significados habitualmente entendidos por los expertos en la materia a la que pertenece la presente invención. En algunos casos, los términos con significados habitualmente entendidos se definen en el presente documento para mayor claridad y/o para referencia sencilla y no debería interpretarse necesariamente que la inclusión de tales definiciones en el presente documento representa una diferencia sustancial frente a lo que se entiende generalmente en la técnica. Muchas de las técnicas y procedimientos descritos o referidos en el presente documento se entienden bien y se emplean habitualmente usando metodología convencional por los expertos en la materia, tal como, por ejemplo, las metodologías de clonación molecular usadas ampliamente descritas en Sambrook y col., Molecular Cloning: A Laboratory Manual 2ª edición (1989) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. Según sea apropiado, los procedimientos que implican el uso de kits y reactivos disponibles en el mercado se llevan a cabo generalmente de acuerdo con los protocolos y/o parámetros definidos por el fabricante a no ser que se indique de otro modo.

35 Las expresiones “cáncer de próstata avanzado”, “cáncer de próstata avanzado localmente”, “enfermedad avanzada” y “enfermedad avanzada localmente” significa cánceres de próstata que se han extendido a través de la cápsula prostática y se pretende que incluyan enfermedad de etapa C según el sistema de la Asociación Urológica Americana (AUA), enfermedad de etapa C1-C2 según el sistema Whitmore-Jewett y enfermedad de etapa T3-T4 y N+ según el sistema TNM (tumor, nódulo, metástasis). En general, no se recomienda cirugía para pacientes con enfermedad localmente avanzada y estos pacientes tienen resultados sustancialmente menos favorables en comparación con pacientes que tienen cáncer de próstata localizado clínicamente (restringido a órgano). La enfermedad localmente avanzada se identifica clínicamente por pruebas palpables de induración más allá del límite lateral de la próstata o asimetría o induración por encima de la base de la próstata.

40 El cáncer de próstata localmente avanzado se diagnostica en la actualidad patológicamente después de prostatectomía radical si el tumor invade o penetra en la cápsula prostática, se extiende al margen quirúrgico o invade las vesículas seminales.

45 “Alterar el patrón de glucosilación nativo” pretende significar para los fines del presente documento suprimir uno o más restos de carbohidrato hallados en la secuencia nativa 238P1B2 (retirando el sitio de glucosilación subyacente o suprimiendo la glucosilación por medios químicos y/o enzimáticos) y/o añadir uno o más sitios de glucosilación que no están presentes en la secuencia nativa de 238P1B2. Además, la frase incluye cambios cualitativos en la glucosilación de las proteínas nativas, que implican un cambio en la naturaleza y proporciones de los diversos restos

de carbohidrato presentes.

5 El término “análogo” se refiere a una molécula que es estructuralmente similar o comparte atributos similares o correspondientes con otra molécula (por ejemplo, una proteína relacionada con 238P1B2). Por ejemplo un análogo de una proteína de 238P1B2 puede unirse específicamente con un anticuerpo o linfocito T que se une específicamente a 238P1B2.

El término “anticuerpo” se usa en el sentido más amplio. Por lo tanto un “anticuerpo” puede ser de origen natural o artificial tal como anticuerpos monoclonales producidos por tecnología de hibridoma convencional. Los anticuerpos anti-238P1B2 comprenden anticuerpos monoclonales y policlonales así como fragmentos que contienen el dominio de unión a antígeno y/o una o más regiones determinantes de complementariedad de estos anticuerpos.

10 Un “fragmento de anticuerpo” se define como al menos una parte de la región variable de la molécula de inmunoglobulina que se une a su diana, es decir, la región de unión al antígeno. Puede abarcar específicamente anticuerpos anti-238P1B2 sencillos y clones de los mismos (incluyendo anticuerpos agonistas, antagonistas y neutralizadores) y composiciones de anticuerpo anti-238P1B2 con especificidad de poliepitópica.

15 La expresión “secuencias con codones optimizados” se refiere a secuencias de nucleótidos que se han optimizado para una especie huésped particular reemplazando cualquier codón que tenga una frecuencia de uso de menos de aproximadamente 20 %. Las secuencias de nucleótidos que se han optimizado para expresión en una especie de huésped dada por eliminación de secuencias de poliadenilación espurias, eliminación de señales de corte y empalme de exones/intrones, eliminación de repeticiones de tipo transposón y/u optimización de contenido de GC además de optimización de codones se denominan en el presente documento “secuencias de expresión
20 potenciada”.

La expresión “agente citotóxico” se refiere a una sustancia que inhibe o evita la actividad de expresión de células, función de células y/o provoca destrucción de células. Se pretende que el término incluya agentes quimioterapéuticos de isótopos radiactivos y toxinas tales como toxinas de moléculas pequeñas o toxinas
25 enzimáticamente activas de origen bacteriano, fúngico, vegetal o animal, incluyendo fragmentos y/o variantes de las mismas. Los ejemplos de agentes citotóxicos incluyen, pero sin limitación maitansinoides, itrio, bismuto, ricina, cadena de ricina A, doxorubicina, daunorrubicina, taxol, bromuro de etidio, mitomicina, etopósido, tenopósido, vincristina, vinblastina, colchicina, dihidroxi antracina, actinomicina, toxina de la difteria, exotoxina de pseudomonas (PE) A, PE40, abrina, cadena de abrina A, cadena de modeccina A, alfa-sarcina, gelonina, mitogelina, retstrictocina, fenomicina, enomicina, curicina, crotina, caliqueamicina, inhibidor de *Saponaria officinalis* y
30 glucocorticoides y otros agentes quimioterapéuticos, así como radioisótopos tales como At²¹¹, I¹³¹, I¹²⁵, Y⁹⁰, Re¹⁸⁶, R¹⁸⁸, S¹⁵³, Bi²¹², P³² e isótopos radiactivos de Lu. Los anticuerpos también pueden conjugarse con una enzima activadora de profármaco antineoplásico capaz de convertir el profármaco a su forma activa.

El término “homólogo” se refiere a una molécula que muestra homología con otra molécula, por ejemplo, que tiene secuencias de restos químicos que son el mismo o similares en posiciones correspondientes.

35 El “antígeno de leucocitos humanos” o “HLA” es una proteína del complejo mayor de histocompatibilidad (MHC) de clase I o clase II (véase, por ejemplo, Stites, y col, INMUNOLOGÍA, 8ª ED., Lange Publishing, Los Altos, CA (1994).

Los términos “hibridar”, “hibridando”, “hibrida” y similares, usados en el contexto de polinucleótidos, pretenden referirse a condiciones de hibridación convencionales, preferentemente tales como hibridación en formamida 50 %/6XSSC/SDS 0,1 %/ADNmc 100 µg/ml, en la que las temperaturas para hibridación están por encima de 37 grados C y la temperatura para lavado en 0,1 XSSC/SDS 0,1 % están por encima de 55 grados C.
40

Las frases “aislado” o “biológicamente puro” se refieren a material que está sustancialmente o esencialmente sin componentes que normalmente acompañan al material como se encuentra en su estado nativo. Por lo tanto, los péptidos aislados descritos en el presente documento preferentemente no contienen materiales normalmente asociados con los péptidos en su ambiente *in situ*. Por ejemplo, se dice que un polinucleótido está “aislado” cuando
45 está sustancialmente separado de polinucleótidos contaminantes que corresponden o son complementarios a genes distintos de los genes de 238P1B2 o que codifican polipéptidos distintos del producto génico de 238P1B2 o fragmentos de los mismos. Un experto en la materia puede emplear fácilmente procedimientos de aislamiento de ácidos nucleicos para obtener un polinucleótido de 238P1B2 aislado.

Se dice que una proteína está “aislada” por ejemplo, cuando se emplean procedimientos físicos, mecánicos o químicos para retirar las proteínas de 238P1B2 de constituyentes celulares que normalmente están asociados con la proteína. Un experto en la materia puede emplear fácilmente procedimientos de purificación convencionales para obtener una proteína 238P1B2 aislada. Como alternativa, puede prepararse una proteína aislada por medios químicos.
50

El término “mamífero” se refiere a cualquier organismo clasificado como un mamífero, incluyendo ratones, ratas, conejos, perros, gatos, vacas, caballos y seres humanos. El mamífero puede ser un ratón. El mamífero puede ser un ser humano.
55

Las expresiones “cáncer de próstata metastásico” y “enfermedad metastásica” significan cánceres de próstata que se han propagado a ganglios linfáticos regionales o a sitios distantes y pretenden incluir enfermedad de etapa D según el sistema AUA y etapa TxNxM + según el sistema TNM. Como es el caso con cáncer de próstata localmente avanzado, la cirugía generalmente no está indicada para pacientes con enfermedad metastásica y la terapia hormonal (ablación de andrógenos) es una modalidad de tratamiento preferida. Los pacientes con cáncer de próstata metastásico con el tiempo desarrollan un estado refractario para andrógenos en un periodo de 12 a 18 meses de inicio del tratamiento. Aproximadamente la mitad de estos pacientes refractarios para andrógenos mueren en un periodo de 6 meses después de desarrollar ese estado. El sitio más habitual para metástasis de cáncer de próstata es el hueso. Las metástasis de hueso de cáncer de próstata son con frecuencia osteoblásticas en lugar de osteolíticas (es decir, dan como resultado formación de hueso neta). Las metástasis de hueso se encuentran más frecuentemente en la columna, seguido del fémur, pelvis, caja torácica, cráneo y húmero. Otros sitios habituales para metástasis incluyen ganglios linfáticos, pulmón, hígado y cerebro. El cáncer de próstata metastásico se diagnostica normalmente por linfadenectomía de pelvis laparoscópica o abierta, exploraciones de radionúclidos del cuerpo completo, radiografía del esqueleto y/o biopsia de lesión de hueso.

La expresión “anticuerpo monoclonal” se refiere a un anticuerpo obtenido de una población de anticuerpos sustancialmente homogénea, es decir, los anticuerpos que comprenden la población son idénticos excepto por posibles mutaciones de origen natural que están presentes en cantidades menores.

Un “motivo”, como en el motivo biológico de una proteína relacionada con 238P 1B2, se refiere a cualquier patrón de aminoácidos que forman parte de la secuencia primaria de una proteína, que se asocia con una función particular (por ejemplo, interacción proteína-proteína, interacción proteína-ADN, etc.) o modificación (por ejemplo, que está fosforilada, glucosilada o amidada), o localización (por ejemplo secuencia secretora, secuencia de localización nuclear, etc.) o una secuencia que se correlaciona con inmunogenicidad, de forma humoral o celular. Un motivo puede ser contiguo o capaz de alinearse con ciertas posiciones que generalmente se correlacionan con una cierta función o propiedad. En el contexto de motivos de HLA, “motivo” se refiere a un patrón de restos en un péptido de longitud definida, habitualmente un péptido de aproximadamente 8 a aproximadamente 13 aminoácidos para un motivo de HLA de clase I y de aproximadamente 6 a aproximadamente 25 aminoácidos para un motivo de HLA de clase II, que se reconoce por una molécula de HLA particular. Los motivos peptídicos para unión de HLA son normalmente diferentes para cada proteína codificada por cada alelo de HLA humano y difieren en el patrón de los restos de anclaje primarios y secundarios.

Un “excipiente farmacéutico” comprende un material tal como un adyuvante, un vehículo, agentes de ajuste de pH y tamponamiento, agentes de ajuste de tonicidad, agentes humectantes, conservante y similares.

“Farmacéuticamente aceptable” se refiere a una composición no tóxica, inerte y/o que es fisiológicamente compatible con seres humanos u otros mamíferos.

El término “polinucleótido” significa una forma polimérica de nucleótidos de al menos 10 bases o pares de bases de longitud, ribonucleótidos o desoxinucleótidos o una forma modificada de uno de los tipos de nucleótidos, y pretende incluir formas mono y bicatenarias de ADN y/o ARN. En la técnica, este término se usa con frecuencia de forma intercambiable con “oligonucleótido”. Un polinucleótido puede comprender una secuencia de nucleótidos desvelada en el presente documento en la que timina (T), como se muestra por ejemplo en la Figura 2, también puede ser uracilo (U); esta definición concierne a las diferencias entre las estructuras químicas de ADN y ARN, en particular la observación de que una de las cuatro bases principales de ARN es uracilo (U) en lugar de timina (T).

El término “polipéptido” significa un polímero de al menos aproximadamente 4, 5, 6, 7 u 8 aminoácidos. A lo largo de la memoria descriptiva, se usan designaciones convencionales de tres letras o una letra para los aminoácidos. En la técnica, este término se usa con frecuencia de forma intercambiable con “péptido” o “proteína”.

Un “resto de anclaje principal” de HLA es un aminoácido en una posición específica a lo largo de una secuencia peptídica que se entiende que proporciona un punto de contacto entre el péptido inmunogénico y la molécula de HLA. De uno a tres, habitualmente dos, restos de anclaje principales dentro de un péptido de longitud definida generalmente definen un “motivo” para un péptido inmunogénico. Se entiende que estos restos se ajustan en contacto cercano con el surco de unión a péptido de una molécula de HLA, con sus cadenas laterales internadas en bolsillos específicos del surco de unión. Por ejemplo, los restos de anclaje principales para una molécula de HLA de clase I pueden localizarse en la posición 2 (desde la posición amino terminal) y en la posición carboxilo terminal de un epítipo peptídico de 8, 9, 10, 11 o 12 restos descrito en el presente documento. Los restos de anclaje principales de un péptido que se unirán a una molécula de HLA de clase II pueden espaciarse entre sí, en lugar de con los extremos de un péptido, cuando el péptido es generalmente de al menos 9 aminoácidos de longitud.

Una molécula de ADN o ARN “recombinante” es una molécula de ADN o ARN que se ha sometido a manipulación molecular *in vitro*.

Los ejemplos no limitantes de moléculas pequeñas incluyen compuestos que se unen o interaccionan con 238P1B2, ligandos que incluyen hormonas, neuropéptidos, quimiocinas, odorantes, fosfolípidos y equivalentes funcionales de los mismos que se unen y preferentemente inhiben la función proteica de 238P1B2. Tales moléculas pequeñas no

limitantes preferentemente tienen un peso molecular de menos de aproximadamente 10 kDa, más preferentemente por debajo de aproximadamente 9, aproximadamente 8, aproximadamente 7, aproximadamente 6, aproximadamente 5 o aproximadamente 4 kDa. Las moléculas pequeñas pueden asociarse físicamente con, o unirse a, proteínas 238P1B2; no se encuentran en rutas metabólicas de origen natural; y/o son más solubles en soluciones acuosas que no acuosas.

La “rigurosidad” de las reacciones de hibridación la puede determinar fácilmente un experto en la materia y generalmente es un cálculo empírico dependiente de la longitud de la sonda, la temperatura de lavado y la concentración salina. En general, sondas más largas requieren temperaturas mayores para hibridación apropiada, mientras que sondas más cortas necesitan temperaturas más bajas. La hibridación depende en general de la capacidad de las secuencias de ácido nucleico desnaturalizadas para rehibridar cuando están presentes hebras complementarias en un ambiente por debajo de su temperatura de fusión. Cuanto mayor sea el grado de homología deseada entre la sonda y la secuencia hibridable, mayor es la temperatura relativa que puede usarse. Como resultado, se entiende que temperaturas relativas mayores tenderían a hacer las condiciones de reacción más rigurosas, mientras que las temperaturas más bajas las harían menos. Para detalles adicionales y explicación de la rigurosidad de reacciones de hibridación, véase Ausubel y col., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, (1995).

“Condiciones rigurosas” o “condiciones de alta rigurosidad”, como se define en el presente documento, se identifican por, pero sin limitación, las que: (1) emplean fuerza iónica baja y alta temperatura para lavar, por ejemplo cloruro sódico 0,015 M/citrato sódico 0,0015 M/dodecilsulfato sódico 0,1 % a 50 °C; (2) emplean durante la hibridación un agente desnaturalizante, tal como formamida, por ejemplo, formamida 50 % (v/v) con albúmina de suero bovino 0,1 %/Ficoll 0,1 %/polivinilpirrolidona 0,1 %/tampón de fosfato sódico 50 mM a pH 6,5 con cloruro sódico 750 mM, citrato sódico 75 mM a 42 °C; o (3) emplean formamida 50 %, SSC 5 x (NaCl 0,75 M, citrato sódico 0,075 M), fosfato sódico 50 mM (pH 6,8), pirofosfato sódico 0,1 %, solución de Denhardt 5 x, ADN de esperma de salmón sonicado (50 µg/ml), SDS 0,1 % y dextran sulfato 10 % a 42 °C, con lavados a 42 °C en SSC 0,2 x (cloruro sódico/citrato sódico) y formamida al 50 % a 55 °C, seguido de un lavado de alta rigurosidad que consiste en SSC 0,1 x que contiene EDTA a 55 °C. Las “condiciones moderadamente rigurosas” se describen por, pero sin limitación, las de Sambrook y col, *Molecular Cloning: A Laboratory Manual*, Nueva York: Cold Spring Harbor Press, 1989, e incluyen el uso de solución de lavado y condiciones de hibridación (por ejemplo, temperatura, fuerza iónica y % de SDS) menos rigurosas que las descritas anteriormente. Un ejemplo de condiciones moderadamente rigurosas es la incubación durante una noche a 37 °C en una solución que comprende: formamida 20 %, SSC 5 x (NaCl 150 mM, citrato trisódico 15 mM), fosfato sódico 50 mM (pH 7,6), solución de Denhardt 5 x, dextran sulfato 10 % y ADN de esperma de salmón roto desnaturalizado 20 mg/ml, seguido de lavado de los filtros en SSC 1 x a aproximadamente 37-50 °C. El experto en la materia reconocerá cómo ajustar la temperatura, fuerza iónica, etc. según sea necesario para adaptar factores tales como longitud de la sonda y similares.

Un “supermotivo” de HLA es una especificidad de unión a péptido compartida por las moléculas de HLA codificadas por dos o más alelos de HLA.

Como se usa en el presente documento “tratar” o “terapéutico” y términos relacionados gramaticalmente, se refieren a cualquier mejora de cualquier consecuencia de enfermedad, tal como supervivencia prolongada, menor morbilidad y/o una reducción de los efectos secundarios que son los productos secundarios de una modalidad terapéutica alternativa; no se requiere erradicación completa de la enfermedad.

Un “animal transgénico” (por ejemplo, un ratón o rata) es un animal que tiene células que contienen un transgén, introduciéndose dicho transgén en el animal o un ancestro del animal en una etapa prenatal, por ejemplo, embrionaria. Un “transgén” es un ADN que está integrado en el genoma de una célula de la que se desarrolla un animal transgénico.

Como se usa en el presente documento, una “vacuna” de respuesta inmune celular o HLA es una composición que contiene o codifica uno o más péptidos descritos en el presente documento. Existen numerosas vacunas tales, tales como un cóctel de uno o más péptidos individuales; uno o más péptidos descritos en el presente documento comprendidos por un péptido poliepitópico; o ácidos nucleicos que codifican tales péptidos o polipéptidos individuales, por ejemplo, un minigen que codifica un péptido poliepitópico. El “uno o más péptidos” pueden incluir cualquier número entero de unidad completa de 1 a 150 o más, por ejemplo, al menos, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145 ó 150 o más péptidos descritos en el presente documento. Los péptidos o polipéptidos pueden opcionalmente modificarse, tal como por lipidación, adición de secuencias de dirección u otras. Los péptidos de HLA de clase I descritos en el presente documento pueden mezclarse con, o ligarse a, péptidos de HLA de clase II, para facilitar la activación de tanto linfocitos T citotóxicos como linfocitos T auxiliares. Las vacunas de HLA también pueden comprender células presentadoras de antígenos pulsadas por péptidos, por ejemplo, células dendríticas.

El término “variante” se refiere a una molécula que muestra una variación de un tipo o norma descrito, tal como una proteína que tiene uno o más restos aminoácidos diferentes en la posición o posiciones correspondientes de una proteína específicamente descrita (por ejemplo, la proteína 238P1B2 mostrada en la Figura 2 o Figura 3). Un

análogo es un ejemplo de una proteína variante. Las isoformas de corte y empalme y polimorfismos de nucleótidos sencillos (SNP) son ejemplos adicionales de variantes.

“Las variantes relacionadas con 238P1B2” incluyen las específicamente identificadas en el presente documento, así como variantes alélicas, variantes de sustitución conservativas, análogos y homólogos que pueden aislarse/generarse y caracterizarse sin experimentación indebida siguiendo los procedimientos descritos en el presente documento o fácilmente disponibles en la técnica. Las proteínas de fusión que combinan partes de diferentes proteínas de 238P1B2 o fragmentos de las mismas, así como proteínas de fusión de una proteína 238P1B2 y un polipéptido heterólogo también están incluidas. Tales proteínas de 238P1B2 se denominan de forma colectiva proteínas relacionadas con 238P1B2, las proteínas descritas en el presente documento o 238P1B2. La expresión “proteína relacionada con 238P1B2” se refiere a un fragmento polipeptídico o una secuencia de proteína 238P1B2 de 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más de 25 aminoácidos; o, al menos 30, 35, 40, 45, 50, 55, 60, 65, 70, 80, 85, 90, 95, 100 o más de 100 aminoácidos.

II) Polinucleótidos de 238P1B2

Los inventores describen polinucleótidos correspondientes o complementarios a todo o parte de un gen, ARNm y/o secuencia codificante de 238P1B2, preferentemente en forma aislada, incluyendo polinucleótidos que codifican una proteína relacionada con 238P1B2 y fragmentos de la misma, ADN, ARN, híbrido ADN/ARN y moléculas relacionadas, polinucleótidos u oligonucleótidos complementarios a un gen 238P1B2 o secuencia de ARNm o una parte de los mismos y polinucleótidos u oligonucleótidos que hibridan con un gen 238P1B2, ARNm o con un polinucleótido que codifica 238P1B2 (colectivamente, “polinucleótidos de 238P1B2”). En todos los casos cuando se haga referencia en esta sección, T también puede ser U en la Figura 2.

Los polinucleótidos de 238P1B2 incluyen: un polinucleótido de 238P1B2 que tiene la secuencia mostrada en la Figura 2, la secuencia de nucleótidos de 238P1B2 como se muestra en la Figura 2 en la que T es U; al menos 10 nucleótidos contiguos de un polinucleótido que tiene la secuencia como se muestra en la Figura 2; o, al menos 10 nucleótidos contiguos de un polinucleótido que tiene la secuencia como se muestra en la Figura 2 en la que T es U. Por ejemplo, los nucleótidos de 238P1B2 comprenden, sin limitación:

(I) un polinucleótido que comprende, que consiste esencialmente en, o consiste en una secuencia como se muestra en la Figura 2, en la que T también puede ser U;

(II) un polinucleótido que comprende, consiste esencialmente en, o consiste en la secuencia como se muestra en la Figura 2A, del número de resto nucleotídico 2133 hasta el número de resto nucleotídico 2894, seguido de un codón de parada, en la que T también puede ser U;

(III) un polinucleótido que comprende, consiste esencialmente en, o consiste en la secuencia como se muestra en la Figura 2B, del número de resto nucleotídico 1947 hasta el número de resto nucleotídico 2894, seguido de un codón de parada, en la que T también puede ser U;

(IV) un polinucleótido que comprende, consiste esencialmente en, o consiste en la secuencia como se muestra en la Figura 2C, del número de resto nucleotídico 2133 hasta el número de resto nucleotídico 2894, seguido de un codón de parada, en la que T también puede ser U;

(V) un polinucleótido que comprende, consiste esencialmente en, o consiste en la secuencia como se muestra en la Figura 2D, del número de resto nucleotídico 2133 hasta el número de resto nucleotídico 2894, seguido de un codón de parada, en la que T también puede ser U;

(VI) un polinucleótido que comprende, consiste esencialmente en, o consiste en la secuencia como se muestra en la Figura 2E, del número de resto nucleotídico 2133 hasta el número de resto nucleotídico 2894, seguido de un codón de parada, en la que T también puede ser U;

(VII) un polinucleótido que comprende, consiste esencialmente en, o consiste en la secuencia como se muestra en la Figura 2F, del número de resto nucleotídico 2133 hasta el número de resto nucleotídico 2894, seguido de un codón de parada, en la que T también puede ser U;

(VIII) un polinucleótido que comprende, consiste esencialmente en, o consiste en la secuencia como se muestra en la Figura 2G, del número de resto nucleotídico 2133 hasta el número de resto nucleotídico 2894, seguido de un codón de parada, en la que T también puede ser U;

(IX) un polinucleótido que codifica una proteína relacionada con 238P1B2 que es al menos 90 % homóloga con una secuencia de aminoácidos completa mostrada en la Figura 2A-G;

(X) un polinucleótido que codifica una proteína relacionada con 238P1B2 que es al menos 90 % idéntica a una secuencia de aminoácidos completa mostrada en la Figura 2A-G;

(XI) un polinucleótido que codifica al menos un péptido expuesto en las Tablas V-XVIII o Tabla XIX;

- 5 (XII) un polinucleótido que codifica una región peptídica de al menos 5 aminoácidos de un péptido de la Figura 3A en cualquier incremento de número entero hasta 254 que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil de Hidrofilia de la Figura 5A o de la Figura 3B en cualquier incremento de número completo hasta 316 que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil de Hidrofilia de la Figura 5B;
- 10 (XIII) un polinucleótido que codifica una región peptídica de al menos 5 aminoácidos de un péptido de la Figura 3A en cualquier incremento de número entero hasta 254 que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil hidropático de la Figura 6A o de la Figura 3B en cualquier incremento de número entero hasta 316, que incluye una posición de aminoácido que tiene un valor menor de 0,5 en el perfil hidropático de la Figura 6B;
- 15 (XIV) un polinucleótido que codifica una región peptídica de al menos 5 aminoácidos de un péptido de la Figura 3A en cualquier aumento de número entero hasta 254 que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil de porcentaje de restos accesibles de la Figura 7A o de la Figura 3B en cualquier incremento de número entero hasta 316, que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil de porcentaje de restos accesibles de la Figura 7B;
- 20 (XV) un polinucleótido que codifica una región peptídica de al menos 5 aminoácidos de un péptido de la Figura 3A en cualquier incremento de número entero hasta 254 que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil de Flexibilidad Media en la Figura 8A o de la Figura 3B en cualquier incremento de número entero hasta 316, que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil de Flexibilidad Media en la Figura 8B;
- 25 (XVI) un polinucleótido que codifica una región peptídica de al menos 5 aminoácidos de un péptido de la Figura 3A en cualquier incremento de número entero hasta 254 que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil de giro Beta de la Figura 9A o de la Figura 3B en cualquier aumento de número entero hasta 316, que incluye una posición de aminoácido que tiene un valor mayor de 0,5 en el perfil de giro Beta de la Figura 9B;
- 30 (XVII) un polinucleótido que codifica una proteína relacionada con 238P1B2 cuya secuencia está codificada por los ADNc contenidos en el plásmido depositado en la Colección Americana de Cultivos Tipo como N° de Acceso PTA-4124 el 7 de marzo de 2002;
- (XVIII) un polinucleótido que es completamente complementario con un polinucleótido de uno cualquiera de (I)-(XVII);
- (XIX) un polinucleótido que hibrida selectivamente en condiciones rigurosas con un polinucleótido de (I)-(XVIII);
- (XX) un péptido que está codificado por cualquiera de (I)-(XIX); y
- (XXI) un polinucleótido de cualquiera de (I)-(XIX) o péptido de (XX) junto con un excipiente farmacéutico y/o en una forma de dosis unitaria humana.
- 35 Como se usa en el presente documento, se entiende que un intervalo desvela específicamente todas las posiciones de unidades enteras del mismo. Los polinucleótidos de 238P1B2 típicos codifican partes específicas de las secuencias de ARNm de 238P1B2 (y las que son complementarias para tales secuencias) tales como las que codifican las proteínas y/o fragmentos de las mismas, por ejemplo:
- 40 (a) 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250 o 254 aminoácidos contiguos de 238P1B2;
- 45 (b) 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250 o 254 aminoácidos contiguos de la variante 1A;
- (c) 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305, 310, 315 o 316 aminoácidos contiguos de la variante 1B; o
- 50 (d) 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250 o 254 aminoácidos contiguos de la variante 2.

55 Por ejemplo, se describen en el presente documento polinucleótidos y sus péptidos codificados, los polinucleótidos representativos pueden codificar de aproximadamente aminoácido 1 a aproximadamente aminoácido 10 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 10 a aproximadamente aminoácido 20 de la proteína 238P1B2 o variantes mostradas

en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 20 a aproximadamente aminoácido 30 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 30 a aproximadamente aminoácido 40 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 40 a aproximadamente aminoácido 50 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 50 a aproximadamente aminoácido 60 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 60 a aproximadamente aminoácido 70 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 70 a aproximadamente aminoácido 80 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 80 a aproximadamente aminoácido 90 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3, polinucleótidos que codifican de aproximadamente aminoácido 90 a aproximadamente aminoácido 100 de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3 o regiones codificantes de aproximadamente aminoácido 100 a aminoácidos posteriores en la secuencia, en aumentos de aproximadamente 10 aminoácidos, terminando en el aminoácido carboxilo terminal expuesto en la Figura 2 o Figura 3. En consecuencia los polinucleótidos que codifican partes de la secuencia de aminoácidos (de aproximadamente 10 aminoácidos), de los aminoácidos 100 hasta el aminoácido carboxilo terminal de la proteína 238P1B2 se describen en el presente documento en el que se entiende que cada posición de aminoácido particular desvela esa posición más o menos cinco restos aminoacídicos.

También se describen en el presente documento polinucleótidos que codifican partes relativamente largas de una proteína 238P1B2. Por ejemplo, pueden generarse polinucleótidos que codifican de aproximadamente aminoácido 1 (o 20, 30 ó 40 etc.) a aproximadamente aminoácido 20 (o 30, 40 ó 50 etc.) de la proteína 238P1B2 o variantes mostradas en la Figura 2 o Figura 3 por diversas técnicas bien conocidas en la materia. Estos fragmentos polinucleotídicos pueden incluir cualquier parte de la secuencia de 238P1B2 o variantes como se muestra en la Figura 2.

Los polinucleótidos ilustrativos adicionales incluyen fragmentos polinucleotídicos de 238P1B2 que codifican uno o más de los motivos biológicos contenidos dentro de una secuencia de proteína 238P1B2 o secuencia variante.

Los inventores describen epítopos de anticuerpo que comprenden una región peptídica o un oligonucleótido que codifica la región peptídica, que tienen dos, tres, cuatro o cinco de las siguientes características:

- i) una región peptídica de al menos 5 aminoácidos de un péptido particular de la Figura 3, en cualquier incremento de número entero hasta la longitud completa de esa proteína en la Figura 3, que incluye una posición de aminoácido que tiene un valor igual a o mayor de 0,5, 0,6, 0,7, 0,8, 0,9 o que tiene un valor igual a 1,0, en el perfil de Hidrofilia de la Figura 5;
- ii) una región peptídica de al menos 5 aminoácidos de un péptido particular de la Figura 3, en cualquier incremento de número entero hasta la longitud completa de esa proteína en la Figura 3, que incluye una posición de aminoácido que tiene un valor igual o menor de 0,5, 0,4, 0,3, 0,2, 0,1 o que tiene un valor igual a 0,0 en el perfil hidropático de la Figura 6;
- iii) una región peptídica de al menos 5 aminoácidos de un péptido particular de la Figura 3, en cualquier incremento de número entero hasta la longitud completa de esa proteína en la Figura 3, que incluye una posición de aminoácido que tiene un valor igual o mayor de 0,5, 0,6, 0,7, 0,8, 0,9 o que tiene un valor igual a 1,0, en el perfil de Porcentaje de Restos Accesibles de la Figura 7;
- iv) una región peptídica de al menos 5 aminoácidos de un péptido particular de la Figura 3, en cualquier incremento de número entero hasta la longitud completa de esa proteína en la Figura 3, que incluye una posición de aminoácido que tiene un valor igual o mayor de 0,5, 0,6, 0,7, 0,8, 0,9 o que tiene un valor igual a 1,0, en el perfil de Flexibilidad Media de la Figura 8; o
- v) una región peptídica de al menos 5 aminoácidos de un péptido particular de la Figura 3, en cualquier incremento de número entero hasta la longitud completa de esa proteína en la Figura 3, que incluye una posición de aminoácido que tiene un valor igual o menor de 0,5, 0,6, 0,7, 0,8, 0,9 o que tiene un valor igual a 1,0, en el perfil de giro Beta de la Figura 9.

Los fragmentos polinucleotídicos típicos codifican una o más regiones de la proteína 238P1B2 o variante que muestra homología con una molécula conocida. Los fragmentos polinucleotídicos típicos pueden codificar uno o más de los sitios de N-glucosilación de la proteína 238P1B2 o variante, sitios de fosforilación de proteína quinasa dependiente de AMPc y GMPc, sitios de fosforilación de caseína quinasa II o sitio de N-miristilación y sitios de amidación.

II.A.) Usos de polinucleótidos de 238P1B2

II.A.1.) Control de anomalías genéticas

Los polinucleótidos de los párrafos precedentes tienen varios usos específicos diferentes. El gen de 238P1B2 humano se mapea en la localización cromosómica expuesta en el Ejemplo 3. Por ejemplo, debido a que el gen 238P1B2 se mapea en este cromosoma, los polinucleótidos que codifican regiones diferentes de las proteínas

238P1B2 se usan para caracterizar anomalías citogenéticas de esta localización cromosómica, tales como anomalías que se identifican como asociadas con diversos cánceres. En ciertos genes, se han identificado diversas anomalías cromosómicas incluyendo reordenamientos como anomalías citogenéticas frecuentes en varios cánceres diferentes (véase por ejemplo Krajcinovic y col, *Mutat. Res* 382(3-4): 81-83 (1998); Johansson y col., *Blood* 86(10): 3905-3914 (1995) y Finger y col, *P.N.A.S.*: 85(23) 9158-9162 (1988)). Por lo tanto, los polinucleótidos que codifican regiones específicas de las proteínas 238P1B2 proporcionan nuevas herramientas que pueden usarse para delinear, con mayor precisión de lo que era posible previamente, anomalías citogenéticas en la región cromosómica que codifica 238P1B2 que pueden contribuir al fenotipo maligno. En este contexto, estos polinucleótidos satisfacen una necesidad en la técnica de expandir la sensibilidad de exploración cromosómica para identificar anomalías cromosómicas más sutiles y menos habituales (véase por ejemplo, Evans y col, *Am J. Obstet. Gynecol* 171(4): 1055-1057 (1994)).

Además, como se demostró que 238P1B2 tenía alta expresión en vejiga y otros cánceres, se usaron polinucleótidos de 238P1B2 en procedimientos que evalúan el estado de los productos génicos de 238P1B2 en tejidos normales frente a cancerosos. Normalmente, los polinucleótidos que codifican regiones específicas de las proteínas 238P1B2 se usan para evaluar la presencia de perturbaciones (tales como deleciones, inserciones, mutaciones puntuales o alteraciones que dan como resultado una pérdida de un antígeno, etc.) en regiones específicas del gen de 238P1B2, tales como regiones que contienen uno o más motivos. Los ensayos ejemplares incluyen tanto ensayos de RT-PCR como análisis de polimorfismo de conformación monocatenaria (SSCP) (véase, por ejemplo, Marrogi y col., *J. Cutan. Pathol.* 26(8): 369-378 (1999)), ambos de los cuales usan polinucleótidos que codifican regiones específicas de una proteína para examinar esas regiones dentro de la proteína.

II.A.2.) Antisentido

Los inventores describen ADN genómicos, ADNc, ribozimas y moléculas antisentido, así como moléculas de ácido nucleico basadas en una cadena principal alternativa, o que incluyen bases alternativas, derivadas de fuentes naturales o sintetizadas, e incluyen moléculas capaces de inhibir el ARN o la expresión proteica de 238P1B2. Por ejemplo, las moléculas antisentido pueden ser ARN u otras moléculas, incluyendo ácidos péptido nucleicos (PNA) o moléculas no de ácido nucleico tales como derivados de fosforotioato, que se unen específicamente a ADN o ARN de una manera dependiente de pares de bases. Un experto en la materia puede obtener fácilmente estas clases de moléculas de ácido nucleico usando los polinucleótidos de 238P1B2 y las secuencias polinucleotídicas desveladas en el presente documento.

La tecnología antisentido implica la administración de oligonucleótidos exógenos que se unen a un polinucleótido diana localizado dentro de las células. El término "antisentido" se refiere al hecho de que tales oligonucleótidos son complementarios para sus dianas intracelulares, por ejemplo, 238P1B2. Véase, por ejemplo, Jack Cohen, *Oligodeoxynucleotides, Antisense Inhibitors of Gene Expression*, CRC Press, 1989; y *Synthesis* 1:1-5 (1988). Los oligonucleótidos antisentido de 238P1B2 descritos en el presente documento incluyen derivados tales como S-oligonucleótidos (derivados de fosforotioato o S-oligos, véase, Jack Cohen, mencionado anteriormente), que muestran acción inhibitoria de crecimiento celular canceroso potenciada. Los S-oligos (nucleósido fosforotioatos) son análogos isoelectrónicos de un oligonucleótido (O-oligo) en el que un átomo de oxígeno no de unión del grupo de fosfato se reemplaza por un átomo de azufre. Los S-oligos descritos en el presente documento pueden prepararse por tratamiento de los O-oligos correspondientes con 3H-1,2-benzoditiol-3-ona-1,1-dióxido, que es un reactivo de transferencia de azufre. Véase, por ejemplo, Iyer, R. P. y col, *J. Org. Chem.* 55:4693-4698 (1990); e Iyer, R. P. y col, *J. Am. Chem Soc.* 112:1253-1254 (1990). Los oligonucleótidos antisentido de 238P1B2 adicionales incluyen oligonucleótidos antisentido morfolinicos conocidos en la técnica (véase, por ejemplo, Partridge y col, 1996, *Antisense & Nucleic Acid Drug Development* 6:169-175).

Los oligonucleótidos antisentido de 238P1B2 descritos en el presente documento normalmente pueden ser ARN o ADN que es complementario e hibrida de forma estable con los primeros 100 codones 5' o los últimos 100 codones 3' de una secuencia genómica de 238P1B2 o el ARNm correspondiente. No se requiere complementariedad absoluta, aunque se prefieren altos grados de complementariedad. El uso de un oligonucleótido complementario a esta región permite la hibridación selectiva con ARNm de 238P1B2 y no con ARNm que especifica otras subunidades reguladoras de proteína quinasa. Los oligonucleótidos antisentido de 238P1B2 pueden ser fragmentos de 15 a 30 mer de la molécula de ADN antisentido que tienen una secuencia que hibrida con ARNm de 238P1B2. Opcionalmente, el oligonucleótido antisentido de 238P1B2 es un oligonucleótido de 30-mer que es complementario a una región en los primeros 10 codones 5' o los últimos 10 codones 3' de 238P1B2. Como alternativa, las moléculas antisentido se modifican para emplear ribozimas en la inhibición de la expresión de 238P1B2, véase, por ejemplo, L. A. Couture & D. T. Stinchcomb; *Trends Genet* 12:510-515 (1996).

II.A.3.) Cebadores y pares de cebadores

Los inventores describen cebadores y pares de cebadores, que permiten la amplificación específica de polinucleótidos descritos en el presente documento o de cualquier parte específica de los mismos y sondas que hibridan selectiva o específicamente con moléculas de ácido nucleico descritas en el presente documento o con cualquier parte de las mismas. Las sondas pueden marcarse con un marcador detectable, tal como, por ejemplo, un radioisótopo, compuesto fluorescente, compuesto bioluminiscente, un compuesto quimioluminiscente, quelante

metálico o enzima. Tales sondas y cebadores se usan para detectar la presencia de un polinucleótido de 238P1B2 en una muestra y como un medio para detectar una célula que exprese una proteína 238P1B2.

5 Ejemplos de tales sondas incluyen polipéptidos que comprenden toda o parte de la secuencia de ADNc de 238P1B2 humana mostrada en la Figura 2. Ejemplos de pares de cebadores capaces de amplificar específicamente ARNm de 238P1B2 también se describen en los ejemplos. Como se entenderá por el experto en la materia, puede prepararse una gran cantidad de cebadores y sondas diferentes basándose en las secuencias proporcionadas en el presente documento y usarse eficazmente para amplificar y/o detectar un ARNm de 238P1B2.

10 Los polinucleótidos de 238P1B2 descritos en el presente documento son útiles para diversos fines, incluyendo pero sin limitación su uso como sondas y cebadores para la amplificación y/o detección del gen o genes de 238P1B2, ARNm o fragmentos de los mismos; como reactivos para el diagnóstico y/o pronóstico de cáncer de próstata y otros cánceres; como secuencias codificantes capaces de dirigir la expresión de polipéptidos de 238P1B2; como herramientas para modular o inhibir la expresión del gen o genes de 238P1B2 y/o traducción del transcrito o transcritos de 238P1B2; y como agentes terapéuticos.

15 Cualquier sonda como se describe en el presente documento puede usarse para identificar y aislar una 238P1B2 o una secuencia de ácido nucleico relacionada con 238P1B2 de una fuente de origen natural, tal como seres humanos u otros mamíferos, así como la secuencia de ácido nucleico aislada por sí misma, que comprendería todas o la mayoría de las secuencias halladas en la sonda usada.

A. II.4) Aislamiento de moléculas de ácido nucleico que codifican 238P1B2

20 Las secuencias de ADNc de 238P1B2 descritas en el presente documento permiten el aislamiento de otros polinucleótidos que codifican producto o productos génicos de 238P1B2, así como el aislamiento de polinucleótidos que codifican homólogos de producto génico de 238P1B2, isoformas de corte y empalme alternativo, variantes alélicas y formas mutantes de un producto génico de 238P1B2 así como polinucleótidos que codifican análogos de proteínas relacionadas con 238P1B2. Se conocen bien diversos procedimientos de clonación molecular que pueden emplearse para aislar ADNc de longitud completa que codifican un gen de 238P1B2 (véase, por ejemplo, Sambrook, J. y col., Molecular Cloning: A Laboratory Manual, 2ª edición, Cold Spring Harbor Press, Nueva York, 1989; Current Protocols in Molecular Biology. Ausubel y col., Eds., Wiley y Sons, 1995). Por ejemplo, pueden emplearse convenientemente metodologías de clonación de fago lambda, usando sistemas de clonación disponibles en el mercado (por ejemplo, Lambda ZAP Express, Stratagene). Los clones de fago que contienen ADNc del gen de 238P1B2 pueden identificarse explorando con un ADNc de 238P1B2 marcado o un fragmento del mismo. Por ejemplo, puede sintetizarse un ADNc de 238P1B2 (por ejemplo, Figura 2) o una parte del mismo y usarse como una sonda para recuperar ADNc solapantes y de longitud completa correspondientes a un gen de 238P1B2. Un gen de 238P1B2 en sí mismo puede aislarse por exploración de genotecas de ADN genómico, bibliotecas de cromosomas artificiales de bacteria (BAC), bibliotecas de cromosomas artificiales de levadura (YAC), y similares, con sondas o cebadores de ADN de 238P1B2.

II.A.5.) Moléculas de ácido nucleico recombinante y sistemas de huésped-vector

40 Las moléculas de ADN o ARN recombinantes pueden contener polinucleótido de 238P1B2, un fragmento, análogo u homólogo del mismo, incluyendo pero sin limitación fagos, plásmidos, fagémidos, cósmidos, YAC, BAC, así como diversos vectores víricos y no víricos bien conocidos en la técnica, y células transformadas o transfectadas con tales moléculas de ADN o ARN recombinantes. Se conocen bien procedimientos para generar tales moléculas (véase, por ejemplo, Sambrook y col, 1989, mencionado anteriormente).

45 Un sistema de huésped-vector puede comprender una molécula de ADN recombinante que contiene un polinucleótido de 238P1B2, fragmento, análogo u homólogo del mismo dentro de una célula huésped procarionota o eucariota adecuada. Los ejemplos de células huésped eucariotas adecuadas incluyen una célula de levadura, una célula vegetal o una célula animal, tales como una célula de mamífero o una célula de insecto (por ejemplo, una célula infectable por baculovirus tal como una célula Sf9 o HighFive). Los ejemplos de células de mamífero adecuadas incluyen diversas líneas celulares de cáncer de próstata tales como DU145 y TsuPr1, otras líneas celulares de cáncer de próstata transfectables o transducibles, células primarias (PrEC), así como varias células de mamífero usadas habitualmente para la expresión de proteínas recombinantes (por ejemplo, células COS, CHO, 293, 293T). Más particularmente, un polinucleótido que comprende la secuencia codificante de 238P1B2 o un fragmento, análogo u homólogo de la misma puede usarse para generar proteínas 238P1B2 o fragmentos de las mismas usando cualquier variedad de sistemas de huésped- vector usadas habitualmente y conocidos ampliamente en la técnica.

55 Está disponible una amplia serie de sistemas de huésped-vector adecuados para la expresión de proteínas 238P1B2 o fragmentos de las mismas, véase, por ejemplo, Sambrook y col., 1989, mencionado anteriormente; Current Protocols in Molecular Biology, de 1995, mencionado anteriormente). Los vectores preferidos para expresión de mamíferos incluyen pero sin limitación pcDNA 3.1 marcador de His-myc (Invitrogen) y el vector retroviral pSRatKneo (Muller y col, 1991, MCB 11:1785). Usando estos vectores de expresión, 238P1B2 puede expresarse en varias líneas celulares de cáncer de próstata y no de próstata, incluyendo por ejemplo, 293, 293T, rat-1, NIH 3T3 y TsuPr1.

Estos sistemas de huésped-vector son útiles para la producción de una proteína de 238P1B2 o fragmento de la misma. Tales sistemas de huésped-vector pueden emplearse para estudiar las propiedades funcionales de 238P1B2 y mutaciones o análogos de 238P1B2.

5 Proteína 238P1B2 humana recombinante o un análogo u homólogo o fragmento de la misma pueden producirla células de mamífero transfectadas con una construcción que codifica un nucleótido relacionado con 238P1B2. Por ejemplo, las células 293T pueden transfectarse con un plásmido de expresión que codifica 238P1B2 o fragmento, análogo u homólogo de la misma, una proteína relacionada con 238P1B2 se expresa en las células 293T y la proteína 238P1B2 recombinante se aísla usando procedimientos de purificación convencionales (por ejemplo, purificación de afinidad usando anticuerpos anti-238P1B2). Puede subclonarse una secuencia codificante de 10 238P1B2 en el vector retroviral pSRaMSVtkneo y usarse para infectar diversas líneas celulares de mamífero, tales como NIH 3T3, TsuPr1, 293 y rat-1 para establecer líneas celulares que expresen 238P1B2. También pueden emplearse diversos otros sistemas de expresión bien conocidos en la técnica. Pueden usarse construcciones de expresión que codifican un péptido líder unido en fase con una secuencia codificante de 238P1B2 para la generación de una forma secretada de proteína 238P1B2 recombinante.

15 Como se analiza en el presente documento, la redundancia en el código genético permite variación en las secuencias génicas de 238P1B2. En particular, se sabe en la técnica que especies de huésped específicas con frecuencia tienen preferencias codónicas específicas y por lo tanto se puede adaptar la secuencia desvelada según se prefiera para un huésped deseado. Por ejemplo, las secuencias codónicas análogas preferidas normalmente tienen codones poco habituales (es decir, codones que tienen una frecuencia de uso de menos de aproximadamente 20 % en secuencias conocidas del huésped deseado) reemplazados con codones de mayor frecuencia. Las 20 preferencias codónicas para una especie específica se calculan, por ejemplo, usando tablas de uso de codones disponibles en INTERNET tales como en la dirección www.dna.affrc.govp/-nakamura/codon.html.

25 Se sabe que modificaciones de secuencia adicionales potencian la expresión proteica en un huésped celular. Estas incluyen eliminación de secuencias que codifican señales de poliadenilación espurias, señales de sitio de corte y empalme de exones/intrones, repeticiones de tipo transposón y/u otras secuencias bien caracterizadas tales que son deletéreas para la expresión génica. El contenido de GC de la secuencia se ajusta a niveles medios para un huésped celular dado, según se calcula por referencia a genes conocidos expresados en la célula huésped. Cuando sea posible, la secuencia se modifica para evitar estructuras de ARNm secundarias en horquilla predichas. Otras modificaciones útiles incluyen la adición de una secuencia consenso de inicio de la traducción al comienzo del marco 30 de lectura abierto, como se describe en Kozak, Mol Cell Biol., 9:5073-5080 (1989). Los expertos en la materia entienden que la regla general de que los ribosomas eucariotas inician la traducción exclusivamente en el codón AUG 5' proximal se anula solamente en condiciones poco habituales (véase, por ejemplo, Kozak PNAS: 92(7) 2662-2666, (1995) y Kozak NAR 15(20): 8125-8148 (1987)).

III.) Proteínas relacionadas con 238P1B2

35 Los inventores describen proteínas relacionadas con 238P1B2. Las proteínas 238P1B2 específicas comprenden un polipéptido que tiene toda o parte de la secuencia de aminoácidos de 238P1B2 humana como se muestra en la Figura 2 o Figura 3. Como alternativa, las proteínas 238P1B2 comprenden polipéptidos variantes, homólogos o análogos que tienen alteraciones en la secuencia de aminoácidos de 238P1B2 mostrada en la Figura 2 o Figura 3.

40 En general, las variantes alélicas de origen natural de 238P1B2 humana comparten un alto grado de identidad estructural y homología (por ejemplo, 90 % o más homología). Normalmente, las variantes alélicas de una proteína 238P1B2 contienen sustituciones de aminoácidos conservativas dentro de las secuencias de 238P1B2 descritas en el presente documento o contienen una sustitución de un aminoácido de una posición correspondiente en un homólogo de 238P1B2. Una clase de variantes alélicas de 238P1B2 son proteínas que comparten un alto grado de homología con al menos una región pequeña de una secuencia de aminoácidos de 238P1B2 particular, pero 45 contienen adicionalmente una diferencia radical de la secuencia, tal como una sustitución no conservativa, truncación, inserción o desplazamiento de fase. En comparaciones de secuencias proteicas, los términos, similitud, identidad y homología tienen cada uno un significado distinto como se aprecia en el campo de la genética. Además, la ortología y paralogía pueden ser conceptos importantes que describen la relación de miembros de una familia proteica dada en un organismo con los miembros de la misma familia en otros organismos.

50 Se proporcionan abreviaturas de aminoácidos en la Tabla II. Pueden realizarse frecuentemente sustituciones de aminoácidos conservativas en una proteína sin alterar la conformación o la función de la proteína. Las proteínas descritas en el presente documento pueden comprender 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 sustituciones conservativas. Tales cambios incluyen sustituir con cualquiera de isoleucina (I), valina (V) y leucina (L) cualquier otro de estos aminoácidos hidrófobos; ácido aspártico (D) a ácido glutámico (E) y viceversa; glutamina (Q) a asparagina (N) y viceversa; y serina (S), a treonina (T) y viceversa. Otras sustituciones también pueden considerarse 55 conservativas, dependiendo del ambiente del aminoácido particular y su papel en la estructura tridimensional de la proteína. Por ejemplo, glicina (G) y alanina (A) pueden ser con frecuencia intercambiables, al igual que alanina (A) y valina (V). La metionina (M), que es relativamente hidrófoba, puede con frecuencia intercambiarse con leucina e isoleucina y en ocasiones con valina. Lisina (K) y arginina (R) son con frecuencia intercambiables en localizaciones 60 en las que la característica significativa del resto aminoacídico es su carga y las pK diferentes de estos dos restos

aminoacídicos no son significativas. Otros cambios más pueden considerarse “conservativos” en ambientes particulares (véase, por ejemplo, Tabla III en el presente documento; páginas 13-15 “Biochemistry” 2ª ED. Lubert Stryer ed (Stanford University); Henikoff et al., PNAS 1992 Vol 89 10915-10919; Lei et al., J Biol Chem 19 mayo 1995; 270 (20): 11882-6).

5 Los inventores describen una amplia diversidad de variantes o análogos aceptados en la técnica de proteínas 238P1B2 tales como polipéptidos que tienen inserciones, deleciones y sustituciones de aminoácidos. Pueden realizarse variantes de 238P1B2 usando procedimientos conocidos en la técnica tales como mutagénesis dirigida, exploración de alanina y mutagénesis por PCR. Puede realizarse mutagénesis dirigida (Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res, 10:6487 (1987)), mutagénesis de casete (Wells et al., Gene, 10
10 34:315 (1985)), mutagénesis de selección de restricción (Wells et al., Philos. Trans. R. Soc London SerA, 317:415 (1986)) u otras técnicas conocidas en el ADN clonado para producir el ADN variante de 238P1B2.

También puede emplearse análisis de exploración de aminoácidos para identificar uno o más aminoácidos junto con una secuencia contigua que está implicada en una actividad biológica específica tal como una interacción proteína-proteína. Entre los aminoácidos de exploración preferidos hay aminoácidos neutros relativamente pequeños. Tales aminoácidos incluyen alanina, glicina, serina y cisteína. La alanina es normalmente un aminoácido de exploración preferido entre este grupo porque elimina la cadena lateral detrás del carbono beta y es menos probable que altere la conformación de la cadena principal de la variante. La alanina también se prefiere normalmente porque es el aminoácido más habitual. Además, se encuentra con frecuencia en posiciones tanto expuestas como internadas (Creighton, The Proteins, (W. H. Freeman & Co., N.Y.); Chothia, J. Mol Biol., 150:1 (1976)). Si la sustitución de alanina no produce cantidades adecuadas de variante, puede usarse un aminoácido isostérico.

Como se define en el presente documento, las variantes, análogos u homólogos de 238P1B2, tienen el atributo distintivo de tener al menos un epítipo que “reacciona de forma cruzada” con una proteína 238P1B2 que tiene una secuencia de aminoácidos de la Figura 3. Como se usa en esta frase, “reacciona de forma cruzada” significa que un anticuerpo o linfocito T que se une específicamente a una variante de 238P1B2 también se une específicamente a una proteína de 238P1B2 que tiene una secuencia de aminoácidos expuesta en la Figura 3. Un polipéptido deja de ser una variante de una proteína mostrada en la Figura 3, cuando ya no contiene ningún epítipo que pueda reconocerse por un anticuerpo o linfocito T que se une específicamente a la proteína 238P1B2 de partida. Los expertos en la materia entienden que los anticuerpos que reconocen proteínas se unen a epítopos de diverso tamaño y un agrupamiento del orden de aproximadamente cuatro o cinco aminoácidos, contiguos o no, se considera un número típico de aminoácidos en un epítipo mínimo. Véase, por ejemplo, Nair et al., J. Immunol 2000 165 (12): 6949-6955; Hebbes et al., Mol Immunol (1989) 26(9): 865-73; Schwartz et al., J Immunol (1985) 135 (4): 2598-608.

Otras clases de variantes de proteína relacionadas con 238P1B2 comparten el 70 %, 75 %, 80 %, 85 % o 90 % o más similitud con una secuencia de aminoácidos de la Figura 3, o un fragmento de la misma. Otra clase específica de variantes o análogos de proteína 238P1B2 comprende uno o más de los motivos biológicos de 238P1B2 descritos en el presente documento o actualmente conocidos en la técnica. Por lo tanto, los análogos de fragmentos de 238P1B2 (nucleicos o aminoacídicos) pueden tener propiedades funcionales alteradas (por ejemplo, inmunogénicas) en relación con el fragmento de partida. También debe apreciarse que los motivos que son ahora o se convierten en parte de la técnica deben aplicarse a las secuencias de ácidos nucleicos o aminoácidos de la Figura 2 o Figura 3.

40 Como se ha analizado en el presente documento, los polipéptidos pueden contener menos de la secuencia de aminoácidos completa de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3. Por ejemplo, los péptidos/proteínas representativos pueden tener cualquiera de 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 o más aminoácidos contiguos de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3.

Además, los polipéptidos representativos pueden consistir en aproximadamente aminoácido 1 a aproximadamente a aproximadamente aminoácido 10 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 10 a aproximadamente aminoácido 20 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 20 a aproximadamente aminoácido 30 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 30 a aproximadamente aminoácido 40 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 40 a aproximadamente aminoácido 50 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 50 a aproximadamente aminoácido 60 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 60 a aproximadamente aminoácido 70 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 70 a aproximadamente aminoácido 80 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 80 a aproximadamente aminoácido 90 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, polipéptidos que consisten en aproximadamente aminoácido 90 a aproximadamente aminoácido a 100 de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3, etc. a lo largo de la totalidad de una secuencia de aminoácidos de 238P1B2. Además, los polipéptidos que consisten en aproximadamente aminoácido 1 (o 20, 30 ó 40 etc.) a aproximadamente aminoácido 20 (o 130, 140 ó 150 etc.) de una proteína 238P1B2 mostrada en la Figura 2 o Figura 3. Debe

apreciarse que las posiciones de partida y parada en este párrafo se refieren a la posición especificada así como esa posición más o menos 5 restos.

Se generan proteínas relacionadas con 238P1B2 usando tecnología de síntesis peptídica convencional o usando procedimientos de escisión química bien conocidos en la técnica. Como alternativa, pueden usarse procedimientos recombinantes para generar moléculas de ácido nucleico que codifiquen una proteína relacionada con 238P1B2. Las moléculas de ácido nucleico pueden proporcionar un medio para generar fragmentos definidos de una proteína 238P1B2 (o variantes, homólogos o análogos de la misma).

III.A.) Proteínas que portan motivo

Los polipéptidos de 238P1B2 adicionales comprenden los restos aminoacídicos de uno o más de los motivos biológicos contenidos dentro de una secuencia polipeptídica de 238P1B2 expuesta en la Figura 2 o Figura 3. Se conocen en la técnica diversos motivos y puede evaluarse una proteína con respecto a la presencia de tales motivos por varios sitios de internet disponibles públicamente (véase, por ejemplo, las direcciones Web: pfam.wustl.edu/; searchlauncher.bcm.tmc.edu/seq-search/struc-predict.html; psort.ims.u-tokyo.ac.jp/; www.cbs.dtu.dk/; www.ebi.ac.uk/interpro/scan.html; www.expasy.ch/tools/scnpsiti.html; Epimatrix™ y Epimer™, Brown University, www.brown.edu/Research/TB-HIVLab/epimatrix-epimatrix.html; y BIMAS, bimas.dcrf.nih.gov/).

Diversas referencias reflejan la técnica con respecto a la identificación y generación de epítomos en una proteína de interés así como análogos de la misma. Véase, por ejemplo, documento WO 9733602 de Chesnut y col; Sette, *Immunogenetics* 1999 50 (3-4): 201-212; Sette y col., *J. Immunol.* 2001 166 (2): 1389-1397; Sidney y col., *Hum. Immunol.* 1997 58(1): 12-20; Kondo y col., *Immunogenetics* 1997 45(4): 249-258; Sidney y col., *J. Immunol.* 1996 157(8): 3480-90; y Falk y col., *Nature* 351: 290-6 (1991); Hunt y col., *Science* 255: 1261-3 (1992); Parker y col., *J. Immunol.* 149: 3580-7 (1992); Parker y col., *J. Immunol.* 152: 163-75 (1994); Kast y col., 1994 152 (8): 3904-12; Borrás-Cuesta y col., *Hum. Immunol.* 2000 61(3): 266-278; Alexander y col., *J. Immunol.* 2000 164(3): 1625-1633; Alexander y col., PMID: 7895164, UI: 95202582; O'Sullivan y col., *J. Immunol.* 1991 147(8): 2663-2669; Alexander y col., *Immunity* 1994 1(9): 751-761 y Alexander y col., *Immunol. Res.* 1998 18(2): 79-92.

Las proteínas relacionadas con 238P1B2 pueden estar en muchas formas, preferentemente en forma aislada. Una molécula proteica de 238P1B2 purificada estará sustancialmente sin otras proteínas o moléculas que perjudican la unión de 238P1B2 con anticuerpo, linfocito T u otro ligando. La naturaleza y grado de aislamiento y purificación dependerá del uso pretendido. Las proteínas relacionadas con 238P1B2 incluyen proteínas relacionadas con 238P1B2 purificadas y proteínas relacionadas con 238P1B2 solubles funcionales. Una proteína de 238P1B2 soluble funcional o fragmento de la misma puede conservar la capacidad de unirse a anticuerpo, linfocito T u otro ligando.

Las proteínas 238P1B2 pueden comprender fragmentos biológicamente activos de una secuencia de aminoácidos de 238P1B2 mostrada en la Figura 2 o Figura 3. Tales proteínas muestran propiedades de la proteína 238P1B2 de partida, tales como la capacidad para inducir la generación de anticuerpos que se unen específicamente a un epítomo asociado con la proteína 238P1B2 de partida; de unirse a tales anticuerpos; de inducir la activación de HTL o CTL; y/o de reconocerse por HTL o CTL que también se unen específicamente a la proteína de partida.

Los polipéptidos relacionados con 238P1B2 que contienen estructuras particularmente interesantes pueden predecirse y/o identificarse usando diversas técnicas analíticas bien conocidas en la materia, incluyendo, por ejemplo, los procedimientos de análisis de Chou - Fasman, Garnier-Robson, Kyte Doolittle, Eisenberg, Karplus Schultz o Jameson-Wolf o basándose en inmunogenicidad. Los fragmentos que contienen tales estructuras son particularmente útiles en la generación de anticuerpos anti-238P1B2 específicos de subunidad o linfocitos T o en la identificación de factores celulares que se unen a 238P1B2. Por ejemplo, pueden generarse perfiles de hidrofilia e identificarse fragmentos peptídicos inmunogénicos, usando el procedimiento de Hopp, T. P. y Woods, K. R., 1981, *Proc. Natl. Acad. Sci. U.S.A.* 78:3824-3828. Pueden generarse perfiles hidropáticos e identificarse fragmentos peptídicos inmunogénicos, usando el procedimiento de Kyte, J y Doolittle, R. F. 1982, *J. Mol Biol.* 157:105-132. Pueden generarse perfiles de porcentaje (%) de restos accesibles e identificarse fragmentos peptídicos inmunogénicos, usando el procedimiento de Janin J., 1979, *Nature* 277: 491-492. Pueden generarse perfiles de flexibilidad media e identificarse fragmentos peptídicos inmunogénicos, usando el procedimiento de Bhaskaran R., Ponnuswamy P. K., 1988, *Int. J. Pept. Protein Res.* 32:242-255. Pueden generarse perfiles de giro beta e identificarse fragmentos peptídicos inmunogénicos, usando el procedimiento de Deleage, G., Roux B., 1987, *Protein Engineering* 1: 289-294.

Pueden determinarse epítomos de CTL usando algoritmos específicos para identificar péptidos dentro de una proteína 238P1B2 que sean capaces de unirse óptimamente a alelos de HLA específicos (por ejemplo, usando el sitio SYFPEITHI en la dirección Web syfpeithi.bmi-heidelberg.com/; Epimatrix™ y Eimer™, Brown University, URL (www.brown.edu/Research/TB-HIV Lab/epimatrix/epimatrix.html); y BIMAS, URL bimas.dcrf.nih.gov/).

El algoritmo de búsqueda de motivo peptídico de HLA lo desarrolló el Dr. Ken Parker basándose en la unión de secuencias peptídicas específicas en el surco de las moléculas de HLA de clase I, en particular HLA-A2 (véase, por ejemplo, Falk y col., *Nature* 351: 290-6(1991); Hunt y col., *Science* 255:1261-3 (1992); Parker y col., *J. Immunol.* 149:3580-7 (1992); Parker y col., *J. Immunol.* 152:163-75 (1994)). Este algoritmo permite la localización y la

clasificación de péptidos 8-mer, 9-mer y 10-mer a partir de una secuencia proteica completa para unión predicha a HLA-A2 así como numerosas otras moléculas de HLA de clase I. Muchos péptidos de unión a HLA de clase I son 8, 9, 10 u 11-mer. Por ejemplo, para HLA-A2 de clase I, los epítomos preferentemente contienen una leucina (L) o metionina (M) en la posición 2 y una valina (V) o leucina (L) en el extremo C terminal (véase, por ejemplo, Parker y col., J. Immunol. 149:3580-7 (1992)).

La puntuación de unión corresponde a la mitad de tiempo estimado de disociación de complejos que contienen el péptido a 37 °C a pH 6,5. Se predice que los péptidos con la mayor puntuación de unión serán los unidos de forma más estrecha a HLA de clase I en la superficie celular durante el período de tiempo mayor y de este modo representan las mejores dianas inmunogénicas para reconocimiento de linfocitos T.

La unión real de péptidos a un alelo de HLA puede evaluarse por estabilización de la expresión de HLA en la línea celular defectuosa para procesamiento de antígenos T2 (véase, por ejemplo, Xue y col., Prostate 30: 73-8 (1997) y Peshwa y col., Prostate 36: 129-38 (1998)). La inmunogenicidad de péptidos específicos puede evaluarse *in vitro* por estimulación de linfocitos T citotóxicos CD8+ (CTL) en presencia de células presentadoras de antígenos tales como células dendríticas.

Debe apreciarse que cada epítomo predicho por el sitio BIMAS, sitios Epimer™ y Epimatrix™ o especificado por los motivos de HLA de clase I o clase II disponibles en la técnica o que forman parte de la técnica tales como los que se exponen en la Tabla IV (o determinados usando el sitio web syfpeithi.bmi-heidelberg.com/ o BIMAS, bimas.dcrf.nih.gov/) deben “aplicarse” a una proteína 238P1B2 descrita en el presente documento. Como se usa en el presente contexto “aplicarse” significa que se evalúa una proteína 238P1B2, por ejemplo, visualmente o por procedimientos de hallazgo de patrones basados en ordenador, como se aprecia por los expertos en la materia relevante. Cada subsecuencia de una proteína 238P1B2 de 8, 9, 10 u 11 restos aminoacídicos que porta un motivo de HLA de clase I o una subsecuencia de 9 o más restos aminoacídicos que portan un motivo de HLA de clase II se describen en el presente documento.

III. B.) Expresión de proteínas relacionadas con 238P1B2

Como se describe en los ejemplos a continuación, 238P1B2 puede expresarse de forma conveniente en células (tales como células 293T) transfectadas con un vector de expresión disponible en el mercado tal como un vector de expresión dirigido por CMV que codifica 238P1B2 con un marcador MYC y 6Xhis C terminal (pcDNA3.1/mycHis, Invitrogen o Tag5, GenHunter Corporation, Nashville TN). El vector Tag5 proporciona una señal de secreción de IgGK que puede usarse para facilitar la producción de una proteína 238P1B2 secretada en células transfectadas. La 238P1B2 marcada con HIS secretada en el medio de cultivo puede purificarse, por ejemplo, usando una columna de níquel usando técnicas convencionales.

III.C.) Modificaciones de proteínas relacionadas con 238P1B2

Se describen en el presente documento modificaciones de proteínas relacionadas con 238P1B2 tales como modificaciones covalentes. Un tipo de modificación covalente incluye hacer reaccionar restos aminoacídicos diana de un polipéptido 238P1B2 con un agente de derivatización orgánico que es capaz de reaccionar con cadenas laterales seleccionada o los restos N o C terminal de una proteína 238P1B2. Otro tipo de modificación covalente de un polipéptido de 238P1B2 descrito en el presente documento comprende alterar el patrón de glucosilación nativo de una proteína descrita en el presente documento. Otro tipo de modificación covalente de 238P1B2 comprende ligar un polipéptido de 238P1B2 con uno o diversos polímeros no proteínicos, por ejemplo, polietilenglicol (PEG), polipropilenglicol o polioxialquilenos, de la manera expuesta en las Patentes de Estados Unidos N° 4.640.835; 4.496.689; 4.301.144; 4.670.417; 4.791.192 o 4.179.337.

Las proteínas relacionadas con 238P1B2 descritas en el presente documento también pueden modificarse para formar una molécula quimérica que comprende 238P1B2 fusionada con otra secuencia polipeptídica o de aminoácidos heteróloga. Una molécula quimérica tal puede sintetizarse químicamente o de forma recombinante. Una molécula quimérica puede tener una proteína descrita en el presente documento fusionada con otro antígeno asociado a tumor o fragmento del mismo. Como alternativa, una proteína descrita en el presente documento puede comprender una fusión de fragmentos de una secuencia de 238P1B2 (de aminoácidos o ácidos nucleicos) de modo que se crea una molécula que no es, en toda su extensión, directamente homóloga a las secuencias de aminoácidos o ácidos nucleicos mostradas en la Figura 2 o la Figura 3. Una molécula quimérica tal puede comprender múltiples de la misma subsecuencia de 238P1B2. Una molécula quimérica puede comprender una fusión de una proteína relacionada con 238P1B2 con un marcador epitópico de polihistidina, que proporciona un epítomo con el que puede unirse de forma selectiva el níquel inmovilizado, con citocinas o con factores de crecimiento. El marcador epitópico se coloca generalmente en el extremo carboxilo o amino terminal de una proteína 238P1B2. Como alternativa, la molécula quimérica puede comprender una fusión de una proteína relacionada con 238P1B2 con una inmunoglobulina o una región particular de una inmunoglobulina. Para una forma bivalente de la molécula quimérica (también denominada “inmunoadhesina”), una fusión tal podría ser con la región Fc de una molécula de IgG. Las fusiones de Ig preferentemente incluyen la sustitución de una forma soluble (dominio transmembrana delecionado o inactivado) de un polipéptido de 238P1B2 en lugar de al menos una región variable dentro de una molécula de Ig. La fusión de inmunoglobulina puede incluir las regiones bisagra, CH2 y CH3 o la bisagra, CH1, CH2 y CH3 de una

molécula de IgG1. Para la producción de fusiones de inmunoglobulina véase, por ejemplo, Patente de Estados Unidos N° 5.428.130 presentada el 27 de junio de 1995.

III.D.) Usos de proteínas relacionadas con 238P1B2

5 Las proteínas descritas en el presente documento tienen varios usos específicos diferentes. Como 238P1B2 está altamente expresada en cánceres de próstata y otros, se usan proteínas relacionadas con 238P1B2 en procedimientos que evalúan el estado de productos génicos de 238P1B2 en tejidos cancerosos frente a normales, elucidando de este modo el fenotipo maligno. Normalmente, se usan polipéptidos de las regiones específicas de una proteína 238P1B2 para evaluar la presencia de perturbaciones (tales como deleciones, inserciones, mutaciones puntuales, etc.) en esas regiones (tales como regiones que contienen uno o más motivos). Los ensayos ejemplares 10 usan anticuerpos o linfocitos T que se dirigen a proteínas relacionadas con 238P1B2 que comprenden los restos aminoacídicos de uno o más de los motivos biológicos contenidos dentro de una secuencia polipeptídica de 238P1B2 para evaluar las características de esta región en tejidos normales frente a cancerosos o para inducir una respuesta inmune al epítipo. Como alternativa, se usan proteínas relacionadas con 238P1B2 que contienen los restos aminoacídicos de uno o más de los motivos biológicos en una proteína 238P1B2 para explorar con respecto a 15 factores que interaccionen con esa región de 238P1B2.

Los fragmentos proteicos/subsecuencias de 238P1B2 son particularmente útiles en la generación y caracterización de anticuerpos específicos de dominio (por ejemplo, anticuerpos que reconocen un epítipo extracelular o intracelular de una proteína 238P1B2), para identificar agentes o factores celulares que se unen a 238P1B2 o un dominio 20 estructural particular de los mismos y en diversos contextos terapéuticos y de diagnóstico, incluyendo pero sin limitación ensayos de diagnóstico, vacunas de cáncer y procedimientos para preparar tales vacunas.

Las proteínas codificadas por los genes de 238P1B2, o por análogos, homólogos o fragmentos de los mismos, tienen diversos usos, incluyendo pero sin limitación generar anticuerpos y en procedimientos para identificar 25 ligandos y otros agentes y constituyentes celulares que se unen a un producto génico de 238P1B2. Los anticuerpos inducidos contra una proteína 238P1B2 o fragmento de la misma son útiles en ensayos de diagnóstico y pronóstico, y metodologías de formación de imágenes en el tratamiento de cánceres humanos caracterizados por la expresión de proteína 238P1B2, tales como los enumerados en la Tabla I. Tales anticuerpos pueden expresarse de forma intracelular y usarse en procedimientos para tratar pacientes con tales cánceres. También se usan ácidos nucleicos o proteínas relacionadas con 238P1B2 en la generación de respuestas de HTL o CTL.

Se usan diversos ensayos inmunológicos útiles para la detección de proteínas 238P1B2, incluyendo pero sin limitación diversos tipos de radioinmunoensayos, ensayos inmunoabsorbentes ligados a enzima (ELISA), ensayos 30 inmunofluorescentes ligados a enzima (ELISA), procedimientos inmunocitoquímicos y similares. Los anticuerpos pueden marcarse y usarse como reactivos de formación de imágenes inmunológicos capaces de detectar células que expresan 238P1B2 (por ejemplo, en procedimientos de formación de imágenes radioescintigráficas). Las proteínas 238P1B2 también son particularmente útiles en la generación de vacunas de cáncer, como se describe 35 adicionalmente en el presente documento.

IV.) Procedimientos para la detección de 238P1B2

Los inventores describen procedimientos para detectar polinucleótidos de 238P1B2 y proteínas relacionadas con 238P1B2, así como procedimientos para identificar una célula que expresa 238P1B2. El perfil de expresión de 40 238P1B2 lo hace un marcador de diagnóstico para enfermedad metastasizada. En consecuencia, el estado de los productos génicos de 238P1B2 proporciona información útil para predecir diversos factores incluyendo susceptibilidad a enfermedad de etapa avanzada, tasa de progresión y/o agresividad del tumor. Como se analiza en detalle en el presente documento, el estado de los productos génicos de 238P1B2 en muestras de pacientes puede analizarse mediante diversos protocolos que se conocen bien en la técnica incluyendo análisis inmunohistoquímico, la diversidad de técnicas de transferencia de Northern incluyendo hibridación *in situ*, análisis por RT-PCR (por 45 ejemplo en muestras microdisecadas por captura de láser) análisis de transferencia de Western y análisis de serie tisular.

De forma más particular, los inventores describen ensayos para la detección de polinucleótidos de 238P1B2 en una muestra biológica, tales como suero, hueso, próstata y otros tejidos, orina, semen, preparaciones celulares y 50 similares. Los polinucleótidos de 238P1B2 detectables incluyen, por ejemplo, un gen de 238P1B2 o fragmento del mismo, ARNm de 238P1B2, ARNm de 238P1B2 de variante de corte y empalme alternativo y moléculas de ADN o ARN recombinantes que contienen un polinucleótido de 238P1B2. Se conocen bien en la técnica varios procedimientos para amplificar y/o detectar la presencia de polinucleótidos de 238P1B2 y pueden emplearse en la práctica de estos ensayos.

Un procedimiento para detectar un ARNm de 238P1B2 en una muestra biológica puede comprender producir ADNc 55 de la muestra por transcripción inversa usando al menos un cebador; amplificar el ADNc producido de este modo usando polinucleótidos de 238P1B2 como cebadores sentido y antisentido para amplificar ADNc de 238P1B2 en el mismo; y detectar la presencia del ADNc de 238P1B2 amplificado. Opcionalmente, puede determinarse la secuencia del ADNc de 238P1B2 amplificado.

Un procedimiento para detectar un gen de 238P1B2 en una muestra biológica puede comprender primero aislar ADN genómico de la muestra; amplificar el ADN genómico aislado usando polinucleótidos de 238P1B2 como cebadores sentido y antisentido; y detectar la presencia del gen de 238P1B2 amplificado. Puede designarse cualquier variedad de combinaciones de sonda sentido y antisentido apropiadas a partir de una secuencia de nucleótidos de 238P1B2 (véase, por ejemplo, Figura 2) y usarse para este fin.

Los inventores también describen ensayos para detectar la presencia de una proteína 238P1B2 en un tejido u otra muestra biológica tal como suero, semen, hueso, próstata, orina, preparaciones celulares y similares. Los procedimientos para detectar una proteína relacionada con 238P1B2 también se conocen bien e incluyen, por ejemplo, inmunoprecipitación, análisis inmunohistoquímico, análisis de transferencia de Western, ensayos de unión molecular, ELISA, ELIFA y similares. Por ejemplo, un procedimiento para detectar la presencia de una proteína relacionada con 238P1B2 en una muestra biológica comprende primero poner en contacto con la muestra con un anticuerpo de 238P1B2, un fragmento sensible a 238P1B2 del mismo o una proteína recombinante que contiene una región de unión a antígeno de un anticuerpo de 238P1B2; y después detectar la unión de la proteína relacionada con 238P1B2 en la muestra.

También se describen procedimientos para identificar una célula que expresa 238P1B2. Un ensayo para identificar una célula que exprese un gen de 238P1B2 puede comprender detectar la presencia de ARNm de 238P1B2 en la célula. Se conocen bien procedimientos para la detección de ARNm particulares en células e incluyen, por ejemplo, ensayos de hibridación usando sondas de ADN complementario (tales como hibridación *in situ* usando ribosondas de 238P1B2 marcadas, transferencia de Northern y técnicas relacionadas) y diversos ensayos de amplificación de ácido nucleico (tales como RT-PCR usando cebadores complementarios específicos para 238P1B2 y otros procedimientos de detección de tipo amplificación, tales como, por ejemplo, ADN ramificado, SISBA, TMA y similares). Como alternativa, un ensayo para identificar una célula que expresa un gen de 238P1B2 comprende detectar la presencia de proteínas relacionadas con 238P1B2 en la célula o secretada por la célula. Se conocen bien en la técnica diversos procedimientos para la detección de proteínas y se emplean para la detección de proteínas relacionadas con 238P1B2 y células que expresan proteínas relacionadas con 238P1B2.

El análisis de expresión de 238P1B2 también es útil como una herramienta para identificar y evaluar agentes que modulan la expresión génica de 238P1B2. Por ejemplo, la expresión de 238P1B2 se regula positivamente de forma significativa en cáncer de próstata y se expresa en cánceres de los tejidos enumerados en la Tabla 1. La identificación de una molécula o agente biológico que inhibe expresión de 238P1B2 o sobreexpresión en células cancerosas es de valor terapéutico. Por ejemplo, un agente tal puede identificarse usando una exploración que cuantifica la expresión de 238P1B2 por RT-PCR, hibridación de ácido nucleico o unión de anticuerpo.

V.) Procedimientos para controlar el estado de genes relacionados con 238P1B2 y sus productos

Se sabe que la oncogénesis es un proceso multietapa en el que el crecimiento celular se desregula progresivamente y las células progresan de un estado fisiológico normal a estados precancerosos y después cancerosos (véase, por ejemplo, Alers y col., Lab Invest. 77(5):437-438 (1997) e Isaacs y col., Cancer Surv. 23:19-32 (1995)). En este contexto, la examinación de una muestra biológica con respecto a pruebas de crecimiento celular desregulado (tal como expresión de 238P1B2 aberrante en cánceres) permite la detección temprana de dicha fisiología aberrante, antes de que un estado patológico tal como cáncer haya progresado a una etapa en la que las opciones terapéuticas son más limitadas y/o el pronóstico es peor. En tales examinaciones, el estado de 238P1B2 en una muestra biológica de interés puede compararse, por ejemplo, con el estado de 238P1B2 en una muestra normal correspondiente (por ejemplo una muestra de ese individuo o como alternativa otro individuo que no está afectado por una patología). Una alteración en el estado de 238P1B2 en la muestra biológica (en comparación con la muestra normal) proporciona pruebas de crecimiento celular desregulado. Además de usar una muestra biológica que no está afectada por una patología como una muestra normal, puede usarse también un valor normativo predeterminado tal como un nivel normal predeterminado de expresión de ARNm (véase, por ejemplo, Grever y col., J. Comp. Neurol. 9 de diciembre de 1996; 376(2):306-14 y Patente de Estados Unidos N° 5.837.501) para comparar el estado de 238P1B2 en una muestra.

El término "estado" en este contexto se usa de acuerdo con su significado aceptado en la técnica y se refiere a la condición o el estado de un gen y sus productos. Normalmente, los expertos en la materia usan varios parámetros para evaluar la condición o estado de un gen y sus productos. Estos incluyen, pero sin limitación la localización de productos génicos expresados (incluyendo la localización de células que expresan 238P1B2) así como el nivel y actividad biológica de productos génicos expresados (tales como ARNm de 238P1B2, polinucleótidos y polipéptidos). Normalmente, una alteración en el estado de 238P1B2 comprende un cambio en la localización de 238P1B2 y/o células que expresan 238P1B2 y/o un aumento en la expresión de ARNm y/o proteína de 238P1B2.

El estado de 238P1B2 en una muestra puede analizarse por varios medios bien conocidos en la técnica, incluyendo sin limitación, análisis inmunohistoquímico, hibridación *in situ*, análisis por RT-PCR en muestras microdisecadas de captura de láser, análisis de transferencia de Western y análisis de serie tisular. Los protocolos típicos para evaluar el estado de un gen de 238P1B2 y productos génicos se encuentran, por ejemplo en Ausubel y col., eds. 1995, Current Protocols In Molecular Biology, unidades 2 (transferencia de Northern), 4 (transferencia de Southern), 15 (inmunotransferencia) y 18 (análisis de PCR). Por lo tanto, el estado de 238P1B2 en una muestra biológica se evalúa

5 por diversos procedimientos usados por los expertos en la materia incluyendo, pero sin limitación análisis de Southern genómico (para examinar, por ejemplo perturbaciones en un gen de 238P1B2), análisis de Northern y/o análisis por PCR de ARNm de 238P1B2 (para examinar, por ejemplo alteraciones en las secuencias polinucleotídicas o niveles de expresión de ARNm de 238P1B2) y, análisis de Western y/o inmunohistoquímicos (para examinar, por ejemplo alteraciones de secuencias polipeptídicas, alteraciones de la localización polipeptídica dentro de una muestra, alteraciones en los niveles de expresión de proteínas 238P1B2 y/o asociaciones de proteínas 238P1B2 con compañeros de unión polipeptídicos). Los polinucleótidos de 238P1B2 detectables incluyen, por ejemplo, un gen de 238P1B2 o fragmento del mismo, ARNm de 238P1B2, variantes de corte y empalme alternativo, ARNm de 238P1B2 y moléculas de ADN o ARN recombinantes que contienen un polinucleótido de 238P1B2.

10 El perfil de expresión de 238P1B2 le hace un marcador de diagnóstico para enfermedad local y/o metastasizada y proporciona información sobre el crecimiento o potencial oncogénico de una muestra biológica. En particular, el estado de 238P1B2 proporciona información útil para predecir la susceptibilidad a etapas de enfermedad particulares, progresión y/o agresividad del tumor. Los inventores describen procedimientos y ensayos para determinar el estado de 238P1B2 y diagnosticar cánceres que expresan 238P1B2, tales como cánceres de los tejidos enumerados en la Tabla I. Por ejemplo, debido a que el ARNm de 238P1B2 está tan altamente expresado en cánceres de próstata y otros en relación con tejido prostático normal, los ensayos que evalúan los niveles de transcritos de ARNm de 238P1B2 o proteínas en una muestra biológica pueden usarse para diagnosticar una enfermedad asociada con desregulación de 238P1B2 y pueden proporcionar información de pronóstico útil en la definición de opciones terapéuticas apropiadas.

15 El estado de expresión de 238P1B2 proporciona información incluyendo la presencia, la etapa y la localización de células precancerosas y cancerosas displásicas, prediciendo la susceptibilidad a diversas etapas de enfermedad y/o para estimar la agresividad del tumor. Además, el perfil de expresión lo hace útil como un reactivo de formación de imágenes para enfermedad metastasizada. En consecuencia, los inventores describen diversos procedimientos de pronóstico y diagnóstico moleculares para examinar el estado de 238P1B2 en muestras biológicas tales como las de individuos que padecen, o se sospecha que padecen una patología caracterizada por crecimiento celular desregulado, tal como cáncer.

20 Como se ha descrito anteriormente, el estado de 238P1B2 en una muestra biológica puede examinarse por varios procedimientos bien conocidos en la técnica. Por ejemplo, el estado de 238P1B2 en una muestra biológica tomada de una localización específica en el cuerpo puede examinarse evaluando la muestra con respecto a la presencia o ausencia de células que expresan 238P1B2 (por ejemplo, las que expresan ARNm o proteínas de 238P1B2). Esta examinación puede proporcionar pruebas de crecimiento celular desregulado, por ejemplo, cuando se encuentran células que expresan 238P1B2 en una muestra biológica que no contiene normalmente tales células (tal como un ganglio linfático), debido a que tales alteraciones en el estado de 238P1B2 en una muestra biológica se asocian con frecuencia con crecimiento celular desregulado. Específicamente, un indicador del crecimiento celular desregulado es las metástasis de células cancerosas de un órgano de origen (tal como la próstata) a un área diferente del cuerpo (tal como un ganglio linfático). En este contexto, las pruebas de crecimiento celular desregulado son importantes por ejemplo debido a que pueden detectarse metástasis de ganglios linfáticos ocultas en una proporción sustancial de pacientes con cáncer de próstata y tales metástasis se asocian con predictores conocidos de progresión de enfermedad (véase, por ejemplo, Murphy y col., Prostate 42 (4): 315-317 (2000); Su y col., Semin. Surg. Oncol 18(1):17-28 (2000) y Freeman y col., J Urol Ago 1995 154 (2 Pt 1): 474-8).

25 Los inventores describen procedimientos para controlar los productos génicos de 238P1B2 determinando el estado de los productos génicos de 238P1B2 expresados por células de un individuo sospechoso de tener una enfermedad asociada con crecimiento celular desregulado (tal como hiperplasia o cáncer) y comparando después el estado determinado de este modo con el estado de los productos génicos de 238P1B2 en una muestra normal correspondiente. La presencia de productos génicos de 238P1B2 aberrantes en la muestra de ensayo en relación con la muestra normal proporciona un indicio de la presencia de crecimiento celular desregulado dentro de las células del individuo.

30 Los inventores describen ensayos útiles en la determinación de la presencia de cáncer en un individuo, que comprenden detectar un aumento significativo de expresión de ARNm o proteínas de 238P1B2 en una muestra celular o tisular de ensayo en relación con niveles de expresión en el tejido o célula normal correspondiente. La presencia de ARNm de 238P1B2 puede, por ejemplo, evaluarse en muestras tisulares incluyendo pero sin limitación las enumeradas en la Tabla I. La presencia de expresión de 238P1B2 significativa en cualquiera de estos tejidos es útil para indicar la aparición, presencia y/o gravedad de un cáncer, puesto que los tejidos normales correspondientes no expresan ARNm de 238P1B2 o lo expresan a niveles más bajos.

35 El estado de 238P1B2 puede determinarse en el nivel proteico en lugar de en el nivel de ácido nucleico. Por ejemplo, un procedimiento tal comprende determinar el nivel de proteína de 238P1B2 expresada por células en una muestra tisular de ensayo y comparar el nivel determinado de este modo con el nivel de 238P1B2 expresado en una muestra normal correspondiente. La presencia de proteína 238P1B2 puede evaluarse, por ejemplo, usando procedimientos inmunohistoquímicos. Los anticuerpos de 238P1B2 o compañeros de unión capaces de detectar expresión de proteína 238P1B2 se usan en diversos formatos de ensayo bien conocidos en la técnica para este fin.

Puede evaluarse el estado de las secuencias de nucleótidos y aminoácidos de 238P1B2 en una muestra biológica para identificar perturbaciones en la estructura de estas moléculas. Estas perturbaciones pueden incluir inserciones, deleciones, sustituciones y similares. Tales evaluaciones son útiles porque las perturbaciones en las secuencias de nucleótidos y aminoácidos se observan en un gran número de proteínas asociadas con un fenotipo desregulado del crecimiento (véase, por ejemplo, Marrogi y col., 1999, J. Cutan. Pathol. 26 (8): 369-378). Por ejemplo, una mutación en la secuencia de 238P1B2 puede ser indicativa de la presencia o promoción de un tumor. Tales ensayos por lo tanto tienen valor diagnóstico y predictivo en el que una mutación de 238P1B2 indica una potencial pérdida de función o aumento de crecimiento tumoral.

Se conoce bien en la técnica una amplia diversidad de ensayos para observar perturbaciones en secuencias de nucleótidos y aminoácidos. Por ejemplo, el tamaño y estructura de las secuencias de ácido nucleico o aminoácidos de los productos génicos de 238P1B2 se observan por los protocolos de Northern, Southern, Western, PCR y secuenciación de ADN analizados en el presente documento. Además, se conocen bien en la técnica otros procedimientos para observar perturbaciones en secuencias de nucleótidos y aminoácidos tales como análisis de polimorfismo de conformación de cadena sencilla (véase, por ejemplo, Patente de Estados Unidos N° 5.382.510 presentada el 7 de septiembre de 1999 y 5.952.170 presentada el 17 de enero de 1995).

Adicionalmente, se puede examinar el estado de metilación de un gen de 238P1B2 en una muestra biológica. Con frecuencia se produce desmetilación aberrante y/o hipermetilación de islas de CpG en regiones reguladoras 5' del gen en células inmortalizadas y transformadas y puede dar como resultado expresión alterada de diversos genes. Por ejemplo, la hipermetilación del promotor de la glutatión S-transferasa de clase pi (una proteína expresada en próstata normal pero no expresada en > del 90 % de carcinomas de próstata) parece silenciar permanentemente la transcripción de este gen y es la alteración genómica más frecuentemente detectada en carcinomas de próstata (De Marzo y col., Am. J. Pathol. 155 (6): 1985-1992 (1999)). Además, esta alteración está presente en al menos el 70 % de los casos de neoplasia intraepitelial prostática de alto grado (PIN) (Brooks y col., Cancer Epidemiol. Biomarkers Prev., 1998, 7:531-536). En otro ejemplo, la expresión del gen específico de tumor LAGE-I (que no se expresa en próstata normal pero se expresa en 25-50 % de los cánceres de próstata) se induce por desoxi-azacitidina en células linfoblastoides, lo que sugiere que la expresión tumoral se debe a desmetilación (Lethe y col., Int J. Cancer 76 (6): 903-908 (1998)). Se conocen bien en la técnica diversos ensayos para examinar el estado de metilación de un gen. Por ejemplo, puede usarse, en enfoques de hibridación de Southern, enzimas de restricción sensibles a metilación que no pueden escindir secuencias que contienen sitios de CpG metilados para evaluar el estado de metilación de las islas de CpG. Además, MSP (PCR específica de metilación) puede perfilar rápidamente el estado de metilación de todos los sitios de CpG presentes en una isla de CpG de un gen dado. Este procedimiento implica modificación inicial de ADN por bisulfito sódico (que convertirá todas las citosinas no metiladas a uracilo) seguido de amplificación usando cebadores específicos para ADN metilado frente a no metilado. También pueden hallarse protocolos que implican interferencia de metilación por ejemplo en Current Protocols in Molecular Biology, unidad 12, Frederick M. Ausubel y col eds., 1995.

La amplificación génica es un procedimiento adicional para evaluar el estado de 238P1B2. La amplificación génica se mide en una muestra directamente, por ejemplo, por transferencia de Southern o transferencia de Northern convencional para cuantificar la transcripción de ARNm (Thomas, 1980, Proc. Natl. Acad. Sci. USA, 77:5201-5205), transferencia puntual (análisis de ADN), o hibridación *in situ*, usando una sonda marcada de forma apropiada, basándose en las secuencias proporcionadas en el presente documento. Como alternativa, se emplean anticuerpos que reconocen dobles cadenas específicas, incluyendo dobles cadenas de ADN, dobles cadenas de ARN, y dobles cadenas híbridas de ADN-ARN o dobles cadenas de ADN-proteína. Los anticuerpos a su vez se marcan y el ensayo se lleva a cabo cuando la doble cadena está unida a una superficie, de modo que tras la formación de doble cadena de superficie, la presencia del anticuerpo unido a la doble cadena pueda detectarse.

Puede analizarse convenientemente tejido biopsiado o sangre periférica con respecto a la presencia de células cancerosas usando por ejemplo, análisis de Northern, de transferencia puntual o RT-PCR para detectar la expresión de 238P1B2. La presencia de ARNm de 238P1B2 amplificable por RT-PCR proporciona un indicio de la presencia de cáncer. Se conocen bien en la técnica ensayos de RT-PCR. En la actualidad se están evaluando los ensayos de detección por RT-PCR para células tumorales en sangre periférica para su uso en el diagnóstico y tratamiento de varios tumores sólidos humanos. En el campo del cáncer de próstata, estos incluyen ensayos de RT-PCR para la detección de células que expresan PSA y PSM (Verkaik y col., 1997, Urol. Res. 25:373-384; Ghossein y col., 1995, J. Clin. Oncol. 13:1195-2000; Heston y col., 1995, Clin. Chem. 41:1687-1688).

Los inventores describen una evaluación de la susceptibilidad que un individuo tiene para desarrollar cáncer. Los inventores describen un procedimiento para predecir la susceptibilidad a cáncer que comprende detectar ARNm de 238P1B2 o proteína 238P1B2 en una muestra tisular, indicando su presencia susceptibilidad a cáncer, en el que el grado de expresión de ARNm de 238P1B2 se correlaciona con el grado de susceptibilidad. La presencia de 238P1B2 en próstata u otro tejido puede examinarse, proporcionando la presencia de 238P1B2 en la muestra un indicio de susceptibilidad a cáncer de próstata (o la aparición o existencia de un tumor de próstata). De forma similar, se puede evaluar la integridad de las secuencias de nucleótidos y aminoácidos de 238P1B2 en una muestra biológica para identificar perturbaciones en la estructura de estas moléculas tales como inserciones, deleciones, sustituciones y similares. La presencia de una o más perturbaciones de los productos génicos de 238P1B2 en la muestra es un indicio de susceptibilidad a cáncer (o la aparición o existencia de un tumor).

Los inventores describen procedimientos para estimar la agresividad del tumor. Un procedimiento para estimar la agresividad de un tumor puede comprender determinar el nivel de ARNm de 238P1B2 o proteína 238P1B2 expresada por células tumorales, comparando el nivel determinado de este modo con el nivel de ARNm de 238P1B2 o proteína 238P1B2 expresado en un tejido normal correspondiente tomado del mismo individuo o una muestra de referencia de tejido normal, en el que el grado de expresión de ARNm de 238P1B2 o proteína 238P1B2 en la muestra de tumor en relación con la muestra normal indica el grado de agresividad. La agresividad de un tumor puede evaluarse determinando el grado en que 238P1B2 se expresa en las células tumorales, indicando mayores niveles de expresión tumores más agresivos. Los inventores describen la evaluación de la integridad de secuencias de nucleótidos y aminoácidos de 238P1B2 en una muestra biológica, para identificar perturbaciones en la estructura de estas moléculas tales como inserciones, deleciones, sustituciones y similares. La presencia de una o más perturbaciones indica tumores más agresivos.

Los inventores describen procedimientos para observar la progresión de un tumor maligno en un individuo a lo largo del tiempo. Los procedimientos para observar la progresión de un tumor maligno en un individuo a lo largo del tiempo pueden comprender determinar el nivel de ARNm de 238P1B2 o proteína 238P1B2 expresado por células en una muestra del tumor, comparando el nivel determinado de este modo con el nivel de ARNm de 238P1B2 o proteína 238P1B2 expresado en una muestra tisular equivalente tomada del mismo individuo en un momento diferente, en el que el grado de expresión de ARNm de 238P1B2 o proteína 238P1B2 en la muestra de tumor a lo largo del tiempo proporciona información sobre la progresión del cáncer. La progresión de un cáncer puede evaluarse determinando la expresión de 238P1B2 en las células tumorales a lo largo del tiempo, indicando un aumento de la expresión a lo largo del tiempo una progresión del cáncer. También, se puede evaluar la integridad de las secuencias de nucleótidos y aminoácidos de 238P1B2 en una muestra biológica para identificar perturbaciones en la estructura de estas moléculas tales como inserciones, deleciones, sustituciones y similares, indicando la presencia de una o más perturbaciones una progresión del cáncer.

Los enfoques de diagnóstico anteriores pueden combinarse con uno cualquiera de una amplia diversidad de protocolos de pronóstico y diagnóstico conocidos en la técnica. Por ejemplo, los inventores describen procedimientos para observar una coincidencia entre la expresión del gen de 238P1B2 y los productos génicos de 238P1B2 (o perturbaciones en el gen de 238P1B2 y los productos génicos de 238P1B2) y un factor que está asociado con tumor maligno, como un medio para diagnosticar y pronosticar el estado de una muestra tisular. Puede usarse una amplia diversidad de factores asociados con tumor maligno, tales como la expresión de genes asociados con tumor maligno (por ejemplo, expresión de PSA, PSCA y PSM para cáncer de próstata etc.), así como observaciones citológicas globales (véase, por ejemplo, Bocking y col., 1984, Anal. Quant. Cytol. (6):74-88; Epstein, 1995, Hum. Pathol. (26):223-9; Thorson y col., 1998, Mod. Pathol. (11):6:543-51; Baisden y col., 1999, Am J. Surg. Pathol. 23(8):918-24). Los procedimientos para observar una coincidencia entre la expresión de gen de 238P1B2 y productos génicos de 238P1B2 (o perturbaciones en el gen de 238P1B2 y productos génicos de 238P1B2) y otro factor que se asocia con tumor maligno son útiles, por ejemplo, porque la presencia de un conjunto de factores específicos que coinciden con la enfermedad proporciona información crucial para diagnosticar y pronosticar el estado de una muestra tisular.

Los procedimientos para observar una coincidencia entre la expresión del gen de 238P1B2 y los productos génicos de 238P1B2 (o perturbaciones en el gen de 238P1B2 y productos génicos de 238P1B2) y otro factor asociado con tumor maligno pueden implicar detectar la sobreexpresión de ARNm o proteína de 238P1B2 en una muestra tisular, detectar la sobreexpresión de ARNm o proteína de PSA en una muestra tisular (o expresión de PSCA o PSM), y observar una coincidencia de sobreexpresión de ARNm o proteína de 238P1B2 y ARNm o proteína de PSA (o expresión de PSCA o PSM). La expresión de ARNm de 238P1B2 y PSA en tejido de próstata puede examinarse, indicando la coincidencia de sobreexpresión de ARNm de 238P1B2 y PSA en la muestra la existencia de cáncer de próstata, susceptibilidad a cáncer de próstata o la aparición o estado de un tumor de próstata.

Se describen en el presente documento procedimientos para detectar y cuantificar la expresión de ARNm o proteína de 238P1B2 y se conocen bien en la técnica tecnologías de detección y cuantificación de proteína y ácido nucleico convencionales. Los procedimientos convencionales para la detección y cuantificación de ARNm de 238P1B2 incluyen hibridación *in situ* usando ribosondas de 238P1B2 marcadas, transferencia de Northern y técnicas relacionadas usando sondas polinucleotídicas de 238P1B2, análisis por RT-PCR usando cebadores específicos para 238P1B2 y otros procedimientos de detección de tipo amplificación, tales como, por ejemplo, ADN ramificado, SISBA, TMA y similares. Puede usarse RT-PCR semicuantitativa para detectar y cuantificar la expresión de ARNm de 238P1B2. Puede usarse cualquier variedad de cebadores capaces de amplificar 238P1B2 para este fin, incluyendo pero sin limitación los diversos conjuntos de cebadores específicamente descritos en el presente documento. Pueden usarse anticuerpos policlonales o monoclonales que pueden ser específicamente reactivos con la proteína 238P1B2 de tipo silvestre en un ensayo inmunohistoquímico de tejido biopsiado.

VI.) Identificación de moléculas que interactúan con 238P1B2

Las secuencias proteica y de ácido nucleico de 238P1B2 desveladas en el presente documento permiten que un experto en la materia identifique proteínas, moléculas pequeñas y otros agentes que interactúan con 238P1B2, así como rutas activadas por 238P1B2 mediante uno cualquiera de los diversos protocolos aceptados en la técnica. Por ejemplo, se puede usar uno de los llamados sistemas de trampa de interacción (también denominado "ensayo de

dos híbridos”). En tales sistemas, las moléculas interaccionan y reconstituyen un factor de transcripción que dirige la expresión de un gen indicador, tras lo cual se ensaya la expresión del gen indicador. Otros sistemas identifican interacciones proteína-proteína *in vivo* a través de reconstitución de un activador transcripcional eucariota, véase, por ejemplo, Patente de Estados Unidos N° 5.955.280 presentada el 21 de septiembre de 1999, 5.925.523 presentada el 20 de julio de 1999, 5.846.722 presentada el 8 de diciembre de 1998 y 6.004.746 presentada el 21 de diciembre de 1999. También están disponibles en la técnica algoritmos para predicciones basadas en genoma de la función proteica (véase, por ejemplo, Marcotte, y col., Nature 402:4 noviembre de 1999, 83-86).

Como alternativa se pueden explorar bibliotecas peptídicas para identificar moléculas que interaccionan con secuencias proteicas de 238P1B2. En tales procedimientos, los péptidos que se unen a 238P1B2 se identifican explorando bibliotecas que codifican una colección controlada o aleatoria de aminoácidos. Los péptidos codificados por las bibliotecas se expresan como proteínas de fusión de proteínas de revestimiento bacteriófago, las partículas de bacteriófago se exploran después frente a la proteína o proteínas 238P1B2.

En consecuencia, se identifican de este modo péptidos que tienen una amplia diversidad de usos, tales como reactivos terapéuticos, de pronóstico o de diagnóstico, sin ninguna información anterior sobre la estructura del ligando o molécula receptora esperados. Se desvelan bibliotecas peptídicas y procedimientos de exploración típicos que pueden usarse para identificar moléculas que interaccionan con secuencias proteicas de 238P1B2 por ejemplo en las Patentes de Estados Unidos N° 5.723.286 presentada el 3 de marzo de 1998 y 5.733.731 presentada el 31 de marzo de 1998.

Como alternativa, se usan líneas celulares que expresan 238P1B2 para identificar interacciones proteína-proteína mediadas por 238P1B2. Tales interacciones pueden examinarse usando técnicas de inmunoprecipitación (véase, por ejemplo, Hamilton B.J., y col. Biochem. Biophys. Res. Commun. 1999, 261: 646-51). La proteína 238P1B2 puede inmunoprecipitarse de líneas celulares que expresan 238P1B2 usando anticuerpos anti-238P1B2. Como alternativa, pueden usarse anticuerpos contra marcador His en una línea celular modificada por ingeniería genética para que exprese fusiones de 238P1B2 y un marcador His (vectores mencionados anteriormente). El complejo inmunoprecipitado puede examinarse con respecto a asociación proteica por procedimientos tales como transferencia de Western, marcaje con ³⁵S-metionina de proteínas, microsecuenciación de proteínas, tinción con plata y electroforesis en gel de dos dimensiones.

Pueden identificarse moléculas pequeñas y ligandos que interaccionan con 238P1B2 a través de ensayos de exploración relacionados. Por ejemplo, pueden identificarse moléculas pequeñas que interfieren con la función proteica, incluyendo moléculas que interfieren con la capacidad de 238P1B2 para mediar en la fosforilación y desfosforilación, interacción con moléculas de ADN o ARN como un indicio de regulación de ciclos celulares, señalización de segundo mensajero o tumorigénesis. De forma similar, se identifican moléculas pequeñas que modulan el canal de iones relacionado con 238P1B2, bomba proteica o funciones de comunicación celular y se usan para tratar pacientes que tengan un cáncer que exprese 238P1B2 (véase, por ejemplo, Hille, B., Ionic Channels of Excitable Membranes 2ª Ed., Sinauer Assoc., Sunderland, MA, 1992). Además, pueden identificarse ligandos que regulan la función de 238P1B2 basándose en su capacidad para unirse a 238P1B2 y activar una construcción indicadora. Se analizan procedimientos típicos por ejemplo en la Patente de Estados Unidos N° 5.928.868 presentada el 27 de julio de 1999 e incluyen procedimientos para formar ligandos híbridos en los que al menos un ligando es una molécula pequeña. En un procedimiento ilustrativo, las células modificadas por ingeniería genética para expresar un proteína de fusión de 238P1B2 y una proteína de unión a ADN se usan para co-expresar una proteína de fusión de un ligando híbrido/molécula pequeña y una proteína activadora de la transcripción de biblioteca de ADNc. Las células contienen adicionalmente un gen indicador, la expresión del cual está condicionada por la proximidad de la primera y segunda proteínas de fusión entre sí, un acontecimiento que se produce solamente si el ligando híbrido se une a sitios diana en ambas proteínas híbridas. Las células que expresan el gen indicador se seleccionan y se identifica la molécula pequeña desconocida o el ligando desconocido. Este procedimiento proporciona un medio para identificar moduladores que activan o inhiben 238P1B2.

Los inventores describen un procedimiento para explorar con respecto a una molécula que interacciona con una secuencia de aminoácidos de 238P1B2 mostrada en la Figura 2 o Figura 3, que comprende las etapas de poner en contacto una población de moléculas con una secuencia de aminoácidos de 238P1B2, permitir que la población de moléculas y la secuencia de aminoácidos de 238P1B2 interaccionen en condiciones que facilitan una interacción, determinar la presencia de una molécula que interacciona con la secuencia de aminoácidos de 238P1B2 y después separar las moléculas que no interaccionan con la secuencia de aminoácidos de 238P1B2 de las moléculas que sí. El procedimiento puede comprender adicionalmente purificar, caracterizar e identificar una molécula que interacciona con la secuencia de aminoácidos de 238P1B2. La molécula identificada puede usarse para modular una función realizada por 238P1B2. La secuencia de aminoácidos de 238P1B2 puede ponerse en contacto con una biblioteca de péptidos.

VII.) Procedimientos de diagnóstico y pronóstico

Como se desvela en el presente documento, pueden usarse polinucleótidos, polipéptidos de 238P1B2, linfocitos T citotóxicos reactivos (CTL), linfocitos T auxiliares reactivos (HTL) y anticuerpos anti-polipéptido en ensayos de diagnóstico, pronóstico y terapéuticos bien conocidos que examinan las afecciones asociadas con crecimiento

celular desregulado tales como cáncer, en particular los cánceres enumerados en la Tabla I (véase, por ejemplo, tanto su patrón específico de expresión tisular como su sobreexpresión en ciertos cánceres como se describe por ejemplo en el Ejemplo 4).

Puede hacerse una analogía de 238P1B2 con un antígeno asociado a próstata PSA, el marcador arquetípico que se ha usado por los facultativos médicos durante años para identificar y controlar la presencia de cáncer de próstata (véase, por ejemplo, Merrill y col., *J. Urol.* 163(2): 503-5120 (2000); Polascik y col., *J. Urol.* Aug; 162(2): 293-306 (1999) y Fortier y col., *J. Nat. Cancer Inst.* 91(19): 1635-1640(1999)). También se usan diversos otros marcadores de diagnóstico en contextos similares incluyendo p53 y K-ras (véase, por ejemplo, Tulchinsky y col., *Int J Mol Med* 1999 Jul 4(1): 99-102 y Minimoto y col., *Cancer Detect Prev* 2000; 24(1): 1-12). Por lo tanto, la presente divulgación de polinucleótidos y polipéptidos de 238P1B2 (así como sondas polinucleotídicas de 238P1B2 y anticuerpos anti-238P1B2 usados para identificar la presencia de estas moléculas) y sus propiedades permite que los expertos en la materia utilicen estas moléculas en procedimientos que son análogos a los usados, por ejemplo, en diversos ensayos de diagnóstico dirigidos a examinar las condiciones asociadas con cáncer.

Los procedimientos de diagnóstico típicos que usan los polinucleótidos, polipéptidos linfocitos T sensibles a y anticuerpos de 238P1B2, son análogos a los procedimientos de ensayos de diagnóstico bien establecidos que emplean, por ejemplo, polinucleótidos, polipéptidos, linfocitos T sensibles a y anticuerpos de PSA. Por ejemplo, igual que los polinucleótidos de PSA se usan como sondas (por ejemplo en análisis de Northern, véase, por ejemplo, Sharief y col., *Biochem. Mol. Biol. Int.* 33(3): 567-74(1994)) y cebadores (por ejemplo en análisis de PCR, véase, por ejemplo, Okegawa y col., *J. Urol.* 163(4): 1189-1190 (2000)) para observar la presencia y/o el nivel de ARNm de PSA en procedimientos de control de la sobreexpresión de PSA o la metástasis de cánceres de próstata, los polinucleótidos de 238P1B2 descritos en el presente documento pueden usarse de la misma manera para detectar sobreexpresión de 238P1B2 o la metástasis de próstata y otros cánceres que expresan este gen. Como alternativa, igual que los polipéptidos de PSA se usan para generar anticuerpos específicos para PSA que pueden después usarse para observar la presencia y/o el nivel de proteínas de PSA en procedimientos para controlar la sobreexpresión de proteína de PSA (véase, por ejemplo, Stephan y col., *Urology* 55(4):560-3 (2000)) o la metástasis de células de próstata (véase, por ejemplo, Alanen y col., *Pathol. Res. Pract.* 192(3):233-7 (1996)), los polipéptidos de 238P1B2 descritos en el presente documento pueden usarse para generar anticuerpos para su uso en la detección de sobreexpresión de 238P1B2 o la metástasis de células de próstata y células de otros cánceres que expresan este gen.

Específicamente, debido a que las metástasis implican el movimiento de células de cáncer de un órgano de origen (tal como el pulmón o la glándula prostática etc.) a un área diferente del cuerpo (tal como un ganglio linfático), pueden usarse ensayos que examinan una muestra biológica con respecto a la presencia de células que expresan polinucleótidos y/o polipéptidos de 238P1B2 para proporcionar pruebas de metástasis. Por ejemplo, cuando se descubre que una muestra biológica de tejido que normalmente no contiene células que expresan 238P1B2 (ganglio linfático) contiene células que expresan 238P1B2 tal como la expresión de 238P1B2 vista en LAPC4 y LAPC9, xenotrasplantes aislados de ganglio linfático y metástasis de hueso, respectivamente, este hallazgo es indicativo de metástasis.

Como alternativa pueden usarse polinucleótidos y/o polipéptidos de 238P1B2 para proporcionar pruebas de cáncer, por ejemplo, cuando se descubre que células en una muestra biológica que normalmente no expresan 238P1B2 o expresan 238P1B2 a un nivel diferente expresan 238P1B2 o tienen una expresión aumentada de 238P1B2 (véase, por ejemplo, la expresión de 238P1B2 en los cánceres enumerados en la Tabla I y en muestras de pacientes etc. mostrados en las Figuras adjuntas). En tales ensayos, los expertos en la materia pueden desear adicionalmente generar pruebas complementarias de metástasis ensayando la muestra biológica con respecto a la presencia de un segundo marcador restringido a tejido (además de 238P1B2) tal como PSA, PSCA etc. (véase, por ejemplo, Alanen y col., *Pathol. Res. Pract.* 192(3): 233-237 (1996)).

Igual que los fragmentos polinucleotídicos y variantes polinucleotídicas de PSA los emplean los expertos en la materia para su uso en procedimientos para controlar PSA, los fragmentos polinucleotídicos y variantes polinucleotídicas de 238P1B2 se usan de una manera análoga. En particular, los polinucleótidos de PSA típicos usados en procedimientos de control de PSA son sondas o cebadores que consisten en fragmentos de la secuencia de ADNc de PSA. Ilustrando esto, los cebadores usados para amplificar por PCR un polinucleótido de PSA deben incluir menos de la secuencia de PSA completa para actuar en la reacción en cadena de la polimerasa. En el contexto de tales reacciones de PCR, los expertos en la materia generalmente crean diversos diferentes fragmentos polinucleotídicos que pueden usarse como cebadores para amplificar diferentes partes de un polinucleótido de interés o para optimizar reacciones de amplificación (véase, por ejemplo, Caetano-Anolles, G. *Biotechniques* 25(3): 472-476, 478-480 (1998); Robertson y col., *Methods Mol. Biol.* 98: 121-154 (1998)). Se proporciona una ilustración adicional del uso de tales fragmentos en el Ejemplo 4, en el que se usa un fragmento polinucleotídico de 238P1B2 como una sonda para mostrar la expresión de ARN de 238P1B2 en células cancerosas. Además, se usan normalmente secuencias polinucleotídicas variantes como sondas y cebadores para los ARNm correspondientes en análisis de PCR y Northern (véase, por ejemplo, Sawai y col., *Fetal Diagn. Ther.* 1996 Nov-Dic 11(6): 407-13 y *Current Protocols In Molecular Biology*, Volumen 2, Unidad 2, Frederick M. Ausubel y col. eds., 1995)). Los fragmentos y variantes polinucleotídicos son útiles en este contexto cuando son capaces de unirse a una secuencia polinucleotídica diana (por ejemplo, un polinucleótido de 238P1B2 mostrado en la Figura 2 o variante del mismo) en

condiciones de alta rigurosidad.

Además, los polipéptidos de PSA que contienen un epítipo que puede reconocerse por un anticuerpo o linfocito T que se une específicamente a ese epítipo se usan en procedimientos de control de PSA. También pueden usarse fragmentos polipeptídicos de 238P1B2 y análogos o variantes polipeptídicas de una manera análoga. Esta práctica de usar fragmentos polipeptídicos o variantes polipeptídicas para generar anticuerpos (tales como anticuerpos o linfocitos T anti PSA) es típica en la técnica usándose una amplia diversidad de sistemas tales como proteínas de fusión por los facultativos (véase, por ejemplo, Current Protocols In Molecular Biology, Volumen 2, Unidad 16, Frederick M. Ausubel y col. eds., 1995). En este contexto, cada epítipo o epítopos actúan para proporcionar la arquitectura con la que un anticuerpo o linfocito T es reactivo. Normalmente, los expertos en la materia crean diversos diferentes fragmentos polipeptídicos que pueden usarse para generar respuestas inmunes específicas para diferentes partes de un polipéptido de interés (véase, por ejemplo, Patente de Estados Unidos N° 5.840.501 y Patente de Estados Unidos N° 5.939.533). Por ejemplo puede ser preferible usar un polipéptido que comprende uno de los motivos biológicos de 238P1B2 analizados en el presente documento o una subsecuencia que porta motivo que se identifica fácilmente por un experto en la materia basándose en motivos disponibles en la técnica. Los fragmentos polipeptídicos, variantes o análogos son normalmente útiles en este contexto siempre que comprendan un epítipo capaz de generar un anticuerpo o linfocito T específico para una secuencia polipeptídica diana (por ejemplo un polipéptido de 238P1B2 mostrado en la Figura 3).

Como se muestra en el presente documento, los polinucleótidos y polipéptidos de 238P1B2 (así como las sondas polinucleotídicas de 238P1B2 y anticuerpos o linfocitos T anti-238P1B2 usados para identificar la presencia de estas moléculas) muestran propiedades específicas que los hacen útiles en el diagnóstico de cánceres tales como los enumerados en la Tabla I. Se usan ensayos de diagnóstico que miden la presencia de productos génicos de 238P1B2, para evaluar la presencia o aparición de una enfermedad descrita en el presente documento, tal como cáncer de próstata, para identificar pacientes para medidas preventivas o control adicional, como se ha hecho tan exitosamente con PSA. Además, estos materiales satisfacen una necesidad de la técnica de moléculas que tengan características similares o complementarias a PSA en situaciones en las que, por ejemplo, no puede realizarse un diagnóstico definitivo de metástasis de origen prostático basándose en un ensayo de PSA solamente (véase, por ejemplo, Alanen y col., Pathol. Res. Pract. 192(3): 233-237 (1996)), y en consecuencia, es necesario emplear materiales tales como polinucleótidos y polipéptidos de 238P1B2 (así como las sondas polinucleotídicas de 238P1B2 y anticuerpos anti-238P1B2 usados para identificar la presencia de estas moléculas) para confirmar una metástasis de origen prostático.

Finalmente, además de su uso en ensayos de diagnóstico, los polinucleótidos de 238P1B2 desvelados en el presente documento tienen varias otras utilidades tales como su uso en la identificación de anomalías cromosómicas asociadas a oncogenes en la región cromosómica en la que se mapea el gen de 238P1B2 (véase Ejemplo 3 posteriormente). Por otra parte, además de su uso en ensayos de diagnóstico, las proteínas y polinucleótidos relacionados con 238P1B2 desvelados en el presente documento tiene otras utilidades tales como su uso en análisis forense de tejidos de origen desconocido (véase, por ejemplo, Takahama K Forensic Sci Int 28 de junio de 1996; 80 (1-2): 63-9).

Adicionalmente, pueden usarse proteínas o polinucleótidos relacionados con 238P1B2 descritos en el presente documento para tratar una afección patológica caracterizada por la sobreexpresión de 238P1B2. Por ejemplo, la secuencia de aminoácidos o ácido nucleico de la Figura 2 o Figura 3, o fragmentos de una de ellas, puede usarse para generar una respuesta inmune a un antígeno de 238P1B2. Pueden usarse anticuerpos u otras moléculas que reaccionan con 238P1B2 para modular la función de esta molécula y de este modo proporcionar un beneficio terapéutico.

VIII.) Kits

Para su uso en las aplicaciones de diagnóstico y terapéuticas descritas en el presente documento, también se describen kits en el presente documento. Tales kits pueden comprender un vehículo, envase o recipiente que se compartimentaliza para recibir uno o más recipientes tales como frascos, tubos y similares, comprendiendo cada uno de los recipientes uno de los elementos separados a usar en el procedimiento. Por ejemplo, el recipiente o recipientes pueden comprender una sonda que está o puede estar marcada de forma detectable. Dicha sonda puede ser un anticuerpo o polinucleótido específico para una proteína relacionada con 238P1B2 o un gen o mensaje de 238P1B2, respectivamente. Cuando el procedimiento usa hibridación de ácido nucleico para detectar el ácido nucleico diana, el kit también puede tener recipientes que contienen nucleótido o nucleótidos para amplificación de la secuencia de ácido nucleico diana y/o un recipiente que comprende un medio indicador, tal como una proteína de unión a biotina, tal como avidina o estreptavidina, unida a una molécula indicadora, tal como un marcador enzimático, fluorescencia o radioisótopo. El kit puede incluir toda o parte de la secuencia de aminoácidos de la Figura 2 o Figura 3 o análogos de la misma, o una molécula de ácido nucleico que codifique tales secuencias de aminoácidos.

El kit normalmente comprenderá el recipiente descrito anteriormente y uno o más otros recipientes que comprenden materiales deseables desde un punto de vista comercial y de usuario, incluyendo tampones, diluyentes, filtros, agujas, jeringas y prospectos con instrucciones para su uso.

Puede estar presente una etiqueta en el recipiente para indicar que la composición se usa para una terapia o aplicación no terapéutica específica y también puede indicar instrucciones para uso *in vivo* o *in vitro*, tal como los descritos anteriormente. También pueden incluirse instrucciones y/u otra información en un prospecto que se incluye con el kit.

5 **Ejemplos:**

Diversos aspectos de la invención se describen adicionalmente y se ilustran por medio de los varios ejemplos que siguen, ninguno de los cuales se pretende que limite el alcance de la invención.

Ejemplo 1: Aislamiento generado por SSH de un fragmento de ADNc del gen de 238P1B2

10 El procedimiento de Hibridación Sustractiva de Supresión (SSH) que usa ADNc derivado de tejidos de cáncer de pacientes se usa para aislar genes que se sobreexpresan en cáncer. La secuencia de ADNc de SSH de 238P1B2 se derivó de un grupo de cáncer de colon sin sustracción de ADNc de tejido normal. Se incluyeron en el conductor ADNc derivados de 10 tejidos normales. El ADNc de 238P1B2 se identificó como altamente expresado en el grupo de tejido de cáncer de próstata, con expresión restringida detectada en tejidos normales.

15 La secuencia de ADN de SSH de 210 pb (Figura 1) es nueva pero muestra homología con ARNm del receptor olfatorio de ratón MOR14-10. Se aisló un ADNc de 238P1B2, clon de 238P1B2 A, de 3754 pb de biblioteca de ADNc de próstata, revelando una ORF de 254 aminoácidos (Figura 2, Figura 3 y Figura 4A).

Materiales y procedimientos

Tejidos humanos:

20 Los tejidos canceroso y normal del paciente se obtuvieron de distintas fuentes tales como el NDRI (Philadelphia, PA). Se obtuvo ARNm para algunos tejidos normales de Clontech, Palo Alto, CA.

Aislamiento de ARN:

Se homogeneizaron tejidos en reactivo de Trizol (Life Technologies, Gibco BRL) usando ARN total aislado de tejido 10 ml/g. Se purificó ARN poli A de ARN total usando kits mini y midi de ARNm Oligotex de Qiagen's. Se cuantificaron el ARNm y total por análisis espectrofotométrico (D.O. 260/280 nm) y se analizaron por electroforesis en gel.

25 Oligonucleótidos:

Se usaron los siguientes oligonucleótidos purificados por HPLC.

DPNCDN (cebador de síntesis de ADNc):

5'TTTTGATCAAGCTT₃₀3'

Adaptador 1:

30 5'CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAG3' (SEC ID N° 708)

3'GGCCGTCCTAG5'

Adaptador 2:

5'GTAATAGCATCACTATAGGGCAGCGTGGTCGGGCCGAG3'

3'CGGCTCCTAG5'

35 Cebador de PCR 1:

5'CTAATACGACTCACTATAGGGC3'

Cebador anidado (PN) 1:

5'TCGAGCGGCCGCCCGGGCAGGA3'

Cebador anidado (PN) 2:

40 5'AGCGTGGTCGCGGCCGAGGA3'

Hibridación sustractiva de supresión

Se usó Hibridación Sustractiva de Supresión (SSH) para identificar ADNc correspondientes a genes que se expresan de forma diferencial en cáncer. La reacción de SSH usó ADNc de cáncer de colon y tejidos normales.

La secuencia génica de 238P1B2 se derivó de un grupo de cáncer de colon sin sustracción de ADNc de tejido normal. Se identificó la secuencia de ADN de SSH (Figura 1).

5 El ADNc derivado de un grupo de tejidos normales se usó como la fuente del ADNc "conductor", mientras que el ADNc de un grupo de tejidos de cáncer de paciente se usó como la fuente del ADNc "de ensayo". Se sintetizaron ADNc bicatenarios correspondientes a ADNc de ensayo y conductor a partir de 2 µg de ARN poli(A)⁺ aislado del tejido de xenotrasplante relevante, como se ha descrito anteriormente, usando Kit de Sustracción de ADNc PCR-Select de CLONTECH y 1 ng de oligonucleótido DPNCNDN como cebador. Se llevó a cabo síntesis de primera y segunda cadena como se describe en el protocolo del manual de usuario del Kit (Protocolo de CLONTECH N° PT1117-1, N° de Catálogo K1804-1). El ADNc resultante se digirió con Dpn II durante 3 horas a 37 °C. El ADNc digerido se extrajo con fenol/cloroformo (1:1) y se precipitó con etanol.

El ADNc conductor se generó combinando en una relación 1:1 ADNc digerido con Dpn II de la fuente tisular relevante (véase anteriormente) con una mezcla de ADNc digeridos derivados de los nueve tejidos normales: estómago, músculo esquelético, pulmón, cerebro, hígado, riñón, páncreas, intestino delgado y corazón.

15 Se generó ADNc de ensayo diluyendo 1 µl de ADNc digerido con Dpn II de la fuente tisular relevante (véase anteriormente) (400 ng) en 5 µl de agua. El ADNc diluido (2 µl, 160 ng) se ligó después con 2 µl de adaptador 1 y adaptador 2 (10 µM), en reacciones de ligación separadas, en un volumen total de 10 µl a 16 °C durante una noche, usando 400 u de ADN ligasa T4 (CLONTECH). La ligación se terminó con 1 µl de EDTA 0,2 M y calentamiento a 72 °C durante 5 minutos.

20 La primera hibridación se realizó añadiendo 1,5 µl (600 ng) de ADNc conductor a cada uno de dos tubos que contenían 1,5 µl (20 ng) de ADNc de ensayo ligado con adaptador 1 y adaptador 2. En un volumen final de 4 µl, las muestras se superpusieron con aceite mineral, desnaturalizado en un termociclador MJ Research a 98 °C durante 1,5 minutos, y después se permitió que hibridaran durante 8 horas a 68 °C. Las dos hibridaciones se mezclaron después juntas con un 1 µl adicional de ADNc conductor desnaturalizado fresco y se permitió que hibridaran durante una noche a 68 °C. La segunda hibridación se diluyó después en 200 µl de Hepes 20 mM, pH 8,3, NaCl 50 mM, EDTA 0,2 mM, se calentó a 70 °C durante 7 minutos y se almacenó a -20 °C.

Amplificación por PCR, clonación y secuenciación de fragmentos génicos generados de SSH:

30 Para amplificar fragmentos génicos resultantes de reacciones de SSH, se realizaron dos ampliaciones por PCR. En la reacción de PCR primaria se añadió 1 µl de la mezcla de hibridación final diluida a 1 µl de cebador de PCR 1 (10 µM), 0,5 µl de mezcla de dNTP (10 µM), 2,5 µl de tampón de reacción 10 x (CLONTECH) y 0,5 µl de mezcla de ADNc polimerasa Advantage 50 x (CLONTECH) en un volumen final de 25 µl. La PCR 1 se realizó usando las siguientes condiciones: 75 °C durante 5 minutos, 94 °C durante 25 segundos, después 27 ciclos de 94 °C durante 10 segundos, 66 °C durante 30 segundos, 72 °C durante 1,5 minutos. Se realizaron cinco reacciones de PCR primaria separadas para cada experimento. Los productos se agruparon y diluyeron 1:10 con agua. Para la reacción de PCR secundaria, se añadió 1 µl de la reacción de PCR primaria diluida y agrupada a la misma mezcla de reacción que se usó para PCR 1, excepto que se usaron cebadores NP1 y NP2 (10 µM) en lugar de cebador de PCR 1. La PCR 2 se realizó usando 10-12 ciclos de 94 °C durante 10 segundos, 68 °C durante 30 segundos y 72 °C durante 1, 5 minutos. Los productos PCR se analizaron usando electroforesis en gel de agarosa 2 %.

40 Los productos de PCR se insertaron en pCR2.1 usando el kit de clonación de vector T/A (Invitrogen). Las *E. coli* transformadas se sometieron a selección de ampicilina y azul/blanco. Se seleccionaron colonias blancas y se dispusieron en placas de 96 pocillos y se dejaron crecer en cultivo líquido durante una noche. Para identificar insertos, se realizó amplificación por PCR en 1 ml de cultivo bacteriano usando las condiciones de PCR1 y NP1 y NP2 como cebadores. Los productos de PCR se analizaron usando electroforesis en gel de agarosa 2 %.

45 Se almacenaron clones bacterianos en glicerol 20 % en un formato de 96 pocillos. Se preparó ADN plasmídico, se secuenció y se sometió a búsquedas de homología de ácidos nucleicos de las bases de datos GenBank, dBest y NCI-CGAP.

Análisis de expresión de RT-PCR:

50 Puede generarse ADNc de primera cadena de 1 µg de ARNm con cebador oligo (dT) 12-18 usando el sistema de preamplificación Gibco-BRL Superscript. Se usó el protocolo del fabricante que incluía una incubación durante 50 minutos a 42 °C con transcriptasa inversa seguido de tratamiento con ARNasa H a 37 °C durante 20 minutos. Después de completar la reacción, el volumen puede aumentarse a 200 µl con agua antes de normalización. Puede obtenerse ADNc de primera cadena de 16 tejidos humanos normales diferentes de Clontech.

55 Se realizó la normalización de los ADNc de primera cadena de múltiples tejidos usando los cebadores 5'atattcgccgctcgtcgtcgacaa3' y 5'agccacacgcagctcattgtagaagg 3' para amplificar β-actina. Se amplificaron ADNc de primera cadena (5 µl) en un volumen total de 50 µl que contenía cebadores 0,4 µM, dNTP cada uno 0,2 µM, tampón de PCR 1X (Clontech, Tris-HCL 10 mM, MgCl₂ 1,5 mM, KCl 50 mM, pH 8.3) y ADN polimerasa Klentaq 1X (Clontech). Pueden retirarse 5 µl de la reacción de PCR a los 18, 20 y 22 ciclos y usarse para electroforesis en gel

de agarosa. Se realizó PCR usando un termociclador MJ Research en las siguientes condiciones: la desnaturalización inicial puede ser a 94 °C durante 15 segundos, seguido de 18, 20 y 22 ciclos de 94 °C durante 15, 65 °C durante 2 minutos, 72 °C durante 5 segundos. Se llevó a cabo una extensión final a 72 °C durante 2 minutos. Después de electroforesis en gel de agarosa, las intensidades de banda de las bandas de β-actina de 283 p.b. de múltiples tejidos se compararon por inspección visual. Se calcularon los factores de dilución para los ADNc de primera cadena para dar como resultado intensidades de banda de β-actina iguales en todos los tejidos después de 22 ciclos de PCR. Pueden requerirse tres rondas de normalización para conseguir intensidades de banda iguales en todos los tejidos después de 22 ciclos de PCR.

Para determinar los niveles de expresión del gen 238P1B2, se analizaron 5 µl de ADNc de primera cadena normalizado por PCR usando 26 y 30 ciclos de amplificación. Puede conseguirse análisis de expresión semicuantitativa comparando los productos de PCR en los números de ciclos que dan intensidades de banda claras. Los cebadores usados para RT-PCR se diseñaron usando la secuencia de SSH de 238P1B2 y se enumeran a continuación:

238P1B2.1

5'-TTGCAGAATATCACCTCCACTTCC -3'

238P1B2.2

5'-GATCAGGCTGTTTCCCAAGAGAG -3'

Se muestra un análisis de expresión de RT-PCR típico en la Figura 14. Se realizó análisis de expresión de RT-PCR en ADNc de primera cadena generados usando grupos de tejidos de múltiples muestras. Se mostró que los ADNc se normalizaban usando PCR de beta-actina. Los resultados muestran expresión fuerte de 238P1B2 en grupo de cáncer de próstata pero no en grupo vital 1 y grupo vital 2.

Ejemplo 2: Clonación de longitud completa de 238P1B2

La secuencia de ADNc de SSH de 238P1B2 derivó de un grupo de cáncer de colon sin sustracción de ADNc de tejido normal. La secuencia de ADNc de SSH (Figura 1) se designó 238P1B2.

La secuencia de ADN de SSH de 210 pb (Figura 1) es nueva y muestra 91 % de identidad con ARNm del receptor olfatorio de ratón MOR14-10.

Se aisló un ADNc de longitud completa (clon de 238P1B2 A) de 3754 pb de biblioteca de próstata, revelando una ORF de 254 aminoácidos (Figuras 2 y 3). El ADNc muestra mayor homología con los receptores olfatorios MOR14-1 y MOR14-10 de ratón, con identidades de 85 % sobre 912 nucleótidos y 83 % sobre 906 nucleótidos respectivamente (Figura 4).

La secuencia proteica revela 7 dominios transmembrana y tiene homología con receptores acoplados a proteína G (GPCR) implicados en la olfacción (Raming y col., 1993, Nature 361:353; Malnic y col., 1999, Cell 96:713). Las proteínas que son miembros de esta familia de receptores muestran un extremo amino terminal extracelular, tres bucles extracelulares adicionales, tres bucles intracelulares y un extremo carboxilo terminal intracelular.

La secuencia más homóloga con 238P1B2 es la proteína MOR14-1 de ratón. Las dos proteínas comparten 83 % de identidad sobre una región de 253 aminoácidos. El alineamiento de las dos proteínas se muestra en la Figura 4.

238P1B2 también muestra homología significativa con el GPCR específico de próstata, PHOR-1. Las dos proteínas comparten 48 % de identidad sobre una región de 250 aminoácidos (Figura 4).

El ADNc de 238P1B2 de longitud completa (clon de 238P1B2 A) se depositó en la Colección Americana de Cultivos Tipo el 7 de marzo de 2002 y se le asignó el número de acceso PTA-4124.

La secuencia proteica de 238P1B2 v.1A codifica una proteína de seis transmembranas. La extensión de la proteína en el extremo amino terminal añade otros 62 aminoácidos a 238P1B2 v.1A que conducen a 238P1B2 v.1B. Este extremo de 62 aminoácidos de 238P1B2 v.1B codifica una región transmembrana adicional y de este modo hace a 238P1B2 v.1B una proteína de siete transmembranas. Además, este extremo amino terminal muestra 84 % de identidad con el receptor de proteína acoplado a G MOR14-10 de ratón, lo que indica que esta parte amino terminal está codificada en la naturaleza.

La expresión natural de 238P1B2 v.1B puede producirse si el sitio de parada cadena arriba de la extensión de 62 aminoácidos se modifica por sustitución de ácido nucleico sencilla que conduce a conversión del codón de parada a un aminoácido codificante. Pueden existir polimorfismos dentro de diferentes tejidos o entre diferentes individuos que mutan el codón de parada y permitiendo de este modo la traducción apropiada de la región amino terminal de 62 aminoácidos de 238P1B2 v.1B. Una variación tal puede añadir una metionina de partida directamente o extender la región codificante de 238P1B2 v.1B 16 aminoácidos adicionales hasta alcanzar una metionina de partida en la posición de ácido nucleico 1896.

Ejemplo 3: Mapeo cromosómico del gen de 238P1B2

La localización cromosómica puede implicar genes en patogénesis de enfermedad. Están disponibles varios enfoques de mapeo de cromosomas incluyendo hibridación fluorescente *in situ* (FISH), paneles de híbrido de radiación humano/hámster (RH) (Walter y col., 1994; Nature Genetics 7:22; Research Genetics, Huntsville AI), paneles de híbrido celular somático de humano-roedor tales como están disponible del Instituto Coriell (Camden, New Jersey), y visores genómicos que usan homologías de BLAST para clones genómicos secuenciados y mapeados (NCBI, Bethesda, Maryland).

Usando la secuencia de 238P1B2 y la herramienta BLAST de NCBI: (véase la dirección Web www.ncbi.nlm.nih.gov/genome/seq/page.cgi?F=HsBlast.html&&ORG=Hs), se situó el 238P1B2 en el cromosoma 11p15.5, una región rica en GPCR.

Debido a que el gen de 238P1B2 humano se mapea en el cromosoma 11p15.5, pueden usarse polinucleótidos que codifican diferentes regiones de la proteína 238P1B2 para caracterizar anomalías citogenéticas en el cromosoma 11, banda p15.5 que se han identificado como asociadas con diversos cánceres. En particular, se han identificado diversas anomalías cromosómicas en 11p15.5 como anomalías citogenéticas frecuentes en varios cánceres diferentes (véase, por ejemplo, Lai y col., 2000, Clin. Cancer Res. 6(8):3172-6; Oya y Schulz, 2000, Br. J. Cancer 83(5):626-31; Svaren y col., Sept. 12, 2000, J. Biol. Chem). En consecuencia, los polinucleótidos que codifican regiones específicas de la proteína 238P1B2 proporcionan nuevas herramientas que pueden usarse para delinear, con mayor precisión de lo que era posible previamente, la naturaleza específica de las anomalías citogenéticas en esta región del cromosoma 11 que contribuye al fenotipo maligno. En este contexto, estos polinucleótidos satisfacen una necesidad en la técnica para expandir la sensibilidad de exploración cromosómica para identificar anomalías cromosómicas más sutiles y menos habituales (véase, por ejemplo, Evans y col., 1994, Am J. Obstet. Gynecol. 171 (4): 1055-1057).

Ejemplo 4: Análisis de expresión de 238P1B2 en tejidos normales y muestras de ensayo de pacientes

El análisis de expresión por RT-PCR demostró que 238P1B2 se expresa fuertemente en muestras de ensayo de pacientes con cáncer de próstata (Figura 14). Se preparó ADNc de primera cadena del grupo vital 1 (hígado, pulmón y riñón), grupo vital 2 (páncreas, colon y estómago) y grupo de cáncer de próstata. Se realizó normalización por PCR usando cebadores para actina y GAPDH. Se realizó PCR semicuantitativa, usando cebadores para 238P1B2 a 26 y 30 ciclos de amplificación. Los resultados muestran fuerte expresión de 238P1B2 en grupo de cáncer de próstata pero no en grupo vital 1 y grupo vital 2.

Se muestra análisis de transferencia de Northern de 238P1B2 en 16 tejidos normales humanos en la Figura 15. Los resultados muestran la ausencia de expresión de 238P1B2 en los 16 tejidos normales ensayados. El análisis exhaustivo de la expresión de 238P1B2 en 76 tejidos humanos muestra expresión restringida de 238P1B2 en placenta (Figura 16).

La expresión de 238P1B2 en muestras de ensayo de cáncer de paciente y tejidos normales humanos se muestra en la Figura 17. Se extrajo ARN de un grupo de tres cánceres de próstata, así como de próstata normal (PN), vejiga normal (VN), riñón normal (RN), colon normal (CN), pulmonar normal (PuN), mama normal (MN) y ovario normal (ON). La transferencia de Northern con 10 µg de ARN total/carril se exploró con secuencia de 238P1B2. Los resultados muestran expresión de un transcrito de 238P1B2 de aproximadamente 4,5 kb en el grupo de cáncer de próstata y ovario pero no en los otros tejidos normales ensayados. El análisis de muestras de ensayo de pacientes individuales muestra expresión fuerte de 238P1B2 en tejidos de cáncer de próstata (Figura 18). La expresión para 238P1B2 detectada en tumores de puntuación de Gleason 7 es significativamente más fuerte que la expresión detectada en tumores de puntuación de Gleason 5. Este resultado indica que 238P1B2 puede usarse como un marcador de pronóstico para cáncer de próstata.

La expresión restringida de 238P1B2 en tejidos normales y la expresión detectada en cáncer de próstata sugiere que 238P1B2 es una diana terapéutica potencial y un marcador de diagnóstico para cánceres humanos.

Ejemplo 5: Variantes de transcrito de 238P1B2

Las variantes de transcrito son variantes de ARNm maduro del mismo gen por transcripción alternativa o corte y empalme alternativo. Los transcritos alternativos son transcritos del mismo gen que comienza la transcripción en puntos diferentes. Las variantes de corte y empalme son variantes de ARNm con corte y empalme diferente del mismo transcrito en eucariotas, cuando se transcribe un gen multiexónico de ADN genómico, el ARN inicial se corta y empalma para producir ARNm funcional, que solamente tiene exones y se usa para traducción en una secuencia de aminoácidos. En consecuencia, un gen dado puede tener de cero a muchos transcritos alternativos y cada transcrito puede tener de cero a muchas variantes de corte y empalme. Cada variante de transcrito tiene una composición exónica única y puede tener diferentes partes codificantes y/o no codificantes (extremo 5' o 3'), del transcrito original. Las variantes de transcrito pueden codificar proteínas similares o diferentes con la misma o similar función o pueden codificar proteínas con diferentes funciones y pueden expresarse en el mismo tejido al mismo tiempo o en diferentes tejidos al mismo tiempo o en el mismo tejido a diferentes tiempos o en diferentes tejidos a diferentes tiempos. Las proteínas codificadas por variantes de transcrito pueden tener localizaciones celulares o

extracelulares similares o diferentes, por ejemplo, secretadas frente a intracelulares.

Se identifican variantes de transcrito por diversos procedimientos aceptados en la técnica. Por ejemplo, se identifican transcritos alternativos y variantes de corte y empalme por experimentos de clonación de longitud completa o mediante el uso de secuencias de EST y transcrito de longitud completa. En primer lugar, todas las EST humanas se agrupan en grupos que muestran identidad directa o indirecta entre sí. En segundo lugar, las EST en el mismo grupo se agrupan adicionalmente en subgrupos y se ordenaron en una secuencia consenso. La secuencia del gen original se compara con la secuencia o secuencias consenso u otras secuencias de longitud completa. Cada secuencia consenso es una variante de corte y empalme potencial para ese gen (véase, por ejemplo, la dirección Web www.doubletwt.com/products/c11_agentsOverview.jhtml). Incluso cuando se identifica una variante que no es un clon de longitud completa, esa parte de la variante es muy útil para generación de antígenos y para clonación adicional de la variante de corte y empalme de longitud completa, usando técnicas conocidas en la materia.

Además, están disponibles programas informáticos en la técnica para identificar variantes de transcrito basándose en secuencias genómicas. Los programas de identificación de variantes de transcrito basados en genoma incluyen FgenesH (A. Salamov y V. Solovyev, "Ab initio gene finding in Drosophila genomic DNA", *Genome Research*. Abril de 2000; 10(4):516-22); Grail (dirección Web compbio.ornl.gov/Grail-bin/EmptyGrailForm) y GenScan (dirección Web genes.mit.edu/GENSCAN.html). Para un análisis general de protocolos de identificación de variantes de corte y empalme véase, por ejemplo, Southan, C., A genomic perspective on human proteases, *FEBS Lett.* 8 de junio de 2001; 498(2-3):214-8; de Souza, S.J., y col., Identification of human chromosome 22 transcribed sequences with ORF expressed sequence tags, *Proc. Natl Acad Sci U S A.* 7 de noviembre de 2000; 97(23):12690-3.

Para confirmar adicionalmente los parámetros de una variante de transcrito, están disponibles diversas técnicas en la materia, tales como clonación de longitud completa, validación proteómica, validación basada en PCR y validación de RACE 5', etc. (véase por ejemplo, Validación Proteómica: Brennan, S.O., y col., Albumin banks peninsula: a new termination variant characterized by electrospray mass spectrometry, *Biochem Biophys Acta.* 17 de agosto de 1999;1433(1-2):321-6; Ferranti P, y col., Differential splicing of premessenger RNA produces multiple forms of mature caprine alpha(s1)-casein, *Eur J Biochem.* 1 de octubre de 1997;249(1):1-7. Para validación basada en PCR: Wellmann S, y col, Specific reverse transcription-PCR quantification of vascular endothelial growth factor (VEGF) splice variants by LightCycler technology, *Clin Chem.* Abril de 2001;47(4):654-60; Jia, H.P., y col, Discovery of new human beta-defensins using a genomics-based approach, *Gene.* 24 de enero de 2001; 263(1-2):211-8. Para validación de RACE 5' y basada en PCR: Brigle, K.E., y col, Organization of the murine reduced folate carrier gene and identification of variant splice forms, *Biochem Biophys Acta.* 7 de agosto de 1997; 1353(2): 191-8).

Se sabe en la técnica que las regiones genómicas están moduladas en los cánceres. Cuando la región genómica en la que se mapea un gen está modulada en un cáncer particular, las variantes de corte y empalme o transcritos alternativos del gen se modulan también. Se desvela en el presente documento que 238P1B2 tiene un perfil de expresión particular. Las variantes de transcritos y corte y empalme alternativos de 238P1B2 pueden compartir este patrón de expresión, sirviendo de este modo como marcadores/antígenos asociados a tumor.

La composición exónica de transcrito original, designado 238P1B2 v.1, se muestra en la tabla XXIII. No se ha identificado variante de transcrito por los procedimientos anteriores.

Ejemplo 6: Polimorfismos de nucleótido sencillo de 238P1B2

Un polimorfismo de nucleótido sencillo (SNP) es una variación de par de bases sencillo en una secuencia de nucleótidos en una localización específica. En cualquier punto dado del genoma, existen cuatro posibles pares de bases de nucleótidos: A/T, C/G, G/C y T/A. El genotipo se refiere a la secuencia de pares de bases específica de una o más localizaciones en el genoma de un individuo. El haplotipo se refiere a la secuencia de pares de bases de más de una localización en la misma molécula de ADN (o el mismo cromosoma en organismos superiores), con frecuencia en el contexto de un gen o en el contexto de varios genes ligados de forma estrecha. Los SNP que se producen en un ADNc se denominan cSNP. Estos cSNP pueden cambiar aminoácidos de la proteína codificada por el gen y cambiar de este modo las funciones de la proteína. Algunos SNP provocan enfermedades heredadas; otros contribuyen a variaciones cuantitativas en el fenotipo y reacciones a factores ambientales incluyendo dieta y fármacos entre los individuos. Por lo tanto, los SNP y/o combinaciones de alelos (denominados haplotipos) tienen muchas aplicaciones, incluyendo diagnóstico de enfermedades heredadas, determinación de reacciones a fármacos y dosificación, identificación de genes responsables de enfermedades y análisis de la relación genética entre los individuos (P. Nowotny, J. M. Kwon y A. M. Goate, " SNP analysis to dissect human traits," *Curr. Opin. Neurobiol.* Octubre de 2001; 11(5):637-641; M. Pirmohamed y B. K. Park, "Genetic susceptibility to adverse drug reactions," *Trends Pharmacol. Sci.* Junio de 2001; 22(6):298-305; J. H. Riley, C. J. Allan, E. Lai y A. Roses, " The use of single nucleotide polymorphisms in the isolation of common disease genes," *Pharmacogenomics.* Febrero de 2000; 1(1):39-47; R. Judson, J. C. Stephens y A. Windemuth, "The predictive power of haplotypes in clinical response," *Pharmacogenomics.* Febrero de 2000; 1(1):15-26).

Los SNP se identifican mediante diversos procedimientos aceptados en la técnica (P. Bean, "The promising voyage of SNP target discovery," *Am. Clin. Lab.* Octubre-noviembre de 2001; 20(9):18-20; K. M. Weiss, "In search of human

variation," *Genome Res.* Julio de 1998; 8(7):691-697; M. M. She, "Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies," *Clin. Chem.* Febrero de 2001; 47(2):164-172). Por ejemplo, se identifican SNP secuenciando fragmentos de ADN que muestran polimorfismo por procedimientos basados en gel tales como polimorfismo de longitud de fragmento de restricción (RFLP) y electroforesis en gel de gradiente desnaturalizante (DGGE). También pueden descubrirse por secuenciación directa de muestras de ADN agrupadas de diferentes individuos o comparando secuencias de diferentes muestras de ADN. Con la rápida acumulación de datos de secuencias en las bases de datos públicas y privadas, pueden descubrirse SNP comparando secuencias usando programas informáticos (Z. Gu, L. Hillier y P. Y. Kwok, "Single nucleotide polymorphism hunting in cyberspace" *Hum. Mutat* 1998; 12(4):221-225). Los SNP puede verificarse y puede determinarse el genotipo o haplotipo de un individuo por diversos procedimientos incluyendo secuenciación directa y microseries de alto rendimiento (P. Y. Kwok, "Methods for genotyping single nucleotide polymorphisms," *Annu. Rev. Genomics Hum. Genet.* 2001; 2:235-258; M. Kokoris, K. Dix, K. Moynihan, J. Mathis, B. Erwin, P. Grass, B. Hines y A. Duesterhoeft, "High-throughput SNP genotyping with the Masscode system," *Mol. Diagn.* Diciembre de 2000; 5(4):329-340).

Se identifican SNP de 238P1B2 por secuenciación directa de los clones de ADNc y por comparación de esas secuencias con secuencias públicas y patentadas. Comparando las secuencias con secuencias públicas o patentadas de alta calidad (por ejemplo, NCBI/GenBank, accesibles en la dirección Web www.ncbi.nlm.nih.gov), se identificaron cinco SNP de 238P1B2 en las posiciones de nucleótidos 274 (T/C), 1268 (T/G), 1299 (T/G), 2806 (T/C) y 3025 (T/C). Los transcritos o proteínas con alelos alternativos se designaron como variantes de 238P1B2 v.2, v.3, v.4, v.5 y v.6. La Figura 10 muestra el alineamiento esquemático de las variantes de nucleótidos. La Figura 11 muestra el alineamiento esquemático de variantes de proteína, correspondientes a variantes de nucleótidos. Las variantes de nucleótidos que codifican la misma secuencia de aminoácidos como variante 1 no se muestran en la Figura 11. Estos alelos de los SNP, aunque se muestran aquí de forma separada, pueden aparecer en diferentes combinaciones (haplotipos).

Ejemplo 7: Producción de 238P1B2 recombinante en sistemas procariotas

Para expresar 238P1B2 recombinante en células procariotas, se clonan las secuencias de ADNc de longitud parcial o completa de 238P1B2 variante 1a, variante 1b y variante 2 en uno cualquiera de diversos vectores de expresión conocidos en la técnica. Una o más de las siguientes regiones de variantes de 238P1B2 se expresan en estas construcciones: aminoácidos 1 a 254 de la variante 1a; aminoácidos 1 a 316 de la variante 1b y aminoácidos 1-254 de la variante 2 o cualesquiera 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 o más aminoácidos contiguos de 238P1B2, variantes o análogos de la misma.

A. Construcciones de transcripción y traducción *in vitro*:

pCRII: para generar sondas de ARN sentido y antisentido de 238P1B2 para investigaciones de ARN *in situ*, se generan construcciones de pCRII (Invitrogen, Carlsbad CA) codificando todo o fragmentos del ADNc de 238P1B2. El vector pCRII tiene los promotores Sp6 y T7 flanqueando el inserto para conducir la transcripción de ARN de 238P1B2 para su uso como sondas en experimentos de hibridación de ARN *in situ*. Estas sondas se usan para analizar la expresión celular y tisular de 238P1B2 al nivel de ARN. Se usa ARN de 238P1B2 transcrito que representa la región codificante de aminoácidos de ADN del gen de 238P1B2 en sistemas de traducción *in vitro* tales como el sistema de reticulolisado acoplado TnT™ (Promega, Corp., Madison, WI) para sintetizar proteína 238P1B2.

B. Construcciones Bacterianas:

Construcciones de pGEX: para generar proteínas 238P1B2 recombinantes en bacterias que se fusionan con la proteína glutatión proteína glutatión S-transferasa (GST); se fusiona toda o partes de la secuencia codificante de proteína de ADNc de 238P1B2 con el gen de GST clonando en pGEX-6P-1 o cualquier otro vector de fusión de GST de la familia pGEX (Amersham Pharmacia Biotech, Piscataway, NJ). Estas construcciones permiten la expresión controlada de secuencias proteicas de 238P1B2 recombinantes con GST fusionada en el extremo amino terminal y un epítipo de seis histidinas (6X His) en el extremo carboxilo terminal. Los marcadores de GST y 6X His permiten la purificación de la proteína de fusión recombinante de bacterias inducidas con la matriz de afinidad apropiada y permiten el reconocimiento de la proteína de fusión con anticuerpos anti-GST y anti-His. El marcador de 6X His se genera añadiendo 6 codones de histidina al cebador de clonación en el extremo 3', por ejemplo, del marco de lectura abierto (ORF). Puede emplearse un sitio de escisión proteolítica, tal como el sitio de reconocimiento de PreScission™ en pGEX-6P-1, para permitir la escisión del marcador de GST de proteína relacionada con 238P1B2. El gen de resistencia a ampicilina y origen de pBR322 permite la selección y mantenimiento de los plásmidos pGEX en *E. coli*.

Construcciones de pMAL: para generar, en bacterias, proteínas 238P1B2 recombinantes que se fusionan con proteína de unión a maltosa (MBP), toda o partes de la secuencia codificante de proteína de ADNc de 238P1B2 se fusiona con el gen de MBP clonando en los vectores pMAL-c2X y pMAL-p2X (New England Biolabs, Beverly, MA). Estas construcciones permiten la expresión controlada de secuencias proteicas de 238P1B2 recombinantes con MBP fusionada en el extremo amino terminal y un marcador epítipo de 6X His en el extremo carboxilo terminal.

Los marcadores MBP y 6X His permiten la purificación de la proteína recombinante de bacterias inducidas con la matriz de afinidad apropiada y permiten el reconocimiento de la proteína de fusión con anticuerpos anti-BMP y anti-His. El marcador epitópico de 6X His se genera añadiendo 6 codones de histidina al cebador de clonación 3'. Un sitio de reconocimiento de factor Xa permite la escisión del marcador de pMAL de 238P1B2. Los vectores pMAL-c2X y pMAL-p2X se optimizan para expresar la proteína recombinante en el citoplasma o periplasma respectivamente. La expresión del periplasma potencia el plegamiento de proteínas con enlaces disulfuro.

Construcciones de pET: para expresar 238P1B2 en células bacterianas, toda o partes de la secuencia codificante de proteínas de ADNc de 238P1B2 se clona en la familia de pET de vectores (Novagen, Madison, WI). Estos vectores permiten expresión estrechamente controlada de proteína 238P1B2 recombinante en bacterias con y sin fusión con proteínas que potencian la solubilidad, tales como NusA y tiorredoxina (Trx), y marcadores epitópicos, tales como 6X His y S-Tag™ que ayudan a la purificación y detección de la proteína recombinante. Por ejemplo, se realizan construcciones usando el sistema de fusión de pET NusA 43.1 de modo que se expresen regiones de la proteína 238P1B2 como fusiones amino terminales con NusA. Se expresó una proteína de fusión ANusA que abarcaba los aminoácidos 412-254 de 238P1B2 con un marcador 6x His C-terminal en *E. coli*, se purificó por cromatografía de afinidad de quelado metálico y se usó como un inmunógeno para generación de anticuerpos.

C. Construcciones de levadura:

Construcciones de pESC: para expresar 238P1B2 en la especie de levadura *Saccharomyces cerevisiae* para generación de proteína recombinante y estudios funcionales, toda o partes de la secuencia codificante de proteína de ADNc de 238P1B2 se clona en la familia pESC de vectores cada uno de los cuales contiene 1 de 4 marcadores seleccionables, HIS3, TRP1, LEU2 y URA3 (Stratagene, La Jolla, CA). Estos vectores permiten la expresión controlada del mismo plásmido de hasta 2 genes diferentes o secuencias clonadas que contienen marcadores epitópicos Flag™ o Myc en la misma célula de levadura. Este sistema es útil para confirmar interacciones proteína-proteína de 238P1B2. Además, la expresión en levadura produce modificaciones post-traduccionales similares, tales como glucosilaciones y fosforilaciones, que se encuentran cuando se expresan en células eucariotas.

Construcciones de pESP: para expresar 238P1B2 en la especie de levaduras *Saccharomyces pombe*, toda o partes de la secuencia codificante de proteína de ADNc de 238P1B2 se clona en la familia pESP de vectores. Estos vectores permiten alto nivel de expresión controlada de una secuencia proteica de 238P1B2 que se fusiona en el extremo amino terminal o en el extremo carboxilo terminal con GST que ayuda a la purificación de la proteína recombinante. Un marcador epitópico Flag™ permite la detección de la proteína recombinante con un anticuerpo anti-Flag™.

Ejemplo 8: Producción de 238P1B2 recombinante en sistemas eucariotas

A. Construcciones de mamíferos

Para expresar 238P1B2 recombinante en células eucariotas, las secuencias de longitud completa o parcial de ADNc de 238P1B2 pueden clonarse en uno cualquiera de diversos vectores de expresión conocidos en la técnica. Se expresa una o más de las siguientes regiones de 238P1B2 en estas construcciones: aminoácidos 1 a 254 de variante 1A o variante 2, aminoácidos 1 a 316 de variante 1B, o cualesquiera 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 o más aminoácidos contiguos de 238P1B2, variantes, o análogos de la misma. Se expresa una región de una variante específica de 238P1B2 que codifica un aminoácido en posición específica que difiere del aminoácido de cualquier otra variante hallado en esa posición. Como alternativa, se expresa una región de una variante de 238P1B2 que queda parcial o completamente dentro de una secuencia que es única de esa variante.

Las construcciones pueden transfectarse en una cualquiera de diversas células de mamífero tales como células 293T. Los lisados de células 293T transfectadas pueden explorarse con el suero policlonal anti-238P1B2, descrito en el presente documento.

Construcciones pcDNA4/HisMax: para expresar 238P1B2 en células de mamífero, se clona una ORF de 238P1B2, o partes de la misma, de 238P1B2 en pcDNA4/HisMax Versión A (Invitrogen, Carlsbad, CA). La expresión proteica se conduce desde el promotor de citomegalovirus (CMV) y el potenciador traduccional SP16. La proteína recombinante tiene epítotos Xpress™ y seis histidinas (6X His) fusionados con el extremo amino terminal. El vector pcDNA4/HisMax también contiene la señal de poliadenilación y secuencia de terminación de la transcripción de la hormona del crecimiento bovina (BGH) para potenciar la estabilidad de ARNm junto con el origen de SV40 para replicación episomal y rescate de vector sencillo en líneas celulares que expresan el antígeno T grande. El gen de resistencia a zeocina permite la selección de células de mamífero que expresan la proteína y el gen de resistencia a ampicilina y el origen ColE1 permite la selección y mantenimiento del plásmido en *E. coli*.

Construcciones pcDNA3.1/MycHis: para expresar 238P1B2 en células de mamífero, se clona una ORF de 238P1B2, o partes de la misma, de 238P1B2 con un sitio de inicio de la traducción de Kozak consenso en pcDNA3.1/MycHis Versión A (Invitrogen, Carlsbad, CA). La expresión proteica se conduce desde el promotor de citomegalovirus (CMV). Las proteínas recombinantes tienen el epítoto myc y el epítoto 6X His fusionados en el

extremo carboxilo terminal. El vector pcDNA3.1/MycHis también contiene la señal de poliadenilación y secuencia de terminación de la transcripción de la hormona de crecimiento bovina (BGH) para potenciar la estabilidad de ARNm, junto con el origen de SV40 para replicación episomal y rescate de vector sencillo en líneas celulares que expresan el antígeno T grande. El gen de resistencia a neomicina puede usarse, puesto que permite la selección de células de mamífero que expresan la proteína y el gen de resistencia a ampicilina y el origen ColE1 permite la selección y mantenimiento del plásmido en *E. coli*.

Construcción pcDNA3.1/CT-GFP-TOPO: para expresar 238P1B2 en células de mamífero y permitir la detección de las proteínas recombinantes usando fluorescencia, una ORF 238P1B2, o partes de la misma, con un sitio de inicio de la traducción de Kozak consenso se clonan en pcDNA3.1/CT-GFP-TOPO (Invitrogen, CA). La expresión proteica se conduce desde el promotor de citomegalovirus (CMV). Las proteínas recombinantes tienen la proteína verde fluorescente (GFP) fusionada con el extremo carboxilo terminal que facilita la detección no invasiva *in vivo* y estudios de biología celular. El vector pcDNA3.1CT-GFP-TOPO también contiene la señal de poliadenilación y secuencia de terminación de la transcripción de la hormona de crecimiento bovina (BGH) para potenciar la estabilidad de ARNm junto con el origen de SV40 para replicación episomal y rescate de vector sencillo en líneas celulares que expresan el antígeno T grande. El gen de resistencia a neomicina permite la selección de células de mamífero que expresan la proteína y el gen de resistencia a ampicilina y origen ColE1 permiten la selección y mantenimiento del plásmido en *E. coli*. Se realizan construcciones adicionales con una fusión de GFP amino terminal en pcDNA3.1/NT-GFP-TOPO que abarca la longitud completa de una proteína 238P1B2.

PAPtag: una ORF de 238P1B2, o partes de la misma, se clona en pAPtag-5 (GenHunter Cop. Nashville, TN). Esta construcción genera una fusión de fosfatasa alcalina en el extremo carboxilo terminal de la proteína 238P1B2 fusionando a la vez la secuencia señal de IgGκ con el extremo amino terminal. También se generan construcciones en las que se fusiona fosfatasa alcalina con una secuencia señal de IgGκ amino terminal con el extremo amino terminal de una proteína 238P1B2. Las proteínas 238P1B2 recombinantes resultantes se optimizan para secreción al medio de células de mamífero transfectadas y pueden usarse para identificar proteínas tales como ligandos o receptores que interactúan con proteínas 238P1B2. La expresión de proteínas se conduce desde el promotor de CMV y las proteínas recombinantes también contienen epítopos myc y 6X His fusionados en el extremo carboxilo terminal que facilitan la detección y purificación. El gen de resistencia a zeocina presente en el vector permite la selección de células de mamífero que expresan la proteína recombinante y el gen de resistencia a ampicilina permite la selección del plásmido en *E. coli*.

ptag5: Una ORF de 238P1B2, o partes de la misma, se clona en pTag-5. Este vector es similar a pAPtag pero sin la fusión de fosfatasa alcalina. Esta construcción genera proteína 238P1B2 con una secuencia señal de IgGκ amino terminal y marcadores epítópicos myc y 6X His en el extremo carboxilo terminal que facilitan la detección y purificación de afinidad. La proteína 238P1B2 recombinante resultante se optimiza para secreción al medio de células de mamífero transfectadas y se usa como inmunógeno o ligando para identificar proteínas tales como ligandos o receptores que interactúan con las proteínas 238P1B2. La expresión proteica se conduce desde el promotor de CMV. El gen de resistencia a zeocina presente en el vector permite la selección de células de mamífero que expresan la proteína y el gen de resistencia a ampicilina permite la selección del plásmido en *E. coli*.

PsecFc: Una ORF de 238P1B2, o partes de la misma, también se clona en psecFc. El vector psecFc se ensambló clonando el Fc de inmunoglobulina G1 (IgG) humana (regiones bisagra, CH2, CH3) en pSecTag2 (Invitrogen, California). Esta construcción genera una fusión de Fc de IgG1 en el extremo carboxilo terminal de las proteínas 238P1B2, fusionando a la vez la secuencia señal de IgGκ en el extremo N-terminal. También se usan fusiones de 238P1B2 que usan la región Fc de IgG1 murina. Las proteínas 238P1B2 recombinantes resultantes se optimizan para secreción al medio de células de mamífero transfectadas y pueden usarse como inmunógenos o para identificar proteínas tales como ligandos o receptores que interactúan con proteína 238P1B2. La expresión proteica se conduce desde el promotor de CMV. El gen de resistencia a higromicina presente en el vector permite la selección de células de mamífero que expresan la proteína recombinante y el gen de resistencia a ampicilina permite la selección del plásmido en *E. coli*.

Construcciones de pSRα: para generar líneas celulares de mamífero que expresen 238P1B2 de forma constitutiva, se clonaron ORF de 238P1B2, o partes de la misma, de 238P1B2 en construcciones pSRα. Se generan retrovirus anfitriónicos y ecotrópicos por transfección de construcciones pSRα en la línea de empaquetamiento 293T-10A1 o co-transfección de pSRα y un plásmido auxiliar (que contiene secuencias de empaquetamiento suprimidas) en las células 293, respectivamente. El retrovirus se usa para infectar diversas líneas celulares de mamífero, dando como resultado la integración del gen clonado, 238P1B2, en las líneas celulares huésped. La expresión proteica se conduce desde una repetición terminal larga (LTR). El gen de resistencia a neomicina presente en el vector permite la selección de células de mamífero que expresan la proteína y el gen de resistencia a ampicilina y origen ColE1 permiten la selección y mantenimiento del plásmido en *E. coli*. Los vectores retrovirales pueden usarse a continuación para infección y generación de diversas líneas celulares usando, por ejemplo, células PC3, NIH 3T3, TsuPr1, 293 o rat-1.

Se realizan construcciones pSRα adicionales que fusionan un marcador epítópico tal como el marcador FLAG™ con el extremo carboxilo terminal de secuencias de 238P1B2 para permitir la detección usando anticuerpo anti-Flag. Por ejemplo, se añade la secuencia de FLAG™ 5' gat tac aaggat gac gac gat aag 3' al cebador de clonación en el

extremo 3' de la ORF. Se realizan construcciones pSR α adicionales para producir proteínas de fusión con GFP y myc/6X His tanto amino terminales como carboxilo terminales de las proteínas de longitud completa 238P1B2.

Vectores víricos adicionales: Se realizan construcciones adicionales para suministro mediado por virus y expresión de 238P1B2. Se consigue alta titulación viral que conduce a alto nivel de expresión de 238P1B2 en sistemas de suministro víricos tales como vectores adenovíricos y vectores de amplición de herpes. Una secuencia codificante de 238P1B2 o fragmentos de la misma se amplifica por PCR y se subclona en el vector lanzadera AdEasy (Stratagene). Se realizan recombinación y empaquetamiento de virus de acuerdo con las instrucciones del fabricante para generar vectores adenovíricos. Como alternativa, las secuencias codificantes de 238P1B2 o fragmentos de las mismas se clonan en el vector HSV-1 (Imgenex) para generar vectores víricos de herpes. Los vectores víricos se usan a continuación para infección de diversas líneas celulares tales como células PC3, NIH 3T3, 293 o rat-1.

Sistemas de expresión regulados: para controlar la expresión de 238P1B2 en células de mamífero, se clonan secuencias codificantes de 238P1B2, o partes de las mismas, en sistemas de expresión de mamíferos regulados tales como el sistema T-Rex (Invitrogen), el sistema GeneSwitch (Invitrogen) y el sistema estrechamente regulado Ecdysone (Stratagene). Estos sistemas permiten el estudio de los efectos dependientes de concentración y temporales de 238P1B2 recombinante. Estos vectores se usan a continuación para controlar la expresión de 238P1B2 en diversas líneas celulares tales como células PC3, NIH 3T3, 293 o rat-1.

B. Sistemas de expresión de baculovirus

Para generar proteínas 238P1B2 recombinantes en un sistema de expresión de baculovirus, se clonan ORF de 238P1B2 OF, o partes de la misma, en el vector de transferencia de baculovirus pBlueBac 4.5 (Invitrogen), que proporciona un marcador His en el extremo N-terminal. Específicamente, pBlueBac-238P1B2 se co-transfecta con el plásmido auxiliar pBac-N-Blue (Invitrogen) en células de insecto SF9 (*Spodoptera frugiperda*) para generar baculovirus recombinante (véase el manual de instrucciones de Invitrogen para detalles). Se recoge después baculovirus del sobrenadante celular y se purifica por ensayo en placa.

Se genera después proteína 238P1B2 recombinante por infección de células de insecto HighFive (Invitrogen) con baculovirus purificado. Puede detectarse proteína 238P1B2 recombinante usando anticuerpo anti-238P1B2 o anti marcador His. La proteína 238P1B2 puede purificarse y usarse en diversos ensayos basados en células o como inmunógeno para generar anticuerpo policlonales y monoclonales específicos para 238P1B2.

Ejemplo 9 perfiles de antigenicidad y estructura secundaria

Figura 5A,B, Figura 6 A,B, Figura 7 A,B, Figura 8 A,B, y Figura 9 A,B representan gráficamente cinco perfiles de aminoácidos de la secuencia de aminoácidos de 238P1B2 variante 1a (SA-9a) y variante 1b (5B-9B), cada evaluación disponible accediendo al sitio web de ProtScale (URL www.expasy.ch/cgi-ben/protscale.pl) en el servidor de biología molecular ExPasy.

Estos perfiles: Figura 5, hidrofilia (Hopp T.P., Woods K.R., 1981. Proc. Natl. Acad. Sci. U.S.A. 78: 3824-3828); Figura 6, hidropático, (Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157:105-132); Figura 7, porcentaje de restos accesibles (Janen J., 1979 Nature 277:491-492); Figura 8, flexibilidad media (Bhaskaran R., y Ponnuswamy P.K., 1988. Ent. J. Pept. Proteen Res. 32:242-255); Figura 9, giro beta (Deleage, G., Roux B. 1987 Proteen Engeneereng 1:289-294); y opcionalmente otros disponibles en la técnica, tales como en el sitio web de ProtScale, se usaron para identificar regiones antigénicas de la proteína 238P1B2. Cada uno de los perfiles aminoacídicos anteriores de 238P1B2 se generaron usando los siguientes parámetros de ProtScale para análisis: 1) un tamaño de ventana de 9; 2) 100 % de peso de los bordes de ventana en comparación con el centro de ventana; y 3) valores del perfil de aminoácidos normalizados para quedar entre 0 y 1.

Se usaron perfiles de hidrofilia (Figura 5), hidropático (Figura 6) y porcentaje de restos accesibles (Figura 7) para determinar tramos de aminoácidos hidrófilos (es decir, valores mayores de 0,5 en el perfil de hidrofilia y porcentaje de restos accesibles y valores menores de 0,5 en el perfil hidropático). Tales regiones probablemente se exponen al ambiente acuoso, están presentes en la superficie de la proteína y por lo tanto están disponibles para reconocimiento inmune, tal como por anticuerpos.

Los perfiles de flexibilidad media (Figura 8) y giro beta (Figura 9) determinan tramos de aminoácidos (es decir valores mayores de 0,5 en el perfil de giro beta y el perfil de flexibilidad media) que no se restringen en estructuras secundarias tales como láminas beta y hélices alfa. Tales regiones también se exponen más probablemente en la proteína y son por lo tanto accesibles para reconocimiento inmune, tal como por anticuerpos.

Se usan secuencias antigénicas de la proteína 238P1B2 variante 1a y variante 1b indicadas, por ejemplo, por los perfiles expuestos en la Figura 5 A,B, Figura 6 A,B, Figura 7 A,B, Figura 8 A,B, y/o Figura 9 A,B para preparar inmunógenos, péptidos o ácidos nucleicos que los codifican, para generar anticuerpos anti-238P1B2 terapéuticos y de diagnóstico. El inmunógeno puede ser cualquiera de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50 o más de 50 aminoácidos contiguos o los ácidos nucleicos correspondientes que los codifican, de las proteínas variantes de 238P1B2. En particular, los inmunógenos peptídicos para 238P1B2 variante

1a pueden comprender una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento de número entero hasta 254 que incluya una posición de aminoácido que tenga un valor mayor de 0,5 en el perfil de hidrofiliía en la Figura 5A; una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier aumento de número entero hasta 254 que incluye una posición de aminoácido que tenga un valor menor de 0,5 en el perfil hidropático de la Figura 6A; una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento del número entero hasta 254 que incluya una posición de aminoácido que tenga un valor mayor de 0,5 en el perfil de porcentaje de restos accesibles de la Figura 7A, una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento de número entero hasta 254 que incluya una posición de aminoácido que tenga un valor mayor de 0,5 en el perfil de flexibilidad media en la Figura 8A; y, una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento de número entero hasta 254 que incluya una posición de aminoácido que tenga un valor mayor de 0,5 en el perfil de giro beta de la Figura 9A.

Además, los inmunógenos peptídicos para 238P1B2 variante 1b pueden comprender una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento de número entero hasta 316 que incluya una posición de aminoácido que tenga un valor mayor de 0,5 en el perfil de hidrofiliía de la Figura 5B; una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento de número entero hasta 316 que incluya una posición de aminoácido que tenga un valor menor de 0,5 en el perfil hidropático de la Figura 6B; una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento del número entero hasta 316 que incluya una posición de aminoácido que tenga un valor mayor de 0,5 en el perfil de porcentaje de restos accesibles de la Figura 7B; una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento de número entero hasta 316 que incluya una posición de aminoácido que tenga un valor mayor de 0,5 en el perfil de flexibilidad media en la Figura 8B; y una región peptídica de al menos 5 aminoácidos de la Figura 2 en cualquier incremento en el número entero hasta 316 que incluya una posición de aminoácido que tenga un valor mayor de 0,5 en el perfil de giro beta de la Figura 9B. Los inmunógenos también pueden comprender ácidos nucleicos que codifiquen cualquiera de los péptidos anteriores.

Todos los inmunógenos descritos en el presente documento, sean péptidos o ácidos nucleicos, pueden incorporarse en una forma de dosis unitaria humana, o estar comprendidos por una composición que incluya un excipiente farmacéutico compatible con la fisiología humana.

La estructura secundaria de 238P1B2, concretamente la presencia y localización predichas de hélices alfa, hebras extendidas y bucles aleatorios, se predice a partir de la secuencia de aminoácidos primaria usando el procedimiento de red neural jerárquica - HNN (Guermeur, 1997, [dirección web pbil.ibcp.fr/cgi-bin/npsa_automat.p1?page_npsa_nn.html](http://direccion_web_pbil.ibcp.fr/cgi-bin/npsa_automat.p1?page_npsa_nn.html)), a la que se accede desde el servidor de biología molecular ExPasy ([dirección web www.expasy.ch/tools/](http://direccion_web_www.expasy.ch/tools/)). El análisis indica que 238P1B2 variante 1a está compuesta de hélice alfa 66,54 %, cadena extendida 6,69 % y bucle aleatorio 26,77 % (Figura 12A). El análisis indica que 238P1B2 variante 1b está compuesta de hélice alfa 61,71 %, cadena extendida 8,86 % y bucle aleatorio 29,43 % (Figura 12B).

Se llevó a cabo análisis con respecto a la presencia potencial de dominios transmembrana en 238P1B2 usando diversos algoritmos de predicción transmembrana a los que se accede desde el servidor de biología molecular ExPasy ([dirección web www.expasy.ch/tools/](http://direccion_web_www.expasy.ch/tools/)). Los programas predicen la presencia de múltiples dominios transmembrana en 238P1B2 variante 1a y variante 1b. La mayor probabilidad de topología para la variante 1a es la de una proteína de superficie celular con 6 dominios transmembrana. La mayor probabilidad de topología para la variante 1b es la de una proteína de superficie celular con 7 dominios transmembrana, que es coherente con la de un receptor acoplado a proteína G. Se muestran gráficamente en la Figura 12C y Figura 12D los resultados del análisis de 238P1B2 variante 1a usando los programas de predicción TMpred (Figura 12C) y TMHMM (Figura 12D) que presentan la localización y topología de los 6 dominios transmembrana. Se muestran en la Figura 12E y Figura 12F los resultados de los programas de predicción para 238P1B2 variante 1b que muestran la localización y topología de los 7 dominios transmembrana. Los resultados de cada programa, concretamente los aminoácidos que codifican el dominio transmembrana, se resumen en la Tabla XXII.

Ejemplo 10: comparación de homología de 238P1B2 con secuencias conocidas

El gen de 238P1B2 es homólogo de un gen clonado y secuenciado, concretamente el receptor olfatorio de ratón MOR14-1 (gi 18479244) (Zhang X, Firestein S. Nat Neurosci. 2002, 5:124), que muestra 83 % de identidad y 90 % de homología con ese producto génico (Figura 4A; Figura 4C). La proteína 238P1B2 muestra el 78 % de identidad y el 87 % de homología con otro receptor olfatorio de ratón, concretamente MOR14-10 (gi 18480766). El homólogo humano más cercano a 238P1B2 es el receptor olfatorio humano 5BETA12 (5112, gi 17456801), con 61 % de identidad y 79 % de homología (Figura 4E). La comparación de 238P1B2 con otro miembro de la familia del receptor olfatorio humano, concretamente 101P3A11 (patente de referencia AGS), la proteína 238P1B2 muestra 48 % de identidad y 70 % de homología con 101P3A11, alineándose el primer aminoácido de 238P1B2 con el aminoácido 62 de 101P3A11. La proteína 238P1B2 variante 1A consiste en 254 aminoácidos, con un peso molecular calculado de 28,5kDa y pl de 9,2. 238P1B2 es una proteína de superficie celular con algo de localización en la mitocondria y retículo endoplásmico. Se han identificado tres formas de la proteína 238P1B2, conteniendo la variante 238P1B2 V2 una mutación puntual de isoleucina a treonina en el aminoácido 225 y conteniendo la variante 238P1B2 V1B 62 aminoácidos adicionales en su extremo amino terminal. (Figura 4F). Aunque se ha proyectado que 238P1B2 V1A tiene 6 dominios transmembrana, 238P1B2 V1B contiene 7 dominios transmembrana, con el extremo N-terminal

orientado extracelularmente y siendo el extremo C-terminal intracelular (Tabla XXII, Figura 13).

El análisis de motivos reveló la presencia de varios motivos conocidos, incluyendo un motivo GPCR de receptor olfatorio de siete transmembranas y una repetición de tipo III de fibronectina. Las proteínas que son miembros de la familia del receptor acoplado a proteína G muestran un extremo amino terminal extracelular, tres bucles extracelulares, tres bucles intracelulares y un extremo carboxilo terminal intracelular. Los receptores acoplados a proteína G son receptores de siete transmembranas que se estimulan por hormonas polipeptídicas, neurotransmisores, quimiocinas y fosfolípidos (Civelli O y col, Trends Neurosci. 2001, 24:230; Vrecl M y col., Mol Endocrinol. 1998, 12:1818). La unión de ligandos se produce tradicionalmente entre el primer y segundo bucles extracelulares del GPCR. Tras la unión de ligando los GPCR transducen señales a través de la membrana de superficie celular asociándose con proteínas G triméricas. Sus señales se transmiten mediante proteínas de unión a nucleótido de guanina triméricas (proteínas G) a receptor de superficie celular, enzimas efectoras o canales iónicos (Simon y col., 1991, Science 252: 802). La transducción de señal y salida biológica mediada por GPCR puede modularse a través de diversos mecanismos incluyendo miméticos peptídicos, moléculas pequeñas inhibitoras y quinasas de GPCR o GRK (Pitcher JA y col, J Biol Chem. 1999, 3; 274:34531; Fawzi AB, y col. 2001, Mol. Pharmacol., 59:30).

Recientemente, también se ha mostrado que los GPCR se unen a rutas de señalización mitogénicas de tirosina quinasas (Luttrell y col., 1999, Science 283: 655; Luttrell y col., 1999 Curr Opin Cell Biol 11: 177). Los GPCR se regulan por fosforilación mediada por GPCR quinasas (GRK), que a su vez se activan indirectamente por los GPCR (Pitcher y col., 1998, Ann. Rev. Biochem. 67: 653). Los GPCR olfatorios transmiten sus señales activando la ruta de AMPc mediante adenilato ciclasa dando como resultado la señalización corriente abajo de proteína quinasa A y activando la ruta de fosfolipasa C generando inositol 1,4,5-trisfosfato (IP3) y diacil-glicerol (DAG) (Breer, 1993, Ciba Found Symp 179: 97; Bruch, 1996, Comp Biochem Physiol B Biochem Mol Biol 113:451). IP3 da como resultado un aumento del calcio intracelular, mientras que DAG activa la proteína quinasa C.

Recientes estudios han asociado los GPCR con la transformación celular. En particular, se descubrió que el receptor acoplado a proteína G KSHV transforma células NIH 3T3 *in vitro* e induce lesiones de tipo KS multifocales en ratones transgénicos KSHV-GPCR (Schwarz M, Murphy PM. J Immunol 2001, 167:505). KSHV-GPCR fue capaz de producir su efecto en células endoteliales y fibroblastos activando rutas de señalización definidas, incluyendo la ruta de supervivencia AKT (Montaner S y col, Cancer Res 2001, 61:2641). Además, KSHV-GPCR indujo la activación de rutas mitogénicas tales como AP-1 y NFkB, dando como resultado la expresión de genes proinflamatorios (Schwarz M, Murphy PM. J Immunol 2001, 167:505). Otros GPCR asociados con formación de tumores incluyen G2A y PAR-1, que se ha descubierto que inducen la transformación de células NIH 3T3 (Whitehead IP y col, Oncogene 2001, 20:1547).

Las regiones de repeticiones de fibronectina son motivos que median la unión a diversas sustancias tales como receptores de fibronectina, heparina, colágeno y fibrina en superficies celulares. Debido a su capacidad de unión las fibronectinas están implicadas en adhesión celular, diferenciación celular, proliferación, migración y metástasis tumoral (Nykqvist P y col, J Biol Chem 2001, 276:38673; Danen EH, Yamada KM. J Cell Physiol 2001, 189:1; Nabeshima K y col, Histol Histopathol 1999, 14:1183). Además, la fibronectina potencia la angiogénesis y formación de vasos sanguíneos *de novo* regulando la migración de células endoteliales, que constituyen componentes esenciales de los vasos sanguíneos (Urbich C y col, Arterioscler Thromb Vasc Biol 2002, 22:69), potenciando de este modo el crecimiento y supervivencia tumoral.

Esta información indica que 238P1B2 desempeña un papel en la transformación de células de mamífero, induce respuestas mitogénicas incluyendo la activación de diversas rutas de señalización y regula la transcripción génica transmitiendo señales de superficie celular al núcleo. En consecuencia, cuando 238P1B2 actúa como un regulador de la transformación celular, formación de tumores o como un modulador de la transcripción implicado en la activación de genes asociados con inflamación, tumorigénesis o proliferación, 238P1B2 se usa para fines terapéuticos, de diagnóstico, de pronóstico y/o preventivos. Además, cuando una molécula, tal como una variante o polimorfismo de 238P1B2 se expresa en tejidos cancerosos, se usa para fines terapéuticos, de diagnóstico, de pronóstico y/o preventivos.

Ejemplo 11: Identificación y confirmación de rutas de transducción de señales potenciales

Se ha indicado que muchas proteínas de mamífero interactúan con moléculas de señalización y participan en la regulación de rutas de señalización (J Neurochem 2001; 76:217-223). En particular, se ha indicado que los GPCR activan cascadas de MAK así como proteínas G y se han asociado con la ruta EGFR en células epiteliales (Naor, Z., y col, Trends Endocrinol Metab. 2000, 11:91; Vacca F y col, Cancer Res. 2000, 60:5310; Della Rocca GJ y col, J Biol Chem. 1999, 274:13978). Usando técnicas de inmunoprecipitación y transferencia de Western, se identifican proteínas que se asocian con 238P1B2 y median en acontecimientos de señalización. Pueden regularse varias rutas que se sabe que desempeñan un papel en la biología del cáncer por 238P1B2, incluyendo rutas de fosfolípidos tales como PI3K, AKT, etc.; rutas de adhesión y migración, incluyendo FAK, Rho, Rac-1, etc.; y cascadas de supervivencia/mitogénicas tales como ERK, p38, etc. (Cell Growth Differ. 2000,11:279; J BiolChem 1999, 274:801; Oncogene. 2000, 19:3003, J. Cell Biol. 1997, 138:913).

Se ha mostrado que varios GPCR transactivan tirosina quinasas receptoras asociadas con la membrana celular, tales como el receptor de EGF (EGFR) (Pierce KL y col, J Biol Chem. 2001, 276:23155; Nath D y col, J Cell Sci. 2001, 114: 1213). Para determinar si la señalización de 238P1B2 da como resultado la activación de EGFR, se dejan crecer células en medio solamente o en presencia del inhibidor de EGFR AG1517. La fosforilación de EGFR se compara en células tratadas y de control. De forma similar, se investiga la comunicación entre las rutas de 238P1B2 y EGFR.

Para confirmar que 238P1B2 activa directa o indirectamente rutas de transducción de señal conocidas en las células, se llevan a cabo ensayos indicadores transcripcionales basados en luciferasa (luc) en células que expresan genes individuales. Estos indicadores transcripcionales contienen sitios de unión a consenso para factores de transcripción conocidos que quedan corriente abajo de rutas de transducción de señal bien caracterizadas. Los indicadores y ejemplos de estos factores de transcripción asociados, rutas de transducción de señal y estímulos de activación se enumeran a continuación

1. NFκB-luc, NFκB/Rel; Ik-quinasa/SAPK; crecimiento/apoptosis/tensión
2. SRE-luc, SRF/TCF/ELK1; MAPK/SAPK; crecimiento/diferenciación
3. AP-1-luc, FOS/JUN; MAPK/SAPK/PKC; crecimiento/apoptosis/tensión
4. ARE-luc, receptor de andrógenos, esteroides/MAPK; crecimiento/diferenciación/apoptosis
5. p53-luc, p53; SAPK; crecimiento/diferenciación/apoptosis
6. CRE-luc, CREB/ATF2; PKA/p38; crecimiento/apoptosis/tensión

Los efectos mediados por gen pueden ensayarse en células que muestran expresión de ARNm. Pueden introducirse plásmidos indicadores de luciferasa por transfección mediada por lípidos (TFX-50, Promega). La actividad luciferasa, un indicador de actividad transcripcional relativa, se mide por incubación de extractos celulares con sustrato de luciferina y la luminiscencia de la reacción se controla en un luminómetro.

Las rutas de señalización activadas por 238P1B2 se mapean y se usan para la identificación y validación de dianas terapéuticas. Cuando 238P1B2 está implicada en la señalización celular, se usa como diana para fines de diagnóstico, de pronóstico, preventivos y/o terapéuticos.

Ejemplo 12: 238P1B2 actúa como un GPCR

El análisis de secuencia y homología de 238P1B2 indica que la proteína 238P1B2 es un miembro de la familia de los receptores olfatorios de GPCR. Se sabe que los receptores olfatorios regulan respuestas biológicas activando adenilato ciclasa. Para confirmar que 238P1B2 actúa como GPCR y media en la activación de adenilato ciclasa, se comprueba la acumulación de AMPc en células PC3 y PC3-238P1B2 en células cultivadas en presencia o ausencia de suero bovino fetal (FBS). Las células se lisan y la concentración intracelular de AMPc se mide usando un inmunoensayo enzimático (EIA) disponible en el mercado. Los cálculos de concentraciones de AMPc se basaron en DO450 de la curva patrón. De forma similar, el mismo ensayo puede usarse para determinar si la inducción de acumulación de AMPc por 238P1B2 se inhibe por inhibidores de GPCR tal como toxina pertussis.

Los GPCR transmiten su señal activando proteínas G triméricas. Una vez que los GPCR se activan, la subunidad Ga asociada se une a GTP, se disocia del receptor y participa en acontecimientos de señalización corriente abajo (Schild D y Restrepo D. Physiol Rev. 1998, 78:429-66). Para determinar que la inhibición de las subunidades Ga tiene un efecto en el crecimiento celular mediado por 238P1B2, se investiga el efecto de los inhibidores de Ga en la proliferación de células 3T3-238P1B2 y PC3-238P1B2. Las células de control y que expresan 238P1B2 se dejan crecer en presencia o ausencia de suramina o su derivado NF 449 (Sigma). Las células se analizan con respecto a proliferación usando un ensayo de tipo MTT.

Cuando 238P1B2 actúa como un GPCR, se usa como diana para fines de diagnóstico, pronóstico, preventivos y/o terapéuticos.

Ejemplo 13: Implicación en la progresión tumoral

El gen 238P1B2 puede contribuir al crecimiento de células cancerosas. El papel de 238P1B2 en el crecimiento tumoral se confirma en diversas líneas celulares primarias y transfectadas incluyendo próstata así como células NIH 3T3 modificadas por ingeniería genética para expresar de forma estable 238P1B2. Se evalúan células parentales sin 238P1B2 y células que expresan 238P1B2 con respecto a crecimiento celular usando un ensayo de proliferación bien documentado (Fraser SP, Grimes JA, Djamgoz MB. Prostate. 2000; 44:61, Johnson DE, Ochieng J, Evans SL. Anticancer Drugs. 1996, 7:288). Para confirmar que 238P1B2 media en la proliferación potenciada por medio de su actividad GPCR, se dejan crecer células de control y células que expresan 238P1B2 en presencia o ausencia de toxina pertussis y se evalúan con respecto a su capacidad proliferativa usando el mismo ensayo descrito anteriormente.

Para confirmar el papel de 238P1B2 en el proceso de transformación, se investiga su efecto en los ensayos de formación de colonias. Se comparan células NIH-3T3 parentales sin 238P1B2 con células NIH-3T3 que expresan 238P1B2, usando un ensayo de agar blando en condiciones rigurosas y más permisivas (Song Z. y col., Cancer Res. (2000) 60:6730.

Para confirmar el papel de 238P1B2 en la invasión y metástasis de células cancerosas, se usa un ensayo bien establecido, por ejemplo, un ensayo de sistema de inserto Transwell (Becton Dickinson) (Cancer Res. 1999; 59:6010). Se comparan células de control, incluyendo líneas celulares de próstata y fibroblastos sin 238P1B2 con células que expresan 238P1B2. Las células se cargan con el colorante fluorescente, calceína, y se siembran en placas en el pocillo superior del inserto Transwell revestido con un análogo de membrana basal. La invasión se determina por fluorescencia de células en la cámara inferior en relación con la fluorescencia de la población celular completa.

238P1B2 también puede desempeñar un papel en el ciclo celular y apoptosis. Las células parentales y células que expresan 238P1B2 se comparan con respecto a diferencias en la regulación del ciclo celular usando un ensayo BrdU bien establecido (Abdel-Malek ZA. J Cell Physiol. 1988, 136:247). En resumen, las células cultivadas en condiciones tanto óptimas (suero completo) como limitantes (suero bajo) se marcan con BrdU y se tiñen con Ab anti-BrdU y yoduro de propidio. Las células se analizan con respecto a entrada en las fases G1, S y G2M del ciclo celular. Como alternativa, el efecto de tensión en la apoptosis se evalúa en células parentales de control y células que expresan 238P1B2, incluyendo células de próstata normales y tumorales. Las células modificadas por ingeniería genética y las parentales se tratan con diversos agentes quimioterapéuticos, tales como etopósido, flutamida, etc., e inhibidores de síntesis proteica tales como cicloheximida. Las células se tiñen con anexina V-FITC y la muerte celular se mide por análisis de FACS. La modulación de muerte celular por 238P1B2 puede desempeñar un papel crítico en la regulación de progresión tumoral y carga tumoral.

Cuando 238P1B2 desempeña un papel en el crecimiento celular, transformación, invasión o apoptosis, se usa como una diana para fines de diagnóstico, pronóstico, preventivos y/o terapéuticos.

Ejemplo 14: Implicación en angiogénesis

La angiogénesis o formación de nuevos vasos sanguíneos capilares es necesaria para el crecimiento tumoral (Hanahan D, Folkman J. Cell. 1996, 86:353; Folkman J. Endocrinology. 1998 139:441). Basándose en el efecto de los inhibidores de fosfodiesterasa en células endoteliales, 238P1B2 desempeña un papel en la angiogénesis (DeFouw L y col, Microvasc Res 2001, 62:263). Se han desarrollado varios ensayos para medir angiogénesis *in vitro* e *in vivo*, tales como los ensayos de cultivo tisular de formación de tubo celular endotelial y proliferación de células endoteliales. Usando estos ensayos así como neovascularización *in vitro*, se confirma el papel de 238P1B2 en angiogénesis, potenciación o inhibición.

Por ejemplo, las células endoteliales modificadas por ingeniería genética para expresar 238P1B2 se evalúan usando ensayos de formación de tubos y proliferación. El efecto de 238P1B2 también se confirma en modelos animales *in vivo*. Por ejemplo, se implantan células que expresan o carecen de 238P1B2 de forma subcutánea en ratones inmunocomprometidos. La migración celular endotelial y angiogénesis se evalúa 5-15 días después usando técnicas de inmunohistoquímica: 238P1B2 afecta a la angiogénesis y se usa como una diana para fines de diagnóstico, pronóstico, preventivos y/o terapéuticos.

Ejemplo 15: Regulación de la transcripción

La localización de superficie celular de 238P1B2 y su similitud con GPCR indica que 238P1B2 se usa eficazmente como un modulador de la regulación transcripcional de genes eucariotas. La regulación de la expresión génica se confirma, por ejemplo, estudiando expresión génica en células que expresan o carecen de 238P1B2. Para este fin, se realizan dos tipos de experimentos.

En el primer conjunto de experimentos, se extrae ARN de células parentales y que expresan 238P1B2 y se hibrida con series génicas disponibles en el mercado (Clontech) (Smid-Koopman E y col. Br J Cancer. 2000. 83:246). Se comparan células en reposo así como células tratadas con FBS o andrógeno. Se identifican genes expresados diferencialmente de acuerdo con procedimientos conocidos en la técnica. Los genes expresados diferencialmente se mapean después en rutas biológicas (Chen K y col. Thyroid. 2001. 11:41.).

En el segundo conjunto de experimentos, se evalúa la activación de ruta transcripcional específica usando construcciones indicadoras de luciferasa disponibles en el mercado (Stratagene) incluyendo: NtKB-luc, SRE-luc, ELK1-luc, ARE-luc, p53-luc y CRE-luc. Estos indicadores transcripcionales contienen sitios de unión consenso para factores de transcripción conocidos que quedan corriente abajo de las rutas de transducción de señal bien caracterizadas y representan una buena herramienta para determinar la activación de la ruta y explorar con respecto a moduladores positivos y negativos de la activación de la ruta.

Por lo tanto, 238P1B2 desempeña un papel en la regulación génica y se usa como una diana para fines de diagnóstico, pronóstico, preventivos y/o terapéuticos.

Ejemplo 16. Implicación en adhesión celular.

La adhesión celular desempeña un papel crítico en la colonización tisular y metástasis. Basándose en la presencia de una repetición de fibronectina en su extremo C-terminal, 238P1B2 puede participar en la organización celular y como consecuencia afecta a la adhesión y motilidad celular. Para confirmar que 238P1B2 regula la adhesión celular,

se comparan células control sin 238P1B2 con células que expresan 238P1B2, usando técnicas previamente descritas (véase, por ejemplo, Haier y col, Br. J. Cancer. 1999, 80:1867; Lehr y Pienta, J. Natl. Cancer Inst. 1998, 90:118). Brevemente, se incuban células marcadas con un indicador fluorescente, tales como calceína, en pocillos de cultivo tisular revestidos con medio solamente o con proteínas de la matriz. Las células adherentes se detectan por análisis fluorométrico y se calcula el porcentaje de adhesión. Como alternativa, se analizan células que carecen de o que expresan 238P1B2 con respecto a su capacidad para mediar en la adhesión célula-célula usando técnicas experimentales similares como se ha descrito anteriormente. Ambos de estos sistemas experimentales se usan para identificar proteínas, anticuerpos y/o moléculas pequeñas que modulan la adhesión celular a matriz extracelular e interacción célula-célula. La adhesión celular desempeña un papel crítico en el crecimiento, progresión y colonización tumoral y 238P1B2 está implicada en estos procesos. Por lo tanto, actúa como una modalidad de diagnóstico, pronóstico, preventiva y/o terapéutica.

Ejemplo 17: Asociación proteína-proteína

Se ha mostrado que varios GPCR interaccionan con otras proteínas, regulando de este modo la transducción de señales, transcripción génica, transformación y adhesión celular (Sexton PM y col, Cell Signal. 2001, 13:73; Turner CE, J Cell Sci. 2000, 23:4139). Usando técnicas de inmunoprecipitación así como dos sistemas de híbrido de levadura, se identifican proteínas que se asocian con 238P1B2. Se comparan inmunoprecipitados de células que expresan 238P1B2 y células sin 238P1B2 con respecto a asociaciones proteína-proteína específicas.

Se realizan estudios para confirmar el alcance de la asociación de 238P1B2 con moléculas efectoras, tales como proteínas nucleares, factores de transcripción, quinasas, fosfatos, etc. Los estudios que comparan células positivas para 238P1B2 y negativas para 238P1B2 así como estudios que comparan células no estimuladas/en reposo y células tratadas con activadores de células epiteliales, tales como citocinas, factores de crecimiento, andrógeno y Ab anti-integrina revelan interacciones únicas.

Además, se confirman las interacciones proteína-proteína usando metodología de dos híbridos de levadura (Curr Opin Chem Biol. 1999, 3:64). Se introduce un vector que porta una biblioteca de proteínas fusionadas con el dominio de activación de un factor de transcripción en levadura que expresa una proteína de fusión de dominio de unión ADN-238P1B2 y una construcción indicadora. La interacción proteína-proteína se detecta por actividad indicadora colorimétrica. La asociación específica con moléculas efectoras y factores de transcripción dirige a un experto en la material al modo de acción de 238P1B2 e identifica de este modo dianas terapéuticas, de pronóstico, preventivas y/o de diagnóstico para cáncer. Este y otros ensayos similares también se usan para identificar y explorar con respecto a moléculas pequeñas que interaccionan con 238P1B2.

Por lo tanto se descubre que 238P1B2 se asocia con proteína y moléculas pequeñas. En consecuencia, 238P1B2 y estas proteínas y moléculas pequeñas se usan para fines de diagnóstico, pronóstico, preventivos y/o terapéuticos.

TABLAS

TABLA I: Tejidos que expresan 238P1B2 cuando son tumorales
Próstata

TABLA II: ABREVIATURAS DE AMINOÁCIDOS

UNA LETRA	TRES LETRAS	NOMBRE COMPLETO
F	Phe	fenilalanina
L	Leu	leucina
s	Ser	serina
Y	Tyr	tirosina
C	Cys	cisteína
w	Trp	triptófano
P	Pro	prolina
H	His	histidina
Q	Gln	glutamina
R	Arg	arginina
I	Ile	isoleucina
M	Met	metionina
T	Thr	treonina

(continuación)

UNA LETRA	TRES LETRAS	NOMBRE COMPLETO
N	Asn	asparagina
K	Lys	lisina
V	Val	valina
A	Ala	alanina
D	Asp	ácido aspártico
E	Glu	ácido glutámico
G	Gly	glicina

TABLA III: MATRIZ DE SUSTITUCIÓN DE AMINOÁCIDOS

5 Adaptada de la matriz de sustitución de aminoácidos de GCG Software 9.0 BLOSUM62 (matriz de sustitución en bloque). Cuanto mayor sea el valor, más probable es que se encuentre una sustitución en proteínas relacionadas naturales. (Véase la dirección web www.ikp.unibe.ch/manual/blosum62.html)

A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	.
4	0	-	-	-2	0	-2	-1	-1	-1	-1	-2	-1	-1	-1	1	0	0	-3	-2	A
		2	1																	
	9	-	-	-2	-3	-3	-1	-3	-1	-1	-3	-3	-3	-3	-1	-1	-1	-2	-2	C
		3	4																	
		6	2	-3	-1	-1	-3	-1	-4	-3	1	-1	0	-2	0	-1	-3	-4	-3	D
			5	-3	-2	0	-3	1	-3	-2	0	-1	2	0	0	-1	-2	-3	-2	E
				6	-3	-1	0	-3	0	0	-3	-4	-3	-3	-2	-2	-1	1	3	F
					6	-2	-4	-2	-4	-3	0	-2	-2	-2	0	-2	-3	-2	-3	G
						8	-3	-1	-3	-2	1	-2	0	0	-1	-2	-3	-2	2	H
							4	-3	2	1	-3	-3	-3	-3	-2	-1	3	-3	-1	I
								5	-2	-1	0	-1	1	2	0	-1	-2	-3	-2	K
									4	2	-3	-3	-2	-2	-2	-1	1	-2	-1	L
										5	-2	-2	0	-1	-1	-1	1	-1	-1	M
											6	-2	0	0	1		-3	-4	-2	N
												7	-1	-2	-1	-1	-2	-4	-3	P
													5	1	0	-1	-2	-2	-1	Q
														5	-1	-1	-3	-3	-2	R
															4	1	-2	-3	-2	S
																5	0	-2	-2	T
																	4	-3	-1	V
																		11	2	W
																			7	Y

Tabla XXI: Motivos y Modificaciones Post-traduccionales de 238P1B2

- 10 Sitio de fosforilación de proteína quinasa dependiente de AMPc y GMPc.
176 - 179 RKeT
- Sitio de fosforilación de proteína quinasa C.
235 - 237 SvK
- Sitio de fosforilación de caseína quinasa II.
9 - 12 SatD
- 15 50 - 53 TvmE
130 - 133 SctD
172 - 175 SpeC
- Sitio de N-miristilación.
14 - 19 GLsiST
- 20 Identificación de familia de receptores acoplados a proteína G 1.
52 - 68 MESsvLlaMAFDRFvaV
Familia de receptores acoplados a proteína G 1.

1-234

TABLA XXII: Características Físicas de 238P1B2

238P1B2 Variante 1A	Programa Bioinformático	URL	Resultado
ORF	ORF finder		3758 pb
Longitud de la proteína			254 aa
Región transmembrana	TM Pred	Dirección web www.ch.embnet.org/	Extremo N terminal intracelular, 6 TM hélices TM en 3-29, 44-62, 86-110, 144-161, 182-203, 217-236aa
	HMMTop	Dirección web www.enzim.hu/hmmtop/	Extremo N terminal intracelular, 6TM hélices TM en 6-28, 43-62, 86-105, 136-158, 180-203, 216-235aa
	Sosui	Dirección web www.genome.ad.jp/SOSui/	Proteína de membrana, 6TM hélices TM en 6-28, 42-64, 87-109, 144-166, 187-209, 217-237aa
	TMHMM	Dirección web www.cbs.dtu.dk/services/TMHMM	Extremo N terminal intracelular, 6TM hélices TM en 7-29, 44-66, 86-108, 142-164, 185-207, 212-234aa
Péptido señal	Signal P	Dirección web www.cbs.dtu.dk/services/SignalP/	Sitio de escisión entre 169-170aa
pl	pl/MW tool	Dirección web www.expasy.ch/tools/	pl 9,24
Peso molecular	pl/MW tool	Dirección web www.expasy.ch/tools/	28,59 kDa
Localización	PSORT	Dirección web psort.nibb.ac.jp/	60 % membrana plasmática, 42,9 % membrana interna mitocondrial, 40 % Golgi, 30 % membrana de retículo endoplásmico
	PSORT II	Dirección web psort.nibb.ac.jp/	44,4 % retículo endoplásmico, 22,2 % membrana plasmática
Motivos	Pfam	Dirección web www.sanger.ac.uk/Pfam/	Receptor 7TM (familia de rodopsina)
	Prints	Dirección web www.biochemucl.ac.uk/	Identificación de receptor olfatorio, proteína R de membrana interna del sistema de secreción de Tipo III, identificación de repetición de tipo III de fibronectina
	Blocks	Dirección web www.blocks.thrcr.org/	Sin motivo significativo

ES 2 375 351 T3

(continuación)

238P1B2 Variante 1A	Programa Bioinformático	URL	Resultado
238P1B2 Variante 1B	Programa Bioinformático	URL	Resultado
ORF	ORF finder		3758 pb
Longitud de la proteína			316 aa
Región transmembrana	TM Pred	Dirección web www.ch.embnet.org/	Extremo N terminal extracelular, 8TM hélices TM en 9-27,33-51, 75-91, 98-125, 145-171, 206-226, 249-265, 276-296aa
	HMMTop	Dirección web www.enzim.hu/hmmtop/	Extremo N terminal extracelular, 7TM hélices TM en 29-53, 66-90, 105-124, 148-171, 202-226, 241-265, 278-297aa
	Sosui	Dirección web www.genome.ad.jp/SOSui/	Proteína de membrana, 8TM hélices TM en 2-24, 33-55, 67-89, 104-126, 149-171, 206-228,245-267,279-299aa
	TMHMM	Dirección web www.cbs.dtu.dk/services/TMHMM	Extremo N terminal intracelular, 7TM hélices TM en 30-52, 65-87, 102-124,145-167,204-226, 247-269, 274-296aa
Péptido señal	Signal P	Dirección web www.cbs.dtu.dk/services/SignalP/	Sitio de escisión entre aa 26 y 27
pl	pl/MW tool	Dirección web www.expasy.ch/tools/	pl 9,03
Peso molecular	pl/MW tool	Dirección web www.expasy.ch/tools/	35,34 kDa
Localización	PSORT	Dirección web psort.nibb.ac.jp/	64 % de membrana plasmática, 46 % Golgi, 37 % membrana del retículo endoplásmico
	PSORT II	Dirección web psort.nibb.ac.jp/	44,4 % retículo endoplásmico, 33,3 % membrana plasmática
Motivos	Pfam	Dirección web www.sanger.ac.uk/Pfam/	Receptor de 7TM (familia de rodopsina), proteína de biosíntesis de polisacáridos
	Prints	Dirección web www.biochem.ucl.ac.uk/	Identificación de receptor olfatorio, proteína R de membrana interna del sistema de secreción de tipo III, identificación de repetición de tipo III de fibronectina
	Blocks	Dirección web www.blocks.fhcrc.org/	Sin motivo significativo

Tabla XXIII. Composiciones Exónicas de 238P1B2 v.1

Número de exones	Comienzo	Fin
Exón 1	1	3758

REIVINDICACIONES

1. Un procedimiento para detectar la presencia de cáncer de próstata en una muestra de ensayo, comprendiendo el procedimiento:

- 5 poner en contacto la muestra con una sonda que se une específicamente a un polinucleótido de la Figura 2B; y detectar la unión del polinucleótido de la muestra con la misma.

Figura 1 Secuencia SSH de 238P1B2 de 210 nucleótidos

```
1 GATCAATACA CAAATGAATG TAGCTTTTTT TGTTTAAGCT CTCATCTCGT TCTGIACITTT
61 TAATTCATCA CGTCAACTTT GCAGAATATC ACCTCCACTT CCATCATTTT CCTGCTCACT
121 GGTGTTCTG GGTGGAAGC CTTCCACACC TGGATCTCCA TTCCTTCTG CTCCTCTCT
181 GTAAGTCTC TCTTGGGAAA CAGCCTGATC
```

Figura 2A. La secuencia de ADNc y de aminoácidos de 238P1B2 v.1 A clon A. La metionina de partida está subrayada. El marco de lectura abierto se extiende desde el ácido nucleico 2133-2897 incluyendo el codón de parada.

```

1  ccgccttccagggtcactccattctcctgcctcagcctcccagtagctggaactacagg
61  cgcccgccaccacgcccggctaatttttttgatatttttagtggagacgaggtttcaccatg
121 ttggccaggatggtctcgatctcctgacctcgtgatccgcccacctcggcctcccaaagt
181 gctgggactacaggcgtgagccaccgcgccagcctgctcaacacttttacacatgccgt
241 tctgaaagcctgggaattcttcgtctcaacttatataaaacacagaaatctttcaagcac
301 tgctttttctagaaagccttctccaaatactacaaaacctactgaatgctctctttctggg
361 aattccttttttagtggtgttttgggttttcaaaatcattttatcttatttctagttgacaa
421 ataatatggtatataatttgtgggtaaaatggtgatgttgttatacatgatgcatgtggtg
481 aatgaccaaataagctaattatcatatccatcacctcaaatatttatcatttctttgtg
541 ctgtgaacatttaaaattctctcttttagttttttgaaatatactatacattaactata
601 attcttatttctctaagaggtttgatcatacatttttagatattttcatcactcgcctgag
661 gacactgtcattactgtaaatttcagaatctaaagctatagtttcattatttggccaat
721 cgtgtaatctataatttcttgtttttcatctagttttctgttaattttagttaaatcaca
781 cccccaccccccttttttttctactaagacttcttcgggtgatccttgagggtgacactt
841 tatcagggtgaatgtgtgagctcggatcacgttgagggtcctctccatcctctcttttgtc
901 aacacttactgtcagccaccagcaccttctcttctctgaagaagtaccctgcaac
961 tggctctctctcattgatgagccctctgcagccacaaccctaatactctgtttttcccta
1021 gacccccgtcttcaactcaaccatacccatctatctacctttctgcaattctccactctc
1081 tgaaaaaagaaatggtgatgtttgcaataaagttatattttccctccgaaacaaaacagt
1141 aacattcaattaggtatctcctaagaagtatgtcttaacaaaatagagagattctgttct
1201 ggagaactagtcttccagcagactgaagagagaactgtgttggaacgtggcagcaccatta
1261 taccaggtggggcagcctcctgtcttccctaagcctgggtgggaggtctgtgttctaacca
1321 tctcttccaaatatgatgccgggtatcatattgtaagggtttctttaatggttagatgct
1381 gaagccatttatttgtcacttctatgtgttactttgtattataatgatctctttcttca
1441 tattttctggcttcccaaattgtctgtaagcttctctgcttccacacttatgaccaatg
1501 taatattgctgttccacagccccagtgatatttttctctgggtgccaaccatagaggaaaaa
1561 taatgcactacctcttttagcacaatttaaacatgccataaagaatgaaatagattggca
1621 aactcagggtaaagtgtttattcttagtttgggtatctatggctagaacgatcagaggag
1681 tggtcacatgatataaacatagctgccagcaaaagggccttagagaagaaagattgtacac
1741 taaggcaaatcaaaaacaaaccaaagtgcccactacagcagttgaaatctgacattttt
1801 taatcaatttaaatgagccattcatagtttcatattatataaaactcatcagttttctg
1861 tacaataaaattattttaattcgatcaatacacaaatgaatgtagctttttttgtttaa
1921 gctctcatctcgttctgtacttttaattcatcacgtcaactttgcagaatatcacctcca
1981 cttccatcatttctctgctcactgggtgttctctgggctggaagcctccacacctggatct
2041 ccattcccttctgcttctctctgtaactgctctcttgggaaacagcctgatcctcttgc

```

1 M Y Y F L S M L S A
 2101 ctaccatcactcagcccagcctccacgaaccaATGTACTATTTCCCTCTCCATGCTGTCGG
 11 T D L G L S I S T L V T M L S I F W F N
 2161 CCACTGACCTCGGCCTGTCCATATCCACTCTGGTCACCATGCTGAGTATATTCTGGTTCA
 31 V R E I S F N A C L S H M F F I K F F T
 2221 ATGTGAGGGAAATCAGCTTTAATGCCTGCTGTGCCACATGTTCTTTATTAAATTCTTCA
 51 V M E S S V L L A M A F D R F V A V S N
 2281 CTGTCATGGAATCCTCAGTGCTGTTGGCCATGGCTTTTGATCGTTTTGTGGCCGCTCTA
 71 P L R Y A M I L T D S R I A Q I G V A S
 2341 ATCCCCTTAGGTATGCCATGATTTTAAGTACTCCAGAATAGCTCAAATTGGAGTGGCAA
 91 V I R G L L M L T P M V A L L I R L S Y
 2401 GTGTCATCAGGGGGCTCCTAATGCTGACACCAATGGTAGCACTTCTTATAAGACTTTCCT
 111 C H S Q V L H H S Y C Y H P D V M K L S
 2461 ACTGCCACAGCCAAGTACTCCACCCTCTACTGCTACCACCCTGATGTGATGAAGCTCT
 131 C T D T R I N S A V G L T A M F S T V G
 2521 CATGCACAGACACCAGAATCAACAGTGCAGTTGGGCTGACTGCCATGTTCTCTACTGTTG
 151 V D L L L I L L S Y V L I I R T V L S V
 2581 GTGTAGACTTACTTCTCATCCTCCTTTCTTATGTTTTGATCATTAGGACTGTCCTTAGCG
 171 A S P E E R K E T F S T C V S H I V A F
 2641 TTGCTTCCCCAGAAGAGAGGAAGGAAACCTTCAGTACATGTGTCTCCACATTGTGGCTT
 191 A I Y Y I P L I S L S I V H R F G K Q A
 2701 TTGCTATATATTACATTCCATTGATCAGTCTGTCCATTGTTACAGATTGGGAAACAAG
 211 P A Y V H T M I A N T Y L L I S P L M N
 2761 CCCCAGCCTATGTACATACTATGATTGCTAACACCTACCTGCTGATCTCCCTTTGATGA
 231 P V I Y S V K T K Q I R R A V I K I L H
 2821 ACCCTGTCTACAGTGTGAAAACCAACAGATACGTAGAGCTGTGATAAAAATTCTCC
 251 S K E T *
 2881 ATTCCAAAGAAACATAAaattgtagaattctagcactagatggaccctcaagatggat
 2941 acttcaggatcatgtttccaattaaaaacaaaaaaccaagaactgggcccattgaa
 3001 gctaagcacataatcatttaaagttaaagctggaatgtatctagaacatcagctagttg
 3061 actactttatttccagatgaggagggtgaagacatgggatggaagagatgaacagcctaa
 3121 aattagccttatataaaaatttctccttctcacaccactgttgggtccaaactgttgtt
 3181 gccttctgtccctgccaacacccttctgttattatctcatctacttcatagcctttaga
 3241 atgacacattacagaaaagtcactcctttgtgagaaataggtagtatttccaactgttc
 3301 tagtaatttcgattgccatttattatttgcctctgacaagtcaggttctaagtggttttt
 3361 gtatgtattaacatattttatatactaaacaatggtataagataatcactttaattatca
 3421 ataagaaaattgaggcatgaaatgtaagtagcttgcttaaacagctagtaactgggtgtt
 3481 gctgatggcagagcacatccatgcagactgactccagagtccccacactcttaatttta
 3541 atgttatattgctacctataatataatataaggtatataatttatatataatttatatt
 3601 tgtaaaaattatgaacttatatataaaacttatattttataaattatgtacttatatg

3661 tgtaaattatatttatattataaaatataaaatagagagaagttcaggcaatcttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaa

3661 tgtaaattatatttatattataaaatataaaatagagagaagttcaggcaatcttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaa

Figura 2B. La secuencia de ADNc y de aminoácidos de 238P1B2 v.1B clon A. El marco de lectura abierto se extiende desde el ácido nucleico 1947-2897 incluyendo el codón de parada.

```

1  cgccttcaggttcactccattctcctgcctcagcctccccagtagctggaactacagg
61  egcccgccaccacgcccggctaattttttgtattttagtgagacgaggttcaccatg
121 ttggccaggatggtctcgatctcctgacctcgtgatccgcccacctcggcctcccaaagt
181 gctgggactacagggctgagccaccgcgcccagcctgctcaacacttttacacatgccgt
241 tctgaaagcctgggaattctctcgtctcaacttatataaaacacagaaatctttcaagcac
301 tgcttttctagaaagccttctccaataactacaaacctactgaatgctctctttctgggt
361 aattccttttttagtggttttgggttttcaaaatcattttatcttattttctagttgacaa
421 ataatatggtatatatttgtggggtaaaaatgtgatggtgttatatacatgtatgcattgtgg
481 aatgaccaaatacaagctaattatcatatccatcacctcaaatatttatcatttctttgtg
541 ctgtgaacatttaaaattctctcttttagttttttgaaatatactatacattaactata
601 attcttatttctcctaagagggttgatcacatcttagatatttcatcactcgcctgag
661 gacactgtcattactgtaaatctcagaatctaaagctatagtttcattattggccaaat
721 cgtgtaatctataatctctgttttcatctagtttctgttaatttagttaaatacaca
781 ccccccccccttttttttcaactaagacttcttccggatccttgagggtgacactt
841 tatcagggtgaatgtgtgagctcggatacgttgagggtcctctccatcctctctttgtc
901 aacacttactgtcagccaccagacccttcttttctctgaagaagttaccctgcaac
961 tggctctctcattgattgagccctctgcagocacaaccctaatactctgtttttcccta
1021 gacccccgtcttcaactcaaccatacccatctatctacctttctgcaattctccactcta
1081 tgaaaaagaaatgtgatggttgcaataaagttatattttccctccgaaacaaaacagt
1141 aacattcaattaggatctcctaagaagtatgtcttaacaaaatagagagattctgttct
1201 ggagaactagtcttccagcagactgaagagagaactgtgttggaacgtggcagcacatta
1261 tatcagggtggggcagcctcctgtcttcttaagcctgggtgggaggtctgtgttctaacca
1321 tctcttccaaatgatgcccgggtatcatattgtaagggttctttaaaggttagatgct
1381 gaagccatttatttgtcacttctatgtgttactttgtattataatgatctctttcttca
1441 tattttctggettcccaaattgtctgtaagcttctctgctttccacacttatgaccaatg
1501 taatattgctgttccacagccccagtgatttttctctggtgccaaccatagaggaaaaa
1561 taatgcactacctctttagcacaatttaaacatgccataaagaatgaaatagattggca
1621 aactcagggtaaaagtgtttattcttagtttggttatctatggctagaacgatcagaggag
1681 tggtcacatgatataaacatagctgccagcaaagggccctagagaagaagattgtacac
1741 taaggcaaatcaaaaacaaaccaaagtgcccactacagcagttgaaatctgacattttt
1801 taatcaatttaaagagccattcatagtttcatattatataaaatactcatcagttttctg
1861 tacaataaaattattttaattcgatcaatacacaatgaatgtagctttttttgtttaa
1          F I T S T L Q N I T S T
1921 gctctcatctcgttctgtacttttaaTTCATCACGTCAACTTTGCAGAATATCACCTCCA
13  S I I F L L T G V P G L E A F H T W I S
1981 CTTCATCATTTTTCTGCTCACTGGTGTTCCTGGGCTGGAAGCCTTCCACACCTGGATCT

```

33 I P F C F L S V T A L L G N S L I L F A
 2041 CCATTCCCTTCTGCTTCCTCTCTGTAACGCTCTCTTGGGAAACAGCCTGATCCTCTTCG
 53 T I T Q P S L H E P M Y Y F L S M L S A
 2101 CTACCATCACTCAGCCCAGCCTCCACGAACCAATGTACTATTTCCCTCCTCCATGCTGTCCG
 73 T D L G L S I S T L V T M L S I F W F N
 2161 CCACTGACCTCGGCCTGTCCATATCCACTCTGGTCACCATGCTGAGTATATTCTGGTTCA
 93 V R E I S F N A C L S H M F F I K F F T
 2221 ATGTGAGGGAAATCAGCTTTAATGCCTGTGTGCCACATGTTCTTTATTAAATCTTCA
 113 V M E S S V L L A M A F D R F V A V S N
 2281 CTGTCATGGAATCCTCAGTGTCTGTTGGCCATGGCTTTTGATCGTTTGTGGCCGCTCTTA
 133 P L R Y A M I L T D S R I A Q I G V A S
 2341 ATCCCCTTAGGTATGCCATGATTTTAACTGACTCCAGAATAGCTCAAATTGGAGTGGCAA
 153 V I R G L L M L T P M V A L L I R L S Y
 2401 GTGTCATCAGGGGCTCCTAATGCTGACACCAATGGTAGCACTTCTTATAAGACTTTCTCT
 173 C H S Q V L H H S Y C Y H P D V M K L S
 2461 ACTGCCACAGCCAAGTACTCCACCCTCTACTGCTACCACCCTGATGTGATGAAGCTCT
 193 C T D T R I N S A V G L T A M F S T V G
 2521 CATGCACAGACACCAGAATCAACAGTGCAGTTGGGCTGACTGCCATGTTCTCTACTGTTG
 213 V D L L L I L L S Y V L I I R T V L S V
 2581 GTGTAGACTTACTTCTCATCCTCCTTTCTTATGTTTTGATCATTAGGACTGTCTTAGCG
 233 A S P E E R K E T F S T C V S H I V A F
 2641 TTGCTTCCCAGAAGAGAGGAAGGAAACCTTCAGTACATGTGTCTCCCACATTGTGGCTT
 253 A I Y Y I P L I S L S I V H R F G K Q A
 2701 TTGCTATATATTACATTCCATTGATCAGTCTGTCCATTGTTACAGATTTGGGAAACAAG
 273 P A Y V H T M I A N T Y L L I S P L M N
 2761 CCCCAGCCTATGTACATACTATGATTGCTAACACCTACCTGCTGATCTCCCCTTTGATGA
 293 P V I Y S V K T K Q I R R A V I K I L H
 2821 ACCCTGTCATCTACAGTGTGAAAACCAAACAGATACGTAGAGCTGTGATAAAAATCTCC
 313 S K E T *
 2881 ATTCCAAAGAAACATAAAaattgtagaattctagcactagatggaccctcaagatggtat
 2941 acttcagggtcatggtttccaattaaaaacaaaaacaaccaagaactgggtgcccattgaa
 3001 gctaagcacataatcatttaaagttaaagctggaatgtatctagaacatcagctagtttg
 3061 actactttatttccagatgaggaggtgaagacatgggatggaagagatgaacagcctaa
 3121 aattagccttatataataaaatttctccttctcacaccactggtgggtccaaactggtggt
 3181 gccttctgtccctgccaacacccttctgttattatctcatctacttcatagcctttaga
 3241 atgacacattacagaaaagtcactcctttgtgagaaataggtagtattttccaactgttc
 3301 tagtaatttcgattgccatttattatttctctgacaagtcaggttctaagtgtttttt
 3361 gtatgtattaacatattttatctctaaaacaatggttataagataatcactttaattatca
 3421 ataagaaaattgaggcatgaaatgtaagtagcttgcttaaacagctagtaactgggtgtt
 3481 gctgatggcagagcacatccatgcagactgactccagagtccccacactcttaatttta

3541 atggttatattgctacctataatataatataataggtatataaatttatatataaatttatatt
3601 tgtaaaaattatatgaacttatatatataaacttatatttataaattatgtacttatatg
3661 tgtaaaattatatttatatttataaatataaatatagagagaagttcaggcaatttttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaaaaaa

3541 atggttatattgctacctataatataatataataggtatataaatttatatataaatttatatt
3601 tgtaaaaattatatgaacttatatatataaacttatatttataaattatgtacttatatg
3661 tgtaaaattatatttatatttataaatataaatatagagagaagttcaggcaatttttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaaaaaa

Figura 2C. La secuencia de ADNc y de aminoácidos de 238P1B2 v.2. La metionina de partida está subrayada. El marco de lectura abierto se extiende desde el ácido nucleico 2133- 2897 incluyendo el codón de parada.

```

1  ccgccttccagggtcactccattctcctgcctcagcctcccagtagctggaactacagg
61  cgccccgccaccacgccccggctaattttttgtatttttagtgagagcagaggtttcaccatg
121 ttggccaggatggtctcgatctcctgacctcgtgatccgcccacctcggcctcccaaagt
181 gctgggactacaggcgtgagccaccgcccagcctgctcaacacttttacacatgccgt
241 tctgaaagcctgggaattcttcgtctcaacttatataaaacacagaaatctttcaagcac
301 tgctttttctagaaagccttctccaaatactacaaaacctactgaatgctctctttctggg
361 aattccttttttagtggtttttgggtttttcaaaatcattttatcttattttctagttgacaa
421 ataatatgggtatataatttggtgggtaaaatgtgatgtgttatacatgtatgcattgtgg
481 aatgaccaaatcaagctaattatcatatccatcacctcaaataattatcattttcttggg
541 ctgtgaacattttaaattctctcttttagttttttgaaatatactatacattaactata
601 attcttatttcttaagaggtttgatcacatttttagatattttcatcactcgcctgag
661 gacactgtcattactgtaaatttcagaatctaaagctatagtttcattatttggccaaat
721 cgtgtaactatataatttcttggtttttcatctagttttctgttaatttagttaaatacaca
781 cccccaccccccttttttttccactaagacttcttccggtgatccttgaggtgacactt
841 tatcaggggtgaatgtgtgagctcggatacgttgaggttcctctccatcctctctttgtc
901 aacacttactgtcagccaccagcacccttcttttctctgaagaagtaccctgcaac
961 tggctctctctcattgattgagccctctgcagccacaaccctaatactctgttttcccta
1021 gacccccgtcttcaactcaaccatacccatctatctacctttctgcaattctccactcta
1081 tgaaaaaagaaatgtgatgtttgcaataaagtatattttccctccgaaacaaaacagt
1141 aacattcaattaggtatctcctaagaagtatgtcttaacaaaatagagagattctgttct
1201 ggagaactagtcttccagcagactgaagagagaactgtgttggaacgtggcagcacatta
1261 tatcaggtgggggcagcctcctgtcttccaaagcctggggagggtctgtgttctaacca
1321 tctcttccaaatatgatgccgggtatcatattgtaagggtttctttaatgggtagatgct
1381 gaagccatttatttgtcacttccatgtgttactttgtattataatgatctctttcttca
1441 tattttctggcttcccaaattgtctgtaagcttctctgctttccacacttatgaccaatg
1501 taatattgctgttccacagccccagtgatttttctctgggtgccaaccatagaggaaaaa
1561 taatgcactacctctttagcacaatttaaacatgccataaagaatgaaatagattggca
1621 aactcagggtaaagtgtttattcttagtttgggtatctatggctagaacgatcagaggag
1681 tggtcacatgatataaacatagctgccagcaaagggccctagagaagaaagattgtacac
1741 taaggcaaatcaaaaacaaaccaaagtgccactacagcagttgaaatctgacattttt
1801 taatcaattttaatgagccattcatagtttcatattatataaatactcatcagtttctg
1861 tacaataaaattattttaattcgatcaatacacaatgaatgtagctttttttgtttaa
1921 gctctcatctcgttctgtacttttaattcatcacgtcaactttgcagaatcacctcca
1981 cttccatcattttctgctcactgggttccctgggctggaagccttccacacctggatct
2041 ccattcccttctgcttccctctctgtaactgctctcttgggaaacagcctgatcctctcg

```

1 M Y Y F L S M L S A
 2101 ctaccatcactcagcccagcctccaacgaaccaATGTACTATTTCTCTCCATGCTGTCCG
 11 T D L G L S I S T L V T M L S I F W F N
 2161 CCACTGACCTCGGCCTGTCCATATCCACTCTGGTCACCATGCTGAGTATATTCTGGTTCA
 31 V R E I S F N A C L S H M F F I K F F T
 2221 ATGTGAGGGAATCAGCTTTAATGCCTGCTGTGCCACATGTTCTTTATTAAMTCTTCA
 51 V M E S S V L L A M A F D R F V A V S N
 2281 CTGTCATGGAATCCTCAGTGCTGTTGGCCATGGCTTTTGATCGTTTTGTGGCCGCTCTA
 71 P L R Y A M I L T D S R I A Q I G V A S
 2341 ATCCCCCTTAGGTATGCCATGATTTTTAACTGACTCCAGAATAGCTCAAATTGGAGTGGCAA
 91 V I R G L L M L T P M V A L L I R L S Y
 2401 GTGTCATCAGGGGGCTCCTAATGTGACACCAATGGTAGCACTTCTTATAAGACTTTCT
 111 C H S Q V L H H S Y C Y H P D V M K L S
 2461 ACTGCCACAGCCAAGTACTCCACCCTCTACTGCTACCACCCTGATGTGATGAAGCTCT
 131 C T D T R I N S A V G L T A M F S T V G
 2521 CATGCACAGACACCAGAATCAACAGTGCAGTTGGGCTGACTGCCATGTTCTCTACTGTTG
 151 V D L L L I L L S Y V L I I R T V L S V
 2581 GTGTAGACTTACTTCTCATCCTCCTTTCTTATGTTTTGATCATTAGGACTGTCTTAGCG
 171 A S P E E R K E T F S T C V S H I V A F
 2641 TTGCTTCCCCAGAAGAGAGGAAGGAAACCTTCAGTACATGTGTCTCCACATTGTGGCTT
 191 A I Y Y I P L I S L S I V H R F G K Q A
 2701 TTGCTATATATTACATTCCATTGATCAGTCTGTCCATTGTTACAGATTGGGAAACAAG
 211 P A Y V H T M I A N T Y L L T S P L M N
 2761 CCCCAGCCTATGTACATACTATGATTGCTAACACCTACCTGCTGACCTCCCCCTTGATGA
 231 P V I Y S V K T K Q I R R A V I K I L H
 2821 ACCCTGTCATCTACAGTGTGAAAACCAAACAGATACGTAGAGCTGTGATAAAAATTCTCC
 251 S K E T *
 2881 ATCCAAAGAAACATAAaattgtagaattctagcactagatggaccctcaagatggat
 2941 acttcaggtcatggttccaattaaaaacaaaaacaaccaagaactggtgcccattgaa
 3001 gctaagcacataatcatttaaagttaaagctggaatgtatctagaacatcagctagttg
 3061 actactttatttccagatgaggaggtgaagacatgggatggaagagatgaacagcctaa
 3121 aattagccttatatataaaaatttctccttctcacaccactggtgggtccaaactggtggt
 3181 gccttctgtccctgccaacaccctttctggtattatctcatctacttcatagcctttaga
 3241 atgacacattacagaaaagtcactcctttggtgagaaataggragtatttccaactggtc
 3301 tagtaatttcgattgccatttattttgctctgacaagtcaggttctaagtggttttt
 3361 gtagtattaacataattttatctaaaacaatggtataagataatcactttaattatca
 3421 ataagaaaattgaggtcatgaaatgtaagtagcttgcttaaacagctagtaactgggtggt
 3481 gctgatggcagagcacatccatgcagactgactccagagtccccacactcttaatttta
 3541 atggtatattgctacctataatataatataataggtatataatttatatataatttatatt
 3601 tgtaaaaatttatatgaacttatatataaaacttatatttataaattatgtacttatatg

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaatcttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaaaaaa

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaatcttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaaaaaa

Figura 2D. La secuencia de ADNc y de aminoácidos de 238P1B2 v.3. La metionina de partida está subrayada. El marco de lectura abierto se extiende desde el ácido nucleico 2133-2897 incluyendo el codón de parada.

```

1  ccgccttcaggtcactccattctcctgcctcagcctcccagtagctggaactacagg
61  cgccccaccacgcccggctaattttttgtattttttagtggagacgaggttcaccatg
121  ttggccaggatggtctcgatctcctgacctcgtgatccgccacctcggcctcccaaagt
181  gctgggactacaggcgtgagccaccgcgccagcctgctcaacacttttacacatgcctg
241  tctgaaagcctgggaattcttctcgtctcaacttatataaaacacagaaatctttcaagcac
301  tgctttttctagaaagccttctccaaatactacaaacctactgaatgctctctttctggt
361  aatccttttttagtgtgttttggtttttcaaaatcatttttatcttatttctagtggaaa
421  ataatatggtatatatttgtggggtaaaatgtgatgttgttatacatgcatgcatgtgtg
481  aatgaccaaatacaagctaattatcatatccatcacctcaaataattatcatttctttgtg
541  ctgtgaacatttaaaatctctcttttttagttttttgaaatatactatacattaactata
601  attcttatttctcctaagagggttgatcatacatttttagatattttcatcactcgcctgag
661  gacactgtcattactgtaaatttcagaatctaaagctatagtttcattatttggccaaat
721  cgtgtaatctataatttcttgtttttcatctagtttctgttaatttttagttaaatcaca
781  ccccccccccttttttttctactaagacttcttccggtgatccttgagggtgacactt
841  tatcagggtgaatgtgtgagctcggatacgttgagggttctctccatcctctcttbtgtc
901  aacacttactgtcagccaccagcaccctttcttttctctgaagaagttaccctgcaac
961  tggtctctctcattgatgagccctctgcagccacaaccctaactctctgtttttcccta
1021 gacccccgtcttcaactcaaccataccatctatctacctttctgcaattctccactcta
1081 tgaaaaaagaaatgtgatgtttgcaataaagtatatattttccctccgaaacaaaacagt
1141 aacattcaattaggtatctcctaagaagtatgtcttaacaaaatagagagattctgttct
1201 ggagaactagtcttccagcagactgaagagagaactgtgttggaaacgtggcagcacatta
1261 taccaggtggggcagcctcctgtcttcttaagcctgggtgggaggtctgtgttctaacca
1321 tctcttccaaatatgatgccgggtatcatattgtaagggttctttaaaggttagatgct
1381 gaagccatttatttgtcacttccatgtgttactttgtattataatgatctcttcttca
1441 tattttctggcttcccaaattgtctgtaagcttctctgctttccacacttatgaccaatg
1501 taatattgctgttccacagccccagtgatattttctctgggtgccaaccatagaggaaaaa
1561 taatgcactacctctttagcacaatttaaacatgccataaagaatgaaatagattggca
1621 aactcagggtaaagtgtttattcttagtttgggttatctatggctagaacgatcagaggag
1681 tggtcacatgatataaacatagctgccagcaaagggccctagagaagaaagattgtacac
1741 taaggcaaattcaaaaacaaaccaaagtgccactacagcagttgaaatctgacattttt
1801 taatcaatttaaatgagccattcatagtttcatattatataaatactcatcagttttctg
1861 tacaataaaattattttaattcgatcaatacacaatgaatgtagctttttttgttttaa
1921 gctctcatctcgttctgtacttttaattcatcacgtcaactttgcagaatatcacctcca
1981 ctccatcattttcctgctcactggtgttctctgggctggaagccttccacacctggatct
2041 ccattcccttctgcttctctctgtaactgctctcttgggaaacagcctgatcctcttcg

```

1 M Y Y F L S M L S A
 2101 ctaccatcactcagcccagcctccacgaaccaATGTACTATTTCCCTCTCCATGCTGTCCG
 11 T D L G L S I S T L V T M L S I F W F N
 2161 CCACTGACCTCGGCCTGTCCATATCCACTCTGGTCACCATGCTGAGTATATTTCTGGTTCA
 31 V R E I S F N A C L S H M F F I K P F T
 2221 ATGTGAGGGAAATCAGCTTTAATGCCTGCTTGCCACATGTTCTTTATTAAATTCTTCA
 51 V M E S S V L L A M A F D R F V A V S N
 2281 CTGTCATGGAATCCTCAGTCTGTTGGCCATGGCTTTTGATCGTTTTGTGGCCGCTCTCTA
 71 P L R Y A M I L T D S R I A Q I G V A S
 2341 ATCCCCCTTAGGTATGCCATGATTTTAAC TGACTCCAGAATAGCTCAAATTGGAGTGGCAA
 91 V I R G L L M L T P M V A L L I R L S Y
 2401 GTGTCATCAGGGGCTCCTAATGCTGACACCAATGGTAGCACTTCTTATAAGACTTTCCT
 111 C H S Q V L H H S Y C Y H P D V M K L S
 2461 ACTGCCACAGCCAAGTACTCCACCCTCTACTGCTACCACCCTGATGTGATGAAGCTCT
 131 C T D T R I N S A V G L T A M F S T V G
 2521 CATGCACAGACACCAGAATCAACAGTGCAGTTGGGCTGACTGCCATGTTCTCTACTGTG
 151 V D L L L I L L S Y V L I I R T V L S V
 2581 GTGTAGACTTACTTCTCATCCTCCTTTCTTATGTTTTGATCATTAGGACTGTCCTTAGCG
 171 A S P E E R K E T F S T C V S H I V A F
 2641 TTGCTTCCCCAGAAGAGAGGAAGGAAACCTCAGTACATGTGTCTCCACATTGTGGCTT
 191 A I Y Y I P L I S L S I V H R F G K Q A
 2701 TTGCTATATATTACATTCCATGATCAGTCTGTCCATTGTTCCACAGATTGGGAAACAAG
 211 P A Y V H T M I A N T Y L L I S P L M N
 2761 CCCCAGCCTATGTACATACTATGATGTCTAACACCTACCTGCTGATCTCCCCCTTGATGA
 231 P V I Y S V K T K Q I R R A V I K I L H
 2821 ACCCTGTCATCTACAGTGTGAAAACCAACAGATACGTAGAGCTGTGATAAAAATTCTCC
 251 S K E T *
 2881 ATTCCAAAGAAACATAAaattgtagaattctagcactagatggaccocctcaagatggat
 2941 acttcaggatcatggttccaattaaaaacaaaaaacaaccaagaactgggccccattgaa
 3001 gctaagcacataatcatttaaagttaaagctggaatgtatctagaacatcagctagtttg
 3061 actactttatttccagatgaggaggtgaagacatgggatggaagagatgaacagcctaa
 3121 aattagccttatataaaaatttctccttctcacaccactggtgggtccaaactgttgtt
 3181 gccttctgtccctgccaacacccttctgttattatctcatctacttcatagcctttaga
 3241 atgacacattacagaaaagtcactccttgtgagaaaataggtagtattttccaactgttc
 3301 tagtaatttcgattgcccatttattatgtgctctgacaagtcaggttctaagtgtttttt
 3361 gtatgtattaacatattttatatactaaaacaatgttataagataatcactttaattatca
 3421 ataagaaaattgaggcatgaaatgtaagtagcttgcttaacagctagtaactgggtgtt
 3481 gctgatggcagagcacatccatgcagactgactccagagtccccacactcttaatttta
 3541 atggttatattgctacctataatataatataataggatataatttatatataatttatatt
 3601 tgtaaaaattatatagaacttatataataaacttatatttataaattatgtacttatatg

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaattttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaaaaaa

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaattttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaaaaaa

Figura 2E. La secuencia de ADNc y de aminoácidos de 238P1B2 v.4. La metionina de partida está subrayada. El marco de lectura abierto se extiende desde el ácido nucleico 2133-2897 incluyendo el codón de parada.

1 ccgccttccaggttcactccattctcctgcctcagcctccccagtagctggaactacagg
 61 cgcccgccaccacgcccggtaatTTTTTGTATTTTtagtggagacgaggtttcaccatg
 121 ttggccaggatggtctcgatctcctgacctcgtgatccgcccacctcggcctcccaaagt
 181 gctgggactacagggctgagccaccgcccagcctgctcaacacttttacacatgccgt
 241 tctgaaagcctgggaattcttcgtctcaacttatataaaacacagaaatctttcaagcac
 301 tgctTTTTTctagaaagccttctccaaatactacaaaacctactgaatgctctctttctggt
 361 aattcctTTTTtagtggttttggTTTTTcaaaatcattttatcttatttctagttgacaa
 421 ataatatggtatatatttgggggtaaaatgtgatggtgtatacatgtatgcattgtgg
 481 aatgaccaaataagctaattatcatatccatcacctcaaataatttatcatttctttgtg
 541 ctgtgaacatttaaaattctctcttttagttttttgaaatatactatacattaactata
 601 attcttatttctctaagaggttgatcacatatttagatattttcatcactcgcctgag
 661 gacactgtcattactgtaaatttcagaatctaaagctatagtttcattattggccaaat
 721 cgtgtaatctataatttcttgTTTTTcatctagttttctgTTAATTTtagttaaatcaca
 781 cccccccccctTTTTTTTTcactaagacttcttccgggtgatccttgaggtgacactt
 841 tatcaggtgaaatgtgtgagctctggatacgttgaggttctctccatcctctctttgtc
 901 aacacttactgtcagccaccagcacccttcttttctctgaagaagtaccctgcaac
 961 tggctctctctcattgattgagccctctgcagccacaaccctaactctctgTTTTTcccta
 1021 gacccccgtcttcaactcaaccatacccatctatctacctttctgcaattctccactcta
 1081 tgaaaaaagaaatgtgatgtttgcaataaagttatattttccctccgaaacaaaacagt
 1141 aacattcaattaggtatctcctaagaagtatgtcttaacaaaatagagagattctgttct
 1201 ggagaactagtcttccagcagactgaagagagaactgtgttggaacgtggcagcacatta
 1261 tatcaggtgggggcagcctcctgtcttcttaagcctggggggaggtctgtgttctaacca
 1321 tctcttccaaatgatgaccgggtatcatattgtaagggttctttaaagggttagatgct
 1381 gaagccatttatttgtcacttctatgtgttactttgtattataatgatctctttctca
 1441 tattttctggcttcccaaattgtctgtaagcttctctgctttccacacttatgaccaatg
 1501 taatattgctgttccacagccccagtgatTTTTTctctgggtgccaaccatagaggaaaa
 1561 taatgcactacctcttagcacaaatttaaacatgccataaagaatgaaatgatggca
 1621 aactcaggttaaagtgtttattcttagtttgggtatctatggctagaacgatcagaggag
 1681 tggtcacatgatataaacatagctgccagcaaaagggcctagagaagaaagattgtacac
 1741 taaggcaaatcaaaaaacaaacaaagtgccactacagcagttgaaatctgacatttt
 1801 taatcaatttaaatgagccattcatagtttcatattatataaatactcatcagttttctg
 1861 tacaataaaattattttaattcgatcaatacaciaaatgaatgtagctTTTTTTTgTTAA
 1921 getctcatctcgttctgtacttttaattcatcacgtcaactttgcagaatatcacctcca
 1981 ctccatcattttctctcactgggtgttctctgggctggaagccttccacacctggatct
 2041 ccattccctctcgttctctctgtaactgctctcttgggaaacagcctgatcctctctc

1 M Y Y F L S M L S A
 2101 ctaccatcactcagcccagcctccacgaaccaATGTACTATTTTCCTCTCCATGCTGTCCG
 11 T D L G L S I S T L V T M L S I F W F N
 2161 CCACTGACCTCGGCCTGTCCATATCCACTCTGGTCACCATGCTGAGTATATTCTGGTTCA
 31 V R E I S F N A C L S H M F F I K F F T
 2221 ATGTGAGGGAAATCAGCTTTAATGCCTGCTTGTCCCACATGTTCTTTATTAATTCCTCA
 51 V M E S S V L L A M A F D R F V A V S N
 2281 CTGTCAATGGAATCCTCAGTCTGTTGGCCATGGCTTTTGATCGTTTTGTGGCCGCTCTA
 71 P L R Y A M I L T D S R I A Q I G V A S
 2341 ATCCCCTTAGGTATGCCATGATTTTAACTGACTCCAGAATAGCTCAAATTGGAGTGGCAA
 91 V I R G L L M L T P M V A L L I R L S Y
 2401 GTGTCAATCAGGGGGCTCCTAATGCTGACACCAATGGTAGCACTTCTTATAAGACTTTCCT
 111 C H S Q V L H H S Y C Y H P D V M K L S
 2461 ACTGCCACAGCCAAGTACTCCACCCTCTACTGCTACCACCCTGATGTGATGAAGCTCT
 131 C T D T R I N S A V G L T A M F S T V G
 2521 CATGCACAGACACCAGAATCAACAGTGCAGTTGGGCTGACTGCCATGTTCTCTACTGTTG
 151 V D L L L I L L S Y V L I I R T V L S V
 2581 GTGTAGACTTACTTCTCATCCTCCTTTCTTATGTTTTGATCATTAGGACTGTCCTTAGCG
 171 A S P E E R K E T F S T C V S H I V A F
 2641 TTGCTTCCCAGAAAGAGAGGAAGGAAACCTTCAGTACATGTGTCTCCACATTGTGGCTT
 191 A I Y Y I P L I S L S I V H R F G K Q A
 2701 TTGCTATATATTACATTCATTGATCAGTCTGTCCATTGTTTACAGATTTGGGAAACAAG
 211 P A Y V H T M I A N T Y L L I S P L M N
 2761 CCCCAGCCTATGTACATACTATGATTGCTAACACCTACCTGCTGATCTCCCCTTGATGA
 231 P V I Y S V K T K Q I R R A V I K I L H
 2821 ACCCTGTCATCTACAGTGTGAAAACCAACAGATACGTAGAGCTGTGATAAAAATTCTCC
 251 S K E T *
 2881 ATTCCAAAGAAACATAAaattgtagaattctagcactagatggaccctcaagatggtat
 2941 acttcagggtcatggttccaattaaaaacaaaaacaaccaagaactgggtgccattgaa
 3001 gctaagcacataatcatttaaagttaaagctggaatgtatctagaacatcagctagtttg
 3061 actactttatctccagatgaggagggtgaagacatgggatggaagagatgaacagcctaa
 3121 aattagccttatataaaaattctctctctcacaccactggtgggtccaaactggtggt
 3181 gcctctgtccctgccaacacccttctgtttattatctcatctacttcatagcctttaga
 3241 atgacacattacagaaaagtcaactccttctgtgagaaataggtagtattttccaactgttc
 3301 tagtaatttcgattgccatttattatttctgctctgacaagtcattggttctaagtgttttt
 3361 gtatgtattaacatattttatatctaaaacaatgttataagataatcactttaattatca
 3421 ataagaaaattgaggcatgaaatgtaagtagcttgcttaaacagctagtaactgggtgtt
 3481 gctgatggcagagcacatccatgcagactgactccagagtcccccacactcttaatttta
 3541 atggttatattgctacctataatataatataataggtatataatttatataatattatatt
 3601 tgtaaaaattatagaaacttatataataaacttatatttataaattatgtacttatatg

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaattttact
3721 aggaacataaagaagtaaaaaaaaaaaaaaaaa

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaattttact
3721 aggaacataaagaagtaaaaaaaaaaaaaaaaa

Figura 2F. La secuencia de ADNc y de aminoácidos de 238P1B2 v.5. La metionina de partida está subrayada. El marco de lectura abierto se extiende desde el ácido nucleico 2133-2897 incluyendo el codón de parada.

```

1  ccgccttccagggttcactccattctcctgcctcagcctcccagtagctggaactacagg
61  cgcgccaccacgcgccggttaattttttgtattttttagtggagacgaggtttcaccatg
121  ttggccaggatggtctcgatctcctgacctcgtgatccgccacctcggcctcccaaagt
181  gctgggactacaggcgtgagccaccgcgccagcctgctcaacacttttacacatgccgt
241  tctgaaagcctgggaattcttcgtctcaacttatataaaacacagaaatctttcaagcac
301  tgctttttctagaaagccttctccaaatactacaaaacctactgaatgctctctttctggg
361  aattccttttttagtgtgttttggtttttcaaaatcattttatcttatttctagttgacaa
421  ataatatggtatatatttgtggggtaaaatgtgatggtgttatacatgtatgcattgtgg
481  aatgaccaaatacaagctaattatcatatccatcacctcaaatatttatcatttctttgtg
541  ctgtgaacatttaaaattctctcttttagttttttgaaatatactatacattaactata
601  attcttatttctctaagaggtttgatcatacatttttagatattttcatcactcgcctgag
661  gacactgtcattactgtaaattcagaatctaaagctatagtttcattatttggccaaat
721  cgtgtaatctataatttcttgtttttcatctagttttctgttaatttttagttaaatcaca
781  cccccaccccccttttttttcaactaagacttcttcgggtgatccttgagggtgacactt
841  taccaggtgaatgtgtgagctcggatacgttgagggtcctctccatcctctcttttgtc
901  aacacttactgtcagccaccagcaccctttcttttctctgaagaagttaccctgcaac
961  tggctctctctcattgattgagccctctgcagccacaaccctaattctctgtttttcccta
1021  gacccccgtcttcaactcaaccatacccatctatctacctttctgcaattctccactcta
1081  tgaaaaaagaaatgtgatgtttgcaataaagttatatttttccctcogaacaaaaacagt
1141  aacattcaattaggtatctcctaagaagtatgtcttaacaaaatagagagattctgttct
1201  ggagaactagtcttccagcagactgaagagagaactgtggtggaacgtggcagcacatta
1261  tatcagggggggcagcctcctgtcttccctaagcctgggtgggaggtctgtgttctaacca
1321  tctcttccaaatgatgcccgggtatcatattgtaagggttctttaaagggttagatgct
1381  gaagccatttatttgtcacttccatgtgttactttgtattataatgatctctttcttca
1441  tattttctggcttcccaaattgtctgtaagcttctctgctttccacacttatgaccaatg
1501  taatattgctgttccacagcccagtgatatttttctctgggtgccaaccatagaggaaaaa
1561  taatgcactacctctttagcacaatttaaacatgccataaagaatgaaatagattggca
1621  aactcagggtaaagtgttattcttagtttggttatctatggctagaacgatcagaggag
1681  tggtcacatgatataaacatagctgccagcaaagggccctagagaagaagattgtacac
1741  taaggcaattcaaaaacaaacaaagtgcccactacagcagttgaaatctgacattttt
1801  taatcaatttaaatgagccattcatagtttcatattatataaatactcatcagttttctg
1861  tacaataaaattatttttaattcgatcaatacacaaatgaatgtagctttttttgtttaa
1921  gctctcatctcgttctgtacttttaattcatcacgtcaactttgcagaatatcacctcca
1981  cttccatcattttcctgctcactgggtgttccctgggctggaagccttccacacctggatct
2041  ccattcccttctgcttccctctctgtaactgctctcttgggaacagcctgatcctctteg

```

1 M Y Y F L S M L S A
 2101 ctaccatcactcagcccagcctccacgaaccaATGTACTATTTCTCTCCATGCTGTCCG
 11 T D L G L S I S T L V T M L S I F W F N
 2161 CCACTGACCTCGGCCTGTCCATATCCACTCTGGTCACCATGCTGAGTATATTCTGGTTCA
 31 V R E I S F N A C L S H M F F I K F F T
 2221 ATGTGAGGGAATCAGCTTTAATGCCTGCTGTGCCACATGTTCTTTATTAAATTCTTCA
 51 V M E S S V L L A M A F D R F V A V S N
 2281 CTGTCATGGAATCCTCAGTGCTGTTGGCCATGGCTTTTGATCGTTTTGTGGCCGTCTCTA
 71 P L R Y A M I L T D S R I A Q I G V A S
 2341 ATCCCCTTAGGTATGCCATGATTTTAACTGACTCCAGAATAGCTCAAATTGGAGTGGCAA
 91 V I R G L L M L T P M V A L L I R L S Y
 2401 GTGTCATCAGGGGGCTCCTAATGCTGACACCAATGGTAGCACTTCTTATAAGACTTTCTCT
 111 C H S Q V L H H S Y C Y H P D V M K L S
 2461 ACTGCCACAGCCAAGTACTCCACCCTCTACTGCTACCACCCTGATGTGATGAAGCTCT
 131 C T D T R I N S A V G L T A M F S T V G
 2521 CATGCACAGACACCAGAATCAACAGTGCACTGGGGCTGACTGCCATGTTCTCTACTGTTG
 151 V D L L L I L L S Y V L I I R T V L S V
 2581 GTGTAGACTTACTTCTCATCCTCCTTTCTTATGTTTTGATCATTAGGACTGTCCTTAGCG
 171 A S P E E R K E T F S T C V S H I V A F
 2641 TTGCTTCCCCAGAAGAGAGGAAGGAAACCTTCAGTACATGTGTCTCCACATTGTGGCTT
 191 A I Y Y I P L I S L S I V H R F G K Q A
 2701 TTGCTATATATTACATTCCATTGATCAGTCTGTCCATTGTTACAGATTGGGAAACAAG
 211 P A Y V H T M I A N T Y L L I S P L M N
 2761 CCCCAGCCTATGTACATACTATGATTGCTAACACCTACCTGCTGATCTCCCCTTGTATGA
 231 P V I Y S V K T K Q I R R A V I K I L H
 2821 ACCCTGTCATCTACAGTGTGAAAACCAACAGATACGTAGAGCTGTGATAAAAAATCTCC
 251 S K E T *
 2881 ATCCAAAGAAACATAAaattgtagaattctagcactagatggacccctcaagatggtat
 2941 acttcaggtcatgtttccaattaaaaacaaaaacaaccaagaactggtgcccattgaa
 3001 gctaagcacataatcatttaaagttaaagctggaatgtatctagaacatcagctagtttg
 3061 actactttatttccagatgaggaggtgaagacatgggatggaagagatgaacagcctaa
 3121 aattagccttatatataaaatttctccttctcacaccactggtgggtccaaactggtggt
 3181 gccttctgtccctgccaacacctttctgttattatctcatctacttcatagcctttaga
 3241 atgacacattacagaaaagtcactcctttgtgagaaataggtagtattttccaactgttc
 3301 tagtaatttcgattgccatttattttgctctgacaagtcaggttctaaggtttttt
 3361 gtatgtattaacatattttatctaaaacaatggtataagataatcactttaattatca
 3421 ataagaaaattgaggtcatgaaatgtaagtagcttgcttaaacagctagtaactggtggt
 3481 gctgatggcagagcacatccatgcagactgactccagagteccccacactcttaatttta
 3541 atgttatattgctacctataataatataataggtatataatttatatataatttatatt
 3601 tgtaaaaatttatatgaacttatatataaaacttatatttataaattatgacttatatg

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaatcttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaa

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaatcttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaa

Figura 2G. La secuencia de ADNc y de aminoácidos de 238P1B2 v.6. La metionina de partida está subrayada. El marco de lectura abierto se extiende desde el ácido nucleico 2133-2897 incluyendo el codón de parada.

```

1  ccgccttcagggttcactccattctcctgectcagcctccccagtagctggaactacagg
61  cgcccgccaccacgcccggetaatTTTTTgtattTTTtagtgagacgaggtTtcaccatg
121  ttggccaggatgggtctcgatctcctgacctcgtgatccgcccacctcggcctcccaaagt
181  gctgggactacagggcgtgagccaccgcgccagcctgctcaacaactTTTtacacatgccgt
241  tctgaaagcctgggaattcttcgtctcaacttacataaaacacagaaatctTtcaagcac
301  tgctTTTTctagaaagccttctccaaatactacaaacctactgaatgctctctTtctggt
361  aattcctTTTTtagtggtTTTTggTTTTTcaaaaccattTTTatcttattTctagttgacaa
421  ataatatggtatatatttgtgggtTaaaatgtgatgttgttatacatgtatgcattgtgg
481  aatgaccaaatacaagctaattatcatatccatcacctcaaatTTTatcatttctTtgtg
541  ctgtgaacattTaaaattctctctTttagtTTTTTgaaatatactatacattaactata
601  attcttatttctcctaagaggttgatcatacattTtagatTTTcatcactcgcctgag
661  gacactgtcattactgtaattTcagaatctaaagctatagttTcattattTggccaaat
721  cgtgtaatctataatttcttgtTTTTcatctagTTTTctgttaattTtagtTaaatcaca
781  cccccccccctTTTTTTTTcactaagacttctcoggtgatccttgaggtgacactT
841  tatcagggTgaatgTgtgagctctggatacgttgaggtTcctctccatcctctctTtgtc
901  aacacttactgtcagccaccagcaccctTctTTTTctctgaagaagTtacccTgcaac
961  TggctctctctcattgattgagcctctgcagccacaaccctaatctctgTTTTTcccta
1021  gacccccgtctTcaactcaaccatacccatctatctacctTtctgcaattctccactcta
1081  TgaaaaaagaaatgtgatgtTtgcaataaagTtatattTTTccctccgaaacaaaacagT
1141  aacattcaattaggTatctcctaagaagTatgtctTaaacaaaatagagagattctgtTct
1201  ggagaactagTctTccagcagactgaagagagaactgtgtTggaacgtggcagcaccatta
1261  tatcaggtgggggcagcctcctgtctTcctaagcctggTgggaggtctgtgtTctaacca
1321  tctctTcaaatatgatgcccggTatcatattgtaaggtTtctTaatggTtagatgct
1381  gaagccattTattTgtcactTcctatgtgtTactTgtattataatgatctctTctTca
1441  tattTctggctTcccaaattgtctgTaaagctTctctgctTtccacactTatgaccaatg
1501  taatattgctgtTccacagccccagTgtattTTTTctctggTgccaaccatagaggaaaaa
1561  taatgcaactacctTtagcacaattTaaacatgccataaagaatgaaatatgattggca
1621  aactcagggTaaagTgtTtattcttagTtTggTtatctatggctagaacgatcagaggag
1681  Tggtcacatgatataaacatagctgccagcaaaagggccctagagaagaaagattgtacac
1741  taaggcaaattcaaaaacaaaccTaaagTgccactacagcagTtgaaatctgacattTt
1801  taatcaattTaaatgagccattcatagTtTcatattatataaaatactcatcagTtTctg
1861  tacaataaaattattTtaattcgatcaatacacaatgaatgtagctTTTTTtTgtTtaa
1921  gctctcatctcgtTctgtactTtTaatTcatcagTcaactTtgcagaatatcactcca
1981  ctTccatcattTtctgctcactggTgtTcctgggctggaagcctTccacacctggatct
2041  ccattccctTctgctTctctgtTaaactgctctctTgggaaacagcctgatcctctctg

```

1 M Y Y F L S M L S A
 2101 ctaccatcactcagcccagcctccacgaaccaATGTACTATTTCCTCTCCATGCTGTCGG
 11 T D L G L S I S T L V T M L S I F W F N
 2161 CCACTGACCTCGGCCTGTCCATATCCACTCTGGTCACCATGCTGAGTATATTCTGGTCA
 31 V R E I S F N A C L S H M F F I K F F T
 2221 ATGTGAGGGAAATCAGCTTTAATGCCTGCTTGTCCACATGTTCTTTATTAAATCTTCA
 51 V M E S S V L L A M A F D R F V A V S N
 2281 CTGTCATGGAATCCTCAGTGTCTGGCCATGGCTTTTGATCGTTTGTGGCCGCTCTA
 71 P L R Y A M I L T D S R I A Q I G V A S
 2341 ATCCCCTTAGGTATGCCATGATTTAACTGACTCCAGAATAGCTCAAATTGGAGTGGCAA
 91 V I R G L L M L T P M V A L L I R L S Y
 2401 GTGTCATCAGGGGGCTCCTAATGCTGACACCAATGGTAGCACTTCTTATAAGACTTTCCT
 111 C H S Q V L H H S Y C Y H P D V M K L S
 2461 ACTGCCACAGCCAAGTACTCCACCCTCTACTGCTACCACCCTGATGTGATGAAGCTCT
 131 C T D T R I N S A V G L T A M F S T V G
 2521 CATGCACAGACACCAGAATCAACAGTGCAGTTGGGCTGACTGCCATGTTCTCTACTGTTG
 151 V D L L L I L L S Y V L I I R T V L S V
 2581 GTGTAGACTTACTTCTCATCTCCTTTCTTATGTTTTGATCATTAGGACTGTCCTTAGCG
 171 A S P E E R K E T F S T C V S H I V A F
 2641 TTGCTTCCCCAGAAGAGAGGAAGGAAACCTTCAGTACATGTGTCTCCCACATTGTGGCTT
 191 A I Y Y I P L I S L S I V H R F G K Q A
 2701 TTGCTATATATTACATTCCATTGATCAGTCTGTCCATTGTTACAGATTTGGGAAACAAG
 211 P A Y V H T M I A N T Y L L I S P L M N
 2761 CCCAGCCTATGTACATACTATGATGCTAACACCTACCTGCTGATCTCCCCTTTGATGA
 231 P V I Y S V K T K Q I R R A V I K I L H
 2821 ACCCTGTCTACTACAGTGTGAAAACCAAACAGATACGTAGAGCTGTGATAAAAATCTCC
 251 S K E T *
 2881 ATTCAAAGAAACATAAaattgtagaattctagcactagatggaccctcaagatggat
 2941 acttcaggatcatggttccaattaaaaacaaaaacaaccaagaactggatgcccattgaa
 3001 gctaagcacataatcatttaaagttaaagctggaatgtatctagaacatcagctagtttg
 3061 actactttatttccagatgaggaggtgaagacatgggatggaagagatgaacagcctaa
 3121 aattagccttatataataaatttctccttctcacaccactggtgggtccaaactggtggt
 3181 gccttctgtccctgccaacacccttctggttattatctcatctacttcatagcctttaga
 3241 atgacacattacagaaaagtcactccttgtgagaaataggtagatatttccaactggtc
 3301 tagtaatttegattgccatttattttgctctgacaagtcatgggtcctaagtgtttttt
 3361 gtatgtattaacatatatttatatctaaacaatggtataagataatcactttaattatca
 3421 ataagaaaattgaggcatgaaatgtaagtagcttgcttaaacagctagtaactgggtgtt
 3481 gctgatggcagagcacatccatgcagactgaactccagagtccccacactcttaatttta
 3541 atggtatattgctacctataatataatataataggtatataatttatatataatttatatt
 3601 tgtaaaaattatatagaacttatataataaacttatatttataaattatgtacttatatg

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaatttttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaa

3661 tgtaaattatatttatatttataaatataaatatagagagaagttcaggcaatttttact
3721 aggaaacataaagaagtaaaaaaaaaaaaaaaaa

Figura 3A. Secuencia de aminoácidos de 238P1B2 v.1A clon A. La proteína 238PIB2 v.1 tiene 254 aminoácidos.

```

1  MYYFLSMLSA  TDGLSISTL  VTMLSIFWFN  VREISFNACL  SHMFFIKFFT  VMESSVLLAM
61  AFDRFVAVSN  PLYAMILTD  SRIAQIGVAS  VIRGLLMLTP  MVALLIRLSY  CHSQVLHHSY
121 CYHPDVMKLS  CTDTRINSAV  GLTAMFSTVG  VDLLLLILSY  VLIIRTVLSV  ASPEERKETF
181 STCVSHIVAF  AIYYIPLISL  SIVHRFGPQA  PAYVHTMIAN  TYLLISPLMN  PVIYSVKTKQ
241 IRRAVIKILH  SKET
    
```

Figura 3B. Secuencia de aminoácidos de 238P1B2 v.1B. La proteína 238P1B2 v.1B tiene 316 aminoácidos.

```

1  FITSTLQINIT  STSIIFLLTG  VPGLEAFHTW  ISIPFCFLSV  TALLGNSLIL  FATITQPSLH
61  EPMYYFLSML  SATDLGLSIS  TLVTMLSIFW  FNVREISFNA  CLSHMPFIKF  FTMESSVLL
121 AMAFDRFVAV  SNPLRYAMIL  TDSRIAQIGV  ASVIRGLLML  TPMVALLIRL  SYCHSQVLHH
181 SYCYHPDVMK  LSCTDTRINS  AVGLTAMFST  VGVLLLLILL  SYVLIIRTVL  SVASPEERKE
241 TFSTCVSHIV  AFAIYYIPLI  SLSIVHRFGK  QAPAYVHTMI  ANTYLLISPL  MNPVIYSVKT
301 KQIRRAVIKI  LHSKET
    
```

Figura 3C. Secuencia de aminoácidos de 238P1B2 v.2. La proteína 238P1B2 v.2 tiene 254 aminoácidos.

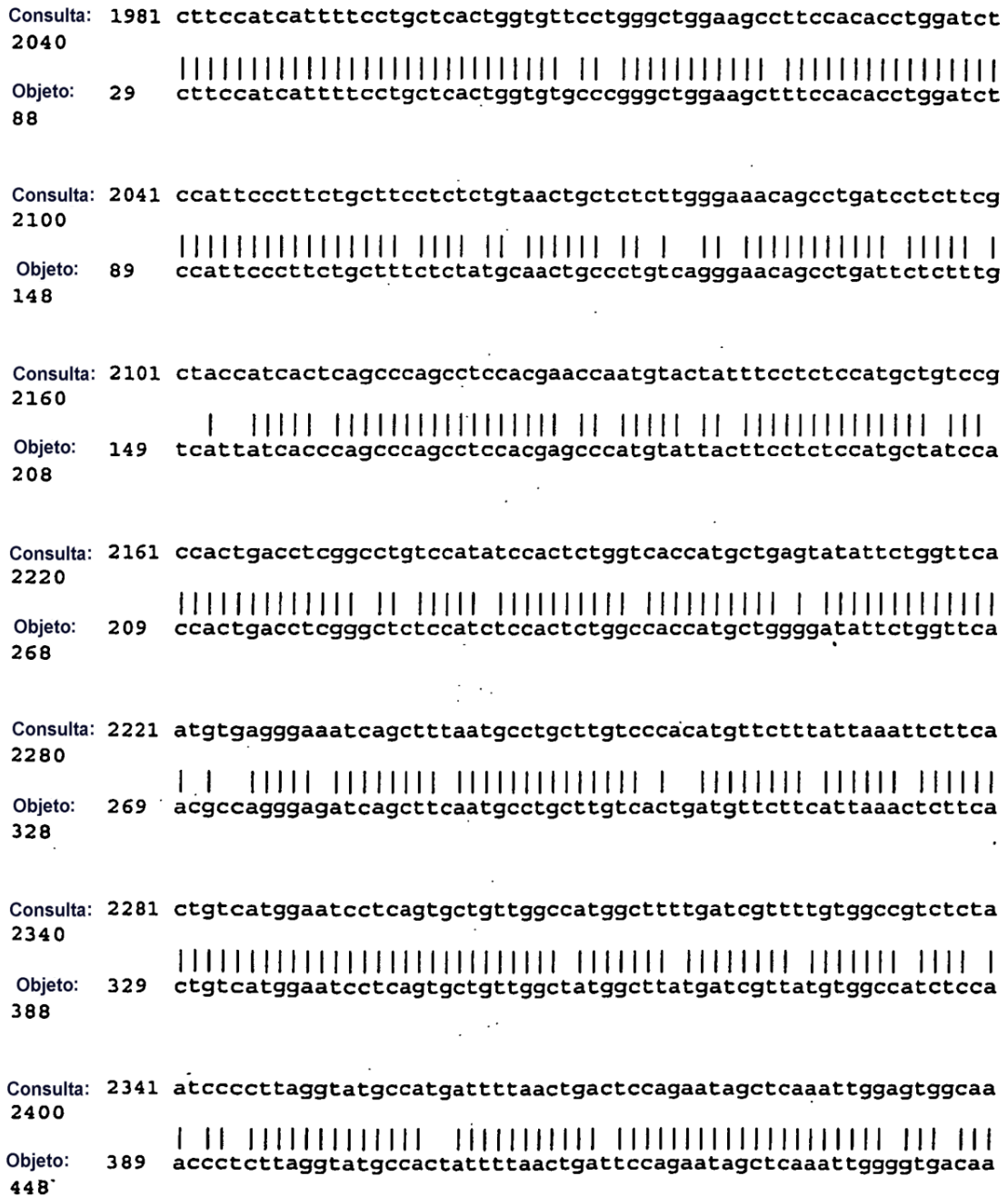
```

1  MYYFLSMLSA  TDGLSISTL  VTMLSIFWFN  VREISFNACL  SHMFFIKFFT  VMESSVLLAM
61  AFDRFVAVSN  PLYAMILTD  SRIAQIGVAS  VIRGLLMLTP  MVALLIRLSY  CHSQVLHHSY
121 CYHPDVMKLS  CTDTRINSAV  GLTAMFSTVG  VDLLLLILSY  VLIIRTVLSV  ASPEERKETF
181 STCVSHIVAF  AIYYIPLISL  SIVHRFGPQA  PAYVHTMIAN  TYLLTSPLMN  PVIYSVKTKQ
241 IRRAVIKILH  SKET
    
```

Figura 4A Alineamiento de secuencias de ácido nucleico de 238P1B2 v.1 con receptor olfatorio de ratón MOR14-1.

>gi|18479243|gb|AY072973.1| Gen del receptor olfatorio MOR14-1 de *Mus musculus*,
 cds completa Longitud =945

Puntuación = 793 bits (400), Expectativa =0,0
 Identities = 784/912 (85 %).
 Cadena = Más / Más

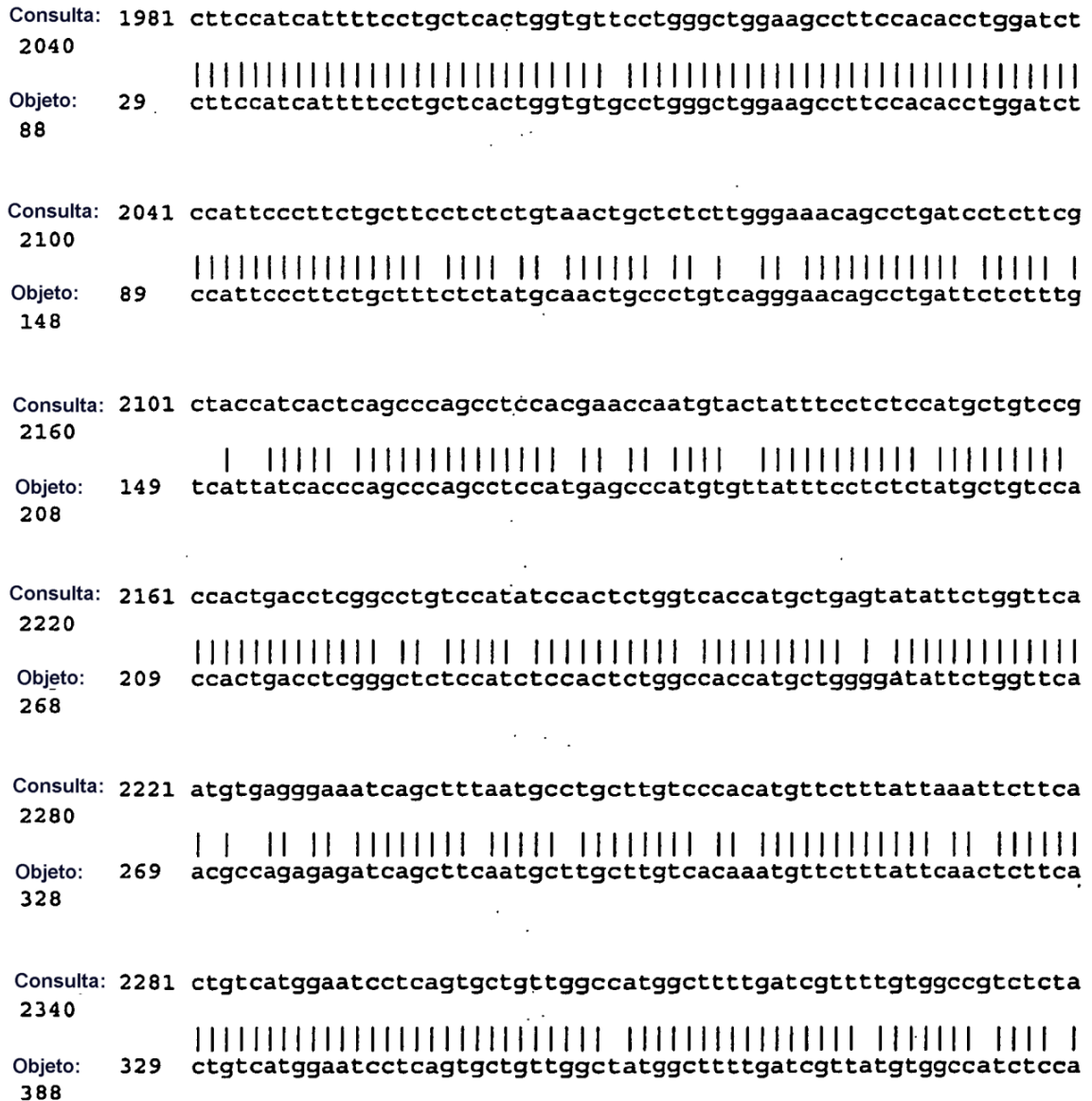


Consulta: 2881 attcceaagaaa 2892
 | | | | | | | | | |
Objeto: 929 aatctaaagaaa 940

Figura 4B Alineamiento de secuencias de ácido nucleico de 238P1B2 v.1 con receptor olfatorio de ratón MOR14-10.

>gi|18480765|gb|AY073734.1| gen del receptor olfatorio MOR14-10 de *Mus musculus*,
 cds completa Longitud = 993

Puntuación = 607 bits (306), Expectativa = e-170
 Identities = 756/906 (83 %)
 Cadena = Más / Más



Consulta: 2821 accctgtcatctacagtgtgaaaaccaaacagatacgtagagctgtgataaaaattctcc
2880

Objeto: 869 accccatcatctacagtgtgaagaccaagaagatacggaggctgtgtgaaagtttcc
928

Consulta: 2881 attcca 2886

Objeto: 929 attcca 934

Figura 4C Alineamiento de secuencias de aminoácidos de 238P1B2 v.1A con receptor olfatorio de ratón MOR14-1.

>gi|18479244|gb|AAL60636.1| (AY072973) receptor olfatorio MOR14-1 [*Mus musculus*]

Longitud =314

Puntuación = 342 bits (878), Expectativa = 2e-93
 Identidades = 210/253 (83 %), Positivos = 230/253 (90 %)

```

238P1B2: 1   MYYFLSMLSATDLGLSISTLVTMLSFVFNVREISFNACLSHMFFIKFFTVMESSVLLAM 60
              MYYFLSMLS TDGLSISTL TML IFWFN REISFNACLS MFFIK FTVMESSVLLAM
MOR14-1: 61   MYYFLSMLSTTDLGLSISTLATMLGIFWFNAREISFNACLSLMFFIKLPTVMESSVLLAM 120

238P1B2: 61   AFDRFVAVSNPLRYAMILTDSRIAQIGVASVIRGLLMLTPMVALLIRLSYCHSQVLHHSY 120
              A+DR+VA+SNPLRYA ILTDSRIAQIGV VIRG +MLTPMVALL RL++C S+VLHHSY
MOR14-1: 121  AYDRYVAISNPLRYATILTDSRIAQIGVTIVIRGTVMLTPMVALLKRLTFCSSRVLHHSY 180

238P1B2: 121  CYHPDVMKLSCTDTRINSAVGLTAMFSTVGVDLLLILLSYVLIIRTVLSVASPEERKETF 180
              C+HPDVMKLSCTDTRIN+AVGLTAM STVGVD +LILLSY+LIIRTVLS+ASPEERK+ F
MOR14-1: 181  CFHPDVMKLSCTDTRINNAVGLTAMISTVGVDLILLSYILIIRTVLSIASPEERKKAF 240

238P1B2: 181  STCVSHIVAFAIYYIPLISLSIVHFRFGQPAPAYVHTMIANTYLLISPLMNPVIYSVKTKQ 240
              STC+SHI A AI+YIPLISLS VHRFG +AP YVHTMIANTYLLI P+MNP+IYSVKTKQ
MOR14-1: 241  STCISHIGAVAIPIYIPLISLSFVHFRFGKRAPPYVHTMIANTYLLIPPVMNPPIIYSVKTKQ 300

238P1B2: 241  IRRAVIKILHSKE 253
              IR+AVIK+ SKE
MOR14-1: 301  IRKAVIKVFQSKE 313
    
```

Figura 4D Alineamiento de secuencias de aminoácidos de 238P1B2 v.1A con GPCR PHOR-1 específico de próstata.

Puntuación = 254 bits (650), Expectativa = 4e-67
 Identidades = 120/250 (48 %), Positivos = 177/250 (70 %), Huecos = 1/250 (0 %)

```

238P1B2: 1  MYFSLMLSATDLGLSISTLVTMLSFVFNREISFNACL SHMFFIKFFTVMESSVLLAM 60
MY FL MLS D+ +S S++ ML+IFWFN I F+ACL +F I + MES+VLLAM
PHOR-1: 62  MYIFLCMLSGIDILISTSSMPKMLAIFWFNSTTIQFDACLLQIFAIHSLSGMESTVLLAM 121

238P1B2: 61  AFDRFVAVSNPLRYAMILTDSRIAQIGVASVIRGLLMLTPMVALLIRLSYCHS QVLHHSY 120
AFDR+VA+ +PLR+A +LT R+ +IGVA+V+RG ++ P+ + +L +C S +L HSY
PHOR-1: 122 AFDRYVAICHPLRHATVLTLPRTKIGVAAVVRGAALMAPLPVFIKQLPFCRSNILSHSY 181

238P1B2: 121 CYHPDVMKLSCTDTRINSAVGLTAMFSTVGVDLLLILLSYVLIIRT VLSVASPEERKETF 180
C H DVMKL+C D R+N GL + S +G+D LLI SY+LI++TVL + + E + + F
PHOR-1: 182 CLHQDVMKLCDDIRVNVVYGLIIVIISAIGLDSLLISFSYLLILKTVLGL-TREAQAKAF 240

238P1B2: 181 STCVSHIVAFAIYYIPLISLSIVHFRFGPQAPAYVHTMIANTYLLISPLMNPVIYSVKTKQ 240
TCVSH+ A I+Y+P I LS+VHRF + + + ++AN YLL+ P++NP++Y VKTK+
PHOR-1: 241 GTCVSHVCAVFIYVPIGLSMVHRFSKRDRSPLPVILANIYLLVPPVLPNPIVYGVKTKE 300

238P1B2: 241 IRRAVIKILH 250
IR+ ++++ H
PHOR-1: 301 IRQRILRLFH 310
    
```

Figura 4E Alineamiento de 238P1B2 variante 1 con receptor olfatorio humano 5112.

Identidades = 154/252 (61 %), Positivos = 201/252 (79 %)

```

Consulta: 1  MYFSLMLSATDLGLSISTLVTMLSFVFNREISFNACL SHMFFIKFFTVMESSVLLAM 60
MYFSLMLSATDL LS+ TL T L +FWF REI+ NAC++ MFF+ FT MES VLLAM
Objeto: 73  MYFSLMLSATDLSLSLCTLTSTLGVWFWEAREINLNACIAQMFFLHGFTFMESGVLLAM 132

Consulta: 61  AFDRFVAVSNPLRYAMILTDSRIAQIGVASVIRGLLMLTPMVALLIRLSYCHS QVLHHSY 120
AFDRFVA+ PLRY ILT++RIA+IG++ +IR + ++ P++ + RLS+C S VL HSY
Objeto: 133 AFDRFVAICYPLRYTTILTNRARIAKIGMSMLIRNVAVMLPVMLFVKRLSFCSSMVLSHSY 192

Consulta: 121 CYHPDVMKLSCTDTRINSAVGLTAMFSTVGVDLLLILLSYVLIIRT VLSVASPEERKETF 180
CYH D+++LSCTD RINS +GL A+ ST G D ILLSY+LIIR+VLS+AS EER++ F
Objeto: 193 CYHVDLIQLSCTDNRINSILGLFALLSTTGDFDCPCILLSYILIIRS VLSIASSEERRKAF 252

Consulta: 181 STCVSHIVAFAIYYIPLISLSIVHFRFGPQAPAYVHTMIANTYLLISPLMNPVIYSVKTKQ 240
+TC SHI A +I+Y+PLISLS+VHR+G AP +VH ++AN +LLI P++NP+IYSVK KQ
Objeto: 253 NTCTSHISAVSIFYLPLISLSLVHRYGHSAPPFVHIIMANVFLLIIPPVLPNPIIYSVKIKQ 312

Consulta: 241 IRRAVIKILH 252
I++A+IK+L K
Objeto: 313 IQKAIKVLIQK 324
    
```

Figura 4F Alineamiento generado por Clustal de las tres variantes de 238P1B2, que representa que 238P1B2 V1B contiene 62 aa adicionales en su extremo N-terminal en relación con V1A, y que 238P1B2 V2 porta una mutación puntual I a T en el aa 225 en relación con V1A.

```

v. 1A -----
v. 1B FITSTLQNITSTSIIFLLTGVPGLEAFHTWISIPFCFLSVTALLGNSLILFATITQPSLH
v. 2 -----

v. 1A --MYYFLSMLSATDLGLSISTLVTMLSIFWFNVREISFNACL SHMFFIKFFTVMESSVLL
v. 1B EPMYYFLSMLSATDLGLSISTLVTMLSIFWFNVREISFNACL SHMFFIKFFTVMESSVLL
v. 2 --MYYFLSMLSATDLGLSISTLVTMLSIFWFNVREISFNACL SHMFFIKFFTVMESSVLL
*****

v. 1A AMAFDRFVAVSNPLRYAMILTDSRIAQIGVASVIRGLLMLTPMVALLIRLSYCHSQVLHH
v. 1B AMAFDRFVAVSNPLRYAMILTDSRIAQIGVASVIRGLLMLTPMVALLIRLSYCHSQVLHH
v. 2 AMAFDRFVAVSNPLRYAMILTDSRIAQIGVASVIRGLLMLTPMVALLIRLSYCHSQVLHH
*****

v. 1A SYCYHPDVMKLSCTDTRINSAVGLTAMFSTVGVDLLLILLSYVLIIRTVLSVASPEERKE
v. 1B SYCYHPDVMKLSCTDTRINSAVGLTAMFSTVGVDLLLILLSYVLIIRTVLSVASPEERKE
v. 2 SYCYHPDVMKLSCTDTRINSAVGLTAMFSTVGVDLLLILLSYVLIIRTVLSVASPEERKE
*****

v. 1A TFSTCVSHIVAFAIYYIPLISLSIVHRFGPQAPAYVHTMIANTYLLISPLMNPVIYSVKT
v. 1B TFSTCVSHIVAFAIYYIPLISLSIVHRFGPQAPAYVHTMIANTYLLISPLMNPVIYSVKT
v. 2 TFSTCVSHIVAFAIYYIPLISLSIVHRFGPQAPAYVHTMIANTYLLTSPLMNPVIYSVKT
*****

v. 1A KQIRRAVIKILHSKET
v. 1B KQIRRAVIKILHSKET
v. 2 KQIRRAVIKILHSKET
*****

```

Figura 5A
Perfil de hidrofilia de 238P1B2 variante 1a
(Hopp T.P., Woods K.R., 1981. Proc. Nati. Acad. Sci. U.S.A. 78:3824-3828)

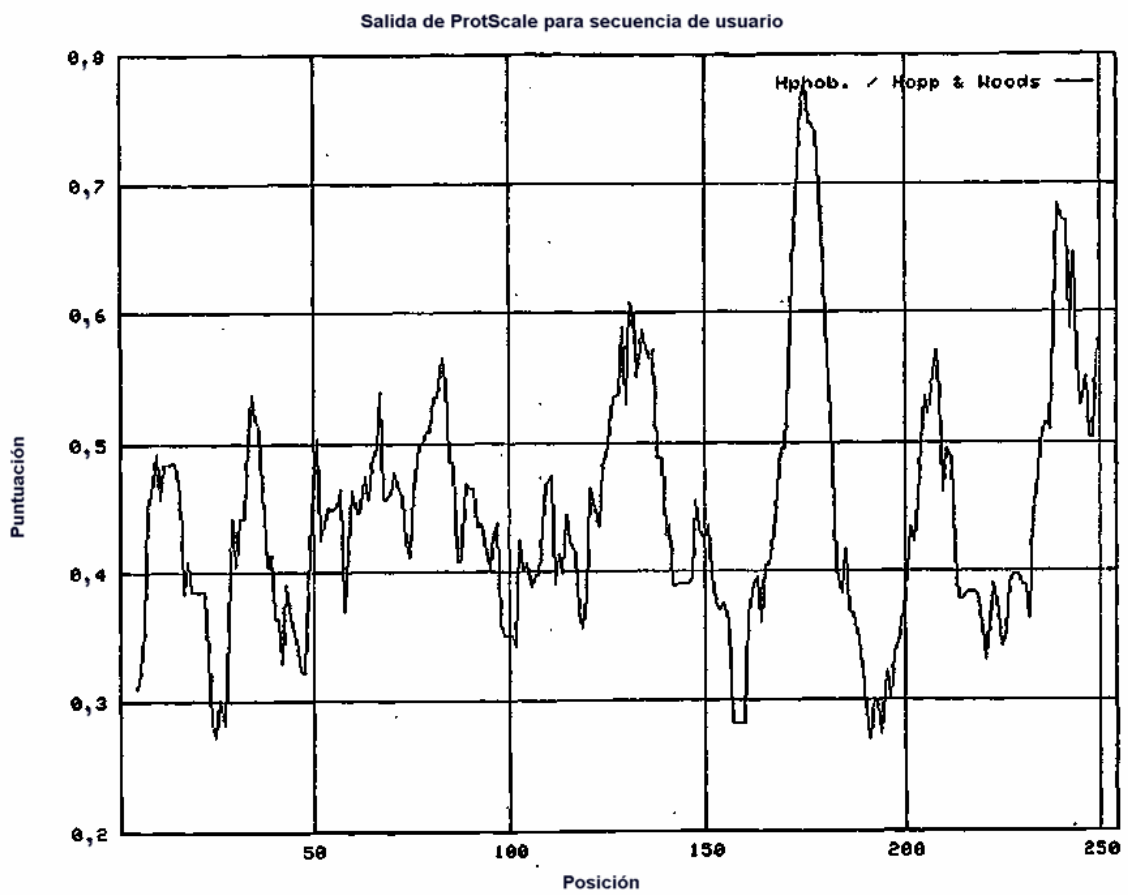


Figura 5B
Perfil de hidrofilia de 238P1B2 variante 1b
(Hopp T.P., Woods K.R., 1981. Proc. Nati. Acad. Sci. U.SA 78: 3824-3828)

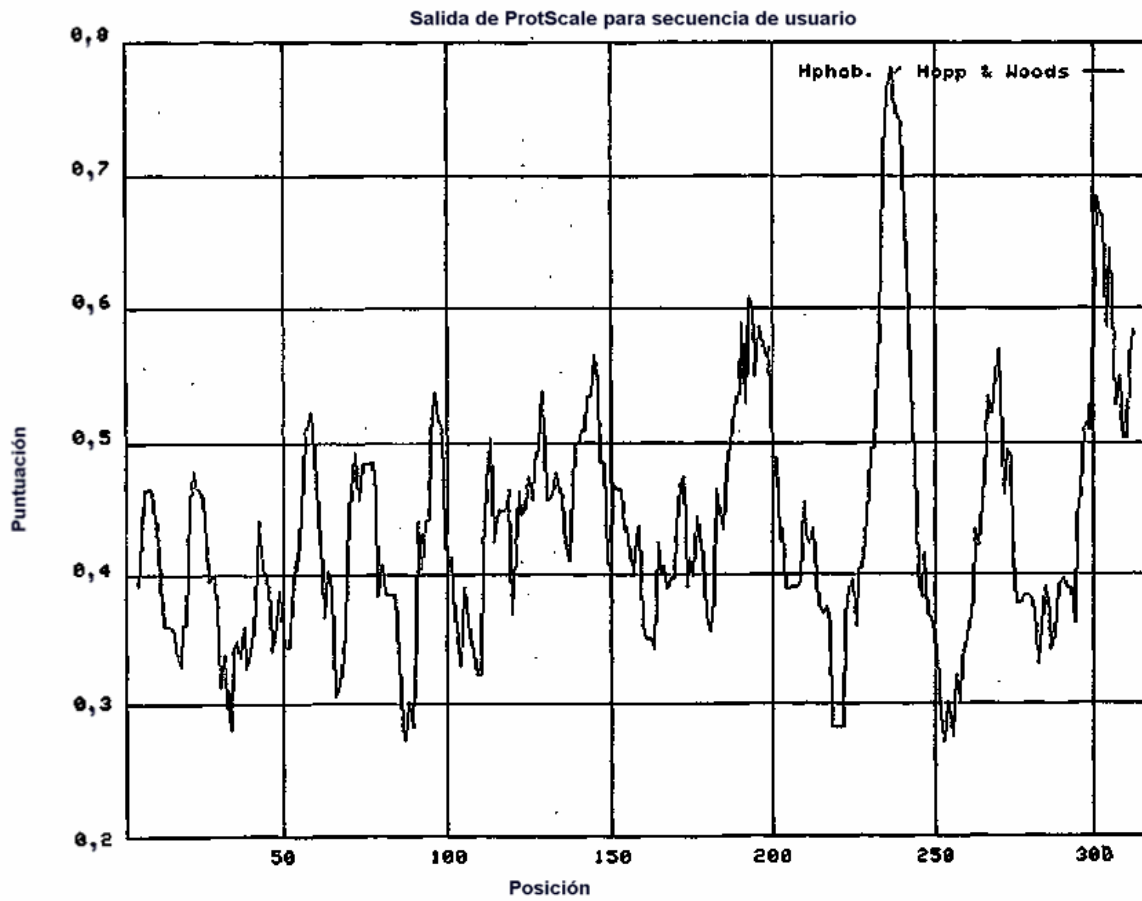


Figura 6A
Perfil hidropático de 238P1B2 variante 1a
(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157: 105-132)

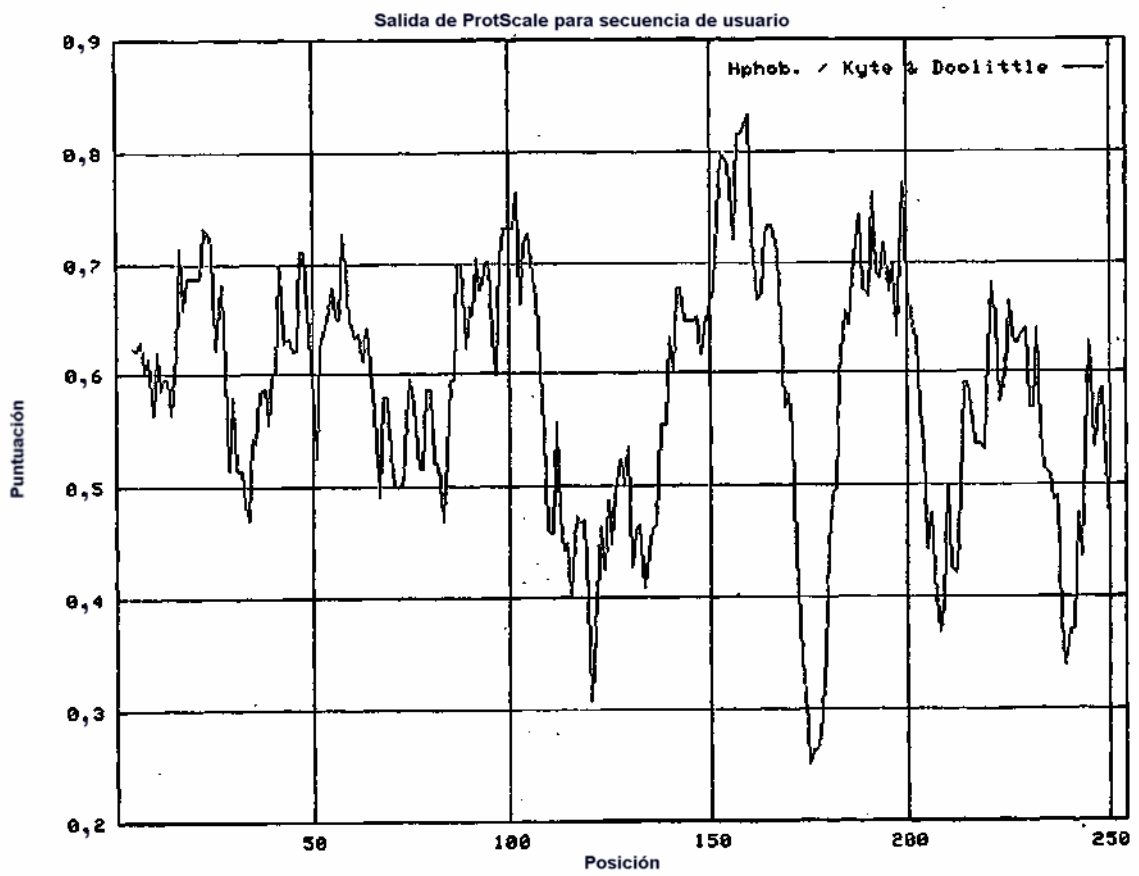


Figura 6B
Perfil hidropático de 238P1B2 variante 1b
(Kyte J., Doolittle R.F., 1982. J. Mol. Biol. 157: 105-132)

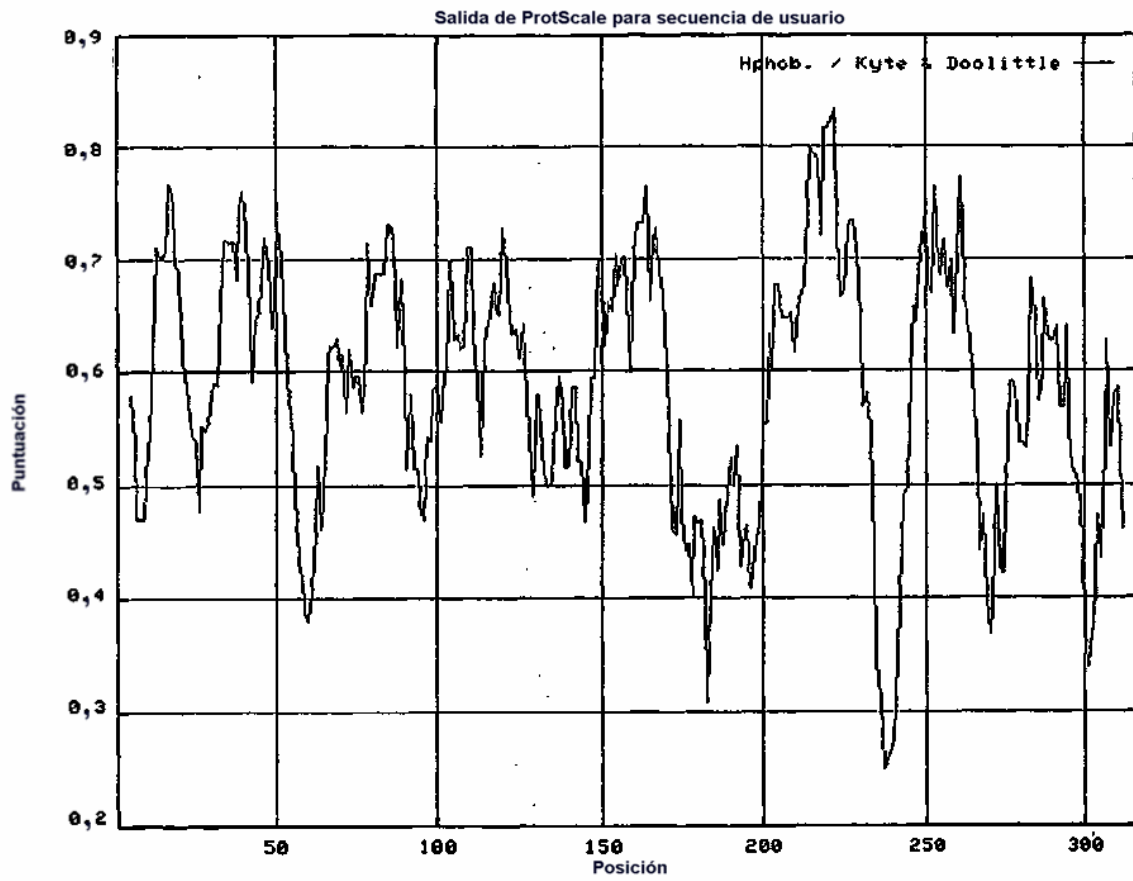


Figura 7A
Perfil de % de restos accesibles de 238P1B2 variante 1a
(Janin J., 1979. Nature 277: 491-492)

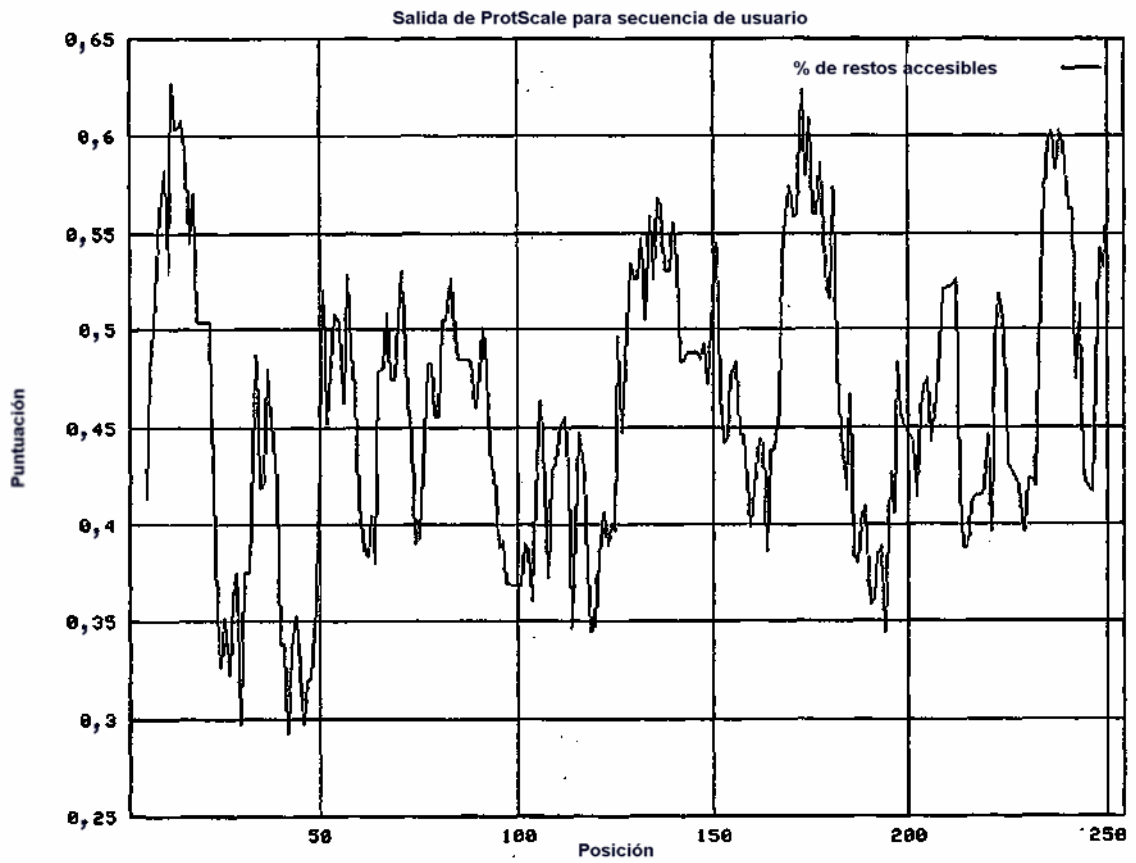


Figura 7B
Perfil de % de restos accesibles de 238P1B2 variante 1b
(Janin J., 1979. Nature 277: 491-492)

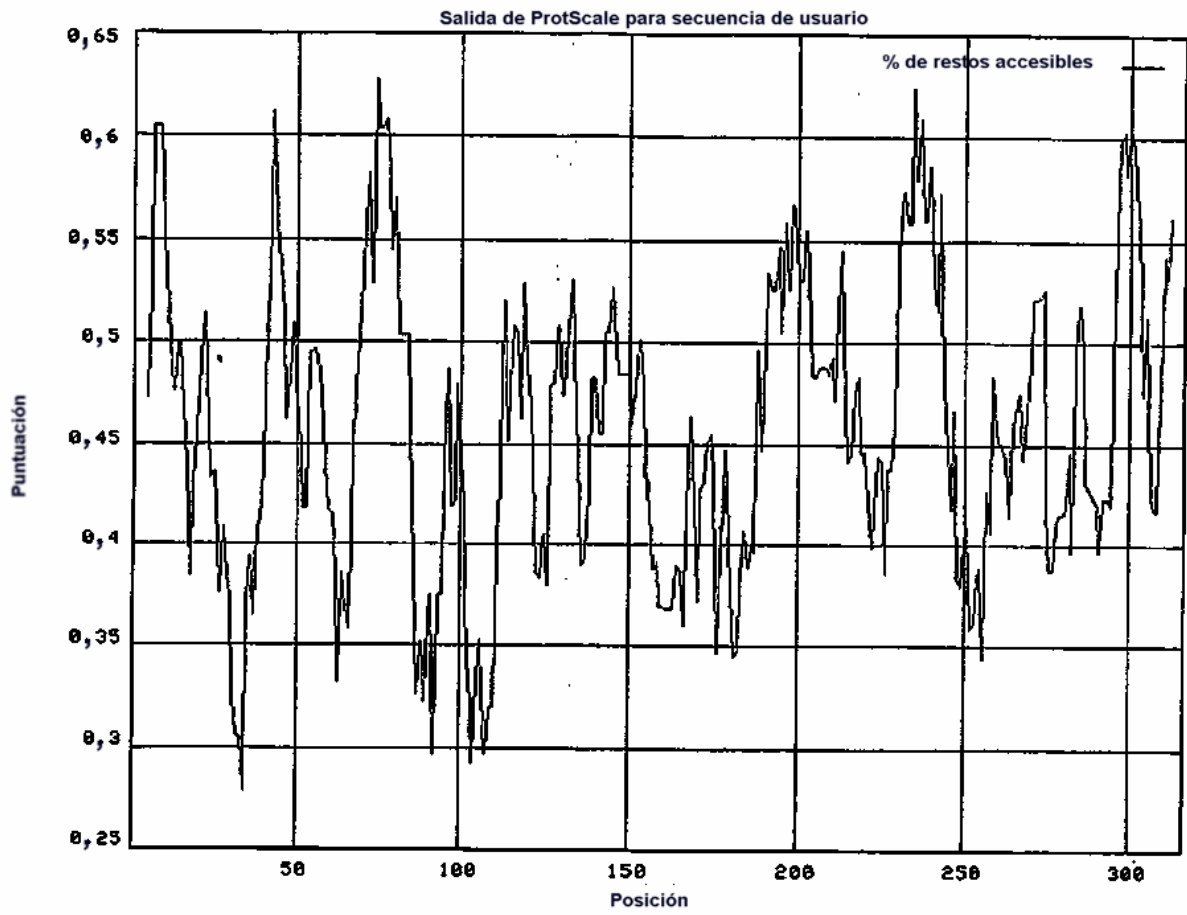


Figura 8A
Perfil de flexibilidad media de 238P1B2 variante 1a
(Bhaskaran R., Ponnuswamy P.K., 1988. Int. J. Pept. Protein Res. 32:242-255)

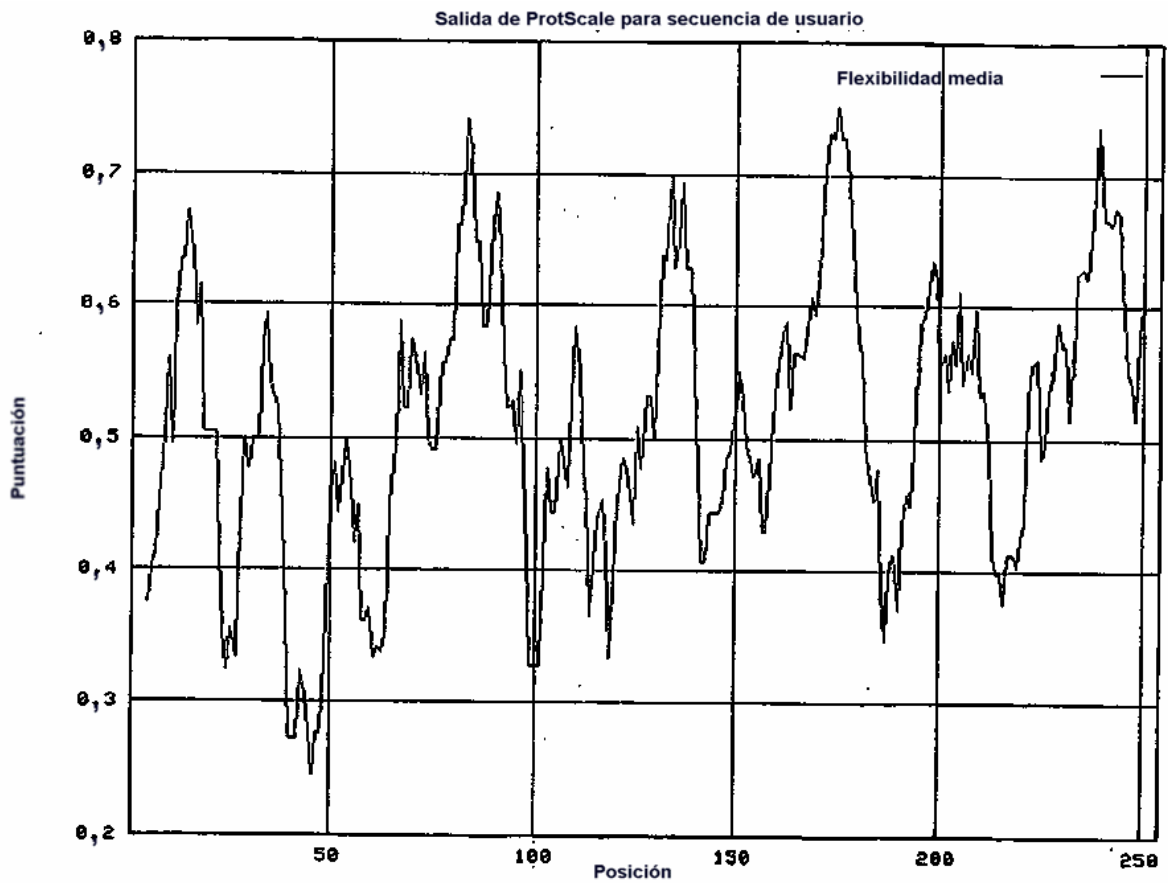


Figura 8B
Perfil de flexibilidad media de 238P1B2 variante 1b
(Bhaskaran R., Ponnuswamy P.K., 1988. Int. J. Pept. Protein Res. 32: 242-255)

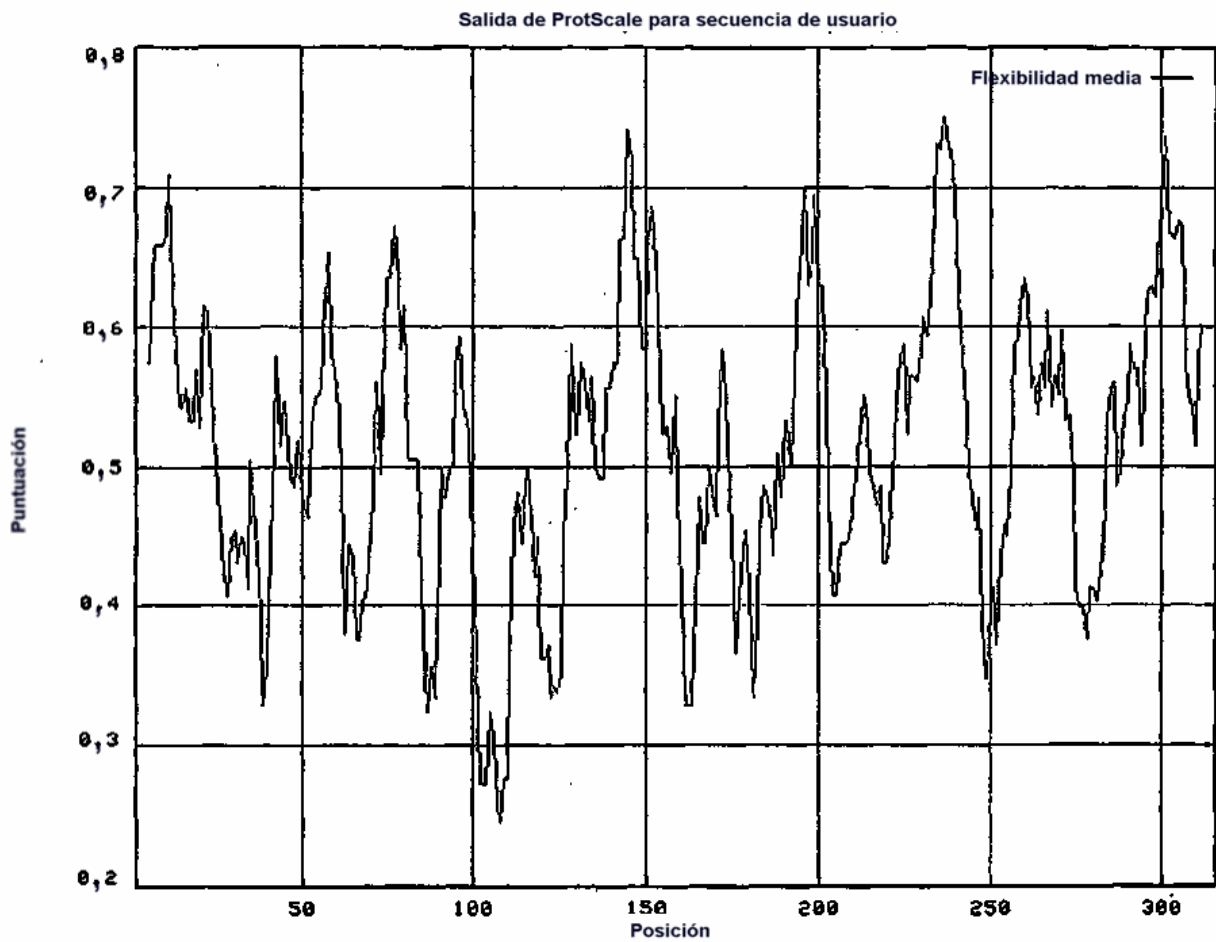


Figura 9A
Perfil de giro beta de 238P1B2 variante 1a
(Deleage, G., Roux B. 1987. Protein Engineering 1:289-294)

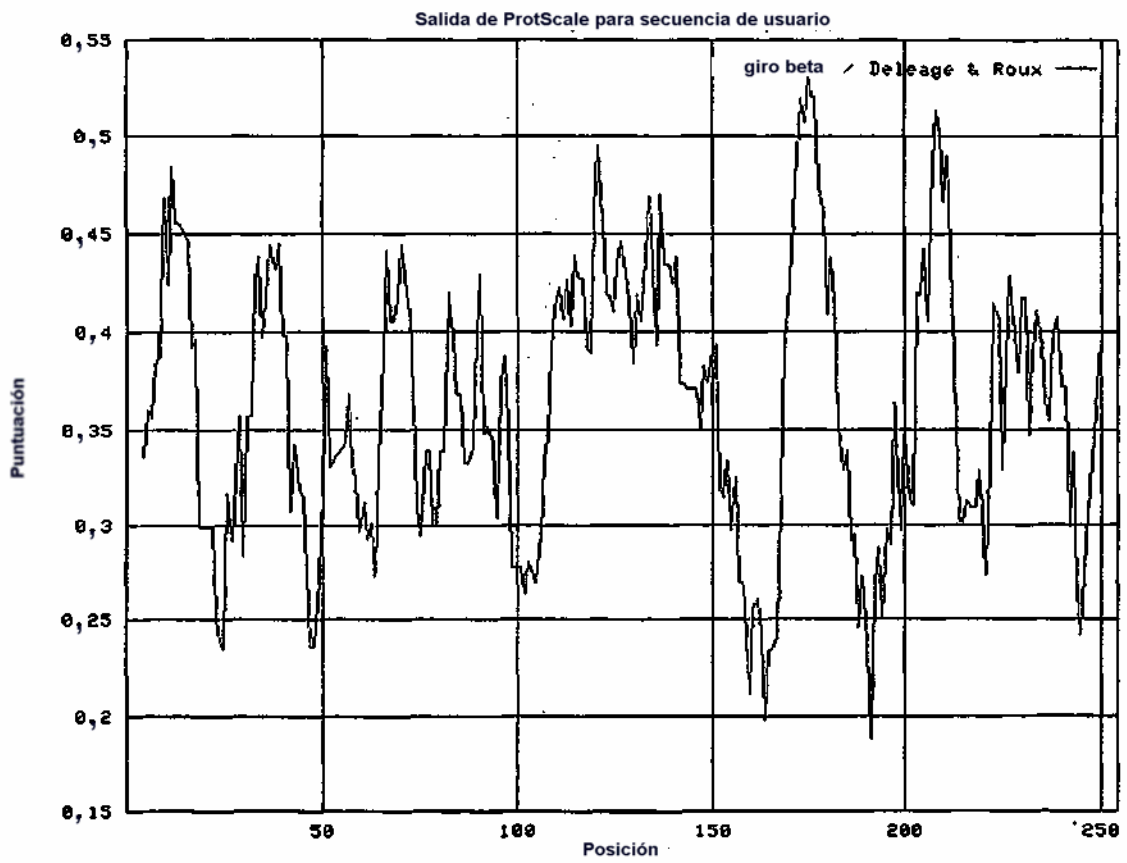


Figura 9B
Perfil de giro beta de 238P1B2 variante 1b
(Deleage, G., Roux B. 1987. Protein Engineering 1: 289-294)

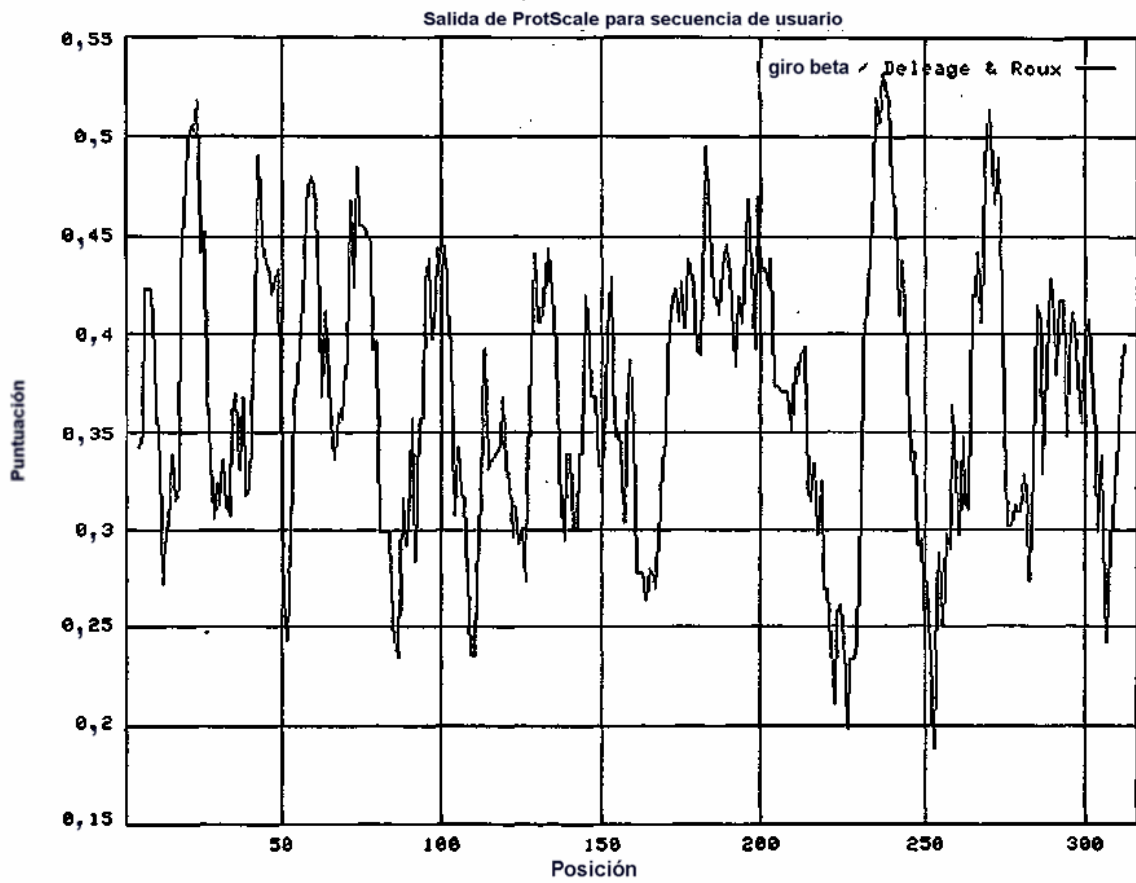


Figura 10

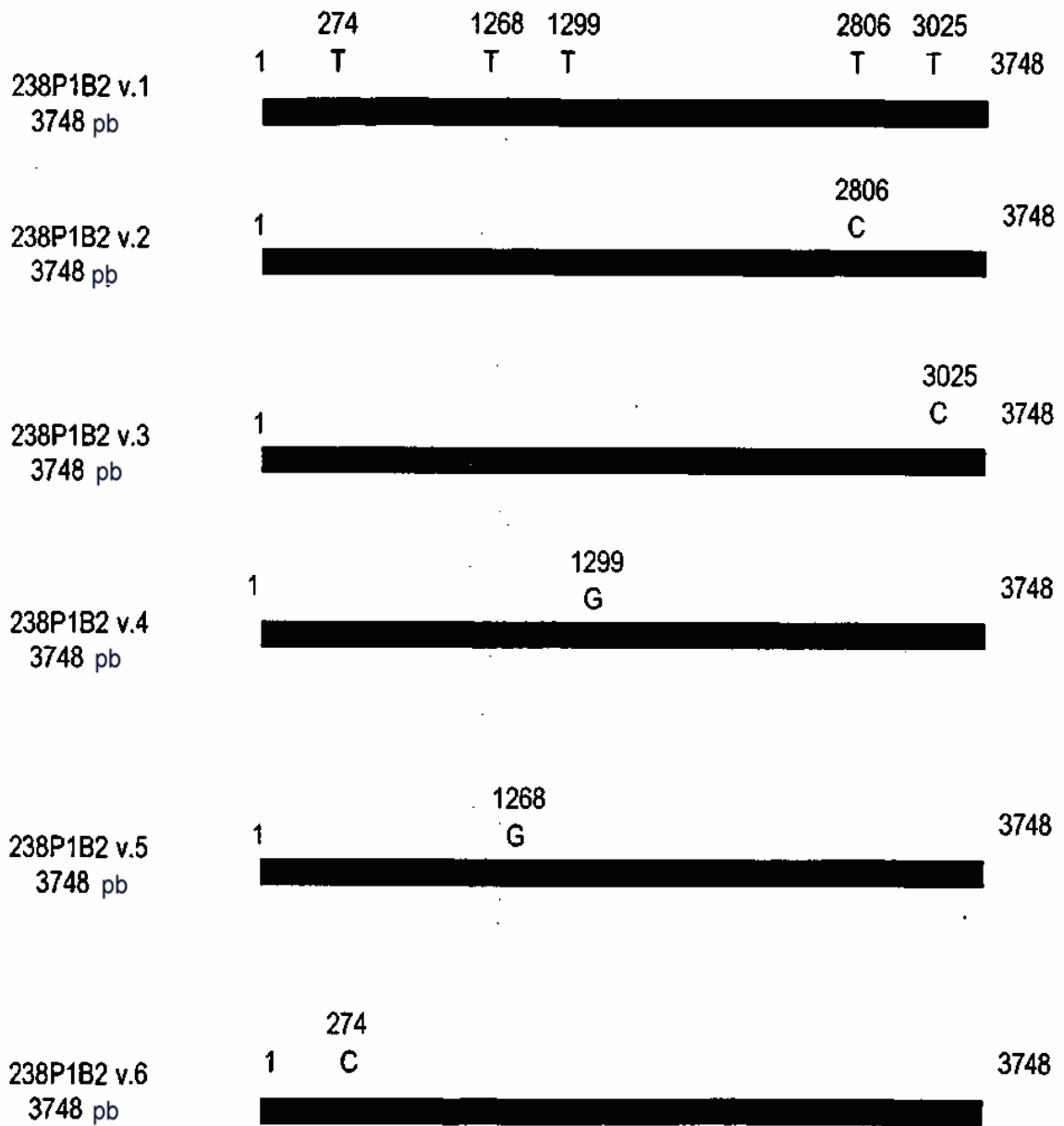


Figura 11

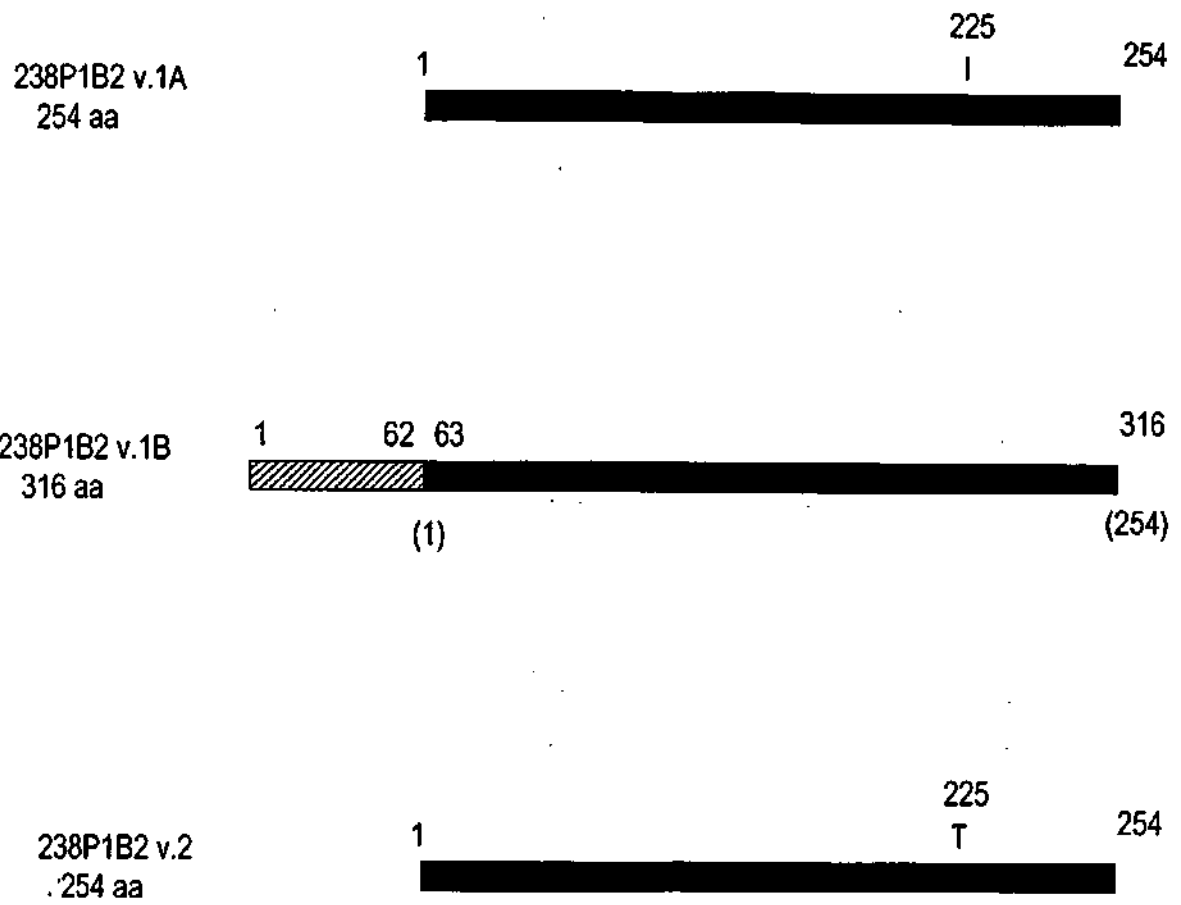
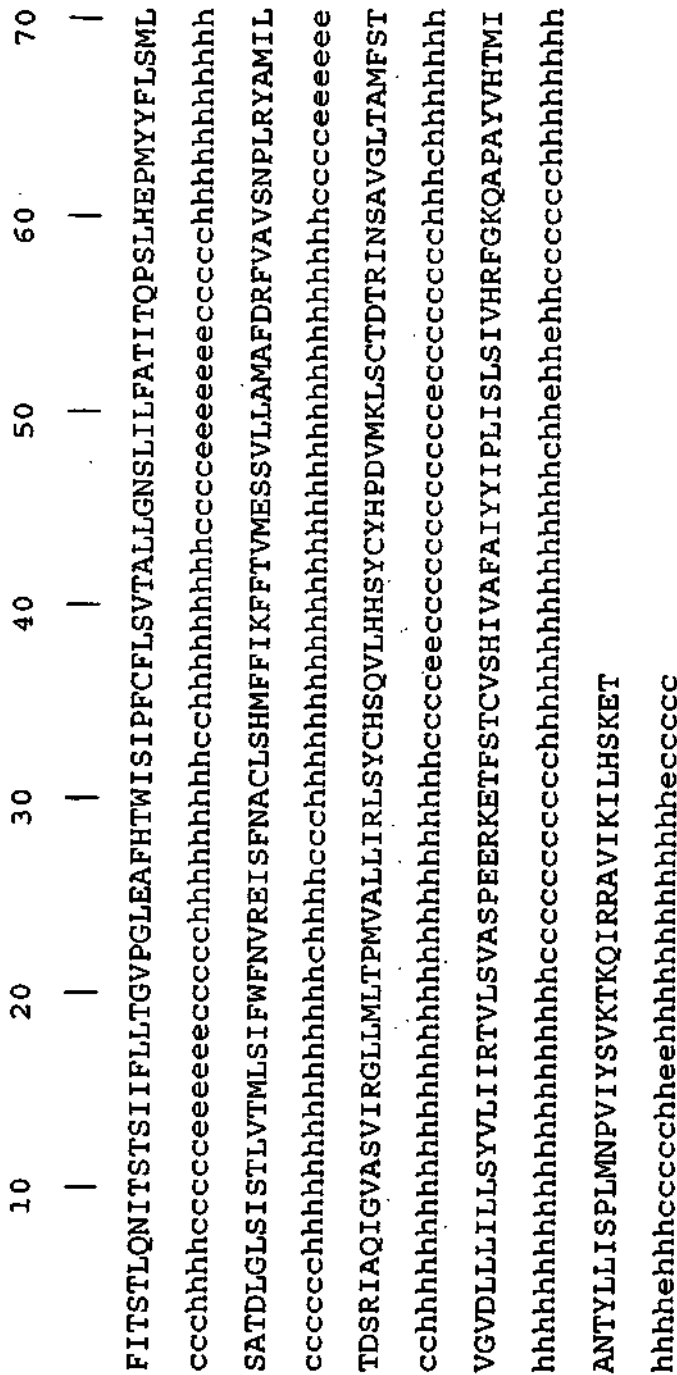


Figura 12B Predicción de estructura secundaria de 238P1B2 variante 1b



Hélice alfa (h): 61,71 %
 Hebra extendida (e): 8,86 %
 Bucle aleatorio (c): 29,43 %

Figura 12 C Predicción transmembrana para 238P1B2 variante 1a

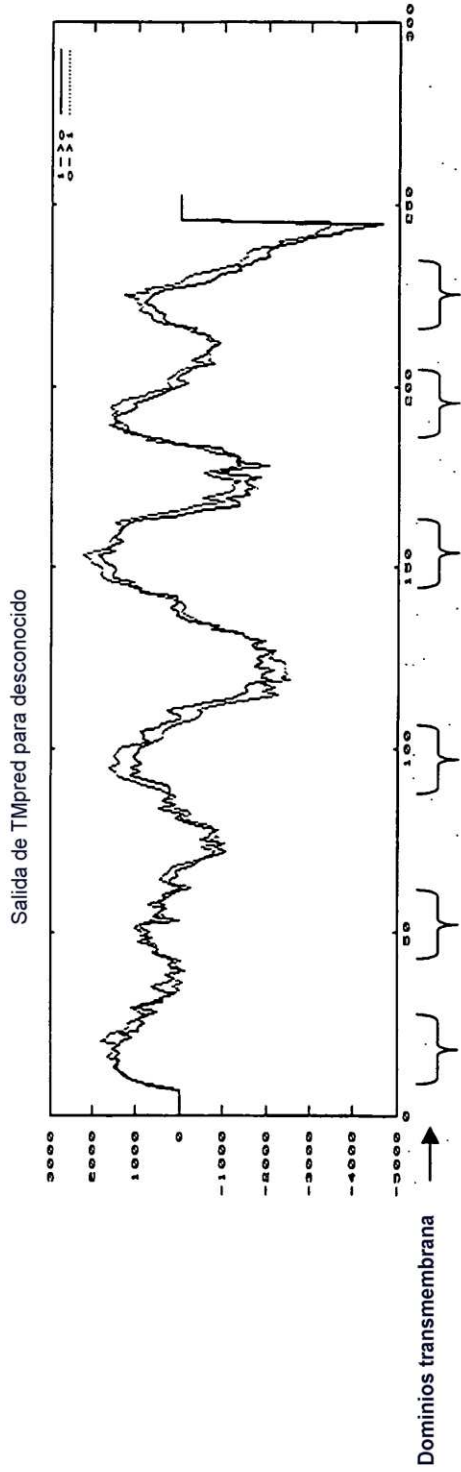


Fig. 12D

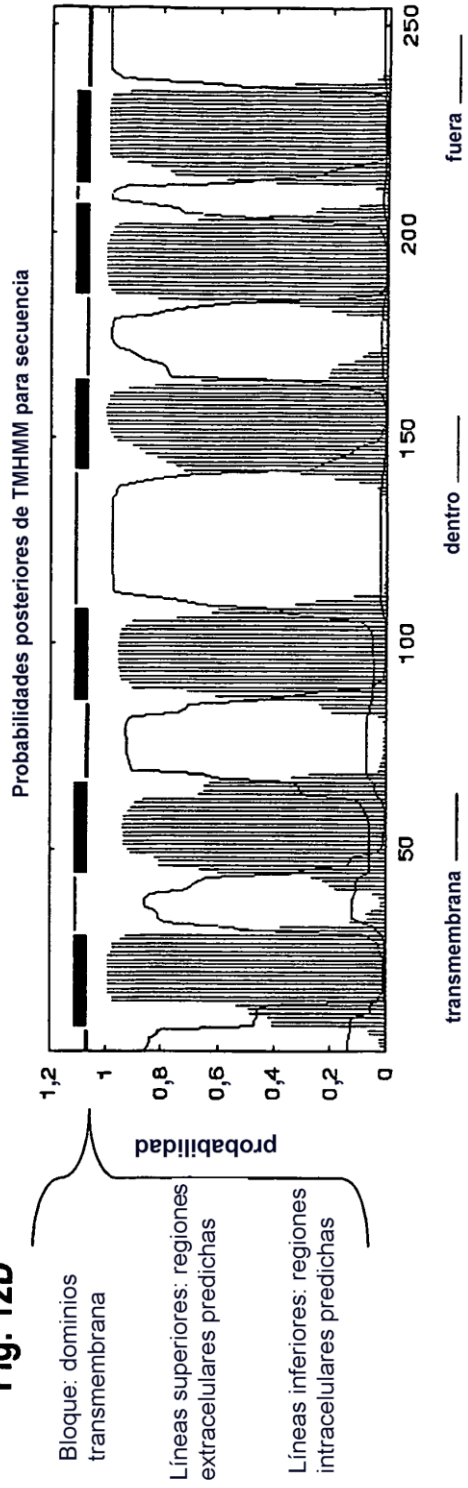


Figura 12E Predicción transmembrana para 238P1B2 variante 1b

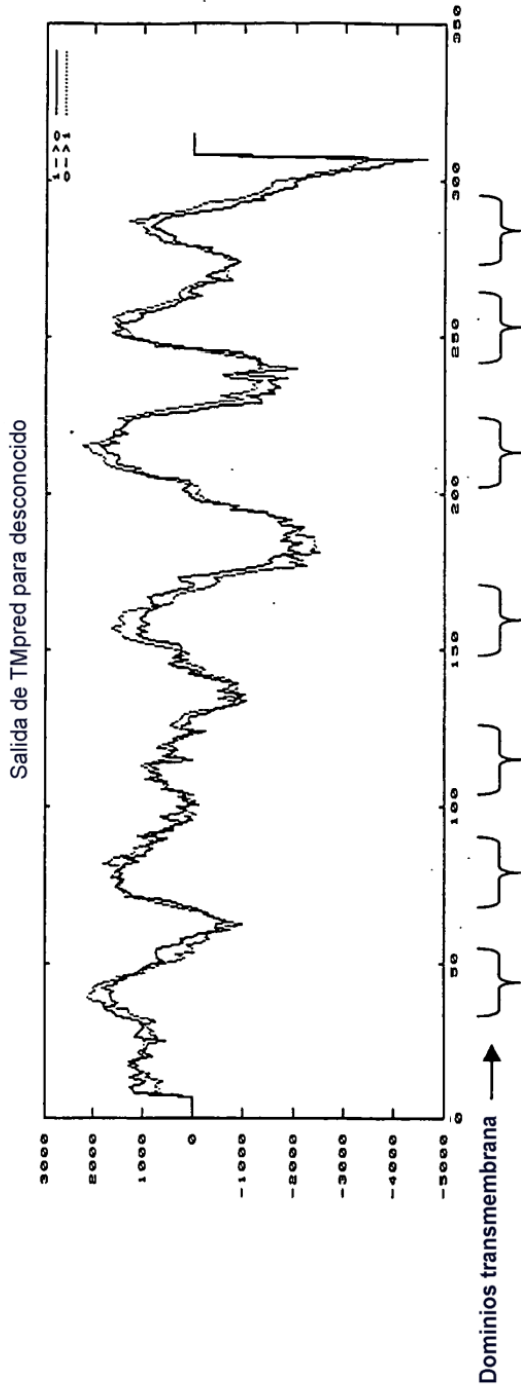


Fig. 12F

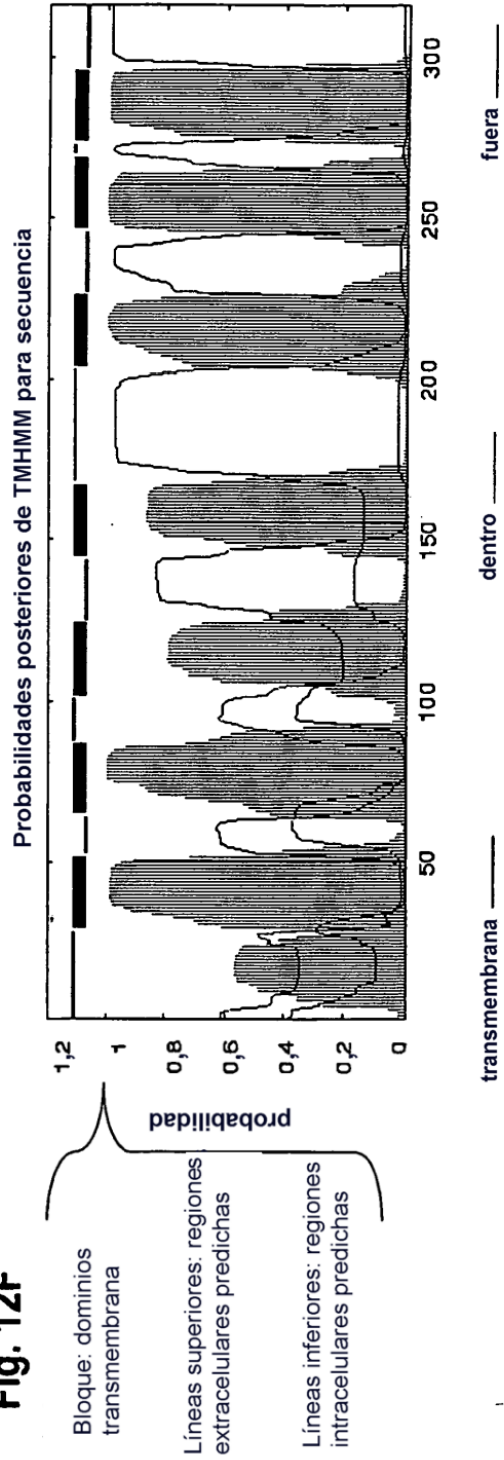
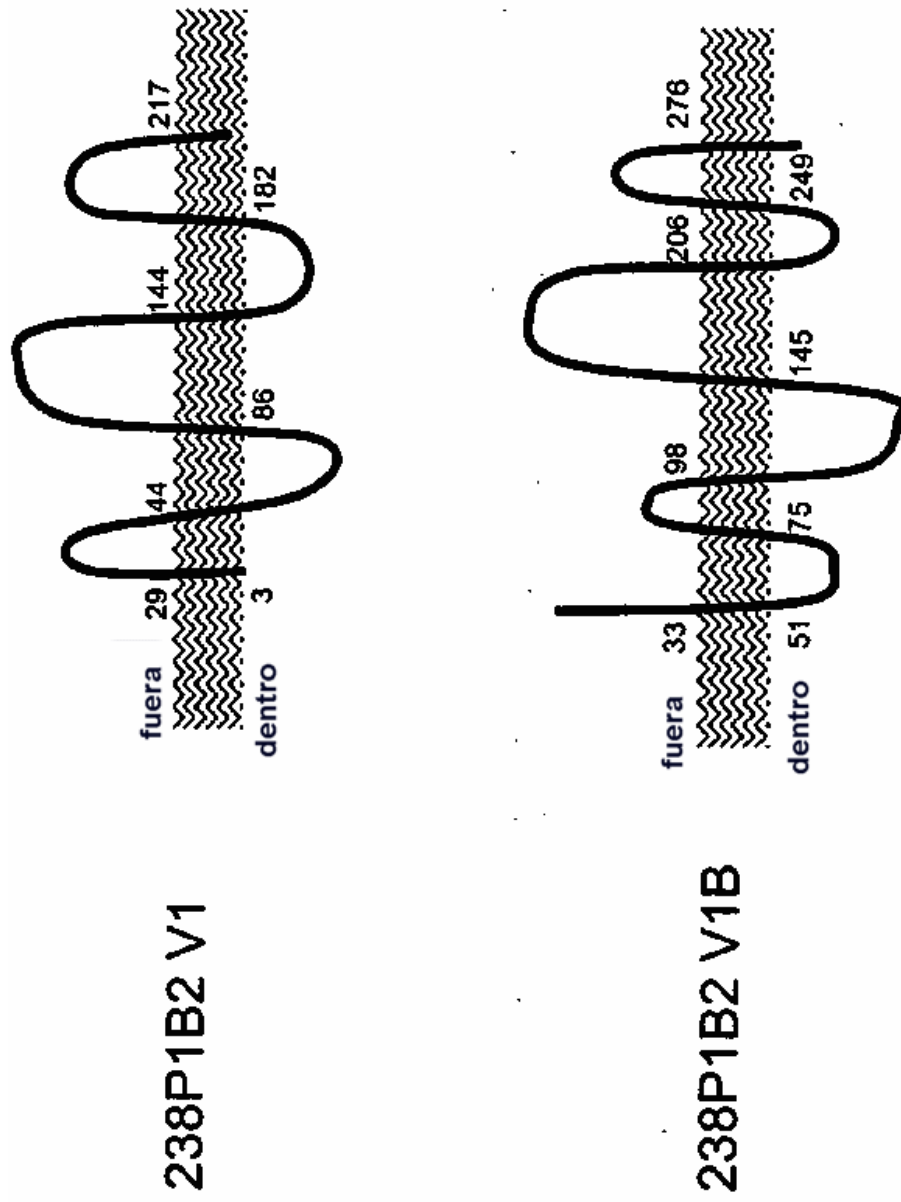
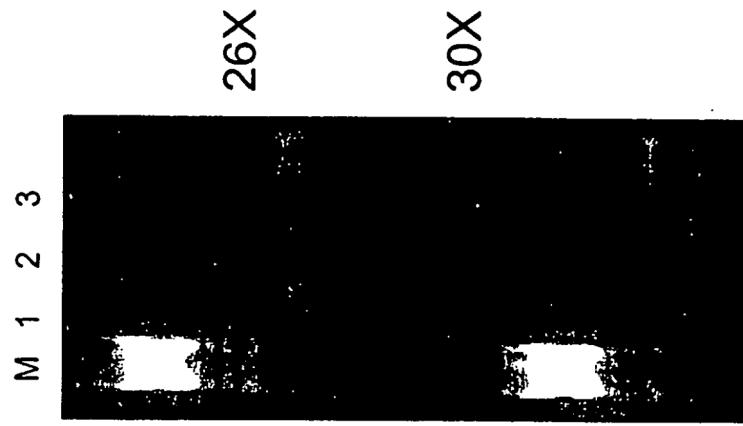


Figura 13: Topología y estructura probables de variantes de 238P1B2



Los números se refieren a los aminoácidos al comienzo de cada dominio transmembrana

Figura 14 Expresión de 238P1B2 por RT-PCR



M = Marcador
1. Grupo vital 1
2. Grupo vital 2
3. Grupo de cáncer de próstata

Figura 15 Expresión de 238P1B2 en tejidos normales

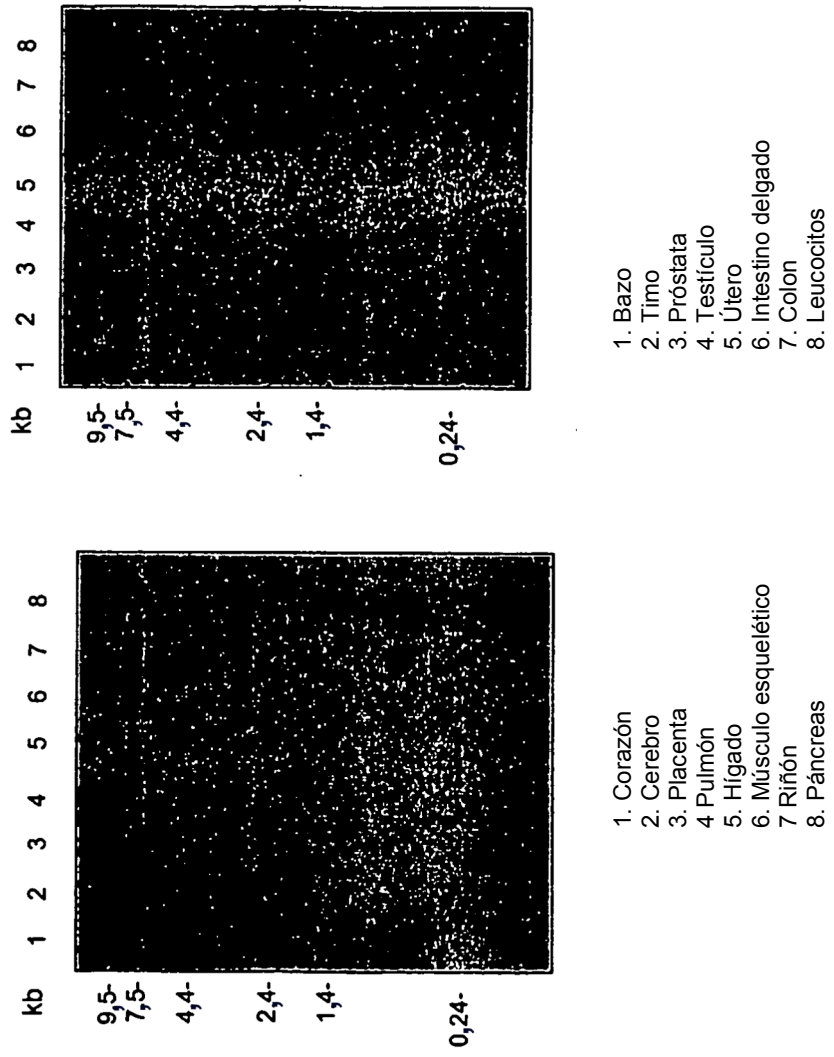


Figura 16 Expresión de 238P1B2 en múltiples tejidos normales

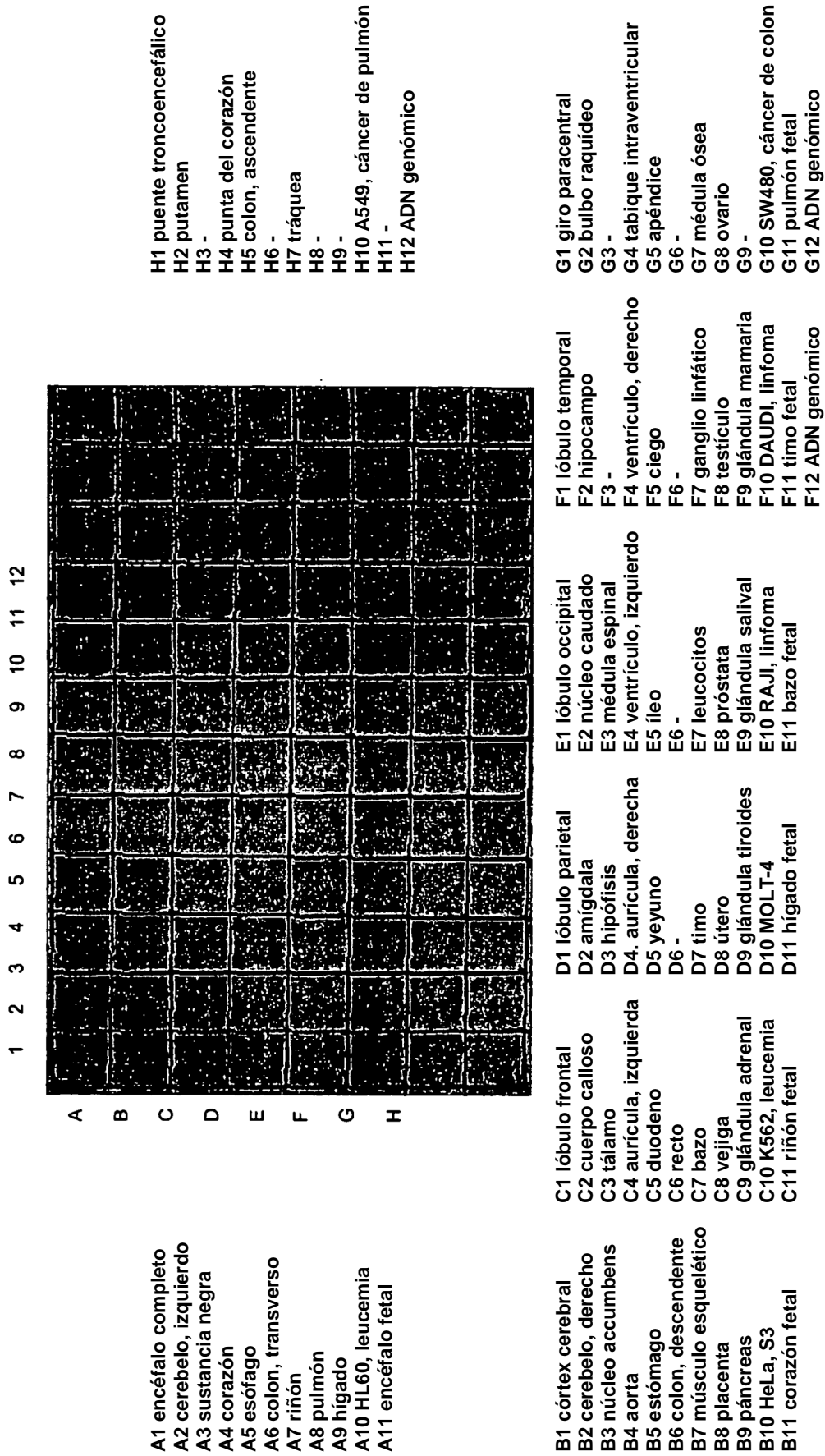


Figura 17 Expresión de 238P1B2 en muestras de ensayo de cáncer de paciente humano y en tejidos normales

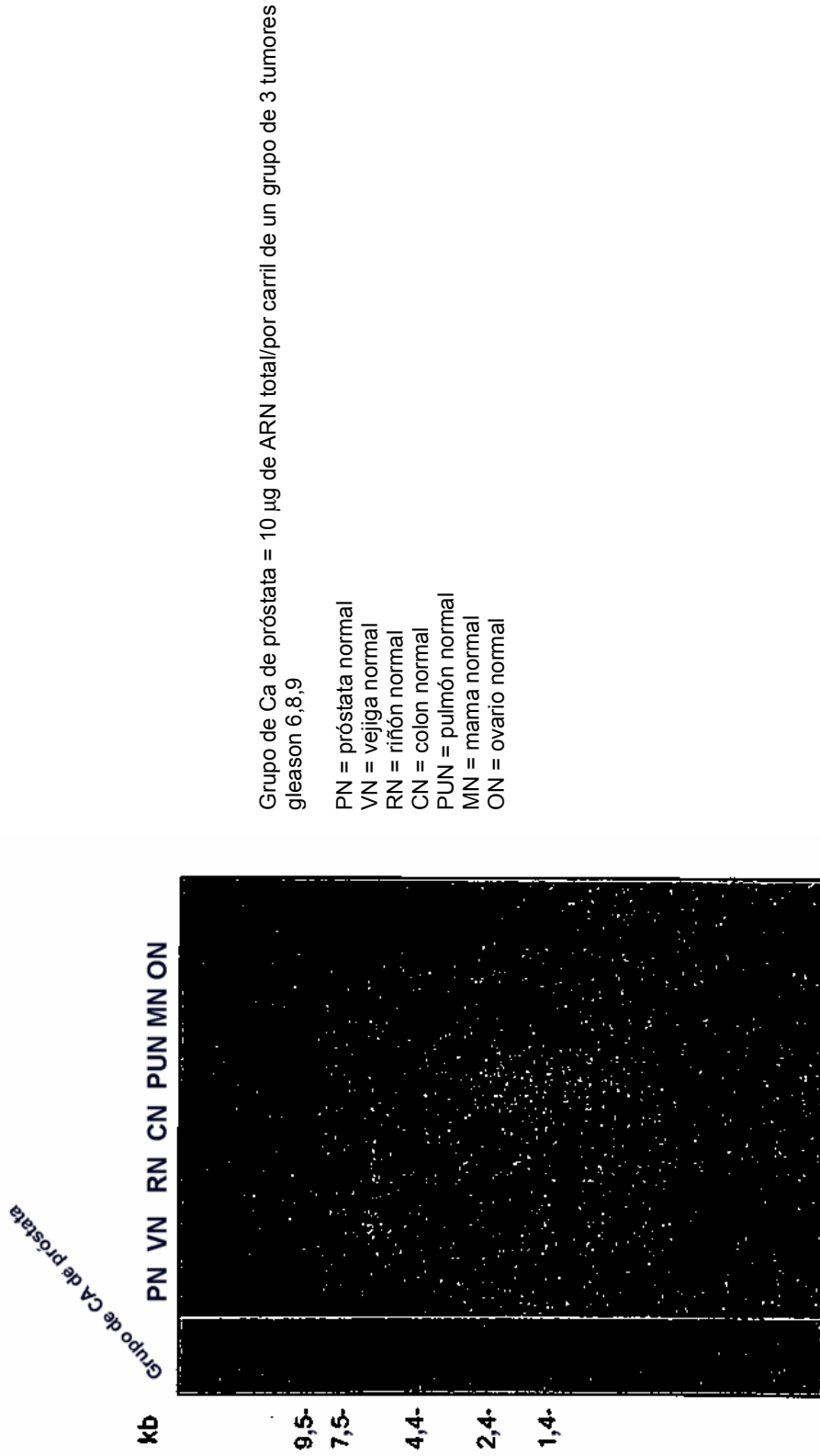


Figura 18 Expresión de 238P1B2 en muestras de ensayo de paciente con cáncer de próstata

