(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 986 963**

(51) Int. Cl.:

C12Q 1/6883 (2008.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **02.02.2018 PCT/IB2018/000181**

(87) Fecha y número de publicación internacional: **07.09.2018 WO18158632**

(96) Fecha de presentación y número de la solicitud europea: **02.02.2018 E 18749101 (4)**

(97) Fecha y número de publicación de la concesión europea: **10.04.2024 EP 3577236**

(54) Título: **Métodos para evaluar el riesgo de desarrollar una enfermedad viral mediante una prueba genética**

(30) Prioridad:

03.02.2017 US 201762454676 P
23.06.2017 US 201762524324 P
30.06.2017 US 201715639591

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
13.11.2024

(73) Titular/es:

PML SCREENING, LLC (25.0%)
23 Corporate Plaza, Suite 150
Newport Beach, CA 92660, US;
UNIVERSITÉ PARIS-SACLAY (25.0%);
THE ASSISTANCE PUBLIQUE-HÔPITAUX DE PARIS (APHP) (25.0%) y
THE INSTITUT NATIONAL DE LA SANTÉ ET DE LARECHERCHE MÉDICALE (INSERM) (25.0%)

(72) Inventor/es:

HATCHWELL, ELI;
EIS, PEGGY S.;
SMITH, III, EDWARD B. y
TAOUFIK, YASSINE

(74) Agente/Representante:

SÁEZ MAESO, Ana

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Métodos para evaluar el riesgo de desarrollar una enfermedad viral mediante una prueba genética

Antecedentes de la divulgación

La leucoencefalopatía multifocal progresiva (PML) es una infección oportunista rara y potencialmente mortal del sistema nervioso central causada por un poliomavirus ubicuo, el virus de JC (JCV). Mientras que el JCV está presente en tasas muy elevadas en la población general, la PML sigue siendo un trastorno poco frecuente, aunque importante debido a la escasa supervivencia y a las graves secuelas neurológicas, así como a la asociación recientemente demostrada con una variedad de terapias útiles, p. ej., el natalizumab en la esclerosis múltiple (MS). Se han descrito varios factores de riesgo de PML, pero es mejor considerarlos necesarios, pero no suficientes. Aunque estos factores de riesgo son muy relevantes, no predicen por sí solos quién desarrollará PML, ya que la gran mayoría de los individuos con estos factores de riesgo no desarrollarán el trastorno. Hay que tener en cuenta otros factores y cada vez hay más pruebas del papel de los factores genéticos del huésped en la susceptibilidad a la PML.

La capacidad de predecir con mayor precisión quién está en riesgo de desarrollar PML será de enorme beneficio en el contexto del tratamiento farmacológico con compuestos que son altamente eficaces en el contexto de su enfermedad (natalizumab en la MS, por ejemplo) pero conllevan el riesgo de un trastorno devastador. Es necesario desarrollar una prueba diagnóstica complementaria para excluir eficazmente a las personas con riesgo de PML y, de paso, tranquilizar a las personas con pruebas negativas sobre la drástica reducción del riesgo de desarrollar PML.

El documento WO 2015/131078 enseña marcadores de expresión génica asociados a PML bajo terapia inmunosupresora.

Breve descripción de los dibujos

Las características novedosas de la divulgación se exponen con particularidad en las reivindicaciones adjuntas. Una mejor comprensión de las características y ventajas de la presente divulgación se obtendrá por referencia a la siguiente descripción detallada que expone realizaciones ilustrativas, en las que se utilizan los principios de la divulgación, y los dibujos adjuntos.

La Figura 1 representa un ejemplo de un gen (PRKCB) afectado por CNV de línea germinal y adquiridas.
La Figura 2 representa un ejemplo de genes (TNFRSF13C y CENPM) afectados por CNV adquiridas.
La Figura 3 representa un ejemplo de un gen (PKHD1) afectado por CNV de línea germinal y adquiridas.
La Figura 4 representa un ejemplo de un gen (BMPR2) afectado por una CNV (pérdidas homocigotas y heterocigotas).
La Figura 5 representa un ejemplo de un gen (COMMD6) afectado por una CNV (p. ej., duplicación homocigota).
La Figura 6 representa un ejemplo de genes (KCTD7, RABGEF1) directa y potencialmente afectados por una CNV (p. ej., duplicación homocigota).
La Figura 7 representa un ejemplo de un gen (FPR2) afectado por una CNV (p. ej., duplicación homocigota).
La Figura 8 representa un ejemplo de un gen (PIK3CD) afectado por una CNV (p. ej., pérdida homocigota).
La Figura 9 representa un ejemplo de un gen (CD 180) potencialmente afectado por una ganancia de CNV intergénica (p. ej., duplicación homocigota).
La Figura 10 representa un ejemplo de un gen (VDAC1) potencialmente afectado por una CNV intergénica (pérdida homocigota).
La Figura 11 representa un ejemplo de genes (EGR1 y ETF1) potencialmente afectados por una CNV intergénica (pérdida homocigota).
La Figura 12 representa un ejemplo de un gen (ITSN2) potencialmente afectado por una CNV intergénica (pérdida homocigota).
La Figura 13 representa un ejemplo de interacciones proteicas conocidas y/o predichas utilizando la base de datos String para 21 de los 43 genes (lista no redundante) indicadas en la Tabla 7. Junto a cada gen se indica el número de casos de PML en los que se han encontrado variantes que afectan a un gen determinado.

Resumen de la invención

En un primer aspecto, la presente invención proporciona un medicamento inmunosupresor para su uso de acuerdo con la reivindicación 1. En un segundo aspecto, la presente invención proporciona un método de detección para un biomarcador de PML de acuerdo con la reivindicación 9. En un tercer aspecto, la presente invención proporciona un método para identificar a un sujeto con riesgo de desarrollar PML de acuerdo con la reivindicación 11. Otras características de la invención se definen en las reivindicaciones dependientes.

En algunas realizaciones, la afección es un cáncer, un trasplante de órganos o una enfermedad autoinmune.

En algunas realizaciones, la afección es una enfermedad autoinmune.

En algunas realizaciones, la enfermedad autoinmune se selecciona del grupo que consiste en enfermedad de Addison, encefalitis por receptor anti-NMDA, síndrome anti sintetasa, anemia aplásica, anemias autoinmunes, anemia hemolítica autoinmune, pancreatitis autoinmune, enfermedad de Behcet, trastornos ampollosos de la piel, enfermedad celíaca - esprúe (enteropatía sensible al gluten), síndrome de fatiga crónica, polineuropatía desmielinizante inflamatoria crónica, leucemia linfocítica crónica, enfermedad de Crohn, dermatomiositis, enfermedad de Devic, eritroblastopenia, síndrome de Evans, glomeruloesclerosis segmentaria focal, granulomatosis con poliangitis, enfermedad de Graves, oftalmopatía de Graves, síndrome de Guillain-Barré, tiroiditis de Hashimoto, púrpura trombocitopénica idiopática (ITP), nefropatía por IgA, enfermedades autoinmunes mediadas por IgA, enfermedad relacionada con IgG4, enfermedad inflamatoria intestinal, artritis idiopática juvenil, esclerosis múltiple, miastenia gravis, mieloma, linfoma no Hodgkin, síndrome opsoclono mioclono (OMS), penfigoide, penfigo, penfigo vulgar, anemia perniciosa, polimiositis, psoriasis, aplasia pura de eritrocitos, artritis reactiva, artritis reumatoide, sarcoidosis, esclerodermia, síndrome de Sjögren, lupus eritematoso sistémico, púrpura trombocitopénica, púrpura trombocitopénica trombótica, diabetes de tipo I, colitis ulcerosa, vasculitis (p. ej., vasculitis asociada a anticuerpos citoplasmáticos antineutrófilos), vitíligo y combinaciones de los mismos.

En algunas realizaciones, la enfermedad autoinmune es la esclerosis múltiple o la enfermedad de Crohn. En algunas realizaciones, la enfermedad autoinmune es la esclerosis múltiple. En algunas realizaciones, la esclerosis múltiple es una forma recidivante de esclerosis múltiple. En algunas realizaciones, la esclerosis múltiple es una esclerosis múltiple remitente-recurrente (RRMS). En algunas realizaciones, la esclerosis múltiple es una esclerosis múltiple primaria progresiva (PPMS). En algunas realizaciones, la esclerosis múltiple es una esclerosis múltiple secundaria progresiva (SPMS).

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden un glucocorticoide, citostático, anticuerpo, fármaco que actúa sobre inmunofilinas, interferón, opioide, proteína de unión a TNF, micofenolato, agente biológico pequeño, molécula pequeña, compuesto orgánico, o cualquier combinación de los mismos.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden abatacept, adalimumab, alefacept, alemtuzumab, anakinra, azatioprina, belimumab, bendamustina, bevacizumab, bortezomib (p. ej., VELCADE), eculizumab (p. ej., SOLIRIS), leflunomida, brentuximab vedotina, capecitabina, carboplatino, cetuximab, clorambucilo, cladribina, ciclofosfamida, ciclosporina, daclizumab, doxorrubicina, efalizumab, etanercept, etopósido, fludarabina, gemcitabina, ibritumomab tiuxetano, imatinib, infliximab, lenalidomida, metotrexato, micofenolato mofetilo, natalizumab, oxaliplatino, rituximab, tocilizumab, tofacitinib, ustekinumab, vedolizumab, vincristina, belatacept, quimioterapia citotóxica, corticosteroides, Ig antitimocítica, basiliximab, muromonab-CD3, ácido micofenólico, prednisona/prednisolona, sirolimus (rapamicina), tacrolimus, fumarato de dimetilo, fingolimod, ruxolitinib, interferón beta-1a, interferón beta-1b, acetato de glatiramero, peginterferón beta-1a, teriflunomida, mitoxantrona, ocrelizumab, asparaginasa, bleomicina, busulfano, carmustina, certolizumab, ibrutinib, idarrubicina, idelalisib, hidrocortisona, ifosfamida, levamisol, mercaptopurina, mizoribina, obinutuzumab, ofatumumab, tegafur/gimeracilo/oteracilo, tiotepa, vinblastina, o cualquier combinación de los mismos.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden interferón beta-1a, interferón beta-1b, acetato de glatiramero, peginterferón beta-1a, teriflunomida, fingolimod, fumarato de dimetilo, alemtuzumab, mitoxantrona, natalizumab, daclizumab, ocrelizumab, o cualquier combinación de los mismos.

En algunas realizaciones, el sujeto no ha tomado uno o más medicamentos inmunosupresores. En algunas realizaciones, el sujeto ha tomado uno o más medicamentos inmunosupresores. En algunas realizaciones, el sujeto está tomando uno o más medicamentos inmunosupresores.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden natalizumab (p. ej., TYSABRI). En algunas realizaciones, se administran al menos aproximadamente 10 mg de natalizumab, p. ej., al menos aproximadamente 10 mg, al menos aproximadamente 15 mg, al menos aproximadamente 20 mg, al menos aproximadamente 30 mg, al menos aproximadamente 40 mg, al menos aproximadamente 50 mg, al menos aproximadamente 60 mg, al menos aproximadamente 70 mg, al menos aproximadamente 80 mg, al menos aproximadamente 90 mg, al menos aproximadamente 100 mg, al menos aproximadamente 150 mg, al menos aproximadamente 200 mg, al menos aproximadamente 250 mg o al menos aproximadamente 300 mg de natalizumab. En algunas realizaciones, al menos aproximadamente 10 mg de natalizumab se administran mediante infusión intravenosa. En algunas realizaciones, al menos aproximadamente 10 mg de natalizumab se administran mediante infusión intravenosa en cuatro semanas.

En algunas realizaciones, se administran aproximadamente 100 mg a aproximadamente 500 mg de natalizumab, p. ej., aproximadamente 100 mg a aproximadamente 200 mg, aproximadamente 100 mg a aproximadamente 300 mg, aproximadamente 100 mg a aproximadamente 400 mg, aproximadamente 100 mg a aproximadamente 500 mg, aproximadamente 200 mg a aproximadamente 300 mg, aproximadamente 200 mg a aproximadamente 400 mg, aproximadamente 200 mg a aproximadamente 500 mg, aproximadamente

300 mg a aproximadamente 400 mg, aproximadamente 300 mg a aproximadamente 500 mg, o aproximadamente 400 mg a aproximadamente 500 mg de natalizumab. En algunas realizaciones, se administran aproximadamente 100 mg a aproximadamente 500 mg de natalizumab mediante infusión intravenosa. En algunas realizaciones, se administran aproximadamente 100 mg a aproximadamente 500 mg de natalizumab mediante infusión intravenosa en cuatro semanas. En algunas realizaciones, se administran aproximadamente 300 mg de natalizumab. En algunas realizaciones, se administran aproximadamente 300 mg de natalizumab mediante infusión intravenosa. En algunas realizaciones, se administran aproximadamente 300 mg de natalizumab mediante infusión intravenosa en cuatro semanas.

En algunas realizaciones, la prueba genética comprende la detección de una o más variaciones genéticas asociadas con un riesgo de desarrollar PML en una muestra de ácido polinucleico del sujeto. En algunas realizaciones, la prueba genética comprende la detección de una o más variaciones genéticas asociadas con un alto riesgo de desarrollar PML en una muestra de ácido polinucleico del sujeto.

En algunas realizaciones, la afección es esclerosis múltiple.

En algunas realizaciones, la afección es la enfermedad de Crohn.

En algunas realizaciones, se identifica que el sujeto no tiene riesgo de desarrollar leucoencefalopatía multifocal progresiva (PML) mediante una prueba genética. En algunas realizaciones, se identifica que el sujeto no tiene un alto riesgo de desarrollar leucoencefalopatía multifocal progresiva (PML) mediante una prueba genética.

En algunas realizaciones, la prueba comprende ensayar una muestra de ácido polinucleico del sujeto para una o más variaciones genéticas.

En algunas realizaciones, el SNV es un SNV heterocigoto.

En algunas realizaciones, el SNV es un SNV homocigoto.

En algunas realizaciones, la prueba genética o la prueba comprende análisis de microarreglos, PCR, secuenciación, hibridación de ácidos nucleicos o cualquier combinación de los mismos.

En algunas realizaciones, la prueba genética o la prueba comprende un análisis de microarreglos seleccionado del grupo que consiste en un análisis de arreglo de hibridación genómica comparativa (CGH) y un análisis de arreglo SNP.

En algunas realizaciones, la prueba genética o la prueba comprende secuenciación, en la que la secuenciación se selecciona del grupo que consiste en secuenciación de firma masivamente paralela (MPSS), secuenciación de Polony, pirosecuenciación 454, secuenciación de Illumina, secuenciación de Illumina (Solexa) utilizando preparación de biblioteca Genomics 10X, secuenciación SOLiD, secuenciación por semiconductores iónicos, secuenciación por nanobolas de ADN, secuenciación heliscópica de molécula única, secuenciación de molécula única en tiempo real (SMRT), secuenciación de RNAP, secuenciación de ADN Nanopore, secuenciación por hibridación y secuenciación microfluídica de Sanger.

En algunas realizaciones, la prueba genética o las pruebas comprenden el análisis de todo el genoma del sujeto.

En algunas realizaciones, la prueba genética o las pruebas comprenden el análisis de un exoma completo del sujeto.

En algunas realizaciones, la prueba genética o la prueba comprende el análisis de información de ácido nucleico que ya se ha obtenido para un genoma o un exoma completos del sujeto.

En algunas realizaciones, la información sobre el ácido nucleico se obtiene a partir de un análisis *in silico*.

En algunas realizaciones, el sujeto es un sujeto humano.

En algunas realizaciones, la muestra de ácido polinucleico comprende un ácido polinucleico de sangre, saliva, orina, suero, lágrimas, piel, tejido o pelo del sujeto.

En algunas realizaciones, el método comprende además tratar al sujeto con un agente que reduce una carga viral en el sujeto.

En algunas realizaciones, el agente inmunosupresor se administra después de reducir la carga viral.

En algunas realizaciones, la carga viral es una carga viral de JCV.

En algunas realizaciones, el agente que reduce la carga viral es un agente dirigido contra el JCV.

En algunas realizaciones, el medicamento inmunosupresor o el método comprende además analizar la presencia de JCV en una muestra biológica del sujeto. En algunas realizaciones, el medicamento o método inmunosupresor comprende una prueba de anticuerpos contra el JCV. En algunas realizaciones, la prueba de anticuerpos contra el JCV tiene un resultado negativo. En algunas realizaciones, la prueba de anticuerpos contra el JCV no detecta la presencia del JCV en la muestra biológica del sujeto. En algunas realizaciones, la prueba de anticuerpos contra el JCV detecta la presencia del JCV en la muestra biológica del sujeto.

En algunas realizaciones, el análisis para detectar la presencia del JCV comprende poner en contacto un reactivo de detección del JCV con la muestra biológica.

En algunas realizaciones, el reactivo de detección del JCV se selecciona del grupo que consiste en un anticuerpo anti-JCV, un cebador específico del JCV y combinaciones de los mismos.

En algunas realizaciones, el medicamento inmunosupresor o el método comprende analizar la presencia de JCV en una muestra biológica del sujeto. En algunas realizaciones, el análisis de la presencia del JCV comprende una prueba de anticuerpos contra el JCV, una prueba de CD62L o una prueba de bandas oligoclonales IgM del CSF. En algunas realizaciones, el análisis para detectar la presencia del JCV se lleva a cabo antes de la prueba genética. En algunas realizaciones, el análisis para detectar la presencia del JCV se lleva a cabo después de la prueba genética. En algunas realizaciones, el análisis para detectar la presencia del JCV se lleva a cabo simultáneamente con la prueba genética. En algunas realizaciones, el análisis de la presencia del JCV identifica al sujeto como vehículo del JCV. En algunas realizaciones, el análisis de la presencia del JCV identifica al sujeto como no portador del JCV. En algunas realizaciones, el resultado de la prueba genética identifica que el sujeto tiene un riesgo o un riesgo aumentado de desarrollar PML. En algunas realizaciones, el resultado de la prueba genética identifica que el sujeto no tiene riesgo o no tiene un riesgo aumentado de desarrollar PML.

En algunas realizaciones, el sujeto está inmunodeprimido. En algunas realizaciones, el sujeto tiene VIH. En algunas realizaciones, el sujeto tiene infección por VIH. En algunas realizaciones, el sujeto está en riesgo de infección por VIH. En algunas realizaciones, el método comprende la administración de una cantidad terapéuticamente eficaz de uno o más medicamentos antivirales, tal como inhibidores de la proteasa (lopinavir/ritonavir {p. ej., KALETRA}, indinavir {p. ej., CRIXIVAN}, ritonavir {p. ej., NORVIR}, nelfinavir {p. ej., VIRACEPT}, saquinavir en cápsulas de gelatina dura {p. ej., INVIRASE}, atazanavir {p. ej., REYATAZ}, amprenavir {p. ej., AGENERASE}, fosamprenavir {p. ej., TELZIR}, tipranavir {p. ej., APTIVUS}), inhibidores de la transcriptasa inversa, incluidos los inhibidores no nucleósidos y nucleósidos/nucleótidos (AZT {zidovudina, p. ej., Retrovir}, ddI {didanosina, p. ej., VIDEX}, 3TC {lamivudina, p. ej., EPIVIR}, d4T {estavudina, p. ej., ZERIT}, abacavir {p. ej., ZIAGEN}, FTC {emtricitabina, p. ej., EMTRIVA}, tenofovir {p. ej., VIREAD}, efavirenz {p. ej., SUSTIVA} y nevirapina {p. ej., VIRAMUNE}), inhibidores de la fusión T20 {enfuvirtida, p. ej., FUZEON}, inhibidores de la integrasa (Raltegravir, p. ej., ISENTRESS, MK-0518; y elvitegravir, p. ej., VITEKTA, GS-9137) e inhibidores de la maduración (bevirimat {PA-457}).

En algunas realizaciones, la afección es un cáncer, una neoplasia hematológica, un trasplante de órganos o una enfermedad autoinmune. En algunas realizaciones, la afección es linfocitopenia CD4+ idiopática (ICL).

En algunas realizaciones, la afección es una enfermedad autoinmune.

En algunas realizaciones, la enfermedad autoinmune se selecciona del grupo que consiste en enfermedad de Addison, enfermedad de Behcet, enfermedad inflamatoria intestinal, enfermedad celíaca - esprúe (enteropatía sensible al gluten), enfermedad de Crohn, dermatomiositis, glomeruloesclerosis segmentaria focal, enfermedad de Graves, tiroiditis de Hashimoto, esclerosis múltiple, miastenia gravis, pénfigo, penfigoide, anemia aplásica, anemia perniciosa, anemia hemolítica autoinmune, eritroblastopenia, púrpura trombocitopénica, síndrome de Evans, vasculitis, granulomatosis con poliangeítis, polineuropatía desmielinizante inflamatoria crónica, síndrome de Guillain-Barré, encefalitis por receptores anti-NMDA, enfermedad de Devic, pancreatitis autoinmune, síndrome opsoclono-mioclono, enfermedad relacionada con IgG4, psoriasis, artritis reactiva, artritis reumatoide, artritis idiopática juvenil, sarcoidosis, síndrome de Sjögren, lupus eritematoso sistémico, diabetes tipo I, vitíligo o colitis ulcerosa.

En algunas realizaciones, la enfermedad autoinmune es la esclerosis múltiple o la enfermedad de Crohn.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden un glucocorticoide, citostático, anticuerpo, fármaco que actúa sobre inmunofilinas, interferón, opioide, proteína de unión a TNF, micofenolato, agente biológico pequeño, molécula pequeña, compuesto orgánico, o cualquier combinación de los mismos.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden un interferón beta-1a, interferón beta-1b, acetato de glatiramero, peginterferón beta-1a, teriflunomida, fingolimod, fumarato de dimetilo, alemtuzumab, mitoxantrona, natalizumab, daclizumab, ocrelizumab, o cualquier combinación de los mismos.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden natalizumab (p. ej., TYSABRI).

En algunas realizaciones, el análisis comprende análisis de microarreglos, PCR, secuenciación, hibridación de ácidos nucleicos o cualquier combinación de los mismos.

En algunas realizaciones, el resultado de la prueba genética comprende un resultado de prueba genética de un análisis de microarreglos, PCR, secuenciación, hibridación de ácidos nucleicos o cualquier combinación de los mismos.

En algunas realizaciones, la detección comprende un análisis de microarreglos, PCR, secuenciación, hibridación de ácidos nucleicos o cualquier combinación de los mismos.

En algunas realizaciones, el análisis de microarreglos se selecciona del grupo que consiste en un análisis de arreglo de hibridación genómica comparativa (CGH) y un análisis de arreglo SNP.

En algunas realizaciones, la secuenciación se selecciona del grupo que consiste en secuenciación de firmas masivas en paralelo (MPSS), secuenciación de Polony, pirosecuenciación 454, secuenciación de Illumina, secuenciación de Illumina (Solexa) utilizando la preparación de bibliotecas Genomics 10X, secuenciación SOLiD, secuenciación por semiconductores iónicos, secuenciación por nanobolas de ADN, secuenciación heliscópica de moléculas individuales, secuenciación de moléculas individuales en tiempo real (SMRT), secuenciación RNAP, secuenciación de ADN Nanopore, secuenciación por hibridación y secuenciación microfluídica Sanger.

En algunas realizaciones, el análisis comprende analizar un genoma completo o un exoma completo del sujeto.

En algunas realizaciones, el análisis comprende analizar información de ácido nucleico que ya se ha obtenido para un genoma completo o un exoma completo del sujeto.

En algunas realizaciones, la información sobre el ácido nucleico se obtiene a partir de un análisis *in silico*.

En algunas realizaciones, el análisis comprende analizar un genoma completo o un exoma completo del sujeto.

En algunas realizaciones, el análisis comprende analizar información de ácido nucleico que ya se ha obtenido para un genoma completo o un exoma completo del sujeto.

En algunas realizaciones, la información sobre el ácido nucleico se obtiene a partir de un análisis *in silico*.

En algunas realizaciones, la detección comprende analizar un genoma completo o un exoma completo del sujeto.

En algunas realizaciones, la detección comprende el análisis de información de ácido nucleico que ya se ha obtenido para un genoma completo o un exoma completo del sujeto.

En algunas realizaciones, la información sobre el ácido nucleico se obtiene a partir de un análisis *in silico*.

En algunas realizaciones, el sujeto es un sujeto humano.

En algunas realizaciones, la muestra de ácido polinucleico comprende un ácido polinucleico de sangre, saliva, orina, suero, lágrimas, piel, tejido o pelo del sujeto.

En algunas realizaciones, el medicamento inmunosupresor o el método comprende además analizar la presencia de JCV en una muestra biológica del sujeto.

En algunas realizaciones, el análisis para detectar la presencia del JCV comprende poner en contacto un reactivo de detección del JCV con la muestra biológica.

En algunas realizaciones, el reactivo de detección del JCV se selecciona del grupo que consiste en un anticuerpo anti-JCV, un cebador específico del JCV y combinaciones de los mismos.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden un glucocorticoide, citostático, anticuerpo, fármaco que actúa sobre inmunofilinas, interferón, opioide, proteína de unión a TNF, micofenolato, agente biológico pequeño, molécula pequeña, compuesto orgánico, o cualquier combinación de los mismos. En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden abatacept, adalimumab, alefacept, alemtuzumab, anakinra, azatioprina, belimumab, bendamustina, bevacizumab, bortezomib, brentuximab vedotina, capecitabina, carboplatino, cetuximab, clorambucilo, cladribina, ciclofosfamida, ciclosporina, daclizumab, doxorrubicina, efalizumab, etanercept, etopósido, fludarabina, gemcitabina, ibritumomab tiuxetano, imatinib, infliximab, lenalidomida, metotrexato, micofenolato mofetilo, natalizumab, oxaliplatino, rituximab, tocilizumab, tofacitinib, ustekinumab, vedolizumab, vincristina, belatacept, quimioterapia citotóxica, corticosteroides, Ig antitimocítica, basiliximab, muromonab-CD3, ácido micofenólico,

prednisona/prednisolona, sirolimus, tacrolimus, fumarato de dimetilo, fingolimod, ruxolitinib, interferón beta-1a, interferón beta-1b, acetato de glatiramero, peginterferón beta-1a, teriflunomida, mitoxantrona, ocrelizumab, asparaginasa, bleomicina, busulfano, carmustina, certolizumab, ibrutinib, idarrubicina, idelalisib, hidrocortisona, ifosfamida, levamisol, mercaptopurina, mizoribina, obinutuzumab, ofatumumab, tegafur/gimeracilo/oteracilo, tiotepa, vinblastina, o cualquier combinación de los mismos. Por ejemplo, uno o más medicamentos inmunosupresores pueden comprender natalizumab.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden una molécula de anticuerpo o un fragmento de la misma. En algunas realizaciones, la molécula de anticuerpo o un fragmento de la misma es una molécula de anticuerpo recombinante o un fragmento de la misma. En algunas realizaciones, la molécula de anticuerpo o un fragmento de la misma es una molécula de anticuerpo humanizada o un fragmento de la misma. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma es una molécula de anticuerpo recombinante humanizada o fragmento de la misma. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma es una molécula de anticuerpo monoclonal recombinante IgG4κ humanizada o fragmento de la misma. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende una secuencia en número de registro CAS: 189261-10-7. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena pesada de anticuerpo. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende dos cadenas pesadas de anticuerpo. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena ligera de anticuerpo. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende dos cadenas ligeras de anticuerpo. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena pesada de anticuerpo y al menos una cadena ligera de anticuerpo.

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma se produce en células de mieloma. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma se produce en células de hibridoma de conejo.

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma se une a un receptor. En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma se une a una integrina. En algunas realizaciones, la integrina se expresa en la superficie de un leucocito. En algunas realizaciones, el leucocito es un neutrófilo. En algunas realizaciones, el leucocito no es un neutrófilo. En algunas realizaciones, la molécula de anticuerpo o un fragmento de la misma se une a la integrina α4β1, a la integrina α4β7 o a ambas. En algunas realizaciones, la molécula de anticuerpo o un fragmento de la misma se une a la subunidad α4 de la integrina α4β1, a la integrina α4β7, o a ambas. En algunas realizaciones, la molécula de anticuerpo o un fragmento de la misma inhibe la adhesión mediada por α4 de un leucocito a su receptor.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden un anticuerpo o un fragmento del mismo, que comprende una secuencia que tiene al menos un 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o 100 % de identidad de secuencia con la SEQ ID NO. 2178 (QVQLVQSGAE VKKPGASVKV SCKASGFNIK DTYIHWVRQA PGQRLEWMGR IDPANGYTKY DPKFQGRVTI TADTSASTAY MELSSLRSED TAVYYCAREG YYGNYGVYAM DYWGQGTLVT VSSASTKGPS VFPLAPCSRS TSESTAALGC LVKDYFPEPV TVSWNSGALT SGVHTFPAVL QSSGLYSLSS VVTVPSSSLG TKTYTCNVDH KPSNTKVDKR VESKYGPPCP SCPAPEFLGG PSVFLFPPKP KDTLMISRTP EVTCVVVDVS QEDPEVQFNW YVDGVEVHNA KTKPREEQFN STYRVVSVLT VLHQDWLNGK EYKCKVSNKG LPSSIEKTIS KAKGQPREPQ VYTLPPSQEE MTKNQVSLTC LVKGFYPSDI AVEWESNGQP ENNYKTTPPV LDSDGSFFLY SRLTVDKSRW QEGNVFSCSV MHEALHNHYT QKSLSLSLGK). En algunas realizaciones, el anticuerpo o fragmento del mismo comprende una secuencia que tiene aproximadamente un 50 %-100 % de identidad, p. ej., aproximadamente un 50 %-60 %, aproximadamente un 50 %-70 %, aproximadamente un 60 %-70 %, aproximadamente un 60 %-80 %, aproximadamente un 70 %-80 %, aproximadamente un 70 %-90 %, aproximadamente un 80 %-90 %, aproximadamente un 80 %-95 %, aproximadamente un 90 %-95 %, aproximadamente un 90 %-99 %, aproximadamente un 90 %-100 %, aproximadamente un 95 %-99 %, o aproximadamente un 99 %-100 % de identidad de secuencia con la SEQ ID NO. 2178.

En algunas realizaciones, uno o más medicamentos inmunosupresores comprenden un anticuerpo o un fragmento del mismo, que comprende una secuencia que tiene al menos aproximadamente 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, o 100 % de identidad de secuencia con la SEQ ID NO. 2179 (DIQMTQSPSS LSASVGDRVT ITCKTSQDIN KYMAWYQQTP GKAPRLLIHY TSALQPGIPS RFSGSGSGRD YTFTISSLQP EDIATYYCLQ YDNLWTFGQG TKVEIKRTVA APSVFIFPPS DEQLKSGTAS VVCLLNNFYP REAKVQWKVD NALQSGNSQE SVTEQDSKDS TYSLSSTLTL SKADYEKHKV YACEVTHQGL SSPVTKSFNRGEC). En algunas realizaciones, el anticuerpo o fragmento del mismo comprende una secuencia que tiene aproximadamente un 50 %-100 % de identidad, p. ej., aproximadamente un 50 %-60 %, aproximadamente un 50 %-70 %, aproximadamente un 60 %-70 %, aproximadamente un 60 %-80 %, aproximadamente un 70 %-80 %, aproximadamente un 70 %-90 %, aproximadamente un 80 %-90 %, aproximadamente un 80 %-95 %, aproximadamente un 90 %-95 %, aproximadamente un 90 %-99 %, aproximadamente un 90 %-100 %, aproximadamente un 95 %-99 %, o aproximadamente un 99 %-100 % de identidad de secuencia con la SEQ ID NO. 2179.

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena pesada de anticuerpo, o un fragmento de unión a α4 de la misma, que comprende CDR no humanas en las posiciones 31-35 (CDR1), 50-65 (CDR2) y 95-102 (CDR3) (numeración de Kabat) de un anticuerpo anti-α4 de ratón y que tiene residuos no humanos en las posiciones marco 27-30 (numeración de Kabat), en las que las posiciones 27-30 tienen la secuencia de aminoácidos Phe 27, Asn 28, Ile 29 y Lys 30.

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena ligera de anticuerpo, o un fragmento de unión a α4 de la misma, que comprende: una CDR1 de cadena ligera (LC) con una secuencia de aminoácidos de la SEQ ID NO: 2180 (KTSQDINKYMA), una CDR2 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2181 (YTSALQP), y una CDR3 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2182 (LQYDNLWT).

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena ligera de anticuerpo, o un fragmento de unión a α4 de la misma, que comprende: una CDR1 de cadena ligera (LC) con una secuencia de aminoácidos de la SEQ ID NO: 2183 (QASQDIIKYLN), una CDR2 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2184 (EASNLQA), y una CDR3 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2185 (QQYQSLPYT).

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena ligera de anticuerpo, o un fragmento de unión a α4 de la misma, que comprende: una CDR1 de cadena ligera (LC) con una secuencia de aminoácidos de la SEQ ID NO: 2186 (KASQSVTNDVA), una CDR2 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2187 (YASNRYT), y una CDR3 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2188 (QQDYSSPYT).

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena pesada de anticuerpo, o un fragmento de unión a α4 de la misma, que comprende: una CDR1 de cadena pesada (HC) con una secuencia de aminoácidos de la SEQ ID NO: 2189 (DTYIH), una CDR2 de HC con una secuencia de aminoácidos de la SEQ ID NO: 2190 (RIDPANGYTKYDPKFQG), y una CDR3 de HC con una secuencia de aminoácidos de la SEQ ID NO: 2191 (EGYYGNYGVYAMDY).

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena pesada de anticuerpo, o un fragmento de unión a α4 de la misma, que comprende: una CDR1 de cadena pesada (HC) con una secuencia de aminoácidos de la SEQ ID NO: 2192 (DTYMH), una CDR2 de HC con una secuencia de aminoácidos de la SEQ ID NO: 2193 (RIDPASGDTKYDPKFQV), y una CDR3 de HC con una secuencia de aminoácidos de la SEQ ID NO: 2194 (DGMWVSTGYALDF).

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende una cadena pesada humanizada, o un fragmento de unión a a4 de la misma, que comprende: una región de cadena pesada variable seleccionada del grupo que consiste en: la SEQ ID NO.: 2195 (MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVRQPPGRGLEWIGRID PASGDTKYDPKFQVKATITADTSSNQFSLRLSSVTAADTAVYYCADGMWVSTGYALDFWGQGTTVTVSSGES, la SEQ ID NO: 2196 (QVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVRQPPGRGLEWIGRIDPASGD TKYDPKF QVRVTMLVDTSSNQFSLRLSSVTSEDTAVYYCADGMWVSTGYALDFWGQGTTVTVSSGES), la SEQ ID NO: 2197 (MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVKQR PGRGLEWIGRIDPASGDTKYDPKFQVRVTMLVDTSSNQFSLRLSSVTAADTAVYYCADGMWVSTGYALDFWG QGTTVTVSSGES), la SEQ ID NO: 2198 (MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTAS GFNIKDTYMHWVRQPGR GLEWIGRIDPASGDTKYDPKFQVRVTMLVDTSSNQFSLRLSSVTAADTAVYCADG MWVSTGYALDFWGQGTTVSSGES), y la SEQ ID NO: 2199 (QYQLYQSGAEYKKPGASYKYSCKASGFNIKDT YIHWVRQAPGQRLEWMGRIDPANGYTKYDPKFQGRVTITADTSASTAYMELSSLRSEDTAVYYCAREGYYGN YGVYAMDYWGQGTLVTVVSS).

En algunas realizaciones, la molécula de anticuerpo o fragmento de la misma comprende una cadena ligera humanizada, o un fragmento de unión a4 de la misma, que comprende una región de cadena ligera variable seleccionada del grupo que consiste en: la SEQ ID NO.: 2200 (MGWSCIILFLVATATGVHSDIQLTQSPSSLSAS VGDRVTITCKASQSVTNDVAWYQQKPGKAPKLLIYYASNRYTGVPSRFSGSGTDFTFTISSLQPEDIATYYCQD YSSPYTFGQGTKVEIKRK), la SEQ ID NO: 2201 (MGWSCIILFLVATATGVHSSIVMTQSPSSLSASVGDRVTIT CKASQSVTNDVAWYQQKPGKAPKLLIYYASNRYTGVPDRFSGSGYGTDFTFTISSLQPEDIATYYCQDYSSPY TFGQGTKVEIKRK), la SEQ ID NO: 2202 (MGWSCIILFLVATATGVHSDIQMTQSPSSLSASVGDRVTITCKAS QSVTNDVAWYQKPGKAPKLLIYYASNRYTGVPDRFSGSGYGTDFTFTISSLQPEDIATYYCQDYSSPYTFGQG TKVEIKRK), y la SEQ ID NO: 2203 (DIQMTQSPSSLSASVGDRVTITCKTSQDINKYMAWYQQTPGKAPRLL IHYTSALQPGIPSRFSGS GSGRDYTFTISSLQPEDIATYYCLQYDNLWTFGQGTKVEIKRTV).

En algunas realizaciones, un producto biológico puede ser un producto biológico aprobado por una agencia reguladora. Por ejemplo, el producto biológico puede estar aprobado por la Administración de Alimentos y Medicamentos de Estados Unidos (FDA) y/o la Agencia Europea de Medicamentos (EMA). En algunas realizaciones, el producto biológico puede ser un producto de referencia. En algunos casos, el producto

biológico puede ser un producto biosimilar. En algunas realizaciones, el producto biológico puede ser un producto intercambiable.

En algunas realizaciones, un producto biosimilar puede ser similar a un producto de referencia. En algunas realizaciones, un producto biosimilar puede no presentar diferencias clínicamente significativas en términos de seguridad y eficacia con respecto al producto de referencia. En algunas realizaciones, un producto biosimilar puede tener los mismos componentes clínicamente inactivos. En algunas realizaciones, un producto biosimilar puede tener diferentes componentes clínicamente inactivos. En algunas realizaciones, un producto biosimilar interactúa específicamente con un sustrato y el producto de referencia interactúa específicamente con el mismo sustrato. En algunas realizaciones, la tasa de respuesta de los sujetos humanos a los que se administra el producto biosimilar puede ser del 50 % al 150 % de la tasa de respuesta de los sujetos humanos a los que se administra el producto de referencia. Por ejemplo, la tasa de respuesta de los sujetos humanos a los que se administra el producto biosimilar puede ser del 50 %-100 %, 50 %-110 %, 50 %-120 %, 50 %-130 %, 50 %-140 %, 50 %-150 %, 60 %-100 %, 60 %-110 %, 60 %-120 %, 60 %-130 %, 60 %-140 %, 60 %-150 %, 70 %-100 %, 70 %-110 %, 70 %-120 %, 70 %-130 %, 70 %-140 %, 70 %-150 %, 80 %-100 %, 80 %-110 %, 80 %-120 %, 80 %-130 %, 80 %-140 %, 80 %-150 %, 90 %-100 %, 90 %-110 %, 90 %-120 %, 90 %-130 %, 90 %-140 %, 90 %-150 %, 100 %-110 %, 100 %-120 %, 100 %-130 %, 100 %-140 %, 100 %-150 %, 110 %-120 %, 110 %-130 %, 110 %-140 %, 110 %-150 %, 120 %-130 %, 120 %-140 %, 120 %-150 %, 130 %-140 %, 130 %-150 %, o 140 %-150 % de la tasa de respuesta de los sujetos humanos a los que se administró el producto de referencia. En algunas realizaciones, un producto biosimilar y un producto de referencia pueden utilizar el mismo mecanismo o mecanismos de acción para la condición o condiciones de uso prescritas, recomendadas o sugeridas en el marcaje propuesto, pero sólo en la medida en que el mecanismo o mecanismos sean conocidos para el producto de referencia.

En algunas realizaciones, un producto intercambiable puede ser un producto biosimilar que cumple normas adicionales de intercambiabilidad. En algunas realizaciones, un producto intercambiable puede producir el mismo resultado clínico que un producto de referencia en todas las condiciones de uso autorizadas del producto de referencia. En algunas realizaciones, un producto intercambiable puede ser sustituido por el producto de referencia por un farmacéutico sin la intervención del profesional sanitario que prescribió el producto de referencia. En algunas realizaciones, cuando se administra más de una vez a un individuo, el riesgo en términos de seguridad o disminución de la eficacia de alternar o cambiar entre el uso del producto biológico y el producto de referencia no es mayor que el riesgo de usar el producto de referencia sin dicha alternancia o cambio. En algunas realizaciones, un producto intercambiable puede ser un producto aprobado por una agencia reguladora. En algunas realizaciones, la tasa de respuesta de los sujetos humanos a los que se administra el producto intercambiable puede ser del 80 % al 120 % de la tasa de respuesta de los sujetos humanos a los que se administra el producto de referencia. Por ejemplo, la tasa de respuesta de los sujetos humanos a los que se administra el producto intercambiable puede ser del 80 %-100 %, 80 %-110 %, 80 %-120 %, 90 %-100 %, 90 %-110 %, 90 %-120 %, 100 %-110 %, 100 %-120 % o 110 %-120 de la tasa de respuesta de los sujetos humanos a los que se administra el producto de referencia.

En algunas realizaciones, la afección es la esclerosis múltiple o la enfermedad de Crohn. En algunos casos, la enfermedad es una forma recidivante de esclerosis múltiple. En algunas realizaciones, el natalizumab se administra mediante infusión intravenosa.

En algunas realizaciones, se administran aproximadamente 100 mg a aproximadamente 500 mg de natalizumab. En algunas realizaciones, se administran aproximadamente 100 mg a aproximadamente 500 mg de natalizumab, p. ej., aproximadamente 100 mg a aproximadamente 200 mg, aproximadamente 100 mg a aproximadamente 300 mg, aproximadamente 100 mg a aproximadamente 400 mg, aproximadamente 100 mg a aproximadamente 500 mg, aproximadamente 200 mg a aproximadamente 300 mg, aproximadamente 200 mg a aproximadamente 400 mg, aproximadamente 200 mg a aproximadamente 500 mg, aproximadamente 300 mg a aproximadamente 400 mg, aproximadamente 300 mg a aproximadamente 500 mg, o aproximadamente 400 mg a aproximadamente 500 mg de natalizumab. En algunas realizaciones, se administran aproximadamente 100 mg a aproximadamente 500 mg de natalizumab mediante infusión intravenosa. En algunas realizaciones, se administran aproximadamente 100 mg a aproximadamente 500 mg de natalizumab mediante infusión intravenosa en cuatro semanas. En algunas realizaciones, se administran aproximadamente 300 mg de natalizumab. En algunas realizaciones, se administran aproximadamente 300 mg de natalizumab mediante infusión intravenosa. En algunas realizaciones, se administran aproximadamente 300 mg de natalizumab mediante infusión intravenosa en cuatro semanas.

En algunas realizaciones, una o más variaciones genéticas están asociadas con un riesgo de desarrollar PML en una muestra de ácido polinucleico del sujeto.

En algunas realizaciones, el medicamento inmunosupresor o método comprende probar al sujeto para una predisposición genética para PML con un ensayo genético. En algunas realizaciones, el ensayo genético tiene un rendimiento diagnóstico de al menos el 5 %. En algunos casos, el ensayo genético tiene un rendimiento diagnóstico de al menos aproximadamente 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 99 %. En algunos casos, el ensayo

genético tiene un rendimiento diagnóstico de aproximadamente 1 %-5 %, 1 %-10 %, 1 %-20 %, 5 %-10 %, 5 %-20 %, 10 %-20 %, 10 %-30 %, 20 %-30 %, 20 %-40 %, 30 %-40 %, 30 %-50 %, 40 %-50 %, 40 %-60 %, 50 %-60 %, 50 %-70 %, 60 %-70 %, 60 %-80 %, 70 %-80 %, 70 %-90 %, 80 %-90 %, 80 %-95 %, 90 %-95 %, 90 %-99 %, 90 %-100 %, 95 %-99 %, o 99 %-100 %. En algunas realizaciones, el ensayo genético tiene un rendimiento diagnóstico de al menos el 20 %.

En algunas realizaciones, la disminución del riesgo del sujeto se debe además a la ausencia de una o más variaciones genéticas que alteran o modulan un gen correspondiente de acuerdo con las Tablas 25A, 25B y 26.

En algunas realizaciones, el ensayo genético comprende análisis de microarreglos, PCR, secuenciación, hibridación de ácidos nucleicos o cualquier combinación de los mismos.

En algunas realizaciones, el medicamento inmunosupresor o el método comprende analizar al sujeto con una prueba de anticuerpos contra el JCV, una prueba de CD62L o una prueba de bandas oligoclonales IgM del CSF. En algunas realizaciones, el medicamento inmunosupresor o el método comprenden analizar al sujeto con la prueba de anticuerpos contra el JCV, en la que la prueba de anticuerpos contra el JCV no detecta la presencia del JCV. En algunas realizaciones, el medicamento inmunosupresor o el método comprende analizar al sujeto con la prueba de anticuerpos contra el JCV, en la que la prueba de anticuerpos contra el JCV detecta la presencia del JCV. En algunas realizaciones, la prueba de anticuerpos contra el JCV comprende poner en contacto un reactivo de detección del JCV con una muestra biológica del sujeto. En algunas realizaciones, el reactivo de detección del JCV se selecciona del grupo que consiste en un anticuerpo anti-JCV, un cebador específico del JCV y combinaciones de los mismos.

Descripción detallada de la divulgación

Los detalles de una o más realizaciones de la invención se exponen en los dibujos adjuntos, en las reivindicaciones y en la presente descripción. Otras características, objetivos y ventajas de las realizaciones inventivas divulgadas y contempladas en el presente documento se desprenderán de la descripción y los dibujos, así como de las reivindicaciones.

Tal como se utiliza en el presente documento, a menos que se indique lo contrario, el artículo "un" significa uno o más, a menos que se disponga explícitamente lo contrario.

Tal como se utilizan en el presente documento, a menos que se indique lo contrario, términos como "contiene", "que contiene", "incluye", "que incluye" y similares significan "que comprende"

Tal como se utiliza en el presente documento, a menos que se indique lo contrario, el término "o" puede ser conjuntivo o disyuntivo. Tal como se utiliza en el presente documento, a menos que se indique lo contrario, cualquier forma de realización puede combinarse con cualquier otra forma de realización.

Tal como se utilizan en el presente documento, a menos que se indique lo contrario, algunas realizaciones inventivas del presente documento contemplan intervalos numéricos. Cuando hay intervalos presentes, los intervalos incluyen los puntos finales de los intervalos. Además, todos los subintervalos y valores dentro del intervalo están presentes como si se hubieran escrito explícitamente.

Tal como se utiliza en el presente documento, a menos que se indique lo contrario, el término "aproximadamente" en relación con un valor numérico de referencia y sus equivalentes gramaticales incluyen un intervalo de valores más o menos el 10 % de dicho valor, tal como un intervalo de valores más o menos el 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 % o 1 % de dicho valor. Por ejemplo, la cantidad "aproximadamente 10" incluye las cantidades de 9 a 11.

Tal como se utiliza en el presente documento, a menos que se indique lo contrario, el término "producto biológico" se refiere a un virus, suero terapéutico, toxina, antitoxina, vacuna, sangre, componente sanguíneo o derivado, producto alergénico, proteína (cualquier polímero de aminoácidos alfa con una secuencia específica definida que tenga un tamaño superior a 40 aminoácidos), o producto análogo, o arsfenamina o derivado de arsfenamina (o cualquier compuesto orgánico trivalente de arsénico), aplicable a la prevención, tratamiento o curación de una enfermedad o afección de los seres humanos.

Tal como se utiliza en el presente documento, a menos que se indique lo contrario, el término "producto biosimilar" se refiere a 1) un producto biológico que tiene una secuencia de aminoácidos que es idéntica a un producto de referencia; 2) un producto biológico que tiene una secuencia de aminoácidos diferente (p. ej., truncamientos del terminal N o C) de un producto de referencia; o 3) un producto biológico que tiene una modificación postraduccional diferente (p. ej., glicosilación o fosforilación) de un producto de referencia, en el que el producto biosimilar y el producto de referencia utilizan el mismo mecanismo o mecanismos de acción para la prevención, tratamiento o curación de una enfermedad o afección.

Como se usa en el presente documento, "mecanismo de acción" se refiere a una interacción o actividad a través de la cual un producto farmacológico (p. ej., un producto biológico) produce un efecto farmacológico.

Tal como se utiliza en el presente documento, a menos que se indique lo contrario, el término "producto intercambiable" se refiere a un producto biosimilar, en el que la tasa de respuesta de un sujeto humano al que se le administra el producto intercambiable es del 80 % al 120 % de la tasa de respuesta del sujeto humano al que se le administra el producto de referencia.

Como se usa en el presente documento, a menos que se indique lo contrario, el término "producto de referencia" se refiere a 1) un producto biológico que tiene una secuencia de aminoácidos que es idéntica a un producto biosimilar; 2) un producto biológico que tiene una secuencia de aminoácidos diferente (p. ej., truncamientos del terminal N o C) de un producto biosimilar; o 3) un producto biológico que tiene una modificación postraduccional diferente (p. ej., glicosilación o fosforilación) de un producto biosimilar, en el que el producto de referencia y el producto biosimilar utilizan el mismo mecanismo o mecanismos de acción para la prevención, tratamiento o curación de una enfermedad o afección.

Tal como se utiliza en el presente documento, a menos que se indique lo contrario, cualquier denominación común o genérica de un producto biológico incluye el producto biológico y cualquier producto biosimilar del mismo. Por ejemplo, la denominación común, filgrastim, se refiere al producto biológico vendido bajo el nombre comercial NEUPOGEN; también incluye el producto biosimilar, filgrastim-sndz, vendido bajo el nombre comercial ZARXIO. En otro ejemplo, la denominación común, natalizumab, se refiere al producto biológico vendido bajo el nombre comercial TYSABRI; también incluye cualquier producto biosimilar del producto biológico.

Leucoencefalopatía multifocal progresiva (PML)

La leucoencefalopatía multifocal progresiva (PML) es una enfermedad viral rara y generalmente mortal que se caracteriza por un daño o inflamación progresiva de la sustancia blanca del cerebro en múltiples localizaciones. La causa de la PML puede ser un tipo de poliomavirus denominado virus de John Cunningham (JC) (o JCV), que puede ser inofensivo salvo en casos de sistemas inmunitarios debilitados. Mientras que el JCV está presente en tasas muy elevadas en la población general, la PML sigue siendo un trastorno poco frecuente, aunque importante debido a las secuelas clínicas.

La PML puede ocurrir en pacientes con inmunodeficiencia grave, que permite la reactivación del virus de JC, tal como: 1) más comúnmente entre pacientes con síndrome de inmunodeficiencia adquirida (SIDA) que resulta de la infección con el virus de inmunodeficiencia humana (VIH), 2) pacientes con medicamentos inmunosupresores como corticosteroides, tacrolimus, sirolimus, y otros medicamentos o biológicos inmunosupresores, para trasplante de órganos (por ej., renal, hepático, pulmonar y cardíaco) y en personas con cáncer (p. ej., enfermedad de Hodgkin, leucemia o linfoma, y neoplasias mieloproliferativas como la mielofibrosis), y 3) individuos con enfermedades autoinmunes (p. ej., esclerosis múltiple, artritis reumatoide, psoriasis y lupus eritematoso sistémico) con terapias que deprimen la respuesta inmunitaria. Se han notificado varios fármacos inmunosupresores en el contexto de la PML inducida por fármacos o PML asociada a fármacos. Por ejemplo, véase: Melis et al., CNS Drugs. 2015; 29(10): 879-91); Maas et al. J Neurol. Octubre de 2016; 263(10): 2004-21; Colin et al., Fundam Clin Pharmacol. Octubre 13 de 2016. Los medicamentos inmunosupresores pueden incluir, entre otros, interferón beta-1a, interferón beta-1b, acetato de glatiramero, peginterferón beta-1a, teriflunomida, mitoxantrona, ocrelizumab, abatacept, adalimumab, alefacept, alemtuzumab, anakinra, bortezomib (p. ej., VELCADE), eculizumab (p. ej., SOLIRIS), leflunomida y varios otros medicamentos para trasplantes como Ig antitimocitos, asparaginasa, azatioprina, basiliximab, belatacept, belimumab, bendamustina, bevacizumab, bleomicina, brentuximab vedotina, busulfano, capecitabina, carboplatino, carmustina, certolizumab, cetuximab, clorambucilo, cladribina, corticosteroides, ciclofosfamida, ciclosporina, quimioterapia citotóxica, daclizumab, fumarato de dimetilo, doxorrubicina, efalizumab, etanercept, etopósido, fingolimod, fludarabina, gemcitabina, hidrocortisona, ibritumomab tiuxetano, ibrutinib, idarrubicina, idelalisib, ifosfamida, imatinib, infliximab, lenalidomida, levamisol, mercaptopurina, metotrexato, mizoribina, muromonab-CD3, micofenolato mofetilo, ácido micofenólico, natalizumab, obinutuzumab, ofatumumab, oxaliplatino, prednisona/prednisolona, rituximab, ruxolitinib, sirolimus (rapamicina), tacrolimus, tegafur/gimeracilo/oteracilo, tiotepa, tocilizumab, tofacitinib, ustekinumab, vedolizumab, vinblastina y vincristina. Algunos ejemplos de medicamentos inmunosupresores de molécula pequeña incluyen el fumarato de dimetilo, el fingolimod y el ruxolitinib. Una terapia inmunosupresora puede clasificarse como un agente terapéutico de Clase 1 (alto riesgo), tal como efalizumab y natalizumab, según se informa en Calabrese L. H. et al., Nat Rev Rheumatol. (2015).

En algunos casos, los medicamentos inmunosupresores pueden ser agentes de entrecruzamiento de ADN y/o ARN, incluyendo, agentes alquilantes, agentes alquilantes de mostaza nitrogenada, inhibidores de topoisomerasa, antraciclinas y fármacos anticancerígenos basados en platino. En algunos casos, los medicamentos inmunosupresores pueden ser inhibidores de la quinasa, incluida la fosfoinositida-3-quinasa. En algunos casos, los medicamentos inmunosupresores pueden ser inhibidores de la tirosina quinasa, incluidos los inhibidores de la región del grupo del punto de ruptura de la proteína de fusión-homólogo 1 del oncogén viral de leucemia murina de Abelson (BCR-ABL), tirosina quinasa de Bruton (BTK), receptor del factor de crecimiento epidérmico (EGFR), quinasa Janus (JAK) y factor de crecimiento endotelial vascular (VEGF). En algunos casos, los medicamentos inmunosupresores pueden ser anticuerpos monoclonales y/o conjugados

anticuerpo-fármaco dirigidos contra proteínas que incluyen grupos de proteínas de diferenciación (CD), tales como CD2, CD3, CD11a, CD20, CD30, CD52; interleucinas (IL), tales como IL-1, IL-2, IL-6, IL-12, IL-23; proteínas de la familia del factor de necrosis tumoral (TNF), tales como TNFα; e integrinas, tales como integrina α4. En algunos casos, los medicamentos inmunosupresores pueden ser inhibidores de mTOR y/o calcineurina. En algunos casos, los medicamentos inmunosupresores pueden ser inhibidores del complemento. En algunos casos, los medicamentos inmunosupresores pueden ser antimetabolitos inmunosupresores, inhibidores metabólicos de nucleósidos, nucleósidos de imidazol, análogos de nucleótidos, inhibidores de la síntesis de nucleósidos, inhibidores de la síntesis de purinas, inhibidores de la síntesis de pirimidinas o inhibidores de la pirimidina sintasa. En algunos casos, los medicamentos inmunosupresores pueden ser proteínas recombinantes, tal como el interferón beta recombinante. En algunos casos, los medicamentos inmunosupresores pueden ser moduladores del receptor de esfingosina-1-fosfato y/o del receptor nicotínico de acetilcolina. En algunos casos, los medicamentos inmunosupresores pueden ser anticuerpos terapéuticos, incluida la inmunoglobulina G. En algunos casos, los medicamentos inmunosupresores pueden ser inhibidores de la asparaginasa. En algunos casos, los medicamentos inmunosupresores pueden ser inhibidores específicos del estimulador de linfocitos B (BLyS). En algunos casos, los medicamentos inmunosupresores pueden ser moduladores de la coestimulación de las células T. En algunos casos, los medicamentos inmunosupresores pueden ser inmunosupresores polipeptídicos cíclicos y/o polipéptidos sintéticos que modifican los procesos inmunitarios. En algunos casos, los medicamentos inmunosupresores pueden ser corticosteroides. En algunos casos, los medicamentos inmunosupresores pueden ser fármacos quimioterapéuticos citotóxicos. En algunos casos, los medicamentos inmunosupresores pueden ser antibióticos glucopéptidos citotóxicos y/o mezclas de los mismos. En algunos casos, los medicamentos inmunosupresores pueden ser moléculas que inhiben la producción de citoquinas proinflamatorias. En algunos casos, los medicamentos inmunosupresores pueden ser análogos de la talidomida. En algunos casos, los medicamentos inmunosupresores pueden ser inhibidores de las vías de degradación celular, tal como los inhibidores del proteasoma.

La PML puede diagnosticarse en un paciente con un curso progresivo de la enfermedad, encontrando ADN del virus de JC en el líquido cefalorraquídeo en el presente documento junto con lesiones consistentes de la sustancia blanca en las imágenes de resonancia magnética (MRI) cerebral; alternativamente, una biopsia cerebral puede ser diagnóstica cuando la histopatología típica de desmielinización, astrocitos extraños y núcleos oligodendrogliales agrandados están presentes, junto con técnicas que muestran la presencia del virus de JC. La evidencia característica de la PML en las imágenes de escaneo de CT cerebral pueden ser lesiones hipodensas multifocales, sin realce de contraste y sin efecto de masa, pero la MRI puede ser más sensible que la CT. La zona más común de afectación puede ser la sustancia blanca cortical de los lóbulos frontal y parieto-occipital, pero las lesiones pueden producirse en cualquier parte del cerebro, tal como los ganglios basales, la cápsula externa y estructuras de la fosa craneal posterior como el tronco encefálico y el cerebelo.

En general, el tratamiento de la PML tiene como objetivo revertir la inmunodeficiencia para ralentizar o detener el progreso de la enfermedad. Los pacientes en régimen de inmunosupresión pueden dejar de tomar el medicamento inmunosupresor o se puede recurrir al recambio plasmático (PLEX) para acelerar la eliminación del medicamento inmunosupresor que puso a la persona en riesgo de PML. Los pacientes infectados por el VIH pueden iniciar un tratamiento antirretroviral de gran actividad (HAART). La aparición de la PML también puede ocurrir en el contexto del síndrome inflamatorio de reconstitución inmune (IRIS), en el que la aparición de la PML puede ocurrir o los síntomas de la PML pueden empeorar después del cese de la inmunosupresión (p. ej., de acuerdo con lo revisado por Pavlovic et al., Ther Adv Neurol Disord. noviembre de 2015; 8(6): 255-73 y Bowen et al., Nat Rev Neurol. Octubre 27 de 2016;12(11): 662-674). Por ejemplo, en pacientes con MS que desarrollan PML durante el tratamiento con natalizumab, a menudo se produce IRIS cuando se interrumpe el tratamiento y se utiliza PLEX para eliminar el natalizumab de la circulación del paciente. El tratamiento del IRIS en pacientes con PML puede incluir la administración de corticosteroides. Otros posibles tratamientos de la PML pueden incluir cidofovir, citarabina, fármaco antipalúdico mefloquina, interleucina-2 y 1-O-hexadeciloxipropil-cidofovir (CMX001, también conocido como brincidofovir). Según la revisión de Pavlovic (Ther Adv Neurol Disord. 2015 noviembre de; 8(6): 255-73), los tratamientos potenciales para la PML incluyen agentes antivirales (p. ej., clorpromazina, citalopram, mirtazapina, risperidona, ziprasidona, retro-2cycl, brefeldina A, cidofovir, brincidofovir, citarabina, ganciclovir, leflunomida, topotecano, mefloquina, 3-aminobenzamida, imatinib y Ag122), moduladores de la respuesta inmune (p. ej., IFN-alfa, IL-2, IL-7, maraviroc y glucocorticoides) e inmunización (p. ej., anticuerpos monoclonales recombinantes humanos VP-1 anti-JCV, terapia con linfocitos T citotóxicos específicos del JCV, IL-7 más vacuna VP1 contra el JCV y vacuna oral contra el JCV).

El término "rendimiento diagnóstico" como se usa en el presente documento se refiere al porcentaje de casos que identificarían la presencia de una o más variaciones genéticas (p. ej., CNV, SNV) en una cohorte de PML usando un ensayo. Por ejemplo, si 40 casos identificaran la presencia de una o más variaciones genéticas (p. ej., CNV, SNV) en una cohorte de 100 pacientes con PML, el rendimiento diagnóstico del ensayo es del 40 %. En algunos casos, los pacientes de la cohorte de PML son diagnosticados clínicamente con PML. En algunos casos, un paciente es diagnosticado clínicamente con PML cuando el ADN del virus de JC está presente en el líquido cefalorraquídeo y se observan lesiones consistentes de la sustancia blanca en las imágenes de

resonancia magnética (MRI) cerebral. En algunos casos, un paciente es diagnosticado clínicamente con PML cuando en una biopsia cerebral se observa la histopatología típica de desmielinización, astrocitos extraños y núcleos oligodendrogliales agrandados, junto con la presencia del virus de JC. En algunos casos, la cohorte de PML tiene al menos 5 casos de PML, p. ej., al menos 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900 o 1000 casos de PML. En algunos casos, la cohorte de PML es una cohorte de las enumeradas en el presente documento. Por ejemplo, la cohorte de PML es la cohorte de pacientes con PML que aparece en la Tabla 7. En algunos casos, el ensayo es un ensayo de anticuerpos contra el JCV. En algunos casos, el ensayo no es un ensayo de anticuerpos contra el JCV. En algunos casos, el ensayo es un ensayo genético. En algunos casos, el ensayo genético comprueba la predisposición genética a la PML.

El ensayo genético puede comprender cualquier método divulgado en el presente documento. En algunos casos, el ensayo genético tiene un rendimiento diagnóstico de al menos aproximadamente 5 %, 6 %, 7 %, 8 %, 9 %, 10 %, 15 %, 20 %, 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 % o 99 %. En algunos casos, el ensayo genético tiene un rendimiento diagnóstico de aproximadamente 1 %-5 %, 1 %-10 %, 1 %-20 %, 5 %-10 %, 5 %-20 %, 10 %-20 %, 10 %-30 %, 20 %-30 %, 20 %-40 %, 30 %-40 %, 30 %-50 %, 40 %-50 %, 40 %-60 %, 50 %-60 %, 50 %-70 %, 60 %-70 %, 60 %-80 %, 70 %-80 %, 70 %-90 %, 80 %-90 %, 80 %-95 %, 90 %-95 %, 90 %-99 %, 90 %-100 %, 95 %-99 %, o 99 %-100 %.

Variaciones genéticas asociadas a la PML

La presente invención implica una variación genómica seleccionada de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. A continuación se ofrece una explicación general de las variaciones genéticas y se describen ejemplos de variaciones que no forman parte de la materia reivindicada en el presente documento, en la medida en que se refieren a variaciones distintas de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. En el presente documento se describen métodos que pueden utilizarse para detectar variaciones genéticas. La detección de variaciones genéticas específicas, p. ej., marcadores polimórficos y/o haplotipos, número de copias, ausencia o presencia de un alelo, o genotipo asociado con una afección (p. ej., enfermedad o trastorno) como se describe en el presente documento, puede llevarse a cabo mediante métodos conocidos en la técnica para analizar ácidos nucleicos y/o detectar secuencias en sitios polimórficos o genéticamente variables, p. ej., técnicas de amplificación, técnicas de hibridación, secuenciación, microarreglos/arreglos, o cualquier combinación de las mismas. De este modo, mediante el uso de estos métodos divulgados en el presente documento u otros métodos disponibles para el experto en la materia, pueden identificarse uno o más alelos en marcadores polimórficos, incluidos microsatélites, polimorfismos de un solo nucleótido (SNP), variaciones de un solo nucleótido (SNV), inserciones/eliminaciones (indels), variaciones del número de copias (CNV) u otros tipos de variaciones genéticas, en una muestra obtenida de un sujeto.

Las secuencias genómicas dentro de las poblaciones muestran variabilidad entre individuos en muchas localizaciones del genoma. Por ejemplo, el genoma humano presenta variaciones de secuencia que se producen en promedio cada 500 pares de bases. Estas variaciones genéticas en las secuencias de ácido polinucleico se denominan comúnmente polimorfismos o sitios polimórficos. Como se utiliza en el presente documento, un polimorfismo, p. ej., una variación genética, incluye una variación en la secuencia del genoma entre una población, tal como variaciones alélicas y otras variaciones que surgen o se observan. De este modo, un polimorfismo se refiere a la aparición de dos o más secuencias o alelos alternativos determinados genéticamente en una población. Estas diferencias pueden darse en porciones codificantes (p. ej., exónicas) y no codificantes (p. ej., intrónicas o intergénicas) del genoma, y pueden manifestarse o detectarse como diferencias en las secuencias de ácidos polinucleicos, la expresión génica, incluyendo, p. ej., la transcripción, el procesamiento, la traducción, el transporte, el procesamiento de proteínas, el tráfico, la síntesis de ADN; las proteínas expresadas, otros productos génicos o productos de vías bioquímicas o en modificaciones postraduccionales y cualquier otra diferencia manifestada entre los miembros de una población. Los polimorfismos que surgen como resultado de un cambio de una sola base, tal como los polimorfismos de un solo nucleótido (SNP) o variaciones de un solo nucleótido (SNV), pueden incluir una inserción, eliminación o cambio en un nucleótido. Un marcador o sitio polimórfico es el locus en el que se produce la divergencia. Estos sitios pueden ser tan pequeños como un par de bases (un SNP o SNV). Los marcadores polimórficos incluyen, entre otros, los polimorfismos de longitud de fragmentos de restricción (RFLP), el número variable de repeticiones en tándem (VNTR), las regiones hipervariables, los minisatélites, las repeticiones dinucleotídicas, las repeticiones trinucleotídicas, las repeticiones tetranucleotídicas y otros patrones de repetición, las repeticiones de secuencia simple y los elementos de inserción, tales como Alu. Las formas polimórficas también se manifiestan como diferentes alelos mendelianos para un gen. Los polimorfismos pueden observarse por diferencias en las proteínas, las modificaciones de las proteínas, la modificación de la expresión del ARN, la

metilación del ADN y el ARN, los factores reguladores que alteran la expresión de los genes y la replicación del ADN, y cualquier otra manifestación de alteraciones en el ácido polinucleico genómico o en los ácidos polinucleicos de los orgánulos. Los expertos en la materia pueden apreciar que los polimorfismos se consideran a veces como una subclase de variaciones, definidas sobre la base de un corte de frecuencia particular en una población. Por ejemplo, se pueden considerar polimorfismos a las variantes/variaciones genéticas que se producen con una frecuencia >1 %, o >5 %, en la población.

Estas variaciones genéticas pueden encontrarse asociadas a uno o más trastornos y/o enfermedades utilizando los métodos divulgados en el presente documento. Estas variaciones genéticas pueden asociarse con la ausencia de uno o más trastornos y/o enfermedades (es decir, una o más variantes son protectoras contra el desarrollo del trastorno y/o enfermedades) utilizando los métodos divulgados en el presente documento.

Estas variaciones genéticas pueden comprender mutaciones puntuales, polimorfismos, polimorfismos de un solo nucleótido (SNP), variaciones de un solo nucleótido (SNV), translocaciones, inserciones, eliminaciones, amplificaciones, inversiones, eliminaciones intersticiales, variaciones del número de copias (CNV), variación estructural (SV), pérdida de heterocigosidad, o cualquier combinación de las mismas. Como variación genética se incluye cualquier eliminación, inserción o sustitución de bases del ADN genómico de uno o más individuos de una primera porción de una población total que, de este modo, da lugar a una diferencia en el lugar de la eliminación, inserción o sustitución de bases con respecto a uno o más individuos de una segunda porción de la población total. Por lo tanto, el término "variación genética" abarca el "tipo silvestre" o la variación que se produce con más frecuencia, y también incluye la "mutante", o la variación que se produce con menos frecuencia. Un alelo de tipo silvestre puede denominarse alelo ancestral.

Tal como se utiliza en el presente documento, una molécula diana que está "asociada con" o "correlacionada con" una variación genética particular es una molécula que puede distinguirse funcionalmente en su estructura, actividad, concentración, compartimentación, degradación, secreción y similares, tal como resultado de dicha variación genética. Los polimorfismos (p. ej., marcadores polimórficos, variaciones genéticas o variantes genéticas) pueden comprender cualquier posición nucleotídica en la que dos o más secuencias son posibles en una población de sujetos. Cada versión de una secuencia de nucleótidos, con respecto al polimorfismo/variación, puede representar un alelo específico del polimorfismo/variación. El ADN genómico de un sujeto puede contener dos alelos para cualquier marcador polimórfico dado, representativos de cada copia del marcador en cada cromosoma. Un alelo puede ser una secuencia de nucleótidos de una localización determinada en un cromosoma. Los polimorfismos/variaciones pueden comprender cualquier número de alelos específicos. Un polimorfismo/variación puede caracterizarse por la presencia de dos o más alelos en una población. El polimorfismo/variación puede caracterizarse por la presencia de tres o más alelos. El polimorfismo/variación puede caracterizarse por cuatro o más alelos, cinco o más alelos, seis o más alelos, siete o más alelos, nueve o más alelos, o diez o más alelos. Un alelo puede estar asociado con una o más enfermedades o trastornos, p. ej., un alelo de riesgo de PML puede ser un alelo que está asociado con un mayor o menor riesgo de desarrollar PML. Las variaciones genéticas y los alelos pueden utilizarse para asociar un fenotipo heredado a un genotipo responsable. Un alelo de riesgo de PML puede ser un alelo variante que está estadísticamente asociado con una detección de PML. Las variaciones genéticas pueden tener cualquier frecuencia medible en la población, p. ej., una frecuencia superior al 10 %, una frecuencia del 5-10 %, una frecuencia del 1-5 %, una frecuencia del 0.1-1 % o una frecuencia inferior al 0.1 %. Como se utiliza en el presente documento, los alelos variantes pueden ser alelos que difieren de un alelo de referencia. Tal y como se utiliza en el presente documento, una variante puede ser un segmento de ADN que difiere del ADN de referencia, tal como una variación genética. Las variaciones genéticas pueden utilizarse para rastrear la herencia de un gen que aún no se ha identificado, pero cuya localización aproximada se conoce.

Tal y como se utiliza en el presente documento, un "haplotipo" puede ser información relativa a la presencia o ausencia de uno o más marcadores genéticos en una región cromosómica determinada en un sujeto. Un haplotipo puede ser un segmento de ADN caracterizado por uno o más alelos dispuestos a lo largo del segmento, p. ej., un haplotipo puede comprender un miembro del par de alelos para cada variación genética o locus. El haplotipo puede comprender dos o más alelos, tres o más alelos, cuatro o más alelos, cinco o más alelos, o cualquier combinación de los mismos, en el que, cada alelo puede comprender una o más variaciones genéticas a lo largo del segmento.

Una variación genética puede ser una aberración funcional que puede alterar la función génica, la expresión génica, la expresión polipeptídica, la función polipeptídica o cualquier combinación de las mismas. Una variación genética puede ser una mutación de pérdida de función, una mutación de ganancia de función, una mutación dominante negativa o una reversión. Una variación genética puede formar parte de la región codificante o de la región reguladora de un gen. Las regiones reguladoras pueden controlar la expresión de los genes y, por lo tanto, la expresión de los polipéptidos. Una región reguladora puede ser un segmento de ADN al que pueden unirse polipéptidos reguladores, p. ej., factores de transcripción o de empalme. Una región reguladora puede situarse cerca del gen que se está regulando, p. ej., posiciones secuencia arriba o secuencia abajo del gen que se está regulando. Una región reguladora (p. ej., un elemento potenciador) puede tener varios miles de pares de bases secuencia arriba o secuencia abajo de un gen.

Las variantes pueden incluir cambios que afectan a un polipéptido, tal como un cambio en el nivel de expresión, secuencia, función, localización, compañeros de unión o cualquier combinación de los mismos. Una variación genética puede ser una mutación con cambio de marco, una mutación no sentido, una mutación de sentido erróneo, una mutación neutra o una mutación silenciosa. Por ejemplo, las diferencias de secuencia, cuando se comparan con una secuencia de nucleótidos de referencia, pueden incluir la inserción o eliminación de un solo nucleótido, o de más de un nucleótido, que da como resultado un desplazamiento de marco; el cambio de al menos un nucleótido, que da como resultado un cambio en el aminoácido codificado; el cambio de al menos un nucleótido, que da como resultado la generación de un codón de parada prematuro; la eliminación de varios nucleótidos, que da como resultado la eliminación de uno o más aminoácidos codificados por los nucleótidos; la inserción de uno o varios nucleótidos, tal como por recombinación desigual o conversión génica, que da como resultado la interrupción de la secuencia codificante de un marco de lectura; la duplicación de toda o parte de una secuencia; la transposición; o el reordenamiento de una secuencia de nucleótidos. Tales cambios de secuencia pueden alterar el polipéptido codificado por el ácido nucleico, p. ej., si el cambio en la secuencia del ácido nucleico causa un desplazamiento de marco, el desplazamiento de marco puede dar lugar a un cambio en los aminoácidos codificados, y/o puede dar lugar a la generación de un codón de parada prematuro, causando la generación de un polipéptido truncado. Una variación genética asociada con PML puede ser un cambio sinónimo en uno o más nucleótidos, p. ej., un cambio que no resulta en un cambio en la secuencia de aminoácidos. Dicho polimorfismo puede, p. ej., alterar los sitios de empalme, afectar a la estabilidad o el transporte del ARNm, o afectar de otro modo a la transcripción o traducción de un polipéptido codificado. Una mutación sinónima puede dar lugar a que el producto polipeptídico tenga una estructura alterada debido al uso de codones poco frecuentes que afectan al plegamiento del polipéptido durante la traducción, lo que en algunos casos puede alterar su función y/o propiedades de unión a fármacos si se trata de una diana farmacológica. Los cambios que pueden alterar el ADN aumentan la posibilidad de que se produzcan cambios estructurales, tal como amplificaciones o eliminaciones, a nivel somático. Un polipéptido codificado por la secuencia de nucleótidos de referencia puede ser un polipéptido de referencia con una secuencia de aminoácidos de referencia particular, y los polipéptidos codificados por secuencias de nucleótidos variantes pueden ser polipéptidos variantes con secuencias de aminoácidos variantes.

Las variantes de secuencia más comunes comprenden variaciones de base en una única posición de base en el genoma, y dichas variantes de secuencia, o polimorfismos, se denominan comúnmente polimorfismos de un solo nucleótido (SNP) o variantes de un solo nucleótido (SNV). Un SNP puede representar una variante genética presente en una frecuencia mayor o igual al 1 % en una población y un SNP o una SNV puede representar una variante genética presente en cualquier nivel de frecuencia en una población. Un SNP puede ser una variación en la secuencia de nucleótidos que se produce cuando un solo nucleótido en un lugar del genoma difiere entre miembros de una especie o entre cromosomas emparejados de un sujeto. Los SNP pueden incluir variantes de un solo nucleótido, p. ej., en una posición de nucleótido determinada, algunos sujetos pueden tener una "G", mientras que otros pueden tener una "C". Los SNP pueden producirse en un único evento mutacional, por lo que puede haber dos alelos posibles en cada sitio del SNP: el alelo original y el alelo mutado. Los SNP que tienen dos bases diferentes en una única posición nucleotídica se denominan SNP bialélicos, los que tienen tres se denominan trialélicos y los que tienen las cuatro bases representadas en la población son cuadralélicos. Los SNP pueden considerarse neutros. Los SNP pueden afectar a la susceptibilidad a una enfermedad (p. ej., PML). Los polimorfismos SNP pueden tener dos alelos, p. ej., un sujeto puede ser homocigoto para un alelo del polimorfismo en el que ambas copias cromosómicas del individuo tienen el mismo nucleótido en la localización del SNP, o un sujeto puede ser heterocigoto en el que los dos cromosomas hermanos del sujeto contienen nucleótidos diferentes. La nomenclatura de SNP presentada en el presente documento es el marcador oficial de identificación ID del SNP de referencia (rs) asignada a cada SNP único por el Centro Nacional de Información Biotecnológica (NCBI).

Otras variaciones genéticas de la divulgación pueden ser las variaciones en el número de copias (CNV). Tal y como se utiliza en el presente documento, las "CNV" incluyen alteraciones del ADN de un genoma que dan lugar a un número anormal de copias de una o más secciones de ADN. Una CNV puede comprender una subregión de CNV. Tal y como se utiliza en el presente documento, una "subregión de la CNV" incluye una secuencia continua de nucleótidos dentro de una CNV. La secuencia de nucleótidos de una subregión de CNV puede ser más corta que la secuencia de nucleótidos de la CNV. La subregión de CNV puede ser equivalente a la CNV (p. ej., tal como en el caso de algunas CNV). Las CNV pueden ser heredadas o causadas por mutaciones de novo y pueden ser responsables de una parte sustancial de la variabilidad fenotípica humana, de los rasgos de comportamiento y de la susceptibilidad a las enfermedades. Las CNV de la presente divulgación pueden asociarse con la susceptibilidad a una o más afecciones, p. ej., PML. Las CNV pueden incluir un único gen o incluir un conjunto contiguo de genes. Las CNV pueden estar causadas por reordenamientos estructurales del genoma, p. ej., translocaciones o inversiones desequilibradas, inserciones, eliminaciones, amplificaciones y eliminaciones intersticiales. Estos reordenamientos estructurales pueden producirse en uno o varios cromosomas. Las repeticiones de baja copia (CSF), que son secuencias repetidas específicas de una región (también conocidas como duplicaciones segmentarias), pueden ser susceptibles a estos reordenamientos estructurales, que da como resultado CNV. Factores como el tamaño, la orientación, el porcentaje de similitud y la distancia entre las copias pueden influir en la susceptibilidad de las CSF al reordenamiento genómico. Además, los reordenamientos pueden estar mediados por la presencia de

repeticiones de alto número de copias, tal como elementos intercalados largos (LINE) y elementos intercalados cortos (SINE), a menudo a través de recombinación no homóloga. Por ejemplo, los reordenamientos cromosómicos pueden surgir de la recombinación homóloga no alélica durante la meiosis o a través de un mecanismo basado en la replicación, tal como el bloqueo de horquilla y el cambio de molde (FoSTeS) (Zhang F. et al., Nat. Genet. (2009)) o la reparación inducida por rotura mediada por microhomología (MMBIR) (Hastings P. J. et al., PLoS Genetics (2009)). Las CNV pueden llamarse variantes estructurales, que son una clase más amplia de variante que también incluye alteraciones neutras del número de copias, tal como inversiones equilibradas y translocaciones equilibradas.

Las CNV pueden explicar la variación genética que afecta a una proporción sustancial del genoma humano, p. ej., las CNV conocidas pueden cubrir más del 15 % de la secuencia del genoma humano (Estivill y Armengol, PLoS Genetics (2007)). Las CNV pueden afectar a la expresión génica, la variación fenotípica y la adaptación al interrumpir o alterar la dosificación génica, y pueden causar enfermedades, p. ej., trastornos por microeliminación y microduplicación, y pueden conferir susceptibilidad a enfermedades y trastornos. Puede encontrarse información actualizada sobre la ubicación, el tipo y el tamaño de las CNV conocidas en una o varias bases de datos, p. ej., la Base de Datos de Variantes Genómicas (véase, MacDonald JR et al., Nucleic Acids Res., 42, D986-92 (2014), que actualmente contiene datos de más de 500,000 CNV (a mayo de 2016).

Otros tipos de variantes de secuencia pueden encontrarse en el genoma humano y pueden asociarse con una enfermedad o trastorno, incluyendo, pero no limitándose a, microsatélites. Los marcadores microsatélite son estables, polimórficos, fáciles de analizar y pueden aparecer regularmente en todo el genoma, lo que los hace especialmente adecuados para el análisis genético. Un microsatélite polimórfico puede comprender múltiples repeticiones pequeñas de bases, p. ej., repeticiones de CA, en un sitio concreto en el que el número de longitudes de repetición varía en una población. Los microsatélites, p. ej., el número variable de repeticiones en tándem (VNTR), pueden ser segmentos cortos de ADN que tienen una o más secuencias repetidas, p. ej., aproximadamente de 2 a 5 nucleótidos de longitud, que pueden ocurrir en el ADN no codificante. Los cambios en los microsatélites pueden producirse durante la recombinación genética de la reproducción sexual, aumentando o disminuyendo el número de repeticiones que se encuentran en un alelo, o cambiando la longitud del alelo.

Las variaciones genéticas divulgadas en el presente documento pueden asociarse con un riesgo de desarrollar PML en un sujeto. En algunos casos, el sujeto puede tener un riesgo disminuido debido a la ausencia de una o más variaciones genéticas que alteren o modulen un gen correspondiente de acuerdo con las Tablas 1 a 26. Por ejemplo, el sujeto puede tener un riesgo disminuido debido a la ausencia de una o más variaciones genéticas que alteren o modulen un gen correspondiente de acuerdo con las Tablas 3 y 6. En algunos casos, el sujeto puede tener un riesgo aumentado debido a la presencia de una o más variaciones genéticas que alteran o modulan un gen correspondiente de acuerdo con las Tablas 1 a 26. Por ejemplo, el sujeto puede tener un riesgo aumentado debido a la presencia de una o más variaciones genéticas que alteren o modulen un gen correspondiente de acuerdo con las Tablas 3 y 6. En algunos casos, uno o más genes enumerados en las Tablas 25A, 25B y 26 pueden eliminarse de cualquiera de las Tablas 1-24. En algunos casos, uno o más genes enumerados en las Tablas 25A, 25B y 26 pueden añadirse a cualquiera de las Tablas 1-24.

Tabla 25A: panel ejemplar de 8 genes

Símbolo del gen RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
BAG3	AR	Base de datos pública	PMID: 19229298, 19282432, 22984599, 27042682	175
BTK	XLR	Base de datos pública	PMID: 18281276, 23765059, 25930993, 26029204	180
CD40LG	XLR	Base de datos pública	PMID: 17360404, 21455173, 23765059, 26008899, 26029204	206
DOCK8	AR	Base de datos pública	PMID: 23765059, 23887241, 26029204, 26454313	242
MAGT1	XLR	Base de datos pública	PMID: 23887241, 25504528, 27873163	326
RAG1	AD_AR	Base de datos pública	PMID: 23122631, 23765059, 23887241, 25976673, 26029204, 26454313, 27484032, 27808398	370
STAT1	AD_AR	Base de datos pública	PMID: 23887241, 25645939, 26029204, 26513235, 26743090, 27821552, 27873163	436

Símbolo del gen RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
WAS	XLR	Ambos	PMID: 12874226, 14647476, 19782549, 20008220, 24753205, 26029204, 26371186	483

Tabla 25B: panel ejemplar de 16 genes

Símbolo del gen RefSeq	Modelo de enfermedad	Fuente del gen	Fuente Anotación	Número de gen (GN)
ADA	AR	Ambos	PMID: 23765059, 24135998, 25930993, 26029204, 26454313	1
BAG3	AR	Base de datos pública	PMID: 19229298, 19282432, 22984599, 27042682	175
BTK	XLR	Base de datos pública	PMID: 18281276, 23765059, 25930993, 26029204	180
CD40LG	XLR	Base de datos pública	PMID: 14647476, 17360404, 21455173, 23765059, 26008899,26029204	206
DNMT3B	AR	Base de datos pública	PMID: 23486536, 23765059, 26029204, 26851945	240
DOCK8	AR	Base de datos pública	PMID: 23765059, 23887241, 26029204, 26454313	242
ITK	AR	Base de datos pública	PMID: 14647476, 23765059, 26029204, 26454313	308
LCK	AR	Base de datos pública	PMID: 14647476, 23765059, 26029204, 26454313	316
PNP	AR	Ambos	PMID: 26029204, 26454313	354
RAG1	AD_AR	Base de datos pública	PMID: 23122631, 23765059, 23887241, 25976673, 26029204, 26454313, 27484032, 27808398	370
STAT1	AD_AR	Base de datos pública	PMID: 23887241, 25645939, 26029204, 26513235, 26743090, 27821552, 27873163	436
STAT3	AD	Base de datos pública	PMID: 23765059, 23887241, 25645939, 25930993, 26029204, 27658964, 27873163	438
STK3	desconocido	Ambos	PMID: 26029204	135
TYK2	AR	Base de datos pública	PMID: 26029204, 26513235, 27821552	144
WAS	XLR	Ambos	PMID: 12874226, 19782549, 20008220, 24753205, 26029204, 26371186	483
WIPF1	AR	Base de datos pública	PMID: 23765059, 26029204, 26453379	485

Tabla 26: panel ejemplar de 2 genes

Símbolo del gen RefSeq	Solapamiento de exones	ID del Gen del NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
ADA	intrónico	100	adenosina desaminasa	Este gen codifica una enzima que cataliza la hidrólisis de la adenosina en inosina. Se han descrito varias mutaciones de este gen y se han relacionado con enfermedades humanas. La deficiencia de esta enzima causa una forma de enfermedad de inmunodeficiencia combinada grave (SCID), en la que hay disfunción de los linfocitos B y T con alteración de la inmunidad celular y disminución de la producción de inmunoglobulinas, mientras que los niveles elevados de esta enzima se han asociado con anemia hemolítica congénita. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor consúltese la ficha de Gene para acceder a publicaciones adicionales ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC040226.1, X02994.1 [ECO:0000332] Intrones de la secuencia de ARN :: una muestra mixta/parcial admite ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	1
STK3	intrónico	6788	serina/treonina -proteína quinasa 3 isoforma 1	Este gen codifica una proteína quinasa serina/treonina activada por moléculas proapoptóticas, lo que indica que la proteína codificada funciona como supresor del crecimiento. La escisión del producto proteico por la caspasa elimina la porción inhibidora del terminal C. La porción del terminal N es transportada al núcleo donde se homodimeriza para formar la quinasa activa que promueve la condensación de la cromatina durante la apoptosis. Se han encontrado múltiples variantes de transcripción que codifican diferentes isoformas para este gen. [proporcionado por RefSeq, enero de 2012]. Variante de transcripción: Esta variante (1) codifica la isoforma 1. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor consúltese la ficha de Gene para acceder a publicaciones adicionales ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: U26424.1, BC010640.2 [ECO:0000332] intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	135

Sujetos

Un "sujeto", tal como se utiliza en el presente documento, puede ser un individuo de cualquier edad o sexo del que se obtiene una muestra que contiene polinucleótidos para su análisis mediante uno o más métodos descritos en el presente documento con el fin de obtener información sobre el ácido polinucleico; p. ej., un adulto de sexo masculino o femenino, un niño, un recién nacido o un feto. Un sujeto puede ser cualquier diana de administración terapéutica. Un sujeto puede ser un sujeto de prueba o un sujeto de referencia.

Tal como se utiliza en el presente documento, una "cohorte" puede representar un grupo étnico, un grupo de pacientes, un grupo de edad particular, un grupo no asociado con una afección particular (p. ej., enfermedad o trastorno), un grupo asociado con una afección particular (p. ej., enfermedad o trastorno), un grupo de sujetos asintomáticos, un grupo de sujetos sintomáticos, o un grupo o subgrupo de sujetos asociados con una respuesta particular a un régimen de tratamiento o inscritos en un ensayo clínico. Un paciente puede ser un sujeto aquejado de una afección (p. ej., enfermedad o trastorno). Un paciente puede ser un sujeto no afectado por una afección (p. ej., enfermedad o trastorno) y considerado aparentemente sano, o un sujeto normal o de control. Un sujeto puede ser un sujeto de prueba, un paciente o un candidato a una terapéutica, en el que el ADN genómico del sujeto, paciente o candidato se obtiene para su análisis mediante uno o más métodos de la presente divulgación, con el fin de obtener información sobre la variación genética del sujeto, paciente o candidato.

La muestra de ácido polinucleico puede obtenerse prenatalmente de un feto o embrión o de la madre, p. ej., de células fetales o embrionarias en la circulación materna. La muestra de ácido polinucleico puede obtenerse con la ayuda de un profesional sanitario, p. ej., para extraer sangre. La muestra de ácido polinucleico puede obtenerse sin la asistencia de un profesional sanitario, p. ej., cuando la muestra de ácido polinucleico se obtiene de forma no invasiva, tal como una muestra de saliva, o una muestra que comprende células bucales que se obtiene utilizando un hisopo o cepillo bucal, o una muestra de enjuague bucal.

La presente divulgación también proporciona métodos para evaluar variaciones genéticas en sujetos que son miembros de una población diana. Dicha población diana puede ser una población o grupo de sujetos en riesgo de desarrollar la afección (p. ej., enfermedad o trastorno), basándose, p. ej., en otros factores genéticos, biomarcadores, parámetros biofísicos, pruebas diagnósticas como las imágenes de resonancia magnética (MRI), antecedentes familiares de la afección, detección previo o historial médico, o cualquier combinación de los mismos.

Las variaciones genéticas de la presente divulgación que se encuentren asociadas con una afección (p. ej., enfermedad o trastorno) pueden mostrar una asociación similar en otras poblaciones humanas. Por lo tanto, también se contemplan poblaciones humanas de sujetos y dentro del alcance de la divulgación. Los sujetos humanos pueden pertenecer a una o más poblaciones humanas, incluidas, entre otras, poblaciones caucásicas, judías asquenazíes, judías sefardíes, europeas, americanas, euroasiáticas, asiáticas, centro/sudasiáticas, asiáticas orientales, de Oriente Medio, africanas, hispanas, caribeñas y oceánicas. Las poblaciones europeas incluyen, entre otras, suecas, noruegas, finlandesas, rusas, danesas, islandesas, irlandesas, celtas, inglesas, escocesas, neerlandesas, belgas, francesas, alemanas, españolas, portuguesas, italianas, polacas, búlgaras, eslavas, serbias, bosnias, checas, griegas y turcas. La contribución étnica en los sujetos también puede determinarse mediante análisis genéticos; p. ej., el análisis genético de la ascendencia puede llevarse a cabo utilizando marcadores de microsatélite no vinculados o polimorfismos de un solo nucleótido (SNP) como los expuestos en Smith *et al.,* (Smith M. W. et al., Am. J. Hum. Genet, 74: 1001 (2004)).

Ciertas variaciones genéticas pueden tener frecuencias poblacionales diferentes en poblaciones diferentes, o ser polimórficas en una población, pero no en otra. Los métodos disponibles y tal como se plantean en el presente documento pueden aplicarse para poner en práctica la presente divulgación en cualquier población humana. Esto puede incluir la evaluación de las variaciones genéticas de la presente divulgación, con el fin de identificar los marcadores que dan la asociación más fuerte dentro de la población específica. De este modo, las variantes de riesgo de la presente divulgación pueden residir en diferentes haplotipos y en diferentes frecuencias en diversas poblaciones humanas.

Afecciones y medicamentos inmunosupresores

Un sujeto puede ser diagnosticado o no diagnosticado con una afección (p. ej., enfermedad o trastorno), puede ser asintomático o sintomático, puede tener susceptibilidad aumentada o disminuida a una afección (p. ej., enfermedad o trastorno), puede estar actualmente bajo o previamente bajo o no bajo un tratamiento para una afección (p. ej., enfermedad o trastorno), o cualquier combinación de los mismos. La afección puede ser sida, cáncer, trasplante de órganos, tuberculosis, sarcoidosis o una enfermedad autoinmune. La afección puede ser PML.

Un sujeto puede ser diagnosticado o no diagnosticado con PML, puede ser asintomático o sintomático, puede tener susceptibilidad aumentada o disminuida a PML, puede estar actualmente bajo o previamente bajo o no bajo un tratamiento para PML, o cualquier combinación de los mismos. Un sujeto puede estar diagnosticado o no de SIDA (p. ej., individuos infectados por el VIH), puede ser asintomático o sintomático, puede tener una susceptibilidad aumentada o disminuida al SIDA, puede estar sometido actual o con anterioridad a un tratamiento contra el SIDA, o no estarlo, o cualquier combinación de los mismos. Un sujeto puede estar diagnosticado o no de cáncer (p. ej., enfermedad de Hodgkin, leucemia, linfoma o mielofibrosis), puede ser asintomático o sintomático, puede tener una susceptibilidad aumentada o disminuida al cáncer, puede estar sometido actual o con anterioridad a un tratamiento contra el cáncer, o no estarlo, o cualquier combinación de los mismos. Un sujeto puede estar actualmente diagnosticado o previamente diagnosticado o no diagnosticado de una enfermedad autoinmune (p. ej., esclerosis múltiple, artritis reumatoide, psoriasis, lupus eritematoso

sistémico), puede ser asintomático o sintomático, puede tener una susceptibilidad aumentada o disminuida a una enfermedad autoinmune, puede estar actualmente sometido o previamente sometido o no a un tratamiento para una enfermedad autoinmune, o cualquier combinación de los mismos.

Se entiende que el término "cáncer" incluye todos los tipos de crecimientos cancerosos o procesos oncogénicos, tejidos metastásicos o células, tejidos u órganos malignamente transformados, independientemente del tipo histopatológico o estadio de invasividad. Un tumor metastásico puede surgir de multitud de tipos de tumores primarios, incluidos, entre otros, los de mama, pulmón, hígado, colon y ovario. Algunos ejemplos de cánceres son, entre otros, fibrosarcoma, miosarcoma, liposarcoma, condrosarcoma, sarcoma osteogénico, cordoma, angiosarcoma, endoteliosarcoma, linfangiosarcoma, linfangioendoteliosarcoma, sinovioma, mesotelioma, tumor de Ewing, leiomiosarcoma, rabdomiosarcoma, cáncer gástrico, cáncer de esófago, cáncer rectal, cáncer de páncreas, cáncer de ovario, cáncer de próstata, cáncer de útero, cáncer de cabeza y cuello, cáncer de piel, cáncer cerebral, carcinoma de células escamosas, carcinoma de glándulas sebáceas, carcinoma papilar, adenocarcinoma papilar, cistadenocarcinoma, carcinoma medular, carcinoma broncogénico, carcinoma de células renales, hepatoma, carcinoma de vías biliares, coriocarcinoma, seminoma, carcinoma embrionario, tumor de Wilm, cáncer de cuello uterino, cáncer de testículo, carcinoma pulmonar de células pequeñas, carcinoma pulmonar de células no pequeñas, carcinoma de vejiga, carcinoma epitelial, glioma, astrocitoma, meduloblastoma, craneofaringioma, ependimoma, pinealoma, hemangioblastoma, neuroma acústico, oligodendroglioma, meningioma, melanoma, neuroblastoma, retinoblastoma, leucemia, linfoma, mielofibrosis o sarcoma de Kaposi.

El término "enfermedad autoinmune" incluye todos los tipos de estados patológicos que surgen de respuestas inmunes anormales del cuerpo a sustancias y tejidos que están normalmente presentes en el cuerpo. Ejemplos de enfermedades autoinmunes son, entre otras, la enfermedad de Addison, la encefalitis por receptores anti-NMDA, el síndrome antisintetasa, la anemia aplásica, las anemias autoinmunes, la anemia hemolítica autoinmune, la pancreatitis autoinmune, la enfermedad de Behcet, los trastornos bullosos de la piel, la enfermedad celíaca - esprúe (enteropatía sensible al gluten), síndrome de fatiga crónica, polineuropatía desmielinizante inflamatoria crónica, leucemia linfocítica crónica, enfermedad de Crohn, dermatomiositis, enfermedad de Devic, eritroblastopenia, síndrome de Evans, glomeruloesclerosis segmentaria focal, granulomatosis con poliangitis, enfermedad de Graves, oftalmopatía de Graves, síndrome de Guillain-Barré, tiroiditis de Hashimoto, púrpura trombocitopénica idiopática (ITP), nefropatía por IgA, enfermedades autoinmunes mediadas por IgA, enfermedad relacionada con IgG4, enfermedad inflamatoria intestinal, artritis idiopática juvenil, esclerosis múltiple, miastenia gravis, mieloma, linfoma no Hodgkin, síndrome opsoclono-mioclono (OMS), penfigoide, pénfigo, pénfigo vulgar, anemia perniciosa, polimiositis, psoriasis, aplasia pura de eritrocitos, artritis reactiva, artritis reumatoide, sarcoidosis, esclerodermia, síndrome de Sjögren, lupus eritematoso sistémico, púrpura trombocitopénica, púrpura trombocitopénica trombótica, diabetes de tipo I, colitis ulcerosa, vasculitis (por ej., vasculitis asociada a anticuerpos citoplasmáticos antineutrófilos) y vitíligo.

Un sujeto puede ser tratado actualmente con un medicamento inmunosupresor. Un sujeto puede ser tratado previamente con un medicamento inmunosupresor. Un sujeto puede no estar aún tratado con un medicamento inmunosupresor. El medicamento inmunosupresor puede incluir, entre otros, glucocorticoides, citostáticos, anticuerpos, fármacos que actúan sobre las inmunofilinas, interferones, opioides, proteínas de unión al TNF, micofenolato u otros agentes biológicos pequeños. Por ejemplo, los glucocorticoides pueden incluir, entre otros, cortisol (hidrocortisona), cortisona, prednisona, prednisolona, metilprednisolona, dexametasona, betametasona, triamcinolona, beclometasona, acetato de fludrocortisona, acetato de desoxicorticosterona (DOCA) o aldosterona. Los citostáticos pueden incluir, entre otros, mostazas nitrogenadas (p. ej., ciclofosfamida), nitrosoureas, compuestos de platino, análogos del ácido fólico como el metotrexato, análogos de las purinas como la azatioprina y la mercaptopurina, análogos de las pirimidinas como el fluorouracilo, inhibidores de la síntesis de proteínas, antibióticos citotóxicos como la dactinomicina, las antraciclinas, la mitomicina C, la bleomicina o la mitramicina. Los anticuerpos pueden incluir, entre otros, anticuerpos policlonales como atgam y timoglobulina, anticuerpos monoclonales como anticuerpos dirigidos contra CD25 y CD3, muromonab-CD3, basiliximab (p. ej., SIMULECT) y daclizumab (p. ej., ZENAPAX). Los fármacos que actúan sobre las inmunofilinas pueden ser, entre otros, ciclosporina, tacrolimus, sirolimus o everolimus. Las proteínas de unión al TNF pueden incluir, entre otras, infliximab (p. ej., REMICADE), etanercept (p. ej., ENBREL) o adalimumab (p. ej., HUMIRA). Otros agentes biológicos pequeños pueden incluir, entre otros, el fingolimod y la miriocina.

El medicamento inmunosupresor puede ser fármacos para el tratamiento de la esclerosis múltiple, entre ellos, interferón beta-1a (p. ej., AVONEX, REBIF), interferón beta-1b (p. ej., BETASERON, EXTAVIA), acetato de glatiramero (p. ej., COPAXONE, GLATOPA), peginterferón beta-1a (p. ej., PLEGRIDY), teriflunomida (p. ej., AUBAGIO), fingolimod (p. ej., GILENYA), fumarato de dimetilo (p. ej., TECFIDERA), alemtuzumab (p. ej., LEMTRADA), mitoxantrona (p. ej., NOVANTRONE), natalizumab (p. ej., TYSABRI), daclizumab (p. ej., ZINBRYTA) u ocrelizumab (p. ej., OCREVUS).

El medicamento inmunosupresor puede ser adalimumab (p. ej., HUMIRA), alemtuzumab (p. ej., LEMTRADA), alentuzumab (p. ej., CAMPATH), azatioprina (p. ej., IMURAN), belimumab (p. ej., BENLYSTA), bevacizumab (p. ej., AVASTATIN), bortezomib (p. ej., VELCADE), eculizumab (p. ej., SOLIRIS), leflunomida, brentuximab

vedotina (p. ej., ADCETRIS), cetuximab (p. ej., ERBITUX), ciclofosfamida, fumarato de dimetilo (p. ej., TECFIDERA), efalizumab (p. ej., RAPTIVA), fingolimod (p. ej., GILENYA), fludarabina (p. ej., FLUDARA), ácido fumárico, imatinib (p. ej., GLEEVEC, GLIVEC), infliximab (p. ej., REMICADE), metotrexato (p. ej., TREXALL, RHEUMATREX), micofenolato de mofetilo (p. ej., CELLCEPT), natalizumab (p. ej., TYSABRI), daclizumab (p. ej., ZINBRYTA), rituximab (p. ej., RITUXIN), vedolizumab (p. ej., ENTYVIO), ruxolitinib (p. ej., JAKAFI, JAKAVI) u ocrelizumab (p. ej., Ocrevus).

Muestras

Las muestras que son adecuadas para su uso en los métodos descritos en el presente documento pueden ser muestras de ácido polinucleico de un sujeto. Una "muestra de ácido polinucleico", tal como se utiliza en el presente documento, puede incluir ARN o ADN, o una combinación de ambos. Una "muestra de polipéptidos" (p. ej., péptidos o proteínas, o fragmentos de los mismos) puede utilizarse para determinar la información de que se ha producido un cambio de aminoácidos, que es el resultado de una variante genética. Los ácidos polinucleicos y los polipéptidos pueden extraerse de una o más muestras, incluidas, entre otras, sangre, saliva, orina, raspados mucosos del revestimiento de la boca, expectorante, suero, lágrimas, piel, tejido o cabello. Una muestra de ácido polinucleico puede analizarse para obtener información sobre el ácido polinucleico. "Información sobre el ácido polinucleico", tal como se utiliza en el presente documento, incluye una secuencia de ácido polinucleico en sí, la presencia/ausencia de variación genética en la secuencia de ácido polinucleico, una propiedad física que varía en función de la secuencia del ácido polinucleico (p. ej., Tm) y la cantidad de ácido polinucleico (p. ej., el número de copias de ARNm). Por "ácido polinucleico" se entiende cualquiera de los siguientes: ADN, ARN, ADN con nucleótidos artificiales o ARN con nucleótidos artificiales. Tal como se utiliza en el presente documento, un "ácido polinucleico purificado" incluye ADNc, fragmentos de ácidos polinucleicos genómicos, ácidos polinucleicos producidos mediante la reacción en cadena de la polimerasa (PCR), ácidos polinucleicos formados por tratamiento con enzimas de restricción de ácidos polinucleicos genómicos, ácidos polinucleicos recombinantes y moléculas de ácido polinucleico sintetizadas químicamente. Una molécula de ácido polinucleico "recombinante" incluye una molécula de ácido polinucleico fabricada mediante una combinación artificial de dos segmentos de secuencia separados de otro modo, p. ej., por síntesis química o por la manipulación de segmentos aislados de ácidos polinucleicos mediante técnicas de ingeniería genética. Tal como se utiliza en el presente documento, un "polipéptido" incluye proteínas, fragmentos de proteínas y péptidos, ya sean aislados de fuentes naturales, producidos mediante técnicas recombinantes o sintetizados químicamente. Un polipéptido puede tener una o más modificaciones, tal como una modificación postraduccional (p. ej., glicosilación, fosforilación, etc.) o cualquier otra modificación (p. ej., pegilación, etc.). El polipéptido puede contener uno o más aminoácidos no naturales (p. ej., tal como un aminoácido con una modificación de la cadena lateral).

La muestra de ácido polinucleico puede comprender células o tejidos, p. ej., líneas celulares. Algunos tipos celulares ejemplares a partir de los cuales pueden obtenerse ácidos nucleicos mediante los métodos descritos en el presente documento incluyen, entre otros, los siguientes: una célula sanguínea, tal como un linfocito B, un linfocito T, un leucocito, un eritrocito, un macrófago o un neutrófilo; una célula muscular, tal como una célula esquelética, una célula muscular lisa o una célula muscular cardiaca; una célula germinal, tal como un espermatozoide o un óvulo; una célula epitelial; una célula del tejido conjuntivo, tal como un adipocito, un condrocito, un fibroblasto o un osteoblasto; una neurona; un astrocito; una célula estromal; una célula específica de un órgano, tal como una célula renal, una célula pancreática, una célula hepática o un queratinocito; una célula madre; o cualquier célula que se desarrolle a partir de ellas. Una célula a partir de la cual pueden obtenerse ácidos nucleicos puede ser una célula sanguínea o un tipo particular de célula sanguínea incluyendo, p. ej., una célula madre hematopoyética o una célula que surge de una célula madre hematopoyética como un glóbulo rojo, un linfocito B, un linfocito T, una célula asesina natural, un neutrófilo, un basófilo, un eosinófilo, un monocito, un macrófago o una plaqueta. En general, puede utilizarse cualquier tipo de célula madre, incluyendo, sin limitación, una célula madre embrionaria, una célula madre adulta o una célula madre pluripotente.

Una muestra de ácido polinucleico puede procesarse para el aislamiento de ARN o ADN, p. ej., el ARN o ADN de una muestra de células o tejido puede separarse de otros componentes de la muestra de ácido polinucleico. Las células pueden extraerse de una muestra de ácido polinucleico mediante técnicas estándar, p. ej., centrifugando una muestra de células y resuspendiendo las células precipitadas, p. ej., en una solución tamponada, p. ej., solución salina tamponada con fosfato (PBS). Tras centrifugar la suspensión celular para obtener un sedimento celular, las células pueden lisarse para extraer el ADN. La muestra de ácido nucleico puede concentrarse y/o purificarse para aislar el ADN. Todas las muestras de ácido nucleico obtenidas de un sujeto, incluidas las sometidas a cualquier tipo de procesamiento posterior, se consideran obtenidas del sujeto. Para extraer ARN o ADN de una muestra de ácido nucleico se pueden utilizar técnicas y kits estándar conocidos en la técnica, tal como p. ej., la extracción con fenol, un kit para tejido QIAAMP® (Qiagen, Chatsworth, California), un kit de purificación de ADN genómico WIZARD® (Promega) o un método Autopure de Qiagen que utiliza la química Puregene, que puede permitir la purificación de ADN altamente estable y adecuado para archivo.

Determinar la identidad de un alelo o determinar el número de copias puede, pero no es necesario, incluir la obtención de una muestra de ácido polinucleico que comprenda ARN y/o ADN de un sujeto, y/o evaluar la

identidad, el número de copias, la presencia o ausencia de una o más variaciones genéticas y sus localizaciones cromosómicas dentro del ADN genómico (es decir, el genoma del sujeto) derivado de la muestra de ácido polinucleico.

El individuo u organización que realiza la determinación no necesita realmente llevar a cabo el análisis físico de una muestra de ácido nucleico de un sujeto. Los métodos pueden incluir el uso de información obtenida mediante el análisis de la muestra de ácido polinucleico por un tercero. Los métodos pueden incluir etapas que tienen lugar en más de un sitio. Por ejemplo, se puede obtener una muestra de ácido polinucleico de un sujeto en un primer lugar, tal como en un centro sanitario o en el domicilio del sujeto en el caso de un kit de autodiagnóstico. La muestra de ácido polinucleico puede analizarse en el mismo lugar o en un segundo lugar, p. ej., en un laboratorio u otro centro de pruebas.

Ácidos nucleicos

Los ácidos nucleicos y polipéptidos descritos en el presente documento pueden utilizarse en métodos y kits de la presente divulgación. Los aptámeros que se unen específicamente a los ácidos nucleicos y polipéptidos descritos en el presente documento pueden utilizarse en los métodos y kits de la presente divulgación. Tal como se utiliza en el presente documento, un ácido nucleico puede comprender un desoxirribonucleótido (ADN) o un ribonucleótido (ARN), ya sea singular o en polímeros, de origen natural o no natural, de doble cadena o de cadena sencilla, codificante, p. ej., un gen traducido, o no codificante, p. ej., una región reguladora, o cualquier fragmento, derivado, mimético o complemento de los mismos. Los ácidos nucleicos pueden comprender oligonucleótidos, nucleótidos, polinucleótidos, secuencias de ácidos nucleicos, secuencias genómicas, ADN complementario (ADNc), ácidos nucleicos antisentido, regiones de ADN, sondas, cebadores, genes, regiones reguladoras, intrones, exones, marcos de lectura abierta, sitios de unión, ácidos nucleicos diana y ácidos nucleicos específicos de alelo.

Una "sonda", tal como se usa en el presente documento, incluye un fragmento de ácido nucleico para examinar un ácido nucleico en un espécimen usando la reacción de hibridación basada en la complementariedad del ácido nucleico.

Un "híbrido", tal como se utiliza en el presente documento, incluye una doble cadena formada entre cualquiera de los ácidos nucleicos antes mencionados, dentro del mismo tipo, o a través de diferentes tipos, incluyendo, ADN-ADN, ADN-ARN, ARN-ARN o similares.

Los ácidos nucleicos "aislados", tal y como se utilizan en el presente documento, están separados de los ácidos nucleicos que normalmente flanquean el gen o la secuencia de nucleótidos (como en las secuencias genómicas) y/o han sido total o parcialmente purificados de otras secuencias transcritas (p. ej., tal como en una biblioteca de ARN). Por ejemplo, los ácidos nucleicos aislados de la divulgación pueden estar sustancialmente aislados con respecto al medio celular complejo en el que se encuentran de forma natural, o al medio de cultivo cuando se producen mediante técnicas recombinantes, o a precursores químicos u otras sustancias químicas cuando se sintetizan químicamente. En algunos casos, el material aislado puede formar parte de una composición, p. ej., un extracto crudo que contenga otras sustancias, un sistema tampón o una mezcla de reactivos. El material puede purificarse hasta alcanzar la homogeneidad esencial utilizando métodos conocidos en la técnica, p. ej., mediante electroforesis en gel de poliacrilamida (PAGE) o cromatografía en columna (p. ej., HPLC). Con respecto al ADN genómico (ADNg), el término "aislado" también puede referirse a los ácidos nucleicos que están separados del cromosoma con el que el ADN genómico está asociado de forma natural. Por ejemplo, la molécula de ácido nucleico aislada puede contener menos de aproximadamente 250 kb, 200 kb, 150 kb, 100 kb, 75 kb, 50 kb, 25 kb, 10 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb o 0.1 kb de los nucleótidos que flanquean la molécula de ácido nucleico en el ADNg de la célula de la que se deriva la molécula de ácido nucleico.

Los ácidos nucleicos que pueden fusionarse a otras secuencias codificantes o reguladoras pueden considerarse aislados. Por ejemplo, el ADN recombinante contenido en un vector se incluye en la definición de "aislado" usado en el presente documento. Los ácidos nucleicos aislados pueden incluir moléculas de ADN recombinante en células huésped heterólogas u organismos heterólogos, así como moléculas de ADN parcial o sustancialmente purificadas en solución. Los ácidos nucleicos aislados también abarcan los transcripciones de ARN *in vivo* e *in vitro* de las moléculas de ADN de la presente divulgación. Una molécula aislada de ácido nucleico o secuencia de nucleótidos puede sintetizarse químicamente o por medios recombinantes. Tales secuencias nucleotídicas aisladas pueden ser útiles, p. ej., en la fabricación del polipéptido codificado, tal como sondas para aislar secuencias homólogas (p. ej., de otras especies de mamíferos), para el mapeo de genes (p. ej., por hibridación *in situ* con cromosomas), o para detectar la expresión del gen, en tejido (p. ej., tejido humano), tal como por análisis de transferencia Northern u otras técnicas de hibridación divulgadas en el presente documento. La divulgación también se refiere a secuencias de ácido nucleico que hibridan en condiciones de hibridación de alta rigurosidad, tal como para hibridación selectiva, a una secuencia nucleotídica descrita en el presente documento. Tales secuencias de ácido nucleico pueden detectarse y/o aislarse mediante hibridación específica de alelo o secuencia (p. ej., en condiciones de alta rigurosidad). Las condiciones de rigurosidad y los métodos para las hibridaciones de ácidos nucleicos son bien conocidos por el

experto (véase, p. ej., Current Protocols in Molecular Biology, Ausubel, F. et al., John Wiley & Sons, (1998), y Kraus, M. y Aaronson, S., Methods Enzymol, 200: 546-556 (1991).

Los cálculos de "identidad" o "porcentaje de identidad" entre dos o más secuencias de nucleótidos o aminoácidos pueden determinarse alineando las secuencias con fines de comparación óptima (p. ej., pueden introducirse huecos en la secuencia de una primera secuencia). A continuación, se comparan los nucleótidos en las posiciones correspondientes, y el porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias (es decir, % de identidad = # de posiciones idénticas/# total de posiciones x 100). Por ejemplo, si una posición de la primera secuencia está ocupada por el mismo nucleótido que la posición correspondiente de la segunda secuencia, las moléculas son idénticas en esa posición. El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de huecos y la longitud de cada hueco, que deben introducirse para una alineación óptima de las dos secuencias.

La longitud de una secuencia alineada con fines de comparación puede ser de al menos el 30 %, al menos el 40 %, al menos el 50 %, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 %, o al menos el 95 %, de la longitud de la secuencia de referencia. La comparación real de las dos secuencias puede realizarse por métodos bien conocidos, p. ej., utilizando un algoritmo matemático. Un ejemplo no limitativo de dicho algoritmo matemático se describe en Karlin, S. y Altschul, S., Proc. Natl. Acad. Sci. USA., 90-5873-5877 (1993). Dicho algoritmo está incorporado en los programas NBLAST y XBLAST (versión 2.0), tal y como se describe en Altschul, S. et al., Nucleic Acids Res., 25: 3389-3402 (1997). Cuando se utilizan los programas BLAST y Gapped BLAST, se puede utilizar cualquier parámetro relevante de los respectivos programas (p. ej., NBLAST). Por ejemplo, los parámetros para la comparación de secuencias pueden establecerse en puntuación = 100, longitud de palabra = 12, o pueden variarse (p. ej., W=5 o W=20). Otros ejemplos son el algoritmo de Myers y Miller, CABIOS (1989), ADVANCE, ADAM, BLAT y FASTA. El porcentaje de identidad entre dos secuencias de aminoácidos puede obtenerse utilizando, p. ej., el programa GAP del paquete de software GCG (Accelrys, Cambridge, Reino Unido).

Las "sondas" o "cebadores" pueden ser oligonucleótidos que se hibridan de forma específica de la base con una cadena complementaria de una molécula de ácido nucleico. Las sondas pueden incluir cebadores, que pueden ser una sonda oligonucleotídica monocatenaria que puede actuar como punto de iniciación de la síntesis de ADN dirigida por molde utilizando métodos que incluyen, entre otros, la reacción en cadena de la polimerasa (PCR) y la reacción en cadena de la ligasa (CSF) para la amplificación de una secuencia diana. Los oligonucleótidos, tal como se describen en el presente documento, pueden incluir segmentos o fragmentos de secuencias de ácido nucleico, o sus complementos. Los segmentos de ADN pueden tener entre 5 y 10,000 bases contiguas, y pueden oscilar entre 5, 10, 12, 15, 20 o 25 nucleótidos y 10, 15, 20, 25, 30, 40, 50, 100, 200, 500, 1000 o 10,000 nucleótidos. Además de ADN y ARN, las sondas y los cebadores pueden incluir ácidos nucleicos polipeptídicos (PNA), tal como se describe en Nielsen, P. et al., Science 254: 1497-1500 (1991). Una sonda o cebador puede comprender una región de secuencia nucleotídica que hibrida con al menos aproximadamente 15, típicamente aproximadamente 20-25, o aproximadamente 40, 50, 60 o 75, nucleótidos consecutivos de una molécula de ácido nucleico.

La presente divulgación también proporciona ácidos nucleicos aislados, p. ej., sondas o cebadores, que contienen un fragmento o porción que puede hibridar selectivamente con un ácido nucleico que comprende, o consiste en, una secuencia de nucleótidos, en la que la secuencia de nucleótidos puede comprender al menos un polimorfismo o alelo polimórfico contenido en las variaciones genéticas descritas en el presente documento o el nucleótido de tipo silvestre que se encuentra en la misma posición, o los complementos de los mismos. La sonda o el cebador pueden ser al menos un 70 % idénticos, al menos un 80 % idénticos, al menos un 85 % idénticos, al menos un 90 % idénticos, o al menos un 95 % idénticos, a la secuencia contigua de nucleótidos o al complemento de la secuencia contigua de nucleótidos.

Una sonda de ácido nucleico puede ser un oligonucleótido capaz de hibridar con una región complementaria de un gen asociado con una afección (p. ej., PML) que contiene una variación genética descrita en el presente documento. Los fragmentos de ácido nucleico de la divulgación pueden utilizarse como sondas o cebadores en ensayos como los descritos en el presente documento.

Los ácidos nucleicos de la divulgación, tal como los descritos anteriormente, pueden identificarse y aislarse utilizando técnicas de biología molecular estándar bien conocidas por el experto. El ADN puede amplificarse y/o marcarse (p. ej., radiomarcado, marcado con fluorescencia) y utilizarse como sonda para la detección, p. ej., de una biblioteca de ADNc derivada de un organismo. El ADNc puede derivarse del ARNm y puede estar contenido en un vector adecuado. Por ejemplo, se pueden aislar los clones correspondientes, obtener ADN tras la escisión *in vivo* y secuenciar el inserto clonado en una o ambas orientaciones mediante métodos reconocidos en la técnica para identificar el marco de lectura correcto que codifica un polipéptido del peso molecular adecuado. Utilizando estos métodos u otros similares, el polipéptido y el ADN que codifica el polipéptido pueden aislarse, secuenciarse y caracterizarse adicionalmente.

El ácido nucleico puede comprender uno o más polimorfismos, variaciones o mutaciones, p. ej., polimorfismos de un solo nucleótido (SNP), variaciones de un solo nucleótido (SNV), variaciones del número de copias (CNV), p. ej., inserciones, eliminaciones, inversiones y translocaciones. Los ácidos nucleicos pueden comprender análogos, p. ej., fosforotioatos, fosforamidatos, fosfonato de metilo, fosfonatos de metilo quirales, ribonucleótidos 2-0-metilo, o ácidos nucleicos modificados, p. ej., residuos o enlaces de cadenas modificadas, o ácidos nucleicos combinados con carbohidratos, lípidos, polipéptidos u otros materiales, o ácidos nucleicos peptídicos (PNA), p. ej., cromatina, ribosomas y transcriptosomas. Los ácidos nucleicos pueden comprender ácidos nucleicos en diversas estructuras, p. ej., ADN A, ADN B, ADN en forma de Z, ARNip, ARNt y ribozimas. El ácido nucleico puede ser natural o no naturalmente polimórfico, p. ej., tener una o más diferencias de secuencia, p. ej., adiciones, supresiones y/o sustituciones, en comparación con una secuencia de referencia. Una secuencia de referencia puede basarse en información disponible públicamente, p. ej., la U.C. Santa Cruz Human Genome Browser Gateway (genome.ucsc.edu/cgi-bin/hgGateway) o el sitio web del NCBI (www.ncbi.nlm.nih.gov). Un practicante de la presente divulgación puede determinar una secuencia de referencia utilizando métodos bien conocidos en la técnica, p. ej., secuenciando un ácido nucleico de referencia.

Una sonda puede hibridar con un alelo, SNP, SNV, o CNV como se describe en el presente documento. La sonda puede unirse a otra secuencia marcadora asociada a PML como se describe en el presente documento.

Un experto en la materia sabría cómo diseñar una sonda de modo que la hibridación específica de secuencia pueda ocurrir sólo si un alelo particular está presente en una secuencia genómica de una muestra de ácido nucleico de prueba. La divulgación también puede reducirse a la práctica utilizando cualquier método de genotipado conveniente, incluidas las tecnologías y métodos disponibles en el mercado para el genotipado de variaciones genéticas concretas.

También pueden usarse sondas de control, p. ej., una sonda que se une a una secuencia menos variable, p. ej., un ADN repetitivo asociado a un centrómero de un cromosoma, puede usarse como control. Las sondas pueden obtenerse de fuentes comerciales. Las sondas pueden sintetizarse, p. ej., químicamente o *in vitro*, o fabricarse a partir de ADN cromosómico o genómico mediante técnicas estándar. Las fuentes de ADN que pueden utilizarse incluyen el ADN genómico, secuencias de ADN clonadas, híbridos de células somáticas que contienen un cromosoma humano, o parte de él, junto con el complemento cromosómico normal del huésped, y cromosomas purificados por citometría de flujo o microdisección. La región de interés puede aislarse mediante clonación o mediante amplificación específica del sitio utilizando PCR.

Uno o más ácidos nucleicos, p. ej., una sonda o cebador, también pueden ser marcados, p. ej., por marcaje directo, para comprender un marcador detectable. Un marcador detectable puede comprender cualquier marcador capaz de detección mediante un proceso físico, químico o biológico, p. ej., un marcador radiactivo, tal como ^{32}P o ^{3}H, un marcador fluorescente, tal como FITC, un marcador cromóforo, un marcador de ligando de afinidad, un marcador enzimático, tal como fosfatasa alcalina, peroxidasa de rábano picante o galactosidasa I2, un marcador de cofactor enzimático, un marcador de conjugado de hapteno, tal como digoxigenina o dinitrofenilo, un marcador generador de señal Raman, un marcador magnético, un marcador de espín, un marcador de epítopo, tal como el epítopo FLAG o HA, un marcador luminiscente, un marcador de átomo pesado, un marcador de nanopartícula, un marcador electroquímico, un marcador de dispersión de luz, un marcador de concha esférica, un marcador de nanocristal semiconductor, tal como los puntos cuánticos (descritos en la patente de los Estados Unidos No. 6,207,392), y sondas marcadoras con cualquier otro marcador generador de señal conocida por los expertos en la materia, en las que un marcador puede permitir que la sonda se visualice con o sin una molécula de detección secundaria. Un nucleótido puede incorporarse directamente a una sonda con técnicas estándar, p. ej., traducción de muesca, cebado aleatorio y marcaje por PCR. Una "señal", tal como se utiliza en el presente documento, incluye una señal detectable y medible por medios apropiados, incluyendo, fluorescencia, radiactividad, quimioluminiscencia y similares.

Ejemplos no limitativos de fracciones de marcaje útiles para la detección incluyen, sin limitación, enzimas adecuadas como peroxidasa de rábano picante, fosfatasa alcalina, beta-galactosidasa o acetilcolinesterasa; miembros de un par de unión que son capaces de formar complejos como estreptavidina/biotina, avidina/biotina o un complejo antígeno/anticuerpo que incluye, p. ej., IgG de conejo e IgG anti-conejo; fluoróforos como umbeliferona, fluoresceína, isotiocianato de fluoresceína, rodamina, tetrametil rodamina, eosina, proteína verde fluorescente, eritrosina, cumarina, metil cumarina, pireno, verde malaquita, estilbeno, amarillo lucifer, Cascade Blue, Texas Red, fluoresceína diclorotriazinilamina, cloruro de dansilo, ficoeritrina, complejos fluorescentes de lantánidos como los que incluyen el europio y el terbio, miembros de la familia de los colorantes de cianina, tal como Cy3 y Cy5, balizas moleculares y sus derivados fluorescentes, así como otros conocidos en la técnica descritos, p. ej., en Principles of Fluorescence Spectroscopy, Joseph R. Lakowicz (Editor), Plenum Pub Corp, 2ª edición (julio de 1999) y la 6ª edición del Molecular Probes Handbook de Richard P. Hoagland; un material luminiscente, tal como el luminol; materiales de dispersión de la luz o resonantes de plasmón, tal como partículas de oro o plata o puntos cuánticos; o material radiactivo, tal como ^{14}C, ^{123}I, ^{124}I, ^{125}I, Tc99m, ^{32}P, ^{33}P, ^{35}S o ^{3}H.

También pueden utilizarse otros marcadores en los métodos de la presente divulgación, p. ej., marcadores de cadena. Los marcadores de cadena comprenden tintes de ácidos nucleicos que se unen a los ácidos nucleicos

de forma independiente de la secuencia. Entre los ejemplos no limitantes se incluyen colorantes intercalantes como las fenantridinas y las acridinas (p. ej., bromuro de etidio, yoduro de propidio, yoduro de hexidio, dihidroetidio, homodímero-1 y -2 de etidio, monoazida de etidio y ACMA); algunos aglutinantes de surco menor como los indoles y los imidazoles (p. ej., Hoechst 33258, Hoechst 33342, Hoechst 34580 y DAPI); y diversas tinciones de ácidos nucleicos como naranja de acridina (también capaz de intercalarse), 7-AAD, actinomicina D, LDS751 e hidroxiestilbamidina. Todas las tinciones de ácidos nucleicos mencionadas están disponibles comercialmente a través de proveedores tales como Molecular Probes, Inc. Otros ejemplos de tinciones de ácidos nucleicos incluyen los siguientes colorantes de Molecular Probes: colorantes de cianina como SYTOX Blue, SYTOX Green, SYTOX Orange, POPO-1, POPO-3, YOYO-1, YOYO-3, TOTO-1, TOTO-3, JOJO-1, LOLO-1, BOBO-1, BOBO-3, PO-PRO-1, PO-PRO-3, BO-PRO-1, BO-PRO-3, TO-PRO-1, TO-PRO-3, TO-PRO-5, JO-PRO-1, LO-PRO-1, YO-PRO-1, YO-PRO-3, PicoGreen, OliGreen, RiboGreen, SYBR Gold, SYBR Green I, SYBR Green II, SYBR DX, SYTO-40, -41, -42, -43, -44, -45 (azul), SYTO-13, -16, -24, -21, -23, -12, -11, -20, -22, -15, - 14, -25 (verde), SYTO-81, -80, -82, -83, -84, -85 (naranja), SYTO-64, -17, -59, -61, -62, -60, -63 (rojo).

Pueden elegirse fluoróforos de diferentes colores, p. ej., ácido 7-amino-4-metilcumarina-3-acético (AMCA), 5-(y -6)-carboxi-X-rodamina, lisamina-rodamina B, 5-(y -6)-carboxifluoresceína, fluoresceína-5-isotiocianato (FITC), ácido 7-dietilaminocumarina-3-carboxílico, tetrametilrodamina-5-(y -6)-isotiocianato, 5-(y -6)-carboxitetrametilrodamina, ácido 7-hidroxicumarina-3-carboxílico, ácido 6-[fluoresceína 5-(y -6)-carboxamido]hexanoico, ácido N-(4,4-difluoro-5,7-dimetil-4-bora-3a,4a diaza-3-indacenopropiónico, eosina-5-isotiocianato, eritrosina-5-isotiocianato, TRITC, rodamina, tetrametilrodamina, R-ficoeritrina, Cy-3, Cy-5, Cy-7, Texas Red, Phar-Red, aloficocianina (APC) y acetilazida azul CASCADE™, de forma que cada sonda de un conjunto o no pueda visualizarse claramente. Las sondas marcadas con fluorescencia pueden visualizarse con un microscopio de fluorescencia y un filtro adecuado para cada fluoróforo, o utilizando conjuntos de filtros de paso de banda doble o triple para observar múltiples fluoróforos. Pueden utilizarse técnicas como la citometría de flujo para examinar el patrón de hibridación de las sondas.

Las sondas pueden marcarse indirectamente, p. ej., con biotina o digoxigenina, o marcarse con isótopos radiactivos como ^{32}P y/o ^{3}H. Como ejemplo no limitativo, una sonda marcada indirectamente con biotina puede detectarse mediante avidina conjugada con un marcador detectable. Por ejemplo, la avidina puede conjugarse con un marcador enzimático como la fosfatasa alcalina o la peroxidasa de rábano picante. Los marcadores enzimáticos pueden detectarse mediante reacciones colorimétricas utilizando un sustrato y/o un catalizador para la enzima. Pueden utilizarse catalizadores de la fosfatasa alcalina, p. ej., 5-bromo-4-cloro-3-indolilfosfato y azul nitro de tetrazolio. Puede utilizarse un catalizador para la peroxidasa de rábano picante, p. ej., diaminobenzoato.

Uno o más genes divulgados en el presente documento pueden estar en condiciones o vías moleculares relacionadas con diversos aspectos de la función inmune, incluyendo, pero no limitado a, la respuesta de interferón tipo I (p. ej., PMID 26052098), la vía del receptor de células B (p. ej., Wikipathways WP23; PMID 22566564), la vía de señalización RANKL/RANK (p. ej., Wikipathways WP2018), la vía de señalización TCR (p. ej., Wikipathways WP69), la vía de señalización NF-kB (p. ej., PMID 28362430), la vía JAK-STAT(p. ej., PMID 28255960), biología de modificaciones postraduccionales como la ubiquitinación vía LUBAC (p. ej., PMID 23104095, 24958845, 25086647, 26085218, 26111062, 26525107, 26848516, 26877205, 27178468, 27786304, 27892465), síndrome de Aicardi-Goutieres (p. ej., PMID 26052098), eosinofilia (por ejemplo, PMID 27222657), neutropenia congénita (p. ej., PMID 24753205), defectos del receptor de células T (p. ej., PMID 25452106, 25636200, 26246585, 26379669, 26453379, 28400082) y defectos de autofagia (p. ej., 19229298, 22984599, 23222957, 26917586, 26953272, 27588602). Uno o más genes divulgados en el presente documento pueden estar relacionados con la biología del virus de JC (p. ej., PMID 15327898, 19282432, 19903823, 22984599, 25910481). Uno o más genes divulgados en el presente documento pueden ser genes de respuesta inmune antiviral.

Tabla 27: Vías ejemplares y biología de los genes de riesgo de PML (panel de 96 genes)*

Genes	Enfermedad auto-inflamatoria	Defectos de autofagia	Defectos de las células B	Vía del receptor de células B	Vía de la des-ubiquitinasa	Inmunodeficiencia asociada a la eosinofilia	Biología del virus de JC	Osteopetrosis	Señalización de PI3K	Vía de RANKL/ RANK	Defectos de las células T	Vía del receptor de células T	Señalización de TLR	Vía de interferencia de tipo I
AP3B1								19782649 24753205						
APOL1							27042682							
ASH1L														
ATM			23785059			27222657	19903823 25910461 27042682							25882705
ATR							19903823 25910461							24799566
BLM														
CARD11			23785059	23785059, WP23		27222657						WP69	25936963	
CERNA1B														
CHD7						27222657								
CLCN7								23877423 24753205						
DCLRE1C														
DDX58														25145756, 26052068, 25763980, 27621552
DOCK8			23785059			27222657					25379669			
EGR1														
EPG5		23222557, 25917586												
ETF1														
FPR2														
GATA2			23785059											

Genes	Enfermedad auto-inflamatoria	Defectos de autofagia	Defectos de las células B	Vía del receptor de células B	Vía de la desubiquitinasa	Inmunodeficiencia asociada a la eosinofilia	Biología del virus de JC	Osteopetrosis	Señalización de PI3K	Vía de RANKL/ RANK	Defectos de las células T	Vía del receptor de células T	Señalización de TLR	Vía de interferencia de tipo I
GFI1								19782549, 24753205						
HIVEP1														
HIVEP2														
HTR2A							27042682							
IL22														
IRH1														26052098, 27821552
IFNGR 2														
IFNLR1														
IGLL1			23765059	23765059										
IKBKB	28466820		23765059	23765059, WP23,	26877205, 28362430			21079651		WP2018		WP69	20404851, 25930993	17047224, 25145756
IL17F														
IL1B					27892465		15327899					WP69	25930993	25145756, 26763950
IL21R			23765059											
IRAK4			23765059	23765059	28362430								20404851, 25930993	25737587
ITSN2														
JUN				WP23			27042682							25888367
KAT6B														
KCTD7														
LIG4						27222657								
LPBA		26707784	23765059								26707784			
MALL														
MAPK8							27042682			WP2018		WP69	25930993	
MAVS														19120474, 22626036, 22801541,

Genes	Enfermedad auto-inflamatoria	Defectos de autofagia	Defectos de las células B	Vía del receptor de células B	Vía de la des-ubiquitinasa	Inmunodeficiencia asociada a la eosinofilia	Biología del virus de JC	Osteopetrosis	Señalización de PI3K	Vía de RANKL/ RANK	Defectos de las células T	Vía del receptor de células T	Señalización de TLR	Vía de interferencia de tipo I
														25145756, 28763980
MCEE														
MKL1														26098208, 25098211
MSD88			23765059	WIP23	28362430			21079651	27616589	WP2018		WP69	20404851, 25930893	18474425, 19573338, 25145756
NBN														
NFKB1	28468620		23765059	WIP23	29877205, 28362430			21079651	27616589	WP2018		WP69	20404851, 25930893	17047224, 25145756
														28763980
NOD2	26421071	26953272			28362430	27222657								28763980
NRIP1														
PIAS1														24036127
PIAS2														21156324, 24036127
PIK3C D									27616589					28231019
PIK3C D-AS1									27616589					
PIK3R1			23765059	WIP23					27616589	WP2018		WP69		28198378
PKHD1														
PLCG2			23765059	WIP23					27616589	WP2018	28479669	WP69	25930893	
PNPT1														
POLA1														27019227, 27821552
POLE			23765059											
PRF1														
PRKCB				WIP23										

Genes	Enfermedad auto-inflamatoria	Defectos de autofagia	Defectos de las células B	Vía del receptor de células B	Vía de la desubiquitinasa	Inmunodeficiencia asociada a la eosinofilia	Biología del virus de JC	Osteopetrosis	Señalización de PI3K	Vía de RANKL/ RANK	Defectos de las células T	Vía del receptor de células T	Señalización de TLR	Vía de interferencia de tipo I
PRKCD				WP23								WP69		11839736
PRKCH														
PRKDC														23251783
PSTPIP1	28421071													
PTEN									27616589					
PTPRC				WP23					27616589					19573588, 25869642
RABGEF1														
RAD51							27042582							
RAG1			23765059			27222657			27616589					
RAG2			23765059			27222657			27616589					
RIPK1	28466630				26877205, 27892465, 28362430								20404851	25145756
RIPK3					26877205, 27892465, 28362430									25145756
RNF168														
RTEL1														
SHARPIN	28466630		23765059		26877205, 27892465, 28362430				25930993				20404851	22901541, 25145756
SKIV2L														25064072, 27821552
SMAD4							27042582							
STIM1			23765059											22144678
STIM2														
STXBP2														
TAP2														

Genes	Enfermedad auto-inflamatoria	Defectos de autofagia	Defectos de las células B	Vía del receptor de células B	Vía de la des-ubiquitinasa	Inmunodeficiencia asociada a la esplicilla	Biología del virus de JC	Osteopetrosis	Señalización de PI3K	Vía de RANKL/ RANK	Defectos de las células T	Vía del receptor de células T	Señalización de TLR	Vía de interferencia de tipo I
TBK1													25930993	16573338 22626038 25145756 26763980 28049150
TCIRG1								23877423 24763206						
TICAM1					28362430								20404851 25930993	19120474 25145756 28049150
TLR3	28469520				28362430								20404851 25930993	19120474 25145756 28049150
TLR4	28469520				28362430								20404851 25930993	25145756
TNFRS F11A	28421071				28362430			21079851 23877423		WP2018			25930993	21527253 25497789
TNFRS F13B			23765069										25930993	
TNFRS F8														
TP53														
TRAF3								21079851		WP2018			25930993	22901541 25723057 26763980
TRAF6 1								25992615					16221974	18949341
TPPMC														
VPS45														
WEE1							27042683							
ZAP70			23765069			27222357			27616589			WP99		

* Se indican los números PMID de las referencias de Pubmed curadas o el número de identificación de Wikipathway

La Tabla 27 contiene un conjunto de vías y biología ejemplares para genes de riesgo de PML basados en el panel de 96 genes listado en la Tabla 19. Los genes divulgados en el presente documento, tal como los genes del panel de 96 genes, pueden agruparse en función de la vía o los procesos biológicos en los que participan.

Métodos de detección

La presente invención consiste en determinar la ausencia o presencia de una variación genómica seleccionada de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4: 103522150, G>A. A continuación se ofrece una explicación general de los métodos de detección que pueden utilizarse para determinar la presencia o ausencia de las variaciones genómicas de las reivindicaciones, y se incluyen ejemplos que no forman parte de la materia reivindicada en el presente documento, en la medida en que se refieren a variaciones distintas de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. Como se usa en el presente documento, la detección de un sujeto comprende el diagnóstico o determinación, teranosis o determinación de la susceptibilidad a desarrollar (pronosticar) una enfermedad, p. ej., PML. En particular, la divulgación proporciona un método para determinar una presencia de, o una susceptibilidad a, PML, mediante la detección de al menos una variación genética en una muestra de un sujeto como se describe en el presente documento. La detección de determinados alelos, marcadores, variaciones o haplotipos puede ser indicativa de la presencia o susceptibilidad a una enfermedad (p. ej., PML).

Si bien existen medios para la detección de PML mediante una prueba de anticuerpos contra el JCV, el riesgo de PML no se evalúa adecuadamente mediante la prueba de anticuerpos contra el JCV por sí sola. Por lo tanto, existe la necesidad de una prueba de detección mejorada para evaluar el riesgo de desarrollar PML. En el presente documento se describen métodos de detección de un individuo para detectar el riesgo de desarrollar PML, incluyendo, entre otros, la determinación de la identidad y localización de variaciones genéticas, tal como variaciones en la secuencia de nucleótidos y en el número de copias, y la presencia o ausencia de alelos o genotipos en una o más muestras de uno o más sujetos utilizando cualquiera de los métodos descritos en el presente documento. La determinación de una asociación a padecer o desarrollar PML puede realizarse detectando variaciones particulares que aparecen con mayor frecuencia en los sujetos de prueba en comparación con los sujetos de referencia y analizando las vías moleculares y fisiológicas a las que estas variaciones pueden afectar.

Dentro de cualquier población dada, puede haber una susceptibilidad absoluta de desarrollar una enfermedad o rasgo, definida como la probabilidad de que una persona desarrolle la enfermedad o rasgo específico durante un periodo de tiempo específico. La susceptibilidad (p. ej., estar en situación de riesgo) suele medirse observando a un gran número de personas, en lugar de a un individuo concreto. Como se describe en el presente documento, ciertas variaciones en el número de copias (variaciones genéticas) y/o variaciones de nucleótido único se consideran útiles para la evaluación de la susceptibilidad a la PML. La evaluación de la susceptibilidad puede implicar la detección de determinadas variaciones genéticas en el genoma de los individuos sometidos a evaluación. Variaciones genéticas particulares se encuentran con mayor frecuencia en individuos con PML, que en individuos sin PML. Por lo tanto, estas variaciones genéticas tienen valor predictivo para detectar la PML, o una susceptibilidad a la PML, en un individuo. Sin pretender estar limitados por la teoría, se cree que las variaciones genéticas descritas en el presente documento que se asocian con la susceptibilidad de la PML representan variantes funcionales que predisponen a la enfermedad. Una variación genética puede conferir una susceptibilidad a la enfermedad, p. ej., los vehículos de la variación genética tienen un riesgo diferente de padecer la enfermedad que los no vehículos. La presencia de una variación genética puede ser indicativa de una mayor susceptibilidad a la PML.

La detección puede realizarse utilizando cualquiera de los métodos divulgados, solos o en combinación. La detección puede realizarse mediante la reacción en cadena de la polimerasa (PCR). La detección puede realizarse mediante hibridación genómica comparativa de arreglos (aCGH) para detectar CNV. La detección puede realizarse mediante secuenciación del exoma para detectar SNV, indel y, en algunos casos, CNV utilizando algoritmos de análisis adecuados. La detección puede realizarse utilizando métodos de secuenciación del genoma completo de alto rendimiento (también conocidos como de próxima generación) y algoritmos adecuados para detectar todas o casi todas las variaciones genéticas presentes en una muestra de ADN genómico. La información sobre la variación genética, tal y como se relaciona con la presente divulgación, puede utilizarse junto con cualquiera de las pruebas de detección sintomáticas mencionadas anteriormente para detección de PML en un sujeto, p. ej., utilizando una combinación de aCGH y/o secuenciación con una prueba de detección del JCV, tal como la prueba de anticuerpos contra el JCV, la prueba de CD62L o la prueba de banda oligoclonal de IgM del CSF. La L-selectina (CD62L) expresada por las células $TCD3^{+}CD4^{+}$ en, p. ej., células mononucleares de sangre periférica (PBMC) crioconservadas, puede ser un biomarcador para la detección del JCV. La expresión de CD62L puede correlacionarse con el riesgo de PML.

La información de cualquiera de los métodos de detección anteriores (p. ej., síntomas específicos, matriz de puntuación o datos de variación genética) puede utilizarse para definir a un sujeto como sujeto de prueba o sujeto de referencia. La información de cualquiera de los métodos de detección anteriores puede utilizarse para asociar un sujeto con una población de prueba o de referencia, p. ej., un sujeto de una población.

Una asociación con PML puede ser determinada por la probabilidad estadística de la presencia de una variación genética en un sujeto con PML, p. ej., un individuo no relacionado o un pariente de primer o segundo grado del sujeto. Una asociación con la PML puede decidirse determinando la probabilidad estadística de la ausencia de una variación genética en un sujeto de referencia no afectado, p. ej., un individuo no emparentado o un pariente de primer o segundo grado del sujeto. Los métodos descritos en el presente documento pueden incluir la obtención y el análisis de una muestra de ácido nucleico de uno o más sujetos de referencia adecuados.

En el presente contexto, el término detección comprende el diagnóstico, el pronóstico y la teranosis. La detección puede referirse a cualquier método de detección disponible, incluidos los mencionados en el presente documento. Como se utiliza en el presente documento, la susceptibilidad puede ser la propensión de un sujeto hacia el desarrollo de la PML, o hacia ser menos capaz de resistir la PML que uno o más sujetos de control. La susceptibilidad puede abarcar el aumento de la susceptibilidad. Por ejemplo, variaciones particulares de ácido nucleico de la divulgación como se describe en el presente documento pueden ser características de una mayor susceptibilidad a PML. Variaciones particulares de ácido nucleico pueden conferir susceptibilidad disminuida, p. ej., variaciones nucleicas particulares de la divulgación como se describe en el presente documento pueden ser características de susceptibilidad disminuida al desarrollo de PML.

Como se describe en el presente documento, una variación genética predictiva de susceptibilidad a o presencia de PML puede ser aquella en la que la variación genética particular está presente con mayor frecuencia en un grupo de sujetos con la enfermedad (afectados), en comparación con la frecuencia de su presencia en un grupo de referencia (control), de tal manera que la presencia de la variación genética es indicativa de susceptibilidad a o presencia de PML. El grupo de referencia puede ser una muestra de ácido nucleico de la población, p. ej., una muestra aleatoria de ácido nucleico de la población general o una mezcla de dos o más muestras de ácido nucleico de una población. Los controles libres de enfermedad pueden caracterizarse por la ausencia de uno o más síntomas específicos asociados a la enfermedad, p. ej., individuos que no han experimentado síntomas asociados a la PML. El grupo de control libre de enfermedad puede caracterizarse por la ausencia de uno o más factores de riesgo específicos de la enfermedad, p. ej., al menos un factor de riesgo genético y/o ambiental. Se puede hacer referencia a una secuencia de referencia para un lugar concreto de variación genética. Un alelo de referencia puede ser un alelo de tipo silvestre y puede elegirse como el primer alelo secuenciado o como el alelo de un individuo de control. Uno o más sujetos de referencia pueden ser emparejados característicamente con uno o más sujetos afectados, p. ej., con edad, sexo o etnia emparejados.

Un experto en la materia puede apreciar que, para variaciones genéticas con dos o más alelos presentes en la población estudiada, y en las que un alelo puede encontrarse con mayor frecuencia en un grupo de individuos con PML en la población, en comparación con los controles, el otro alelo del marcador puede encontrarse con menor frecuencia en el grupo de individuos con el rasgo o enfermedad, en comparación con los controles. En tal caso, un alelo del marcador, p. ej., el alelo encontrado con mayor frecuencia en individuos con PML, puede ser el alelo de riesgo, mientras que el otro alelo o alelos pueden ser un alelo neutro o protector.

Una variante genética asociada con PML puede ser usada para predecir la susceptibilidad de la enfermedad para un genotipo dado. Para cualquier variación genética, puede haber uno o más genotipos posibles, p. ej., homocigoto para la variante de riesgo (p. ej., en trastornos autosómicos recesivos), heterocigoto y no vehículo de la variante de riesgo. Los trastornos autosómicos recesivos también pueden ser el resultado de dos variantes genéticas distintas que afectan al mismo gen, de forma que el individuo es un heterocigoto compuesto (p. ej., el alelo materno contiene una mutación diferente del alelo paterno). La heterocigosidad compuesta puede ser el resultado de dos SNV diferentes, dos CNV diferentes, una SNV y una CNV, o cualquier combinación de dos variantes genéticas diferentes pero cada una presente en un alelo diferente del gen. En el caso de los genes ligados al cromosoma X, los varones que poseen una copia de un gen con una variante pueden verse afectados, mientras que las mujeres portadoras, que también poseen un gen de tipo silvestre, pueden no verse afectadas. La susceptibilidad asociada a variantes en múltiples loci puede utilizarse para estimar la susceptibilidad global. Para múltiples variantes genéticas, puede haber k (k = 3^n * 2^P) genotipos posibles; donde n puede ser el número de loci autosómicos y p puede ser el número de loci gonosómicos (cromosomas sexuales). Los cálculos de evaluación de la susceptibilidad global pueden suponer que las susceptibilidades relativas de las distintas variantes genéticas se multiplican; p. ej., la susceptibilidad global asociada a una determinada combinación de genotipos puede ser el producto de los valores de susceptibilidad del genotipo en cada locus. Si la susceptibilidad presentada es la susceptibilidad relativa para una persona, o un genotipo específico para una persona, en comparación con una población de referencia, entonces la susceptibilidad combinada puede ser el producto de los valores de susceptibilidad específicos del locus y puede corresponder a una estimación de susceptibilidad global en comparación con una población. Si la susceptibilidad de una persona se basa en una comparación con los no vehículos del alelo de riesgo, entonces la susceptibilidad combinada puede corresponder a una estimación que compara a la persona con una combinación dada de genotipos en todos los loci con un grupo de individuos que no son vehículos de variantes de riesgo en ninguno

de esos loci. El grupo de no vehículos de cualquier variante de riesgo puede tener la susceptibilidad estimada más baja y puede tener una susceptibilidad combinada, comparada consigo mismo, p. ej., con los no vehículos, de 1.0, pero puede tener una susceptibilidad global, comparada con la población, de menos de 1.0.

El riesgo global para múltiples variantes de riesgo puede realizarse utilizando una metodología estándar. Las variaciones genéticas descritas en el presente documento pueden constituir la base de un análisis de riesgo que combine otras variaciones genéticas de las que se sabe que aumentan el riesgo de PML, u otras variantes genéticas de riesgo de PML. Se puede utilizar una pluralidad de variantes (variaciones genéticas, alelos variantes y/o haplotipos) para la evaluación global del riesgo. Estas variantes pueden seleccionarse de entre las variaciones genéticas divulgadas en el presente documento. Las variantes de la presente divulgación pueden utilizarse en combinación con otras variantes que se sabe que son útiles para detectar una susceptibilidad a PML. El estado genotípico de una pluralidad de variaciones genéticas, marcadores y/o haplotipos puede determinarse en un individuo, y el estado del individuo puede compararse con la frecuencia poblacional de las variantes asociadas, o con la frecuencia de las variantes en sujetos clínicamente sanos, tal como sujetos emparejados por edad y sexo.

Métodos tales como el uso de algoritmos y software disponibles pueden ser utilizados para identificar, o llamar, variaciones genéticas significativas, incluyendo, pero no limitado a, algoritmos de DNA Analytics o DNAcopy, iPattern y/o QuantiSNP. Para determinar las pérdidas y las ganancias puede utilizarse un valor de umbral de relación logarítmica para determinar pérdidas y ganancias. Por ejemplo, utilizando DNA Analytics, se puede utilizar un corte de relación $\log_2$ de ≥0.5 y <0.5 para clasificar las ganancias y pérdidas de CNV respectivamente. Por ejemplo, utilizando DNA Analytics, se puede utilizar un corte de relación $\log_2$ de ≥0.25 y <0.25 para clasificar las ganancias y pérdidas de CNV respectivamente. Como ejemplo adicional, utilizando DNAcopy, se puede utilizar un corte de relación $\log_2$ de ≥0.35 y <0.35 para clasificar las ganancias y pérdidas de CNV respectivamente. Por ejemplo, se puede utilizar un algoritmo de módulo de detección de aberraciones 2 (ADM2), tal como aquel de DNA Analytics 4.0.85, para identificar, o llamarse, variaciones genéticas significativas. Se pueden utilizar dos o más algoritmos para identificar, o llamarse, variaciones genéticas significativas. Por ejemplo, pueden utilizarse 2, 3, 4, 5, 6, 7, 8, 9, o 10 o más algoritmos para identificar, o llamarse, variaciones genéticas significativas. La relación log2 de una o más sondas individuales de una micromatriz puede utilizarse para identificar variaciones genéticas significativas, tal como la presencia de regiones homocigóticamente eliminadas en el genoma de un sujeto. Las variaciones genéticas significativas pueden ser CNV.

Las CNV detectadas por dos o más algoritmos pueden definirse como estrictas y pueden utilizarse para análisis posteriores. La información y las llamadas de dos o más de los métodos descritos en el presente documento pueden compararse entre sí para identificar variaciones genéticas significativas con mayor o menor rigor. Por ejemplo, las llamadas de CNV generadas por dos o más de los algoritmos DNA Analytics, módulo de detección de aberración 2 (ADM2) y DNAcopy pueden definirse como CNV estrictas. Las variaciones genéticas significativas o estrictas se pueden marcar como identificados o denominados si se puede encontrar que tienen un solapamiento recíproco mínimo con una variación genética detectada por una o más plataformas y/o métodos descritos en el presente documento. Por ejemplo, se puede utilizar un solapamiento recíproco mínimo del 50 % para marcar las CNV como identificadas o denominadas. Por ejemplo, las variaciones genéticas significativas o estrictas se pueden marcar como identificadas o llamadas si se puede encontrar que tienen un solapamiento recíproco de más de aproximadamente 50 %, 55 % 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, o 95 %, 99 %, o igual a 100 %, a una variación genética detectada por una o más plataformas y/o métodos descritos en el presente documento. Por ejemplo, las variaciones genéticas significativas o estrictas se pueden marcar como identificadas o denominadas si se puede encontrar que tienen un solapamiento recíproco de más de aproximadamente el 50 % de solapamiento recíproco con una variación genética detectada por una o más plataformas y/o métodos descritos en el presente documento. Las variaciones genéticas pueden detectarse a partir de los valores de relación log2 calculados para sondas individuales presentes en un microarreglo aCGH mediante una comparación estadística del valor de relación log2 de la sonda en una cohorte de sujetos con PML con el valor de relación log2 de la sonda en una cohorte de sujetos sin PML.

Se puede utilizar un valor umbral de relación logarítmica para determinar las pérdidas y ganancias. Un valor de relación logarítmica puede ser cualquier valor de relación logarítmica; p. ej., un valor de relación logarítmica puede ser una relación log2 o una relación log10. Un segmento de CNV cuya mediana de la relación log2 sea inferior o igual a un valor umbral de la relación log2 puede clasificarse como pérdida. Por ejemplo, cualquier segmento cuya mediana de la relación log2 sea inferior o igual a -0.1, -0.11, -0.12, -0.13, -0.14, -0.15, -0.16, -0.17, -0.18, -0.19, -0,2, -0,21, -0,22, -0,23, -0,24, -0,25, -0,26, -0.27, -0.28, -0.29, -0.3, -0.31, -0.32, -0.33, -0.34, -0.35, -0.36, -0.37, -0.38, -0.39, -0.4, -0.41, -0.42, -0.43, -0.44, -0.45, -0.46, -0.47, - 0.48, -0.49, -0.5, -0.55, -0.6, -0.65, -0.7, -0.75, -0.8, -0.85, -0.9, -0.95, -1, -1.1, -1.2, -1.3, -1.4, -1.5, -1.6, - 1.7, -1.8, -1.9, -2, -2.1, -2.2, -2.3, -2.4, -2.5, -2.6, -2.7, -2.8, -2.9, -3, -3.1, -3.2, -3.3, -3.4, -3.5, -3.6, -3.7, - 3.8, -3.9, -4, -4.1, -4.2, -4.3, -4.4, -4.5, -4.6, -4.7, -4.8, -4.9, -5, -5.5, -6, -6.5, -7, -7.5, -8, -8.5, -9, -9.5, -10, -11, -12, -13, -14, -15, -16, -17, -18, -19, -20 o menos, puede clasificarse como pérdida.

Un algoritmo puede utilizarse para llamar o identificar variaciones genéticas significativas, en las que cualquier segmento cuya mediana de la relación log2 fuera menor o igual a -0.1, -0.11, -0.12, -0.13, -0.14, -0.15, -0.16, -

0.17, -0.18, -0.19, -0.2, -0.21, -0.22, -0.23, -0.24, -0.25, -0.26, -0.27, -0.28, -0.29, -0.3, -0.31, -0.32, -0.33, -0.34, -0.35, -0.36, -0.37, -0.38, -0.39, -0.4, -0.41, -0.42, -0.43, -0.44, -0.45, -0.46, -0.47, -0.48, -0.49, -0.5, - 0.55, -0.6, -0.65, -0.7, -0.75, -0.8, -0.85, -0.9, -0.95, -1, -1.1, -1.2, -1.3, -1.4, -1.5, -1.6, -1.7, -1.8, -1.9, -2, -2.1, -2.2, -2.3, -2.4, -2.5, -2.6, -2.7, -2.8, -2.9, -3, -3.1, -3.2, -3.3, -3.4, -3.5, -3.6, -3.7, -3.8, -3.9, -4, -4.1, -4.2, -4.3, -4.4, -4.5, -4.6, -4.7, -4.8, -4.9, -5, -5.5, -6, -6.5, -7, -7.5, -8, -8.5, -9, -9.5, -10, -11, -12, -13, -14, -15, -16, -17, -18, -19, -20 o menos, puede clasificarse como pérdida. Por ejemplo, cualquier segmento de CNV cuya mediana de la relación log2 sea inferior a -0.35, de acuerdo con lo determinado por DNAcopy, puede clasificarse como pérdida. Por ejemplo, las pérdidas pueden determinarse en función de una relación log2 umbral, que puede fijarse en -0.35. Las pérdidas pueden determinarse en función de una relación log2 umbral, que puede fijarse en -0.5.

Pueden utilizarse dos algoritmos para llamar o identificar variaciones genéticas significativas, en las que cualquier segmento cuya mediana de la relación log2 sea inferior o igual a -0.1, -0.11, -0.12, -0.13, -0.14, -0.15, -0.16, -0.17, -0.18, - 0.19, -0.2, -0.21, -0.22, -0.23, -0.24, -0.25, -0.26, -0.27, -0.28, -0.29, -0.3, -0.31, -0.32, -0.33, -0.34, -0.35, -0.36, -0.37, -0.38, -0.39, -0.4, -0.41, -0.42, -0.43, -0.44, -0.45, -0.46, -0.47, -0.48, -0.49, -0.5, -0.55, -0.6, -0.65, -0.7, -0.75, -0.8, -0.85, -0.9, -0.95, -1, -1.1, -1.2, -1.3, -1.4, -1.5, -1.6, -1.7, -1.8, -1.9, -2, -2.1, -2.2, - 2.3, -2.4, -2.5, -2.6, -2.7, -2.8, -2.9, -3, -3.1, -3.2, -3.3, -3.4, -3.5, -3.6, -3.7, -3.8, -3.9, -4, -4.1, -4.2, -4.3, - 4.4, -4.5, -4.6, -4.7, -4.8, -4.9, -5, -5.5, -6, -6.5, -7, -7.5, -8, -8.5, -9, -9.5, -10, -11, -12, -13, -14, -15, -16, -17, -18, -19, -20 o menos, según lo determinado por un algoritmo, y en el que cualquier segmento cuya mediana de la relación log2 sea menor o igual que -0.1, -0.11, -0.12, -0.13, -0.14, -0.15, -0.16, -0.17, -0.18, -0.19, -0.2, -0.21, - 0.22, -0.23, -0.24, -0.25, -0.26, -0.27, -0.28, -0.29, -0.3, -0.31, -0.32, -0.33, -0.34, -0.35, -0.36, -0.37, -0.38, -0.39, -0.4, -0.41, -0.42, -0.43, -0.44, -0.45, -0.46, -0.47, -0.48, -0.49, -0.5, -0.55, -0.6, -0.65, -0.7, - 0.75, -0.8, -0.85, -0.9, -0.95, -1, -1.1, -1.2, -1.3, -1.4, -1.5, -1.6, -1.7, -1.8, -1.9, -2, -2.1, -2.2, -2.3, -2.4, - 2.5, -2.6, -2.7, -2.8, -2.9, -3, -3.1, -3.2, -3.3, -3.4, -3.5, -3.6, -3.7, -3.8, -3.9, -4, -4.1, -4.2, -4.3, -4.4, -4.5, - 4.6, -4.7, -4.8, -4.9, -5, -5.5, -6, -6.5, -7, -7.5, -8, -8.5, -9, -9.5, -10, -11, -12, -13, -14, -15, -16, -17, -18, -19, -20, o menos, según lo determinado por el otro algoritmo puede clasificarse como pérdida. Por ejemplo, la denominación de CNV puede comprender el uso del algoritmo del módulo de detección de aberraciones 2 (ADM2) y el algoritmo de DNAcopy, en el que las pérdidas pueden determinarse de acuerdo con dos relaciones log2 umbral, en las que la relación log2 del algoritmo del módulo de detección de aberraciones 2 (ADM2) puede ser -0.25 y la relación log2 del algoritmo de DNAcopy puede ser -0.41.

El uso de dos algoritmos para llamar o identificar variaciones genéticas significativas puede ser un método riguroso. El uso de dos algoritmos para llamar o identificar variaciones genéticas significativas puede ser un método más estricto en comparación con el uso de un algoritmo para llamar o identificar variaciones genéticas significativas.

Cualquier segmento de CNV cuya mediana de relación log2 sea mayor que un valor umbral de relación log2 puede clasificarse como ganancia. Por ejemplo, cualquier segmento cuya mediana de la relación log2 sea superior a 0.1, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.2, 0.21, 0.22, 0.23, 0.24, 0.25, 0.26, 0.27, 0.28, 0.29, 0.3, 0.31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4, 0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48, 0.49, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3 o más pueden clasificarse como ganancia.

Puede utilizarse un algoritmo para llamar o identificar variaciones genéticas significativas, en el que cualquier segmento cuya mediana de la relación log2 sea mayor o igual a 0.1, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0,2, 0,21, 0,22, 0,23, 0,24, 0,25, 0,26, 0,27, 0,28, 0,29, 0,3, 0,31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4, 0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48, 0.49, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3 o más pueden clasificarse como ganancia. Por ejemplo, cualquier segmento de CNV cuya mediana de la relación log2 sea superior a 0.35 de acuerdo con la DNAcopy puede clasificarse como ganancia. Por ejemplo, las ganancias pueden determinarse en función de una relación log2 umbral, que puede fijarse en 0.35. Las ganancias pueden determinarse en función de una relación log2 umbral, que puede fijarse en 0.5.

Pueden utilizarse dos algoritmos para llamar o identificar variaciones genéticas significativas, en las que cualquier segmento cuya mediana de la relación log2 sea mayor o igual a 0.1, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0,2, 0,21, 0,22, 0,23, 0,24, 0,25, 0,26, 0,27, 0,28, 0.29, 0.3, 0.31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4, 0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48, 0.49, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3 o más, según lo determinado por un algoritmo, y en el que cualquier segmento cuya mediana de la relación log2 sea mayor o igual que 0.1, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.2, 0.21, 0.22, 0.23, 0.24, 0.25, 0.26, 0.27, 0.28, 0.29, 0.3, 0.31, 0.32, 0.33, 0.34, 0.35, 0.36, 0.37, 0.38, 0.39, 0.4, 0.41, 0.42, 0.43, 0.44, 0.45, 0.46, 0.47, 0.48, 0.49, o 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 3, o más, según lo determinado por el otro algoritmo puede clasificarse como ganancia. Por ejemplo, la llamada de CNV puede comprender el uso del algoritmo del módulo de detección de aberraciones 2 (ADM2) y el algoritmo de DNAcopy, en el que las ganancias pueden determinarse de acuerdo con dos relaciones log2 de umbral, en el que la relación log2 del algoritmo del módulo de detección de aberraciones 2 (ADM2) puede ser 0.25 y la relación log2 del algoritmo de DNAcopy puede ser 0.32.

Cualquier segmento CNV cuyo valor absoluto (mediana de la relación log-ratio/mad) sea inferior a 2 puede ser excluido (no identificado como una variación genética significativa). Por ejemplo, puede excluirse cualquier segmento de CNV cuyo valor absoluto (mediana de la relación log-ratio/mad) sea inferior a 2, 1.9, 1.8, 1.7, 1.6, 1.5, 1.4, 1.3, 1.2, 1.1, 1, 0.9, 0.8, 0.7, 0.6 o 0.5 o menos.

Los análisis multivariantes o los análisis de riesgo conjuntos, incluido el uso de un modelo multiplicativo para la evaluación del riesgo global, pueden utilizarse posteriormente para determinar el riesgo global conferido en función del estado del genotipo en los loci múltiples. El uso de un modelo multiplicativo, p. ej., asumiendo que el riesgo de las variantes de riesgo individuales se multiplica para establecer el efecto global, permite un cálculo directo del riesgo global para múltiples marcadores. El modelo multiplicativo es un modelo parsimonioso que suele ajustarse razonablemente bien a los datos de rasgos complejos. Las desviaciones de la multiplicidad se han descrito raramente en el contexto de variantes comunes para enfermedades comunes, y si se comunican suelen ser sólo sugestivas, ya que pueden requerirse tamaños de muestra muy grandes para poder demostrar interacciones estadísticas entre loci. La evaluación del riesgo basada en dicho análisis puede utilizarse posteriormente en los métodos, usos y kits de la divulgación, tal como se describe en el presente documento.

La importancia del aumento o disminución de la susceptibilidad puede medirse mediante un porcentaje. Un aumento significativo de la susceptibilidad puede medirse como una susceptibilidad relativa de al menos 1.2, incluyendo, pero no limitándose a: al menos 1.3, al menos 1.4, al menos 1.5, al menos 1.6, al menos 1.7, al menos 1.8, al menos 1.9, al menos 2.0, al menos 2.5, al menos 3.0, al menos 4.0, al menos 5.0, al menos 6.0, al menos 7.0, al menos 8.0, al menos 9.0, al menos 10.0 y al menos 15.0. Una susceptibilidad relativa de al menos 2.0, al menos 3.0, al menos 4.0, al menos 5.0, al menos 6.0 o al menos 10.0 es significativa. También se contemplan otros valores para la susceptibilidad significativa, p. ej., al menos 2.5, 3.5, 4.5, 5.5, o cualquier otro valor numérico adecuado, donde los valores también están dentro del alcance de la presente divulgación. Un aumento significativo de la susceptibilidad puede ser de al menos un 20 %, incluyendo, pero sin limitarse a aproximadamente 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 100 %, 150 %, 200 %, 300 %, 400 %, 500 %, 600 %, 700 %, 800 %, 900 %, 1000 % y 1500 %. Un aumento significativo de la susceptibilidad puede ser de al menos el 100 %. Un aumento significativo de la susceptibilidad puede ser de al menos 200 %, al menos 300 %, al menos 400 %, al menos 500 %, al menos 700 %, al menos 800 %, al menos 900 % y al menos 1000 %. También se contemplan otros límites o intervalos que el experto en la materia considere adecuados para caracterizar la divulgación, y que también están dentro del alcance de la presente divulgación. Un aumento significativo de la susceptibilidad puede caracterizarse por un valor p, tal como un valor p inferior a 0.5, inferior a 0.4, inferior a 0.3, inferior a 0.2, inferior a 0.1, inferior a 0.05, inferior a 0.01, inferior a 0.001, inferior a 0.0001, inferior a 0.00001, inferior a 0.000001, inferior a 0.0000001, inferior a 0.00000001 o inferior a 0.000000001.

Un individuo que tiene una susceptibilidad disminuida o la ausencia de presencia de una afección (p. ej., PML) puede ser un individuo en el que se identifica al menos una variación genética que confiere una susceptibilidad disminuida o la ausencia de presencia de la afección. Las variaciones genéticas que confieren una menor susceptibilidad también pueden ser protectoras. En un aspecto, las variaciones genéticas pueden conferir una disminución significativa de la susceptibilidad o la ausencia de presencia de PML.

La disminución significativa de la susceptibilidad puede medirse como una susceptibilidad relativa de menos de 0.9, incluyendo, pero no limitado a menos de 0.9, menos de 0.8, menos de 0.7, menos de 0.6, menos de 0.5, menos de 0.4, menos de 0.3, menos de 0.2 y menos de 0.1. La disminución de la susceptibilidad puede ser de al menos un 20 %, incluyendo, pero no limitándose a al menos un 25 %, al menos un 30 %, al menos un 35 %, al menos un 40 %, al menos un 45 %, al menos un 50 %, al menos un 55 %, al menos un 60 %, al menos un 65 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 95 % y al menos un 98 %. Sin embargo, también se contemplan otros límites o intervalos que el experto en la materia considere adecuados para caracterizar la divulgación, y que también están dentro del alcance de la presente divulgación. Una disminución significativa de la susceptibilidad puede caracterizarse por un valor p, tal como un valor p inferior a 0.05, inferior a 0.01, inferior a 0.001, inferior a 0.0001, inferior a 0.00001, inferior a 0.000001, inferior a 0.0000001, inferior a 0.00000001 o inferior a 0.000000001. Pueden utilizarse otras pruebas de significancia, p. ej., una prueba exacta de Fisher. También se contemplan otras pruebas estadísticas de significancia conocidas por el experto y que también están dentro del alcance de la divulgación.

La importancia de la susceptibilidad aumentada o disminuida puede determinarse en función de la relación entre las mediciones de un sujeto de prueba y un sujeto de referencia. Las pérdidas o ganancias de uno o más CNV pueden determinarse de acuerdo con una relación $\log_2$ umbral determinado por estas mediciones. Un valor de relación $\log_2$ superior a 0.35, o 0.5, puede ser indicativo de una ganancia de una o más CNV. Un valor de la relación $\log_2$ inferior a -0.35, o -0.5, puede ser indicativo de una pérdida de una o más CNV. La relación entre las mediciones de un sujeto de prueba y un sujeto de referencia puede invertirse de forma que las relaciones log2 de las ganancias de número de copias sean negativas y las relaciones log2 de las pérdidas de número de copias sean positivas.

La susceptibilidad combinada o global asociada con una pluralidad de variantes asociadas con la PML también puede ser evaluada; p. ej., las variaciones genéticas descritas en el presente documento para ser asociadas

con la susceptibilidad a la PML pueden ser combinadas con otros factores de riesgo genéticos comunes. El riesgo combinado de dichas variantes genéticas puede estimarse de forma análoga a los métodos descritos en el presente documento.

El cálculo del riesgo conferido por un genotipo particular para el individuo puede basarse en la comparación del genotipo del individuo con el riesgo previamente determinado expresado, p. ej., tal como un riesgo relativo (RR) o una razón de probabilidades (OR), para el genotipo, p. ej., para un vehículo heterocigoto de una variante de riesgo para PML. Una razón de probabilidades puede ser una medida estadística utilizada como indicador de causalidad. Por ejemplo, en la investigación de enfermedades genéticas puede utilizarse para transmitir la importancia de una variante en una cohorte de enfermedad en relación con una cohorte no afectada/normal. El riesgo calculado para el individuo puede ser el riesgo relativo para un sujeto, o para un genotipo específico de un sujeto, en comparación con la media de la población. El riesgo medio de la población puede expresarse como una media ponderada de los riesgos de los diferentes genotipos, utilizando los resultados de una población de referencia, y a continuación pueden realizarse los cálculos apropiados para calcular el riesgo de un grupo de genotipos en relación con la población. Alternativamente, el riesgo para un individuo puede basarse en una comparación de genotipos particulares, p. ej., vehículos heterocigotos y/u homocigotos de un alelo de riesgo de un marcador comparados con no vehículos del alelo de riesgo (o par de alelos en el caso de variantes heterocigotas compuestas, en las que una variante afecta al alelo heredado por vía materna y la otra afecta al alelo heredado por vía paterna). Utilizar la media de la población puede ser más conveniente, ya que proporciona una medida que puede ser fácil de interpretar para el usuario, p. ej., una medida que da el riesgo para el individuo, basado en su genotipo, comparado con la media de la población.

El valor OR puede calcularse como sigue: OR = (A/(N1-A))/(U/(N2-U)), donde A = número de casos afectados con variante, N1 = número total de casos afectados, U = número de casos no afectados con variante y N2 = número total de casos no afectados. En circunstancias en las que U = 0, es convencional fijar U=1, para evitar infinitos. La OR puede calcularse esencialmente como en el caso anterior, salvo que cuando U o A = 0, se añade 0.5 a todos los valores de A, N1, U, N2. La prueba exacta de Fisher (FET) puede calcularse utilizando métodos estándar. Los valores p pueden corregirse en función de la tasa de falsos descubrimientos (FDR) mediante el método de Benjamini-Hochberg (Benjamini Y. y Hochberg Y., J. Royal Statistical Society 57:289 (1995); Osborne J. A. y Barker C. A. (2007)).

Una variación genética puede correlacionarse con PML referenciando los datos de la variación genética a una tabla de búsqueda que comprende correlaciones entre la variación genética y PML. La variación genética puede comprender al menos una indicación de la variación genética. La tabla puede incluir una correlación para una variación genética. La tabla puede comprender una correlación para una pluralidad de variaciones genéticas en ambos escenarios, haciendo referencia a una tabla de búsqueda que da una indicación de una correlación entre una variación genética y la PML, se puede identificar un riesgo de PML, o una susceptibilidad a la PML, en el individuo del que se deriva la muestra de ácido nucleico.

La presente divulgación también se refiere a métodos de detección clínicos, p. ej., diagnóstico, pronóstico o teranosis de un sujeto realizados por un profesional médico utilizando los métodos divulgados en el presente documento. La divulgación puede referirse a métodos de detección realizados por un profano. El profano puede ser cliente de un proveedor de servicios de genotipado, micromatrices, secuenciación del exoma o secuenciación del genoma completo. El profano también puede ser un proveedor de servicios de genotipado, microarreglos, secuenciación del exoma o secuenciación del genoma completo, que realiza análisis genéticos de una muestra de ADN de un individuo, con el fin de prestar un servicio relacionado con los factores de riesgo genético de rasgos o enfermedades particulares, basándose en el estado genotípico del sujeto obtenido a partir del uso de los métodos descritos en el presente documento. El genotipo resultante o la información genética puede ponerse a disposición del individuo y puede compararse con la información sobre la PML o el riesgo de desarrollar PML asociada a una o varias variaciones genéticas, incluyendo, pero no limitándose a, información de bases de datos públicas o privadas de variaciones genéticas o literatura y publicaciones científicas. Las aplicaciones de detección de las variaciones genéticas asociadas a la PML, tal y como se describen en el presente documento, pueden ser realizadas, p. ej., por un individuo, un profesional sanitario o un tercero, p. ej., un proveedor de servicios que interprete la información del genotipo del sujeto. El análisis genético puede realizarse en un laboratorio con certificación CLIA (es decir, las normas reguladoras federales de EE.UU. que se especifican en las Enmiendas para la Mejora de los Laboratorios Clínicos, administradas por los Centros de Servicios de Medicare y Medicaid) o en laboratorios equivalentes de Europa y otras partes del mundo.

La información derivada del análisis de los datos de secuencia puede comunicarse a cualquier organismo en particular, incluido el individuo del que se deriva la muestra de ácido nucleico o los datos de secuencia, un tutor o representante del individuo, un clínico, un profesional de la investigación, un profesional médico, un proveedor de servicios y un asegurador médico o una compañía de seguros. Los profesionales médicos pueden ser, p. ej., médicos, enfermeros, técnicos de laboratorio médico y farmaceutas. Los profesionales de la investigación pueden ser, p. ej., investigadores principales, técnicos de investigación, becarios posdoctorales y estudiantes de posgrado.

Un profesional puede ser asistido determinando si variantes genéticas específicas están presentes en una muestra de ácido nucleico de un sujeto, y comunicando información sobre variantes genéticas a un profesional. Una vez comunicada la información sobre variantes genéticas específicas, un profesional médico puede tomar una o varias medidas que pueden afectar a la atención del sujeto. Por ejemplo, un profesional médico puede registrar información en la historia clínica del sujeto (p. ej., registros de salud electrónicos o registros médicos electrónicos, incluidos, entre otros, los servicios sanitarios a escala nacional, tal como el Servicio Nacional de Salud del Reino Unido) sobre el riesgo del sujeto de desarrollar PML. Un profesional médico puede registrar información relativa a la evaluación de riesgos, o transformar de otro modo la historia clínica del sujeto, para reflejar su estado de salud actual. Un profesional médico puede revisar y evaluar el historial médico completo de un sujeto y valorar múltiples estrategias de tratamiento para la intervención clínica de la enfermedad de un sujeto. La información puede registrarse en el contexto del sistema desarrollado por la Organización Mundial de la Salud (OMS), la Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la Salud (ICD), que actualmente utiliza la 10ª revisión (códigos ICD-10). Por ejemplo, el código ICD-10 para la PML es A81.2, mientras que el código ICD-10 para la esclerosis múltiple es G35.

Un profesional médico puede iniciar o modificar el tratamiento tras recibir información sobre la detección de PML en un sujeto, por ejemplo. Un profesional médico puede recomendar un cambio de terapia o excluir una terapia. Un profesional médico puede inscribir a un sujeto en un ensayo clínico para, a modo de ejemplo, detectar correlaciones entre un haplotipo como el descrito en el presente documento y cualquier parámetro medible o cuantificable relacionado con el resultado del tratamiento como el descrito anteriormente.

Un profesional médico puede comunicar al sujeto o a su familia información relativa a la probabilidad de un sujeto de desarrollar PML. Un profesional médico puede proporcionar a un sujeto y/o a su familia información respecto a la PML y la información sobre la evaluación de riesgos, incluyendo, opciones de tratamiento y derivaciones a especialistas. Un profesional médico puede facilitar una copia del historial médico de un sujeto a un especialista. Un profesional de la investigación puede aplicar la información relativa al riesgo de un sujeto de desarrollar PML para avanzar en la investigación científica. Un profesional de la investigación puede obtener el haplotipo de un sujeto tal y como se describe en el presente documento para evaluar su inscripción, o su participación continuada, en un estudio de investigación o ensayo clínico. Un profesional de la investigación puede comunicar la información relativa a la detección de PML en un sujeto a un profesional médico. Un profesional de la investigación puede remitir a un sujeto a un profesional médico.

Se puede utilizar cualquier método apropiado para comunicar información a otra persona. Por ejemplo, la información puede facilitarse directa o indirectamente a un profesional y un técnico de laboratorio puede introducir la variación genética de un sujeto, tal como se describe en el presente documento, en un registro informático. Se puede comunicar información alterando físicamente los historiales médicos o de investigación. Por ejemplo, un profesional médico puede hacer una anotación permanente o marcar un historial médico para comunicar la evaluación de riesgos a otros profesionales médicos que revisen el historial. Además, se puede utilizar cualquier tipo de comunicación para transmitir la información de la evaluación de riesgos. Por ejemplo, se puede utilizar el correo postal, el correo electrónico, el teléfono y las interacciones cara a cara. La información también puede comunicarse a un profesional poniéndola a su disposición por vía electrónica. Por ejemplo, la información puede comunicarse a un profesional colocándola en una base de datos informática de forma que el profesional pueda acceder a ella. Además, la información puede comunicarse a un hospital, clínica o centro de investigación que actúe como agente del profesional.

Los resultados de estas pruebas, y opcionalmente la información interpretativa, pueden ser devueltos al sujeto, al proveedor de atención médica o a un tercero. Los resultados pueden comunicarse al sujeto sometido a la prueba, p. ej., con un pronóstico y, opcionalmente, con materiales interpretativos que pueden ayudar al sujeto a comprender los resultados de la prueba y el pronóstico; utilizados por un proveedor de asistencia sanitaria, p. ej., para determinar si debe administrar un fármaco específico, o si un sujeto debe ser asignado a una categoría específica, p. ej., una categoría asociada con un endofenotipo de enfermedad específico, o con la respuesta o no respuesta al fármaco; utilizado por un tercero, tal como un pagador de asistencia sanitaria, p. ej., una compañía de seguros o una HMO, u otro organismo, para determinar si debe o no reembolsar a un proveedor de asistencia sanitaria los servicios prestados al sujeto, o si debe aprobar la prestación de servicios al sujeto. Por ejemplo, el pagador de la asistencia sanitaria puede decidir reembolsar a un proveedor de asistencia sanitaria los tratamientos para la PML si el sujeto padece PML o tiene un mayor riesgo de desarrollarla.

También se proporcionan en el presente documento bases de datos que incluyen una lista de variaciones genéticas como se describe en el presente documento, y en el que la lista puede ser en gran parte o totalmente limitada a variaciones genéticas identificadas como útiles para la detección de PML como se describe en el presente documento. La lista puede almacenarse, p. ej., en un archivo plano o en un vehículo legible por ordenador. Las bases de datos pueden incluir además información relativa a uno o más sujetos, p. ej., si un sujeto está afectado o no, información clínica como el endofenotipo, la edad de inicio de los síntomas, cualquier tratamiento administrado y los resultados, p. ej., datos relevantes para la farmacogenómica, el diagnóstico, el pronóstico o la teranóstica, y otros detalles, p. ej., datos sobre el trastorno en el sujeto, o factores ambientales (p. ej., incluyendo, entre otros, la infección o un historial de infección por VIH o JCV) u otros factores genéticos.

Las bases de datos pueden utilizarse para detectar correlaciones entre un haplotipo concreto y la información relativa al sujeto.

Los métodos descritos en el presente documento también pueden incluir la generación de informes para su uso, p. ej., por un sujeto, un cuidador o un investigador, que incluyan información relativa a las variaciones genéticas de un sujeto y, opcionalmente, información adicional como los tratamientos administrados, el historial de tratamientos, el historial médico, la respuesta prevista y la respuesta real. Los informes pueden grabarse en un medio tangible, p. ej., un disco legible por ordenador, un dispositivo de memoria de estado sólido o un dispositivo de almacenamiento óptico.

Métodos de detección mediante variaciones de ARN y/o polipéptidos

La presente invención consiste en determinar la ausencia o presencia de una variación genómica seleccionada de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. A continuación se ofrece una explicación general de los métodos de detección que pueden utilizarse para determinar la presencia o ausencia de las variaciones genómicas de las reivindicaciones, y se incluyen ejemplos que no forman parte de la materia reivindicada en el presente documento, en la medida en que se refieren a variaciones distintas de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. La detección de PML puede realizarse examinando o comparando cambios en la expresión, localización, compañeros de unión y composición de un polipéptido codificado por una variante de ácido nucleico asociada con PML, p. ej., en aquellos casos en los que las variaciones genéticas de la presente divulgación den lugar a un cambio en la composición o expresión del polipéptido y/o ARN, p. ej., ARNm, microARN (miARN) y otros ARN no codificantes (ARNnc). De este modo, la detección de PML puede hacerse examinando la expresión y/o composición de uno de estos polipéptidos y/o ARN, u otro polipéptido y/o ARN codificado por un ácido nucleico asociado con PML, en aquellos casos en los que la variación genética de la presente divulgación resulte en un cambio en la expresión, localización, compañeros de unión, y/o composición del polipéptido y/o ARN. La detección puede comprender el diagnóstico de un sujeto. La detección puede comprender la determinación del pronóstico de un sujeto, p. ej., la determinación de la susceptibilidad de desarrollar PML. La detección puede comprender la teranosis de un sujeto.

Las variaciones genéticas descritas en el presente documento que muestran asociación a PML pueden jugar un papel a través de su efecto en uno o más de estos genes, ya sea impactando directamente uno o más genes o influenciando la expresión de uno o más genes cercanos. Por ejemplo, sin pretender estar limitado por la teoría, generalmente se espera que la eliminación de un segmento cromosómico que comprende un gen particular, o un fragmento de un gen, pueda dar lugar a una composición o expresión alterada, o ambas, del polipéptido y/o ARNm codificado. Del mismo modo, se espera que las duplicaciones, o las variaciones en el número de copias con un número elevado, den lugar en general a un aumento de la expresión del polipéptido y/o ARN codificado si el gen del que se expresan está totalmente englobado dentro del segmento genómico duplicado (o triplicado, o incluso con ganancias en el número de copias superiores), o por el contrario pueden dar lugar a una disminución de la expresión o a un ARN o polipéptido alterado si uno o ambos puntos de rotura de la ganancia en el número de copias alteran un gen determinado. Otros posibles mecanismos que afectan a los genes dentro de una región de variación genética incluyen, p. ej., efectos sobre la transcripción, efectos sobre el empalme del ARN, alteraciones en las cantidades relativas de formas alternativas de empalme del ARNm, efectos sobre la estabilidad del ARN, efectos sobre el transporte del núcleo al citoplasma y efectos sobre la eficacia y precisión de la traducción. De este modo, las variaciones de ADN pueden detectarse directamente, utilizando el ADN genómico amplificado o no amplificado de los sujetos, o indirectamente, utilizando ARN o ADN obtenido del tejido o tejidos del sujeto que estén presentes en una forma o nivel de expresión aberrante como resultado de las variaciones genéticas de la divulgación que muestran asociación con la PML. Las variaciones en el ADN pueden detectarse indirectamente utilizando un polipéptido o proteína obtenido del tejido o tejidos del sujeto que esté presente en una forma o nivel de expresión aberrante como resultado de las variaciones genéticas de la divulgación que muestren asociación con la PML. Una forma aberrante o nivel de expresión de un polipéptido o proteína que resulta de una o más variaciones genéticas de la divulgación que muestra asociación a PML puede detectarse indirectamente a través de otro polipéptido o proteína presente en la misma vía biológica/celular que se modula o interactúa con dicho polipéptido o proteína que resulta de una o más variaciones genéticas de la divulgación. Las variaciones genéticas de la divulgación que muestran asociación con la PML pueden afectar a la expresión de un gen dentro de la región de variación genética. Una variación genética que afecte a una región exónica de un gen puede afectar, alterar o modular la expresión del gen. Una variación genética que afecte a una región intrónica o intergénica de un gen puede afectar, alterar o modular la expresión del gen.

Ciertas regiones de variación genética pueden tener segmentos duplicados flanqueantes, y los genes dentro de tales segmentos pueden tener una expresión y/o composición alterada como resultado de tales alteraciones genómicas. Los elementos reguladores que afectan a la expresión génica pueden estar situados lejos, incluso a decenas o centenares de kilobases de distancia, del gen regulado por dichos elementos reguladores. Por lo tanto, los elementos reguladores de los genes que se encuentran fuera del gen (p. ej., secuencia arriba o secuencia abajo del gen) pueden estar ubicados dentro de la variación genética y, por lo tanto, verse afectados por la variación genética. Por lo tanto, se contempla que la detección de las variaciones genéticas descritas en el presente documento puede utilizarse para evaluar la expresión de uno o más genes asociados no afectados directamente por las variaciones genéticas. Una variación genética que afecte a una región intergénica de un gen puede afectar, alterar o modular la expresión de un gen situado en otra parte del genoma, tal como se ha descrito anteriormente. Por ejemplo, una variación genética que afecte a una región intergénica de un gen puede afectar, interrumpir o modular la expresión de un factor de transcripción, situado en otra parte del genoma, que regula el gen. Los elementos reguladores también pueden localizarse dentro de un gen, tal como en las regiones intrónicas, y afectar de forma similar al nivel de expresión del gen y, en última instancia, al nivel de expresión de la proteína sin cambiar la estructura de la proteína. Los efectos de las variantes genéticas sobre los elementos reguladores pueden manifestarse de forma específica para cada tejido; p. ej., uno o más factores de transcripción que se unen al elemento regulador afectado por una o más variaciones genéticas pueden expresarse en mayor concentración en las neuronas que en las células de la piel (es decir, el impacto de una o más variaciones genéticas puede ser evidente principalmente en las células neuronales).

Las variaciones genéticas de la divulgación que muestran asociación a PML pueden afectar la expresión de la proteína a nivel traslacional. Los expertos en la materia pueden apreciar que esto puede ocurrir por el aumento o la disminución de la expresión de uno o más microARN (miARN) que regula la expresión de una proteína que se sabe que es importante, o está implicada, en la causa, aparición o progresión de la PML. El aumento o disminución de la expresión de uno o más miARN puede ser el resultado de la ganancia o pérdida de todo el gen de miARN, la interrupción o deterioro de una porción del gen (p. ej., por un indel o CNV), o incluso un cambio de una sola base (SNP o SNV) que produce una secuencia de miARN alterada, no funcional o de funcionamiento aberrante. También puede ser apreciado por los expertos en la materia que la expresión de la proteína, p. ej., una conocida por causar PML por aumento o disminución de la expresión, puede resultar debido a una variación genética que resulta en la alteración de un sitio de unión de miARN existente dentro de la transcripción de ARNm del polipéptido, o incluso crea un nuevo sitio de unión de miARN que conduce a la expresión aberrante del polipéptido.

Se puede utilizar una variedad de métodos para detectar la composición y/o los niveles de expresión de polipéptidos, incluyendo, pero sin limitarse a ensayos inmunoabsorbentes ligados a enzimas (ELISA), transferencias Western, espectroscopia, espectrometría de masas, matrices peptídicas, colorimetría, electroforesis, enfoque isoeléctrico, inmunoprecipitaciones, inmunoensayos e inmunofluorescencia y otros métodos bien conocidos en la técnica. Una muestra de ácido nucleico de prueba de un sujeto puede evaluarse para detectar la presencia de una alteración en la expresión y/o una alteración en la composición del polipéptido codificado por un ácido nucleico asociado a PML. Una "alteración" en la expresión o composición del polipéptido, tal como se utiliza en el presente documento, se refiere a una alteración en la expresión o composición en una muestra de ácido nucleico de prueba, en comparación con la expresión o composición del polipéptido en una muestra de ácido nucleico de control. Dicha alteración puede ser, p. ej., una alteración en la expresión cuantitativa del polipéptido o puede ser una alteración en la expresión cualitativa del polipéptido, p. ej., la expresión de un polipéptido mutante o de una variante de empalme diferente, o una combinación de las mismas. La detección de PML puede realizarse detectando una variante de empalme particular codificada por un ácido nucleico asociado con PML, o un patrón particular de variantes de empalme.

Los anticuerpos pueden ser policlonales o monoclonales y pueden estar marcados o no. Puede utilizarse un anticuerpo intacto o un fragmento del mismo. El término "marcado", en relación con la sonda o el anticuerpo, pretende abarcar el marcado directo de la sonda o el anticuerpo mediante el acoplamiento de una sustancia detectable a la sonda o el anticuerpo, así como el marcado indirecto de la sonda o el anticuerpo mediante reactividad con otro reactivo marcado directamente como se ha descrito anteriormente en el presente documento. Otros ejemplos no limitantes de marcaje indirecto incluyen la detección de un anticuerpo primario utilizando un anticuerpo secundario marcado, p. ej., un anticuerpo secundario marcado con fluorescencia y el marcaje final de una sonda de ADN con biotina de forma que pueda detectarse con estreptavidina marcada con fluorescencia.

Métodos de detección de variaciones genéticas

La presente invención consiste en determinar la ausencia o presencia de una variación genómica seleccionada de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4: 103522150, G>A. A continuación se ofrece una explicación general de los métodos de detección que pueden utilizarse para determinar la presencia o ausencia de las variaciones genómicas de las reivindicaciones, y se incluyen ejemplos que no forman parte de la materia

reivindicada en el presente documento, en la medida en que se refieren a variaciones distintas de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. Pueden utilizarse técnicas estándar de genotipado para detectar la presencia de variaciones genéticas, p. ej., la amplificación. La amplificación de ácidos nucleicos puede llevarse a cabo mediante métodos conocidos en la técnica. Generalmente, la información de la secuencia de la región de interés puede utilizarse para diseñar cebadores oligonucleotídicos que pueden ser idénticos o similares en secuencia a las cadenas opuestas de un molde a amplificar. Los métodos de amplificación pueden incluir, entre otros, técnicas basadas en la fluorescencia que utilizan la PCR, p. ej., la reacción en cadena de la ligasa (CSF), la PCR anidada, la amplificación por transcripción, la replicación de secuencias autosostenida, la amplificación de secuencias basada en ácidos nucleicos (NASBA) y la amplificación de sondas dependiente de ligadura múltiple (MLPA). Las directrices para la selección de cebadores para la amplificación por PCR son bien conocidas en la técnica. Se puede utilizar un programa informático para diseñar los cebadores, p. ej., Oligo (National Biosciences, Inc., Plymouth Minn), MacVector (Kodak/IBI) y suite GCG de programas de análisis de secuencias.

Se pueden utilizar metodologías comerciales disponibles para el genotipado, p. ej., genotipado SNP, pero no se limitan a, ensayos de genotipado TaqMan (Applied Biosystems), plataformas SNPlex (Applied Biosystems), electroforesis en gel, electroforesis capilar, cromatografía de exclusión por tamaño, espectrometría de masas, p. ej., sistema MassARRAY (Sequenom), métodos de minisecuenciación, reacción en cadena de la polimerasa (PCR) en tiempo real, sistema Bio-Plex (BioRad), sistemas CEQ y SNPstream (Beckman), tecnología de hibridación de arreglos, p. ej., Affymetrix GeneChip (Perlegen), BeadArray Technologies, p. ej., ensayos Illumina GoldenGate e Infinium, tecnología de marcadores de arreglos, amplificación de sonda dependiente de ligadura múltiple (MLPA), y tecnología de hibridación por fluorescencia basada en endonucleasas (ensayo Invader, utilizando ADN genómico amplificado o no amplificado, o ARN total no amplificado, o ADNc amplificado o no amplificado; Third Wave/Hologic). La PCR puede ser un procedimiento en el que el ácido nucleico diana se amplifica de forma similar a la descrita en la patente de los Estados Unidos No. 4,683,195 y modificaciones posteriores del procedimiento descrito en la misma. La PCR puede incluir un ciclo de tres fases de temperatura de desnaturalización del ADN en cadenas simples, hibridación de los cebadores con las cadenas desnaturalizadas y extensión de los cebadores por una enzima ADN polimerasa termoestable. Este ciclo puede repetirse hasta que haya suficientes copias para ser detectadas y analizadas. La PCR cuantitativa en tiempo real puede utilizarse para determinar variaciones genéticas, en las que la PCR cuantitativa puede permitir tanto la detección como la cuantificación de una secuencia de ADN en una muestra de ácido nucleico, p. ej., tal como un número absoluto de copias o como una cantidad relativa cuando se normaliza con respecto a la entrada de ADN u otros genes normalizadores. Los métodos de cuantificación pueden incluir el uso de tintes fluorescentes que pueden intercalarse con ADN de doble cadena, y sondas de oligonucleótidos de ADN modificados que pueden emitir fluorescencia cuando se hibridan con un ADN complementario.

Una muestra de ácido nucleico obtenida del sujeto puede ser recogida y la PCR puede ser usada para amplificar un fragmento de ácido nucleico que comprende una o más variaciones genéticas que pueden ser indicativas de una susceptibilidad a PML. La detección de variaciones genéticas puede realizarse mediante análisis de expresión, p. ej., utilizando la PCR cuantitativa. Esta técnica puede evaluar la presencia o ausencia de una alteración genética en la expresión o composición de uno o más polipéptidos o variantes de empalme codificados por un ácido nucleico asociado a PML.

La muestra de ácido nucleico de un sujeto que contiene un SNP puede amplificarse por PCR antes de la detección con una sonda. El ADN amplificado puede servir de molde para una sonda de detección y, en algunos casos, para una sonda potenciadora. La sonda de detección, la sonda potenciadora y/o los cebadores utilizados para la amplificación del molde por PCR pueden comprender el uso de bases modificadas, p. ej., A, T, C, G y U modificadas, en las que el uso de bases modificadas puede ser útil para ajustar la temperatura de fusión de la sonda de nucleótidos y/o el cebador al ADN molde. Las bases modificadas pueden utilizarse en el diseño de la sonda de nucleótidos de detección. En estos métodos puede seleccionarse cualquier base modificada conocida por el experto, y la selección de bases adecuadas está dentro del alcance del experto basándose en las enseñanzas del presente documento y en las bases conocidas disponibles en fuentes comerciales conocidas por el experto.

La identificación de variaciones genéticas puede lograrse utilizando métodos de hibridación. La presencia de un alelo marcador específico o de un segmento genómico particular que comprende una variación genética, o representativo de una variación genética, puede indicarse mediante la hibridación específica de secuencia de una sonda de ácido nucleico específica para el alelo particular o la variación genética en una muestra de ácido nucleico que ha sido o no amplificada, por métodos descritos en el presente documento. La presencia de más de un alelo marcador específico o de varias variaciones genéticas puede indicarse utilizando dos o más sondas de ácido nucleico de secuencia específica, en las que cada una es específica para un alelo y/o variación genética particular.

La hibridación puede realizarse por métodos bien conocidos por el experto en la materia, p. ej., técnicas de hibridación como la hibridación fluorescente *in situ* (FISH), el análisis Southern, el análisis Northern o la hibridación *in situ*. La hibridación puede referirse a la hibridación específica, en la que la hibridación puede realizarse sin discordancias. La hibridación específica, si existe, puede realizarse mediante métodos estándar. Si se produce hibridación específica entre una sonda de ácido nucleico y el ácido nucleico de la muestra de ácido nucleico, la muestra de ácido nucleico puede contener una secuencia que puede ser complementaria a un nucleótido presente en la sonda de ácido nucleico. Si una sonda de ácido nucleico puede contener un alelo particular de un marcador polimórfico, o alelos particulares para una pluralidad de marcadores, la hibridación específica puede ser indicativa de que el ácido nucleico es completamente complementario a la sonda de ácido nucleico, incluyendo, los alelos particulares en los marcadores polimórficos dentro de la sonda. Una sonda puede contener más de un alelo marcador de un haplotipo particular, p. ej., una sonda puede contener alelos complementarios a 2, 3, 4, 5 o todos los marcadores que componen un haplotipo particular. La detección de uno o más marcadores particulares del haplotipo en la muestra de ácido nucleico puede ser indicativa de que la fuente de la muestra de ácido nucleico tiene el haplotipo particular.

Se pueden desarrollar condiciones de PCR y cebadores que amplifiquen un producto sólo cuando el alelo variante está presente o sólo cuando el alelo de tipo silvestre está presente, p. ej., PCR específica de alelo. Con la PCR específica de alelo, se puede emplear un método que utiliza una sonda de oligonucleótido de detección que comprende una fracción o grupo fluorescente en su terminal 3' y un inactivador en su terminal 5', y un oligonucleótido potenciador(véase, p. ej., Kutyavin et al., Nucleic Acid Res. 34:Pe128 (2006)).

Un cebador/sonda específico de alelo puede ser un oligonucleótido que es específico para un polimorfismo particular puede ser preparado usando métodos estándar. Las sondas oligonucleótidas específicas de alelo pueden hibridarse específicamente con una región de ácido nucleico que contenga una variación genética. Las condiciones de hibridación pueden seleccionarse de forma que una sonda de ácido nucleico pueda unirse específicamente a la secuencia de interés, p. ej., la secuencia de ácido nucleico variante.

El análisis de digestión de restricción específico de alelo puede ser usado para detectar la existencia de una variante polimórfica de un polimorfismo, si variantes polimórficas alternas del polimorfismo pueden resultar en la creación o eliminación de un sitio de restricción. Las digestiones de restricción específicas de alelo pueden realizarse, p. ej., con la enzima de restricción particular que pueda diferenciar los alelos. La PCR puede utilizarse para amplificar una región que comprenda el sitio polimórfico, y puede realizarse un análisis de polimorfismo de longitud de fragmentos de restricción. Para las variantes de secuencia que no alteran un sitio de restricción común, se pueden diseñar cebadores mutagénicos que pueden introducir uno o más sitios de restricción cuando el alelo variante está presente o cuando el alelo de tipo silvestre está presente.

La incorporación de colorante-terminador dirigida por molde de polarización de fluorescencia (FP-TDI) puede utilizarse para determinar cuál de las múltiples variantes polimórficas de un polimorfismo puede estar presente en un sujeto. A diferencia del uso de sondas o cebadores específicos de alelos, este método puede emplear cebadores que pueden terminar adyacentes a un sitio polimórfico, de modo que la extensión del cebador por un solo nucleótido puede resultar en la incorporación de un nucleótido complementario a la variante polimórfica en el sitio polimórfico.

El ADN que contiene una porción amplificada puede ser sometido a transferencia de puntos, utilizando métodos estándar y poniendo en contacto la transferencia con la sonda de oligonucleótidos. A continuación, puede detectarse la presencia de hibridación específica de la sonda con el ADN. Los métodos pueden incluir la determinación del genotipo de un sujeto con respecto a ambas copias del sitio polimórfico presente en el genoma, en el que, si existen múltiples variantes polimórficas en un sitio, esto puede indicarse adecuadamente especificando qué variantes están presentes en un sujeto. Cualquiera de los medios de detección descritos en el presente documento puede utilizarse para determinar el genotipo de un sujeto con respecto a una o ambas copias del polimorfismo presente en el genoma del sujeto.

Se puede utilizar una sonda de ácido nucleico peptídico (PNA) además de, o en lugar de, una sonda de ácido nucleico en los métodos descritos en el presente documento. Un PNA puede ser un imitador del ADN que tiene una cadena inorgánica similar a un péptido, p. ej., unidades de N-(2-aminoetil) glicina con una base orgánica (A, G, C, T o U) unida al nitrógeno de la glicina mediante un enlazador de metileno carbonilo.

El análisis de la secuencia del ácido nucleico también puede utilizarse para detectar variaciones genéticas, p. ej., las variaciones genéticas pueden detectarse secuenciando exones, intrones, secuencias no traducidas 5' o secuencias no traducidas 3'. Para detectar variaciones genéticas, pueden utilizarse uno o varios métodos de análisis de ácidos nucleicos conocidos por los expertos en la materia, tal como la secuenciación manual directa, la secuenciación fluorescente automatizada, los ensayos de polimorfismo de conformación monocatenario (SSCP), la electroforesis en gel desnaturalizante con pinzas (CDGE), la electroforesis en gel de gradiente desnaturalizante (DGGE), la electroforesis en gel bidimensional (2DGE o TDGE), la electroforesis en gel sensible a la conformación (CSGE); cromatografía líquida desnaturalizante de alto rendimiento (DHPLC), espectrometría de masas por desorción/ionización láser asistida por matriz de infrarrojos (IR-MALDI), análisis de desplazamiento de movilidad, PCR cuantitativa en tiempo real, análisis de enzimas de restricción, análisis

de heterodúplex; escisión química de desajustes (CMC), ensayos de protección de RNasa, uso de polipéptidos que reconocen desajustes de nucleótidos, PCR específica de alelo, secuenciación de ADN con pirofosfato en tiempo real, amplificación de PCR en combinación con cromatografía líquida desnaturalizante de alto rendimiento (dHPLC) y combinaciones de dichos métodos.

La secuenciación puede realizarse mediante métodos clásicos de secuenciación de Sanger, conocidos en la técnica. La secuenciación puede realizarse utilizando métodos de secuenciación de alto rendimiento, algunos de los cuales permiten la detección de un nucleótido secuenciado inmediatamente después o en el momento de su incorporación a una cadena en crecimiento, p. ej., la detección de la secuencia en tiempo sustancialmente real o en tiempo real. En algunos casos, la secuenciación de alto rendimiento genera al menos 1,000, al menos 5,000, al menos 10,000, al menos 20,000, al menos 30,000, al menos 40,000, al menos 50,000, al menos 100,000 o al menos 500,000 lecturas de secuencias por hora; siendo cada lectura de al menos 50, al menos 60, al menos 70, al menos 80, al menos 90, al menos 100, al menos 120 o al menos 150 bases por lectura (o 500 - 1,000 bases por lectura para 454).

Los métodos de secuenciación de alto rendimiento pueden incluir, entre otros, la secuenciación de firmas en paralelo masivo (MPSS, Lynx Therapeutics), la secuenciación Polony, la pirosecuenciación 454, la secuenciación Illumina (Solexa), la secuenciación Illumina (Solexa) utilizando la preparación de bibliotecas Genomics 10X, la secuenciación SOLiD, la secuenciación en semiconductor, Secuenciación de nanobolas de ADN, secuenciación de molécula única Helioscope™, secuenciación de molécula única SMRT™, secuenciación de molécula única en tiempo real (RNAP), secuenciación de ADN Nanopore, y/o secuenciación por hibridación, p. ej., un método no enzimático que utiliza una micromatriz de ADN, o secuenciación de Sanger de microfluidos.

La secuenciación de alto rendimiento puede implicar el uso de tecnología disponible por Helicos BioSciences Corporation (Cambridge, Mass.) como el método de secuenciación de molécula única por síntesis (SMSS). El SMSS es única porque permite secuenciar todo el genoma humano en un máximo de 24 horas. Este método de secuenciación rápida también permite detectar un SNP/nucleótido en una secuencia en tiempo sustancialmente real o en tiempo real. Por último, el SMSS es potente porque, al igual que la tecnología MIP, no utiliza una etapa de preamplificación previa a la hibridación. El SMSS no utiliza ningún tipo de amplificación. El SMSS se describe en las publicaciones de solicitudes de patente de los Estados Unidos Nos. 20060024711; 20060024678; 20060012793; 20060012784 y 20050100932. La secuenciación de alto rendimiento puede implicar el uso de la tecnología disponible por 454 Life Sciences, Inc. (una empresa de Roche, Branford, Conn.) tal como el dispositivo PicoTiterPlate que incluye una placa de fibra óptica que transmite la señal quimioluminiscente generada por la reacción de secuenciación para ser registrada por una cámara CCD en el instrumento. Este uso de la fibra óptica permite detectar un mínimo de 20 millones de pares de bases en 4.5 horas.

El ácido nucleico monocatenario amplificado por PCR puede hibridarse con un cebador e incubarse con una polimerasa, ATP sulfurilasa, luciferasa, apirasa y los sustratos luciferina y adenosina 5' fosfosulfato. A continuación, pueden añadirse secuencialmente trifosfatos de desoxinucleótidos correspondientes a las bases A, C, G y T (U). La incorporación de una base puede ir acompañada de la liberación de pirofosfato, que puede ser convertido en ATP por la sulfurilasa, lo que puede impulsar la síntesis de oxiluciferina y la liberación de luz visible. Dado que la liberación de pirofosfato puede ser equimolar con el número de bases incorporadas, la luz emitida puede ser proporcional al número de nucleótidos añadidos en cualquier etapa. El proceso puede repetirse hasta que pueda determinarse la secuencia completa. La pirosecuenciación puede utilizarse para analizar los amplicones para determinar si existen puntos de rotura. La pirosecuenciación puede mapear secuencias circundantes como control de calidad interno.

Los métodos de análisis de pirosecuenciación son conocidos en la técnica. El análisis de secuencias puede incluir un esquema de secuenciación por ligadura de cuatro colores (ligadura degenerada), que consiste en hibridar un cebador de anclaje en una de las cuatro posiciones. A continuación, se puede llevar a cabo una reacción de ligación enzimática del cebador de anclaje con una población de nonámeros degenerados marcados con colorantes fluorescentes. En cualquier ciclo dado, la población de nonámeros que se utiliza puede estructurarse de tal forma que la identidad de una de sus posiciones pueda correlacionarse con la identidad del fluoróforo unido a ese nonámero. En la medida en que la ligasa discrimina por complementariedad en esa posición consultada, la señal fluorescente puede permitir inferir la identidad de la base. Después de realizar la ligadura y la obtención de imágenes de cuatro colores, los complejos cebador de anclaje: nonámero pueden despojarse y comienza un nuevo ciclo. Los métodos para obtener imágenes de la información de la secuencia después de realizar la ligadura son conocidos en la técnica.

El análisis por digestión de enzimas de restricción puede utilizarse para detectar una variación genética particular si la variación genética resulta en la creación o eliminación de uno o más sitios de restricción en relación con una secuencia de referencia. Puede realizarse un análisis de polimorfismo de longitud de fragmentos de restricción (RFLP), en el que el patrón de digestión del fragmento de ADN pertinente indica la presencia o ausencia de la variación genética concreta en la muestra de ácido nucleico.

Las matrices de sondas de oligonucleótidos que pueden ser complementarias a segmentos de secuencia de ácido nucleico diana de un sujeto pueden utilizarse para identificar variaciones genéticas. Un arreglo de sondas de oligonucleótidos puede comprender una matriz de oligonucleótidos, p. ej., una micromatriz. La presente divulgación presenta matrices que incluyen un sustrato que tiene una pluralidad de áreas direccionables, y métodos para utilizarlas. Al menos una zona de la pluralidad incluye una sonda de ácido nucleico que se une específicamente a una secuencia que comprende una variación genética, y puede utilizarse para detectar la ausencia o presencia de la variación genética, p. ej., uno o más SNP, microsatélites o CNV, tal como se describe en el presente documento, para determinar o identificar un alelo o genotipo. Por ejemplo, el arreglo puede incluir una o más sondas de ácido nucleico que pueden utilizarse para detectar una variación genética asociada a un gen y/o producto génico. La matriz puede comprender además al menos un área que incluya una sonda de ácido nucleico que puede utilizarse para detectar específicamente otro marcador asociado con PML como se describe en el presente documento.

La hibridación de microarreglos puede realizarse hibridando un ácido nucleico de interés, p. ej., un ácido nucleico que engloba una variación genética, con el arreglo y detectando la hibridación utilizando sondas de ácido nucleico. El ácido nucleico de interés puede amplificarse antes de la hibridación. La hibridación y la detección pueden llevarse a cabo de acuerdo con los métodos estándar descritos en las solicitudes PCT publicadas: WO 92/10092 y WO 95/11995 y la patente de los Estados Unidos No. 5,424,186. Por ejemplo, se puede escanear una matriz para determinar la posición en la matriz con la que se hibrida el ácido nucleico. Los datos de hibridación obtenidos del escaneo pueden ser, p. ej., en forma de intensidades de fluorescencia en función de la localización en la matriz.

Las matrices pueden formarse sobre sustratos fabricados con materiales como papel; vidrio; plástico, p. ej., polipropileno, nailon o poliestireno; poliacrilamida; nitrocelulosa; silicio; fibra óptica; o cualquier otro vehículo sólido o semisólido adecuado; y pueden configurarse en una configuración plana, p. ej., placas de vidrio o chips de silicio); o tridimensional, p. ej., alfileres, fibras, cuentas, partículas, pocillos de microtitulación y capilares.

Los métodos para generar matrices son conocidos en la técnica y pueden incluir, p. ej., métodos fotolitográficos (patentes de los Estados Unidos Nos. 5,143,854, 5,510,270 y 5,527,681); métodos mecánicos, p. ej., métodos de flujo dirigido (patente de los Estados Unidos No. 5,384,261); métodos basados en pasadores (patente de los Estados Unidos No. 5,288,514); técnicas basadas en cuentas (PCT US/93/04145); métodos de síntesis de oligonucleótidos en fase sólida; o mediante otros métodos conocidos por un experto en la materia (véase, por ejemplo Bier, F.F., et al., Adv Biochem Eng Biotechnol 109: 433-53 (2008); Hoheisel, J. D., Nat Rev Genet 7: 200-10 (2006); Fan, J. B., et al., Methods Enzymol 410: 57-73 (2006); Raqoussis, J. & Elvidge, G., Expert Rev Mol Design 6: 145-52 (2006); Mockler, T.C., et al., Genomics 85: 1-15 (2005), y las referencias citadas). Pueden encontrarse muchas descripciones adicionales de la preparación y uso de arreglos de oligonucleótidos para la detección de polimorfismos, p. ej., en los documentos US 6,858,394, US 6,429,027, US 5,445,934, US 5,700,637, US 5,744,305, US 5,945,334, US 6,054,270, US 6,300,063, US 6,733,977, US 7,364,858, EP 619 321 y EP 373 203. Los métodos de producción, hibridación y análisis de arreglos también se describen en Snijders et al. Genetics 29: 263-264 (2001); Klein et al. Natl. Acad. Sci. USA 96: 4494-4499 (1999); Albertson et al., Breast Cancer Research and Treatment 78: 289-298 (2003); y Snijders et al., "BAC microarray based comparative genomic hybridization", en: Zhao et al., (eds), Bacterial Artificial Chromosomes: Methods and Protocols, Methods in Molecular Biology, Humana Press (2002).

Las sondas de oligonucleótidos que forman una matriz pueden fijarse a un sustrato por cualquier cantidad de técnicas, incluyendo, pero sin limitarse a, síntesis *in situ*, p. ej., matrices de oligonucleótidos de alta densidad, utilizando técnicas fotolitográficas; manchado/impresión de una densidad media a baja sobre vidrio, nylon o nitrocelulosa; por enmascaramiento; y por transferencia de puntos sobre una membrana de hibridación de nylon o nitrocelulosa. Los oligonucleótidos pueden inmovilizarse mediante un enlazador, entre otros, por enlace covalente, iónico o físico. Los enlazadores para inmovilizar ácidos nucleicos y polipéptidos, incluidos los enlazadores reversibles o escindibles, son conocidos en la técnica (patente de los Estados Unidos No. 5,451,683 y el documento WO98/20019). Los oligonucleótidos pueden inmovilizarse de forma no covalente en un sustrato por hibridación con anclajes, mediante perlas magnéticas, o en una fase fluida, p. ej., en pocillos o capilares.

Una matriz puede comprender sondas de hibridación de oligonucleótidos capaces de hibridarse específicamente con diferentes variaciones genéticas. Las matrices de oligonucleótidos pueden comprender una pluralidad de sondas de oligonucleótidos diferentes acopladas a una superficie de un sustrato en diferentes ubicaciones conocidas. Las sondas de oligonucleótidos pueden presentar una unión diferencial o selectiva a sitios polimórficos, y pueden ser diseñadas fácilmente por un experto en la materia, p. ej., un oligonucleótido que sea perfectamente complementario a una secuencia que abarque un sitio polimórfico, p. ej., una secuencia que incluya el sitio polimórfico, dentro de ella, o en un extremo, puede hibridar preferentemente con un ácido nucleico que comprenda esa secuencia, frente a un ácido nucleico que comprenda una variante polimórfica alternativa.

Las matrices pueden incluir múltiples bloques de detección, p. ej., múltiples grupos de sondas diseñadas para la detección de polimorfismos particulares. Estas matrices pueden utilizarse para analizar múltiples

polimorfismos diferentes. Los bloques de detección pueden agruparse en una única matriz o en múltiples matrices separadas, en las que pueden utilizarse condiciones variables, p. ej., condiciones optimizadas para polimorfismos concretos, durante la hibridación. Pueden encontrarse descripciones generales del uso de arreglos de oligonucleótidos para la detección de polimorfismos, p. ej., en las patentes de los Estados Unidos Nos. 5,858,659 y 5,837,832. Además de las matrices de oligonucleótidos, también pueden utilizarse matrices de ADNc.

Los métodos descritos en el presente documento pueden incluir, entre otros, proporcionar una matriz como la descrita en el presente documento; poner en contacto la matriz con una muestra de ácido nucleico y detectar la unión de un ácido nucleico de la muestra de ácido nucleico a la matriz. El método puede comprender amplificar ácido nucleico de la muestra de ácido nucleico, p. ej., una región asociada con PML o una región que incluye otra región asociada con PML. Los métodos descritos en el presente documento pueden incluir el uso de una matriz que puede identificar patrones de expresión diferencial o números de copias de uno o más genes en muestras de ácido nucleico de individuos de control y afectados. Por ejemplo, las matrices de sondas a un marcador descrito en el presente documento pueden utilizarse para identificar variaciones genéticas entre el ADN de un sujeto afectado y el ADN de control obtenido de un individuo que no tiene PML. Dado que los nucleótidos de la matriz pueden contener marcadores de secuencia, sus posiciones en la matriz pueden conocerse con precisión en relación con la secuencia genómica.

Puede ser deseable emplear métodos que puedan detectar la presencia de múltiples variaciones genéticas, p. ej., variantes polimórficas en una pluralidad de sitios polimórficos, en paralelo o sustancialmente simultáneamente. Estos métodos pueden comprender matrices de oligonucleótidos y otros métodos, incluyendo, métodos en los que las reacciones, p. ej., amplificación e hibridación, pueden realizarse en recipientes individuales, p. ej., dentro de pocillos individuales de una placa de pocillos múltiples u otro recipiente.

Determinar la identidad de una variación genética también puede incluir o consistir en revisar el historial médico de un sujeto, donde el historial médico incluye información relativa a la identidad, número de copias, presencia o ausencia de uno o más alelos o SNP en el sujeto, p. ej., resultados de una prueba genética.

Los recorridos de homocigosidad extendidos (ROH) pueden ser útiles para mapear genes de enfermedades recesivas en poblaciones exógenas. Además, incluso en trastornos complejos, un elevado número de individuos afectados puede tener el mismo haplotipo en la región que rodea a la mutación de una enfermedad. Por lo tanto, una variante patogénica rara y el haplotipo circundante pueden estar enriquecidos en frecuencia en un grupo de individuos afectados en comparación con la frecuencia del haplotipo en una cohorte de controles no afectados. Los haplotipos homocigotos (HH) compartidos por múltiples individuos afectados pueden ser importantes para el descubrimiento de genes recesivos de la enfermedad en una afección como la PML. El método tradicional de mapeo de homocigosidad puede ampliarse analizando el haplotipo dentro de las regiones ROH compartidas para identificar segmentos homocigotos de haplotipo idéntico que estén presentes de forma única o con una mayor frecuencia en los probandos de PML en comparación con los controles parentales. Tales regiones se denominan haplotipos homocigotos de riesgo (rHH), que pueden contener variantes recesivas de baja frecuencia que contribuyen al riesgo de PML en un subconjunto de pacientes con PML.

Las variaciones genéticas también pueden identificarse utilizando cualquiera de los métodos conocidos en la técnica. Por ejemplo, pueden utilizarse variaciones genéticas disponibles en bases de datos públicas, que pueden buscarse utilizando métodos y algoritmos personalizados o algoritmos conocidos en la técnica. Una secuencia de referencia puede proceder, p. ej., del borrador de la secuencia del genoma humano, disponible públicamente en varias bases de datos, o de una secuencia depositada en una base de datos como GenBank.

Una comparación de uno o más genomas en relación con otro u otros genomas con matriz CGH, o una variedad de otros métodos de detección de variaciones genéticas, puede revelar el conjunto de variaciones genéticas entre dos genomas, entre un genoma en comparación con múltiples genomas, o entre un conjunto de genomas en comparación con otro conjunto de genomas. Un experimento con una matriz CGH puede realizarse hibridando un único genoma de prueba contra una muestra de ácido nucleico agrupada de dos o más genomas, lo que puede resultar en minimizar la detección de variantes de mayor frecuencia en el experimento. Un genoma de prueba puede hibridarse solo (es decir, detección de un solo color) con una micromatriz, p. ej., utilizando métodos de genotipado con la matriz CGH o SNP, y la etapa de comparación con uno o más genomas de referencia puede realizarse *in silico* para revelar el conjunto de variaciones genéticas en el genoma de prueba en relación con uno o más genomas de referencia. Un único genoma de prueba puede compararse con un único genoma de referencia en un experimento de 2 colores en el que ambos genomas se hibridan juntamente con la micromatriz, puede analizarse el genoma completo o el exoma completo de uno o más sujetos. La información sobre ácidos nucleicos ya se ha obtenido para el genoma completo o el exoma completo de uno o más individuos y la información sobre ácidos nucleicos puede obtenerse a partir de un análisis *in silico.*

Cualquiera de los polinucleótidos descritos, incluidos los polinucleótidos que comprenden una variación genética, puede fabricarse sintéticamente utilizando métodos conocidos en la técnica.

Medicamento inmunosupresor para uso terapéutico

En un aspecto la invención proporciona un medicamento inmunosupresor para uso en el tratamiento de una condición en un sujeto que necesita terapia con medicamentos inmunosupresores. El medicamento inmunosupresor para uso de la presente invención implica determinar la ausencia de una variación genómica seleccionada de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. A continuación se ofrece una explicación general del uso de un medicamento inmunosupresor, y se incluyen ejemplos que no forman parte de la materia reivindicada en el presente documento, en la medida en que se refieren a variaciones distintas de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. También se describen en el presente documento métodos, composiciones farmacéuticas y kits para el tratamiento de una afección en sujetos animales. La afección puede ser VIH/SIDA, cáncer o una enfermedad autoinmune. La afección puede ser PML. Por ejemplo, puede tratarse de esclerosis múltiple. Los métodos pueden comprender la administración de uno o más medicamentos inmunosupresores. Las composiciones farmacéuticas y los kits pueden comprender uno o más medicamentos inmunosupresores. El uno o más medicamentos inmunosupresores pueden ser adalimumab (p. ej., HUMIRA), alemtuzumab (p. ej., LEMTRADA), alentuzumab (p. ej., CAMPATH), azatioprina (p. ej., IMURAN), belimumab (p. ej., BENLYSTA), bevacizumab (p. ej., AVASTATIN), bortezomib (p. ej., VELCADE), eculizumab (p. ej., SOLIRIS), leflunomida, brentuximab vedotina (p. ej., ADCETRIS), cetuximab (p. ej., ERBITUX), ciclofosfamida, fumarato de dimetilo (p. ej., TECFIDERA), efalizumab (p. ej., RAPTIVA), fingolimod (p. ej., GILENYA), fludarabina (p. ej., FLUDARA), ácido fumárico, imatinib (p. ej., GLEEVEC, GLIVEC), infliximab (p. ej., REMICADE), metotrexato (p. ej., TREXALL, RHEUMATREX), micofenolato de mofetilo (p. ej., CELLCEPT), natalizumab (p. ej., TYSABRI), rituximab (p. ej., RITUXIN), daclizumab (p. ej., ZINBRYTA), vedolizumab (ENTYVIO), ruxolitinib (p. ej., JAKAFI, JAKAVI), ocrelizumab (p. ej., OCREVUS), o cualquier combinación de los mismos. El término "sujeto animal", tal y como se utiliza en el presente documento, incluye tanto a los seres humanos como a otros mamíferos. El término "tratar", tal y como se utiliza en el presente documento, incluye lograr un beneficio terapéutico y/o un beneficio profiláctico. Por beneficio terapéutico se entiende la erradicación o mejora de la infección viral subyacente (p. ej., el VIH), el cáncer o la enfermedad autoinmune.

Un sujeto puede ser tratado actualmente con un medicamento antirretroviral. Un sujeto puede ser tratado previamente con un medicamento antirretroviral. Un sujeto puede no estar aún tratado con un medicamento antirretroviral. El medicamento antirretroviral puede incluir, entre otros, inhibidores nucleósidos de la transcriptasa inversa (NRTI), inhibidores no nucleósidos de la transcriptasa inversa (NNRTI), inhibidores de la proteasa (PI), inhibidores de la fusión, inhibidores de la entrada, inhibidores de la integrasa, potenciadores farmacocinéticos y medicamentos combinados contra el VIH. En algunos casos, los Inhibidores nucleósidos de la transcriptasa inversa pueden incluir, entre otros, abacavir, didanosina, emtricitabina, lamivudina, estavudina, tenofovir disoproxil fumarato y zidovudina. En algunos casos, los inhibidores no nucleósidos de la transcriptasa inversa pueden incluir, entre otros, efavirenz, etravirina, nevirapina y rilpivirina. En algunos casos, los inhibidores de la proteasa pueden incluir, entre otros, atazanavir, darunavir, fosamprenavir, indinavir, nelfinavir, ritonavir, saquinavir y tipranavir. En algunos casos, los inhibidores de la fusión pueden incluir, entre otros, la enfuvirtida. En algunos casos, los inhibidores de entrada pueden incluir, pero no limitarse a maraviroc. En algunos casos, los inhibidores de la integrasa pueden incluir, entre otros, dolutegravir, elvitegravir y raltegravir. En algunos casos, los potenciadores farmacocinéticos pueden incluir, entre otros, cobicistat. En algunos casos, los medicamentos combinados contra el VIH pueden incluir, entre otros, abacavir y lamivudina, abacavir, dolutegravir y lamivudina, abacavir, lamivudina y zidovudina, atazanavir y cobicistat, darunavir y cobicistat, efavirenz, emtricitabina y tenofovir disoproxil fumarato, elvitegravir, cobicistat, emtricitabina y tenofovir alafenamida fumarato, elvitegravir, cobicistat, emtricitabina y tenofovir disoproxil fumarato, emtricitabina, rilpivirina y tenofovir alafenamida, emtricitabina, rilpivirina y tenofovir disoproxil fumarato, emtricitabina y tenofovir alafenamida, emtricitabina y tenofovir disoproxil fumarato, lamivudina y zidovudina, lopinavir y ritonavir, y cualquier combinación de medicamentos antirretrovirales enumerados anteriormente.

Cuando se identifica que un sujeto tiene al menos una de las variantes genéticas descritas en el presente documento, se puede administrar al sujeto un agente dirigido contra el virus JC. Puede administrarse a un sujeto un medicamento que impida el desarrollo de la PML, o que reduzca, disminuya, acorte y/o mejore de otro modo la progresión de la PML, o los síntomas que se desarrollen. La composición farmacéutica puede modular o dirigirse contra el virus de JC. A un sujeto identificado con PML se le puede administrar un agente que reduzca la carga viral en el sujeto. Puede administrarse un agente inmunosupresor antes o junto con un agente que reduzca la carga viral en el sujeto. A un sujeto identificado con riesgo de desarrollar PML se le puede administrar un agente que impida el aumento de la carga viral en el sujeto. A un sujeto identificado como de alto riesgo de desarrollar PML se le puede administrar un agente que prevenga el aumento de la carga viral en el sujeto. Puede administrarse un agente inmunosupresor antes o junto con un agente que impida el aumento de la carga viral en el sujeto. El agente que reduce la carga viral en el sujeto o que previene el aumento de la carga viral en el sujeto puede ser, p. ej., un agente dirigido contra el virus de JC. Los agentes ejemplares incluyen anticuerpos, tal como los anticuerpos ampliamente neutralizantes del JCV. Por ejemplo, un agente

puede ser un anticuerpo monoclonal humano ampliamente neutralizante VP-1 específico del poliomavirus de JC (véase, p. ej., Jelcic et al., Science Translational Medicine, vol. 7, número 306, pp. 306ra150 (2015) y Ray et al., Science Translational Medicine, vol. 7, número 306, pp 306ra151 (2015)). Otros agentes ejemplares son los antirretrovirales, el cidofovir, el hexadeciloxipropil-cidofovir (un derivado de éster lipídico), la citarabina (p. ej., arabinósido de citosina), los agentes que bloquean el receptor 5HT2a (p. ej., olanzapina, zisprasidona, mirtazapina, ciproheptadina y risperidona), los inhibidores de la topoisomerasa (p. ej., topotecano) y la mefloquina.

Una composición farmacéutica de la divulgación puede ser administrada a un sujeto en riesgo de desarrollar PML, o a un sujeto que reporte uno o más de los síntomas fisiológicos de PML, aun cuando no pueda haberse hecho una detección de la condición. Una composición farmacéutica de la divulgación puede ser administrada a un sujeto no identificado por tener un riesgo de desarrollar PML, o a un sujeto no identificado por tener uno o más de los síntomas fisiológicos de PML, aunque una detección de la condición no pueda haber sido hecha.

La presente divulgación también incluye kits que pueden utilizarse para tratar una afección en sujetos animales. Estos kits comprenden uno o más medicamentos inmunosupresores y, opcionalmente, instrucciones que enseñan el uso del kit de acuerdo con los diversos métodos y enfoques descritos en el presente documento. Dichos kits también pueden incluir información, tal como referencias bibliográficas científicas, prospectos, resultados de ensayos clínicos y/o resúmenes de éstos y similares, que indiquen o establezcan las actividades y/o ventajas (o riesgos y/o desventajas) del agente. Dicha información puede basarse en los resultados de diversos estudios, p. ej., estudios con animales de experimentación con modelos *in vivo* y estudios basados en ensayos clínicos con seres humanos. Los kits descritos en el presente documento pueden proporcionarse, comercializarse y/o promocionarse a proveedores de servicios sanitarios, incluidos médicos, enfermeros, farmacéuticos, responsables de formularios y similares.

En algunos aspectos, una célula huésped puede utilizarse para probar o administrar compuestos terapéuticos. Una célula huésped puede comprender un ácido nucleico que comprenda secuencias de control de expresión enlazadas operablemente a una región codificante. La célula huésped puede ser natural o no natural. El hospedador no natural utilizado en aspectos del método puede ser cualquier célula capaz de expresar un ácido nucleico de la divulgación, incluyendo, células bacterianas, células fúngicas, células de insectos, células de mamíferos y células vegetales. En algunos aspectos, el hospedador natural es una célula de tejido de mamífero y el hospedador no natural es una célula de tejido de mamífero diferente. Otros aspectos del método incluyen un hospedador natural que es una primera célula que reside normalmente en una primera especie de mamífero y el hospedador no natural es una segunda célula que reside normalmente en una segunda especie de mamífero. En otro aspecto alternativo, el método utiliza una primera célula y una segunda célula que proceden del mismo tipo de tejido. En aquellos aspectos del método en los que la región codificante codifica un polipéptido de mamífero, el polipéptido de mamífero puede ser una hormona. En otros aspectos, la región codificante puede codificar un neuropéptido, un anticuerpo, un antimetabolito o un polipéptido o nucleótido terapéutico.

Las secuencias de control de expresión pueden ser aquellas secuencias de nucleótidos, tanto 5' como 3' a una región codificante, que se requieren para la transcripción y traducción de la región codificante en un organismo huésped. Las secuencias reguladoras incluyen un promotor, un sitio de unión al ribosoma, elementos inducibles opcionales y elementos de secuencia necesarios para un procesamiento en 3' eficaz, incluida la poliadenilación. Cuando el gen estructural se ha aislado del ADN genómico, las secuencias reguladoras también incluyen las secuencias intrónicas necesarias para el empalme de los intrones como parte de la formación del ARNm en el huésped diana.

Formulaciones, vías de administración y dosis eficaces

En un aspecto la invención proporciona un medicamento inmunosupresor para uso en el tratamiento de una condición en un sujeto que necesita terapia con medicamentos inmunosupresores. A continuación, se ofrece una explicación general de las composiciones, formulaciones, vías de administración y dosis eficaces útiles para un medicamento inmunosupresor de este tipo, pero que no forma parte de la materia que se reivindica en el presente documento. Además, el medicamento inmunosupresor para uso de la presente invención implica determinar la ausencia de una variación genómica seleccionada de la chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. A continuación se ofrece una explicación general de las variaciones y, por lo tanto, se incluyen ejemplos que no forman parte de la materia reivindicada en el presente documento, en la medida en que se refieren a variaciones distintas de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. Otro aspecto de la presente divulgación se refiere a formulaciones, vías de administración y dosis eficaces para composiciones farmacéuticas que comprenden un agente o combinación de agentes de la presente

divulgación. Dichas composiciones farmacéuticas pueden utilizarse para tratar una afección (p. ej., esclerosis múltiple) como se ha descrito anteriormente.

Los compuestos de la divulgación pueden administrarse como formulaciones farmacéuticas, incluyendo, aquellas adecuadas para administración oral (incluyendo, bucal y sublingual), rectal, nasal, tópica, parche transdérmico, pulmonar, vaginal, supositorio o parenteral (incluyendo, intramuscular, intraarterial, intratecal, intradérmica, intraperitoneal, subcutánea e intravenosa) o en una forma adecuada para administración por mediante aerosol, inhalación o insuflación. Puede encontrarse información general sobre los sistemas de administración de fármacos en Ansel et al., Pharmaceutical Dosage Forms and Drug Delivery Systems (Lippencott Williams & Wilkins, Baltimore Md. (1999)).

La composición farmacéutica puede incluir vehículos y excipientes (incluidos, entre otros, tampones, carbohidratos, manitol, polipéptidos, aminoácidos, antioxidantes, bacteriostáticos, agentes quelantes, agentes de suspensión, agentes espesantes y/o conservantes), agua, aceites, incluidos los de origen petrolífero, animal, vegetal o sintético, tal como aceite de cacahuete, aceite de soja, aceite mineral, aceite de sésamo y similares, soluciones salinas, soluciones acuosas de dextrosa y glicerol, agentes saborizantes, colorantes, separadores y otros aditivos, adyuvantes o aglutinantes aceptables, aceite de cacahuete, aceite de soja, aceite mineral, aceite de sésamo y similares, soluciones salinas, soluciones acuosas de dextrosa y glicerol, agentes saborizantes, colorantes, separadores y otros aditivos, adyuvantes o aglutinantes aceptables, otras sustancias auxiliares farmacéuticamente aceptables para aproximarse a las condiciones fisiológicas, tal como agentes tamponadores del pH, agentes ajustadores de la tonicidad, agentes emulsionantes, agentes humectantes y similares. Algunos ejemplos de excipientes son almidón, glucosa, lactosa, sacarosa, gelatina, malta, arroz, harina, tiza, gel de sílice, estearato de sodio, monoestearato de glicerol, talco, cloruro de sodio, leche desnatada en polvo, glicerol, propilenglicol, agua, etanol y similares. La preparación farmacéutica puede estar sustancialmente exenta de conservantes. La preparación farmacéutica puede contener al menos un conservante. La metodología general sobre formas farmacéuticas se encuentra en Ansel et al., Pharmaceutical Dosage Forms and Drug Delivery Systems (Lippencott, Williams, & Wilkins, Baltimore Md. (1999)). Puede reconocerse que, si bien puede emplearse cualquier vehículo adecuado conocido por los expertos en la materia para administrar las composiciones de la presente divulgación, el tipo de vehículo puede variar en función del modo de administración.

Los compuestos también pueden encapsularse dentro de liposomas utilizando una tecnología bien conocida. Las microesferas biodegradables también pueden emplearse como vehículos de las composiciones farmacéuticas de la presente divulgación. Las microesferas biodegradables adecuadas se divulgan, p. ej., en las patentes de los Estados Unidos Nos. 4,897,268, 5,075,109, 5,928,647, 5,811,128, 5,820,883, 5,853,763, 5,814,344 y 5,942,252.

El compuesto puede administrarse en liposomas o microesferas (o micropartículas). Los métodos de preparación de liposomas y microesferas para su administración a un sujeto son bien conocidos por los expertos en la materia. La patente de los Estados Unidos No. 4,789,734, describe métodos para encapsular materiales biológicos en liposomas. Esencialmente, el material se disuelve en una solución acuosa, se añaden los fosfolípidos y lípidos adecuados, junto con tensioactivos si es necesario, y el material se dializa o se somete a ultrasonido, según sea necesario. Una revisión de métodos conocidos es proporcionada por G. Gregoriadis, Capítulo 14, "Liposomes," Drug Carriers in Biology and Medicine, pp. 2.sup. 87-341 (Academic Press, 1979).

Las microesferas formadas por polímeros o polipéptidos son bien conocidas por los expertos en la materia, y pueden adaptarse para su paso a través del tracto gastrointestinal directamente al torrente sanguíneo. Alternativamente, el compuesto puede incorporarse y las microesferas, o el compuesto de microesferas, implantarse para una liberación lenta durante un periodo de tiempo que puede variar de días a meses. Véase, p. ej., las patentes de los Estados Unidos Nos. 4,906,474, 4,925,673 y 3,625,214 y Jein, TIPS 19:155-157 (1998).

La concentración del fármaco puede ajustarse, el pH de la solución puede tamponarse y la isotonicidad puede ajustarse para que sea compatible con la inyección intravenosa, tal como es bien sabido en la técnica.

Los compuestos de la divulgación pueden formularse como una solución o suspensión estéril, en vehículos adecuados, bien conocidos en la técnica. Las composiciones farmacéuticas pueden esterilizarse mediante técnicas de esterilización convencionales bien conocidas, o pueden filtrarse de forma estéril. Las soluciones acuosas resultantes pueden envasarse para su uso tal cual, o liofilizarse, combinándose la preparación liofilizada con una solución estéril antes de su administración. Las formulaciones adecuadas y los vehículos adicionales se describen en Remington "The Science and Practice of Pharmacy" (20th Ed., Lippincott Williams & Wilkins, Baltimore MD).

Los agentes o sus sales farmacéuticamente aceptables pueden proporcionarse solos o en combinación con uno o más agentes o con una o más formas. Por ejemplo, una formulación puede comprender uno o más agentes en proporciones particulares, dependiendo de las potencias relativas de cada agente y de la indicación prevista. Por ejemplo, en composiciones dirigidas a dos dianas hospedadoras diferentes, y en las que las

potencias son similares, puede utilizarse aproximadamente una proporción 1:1 de agentes. Las dos formas pueden formularse juntas, en la misma unidad de dosificación, p. ej., en una crema, supositorio, comprimido, cápsula, aerosol o paquete de polvo para disolver en una bebida; o cada forma puede formularse en una unidad separada, p. ej., dos cremas, dos supositorios, dos comprimidos, dos cápsulas, un comprimido y un líquido para disolver el comprimido, dos aerosoles, o un paquete de polvo y un líquido para disolver el polvo, etc.

El término "sal farmacéuticamente aceptable" significa aquellas sales que retienen la eficacia biológica y las propiedades de los agentes utilizados en la presente divulgación, y que no son biológicamente o de otro modo indeseables.

Las sales típicas son las de los iones inorgánicos, tales como, p. ej., iones de sodio, potasio, calcio, magnesio y similares. Tales sales incluyen sales con ácidos inorgánicos u orgánicos, tal como ácido clorhídrico, ácido bromhídrico, ácido fosfórico, ácido nítrico, ácido sulfúrico, ácido metanosulfónico, ácido p toluenosulfónico, ácido acético, ácido fumárico, ácido succínico, ácido láctico, ácido mandélico, ácido málico, ácido cítrico, ácido tartárico o ácido maleico. Además, si el agente o agentes contienen un grupo carboxilo u otro grupo ácido, puede convertirse en una sal de adición farmacéuticamente aceptable con bases inorgánicas u orgánicas. Ejemplos de bases adecuadas incluyen hidróxido de sodio, hidróxido de potasio, amoníaco, ciclohexilamina, diciclohexilamina, etanolamina, dietanolamina, trietanolamina y similares.

Un éster o amida farmacéuticamente aceptable se refiere a aquellos que retienen la eficacia biológica y las propiedades de los agentes utilizados en la presente divulgación, y que no son biológicamente o de otro modo indeseables. Los ésteres típicos incluyen etilo, metilo, isobutilo, etilenglicol y similares. Las amidas típicas incluyen amidas no sustituidas, alquilamidas, dialquilamidas y similares.

Un agente puede administrarse en combinación con uno o más compuestos, formas y/o agentes, p. ej., tal como se ha descrito anteriormente. Las composiciones farmacéuticas con otro u otros agentes activos pueden formularse para comprender determinadas proporciones molares. Por ejemplo, pueden utilizarse proporciones molares de aproximadamente 99:1 a aproximadamente 1:99 de un primer agente activo con respecto al otro agente activo. El intervalo de proporciones molares de un primer agente activo: otros agentes activos puede seleccionarse entre aproximadamente 80:20 y aproximadamente 20:80; aproximadamente 75:25 y a aproximadamente 25:75, a aproximadamente 70:30 y aproximadamente 30:70, a aproximadamente 66:33 y aproximadamente 33:66, aproximadamente 60:40 y aproximadamente 40:60; aproximadamente 50:50 y aproximadamente 90: 10 a aproximadamente 10:90. La proporción molar de un primer activo: otros agentes activos puede ser de aproximadamente 1:9, o puede ser de aproximadamente 1:1. Los dos agentes, formas y/o compuestos pueden formularse juntos, en la misma unidad de dosificación, p. ej., en una crema, supositorio, comprimido, cápsula o paquete de polvo para disolver en una bebida; o cada agente, forma y/o compuesto puede formularse en unidades separadas, p. ej., dos cremas, supositorios, comprimidos, dos cápsulas, un comprimido y un líquido para disolver el comprimido, un aerosol, un paquete de polvo y un líquido para disolver el polvo, etc.

Si es necesario o deseable, los agentes y/o combinaciones de agentes pueden administrarse con aún otros agentes. La elección de los agentes que pueden coadministrarse con los agentes y/o combinaciones de los agentes de la presente divulgación puede depender, al menos en parte, de la afección que se esté tratando. Los agentes de uso particular en las formulaciones de la presente divulgación incluyen, p. ej., cualquier agente que tenga un efecto terapéutico para una infección viral, incluyendo, p. ej., fármacos utilizados para tratar afecciones inflamatorias. Por ejemplo, en tratamientos para la gripe, las formulaciones de la presente divulgación pueden contener adicionalmente uno o más fármacos antiinflamatorios convencionales, tal como un NSAID, p. ej., ibuprofeno, naproxeno, acetaminofén, ketoprofeno o aspirina. Alternativamente, para el tratamiento de la gripe, las formulaciones de la presente divulgación pueden contener adicionalmente uno o más agentes antivirales de la gripe convencionales, tal como amantadina, rimantadina, zanamivir y oseltamivir. En los tratamientos de infecciones retrovirales, tal como el VIH, las formulaciones de la presente divulgación pueden contener además uno o más fármacos antivirales convencionales, tal como los inhibidores de la proteasa (lopinavir/ritonavir {p. ej., KALETRA}, indinavir {p. ej., CRIXIVAN}, ritonavir {p. ej., NORVIR}, nelfinavir {p. ej., VIRACEPT}, saquinavir en cápsulas de gelatina dura {p. ej., INVIRASE}, atazanavir {p. ej., REYATAZ}, amprenavir {p. ej., AGENERASE}, fosamprenavir {p. ej., TELZIR}, tipranavir {p. ej., AITPVUS}), inhibidores de la transcriptasa inversa, incluidos los inhibidores no nucleósidos y nucleósidos/nucleótidos (AZT {zidovudina, p. ej., Retrovir}, ddI {didanosina, p. ej., VIDEX}, 3TC {lamivudina, p. ej., EPIVIR}, d4T {stavudina, p. ej., ZERIT}, abacavir {p. ej., ZIAGEN}, FTC {emtricitabina, p. ej., EMTRIVA}, tenofovir {p. ej., VIREAD}, efavirenz {p.ej., SUSTIVA} y nevirapina {p. ej., VIRAMUNE}), inhibidores de la fusión T20 {enfuvirtida, p. ej., FUZEON}, inhibidores de la integrasa (Raltegravir, p. ej., ISENTRESS, MK-0518; y elvitegravir, p. ej., VITEKTA, GS-9137), e inhibidores de la maduración (bevirimat {PA-457}). Como otro ejemplo, las formulaciones pueden contener adicionalmente uno o más suplementos, tal como vitamina C, E u otros antioxidantes.

El(los) agente(s) (o sales, ésteres o amidas farmacéuticamente aceptables de los mismos) puede(n) administrarse per se o en forma de una composición farmacéutica en la que el(los) agente(s) activo(s) se encuentra(n) en una mezcla con uno o más vehículos farmacéuticamente aceptables. Una composición farmacéutica, tal como se utiliza en el presente documento, puede ser cualquier composición preparada para

su administración a un sujeto. Las composiciones farmacéuticas para uso de acuerdo con la presente divulgación pueden formularse de manera convencional utilizando uno o más vehículos fisiológicamente aceptables, que comprenden excipientes, diluyentes y/o auxiliares, p. ej., que facilitan la transformación de los agentes activos en preparaciones que pueden administrarse. Una formulación adecuada puede depender, al menos en parte, de la vía de administración elegida. El (los) agente(s) útil(es) en la presente divulgación, o las sales, ésteres o amidas farmacéuticamente aceptables de los mismos, pueden administrarse a un sujeto utilizando diversas vías o modos de administración, incluyendo, aplicaciones orales, bucales, tópicas, rectales, transdérmicas, transmucosas, subcutáneas, intravenosas e intramusculares, así como por inhalación.

Para administración oral, los agentes pueden formularse fácilmente combinando el(los) agente(s) activo(s) con vehículos farmacéuticamente aceptables bien conocidos en la técnica. Tales vehículos permiten que los agentes de la divulgación se formulen como comprimidos, incluidos comprimidos masticables, píldoras, grageas, cápsulas, pastillas, caramelos duros, líquidos, geles, jarabes, lechadas, polvos, suspensiones, elixires, obleas y similares, para su ingestión oral por un sujeto a tratar. Tales formulaciones pueden comprender vehículos farmacéuticamente aceptables, incluidos diluyentes sólidos o rellenos, medios acuosos estériles y diversos disolventes orgánicos no tóxicos. Un vehículo sólido puede ser una o más sustancias que también pueden actuar como diluyentes, agentes saborizantes, solubilizantes, lubricantes, agentes de suspensión, aglutinantes, conservantes, agentes desintegradores de comprimidos o un material encapsulante. En los polvos, el vehículo suele ser un sólido finamente dividido que se mezcla con el componente activo finamente dividido. En los comprimidos, el componente activo se mezcla generalmente con el vehículo que tiene la capacidad aglutinante necesaria en proporciones adecuadas y se compacta en la forma y tamaño deseados. Los polvos y comprimidos contienen preferentemente entre aproximadamente uno (1) y aproximadamente setenta (70) por ciento del compuesto activo. Los vehículos adecuados incluyen, entre otros, carbonato de magnesio, estearato de magnesio, talco, azúcar, lactosa, pectina, dextrina, almidón, gelatina, tragacanto, metilcelulosa, carboximetilcelulosa sódica, cera de baja fusión, manteca de cacao y similares. Generalmente, los agentes de la divulgación pueden incluirse en niveles de concentración que van desde aproximadamente 0.5 %, aproximadamente 5 %, aproximadamente 10 %, aproximadamente 20 % o aproximadamente 30 % hasta aproximadamente 50 %, aproximadamente 60 %, aproximadamente 70 %, aproximadamente 80 % o aproximadamente 90 % en peso de la composición total de las formas de dosificación oral, en una cantidad suficiente para proporcionar una unidad de dosificación deseada.

Las suspensiones acuosas para uso oral pueden contener agente(s) de esta divulgación con excipientes farmacéuticamente aceptables, tal como un agente de suspensión (p. ej., metilcelulosa), un agente humectante (p. ej., lecitina, lisolecitina y/o un alcohol graso de cadena larga), así como agentes colorantes, conservantes, agentes saborizantes y similares.

Pueden utilizarse aceites o disolventes no acuosos para disolver los agentes, debido, p. ej., a la presencia de grandes moléculas lipofílicas. Como alternativa, pueden utilizarse emulsiones, suspensiones u otros preparados, p. ej., preparados liposomales. Con respecto a las preparaciones liposomales, puede utilizarse cualquier método conocido para preparar liposomas para el tratamiento de una afección. Véase, p. ej., Bangham et al., J. Mol. Biol. 23: 238-252 (1965) y Szoka et al., Proc. Natl Acad. Sci. USA 75: 4194-4198 (1978). También se pueden unir ligandos a los liposomas para dirigir estas composiciones a determinados lugares de acción. Los agentes de esta divulgación también pueden integrarse en productos alimenticios, p. ej., queso crema, mantequilla, aderezo para ensaladas o helado para facilitar la solubilización, administración y/o cumplimiento en ciertas poblaciones de sujetos.

Las preparaciones farmacéuticas para uso oral pueden obtenerse como excipiente sólido, moliendo opcionalmente una mezcla resultante, y procesando la mezcla de gránulos, tras añadir auxiliares adecuados, si se desea, para obtener comprimidos o núcleos de grageas. Los excipientes adecuados son, en particular, rellenos como azúcares, incluyendo, lactosa, sacarosa, manitol o sorbitol; elementos saborizantes, preparados de celulosa como, p. ej., almidón de maíz, almidón de trigo, almidón de arroz, almidón de patata, gelatina, goma tragacanto, metilcelulosa, hidroxipropilmetilcelulosa, carboximetilcelulosa sódica y/o polivinilpirrolidona (PVP). Si se desea, pueden añadirse agentes desintegrantes, tal como la polivinilpirrolidona reticulada, el agar o el ácido algínico o una sal del mismo, tal como el alginato sódico. Los agentes también pueden formularse como preparados de liberación sostenida.

Los núcleos de grageas pueden estar provistos de revestimientos adecuados. Para ello, pueden utilizarse soluciones concentradas de azúcar, que pueden contener opcionalmente goma arábiga, talco, polivinilpirrolidona, gel de carbopol, polietilenglicol y/o dióxido de titanio, soluciones de laca y disolventes orgánicos o mezclas de disolventes adecuados. Pueden añadirse colorantes o pigmentos a los comprimidos o recubrimientos de grageas para su identificación o para caracterizar diferentes combinaciones de agentes activos.

Las preparaciones farmacéuticas que pueden usarse por vía oral incluyen cápsulas de gelatina con cierre a presión, así como cápsulas blandas y selladas hechas de gelatina y un plastificante, tal como glicerol o sorbitol. Las cápsulas encajables pueden contener los principios activos mezclados con rellenos como lactosa, aglutinantes como almidón y/o lubricantes como talco o estearato de magnesio y, opcionalmente, estabilizantes.

En las cápsulas blandas, los agentes activos pueden disolverse o suspenderse en líquidos adecuados, tal como aceites grasos, parafina líquida o polietilenglicoles líquidos. Además, se pueden añadir estabilizantes. Todas las formulaciones para administración oral deben estar en dosis adecuadas para su administración.

Otras formas adecuadas para la administración oral incluyen preparaciones en forma líquida incluyendo, emulsiones, jarabes, elixires, soluciones acuosas, suspensiones acuosas, o preparaciones en forma sólida que están destinadas a ser convertidas poco antes de su uso en preparaciones en forma líquida. Las emulsiones pueden prepararse en soluciones, p. ej., en soluciones acuosas de propilenglicol o pueden contener agentes emulsionantes, p. ej., tal como lecitina, monooleato de sorbitán o acacia. Las soluciones acuosas pueden prepararse disolviendo el componente activo en agua y añadiendo colorantes, saborizantes, estabilizantes y espesantes adecuados. Las suspensiones acuosas pueden prepararse dispersando el componente activo finamente dividido en agua con material viscoso, tal como gomas naturales o sintéticas, resinas, metilcelulosa, carboximetilcelulosa sódica y otros agentes de suspensión bien conocidos. Los rellenos o vehículos adecuados con los que pueden administrarse las composiciones incluyen agar, alcohol, grasas, lactosa, almidón, derivados de la celulosa, polisacáridos, polivinilpirrolidona, sílice, solución salina estéril y similares, o mezclas de los mismos utilizadas en cantidades adecuadas. Los preparados en forma sólida incluyen soluciones, suspensiones y emulsiones, y pueden contener, además del componente activo, colorantes, saborizantes, estabilizantes, tampones, edulcorantes artificiales y naturales, dispersantes, espesantes, agentes solubilizantes y similares.

Se puede elaborar un jarabe o suspensión añadiendo el compuesto activo a una solución acuosa concentrada de un azúcar, p. ej., sacarosa, a la que también se puede añadir cualquier ingrediente accesorio. Tales ingredientes accesorios pueden incluir saborizantes, un agente para retardar la cristalización del azúcar o un agente para aumentar la solubilidad de cualquier otro ingrediente, p. ej., tal como un alcohol polihídrico, p. ej., glicerol o sorbitol.

Cuando se formulan compuestos de la divulgación para administración oral, puede ser deseable utilizar formulaciones retenedoras en el sistema gástrico para mejorar la absorción desde el tracto gastrointestinal (GI). Una formulación que se retiene en el estómago durante varias horas puede liberar los compuestos de la divulgación lentamente y proporcionar una liberación sostenida que puede ser preferible en algunos casos. La divulgación de tales formulaciones retenedoras en el sistema gástrico se encuentra en Klausner E.A., et al., Pharm. Res. 20, 1466-73 (2003); Hoffman, A. et al., Int. J. Pharm. 11, 141-53 (2004), Streubel, A., et al. Expert Opin. Drug Deliver. 3, 217-3 y Chavanpatil, M.D. et al., Int. J. Pharm. (2006). Pueden utilizarse técnicas expandibles, flotantes y bioadhesivas para maximizar la absorción de los compuestos de la divulgación.

Los compuestos de la divulgación pueden formularse para administración parenteral (p. ej., mediante inyección, p. ej., inyección en bolo o infusión continua) y pueden presentarse en forma de dosis unitaria en ampollas, jeringas precargadas, infusión de pequeño volumen o en envases multidosis con un conservante añadido. Las composiciones pueden adoptar formas tales como suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos, p. ej., soluciones en polietilenglicol acuoso.

Para formulaciones inyectables, el vehículo puede elegirse entre los conocidos en el arte por ser adecuados, incluyendo, soluciones acuosas o suspensiones oleosas, o emulsiones, con aceite de sésamo, aceite de maíz, aceite de semilla de algodón, o aceite de cacahuete, así como elixires, manitol, dextrosa, o una solución acuosa estéril, y vehículos farmacéuticos similares. La formulación también puede incluir composiciones poliméricas biocompatibles y biodegradables, tal como el ácido poli(láctico-co-glicólico). Estos materiales pueden fabricarse en microesferas o nanoesferas, cargarse con el fármaco y recubrirse o derivatizarse para proporcionar un rendimiento superior de liberación sostenida. Los vehículos adecuados para la inyección periocular o intraocular incluyen, p. ej., suspensiones del agente terapéutico en agua para inyección, liposomas y vehículos adecuados para sustancias lipofílicas. Otros vehículos para la inyección periocular o intraocular son bien conocidos en la técnica.

La composición puede formularse de acuerdo con procedimientos rutinarios como una composición farmacéutica adaptada para la administración intravenosa a seres humanos. Típicamente, las composiciones para administración intravenosa son soluciones en tampón acuoso isotónico estéril. Cuando sea necesario, la composición también puede incluir un agente solubilizante y un anestésico local como la lidocaína para aliviar el dolor en el lugar de la inyección. Por lo general, los ingredientes se suministran por separado o mezclados entre sí en forma de dosificación unitaria, p. ej., tal como polvo seco liofilizado o concentrado sin agua en un envase herméticamente cerrado, tal como una ampolla o sobre que indique la cantidad de agente activo. Cuando la composición vaya a administrarse por infusión, puede dispensarse con un frasco de infusión que contenga agua o solución salina estéril de calidad farmacéutica. Cuando la composición se administra por inyección, puede proporcionarse una ampolla de agua estéril para inyección o solución salina para que los ingredientes puedan mezclarse antes de la administración.

Cuando la administración es por inyección, el compuesto activo puede formularse en soluciones acuosas, específicamente en tampones fisiológicamente compatibles como la solución de Hank, la solución de Ringer o el tampón salino fisiológico. La solución puede contener agentes formuladores tal como agentes de suspensión,

estabilización y/o dispersión. Alternativamente, el compuesto activo puede estar en forma de polvo para su constitución con un vehículo adecuado, p. ej., agua estéril libre de pirógenos, antes de su uso. La composición farmacéutica puede no comprender un adyuvante o cualquier otra sustancia añadida para potenciar la respuesta inmunitaria estimulada por el péptido. La composición farmacéutica puede comprender una sustancia que inhiba una respuesta inmunitaria al péptido. Los métodos de formulación son conocidos en la técnica, p. ej., los descritos en Remington's Pharmaceutical Sciences, última edición, Mack Publishing Co., Easton P.

Además de las formulaciones descritas anteriormente, los agentes también pueden formularse como una preparación de depósito. Estas formulaciones de acción prolongada pueden administrarse mediante implantación o administración transcutánea (p. ej., subcutánea o intramuscular), inyección intramuscular o uso de un parche transdérmico. De este modo, p. ej., los agentes pueden formularse con materiales poliméricos o hidrófobos adecuados (p. ej., tal como una emulsión en un aceite aceptable) o resinas de intercambio iónico, o como derivados poco solubles, p. ej., tal como una sal poco soluble.

Las composiciones farmacéuticas que comprenden uno o más agentes de la presente divulgación pueden ejercer efectos locales y regionales cuando se administran tópicamente o se inyectan en o cerca de sitios particulares de infección. La aplicación tópica directa, p. ej., de un líquido viscoso, solución, suspensión, soluciones a base de dimetilsulfóxido (DMSO), formulaciones liposomales, gel, jalea, crema, loción, pomada, supositorio, espuma o aerosol, puede utilizarse para la administración local, para producir, p. ej., efectos locales y/o regionales. Los vehículos farmacéuticamente apropiados para dicha formulación incluyen, p. ej., alcoholes alifáticos inferiores, poliglicoles (p. ej., glicerol o polietilenglicol), ésteres de ácidos grasos, aceites, grasas, siliconas y similares. Dichos preparados también pueden incluir conservantes (p. ej., ésteres del ácido p-hidroxibenzoico) y/o antioxidantes (p. ej., ácido ascórbico y tocoferol). Véase también Dermatological Formulations: Percutaneous absortion, Barry (Ed.), Marcel Dekker Inc., 1983.

Las composiciones farmacéuticas de la presente divulgación pueden contener un vehículo cosmético o dermatológicamente aceptable. Dichos vehículos son compatibles con la piel, las uñas, las membranas mucosas, los tejidos y/o el cabello, y pueden incluir cualquier vehículo cosmético o dermatológico de uso convencional que cumpla estos requisitos. Tales vehículos pueden ser fácilmente seleccionados por un experto en la materia. En la formulación de ungüentos para la piel, un agente o combinación de agentes de la presente divulgación puede formularse en una base de hidrocarburo oleaginoso, una base de absorción anhidra, una base de absorción de agua en aceite, una base de aceite en agua removible en agua y/o una base soluble en agua. Ejemplos de tales vehículos y excipientes incluyen, pero no se limitan a, humectantes (p. ej., urea), glicoles (p. ej., propilenglicol), alcoholes (p. ej., etanol), ácidos grasos (p. ej., ácido oleico), tensioactivos (p. ej., miristato de isopropilo y lauril sulfato de sodio), pirrolidonas, monolaurato de glicerol, sulfóxidos, terpenos (p. ej., mentol), aminas, amidas, alcanos, alcanoles, agua, carbonato cálcico, fosfato cálcico, diversos azúcares, almidones, derivados de la celulosa, gelatina y polímeros como los polietilenglicoles.

Las pomadas y cremas pueden, p. ej., formularse con una base acuosa o aceitosa con la adición de agentes espesantes y/o gelificantes adecuados. Las lociones pueden formularse con una base acuosa u oleosa y, en general, también pueden contener uno o más agentes emulsionantes, agentes estabilizantes, agentes dispersantes, agentes de suspensión, agentes espesantes o agentes colorantes. La construcción y el uso de parches transdérmicos para la administración de agentes farmacéuticos son bien conocidos en la técnica. Véase, p. ej., las patentes de los Estados Unidos Nos. 5,023,252, 4,992,445 y 5,001,139. Dichos parches pueden fabricarse para la administración continua, pulsátil o por demanda de agentes farmacéuticos.

Los lubricantes que pueden usarse para formar composiciones farmacéuticas y formas de dosificación de la divulgación incluyen, pero no se limitan a, estearato de calcio, estearato de magnesio, aceite mineral, aceite mineral ligero, glicerina, sorbitol, manitol, polietilenglicol, otros glicoles, ácido esteárico, lauril sulfato de sodio, talco, aceite vegetal hidrogenado (p. ej., aceite de cacahuete, aceite de semilla de algodón, aceite de girasol, aceite de sésamo, aceite de oliva, aceite de maíz y aceite de soja), estearato de zinc, oleato de etilo, laurato de etilo, agar o mezclas de los mismos. Entre los lubricantes adicionales se incluyen, p. ej., un gel de sílice siloide, un aerosol coagulado de sílice sintética o mezclas de los mismos. Opcionalmente, puede añadirse un lubricante, en una cantidad inferior a aproximadamente el 1 por ciento en peso de la composición farmacéutica.

Las composiciones de acuerdo con la presente divulgación pueden presentarse en cualquier forma adecuada para la aplicación tópica, incluyendo, soluciones acuosas, acuosas-alcohólicas u oleosas, dispersiones en loción o suero, geles acuosos, anhidros u oleosos, emulsiones obtenidas por dispersión de una fase grasa en una fase acuosa (O/W o aceite en agua) o, a la inversa, (W/O o agua en aceite), microemulsiones o alternativamente microcápsulas, micropartículas o dispersiones de vesículas lipídicas de tipo iónico y/o no iónico. Estas composiciones pueden prepararse de acuerdo con métodos convencionales. Aparte de los agentes de la divulgación, las cantidades de los diversos constituyentes de las composiciones de acuerdo con la divulgación son las utilizadas convencionalmente en la técnica. Estas composiciones constituyen, en particular, cremas, leches, lociones, geles o espumas de protección, tratamiento o cuidado para la cara, las manos, el cuerpo y/o las mucosas, o para la limpieza de la piel. Las composiciones también pueden consistir en preparados sólidos que constituyen jabones o barras limpiadoras.

Las composiciones de la presente divulgación también pueden contener adyuvantes comunes en los campos cosmético y dermatológico, tales como agentes gelificantes hidrofílicos o lipofílicos, agentes activos hidrofílicos o lipofílicos, agentes conservantes, antioxidantes, disolventes, fragancias, rellenos, filtros solares, absorbentes de olores y colorantes. Las cantidades de estos diversos adyuvantes son las utilizadas convencionalmente en los campos considerados y, p. ej., oscilan entre aproximadamente el 0.01 % y aproximadamente el 20 % del peso total de la composición. Según su naturaleza, estos adyuvantes pueden introducirse en la fase grasa, en la fase acuosa y/o en las vesículas lipídicas.

Las infecciones virales oculares pueden tratarse eficazmente con soluciones oftálmicas, suspensiones, ungüentos o insertos que comprenden un agente o combinación de agentes de la presente divulgación. Los colirios pueden prepararse disolviendo el principio activo en una solución acuosa estéril, tal como solución salina fisiológica, solución amortiguadora, etc., o combinando composiciones en polvo que se disolverán antes de su uso. Pueden elegirse otros vehículos, tal como es conocido en la técnica, incluyendo, pero no limitándose a: solución salina de equilibrio, solución salina, poliéteres solubles en agua como el polietilenglicol, polivinilos, tal como el alcohol polivinílico y la povidona, derivados de la celulosa como la metilcelulosa y la hidroxipropilmetilcelulosa, derivados del petróleo, tal como el aceite mineral y la vaselina blanca, grasas animales, tal como la lanolina, polímeros de ácido acrílico, tal como el gel de carboxipolimetileno, grasas vegetales, tal como el aceite de cacahuete, y polisacáridos, tal como los dextranos, y glicosaminoglicanos, tal como el hialuronato sódico. Si se desea, pueden añadirse aditivos utilizados habitualmente en los colirios. Dichos aditivos incluyen agentes isotonizantes (p. ej., cloruro de sodio, etc.), agente tampón (p. ej., ácido bórico, monohidrógeno fosfato de sodio, dihidrógeno fosfato de sodio, etc.), conservantes (p. ej., cloruro de benzalconio, cloruro de bencetonio, clorobutanol, etc.), espesantes (p. ej., sacárido como lactosa, manitol, maltosa, etc.; p. ej., ácido hialurónico o su sal tal como hialuronato sódico, hialuronato potásico, etc.; p. ej., mucopolisacárido tal como el sulfato de condroitina, etc.; p. ej., poliacrilato sódico, polímero de carboxivinilo, poliacrilato reticulado, alcohol polivinílico, polivinilpirrolidona, metilcelulosa, hidroxipropilmetilcelulosa, hidroxietilcelulosa, carboximetilcelulosa, hidroxipropilcelulosa u otros agentes conocidos por los expertos en la materia).

La solubilidad de los componentes de las presentes composiciones puede mejorarse mediante un tensioactivo u otro cosolvente apropiado en la composición. Tales cosolventes incluyen polisorbato 20, 60 y 80, Pluronic F68, F-84 y P-103, ciclodextrina u otros agentes conocidos por los expertos en la materia. Tales cosolventes pueden emplearse a un nivel de aproximadamente 0.01 % a 2 % en peso.

Las composiciones de la divulgación pueden envasarse en forma multidosis. Pueden preferirse conservantes para evitar la contaminación microbiana durante el uso. Entre los conservantes adecuados se incluyen: cloruro de benzalconio, timerosal, clorobutanol, metilparabeno, propilparabeno, alcohol feniletílico, edetato disódico, ácido sórbico, Onamer M u otros agentes conocidos por los expertos en la materia. En los productos oftálmicos de la técnica anterior, dichos conservantes pueden emplearse a un nivel del 0.004 % al 0.02 %. En las composiciones de la presente solicitud, el conservante, preferentemente cloruro de benzalconio, puede emplearse a un nivel del 0.001 % a menos del 0.01 %, p. ej., del 0.001 % al 0.008 %, preferentemente aproximadamente del 0.005 % en peso. Se ha descubierto que una concentración de cloruro de benzalconio del 0.005 % puede ser suficiente para preservar las composiciones de la presente divulgación del ataque microbiano.

Los agentes de la presente divulgación pueden administrarse en forma soluble en lugar de suspensión, lo que permite una absorción más rápida y cuantitativa a los sitios de acción. En general, las formulaciones como jaleas, cremas, lociones, supositorios y ungüentos pueden proporcionar un área con una exposición más prolongada a los agentes de la presente divulgación, mientras que las formulaciones en solución, p. ej., aerosoles, proporcionan una exposición más inmediata y a corto plazo.

Cuando se relacionan con la aplicación tópica/local, las composiciones farmacéuticas pueden incluir uno o más potenciadores de la penetración. Por ejemplo, las formulaciones pueden comprender vehículos o excipientes adecuados en fase sólida o gel que aumenten la penetración o ayuden al suministro de agentes o combinaciones de agentes de la divulgación a través de una barrera de permeabilidad, p. ej., la piel. Muchos de estos compuestos potenciadores de la penetración son conocidos en la técnica de la formulación tópica, e incluyen, p. ej., agua, alcoholes (p. ej., terpenos como metanol, etanol, 2-propanol), sulfóxidos (p. ej., dimetilsulfóxido, decilmetilsulfóxido, tetradecilmetilsulfóxido), pirrolidonas (p. ej., 2-pirrolidona, N-metil-2-pirrolidona, N-(2-hidroxietil)pirrolidona), laurocapram, acetona, dimetilacetamida, dimetilformamida, alcohol tetrahidrofurfurílico, L-α-aminoácidos, tensioactivos aniónicos, catiónicos, anfóteros o no iónicos (p. ej., miristato de isopropilo y lauril sulfato de sodio), ácidos grasos, alcoholes grasos (p. ej., ácido oleico), aminas, amidas, amidas de ácido clofíbrico, lauramida de hexametileno, enzimas proteolíticas, α-bisabolol, d-limoneno, urea y N,N-dietil-m-toluamida, y similares. Otros ejemplos incluyen humectantes (p. ej., urea), glicoles (p. ej., propilenglicol y polietilenglicol), monolaurato de glicerol, alcanos, alcanoles, ORGELASA, carbonato cálcico, fosfato cálcico, diversos azúcares, almidones, derivados de la celulosa, gelatina y/u otros polímeros. Las composiciones farmacéuticas pueden incluir uno o más de estos potenciadores de la penetración.

Las composiciones farmacéuticas para aplicación local/tópica pueden incluir uno o más conservantes antimicrobianos como compuestos de amonio cuaternario, mercuriales orgánicos, p-hidroxi benzoatos, alcoholes aromáticos, clorobutanol y similares.

Las composiciones farmacéuticas pueden ser soluciones, suspensiones, ungüentos, enemas y/o supositorios administrados por vía oral o rectal que comprenden un agente o combinación de agentes de la presente divulgación.

Las composiciones farmacéuticas pueden ser soluciones en aerosol, suspensiones o polvos secos que comprenden un agente o combinación de agentes de la presente divulgación. El aerosol puede administrarse a través del sistema respiratorio o los pasajes nasales. Por ejemplo, un experto en la materia puede reconocer que una composición de la presente divulgación puede suspenderse o disolverse en un vehículo apropiado, p. ej., un propelente farmacéuticamente aceptable, y administrarse directamente en los pulmones mediante un aerosol nasal o inhalante. Por ejemplo, una formulación en aerosol que comprende un agente puede disolverse, suspenderse o emulsionarse en un propelente o una mezcla de disolvente y propelente, p. ej., para su administración como aerosol nasal o inhalante. Las formulaciones en aerosol pueden contener cualquier propelente aceptable bajo presión, tal como un propelente cosmético o dermatológico o farmacéuticamente aceptable, tal como se utiliza convencionalmente en la técnica.

Una formulación en aerosol para administración nasal es generalmente una solución acuosa diseñada para ser administrada a los conductos nasales en gotas o aerosoles. Las soluciones nasales pueden ser similares a las secreciones nasales en el sentido de que generalmente son isotónicas y están ligeramente tamponadas para mantener un pH de entre 5.5 y aproximadamente 6.5, aunque también pueden utilizarse valores de pH fuera de este intervalo. También pueden incluirse en la formulación agentes antimicrobianos o conservantes.

Una formulación en aerosol para inhalaciones e inhalantes puede diseñarse de modo que el agente o la combinación de agentes de la presente divulgación sea transportado al árbol respiratorio del sujeto cuando se administra por vía respiratoria nasal u oral. Las soluciones para inhalación pueden administrarse, p. ej., mediante un nebulizador. Las inhalaciones o insuflaciones, que comprenden fármacos finamente pulverizados o líquidos, pueden administrarse al sistema respiratorio como un aerosol farmacéutico de una solución o suspensión del agente o combinación de agentes en un propelente, p. ej., para ayudar a su disposición. Los propelentes pueden ser gases licuados, incluidos los compuestos halocarbonados, p. ej., los fluorocarbonados, tal como los hidrocarburos clorados fluorados, los hidroclorofluorocarbonados y los hidroclorocarbonados, así como los hidrocarburos y los éteres de hidrocarburos.

Los propelentes halocarbonados útiles en la presente divulgación incluyen propelentes fluorocarbonados en los que todos los hidrógenos se sustituyen por flúor, propelentes clorofluorocarbonados en los que todos los hidrógenos se sustituyen por cloro y al menos un flúor, propelentes fluorocarbonados que contienen hidrógeno y propelentes clorofluorocarbonados que contienen hidrógeno. Los propelentes de halocarbonados se describen en Johnson, patentes de los Estados Unidos No. 5,376,359; Byron et al., patente de los Estados Unidos No. 5,190,029 y Purewal et al., patente de los Estados Unidos No. 5,776,434. Los propelentes de hidrocarburos útiles en la divulgación incluyen, p. ej., propano, isobutano, n-butano, pentano, isopentano y neopentano. También puede utilizarse una mezcla de hidrocarburos como propelente. Los propelentes de éter incluyen, p. ej., el dimetil éter, así como los éteres. Una formulación en aerosol de la divulgación también puede comprender más de un propelente. Por ejemplo, la formulación del aerosol puede comprender más de un propelente de la misma clase, tal como dos o más fluorocarbonados; o más de uno, más de dos, más de tres propelentes de clases diferentes, tal como un fluorohidrocarbonados y un compuesto hidrocarbonado. Las composiciones farmacéuticas de la presente divulgación también pueden dispensarse con un gas comprimido, p. ej., un gas inerte como el dióxido de carbono, el óxido nitroso o el nitrógeno.

Las formulaciones en aerosol también pueden incluir otros componentes, p. ej., etanol, isopropanol, propilenglicol, así como tensioactivos u otros componentes como aceites y detergentes. Estos componentes pueden servir para estabilizar la formulación y/o lubricar los componentes de la válvula.

La formulación en aerosol puede envasarse a presión y puede formularse como aerosol utilizando soluciones, suspensiones, emulsiones, polvos y preparaciones semisólidas. Por ejemplo, una formulación de aerosol en solución puede comprender una solución de un agente de la divulgación en propelente (sustancialmente) puro o como mezcla de propelente y disolvente. El disolvente puede utilizarse para disolver el agente y/o retardar la evaporación del propelente. Los disolventes útiles en la divulgación incluyen, p. ej., agua, etanol y glicoles. Puede utilizarse cualquier combinación de disolventes adecuados, opcionalmente combinados con conservantes, antioxidantes y/u otros componentes del aerosol.

Una formulación en aerosol también puede ser una dispersión o suspensión. Una formulación de aerosol en suspensión puede comprender una suspensión de un agente o combinación de agentes de la presente divulgación. Los agentes dispersantes útiles en la divulgación incluyen, p. ej., trioleato de sorbitán, alcohol oleílico, ácido oleico, lecitina y aceite de maíz. Una formulación de aerosol en suspensión también puede incluir lubricantes, conservantes, antioxidantes y/u otros componentes del aerosol.

Una formulación en aerosol puede formularse de forma similar como una emulsión. Una formulación de emulsión en aerosol puede incluir, p. ej., un alcohol como el etanol, un tensioactivo, agua y un propelente, así como un agente o combinación de agentes de la divulgación. El tensioactivo utilizado puede ser no iónico, aniónico o catiónico. Un ejemplo de formulación de aerosol en emulsión comprende, p. ej., etanol, tensioactivo, agua y propelente. Otro ejemplo de formulación de aerosol en emulsión comprende, p. ej., aceite vegetal, monoestearato de glicerilo y propano.

Los compuestos de la divulgación pueden formularse para su administración como supositorios. Primero se funde una cera de bajo punto de fusión, tal como una mezcla de triglicéridos, glicéridos de ácidos grasos, Witepsol S55 (marca comercial de Dynamite Nobel Chemical, Alemania) o manteca de cacao, y el componente activo se dispersa homogéneamente, p. ej., mediante agitación. A continuación, la mezcla homogénea fundida se vierte en moldes de tamaño adecuado, se deja enfriar y se solidifica.

Los compuestos de la divulgación pueden formularse para administración vaginal. Pesarios, tampones, cremas, geles, pastas, espumas o aerosoles que contengan, además del ingrediente activo, vehículos conocidos en la técnica.

Se prevé adicionalmente, que los compuestos de la divulgación se puedan unir de forma liberable a polímeros biocompatibles para su uso en formulaciones de liberación sostenida sobre, en o unidas a insertos para administración tópica, intraocular, periocular o sistémica. La liberación controlada de un polímero biocompatible puede utilizarse también con un polímero soluble en agua para formar una formulación instilable. La liberación controlada a partir de un polímero biocompatible, tal como p. ej., microesferas o nanoesferas de PLGA, puede utilizarse en una formulación adecuada para la implantación o inyección intraocular para la administración de liberación sostenida, así como cualquier polímero biodegradable y biocompatible adecuado.

En un aspecto de la divulgación, el estado de vehículo del sujeto de cualquiera de las variantes genéticas de riesgo de variación descritas en el presente documento, o las variantes genéticas identificadas mediante otros métodos de análisis dentro de los genes o loci reguladores que son identificados por las CNV o SNV descritas en el presente documento, pueden utilizarse para ayudar a determinar si debe administrarse una modalidad de tratamiento particular, tal como cualquiera de las anteriores, o una combinación de las mismas. La administración de una opción de tratamiento, tal como cualquiera de las opciones de tratamiento mencionadas anteriormente, puede determinarse en función de la presencia o ausencia de una variante de riesgo de variación genética particular en el individuo, o mediante la monitorización de la expresión de genes que están asociados con las variantes de la presente divulgación. Los niveles de expresión y/o los niveles de ARNm pueden así determinarse antes y durante el tratamiento para controlar su eficacia. Alternativamente, o de forma concomitante, el estado con respecto a una variación genética, y o el genotipo y/o el estado del haplotipo de al menos una variante de riesgo para la PML presentada en el presente documento puede determinarse antes y durante el tratamiento para monitorizar su eficacia. Los expertos en la materia también pueden apreciar que los niveles de expresión aberrante de un gen afectado por una CNV u otras mutaciones encontradas como consecuencia de la secuenciación dirigida del gen identificado por la CNV pueden ensayarse o probarse diagnósticamente midiendo el nivel de expresión polipeptídica de dicho gen expresado aberrantemente. Los niveles de expresión aberrantes de un gen pueden ser el resultado de una CNV que afecte a una secuencia de ADN (p. ej., un sitio de unión a un factor de transcripción) que regule un gen cuyo nivel de expresión aberrante esté implicado en la PML o la provoque, o de otras mutaciones encontradas como consecuencia de la secuenciación dirigida de la secuencia reguladora del gen identificada por la CNV, que pueden analizarse o someterse a pruebas diagnósticas midiendo el nivel de expresión polipeptídica del gen implicado en la PML o la provoque. Una mutación específica de la CNV dentro de un gen, u otras mutaciones específicas encontradas en la secuenciación dirigida de un gen identificado por las CNV que se ha descubierto que está implicado o es causante de PML, puede causar un cambio estructural aberrante en el polipéptido expresado que resulta de dichas mutaciones genéticas, y la(s) estructura(s) polipeptídica(s) alterada(s) puede(n) analizarse mediante diversos métodos conocidos por los expertos en la materia.

Alternativamente, las redes biológicas o vías metabólicas relacionadas con los genes dentro de, o asociadas con, las variaciones genéticas descritas en el presente documento pueden ser monitorizadas determinando los niveles de ARNm y/o polipéptidos. Esto puede hacerse, p. ej., controlando los niveles de expresión de los polipéptidos de varios genes pertenecientes a la red y/o vía en muestras de ácido nucleico tomadas antes y durante el tratamiento. Alternativamente, los metabolitos pertenecientes a la red biológica o a la vía metabólica pueden determinarse antes y durante el tratamiento. La eficacia del tratamiento se determina comparando los cambios observados en los niveles de expresión/niveles de metabolitos durante el tratamiento con los datos correspondientes de sujetos sanos.

Las variaciones genéticas descritas en el presente documento y/o las encontradas posteriormente (p. ej., mediante otros métodos de análisis genético como la secuenciación) a través del análisis dirigido de aquellos genes identificados inicialmente por las variaciones genéticas descritas en el presente documento, pueden utilizarse para prevenir los efectos adversos asociados a un agente terapéutico, tal como durante los ensayos clínicos. Por ejemplo, los individuos que son portadores de al menos una variación genética de riesgo pueden tener más probabilidades de responder negativamente a un agente terapéutico, tal como un agente

inmunosupresor. Por ejemplo, los portadores de determinadas variantes genéticas pueden ser más propensos a mostrar una respuesta adversa al agente terapéutico. Una o más de las variaciones genéticas empleadas durante los ensayos clínicos para un agente terapéutico dado pueden utilizarse en una prueba diagnóstica complementaria que se administra al paciente antes de la administración del agente terapéutico para determinar si es probable que el paciente tenga una respuesta favorable o adversa al agente terapéutico.

Las variaciones genéticas descritas en el presente documento pueden utilizarse para determinar si a un sujeto se le administra un agente farmacéutico, tal como un fármaco inmunosupresor. Ciertas combinaciones de variantes, incluyendo, las descritas en el presente documento, pero también combinaciones con otras variantes de riesgo de PML, pueden ser adecuadas para una selección de opciones de tratamiento, mientras que otras combinaciones de variantes pueden ser adecuadas para la selección de otras opciones de tratamiento. Dichas combinaciones de variantes pueden incluir una variante, dos variantes, tres variantes o cuatro o más variantes, según sea necesario para determinar con precisión clínicamente fiable la selección del módulo de tratamiento. La información de las pruebas para las variaciones genéticas descritas en el presente documento, u otras variaciones genéticas raras en o cerca de los genes descritos en el presente documento, puede combinarse con la información de otros tipos de pruebas (p. ej., una prueba de anticuerpos contra el JCV, una prueba de CD62L o una prueba de bandas oligoclonales de IgM del CSF) para la selección de opciones de tratamiento.

Kits

La presente invención se refiere a un medicamento inmunosupresor para su uso o a métodos tal como se definen en las reivindicaciones, para los cuales lo que sigue proporciona una divulgación útil relativa a los kits (en el sentido de que, p. ej., el medicamento inmunosupresor para su uso puede estar comprendido en un kit, o los reactivos para llevar a cabo el método reivindicado pueden proporcionarse en un kit), pero tales kits no constituyen el objeto que se reivindica en el presente documento. Los kits útiles en los métodos de la divulgación comprenden componentes útiles en cualquiera de los métodos descritos en el presente documento, incluyendo, p. ej., cebadores para amplificación de ácido nucleico, sondas de hibridación para detectar variación genética, u otra detección de marcadores, enzimas de restricción, sondas de ácido nucleico, opcionalmente marcadas con marcadores adecuados, oligonucleótidos específicos de alelo, anticuerpos que se unen a un polipéptido alterado codificado por un ácido nucleico de la divulgación según se describe en el presente documento o a un polipéptido de tipo silvestre codificado por un ácido nucleico de la divulgación según se describe en el presente documento, medios para la amplificación de variaciones genéticas o fragmentos de las mismas, medios para analizar la secuencia de ácido nucleico de ácidos nucleicos que comprenden variaciones genéticas según se describe en el presente documento, medios para analizar la secuencia de aminoácidos de un polipéptido codificado por una variación genética, o un ácido nucleico asociado con una variación genética, etc. Los kits pueden incluir, p. ej., los tampones necesarios, cebadores de ácidos nucleicos para amplificar ácidos nucleicos, y reactivos para la detección específica de alelo de los fragmentos amplificados utilizando dichos cebadores y las enzimas necesarias (p. ej., ADN polimerasa). Además, los kits pueden proporcionar reactivos para ensayos que se utilizarán en combinación con los métodos de la presente divulgación, p. ej., reactivos para su uso con otros ensayos de detección de PML.

La divulgación puede pertenecer a un kit para ensayar una muestra de ácido nucleico de un sujeto para detectar la presencia de una variación genética, donde el kit comprende reactivos necesarios para detectar selectivamente al menos una variación genética particular en el genoma del individuo. La divulgación puede pertenecer a un kit para ensayar una muestra de ácido nucleico de un sujeto para detectar la presencia de al menos un alelo particular de al menos un polimorfismo asociado con una variación genética en el genoma del sujeto. Los reactivos pueden comprender al menos un oligonucleótido contiguo que se hibrida con un fragmento del genoma del individuo que comprende al menos la variación genética. Los reactivos pueden comprender al menos un par de oligonucleótidos que hibridan con cadenas opuestas de un segmento genómico obtenido de un sujeto, en el que cada par de cebadores oligonucleotídicos está diseñado para amplificar selectivamente un fragmento del genoma del individuo que incluye al menos una variación genética, o un fragmento de una variación genética. Dichos oligonucleótidos o ácidos nucleicos pueden diseñarse utilizando los métodos descritos en el presente documento. El kit puede comprender uno o más ácidos nucleicos marcados capaces de detección específica de alelo de uno o más marcadores polimórficos específicos o haplotipos con una variación genética, y reactivos para la detección del marcado. Un kit para detectar marcadores de SNP pueden comprender una sonda de oligonucleótido de detección, que se hibrida con un segmento de ADN molde que contiene un polimorfismo SNP que se va a detectar, una sonda de oligonucleótido potenciador, una sonda de detección, un cebador y/o una endonucleasa, p. ej., de acuerdo con lo descrito por Kutyavin et al., (Nucleic Acid Res. 34:el28 (2006)). El kit puede contener reactivos para detectar SNV y/o CNV.

El molde de ADN puede ser amplificado por cualquier medio de la presente divulgación, antes de la evaluación de la presencia de variaciones genéticas específicas como se describe en el presente documento. Pueden utilizarse métodos estándar bien conocidos por el experto para llevar a cabo estos métodos, y están dentro del alcance de la divulgación. Los reactivos para llevar a cabo estos métodos pueden incluirse en el kit de reactivos.

En otro aspecto de la presente divulgación, se proporciona un envase (kit) farmacéutico, que comprende un agente terapéutico y un conjunto de instrucciones para la administración del agente terapéutico a seres

humanos sometidos a detección para una o más variantes de la presente divulgación, como se divulga en el presente documento. El agente terapéutico puede ser un fármaco de molécula pequeña, un anticuerpo, un péptido, una molécula antisentido o ARNi, u otras moléculas terapéuticas como las descritas en el presente documento. A un individuo identificado como no portador de al menos una variante de la presente divulgación se le puede indicar que tome el agente terapéutico. A un individuo identificado como no portador de al menos una variante de la presente divulgación se le puede indicar que tome una dosis prescrita del agente terapéutico. A un individuo identificado como portador de al menos una variante de la presente divulgación se le puede indicar que no tome el agente terapéutico. Un individuo identificado como portador de al menos una variante de la presente divulgación puede recibir instrucciones de no tomar una dosis prescrita del agente terapéutico. A un individuo identificado como portador de al menos una variante de la presente divulgación se le puede indicar que tome un agente dirigido contra el virus de JC. Por ejemplo, a un individuo identificado como portador de al menos una variante de la presente divulgación se le puede indicar que tome un agente dirigido contra el virus de JC antes o junto con la toma de un agente inmunosupresor.

También se proporcionan en el presente documento artículos de fabricación, que comprenden una sonda que hibrida con una región del cromosoma humano como se describe en el presente documento y se puede utilizar para detectar un polimorfismo descrito en el presente documento. Por ejemplo, cualquiera de las sondas para detectar polimorfismos o variaciones genéticas descritas en el presente documento puede combinarse con material de envasado para generar artículos de fabricación o kits. El kit puede incluir uno o más elementos que incluyan: instrucciones de uso; y otros reactivos como un marcador o un agente útil para fijar un marcador a la sonda. Las instrucciones de uso pueden incluir instrucciones para aplicaciones de detección de la sonda para realizar un diagnóstico, pronóstico o teranosis a PML en un método descrito en el presente documento. Otras instrucciones pueden incluir instrucciones para adherir un marcador a la sonda, instrucciones para realizar un análisis *in situ* con la sonda, y/o instrucciones para obtener una muestra de ácido nucleico para ser analizada de un sujeto. En algunos casos, el kit puede incluir una sonda marcada que hibrida con una región del cromosoma humano como se describe en el presente documento.

El kit también puede incluir una o más sondas de referencia o control adicionales que hibridan con el mismo cromosoma u otro cromosoma o porción del mismo que puede tener una anomalía asociada con un endofenotipo particular. Un kit que incluye sondas adicionales puede incluir además marcadores, p. ej., una o más marcadores iguales o diferentes para las sondas. La sonda o sondas adicionales suministradas con el kit pueden ser una sonda o sondas marcadas. Cuando el kit incluye además una o más sondas adicionales, el kit puede proporcionar además instrucciones para el uso de la sonda o sondas adicionales. También pueden proporcionarse kits para autodiagnóstico. Dichos kits de prueba pueden incluir dispositivos e instrucciones que un sujeto puede utilizar para obtener una muestra de ácido nucleico (p. ej., células bucales, sangre) sin la ayuda de un profesional sanitario. Por ejemplo, las células bucales pueden obtenerse utilizando un hisopo o cepillo bucal, o utilizando enjuague bucal.

Los kits tal como se proporcionan en el presente documento también pueden incluir un remitente (p. ej., un sobre con franqueo pagado o un paquete postal) que se puede utilizar para devolver la muestra de ácido nucleico para su análisis, p. ej., a un laboratorio. El kit puede incluir uno o más contenedores para la muestra de ácido nucleico, o la muestra de ácido nucleico puede estar en un vial estándar de extracción de sangre. El kit también puede incluir uno o más de un formulario de consentimiento informado, un formulario de solicitud de prueba e instrucciones sobre cómo utilizar el kit en un método descrito en el presente documento. También se incluyen métodos para utilizar dichos kits. Uno o más de los formularios (p. ej., el formulario de solicitud de la prueba) y el recipiente que contiene la muestra de ácido nucleico pueden codificarse, p. ej., con un código de barras para identificar al sujeto que proporcionó la muestra de ácido nucleico.

Una prueba de detección *in vitro* puede comprender uno o más dispositivos, herramientas y equipos configurados para recoger una muestra de ácido nucleico de un individuo. Con una prueba de detección *in vitro*, las herramientas para recoger una muestra de ácido nucleico pueden incluir uno o más de un hisopo, un bisturí, una jeringa, un raspador, un recipiente y otros dispositivos y reactivos diseñados para facilitar la recogida, el almacenamiento y el transporte de una muestra de ácido nucleico. Una prueba de detección *in vitro* puede incluir reactivos o soluciones para recoger, estabilizar, almacenar y procesar una muestra de ácido nucleico.

Tales reactivos y soluciones para la recolección, estabilización, almacenamiento y procesamiento de nucleótidos son bien conocidos por los expertos en la materia y pueden ser indicados por métodos específicos utilizados por una prueba de detección *in vitro* como se describe en el presente documento. Un ensayo de detección *in vitro*, tal como se describe en el presente documento, puede comprender un aparato de microarreglos y reactivos, un aparato de celdas de flujo y reactivos, un secuenciador multiplex de nucleótidos y reactivos, y el hardware y software adicionales necesarios para analizar una muestra de ácido nucleico en busca de determinados marcadores genéticos y para detectar y visualizar determinados marcadores genéticos.

La presente divulgación se refiere además a kits para utilizar anticuerpos en los métodos descritos en el presente documento. Esto incluye, entre otros, kits para detectar la presencia de un polipéptido variante en una muestra de ácido nucleico de prueba. Preferentemente se incluyen anticuerpos tales como un anticuerpo

marcado o marcable y un compuesto o agente para detectar polipéptidos variantes en una muestra de ácido nucleico, medios para determinar la cantidad o la presencia y/o ausencia de polipéptido variante en la muestra de ácido nucleico, y medios para comparar la cantidad de polipéptido variante en la muestra de ácido nucleico con un estándar, así como instrucciones de uso del kit. El kit puede comprender además un conjunto de instrucciones para utilizar los reactivos que lo componen.

Aspectos informáticos

La presente invención se refiere a un medicamento inmunosupresor para su uso o a métodos como los definidos en las reivindicaciones, para los cuales lo que sigue proporciona una divulgación útil relativa a aspectos informáticos (que pueden, p. ej., utilizarse para determinar el riesgo de PML), pero tales aspectos informáticos no constituyen el objeto que se reivindica en el presente documento. Como comprenderán los expertos en la materia, los métodos y la información descritos en el presente documento (asociación de variación genética con PML) pueden implementarse, en todo o en parte, tal como instrucciones ejecutables por ordenador en medios legibles por ordenador conocidos. Por ejemplo, los métodos descritos en el presente documento pueden implementarse en hardware. Alternativamente, el método puede implementarse en software almacenado en, p. ej., una o más memorias u otro medio legible por ordenador e implementado en uno o más procesadores. Como es sabido, los procesadores pueden estar asociados a uno o más controladores, unidades de cálculo y/u otras unidades de un sistema informático, o implantados en firmware según se desee. Si se implementan en software, las rutinas pueden almacenarse en cualquier memoria legible por ordenador, tal como en RAM, ROM, memoria flash, un disco magnético, un disco láser u otro medio de almacenamiento, como también es conocido. Del mismo modo, este software puede suministrarse a un dispositivo informático a través de cualquier método de suministro conocido, incluyendo, p. ej., a través de un canal de comunicación como una línea telefónica, Internet, una conexión inalámbrica, etc., o a través de un medio transportable, tal como un disco legible por ordenador, una unidad flash, etc.

De manera más general, y tal como lo entienden aquellos con conocimientos ordinarios en la materia, los diversas etapas descritas anteriormente pueden implementarse como diversos bloques, operaciones, herramientas, módulos y técnicas que, a su vez, pueden implementarse en hardware, firmware, software o cualquier combinación de hardware, firmware y/o software. Cuando se implementan en hardware, algunos o todos los bloques, operaciones, técnicas, etc. pueden implementarse, p. ej., en un circuito integrado (IC) personalizado, un circuito integrado de aplicación específica (ASIC), una matriz lógica programable de campo (FPGA), una matriz lógica programable (PLA), etc.

Los resultados de dicho genotipado pueden almacenarse en una unidad de almacenamiento de datos, tal como un portador de datos, incluyendo, bases de datos informáticas, discos de almacenamiento de datos, o por otros medios convenientes de almacenamiento de datos. La base de datos informática puede ser una base de datos de objetos, una base de datos relacional o una base de datos posrelacional. Los datos pueden recuperarse de la unidad de almacenamiento de datos utilizando cualquier método de consulta de datos conveniente.

Cuando se implementa en software, el software puede almacenarse en cualquier medio legible por ordenador conocido, tal como en un disco magnético, un disco óptico u otro medio de almacenamiento, en una memoria RAM o ROM o flash de un ordenador, procesador, unidad de disco duro, unidad de disco óptico, unidad de cinta, etc. Del mismo modo, el software puede suministrarse a un usuario o a un sistema informático a través de cualquier método de suministro conocido, incluyendo, p. ej., en un disco legible por ordenador u otro mecanismo de almacenamiento informático transportable.

Las etapas de los métodos reivindicados pueden ser operativas con otros numerosos entornos o configuraciones de sistemas informáticos de propósito general o especial. Ejemplos de sistemas informáticos conocidos, entornos y/o configuraciones que pueden ser adecuados para su uso con los métodos o sistemas de las reivindicaciones incluyen, pero no se limitan a, ordenadores personales, ordenadores de servidor, dispositivos de mano o portátiles, sistemas multiprocesador, sistemas basados en microprocesador, decodificadores, electrónica de consumo programable, ordenadores personales en red, miniordenadores, ordenadores centrales, entornos informáticos distribuidos que incluyen cualquiera de los sistemas o dispositivos anteriores, y similares.

Las etapas del método y sistema reivindicados pueden describirse en el contexto general de instrucciones ejecutables por ordenador, tales como módulos de programa, que son ejecutados por un ordenador. Generalmente, los módulos de programa incluyen rutinas, programas, objetos, componentes y/o estructuras de datos que realizan tareas particulares o implementan tipos de datos abstractos particulares. Los métodos y aparatos también pueden practicarse en entornos informáticos distribuidos en los que las tareas son realizadas por dispositivos de procesamiento remotos que están conectados a través de una red de comunicaciones. Tanto en entornos informáticos integrados como distribuidos, los módulos de programa pueden ubicarse en medios de almacenamiento informático locales y remotos, incluidos dispositivos de almacenamiento de memoria. Se contemplan numerosas implementaciones alternativas, utilizando tecnología actual o tecnología desarrollada después de la fecha de presentación de esta solicitud.

Aunque el sistema y el método de evaluación de riesgos, y otros elementos, se han descrito como preferiblemente implementados en software, pueden implementarse en hardware, firmware, etc., y pueden ser implementados por cualquier otro procesador. Así pues, los elementos descritos en el presente documento pueden implementarse en una CPU multipropósito estándar o en hardware o firmware diseñado específicamente, tal como un circuito integrado de aplicación específica (ASIC) u otro dispositivo cableado, según se desee. Cuando se implementa en software, la rutina de software puede almacenarse en cualquier memoria legible por ordenador, tal como un disco magnético, un disco láser u otro medio de almacenamiento, en una RAM o ROM de un ordenador o procesador, en cualquier base de datos, etc. Asimismo, este software puede suministrarse a un usuario o a un sistema de detección a través de cualquier método de suministro conocido o deseado, incluyendo, p. ej., en un disco legible por ordenador u otro mecanismo de almacenamiento informático transportable o a través de un canal de comunicación, p. ej., una línea telefónica, Internet o una comunicación inalámbrica.

A menos que se explique lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que comúnmente entiende una persona con conocimientos ordinarios en la técnica a la que pertenece esta divulgación. Las siguientes referencias contienen ejemplos de los métodos y composiciones que pueden utilizarse en el presente documento: The Merck Manual of Diagnosis and Therapy, 18th edition, publicado por Merck Research Laboratories, 2006 (ISBN 0-911910-18-2); Benjamin Lewin, Genes IX, publicado por Jones & Bartlett Publishing, 2007 (ISBN-13: 9780763740634); Kendrew et al., (eds.), The Encyclopedia of Molecular Biology, publicado por Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); y Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, publicado por VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

Los procedimientos estándar de la presente divulgación se describen, p. ej., en Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA (1982); Sambrook et al., Molecular Cloning: A Laboratory Manual (2 ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., USA (1989); Davis et al., Basic Methods in Molecular Biology, Elsevier Science Publishing, Inc., Nueva York, USA (1986); o Methods in Enzymology: Guide to Molecular Cloning Techniques Vol. 152, S. L. Berger y A. R. Kimmerl (eds.), Academic Press Inc., San Diego, USA (1987)). Current Protocols in Molecular Biology (CPMB) (Fred M. Ausubel, et al., ed., John Wiley and Sons, Inc.), Current Protocols in Protein Science (CPPS) (John E. Coligan, et al., ed., John Wiley and Sons, Inc.), Current Protocols in Immunology (CPI) (John E. Coligan, et al., ed., John Wiley and Sons, Inc. John Wiley and Sons, Inc.), Current Protocols in Cell Biology (CPCB) (Juan S. Bonifacino et al., ed., John Wiley and Sons, Inc.), Culture of Animal Cells: A Manual of Basic Technique por R. Ian Freshney, Editorial: Wiley-Liss; 5th edition (2005), y Animal Cell Culture Methods (Methods in Cell Biology, Vol. 57, Jennie P. Mather y David Barnes editores, Academic Press, 1st edition, 1998).

Debe entenderse que los siguientes ejemplos no deben interpretarse como limitativos de la metodología, protocolos y composiciones particulares, etc., descritos en el presente documento y, como tales, pueden variar. Los siguientes términos utilizados en el presente documento se utilizan únicamente para describir realizaciones concretas y no pretenden limitar el alcance de las realizaciones descritas en el presente documento.

Se divulgan en el presente documento moléculas, materiales, composiciones y componentes que pueden usarse para, pueden usarse en conjunción con, pueden usarse en preparación para, o son productos de métodos y composiciones divulgados en el presente documento.

Se entiende que los métodos y composiciones divulgados no se limitan a la metodología, protocolos y reactivos particulares descritos, ya que éstos pueden variar. También debe entenderse que la terminología empleada en el presente documento tiene por objeto describir únicamente realizaciones particulares, y no pretende limitar el alcance de la presente divulgación, que sólo puede limitarse mediante las reivindicaciones adjuntas.

A menos que se definan de otro modo, todos los términos técnicos y científicos utilizados en el presente documento tienen el significado que comúnmente entendería un experto en la materia en el contexto de la presente especificación.

Cabe señalar que, tal como se utilizan en el presente documento y en las reivindicaciones adjuntas, las formas singulares "un", "uno, una" y "el, la" incluyen una referencia plural a menos que el contexto indique claramente lo contrario. De este modo, p. ej., la referencia a "un nucleótido" incluye una pluralidad de tales nucleótidos; la referencia a "el nucleótido" es una referencia a uno o más nucleótidos, y así sucesivamente.

En el presente contexto, se entenderá que el término "y/o" indica que se trata de uno o de ambos elementos conectados por él. Aunque se han mostrado y descrito en el presente documento realizaciones preferidas de la presente divulgación, puede resultar obvio para los expertos en la materia que dichas realizaciones se proporcionan únicamente a modo de ejemplo. La invención puede modificarse dentro del ámbito de las reivindicaciones adjuntas.

Ejemplos

La presente invención implica una variación genómica seleccionada de chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A. Los siguientes Ejemplos, en la medida en que no están relacionados con las variaciones genómicas chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A, se incluyen como referencia.

Ejemplo 1 - Enfoque experimental

En el presente estudio, un conjunto de genes fue identificado, variantes deletéreas dentro de los cuales aumentan la susceptibilidad a PML. Los genes relevantes se descubrieron sobre la base de un enfoque combinado de CNV más análisis de secuencias. Se recopilaron dos conjuntos de genes (véase la Tabla 6 y la descripción correspondiente):

A. Un conjunto basado en una revisión bibliográfica detallada de genes implicados en el sistema inmunitario y la biología del virus de JC, junto con genes descritos en el contexto de la PML a través de informes de casos.
B. Un conjunto basado en la observación de CNV raras dentro de la cohorte de PML.

Se generó una lista no redundante de 419 genes ((véase la Tabla 6), que contiene 245 curados de revisiones de deficiencia inmunitaria (inmunodeficiencia) (Tabla 6, 'Base de datos pública'), 169 identificados a través de CNV raras usando los métodos descritos en el presente documento (Tabla 6, 'PBio'), y 6 genes que fueron encontrados usando ambos métodos (Tabla 6, 'Ambos'). Para más información, véase la Tabla 6 y la descripción que figura a continuación).

Utilizando este conjunto de 419 genes, se determinó si:

- Se observaron CNV raras que podrían explicar la susceptibilidad a la PML;
- Se observaron variantes raras de la secuencia (determinadas mediante el análisis de secuenciación del exoma completo - WES) que podrían explicar la susceptibilidad a la PML;
- Las combinaciones de CNV, SNV y/o CNV y SNV podrían explicar la susceptibilidad;
- Las variantes individuales podrían estar presentes con mayor frecuencia en la cohorte de PML (análisis de la carga de variantes - Tablas 14, 15);
- El número total de variantes heterocigotas dañinas fue elevado para algunos genes específicos (análisis de la carga génica - Tabla 13).

En todos los casos, se tuvo debidamente en cuenta:

- Naturaleza patógena/deletérea de las variantes observadas (p. ej., si era muy probable que la función del gen se viera afectada);
- Rareza de las variantes o combinaciones de variantes (p. ej., se consideraron aquellas que se esperaría que estuvieran presentes en el 1 % o menos de la población normal);
- Origen étnico de los casos de PML para tener en cuenta las posibles diferencias de frecuencia en un subgrupo de población frente a otro. Las etnias (p. ej., ascendencia) de los pacientes con PML se indican en la Tabla 7. En el caso de los identificadores de las muestras que empezaban por `MVGS', no se comunicaron las etnias, pero todos los pacientes procedían de EE.UU. y se asumió que sus etnias eran de ascendencia europea (EUR). Sin embargo, el caso de PML MVGS9 11-13a es potencialmente de ascendencia africana (AFR) sobre la base de SNV comunes que también se encuentran en casos de PML conocidos por ser de ascendencia AFR. En un caso, se utilizaron los datos de frecuencia específicos de la etnia de la base de datos del ExAC para evaluar las frecuencias relativas de las variantes encontradas en pacientes con PML frente a una población no seleccionada (sujetos del ExAC). Las etnias del ExAC se designaron de la siguiente manera: africanos/afroamericanos (AFR), latinos (LAT, también conocidos como AMR), asiáticos orientales (EAS), finlandeses (FIN), europeos no finlandeses (EUR, también conocidos como NFE), sudasiáticos (SAS) y otros (OTH). En algunos de los casos de PML recogidos en la Tabla 7, las etnias eran alternativamente subsahariana, norteafricana (MGB), caribeña (CAR) o hispana (HISP). Para la interpretación de las variantes encontradas en estos pacientes, las asignaciones de ascendencia utilizando las designaciones de las bases de datos del ExAC fueron las siguientes: AFR = MGB o Subsahariana; LAT = CAR o HISP. Se desconocía la ascendencia de dos casos de PML (PML02 y PML28) y, a efectos de interpretación de frecuencias (utilizando las bases de datos del ExAC), se asumió que eran de ascendencia europea (EUR).

Aunque el principal mecanismo genético que se consideró fue la herencia autosómica recesiva (AR), varias soluciones se basaron en la herencia autosómica dominante (AD), pero sólo en casos para los que se encontró evidencia previa de que las variantes heterocigotas en el gen relevante se habían asociado previamente con un síndrome de inmunodeficiencia. Los expertos en la materia pueden apreciar que algunos genes pueden

contener variantes patogénicas tanto del modelo AR como del modelo AD (p. ej., véanse las entradas de la Tabla 6 marcadas como "AD_AR" en la columna "Modelo de Enfermedad").

Para la herencia AR (~40 % de los genes en la Tabla 6 caen en esta categoría, AR o AD_AR), se consideró lo siguiente:

- CNV disruptivas homocigotas o heterocigotas compuestas;
- Variantes de secuencia homocigotas o heterocigotas compuestas; es decir, variantes de un solo nucleótido (SNV). La heterocigosidad compuesta sólo se dedujo cuando se disponía de desfase o uno de los pares de SNV era homocigótico;
- Heterocigosidad compuesta para una CNV y una SNV. Tales llamadas sólo fueron posibles en los casos en los que el SNV estaba en trans para una eliminación (p. ej., el SNV DUSP16 en la Tabla 10 y el CNV en la Tabla 1).

Ejemplo 2 - Análisis de variantes del número de copias (CNV)

Los datos presentados en el presente documento se generaron a partir de una comparación de variantes en el número de copias (CNV) identificadas en 2 cohortes:

1) 1,005 individuos normales (motor de variación normal - NVE);
2) 71 casos de leucoencefalopatía multifocal progresiva (PML) junto con 6 casos de virus de inmunodeficiencia humana (VIH) sin diagnóstico de PML (para ayudar a distinguir las variantes de línea germinal frente a las variantes adquiridas que resultan de la infección por VIH). Tamaño total de la cohorte = 77.

Hibridación de muestras de ADN genómico - NVE y PML, cohortes de VIH

Muestras de ADN genómico de individuos dentro de la cohorte normal (sujetos de "prueba" NVE, también denominados como "casos NVE" en algunas tablas en el presente documento) y de la cohorte de PML, VIH (sujetos de "prueba" de PML, VIH) fueron hibridadas contra un solo individuo de referencia emparejado por sexo. Las muestras de ADN de referencia se marcaron con Cy5 y las muestras de ADN de los sujetos de prueba se marcaron con Cy3. Tras el marcaje, las muestras se combinaron e hibridaron juntamente con microarreglos de oligonucleótidos Agilent Feature 1M, ID de diseño 021529 (número de producto Agilent G4447A) utilizando condiciones estándar (hibridación genómica comparativa de arreglos - aCGH). Tras la hibridación, las matrices se escanearon con una resolución de 2 µm, utilizando el escáner de microarreglos de ADN de Agilent, generando imágenes tiff para su posterior análisis.

Todas las imágenes tiff fueron analizadas utilizando el software Feature Extraction (FE) de Agilent, con la siguiente configuración:

- Congelación del genoma humano:	hg18:NCBI36:Mar2006
- Versión FE:	10.7.3.1
- Rejilla/archivo de diseño:	021529_D_F_20091001
- Protocolo:	CGH_107_Sep09

Este procedimiento genera varios archivos de salida, uno de los cuales es un archivo de texto delimitado por pestañas, que contiene ~1,000,000 de filas de datos, cada una de las cuales corresponde a una característica específica de la matriz. Este archivo *.txt se utilizó para realizar llamadas de CNV utilizando DNAcopy, un paquete de software de código abierto implementado en R a través de BioConductor (http://www.bioconductor.org/packages/release/bioc/html/DNAcopy.html). Las pérdidas heterocigotas (het_loss), las pérdidas homocigotas (hom_loss) o las ganancias se determinaron en función de una relación log2 de umbral, que se fijó en:

- hom_loss mínima = -1000;
- hom_loss máxima = -2;
- het_loss mínima = -2;
- het_loss máxima = -0.5;
- ganancia mínima = 0.5;
- ganancia máxima = 1000;

Con muy pocas excepciones, todas las CNV con un valor de relación log2 entre -0.5 y +0.5 no fueron consideradas. Todos los valores de relación log2 se determinaron de acuerdo con Cy3/Cy5 (Prueba/Referencia). Se fijó en 2 el umbral mínimo de sondas para la llamada de CNV (2 sondas consecutivas eran suficientes para llamar una CNV). Se generó una lista de CNV para cada individuo de las 3 cohortes (NVE, PML y VIH).

Usando guiones personalizados, las CNV identificadas en las cohortes NVE y PML (muchas de las cuales aparecieron en múltiples individuos) fueron (separadamente) 'fusionadas' en listas maestras de subregiones de CNV no redundantes, de acuerdo con la presencia o ausencia de la subregión de CNV en individuos dentro de la cohorte. Usando este enfoque, las listas maestras de NVE tienen:

- 7778 het_loss
- 653 hom_loss
- 4862 ganancia

distintas subregiones de CNV, respectivamente. La cohorte de PML + VIH de 77 listas maestras de individuos contenía:

- 2523 het_loss
- 314 hom_loss
- 1639 ganancia

distintas subregiones de CNV, respectivamente.

Los expertos en la materia pueden apreciar que se pueden adquirir CNV en el genoma de un individuo que no se heredan. Estas "CNV adquiridas" se producen a menudo en tejidos específicos, tal como los tumores sólidos, en comparación con el tejido normal del paciente. En las muestras de ADN genómico derivado de la sangre, que son las que se utilizaron para los sujetos con NVE y PML en los estudios descritos en el presente documento, las CNV adquiridas pueden ser el resultado de cánceres sanguíneos como la leucemia y el linfoma, pero también deberse a la infección por VIH. Muchos de los casos de PML de este estudio tenían el VIH como enfermedad primaria (véase la Tabla 7). Para ayudar en la interpretación de las CNV adquiridas frente a las de línea germinal, se incluyó una subcohorte de 6 casos de VIH en la comparación primaria de CNV de todo el genoma, pero las CNV raras en los 6 casos de VIH (no PML) no se consideraron relevantes para la susceptibilidad a la PML. El propósito de generar datos sobre los 6 casos de VIH era determinar si algunos cambios observados en pacientes con PML que desarrollaron el trastorno sobre un fondo de VIH (PML/VIH) estaban probablemente relacionados con el VIH subyacente y no con la propia susceptibilidad a la PML. En otras palabras, los casos de VIH sirvieron de control general para el gran número de casos de PML/VIH.

Por ejemplo, considere 3 individuos dentro de la cohorte de NVE con las siguientes CNV hipotéticas:

Chrl: 1-100,000; Chrl: 10,001-100,000; y Chrl: 1-89,999. En la lista maestra, éstas se fusionarían en 3 subregiones de CNV distintas, tal como sigue:

- subregión 1 de CNV	Cr1:1-10,000	Sujetos A, C
- subregión 2 de CNV	Cr1:10,001-89,999	Sujetos A, B, C
- subregión 3 de CNV	Chr90,000:1-100,000	Sujetos A, B

Se realizó la comparación de las correspondientes listas maestras de NVE y PML de las subregiones de CNV (het_loss frente a het_loss, hom_loss frente a hom_loss y ganancia frente a ganancia), lo que dio como resultado un archivo combinado con los totales de NVE y PML para cada subregión distinta de CNV del estudio.

Los datos son subsecuentemente curados como sigue (el ejemplo de cálculo abajo fue basado en una cohorte original de PML de 80 casos, de los cuales 6 son controles de VIH no-PML y 3 casos de PML que eran muestras duplicadas. En algunos casos, los valores OR y FET indicados en la Tabla 2 se utilizaron como directrices "relativas" a la hora de considerar la relevancia de una CNV. En casi todos los casos, una CNV se consideró como causa potencial o factor contribuyente a la PML si estaba ausente de la base de datos de NVE de CNV).

- Anotación mediante secuencias de comandos diseñadas a medida con el fin de adjuntar información relevante a cada región de CNV en relación con el solapamiento con genes y exones conocidos, el solapamiento con genes implicados en el sistema inmunitario y el solapamiento con regiones reguladoras, incluidos los sitios de unión de factores de transcripción.
- Un cálculo de la razón de probabilidades (OR) y de la prueba exacta de Fisher (FET) para cada subregión de CNV, de acuerdo con la siguiente fórmula:
  - $$OR = (PML/(77\text{-}PML))/(NVE/(1005\text{-}NVE))$$
  - donde:

◦ PML = número de individuos con PML con subregión de CNV de interés
◦ NVE = número de individuos con NVE con subregión de CNV de interés

Como un ejemplo ilustrativo, considere la ganancia de subregión de CNV que involucra chr2:55764753-55771586, que se encuentra en 3 individuos en la cohorte de PML y 1 individuo en la cohorte de NVE ((véase la Tabla 2). La OR es: (3/74)/(1/1004) = 40.7. Téngase en cuenta que, de acuerdo con una convención, si cualquiera de NVE o PML = 0, se añade un valor de 0.5 a las 4 entradas de la fórmula principal anterior, para evitar tratar con infinitos (véase Deeks y Higgins, Statistical algorithms in Review Manager 5, Statistical Methods Group of The Cochrane Collaboration, (2010)). Esto tiene el efecto de reducir artificialmente los valores de OR en los casos en los que ningún individuo dentro del NVE tiene la CNV. Este método es aplicable a todos los cálculos de la Tabla 2. Este método también se utiliza al calcular la prueba exacta de Fisher de 2 colas (FET) en caso de que alguna de las variables sea cero. Por conveniencia en el análisis, en la Tabla 2 se retuvo la subcohorte de 6 casos de VIH (no PML). Por lo tanto, los valores de OR indicados en la Tabla 2 son ligeramente diferentes de los cálculos de OR para el número real de casos de PML (n = 71). Utilizando el ejemplo anterior para una ganancia de la subregión de CNV que implica chr2:55764753-55771586, la OR real utilizando 71 casos de PML frente a 1005 sujetos de NVE fue: (3/68)/(1/(1004) = 44.29. En algunos casos, se encuentra que un control de VIH no-PML (véase la Tabla 11, identificado como 3280, 3281, 3283, 3284, 3285 y 3286) tiene una CNV de relevancia potencial en sujetos con PML. Esto también puede influir en el cálculo de OR. Por ejemplo, para la pérdida de la subregión de CNV chr19:55247874-55250186 la OR en la Tabla 2 aparece como 17.38 pero un caso es un control de VIH no-PML (Tabla 11, control de PML70 = 3280). Para este ejemplo, la OR real utilizando 71 casos de PML frente a 1005 sujetos de NVE, y excluyendo el caso de VIH sin PML, fue: (4/67)/(4/(1001) = 14.94.

Las subregiones de CNV/genes que se enumeran en el presente documento (p. ej., en una o más de las Tablas 1-4), cumplen uno de los siguientes criterios:

- Biología sólida que vincule el gen al que afecta o cerca del cual se encuentra la subregión de la CNV, con vías/mecanismos conocidos de inmunodeficiencia o biología en la PML (p. ej., biología relacionada con el virus de JC). Es decir, en algunos casos, faltan pruebas estadísticas, pero no excluyen la subregión de la CNV como candidata;
- Análisis estadístico combinado con biología media a fuerte (p. ej., vínculos en la literatura revisada por pares con PML, virus de JC, defensa del huésped, inmunodeficiencia o neuropatología) sin conexión biológica obvia (el mejor FET en esta categoría fue de 3.25E-10);

Puede ser apreciado por los expertos en la materia que el número de subregiones de CNV candidatas a PML, independientemente de la categoría, puede aumentar o disminuir a medida que se analizan cohortes de PML adicionales.

Ejemplo 3 - Secuenciación del exoma completo (WES) y análisis a nivel de caso.

Los datos de WES fueron obtenidos en un total de 70 casos de PML (los casos de VIH no-PML no fueron secuenciados - fueron usados simplemente para ayudar en la interpretación de CNV complejos observados en pacientes con PML que también tenían VIH).

Los informes de anotación de variantes se contrastaron además con el conjunto completo de genes detallados anteriormente. No se tuvieron en cuenta las variantes sinónimas ni las que se preveía que fueran modificadoras (fuera de las regiones codificantes). Para todas las demás variantes, se realizó un filtrado adicional, de modo que sólo se consideraron patogénicas para su posterior evaluación aquellas que al menos un algoritmo de predicción *in silico* (p. ej., Polyphen2, SIFT, MutationTaster) predijo que lo eran. Por último, sólo se evaluaron para el análisis a nivel de caso las variantes o combinaciones de variantes que se esperaría que estuvieran presentes en el 1 % o menos de la población normal (Tablas 7-10). Se utilizaron datos del Exome Aggregation Consortium (ExAC) para obtener datos de frecuencia específicos de cada etnia para las variantes consideradas (véase, Lek et al., Nature, 17; 536(7616): 285-91) (2016)).

Ejemplo 4 - Descripción de los datos de secuencia

El archivo de secuencia 33655-710.601_ST25.txt contiene información de secuencia genómica para (en el siguiente orden):

A. Todas las CNV distintas enumeradas en la Tabla 1;
B. La extensión genómica completa de los transcripciones enumerados en la Tabla 4;
C. Variantes de secuencia detalladas en la Tabla 5.
D. La extensión genómica completa de los transcripciones enumerados en la Tabla 12

Téngase en cuenta que:

1. SEQ_ID 1-172 son las secuencias de CNV de la Tabla 1;
2. SEQ_ID 173-455 son las secuencias de transcripciones de la Tabla 4;

3. SEQ_ID 1000-1329 son las variantes de secuencia de la Tabla 5;
4. SEQ_ID 1500-2177 son las secuencias de transcripciones de la Tabla 12.

Ejemplos de secuencias presentadas:

Se inicia la entrada de secuencia:

Tabla 1: SEQ_ID 1 = CNV 49,653 pb (het loss) en chr1:1086119-1135772 involucrando los genes MIR200A,MIR200B,MIR429,TNFRSF18,TTLL10:

<210> 1

<211> 49654

<212> ADN

<213> Homo sapiens

<400> 1

cttctggggt ctaaggccag aagtgacctt tctctcacg gaggcacccc cacatcacag 60
gccccaagct cccaccagga gtcccaggc agcaggtttt ccaccacagc cgggaagagc 120
cccgccttca ccacccacca ccagccaatc ccgagaccac cgaagccccc agaccgggcc 180
.....(sequence truncated for brevity)
gattcccgca cggccgggga cggccccagg gccttgggag cgtctgtgga cacctgtggt 49560
gtgggccgag gagctgggag ctcatctgaa cacgccagca ctcgcgcatc cacgctgctg 49620
gcggatgcct gggtttctcc actgtggggc cacg 49654

Finaliza la entrada de secuencia.

Se inicia la entrada en secuencia:

Tabla 4: SEQ_ID 173 = MIR200B, transcripción NR_029639, que tiene 95 pb de longitud:

<210> 173

<211> 95

<212> ADN

<213> Homo sapiens

<400> 173

ccagctcggg cagccgtggc catcttactg ggcagcattg gatggagtca ggtctctaat 60
actgcctggt aatgatgacg gcggagccct gcacg 95

Finaliza la entrada de secuencia.

Se inicia la entrada en secuencia:

Tabla 5: SEQ_ID 1148 = alelo de referencia chr9:304628 = G; alelo alternativo = A

<210> 1148

<211> 40

<212> ADN

<213> Homo sapiens

<220>

<221> variante

<222> (20)..(20)

<223> G->A

<400> 1148

tttaaaaaga ctggatctcg aaaagatttt cacaagacgc 40

Finaliza la entrada de secuencia.

Se inicia la entrada en secuencia:

Tabla 12: SEQ_ID 1500 = ACADM, transcripción NM _000016, que tiene 39,313 pb de longitud:

<210> 1500

<211> 39313

<212> ADN

<213> Homo sapiens

<400> 1500

cgcaagtccc cccaccgttc agcgccaaccg ggccctccca gccccgccgc cgtccccctc 60
ccccgccctg gctctcttlc cgcgctgcgg tcagcctcgg cgtcccacag agagggccag 120
.....(sequence truncated for brevity)
gtaatagtgt atatttctlg tatttactat gatgaaaaaa ggtcgtttta attttgaatt 39240
gaataaagtt accigttcat ttttattag atattttaaa gacttcagaa aatataaata 39300
tgaaataatt taa 39313

Finaliza la entrada de secuencia.

Ejemplo 5

Los expertos en la materia pueden apreciar que los genes pueden verse afectados por variantes genéticas adquiridas o de línea germinal (p. ej., CNV), en las que cada gen tiene el potencial de contener variantes genéticas adquiridas (p. ej., a través de un proceso de enfermedad como la infección por VIH, o cánceres como la leucemia y el linfoma) o presentes en la línea germinal (p. ej., heredadas de un progenitor o de novo, es decir, no heredadas de un progenitor). En la Figura 1, el gen PRKCB estaba afectado por variantes de línea germinal en 2 casos de PML y por variantes adquiridas en 6 casos de PML. La invención descrita en el presente documento se centra en la detección de variantes de la línea germinal que están presentes en los genomas de pacientes con PML. Por lo tanto, ninguna solución/explicación para la PML de un paciente dado se basó en una CNV adquirida, aunque otro paciente con PML podría potencialmente ser "resuelto" por una o dos variantes raras de la línea germinal que afecten al gen.

Para este ejemplo PRKCB, no se encontraron soluciones basadas en CNV (se asumió un modelo AR), pero 1 solución de SNV se reporta en la Tabla 8 (SNV het, se asumió un modelo AD para este caso de PML). Se evaluaron más pruebas de apoyo para el gen PRKCB realizando un análisis String (alta confianza = 0.7, 1ª capa = hasta 10 interactuadores; string-db.org; véase Szklarczyk et al., (2015), y sus referencias). El análisis String mostró que PRKCB interactúa con los genes de PML-419 CARD11, IKBKB, y RBCK1 (véase la Tabla 6).

En la Figura 2, tanto TNFRSF13C como CENPM están interrumpidos y/o ganados por un conjunto de ganancias de CNV adquiridas. Las CNV adquiridas pueden ser muy complejas, tal como las elevadas ganancias en el número de copias que suelen identificarse en las muestras de ADN derivadas de tumores (en comparación con el genoma normal del paciente). En el descubrimiento del gen de PML descrito en el presente documento, el ADN genómico derivado de la sangre obtenido de varios pacientes con VIH diagnosticados con PML, o casos de PML con una enfermedad primaria de leucemia y linfoma (reportados como 'Otros' en la Tabla 7), mostraron cambios genómicos complejos (p. ej., ganancias que exhibían un patrón dup-trip-dup). En algunos casos de PML, las ganancias adquiridas superaron el punto de corte de la relación log2 (>0.5) que se seleccionó para este estudio, pero en otros casos de PML las relaciones log2 para las ganancias fueron <0.5 y estos datos se filtraron de los análisis principales que se realizaron para determinar las CNV de línea germinal rara.

En un caso, se utilizó un conjunto de 6 casos de VIH sin PML (3 de ascendencia africana, 3 de ascendencia europea) para ayudar en la interpretación de si una CNV era un evento adquirido o de línea germinal. Los "casos de PML" sin PML están marcados con "_control" en la Tabla 11 y corresponden a los números "ID del caso de PML" 3280, 3281, 3283, 3284, 3285 y 3286. Mientras que algunas CNV se reportan en las Tablas 1 y 2 para este conjunto de sujetos con VIH control sin PML, ninguno de estos hallazgos genéticos se utilizó para nominar un gen descubierto sobre la base de CNV raras (en comparación con la base de datos de NVE) como un gen potencial de PML (genes PBio reportados en la Tabla 6). En otras palabras, estas CNV raras sólo se utilizaron para ayudar a determinar si una región genómica concreta que contenía múltiples CNV superpuestas se debía potencialmente a un evento genético adquirido. Los expertos en la materia pueden apreciar que el conjunto de experimentos descritos en el presente documento no necesariamente descarta por completo que una región genómica dada contenga sólo CNV adquiridas frente a sólo CNV de línea germinal (es decir, es posible que la misma región pueda contener una CNV adquirida en un individuo y una CNV de línea germinal en otro).

Para los datos de CNV mostrados en la Figura 2, ambos genes TNFRSF13C y CENPM fueron incluidos en la lista de genes de PML-419 (Tabla 6) sobre la base de su biología inmune o neurológica relacionada reportada

en la bibliografía. No se encontraron soluciones de PML para CNV o SNV para estos dos genes, pero el análisis String (alta confianza = 0.7, 1ª capa = hasta 10 interactuadores) muestra que TNFRSF13C interactúa con los genes de PML-419 TRAF3 (solución de la Tabla 7) y TNFRSF13B (solución de la Tabla 8), así como con BTK (un gen de PML conocido, véase la Tabla 6).

La Figura 3 muestra otro ejemplo de un gen que es afectado tanto por CNV de línea germinal como adquiridas. Mientras que ningún caso de PML fue resuelto sobre la base de las CNV adquiridas o de línea germinal que mostraron impactar el gen PKHD 1, la nominación de este gen a la Tabla 6 sobre la base de su biología resultó en el hallazgo de 3 potenciales soluciones alternativas (modelo AR) para otros 3 casos de PML (véase la Tabla 8). Sin embargo, el análisis String (alta confianza = 0.7, 1ª capa = hasta 10 interactuadores) no reveló ninguna interacción del gen de PML-419 con PKHD1.

Ejemplo 6

Los expertos en la materia pueden apreciar que un modelo de enfermedad AR implicaría determinar si ambos alelos (para un gen o locus genético) están afectados por una variante genética en individuos afectados por el trastorno. Los tipos de variantes genéticas pueden ser SNV, CNV, indels, etc. En el estudio descrito en el presente documento, si se invocaba un modelo de enfermedad AR para un gen (véase la Tabla 6), evaluábamos los datos CGH del paciente con PML en busca de CNV (heterocigotos u homocigotos) y sus datos de exoma en busca de SNV (heterocigotos u homocigotos). De este modo, cada paciente puede ser resuelto para uno de los genes de PML-419 (Tabla 6) con uno de los siguientes escenarios: eliminación homocigota, duplicación homocigota (la relación log2 aparecerá comparable a la encontrada típicamente para triplicaciones), SNV homocigota, SNV heterocigota compuesta, CNV heterocigota compuesta, o SNV heterocigota compuesta y CNV. Los expertos en la materia saben que, para un mecanismo de enfermedad AR, una SNV o CNV patogénica puede tener una frecuencia apreciable en la población general (p. ej., hasta un 1 % de frecuencia) con poco o ningún impacto en la salud del individuo, pero cuando está presente con una segunda variante patogénica en el otro alelo, puede causar la enfermedad.

La Figura 4 muestra un ejemplo de una pérdida intrónica que afecta al gen BMPR2. Se descubrió que el paciente PML29 tenía una eliminación homocigota, mientras que los pacientes PML58 y MVGS811-13a tenían una eliminación heterocigota. Asumiendo un modelo de enfermedad AR, no se encontraron soluciones de SNV para este gen; sin embargo, PML29 está potencialmente resuelto debido a la eliminación homocigota que se detectó. Mientras que la biología relacionada con el sistema inmune es reportada para estudios sobre BMPR2 (véase la Tabla 6), el análisis String (alta confianza = 0.7, 1ª capa = hasta 10 interactuadores) no reveló ninguna interacción del gen de PML-419 con BMPR2.

La Figura 5 muestra un ejemplo de una ganancia exónica que interrumpe el gen COMMD6. Se encontró que dos pacientes con PML tenían duplicaciones homocigotas de esta CNV. Curiosamente, mientras que el análisis String (alta confianza = 0.7, 1a capa = hasta 10 interactuadores) no reveló ninguna interacción del gen de PML-419 con COMMD6, estudios recientes (véase la Tabla 6, PMID 25355947 y 27441653) muestran un vínculo potencial entre COMMD6 y el conocido gen de PML WAS a través del gen WASH.

La Figura 6 muestra un ejemplo de una ganancia exónica que interrumpe el gen KCTD7 y su punto de ruptura derecho está secuencia arriba de RABGEF1 (es decir, uno o ambos genes pueden estar causando/contribuyendo a PML). Un ARN no codificante recientemente anotado (véase el ensamblaje hg19, LOC100996437) también puede estar afectado por esta CNV. Ambos genes tienen vínculos inmunológicos y neurológicos (véase la Tabla 6) y como el paciente PML29 tiene una duplicación homocigota, se añadió como solución de PML en la Tabla 7. El análisis String (confianza alta = 0.7, 1ª capa = hasta 10 interactuadores) no reveló ninguna interacción del gen de PML-419 para ninguno de los dos genes, pero están vinculados en un análisis String conjunto.

La Figura 7 muestra un ejemplo de una ganancia que interrumpe FPR2 (punto de ruptura izquierdo) y ZNF616 (punto de ruptura derecho, gen no marcado), y otros genes son totalmente abarcados por esta CNV. Hay un fuerte apoyo biológico para el FPR2 (véase la Tabla 6) y está listado como una solución de PML en la Tabla 7. El análisis String (alta confianza = 0.7, 1a capa = hasta 10 interactuadores) no reveló ninguna interacción del gen de PML-419 para FPR2, pero un análisis conjunto de los genes de la Tabla 7 sí reveló una interacción (véase la Figura 13).

La Figura 8 muestra un ejemplo de una pérdida exónica que afecta a los genes PIK3CD y PIK3CD-AS 1. El paciente MVGS811-13a tiene una eliminación homocigota y se informa como solución en la Tabla 7 basándose en la fuerte biología relacionada con el sistema inmune para PIK3CD (véase la Tabla 6). El análisis String (alta confianza = 0.7, 1ª capa = hasta 10 interactuadores) revela interacciones del gen de PML-419 para PTEN y PIK3R1.

Ejemplo 7

Un subconjunto de las CNV raras encontradas en nuestro estudio de PML estaban localizadas en regiones intergénicas. Aunque los expertos en la materia pueden apreciar que las variantes intergénicas (CNV, SNV,

etc.) pueden tener efectos de largo alcance sobre la expresión de los genes (p. ej., los elementos reguladores de los genes pueden estar localizados a varios kilobases de distancia de los genes bajo su influencia), en nuestro estudio asumimos que las CNV intergénicas tenían un impacto potencial sobre uno o ambos genes adyacentes si estaban localizadas a <~100 Kb de distancia, ya sea secuencia arriba o secuencia abajo. El proyecto ENCODE ha revelado una gran cantidad de información, tal como los sitios de unión a factores de transcripción, y las CNV raras que se identificaron en el presente estudio se comprobaron por su posible impacto en estos sitios (se comprobó la anotación ENCODE del ensamblaje hg19) y a menudo se encontró que afectaban a los sitios de unión a factores de transcripción y/o estaban localizadas en regiones de ADN conservadas.

La Figura 9 muestra una ganancia intergénica que se encuentra secuencia arriba de CD180. El paciente MVGS995-4a tiene una duplicación homocigota y, aunque no se considera como una solución de PML en la Tabla 7, es potencialmente una solución alternativa que puede estar causando o contribuyendo a la PML del paciente basándose en la expresión alterada de CD180. El gen tiene una biología relacionada con la inmunidad (véase la Tabla 6) y el análisis String (alta confianza = 0.7, 1a capa = hasta 10 interactuadores) revela una interacción del gen de PML-419 con PLCG2 (véase la Tabla 7, 2 casos de PML tienen una solución para este gen).

La Figura 10 muestra una pérdida intergénica que se encuentra secuencia arriba de VDAC1. El paciente PML30 tiene una eliminación homocigota y, aunque no se considera como una solución de PML en la Tabla 7, es potencialmente una solución alternativa que puede estar causando o contribuyendo a la PML del paciente basada en la expresión alterada de VDAC1. El análisis String (alta confianza = 0.7, 1ª capa = hasta 10 interactuadores) no reveló ninguna interacción del gen de PML-419 para VDAC1.

La Figura 11 muestra una pérdida intergénica que está secuencia abajo de EGR1 y ETF1. El paciente PML69 tiene una eliminación homocigota y, basándose en los vínculos de EGR1 con los genes de PML-419 (Tabla 6) y su proximidad a EGR1 (~4 Kb de distancia), se añadió como una posible solución de PML en la Tabla 7. El análisis String (confianza alta = 0.7, 1ª capa = hasta 10 interactuadores) revela interacciones del gen de PML-419 con JUN, PTEN y TP53), pero no se encontró nada digno de mención en el análisis String de ETF1.

La Figura 12 muestra una pérdida intergénica que está secuencia arriba de ITSN2. El paciente PML65 tiene una eliminación homocigota y, basándose en los enlaces de ITSN2 con un gen de PML conocido (WAS) en la lista de genes de PML-419 (Tabla 6), se añadió como una posible solución de PML en la Tabla 7. Curiosamente, se encontró que otro caso de PML tenía una rara SNV homocigota en ITSN2, por lo que este gen tiene 2 soluciones de PML reportadas en la Tabla 7. El análisis String (alta confianza = 0.7, 1ª capa = hasta 10 interactuadores) no reveló ninguna interacción del gen de PML-419.

Ejemplo 8

Los análisis de rutas, tal como las interacciones proteína-proteína, están proporcionando información valiosa sobre la biología subyacente de enfermedades complejas. Mientras que la PML es una enfermedad muy rara que requiere varios factores concurrentes (p. ej., la infección por el virus de JC), múltiples genes pueden estar causando o aumentando independientemente el riesgo de desarrollar este trastorno neurodegenerativo basado en la presencia de una variante genética en un gen dado (p. ej., una variante heterocigótica en la que una variante deletérea está presente en el alelo heredado maternal o paternalmente, una variante homocigótica en la que la misma variante deletérea está presente en ambos alelos, o variantes heterocigotas compuestas en las que un par de variantes deletéreas están presentes pero una se encuentra en el alelo heredado por vía materna y la otra en el alelo heredado por vía paterna). Según la hipótesis planteada, la presencia de un trastorno genético por inmunodeficiencia era otro requisito previo. De hecho, en el estudio de PML descrito en el presente documento, 43 genes fueron propuestos como soluciones para 61 de 71 casos de PML (véase la Tabla 7) que fueron evaluados usando arreglo CGH y secuenciación del exoma completo. Se han desarrollado numerosos algoritmos y bases de datos asociadas para investigar vías moleculares, tal como String (véase, Szklarczyk et al., (2015), y las referencias citadas allí).

La Figura 13 muestra un ejemplo de análisis String realizado en los 43 genes considerados como soluciones de PML sobre la base de un modelo de enfermedad AD o AR. Se encontró una serie de interacciones para 21 de 43 genes, y en varios casos esto incluyó interacciones para genes implicados en 2 o más casos de PML que se reportan en la Tabla 7 (9 casos para TNFRSF11A, 4 casos para PLCG2, 3 casos para ZAP70 y NOD2, y 2 casos de PML para TICAM1).

Ejemplo 9

Para determinar la probabilidad de que un individuo seleccionado al azar albergara una de las variantes descritas en el presente documento, se realizó el siguiente análisis: para cada variante o combinación de variantes, se utilizó la frecuencia específica de la etnia citada en la Tabla 7 para determinar la probabilidad de que un individuo seleccionado al azar de la misma etnia no albergara la variante o combinación de variantes. Se calculó el producto de todas esas probabilidades (p. ej., la probabilidad de que un individuo seleccionado al

azar no albergara ninguna de las variantes) y se restó de 1, obteniéndose la probabilidad de que un individuo seleccionado al azar albergara al menos una de las variantes. Se descubrió que, para los casos de VIH, la probabilidad de que un individuo aleatorio albergara al menos una de las variantes era del ~5 %, lo que concuerda con el riesgo de PML previo a la HAART en el contexto del VIH. En los casos no relacionados con el VIH (principalmente MS/NTZ), el riesgo fue de ~ 1 %, lo que, de nuevo, es coherente con el riesgo de PML en la MS/NTZ, especialmente tras una terapia a largo plazo.

Estos análisis apoyan la noción de que las frecuencias de las variantes identificadas como relevantes para el riesgo de PML son consistentes con los riesgos reales observados para individuos no seleccionados. Los análisis se basan en la suposición razonable de que no existe ninguna conexión relevante para la PML con el riesgo de desarrollar VIH (una infección adquirida) y/o MS (p. ej., esto implica que el tratamiento de individuos sanos con Natalizumab, p. ej., resultaría en riesgos similares de PML). Cualquier desviación (p. ej., variantes encontradas en un número ligeramente mayor de individuos normales de lo esperado de acuerdo con el número realmente observado de afectados por PML) puede deberse a: penetrancia (p. ej., no todas las personas con las variantes tendrán el máximo riesgo de PML); la suposición de que los individuos con MS, VIH y otras condiciones subyacentes representaban una sección transversal normal (p. ej., con respecto al riesgo de PML) de la población general, antes de desarrollar los trastornos subyacentes VIH, MS, etc.; y la infravaloración de la PML, incluso en pacientes con VIH, MS/NTZ.

Ejemplo 10 - Tablas referenciadas en este estudio

Tabla 1: CNV de interés en este estudio

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
1	1086119	1135772	49653	het loss	3009	MIR200A,MIR200B,MIR429,TNFRSF1 8, TTLL10	1
1	9634094	9635206	1112	hom loss	3009	PIK3CD	2
1	12018512	12032581	14069	ganancia	3205		3
1	19593401	19602807	9406	het loss	3203	CAPZB	4
1	21695957	21700243	4286	het loss	3161		5
1	24364786	24391166	26380	ganancia	3199	IFNLR1	6
1	28666669	28737671	71002	ganancia	3161	PHACTR4,RCC1,SNHG3	7
1	49372054	49380088	8034	het loss	3145	AGBL4	8
1	153816159	15382769 8	11539	het loss	3168		9
1	205607255	20561034 1	3086	ganancia	3007		10
1	215760485	21576245 1	1966	het loss	3117	GPATCH2	11
1	215866737	21586990 0	3163	het loss	3151	GPATCH2	12
2	10352668	10356083	3415	het loss	3007		13
2	24457024	24462631	5607	hom loss	3204		14
2	38468717	38471950	3233	het loss	3175		15
2	38516138	38524237	8099	het loss	3151		16
2	38726517	38731845	5328	het loss	3159		17
2	40620890	40624089	3199	het loss	3202		18
2	46631006	46643501	12495	ganancia	3145	RHOQ	19
2	55764753	55790559	25806	ganancia	3143	PNPT1	20
2	55764753	55790559	25806	ganancia	3193	PNPT1	20
2	55764753	55790559	25806	ganancia	3282	PNPT1	20
2	55764753	55790559	25806	ganancia	3143	PNPT1	20
2	55764753	55790559	25806	ganancia	3193	PNPT1	20
2	55764753	55790559	25806	ganancia	3282	PNPT1	20

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
2	55764753	55790559	25806	ganancia	3143	PNPT1	20
2	55764753	55790559	25806	ganancia	3193	PNPT1	20
2	55764753	55790559	25806	ganancia	3282	PNPT1	20
2	71190677	71200120	9443	het loss	3175	MCEE	21
2	71190677	71200120	9443	het loss	3175	MCEE	21
2	71191311	71200120	8809	het loss	3204	MCEE	22
2	71198108	71200120	2012	het loss	3143	MCEE	23
2	71190677	71200120	9443	het loss	3175	MCEE	21
2	71198108	71200120	2012	het loss	3193	MCEE	23
2	71198108	71200120	2012	het loss	3200	MCEE	23
2	71191311	71200120	8809	het loss	3204	MCEE	22
2	74773432	74913493	140061	ganancia	3118	HK2	24
2	105418748	105435274	16526	het loss	3193	FHL2	25
2	110182348	110210249	27901	ganancia	3174	MALL,MIR4267,MIR4436B1,MIR4436B2	26
2	127823042	127828410	5368	het loss	3273		27
2	134911636	134914254	2618	het loss	3273	MGAT5	28
2	203005216	203019933	14717	het loss	3009	BMPR2	29
2	203005216	203019933	14717	het loss	3192	BMPR2	29
2	203005216	203019933	14717	hom loss	3152	BMPR2	29
2	230212897	230216339	3442	het loss	3154	DNER	30
3	122979920	122994402	14482	ganancia	3202	IQCB1	31
4	26565071	26566345	1274	het loss	3010	STIM2	32
4	26565071	26566345	1274	het loss	3125	STIM2	32
4	26565071	26566345	1274	het loss	3168	STIM2	32
4	26565071	26566345	1274	het loss	3282	STIM2	32
4	26565071	26566345	1274	het loss	3284	STIM2	32

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
4	26565071	26566345	1274	hom loss	3273	STIM2	32
4	54838623	54873909	35286	ganancia	3153	PDGFRA	33
4	90791460	90843887	52427	ganancia	3168		34
4	90800863	90808258	7395	het loss	3009		35
4	90800863	90808258	7395	het loss	3284		35
5	45331278	46150784	819506	ganancia	3157	HCN1	36
5	49771219	49774457	3238	ganancia	3273	MSB	37
5	66619415	66636116	16701	ganancia	3010		38
5	78480194	78497296	17102	ganancia	3205		39
5	78497296	78531091	33795	ganancia	3132		40
5	78497296	78521408	24112	ganancia	3185		41
5	78497296	78531091	33795	ganancia	3132		40
5	78497296	78521408	24112	ganancia	3185		41
5	78500552	78526637	26085	ganancia	3205		42
5	78497296	78531091	33795	ganancia	3132		40
5	78500552	78526637	26085	ganancia	3205		42
5	78497296	78531091	33795	ganancia	3132		40
5	83490494	83495169	4675	het loss	3204	EDIL3	43
5	133372071	133379727	7656	hom loss	3153		44
5	137836466	137843309	6843	hom loss	3279		45
5	150159466	150202601	43135	het loss	3117		46
5	150159466	150204134	44668	het loss	3180		47
5	150159466	150202601	43135	het loss	3199		46
5	150159466	150204134	44668	het loss	3278		47
5	150159466	150202601	43135	het loss	3117		46
5	150159466	150204134	44668	het loss	3180		47

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
5	150159466	150202601	43135	het loss	3199		46
5	150159466	150204134	44668	het loss	3278		47
5	150159466	150202601	43135	het loss	3117		46
5	150159466	150204134	44668	het loss	3180		47
5	150159466	150202601	43135	het loss	3199		46
5	150159466	150204134	44668	het loss	3278		47
5	150185190	150201145	15955	hom loss	3009		48
5	150185190	150201145	15955	hom loss	3143		48
5	150185190	150202601	17411	hom loss	3152		49
5	150185190	150202601	17411	hom loss	3154		49
5	150185190	150202601	17411	hom loss	3193		49
5	150159466	150201145	41679	hom loss	3196		50
5	150185190	150201145	15955	hom loss	3281		48
5	150185190	150201145	15955	hom loss	3009		48
5	150185190	150201145	15955	hom loss	3143		48
5	150185190	150202601	17411	hom loss	3152		49
5	150185190	150202601	17411	hom loss	3154		49
5	150185190	150202601	17411	hom loss	3193		49
5	150159466	150201145	41679	hom loss	3196		50
5	150185190	150201145	15955	hom loss	3281		48
5	150185190	150202601	17411	hom loss	3152		49
5	150185190	150202601	17411	hom loss	3154		49
5	150185190	150202601	17411	hom loss	3193		49
5	150185190	150204134	18944	het loss	3132		51
5	150159466	150204134	44668	het loss	3180		47
5	150202601	150204134	1533	het loss	3196		52

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
5	150191322	150204134	12812	het loss	3273		53
5	150185190	150204134	18944	het loss	3277		51
5	150159466	150204134	44668	het loss	3278		47
5	150185190	150204134	18944	het loss	3280		51
5	150185190	150204134	18944	het loss	3282		51
5	179590681	179626660	35979	het loss	3172	MAPK9	54
6	2882577	2947403	64826	het loss	3196	DKFZP686I15217,NQO2,SERPINB6	55
6	2964646	2966011	1365	het loss	3193	HTATSF1P2,NQO2	56
6	51766024	51773250	7226	het loss	3167	PKHD1	57
6	51952217	51969378	17161	ganancia	3127	PKHD1	58
6	51952217	51969378	17161	ganancia	3127	PKHD1	58
6	51953476	51965723	12247	ganancia	3205	PKHD1	59
6	51952217	51969378	17161	ganancia	3127	PKHD1	58
6	74396294	74404837	8543	het loss	3009	SLC17A5	60
6	74396294	74398409	2115	het loss	3160	SLC17A5	61
6	74396294	74404837	8543	het loss	3009	SLC17A5	60
6	86416979	86431527	14548	het loss	3197		62
6	91131823	91135670	3847	het loss	3171		63
6	107882367	107890605	8238	het loss	3201	PDSS2	64
6	166418511	166422386	3875	het loss	3125		65
6	166418511	166422386	3875	het loss	3163		65
6	166418511	166422386	3875	het loss	3192		65
6	166418511	166422386	3875	het loss	3193		65
6	166418511	166422386	3875	het loss	3194		65
6	166418511	166422386	3875	het loss	3200		65
6	166418511	166422386	3875	het loss	3205		65

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
6	166418511	166422386	3875	het loss	3280		65
6	166418511	166422386	3875	het loss	3281		65
6	166418511	166422386	3875	het loss	3284		65
6	166418511	166422386	3875	hom loss	3009		65
6	166418511	166422386	3875	hom loss	3152		65
6	166418511	166422386	3875	hom loss	3175		65
7	65741238	65768682	27444	ganancia	3152	KCTD7	66
7	65741238	65768682	27444	ganancia	3202	KCTD7	66
7	157174966	157177843	2877	het loss	3009	PTPRN2	67
7	157425841	157496238	70397	ganancia	3189	PTPRN2	68
7	158000082	158024569	24487	het loss	3279	PTPRN2	69
7	158000082	158024569	24487	het loss	3279	PTPRN2	69
7	158000082	158024569	24487	het loss	3279	MIR595,PTPRN2	69
8	23103186	23125443	22257	het loss	3140	TNFRSF10A	70
8	39914488	39919594	5106	het loss	3126	IDO2	71
8	79905654	79910286	4632	het loss	3159		72
8	99790200	99799839	9639	het loss	3006	STK3	73
8	102049360	102064431	15071	het loss	3173		74
8	102049360	102064431	15071	het loss	3175		74
8	102049360	102064431	15071	het loss	3282		74
9	571398	584647	13249	het loss	3006	KANK1	75
9	571398	584647	13249	het loss	3006	KANK1	75
9	580722	598488	17766	het loss	3200	KANK1	76
9	580722	598488	17766	het loss	3282	KANK1	76
9	580722	598488	17766	het loss	3200	KANK1	76
9	580722	598488	17766	het loss	3282	KANK1	76

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
9	634039	637589	3550	het loss	3273	KANK1	77
9	634039	637589	3550	het loss	3282	KANK1	77
9	74050088	74059447	9359	het loss	3165	GDA	78
9	93140394	93447826	307432	ganancia	3198	AUH,MIR3163,MIR3910-1,MIR3910-2,NFIL3	79
9	118564159	118575633	11474	ganancia	3193	ASTN2	80
9	118612694	118664593	51899	het loss	3144	ASTN2	81
9	119220847	119233078	12231	ganancia	3005		82
10	899657	1071401	171744	ganancia	3161	GTPBP4,IDI2,IDI2-AS1,LARP4B	83
10	76217585	76411591	194006	ganancia	3179	KAT6B	84
10	116000069	116004388	4319	ganancia	3010	VWA2	85
11	14677012	14689025	12013	het loss	3199	PDE3B	86
11	34608313	34615878	7565	het loss	3117	EHF	87
11	62382087	62398462	16375	het loss	3205	SLC3A2	88
11	76631014	76643625	12611	het loss	3193	GDPD4	89
12	11616557	12422129	805572	het loss	3126	ETV6	90
12	12435301	12778142	342841	het loss	3126	APOLD1,CDKN1B,CREBL2,DUSP16,GPR19,LOH12CR1	91
12	12968705	12971310	2605	ganancia	3127		92
12	91786998	94313682	2526684	het loss	3126	EEA1,LOC643339	93
12	91786998	94313682	2526684	het loss	3126	LOC643339,MRPL42,NUDT4,NUDT4P1,SOCS2,SOCS2-AS1,UBE2N	93
12	91786998	94313682	2526684	het loss	3126	CCDC41,CRADD,PLXNC1	93
12	111061085	111064486	3401	het loss	3004	TRAFD1	94
13	40939924	41026908	86984	ganancia	3140	RGCC	95
13	75006025	75016304	10279	ganancia	3009	COMD6	96
13	75006025	75016304	10279	ganancia	3152	COMD6	96

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
13	91811087	91814369	3282	het loss	3143	GPC5	97
13	91811087	91811118	31	hom loss	3173	GPC5	98
13	110754499	110778301	23802	ganancia	3006	ARHGEF7,TEX29	99
14	20021118	20055469	34351	ganancia	3205	RNASE10	100
14	20426824	20481852	55028	hom loss	3200	ECRP,RNASE3	101
14	20430810	20490129	59319	het loss	3192	ECRP	102
14	20430810	20490129	59319	het loss	3192		102
14	20430810	20490129	59319	het loss	3192		102
14	21096689	21105611	8922	het loss	3125		103
14	21096689	21105611	8922	het loss	3175		103
14	21096689	21105611	8922	het loss	3194		103
14	21096689	21105611	8922	het loss	3204		103
14	21096689	21105611	8922	het loss	3273		103
14	21120750	21125513	4763	ganancia	3143		104
14	21120750	21125513	4763	ganancia	3173		104
14	60901636	60909492	7856	het loss	3193	PRKCH	105
14	60912874	60921269	8395	het loss	3174	PRKCH	106
14	63937192	63944459	7267	ganancia	3205	MTHFD1	107
14	95754535	95759056	4521	het loss	3009	BDKRB2	108
14	95754535	95759056	4521	het loss	3173	BDKRB2	108
14	95754535	95759056	4521	het loss	3202	BDKRB2	108
15	66065925	66082418	16493	het loss	3010		109
15	70432627	70443017	10390	ganancia	3169	HEXA	110
15	75096101	75128723	32622	ganancia	3200	PSTPIP1	111
15	75101524	75115806	14282	ganancia	3132	PSTPIP1	112
15	75096101	75128723	32622	ganancia	3200	PSTPIP1	111

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
15	75105789	75115806	10017	ganancia	3127	PSTPIP1	113
15	75101524	75115806	14282	ganancia	3132	PSTPIP1	112
15	75105789	75115806	10017	ganancia	3199	PSTPIP1	113
15	75096101	75128723	32622	ganancia	3200	PSTPIP1	111
15	75105789	75115806	10017	ganancia	3279	PSTPIP1	113
15	75105789	75115806	10017	ganancia	3127	PSTPIP1	113
15	75101524	75115806	14282	ganancia	3132	PSTPIP1	112
15	75105789	75115806	10017	ganancia	3199	PSTPIP1	113
15	75096101	75128723	32622	ganancia	3200	PSTPIP1	111
15	75105789	75115806	10017	ganancia	3279	PSTPIP1	113
15	75096101	75128723	32622	ganancia	3200	PSTPIP1	111
15	88999998	89016848	16850	het loss	3172		114
16	6823677	6932753	109076	het loss	3126	RBFOX1	115
16	6823677	6932753	109076	het loss	3126	RBFOX1	115
16	6942078	6945539	3461	ganancia	3173	RBFOX1	116
16	6942078	6945539	3461	ganancia	3175	RBFOX1	116
16	6942078	6945539	3461	ganancia	3282	RBFOX1	116
16	23842653	23848772	6119	het loss	3198	PRKCB	117
16	23892842	23903495	10653	ganancia	3199	PRKCB	118
16	23892842	23903495	10653	ganancia	3199	PRKCB	118
16	23893969	23908248	14279	ganancia	3205	PRKCB	119
16	23893969	23908248	14279	ganancia	3205	PRKCB	119
16	69044235	69050151	5916	ganancia	3174	FUK	120
16	69044235	69050151	5916	ganancia	3185	FUK	120
16	69052450	69081640	29190	het loss	3197	COG4,FUK	121
16	70653499	70665447	11948	ganancia	3143	HPR	122

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
16	70653499	70665447	11948	ganancia	3152	HPR	122
16	70653499	70665447	11948	ganancia	3192	HPR	122
16	70653499	70665447	11948	ganancia	3200	HPR	122
16	70653499	70665447	11948	ganancia	3282	HPR	122
16	70653499	70665447	11948	ganancia	3284	HPR	122
17	69341925	70202523	860598	ganancia	3183	BTBD17,C17orf77,CD300A,CD300C,CD300E,CD300LB,CD300LD,CD300LF,DNAI2,GPR142,GPRC5C,KIF19,MGC16275,RAB37,RPL38,TTSH2	123
17	75608151	75615433	7282	het loss	3144	TBC1D16	124
17	75608151	75615433	7282	het loss	3152	TBC1D16	124
17	75608151	75615433	7282	het loss	3163	TBC1D16	124
17	75608151	75611602	3451	het loss	3192	TBC1D16	125
17	75608151	75615433	7282	het loss	3200	TBC1D16	124
17	75608151	75611602	3451	het loss	3204	TBC1D16	125
17	75608151	75611602	3451	het loss	3284	TBC1D16	125
17	75608151	75611602	3451	hom loss	3009	TBC1D16	125
17	75611602	75615433	3831	hom loss	3175	TBC1D16	126
17	75608151	75615433	7282	het loss	3144	TBC1D16	124
17	75608151	75615433	7282	het loss	3152	TBC1D16	124
17	75608151	75615433	7282	het loss	3163	TBC1D16	124
17	75608151	75615433	7282	het loss	3200	TBC1D16	124
17	76241510	76267844	26334	ganancia	3205	RPTOR	127
17	76247305	76265683	18378	ganancia	3127	RPTOR	128
17	76241510	76267844	26334	ganancia	3205	RPTOR	127
17	76241510	76267844	26334	ganancia	3205	RPTOR	127
18	9985530	10125331	139801	ganancia	3175		129

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
18	12764095	12781985	17890	ganancia	3191	PTPN2	130
18	27026203	27029351	3148	het loss	3125		131
18	27026203	27029351	3148	het loss	3143		131
18	27026203	27029351	3148	het loss	3175		131
18	42537949	42663605	125656	ganancia	3125	PIAS2,ST8SIA5	132
18	46917195	46945018	27823	het loss	3161		133
18	59457622	59465699	8077	het loss	3145	SERPINB4	134
19	3270755	3291144	20389	ganancia	3205		135
19	46386511	46388364	1853	hom loss	3175		136
19	52496536	52501292	4756	ganancia	3124		137
19	55247874	55252420	4546	het loss	3163	FLJ26850	138
19	55247874	55252420	4546	het loss	3173	FLJ26850	138
19	55247874	55252420	4546	het loss	3192	FLJ26850	138
19	55247874	55252420	4546	het loss	3200	FLJ26850	138
19	55247874	55252420	4546	het loss	3280	FLJ26850	138
19	55247874	55252420	4546	het loss	3163	FLJ26850	138
19	55247874	55252420	4546	het loss	3173	FLJ26850	138
19	55247874	55252420	4546	het loss	3192	FLJ26850	138
19	55250187	55252420	2233	het loss	3194	FLJ26850	139
19	55247874	55252420	4546	het loss	3200	FLJ26850	138
19	55247874	55252420	4546	het loss	3280	FLJ26850	138
19	55250187	55252420	2233	hom loss	3175	FLJ26850	139
19	55250187	55252420	2233	hom loss	3202	FLJ26850	139
19	56964168	57308449	344281	ganancia	3155	FPR2,FPR3,ZNF350,ZNF432,ZNF577,ZNF613,ZNF614,ZNF615,ZNF649,ZNF841	140

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
19	56964168	57308449	344281	ganancia	3157	FPR2,FPR3,ZNF350,ZNF432,ZNF577,ZNF613,ZNF614,ZNF615,ZNF649,ZNF841	140
19	59013780	59023850	10070	het loss	3117	NLRP12	141
19	59249279	59251831	2552	hom loss	3160	VSTM1	142
19	59249279	59251831	2552	hom loss	3164	VSTM1	142
19	59250742	59251831	1089	hom loss	3117	VSTM1	143
19	59249279	59251831	2552	hom loss	3160	VSTM1	142
19	59249279	59251831	2552	hom loss	3164	VSTM1	142
19	59250742	59251831	1089	hom loss	3277	VSTM1	143
20	17844577	17954650	110073	ganancia	3166	MGME1,OVOL2,SNORD17,SNX5	144
20	42706680	42711434	4754	het loss	3125	ADA	145
21	15234620	15312960	78340	ganancia	3009	NRIP1	146
21	29643302	29647950	4648	het loss	3202	BACH1	147
21	44634707	44666832	32125	ganancia	3200	TRPM2	148
21	44634707	44641658	6951	ganancia	3205	TRPM2	149
21	44634707	44671482	36775	ganancia	3279	TRPM2	150
21	44637544	44669596	32052	ganancia	3127	TRPM2	151
21	44637544	44657372	19828	ganancia	3185	TRPM2	152
21	44634707	44666832	32125	ganancia	3200	TRPM2	148
21	44634707	44641658	6951	ganancia	3205	TRPM2	149
21	44634707	44671482	36775	ganancia	3279	TRPM2	150
21	44637544	44669596	32052	ganancia	3127	TRPM2	151
21	44637544	44657372	19828	ganancia	3185	TRPM2	152
21	44634707	44666832	32125	ganancia	3200	TRPM2	148
21	44634707	44671482	36775	ganancia	3279	TRPM2	150
21	44643974	44657372	13398	het loss	3161	TRPM2	153

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
21	44637544	44669596	32052	ganancia	3127	TRPM2	151
21	44637544	44657372	19828	ganancia	3185	TRPM2	152
21	44634707	44666832	32125	ganancia	3200	TRPM2	148
21	44643974	44657372	13398	ganancia	3205	TRPM2	153
21	44634707	44671482	36775	ganancia	3279	TRPM2	150
21	44637544	44669596	32052	ganancia	3127	TRPM2	151
21	44634707	44666832	32125	ganancia	3200	TRPM2	148
21	44634707	44671482	36775	ganancia	3279	TRPM2	150
21	44637544	44669596	32052	ganancia	3127	TRPM2	151
21	44634707	44666832	32125	ganancia	3200	TRPM2	148
21	44660199	44681194	20995	ganancia	3205	TRPM2	154
21	44634707	44671482	36775	ganancia	3279	TRPM2	150
21	44637544	44669596	32052	ganancia	3127	TRPM2	151
21	44660199	44681194	20995	ganancia	3205	TRPM2	154
21	44634707	44671482	36775	ganancia	3279	TRPM2	150
21	44660199	44681194	20995	ganancia	3205	TRPM2	154
21	44634707	44671482	36775	ganancia	3279	TRPM2	150
21	44660199	44681194	20995	ganancia	3205	TRPM2	154
21	45348895	45354820	5925	het loss	3179	ADARB1	155
22	37689058	37715385	26327	ganancia	3169	APOBEC3A,APOBEC3A B,APOBEC3B	156
22	39257585	39261621	4036	het loss	3005	MKL1	157
22	40642402	40655210	12808	ganancia	3205	TNFRSF13C	158
22	40656820	40673250	17430	ganancia	3185		159
22	40656820	40675788	19968	ganancia	3205		160
22	40659633	40671866	12233	ganancia	3127		161
22	40656820	40673250	17430	ganancia	3185		159

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
22	40655820	40675788	19968	ganancia	3205		160
22	40659633	40671866	12233	ganancia	3127	CENPM	161
22	40655820	40673250	17430	ganancia	3185	CENPM	159
22	40663050	40668079	5029	ganancia	3190	CENPM	162
22	40663050	40668079	5029	ganancia	3202	CENPM	162
22	40655820	40675788	19968	ganancia	3205	CENPM	160
22	40659633	40671866	12233	ganancia	3127	CENPM	161
22	40655820	40673250	17430	ganancia	3185	CENPM	159
22	40655820	40675788	19968	ganancia	3205	CENPM	160
22	40655820	40673250	17430	ganancia	3185	CENPM	159
22	40655820	40675788	19968	ganancia	3205	CENPM	160
22	40655820	40675788	19968	ganancia	3205		160
23	232907	244684	11777	het loss	3007	PPP2R3B	163
23	7585301	7830994	245693	ganancia	3172		164
23	7585301	7830994	245693	ganancia	3172	VCX	164
23	7769323	7779354	10031	het loss	3132		165
23	6465033	8093113	1628080	het loss	3171		166
23	7769323	7779354	10031	het loss	3204		165
23	7585301	7830994	245693	ganancia	3172		164
23	7585301	7830994	245693	ganancia	3172		164
23	6465033	8093113	1628080	het loss	3171	MIR651,PNPLA4	166
23	7585301	7830994	245693	ganancia	3172	PNPLA4	164
23	48350646	48408854	50208	het loss	3009		167
23	64710574	64725828	15254	ganancia	3125		168
23	73083877	73086192	2315	hom loss	3193	JPX	169
23	73083877	73086192	2315	hom loss	3200	JPX	169

Chr	Inicio de CNV original	Parada de CNV original	Tamaño original de la CNV	Tipo de CNV	ID de caso PML	Símbolo del gen de RefSeq	SEQ ID
23	122337025	122340879	3854	hom loss	3125	GRIA3	170
23	148452844	148461889	9045	het loss	3163		171
23	148452844	148461889	9045	het loss	3205		171
23	148452844	148461889	9045	hom loss	3144		171
23	148452844	148461889	9045	hom loss	3193		171
23	149901706	149904265	2559	ganancia	3117	HMGB3	172
23	149901706	149904265	2559	ganancia	3118	HMGB3	172

La Tabla 1 enumera todas las CNV de interés, obtenidas como se describe en el texto, con la excepción de que, para cada entrada, se anotan las posiciones originales de inicio y parada de la CNV, junto con el tamaño original de la CNV, tipo (pérdida heterocigota, pérdida o ganancia homocigota), ID del caso y anotación del gen (para la subregión de CNV NO CNV original). La última columna contiene los números de la SEQ_ID. En la Tabla 1 se utiliza la numeración cromosómica estándar utilizada por los expertos en la materia para los cromosomas autosómicos (1-22) pero, por comodidad con los métodos de análisis, el cromosoma X se designa en el presente documento como cromosoma 23. Todas las coordenadas se basan en hg18.

Tabla 2: subregiones de CNV de interés en este estudio

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
1	1086119	1135772	49653	het_loss	3009	MIR200A,MIR200B, MIR429,TNF RSF18, TTLL10	S	0	1	0.005115965	39.43	1
1	9634094	9635206	1112	hom_loss	3009	PIK3CD	S	0	1	0.005115965	39.43	2
1	12018512	12032581	14069	ganancia	3205		N	0	1	0.005115965	39.43	3
1	19593401	19602807	9406	het_loss	3203	CAPZB	N	0	1	0.005115965	39.43	4
1	21698753	21700243	1490	het_loss	3161		N	0	1	0.005115965	39.43	5
1	24364786	24391166	26380	ganancia	3199	IFNLR1	S	0	1	0.005115965	39.43	6
1	28666669	28737671	71002	ganancia	3161	PHACTR4,RCC1,SNHG3	S	0	1	0.005115965	39.43	7
1	49372054	49380088	8034	het_loss	3145	AGBL4	N	0	1	0.005115965	39.43	8
1	153816159	153827698	11539	het_loss	3168		N	0	1	0.005115965	39.43	9
1	205607255	205610341	3086	ganancia	3007		N	0	1	0.005115965	39.43	10
1	215760485	215762451	1966	het_loss	3117	GPATCH2	N	0	1	0.005115965	39.43	11
1	215866737	215869900	3163	het_loss	3151	GPATCH2	N	0	1	0.005115965	39.43	12
2	10352668	10356083	3415	het_loss	3007		N	0	1	0.005115965	39.43	13
2	24457024	24462631	5607	hom_loss	3204		N	0	1	0.005115965	39.43	14
2	38468717	38471950	3233	het_loss	3175		N	0	1	0.005115965	39.43	15
2	38516138	38524237	8099	het_loss	3151		N	0	1	0.005115965	39.43	16
2	38726517	38731845	5328	het_loss	3159		N	0	1	0.005115965	39.43	17
2	40620890	40624089	3199	het_loss	3202		N	0	1	0.005115965	39.43	18
2	46631006	46643501	12495	ganancia	3145	RHOQ	N	0	1	0.005115965	39.43	19
2	55764753	55771586	6833	ganancia	3143	PNPT1	S	1	3	0.001318303	40.7	20
2	55764753	55771586	6833	ganancia	3193	PNPT1	S	1	3	0.001318303	40.7	21
2	55764753	55771586	6833	ganancia	3282	PNPT1	S	1	3	0.001318303	40.7	22
2	55771587	55772965	1378	ganancia	3143	PNPT1	N	2	3	0.003126725	20.33	23
2	55771587	55772965	1378	ganancia	3193	PNPT1	N	2	3	0.003126725	20.33	24

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
2	55771587	55772965	1378	ganancia	3282	PNPT1	N	2	3	0.003126725	20.33	25
2	55772966	55790559	17593	ganancia	3143	PNPT1	S	1	3	0.001318303	40.7	26
2	55772966	55790559	17593	ganancia	3193	PNPT1	S	1	3	0.001318303	40.7	27
2	55772966	55790559	17593	ganancia	3282	PNPT1	S	1	3	0.001318303	40.7	28
2	71190677	71191310	633	het_loss	3175	MCEE	S	0	1	0.005115965	39.43	29
2	71191311	71198107	6796	het_loss	3175	MCEE	N	1	2	0.014314826	26.77	30
2	71191311	71198107	6796	het_loss	3204	MCEE	N	1	2	0.014314826	26.77	31
2	71198108	71200120	2012	het_loss	3143	MCEE	N	2	5	3.02E-05	34.83	32
2	71198108	71200120	2012	het_loss	3175	MCEE	N	2	5	3.02E-05	34.83	33
2	71198108	71200120	2012	het_loss	3193	MCEE	N	2	5	3.02E-05	34.83	34
2	71198108	71200120	2012	het_loss	3200	MCEE	N	2	5	3.02E-05	34.83	35
2	71198108	71200120	2012	het_loss	3204	MCEE	N	2	5	3.02E-05	34.83	36
2	74827730	74913493	85763	ganancia	3118	HK2	S	0	1	0.005115965	39.43	37
2	105418748	105435274	16526	het_loss	3193	FHL2	S	0	1	0.005115965	39.43	38
2	110182348	110210249	27901	ganancia	3174	MALL,MIR4267, MIR4436B1,MIR 4436B2	S	2	1	0.198831257	6.6	39
2	127823042	127828410	5368	het_loss	3273		N	0	1	0.005115965	39.43	40
2	134911636	134914254	2618	het_loss	3273	MGAT5	N	0	1	0.005115965	39.43	41
2	203005216	203019933	14717	het_loss	3009	BMPR2	N	2	2	0.02731135	13.37	42
2	203005216	203019933	14717	het_loss	3192	BMPR2	N	2	2	0.02731135	13.37	43
2	203005216	203019933	14717	hom_loss	3152	BMPR2	N	0	1	0.005115965	39.43	44
2	230212897	230216339	3442	het_loss	3154	DNER	N	0	1	0.005115965	39.43	45
3	122979920	122994402	14482	ganancia	3202	IQCB1	S	0	1	0.005115965	39.43	46
4	26565071	26566345	1274	het_loss	3010	STIM2	N	85	5	0.671895631	0.75	47
4	26565071	26566345	1274	het_loss	3125	STIM2	N	85	5	0.671895631	0.75	48
4	26565071	26566345	1274	het_loss	3168	STIM2	N	85	5	0.671895631	0.75	49

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
4	26565071	26566345	1274	het_loss	3282	STIM2	N	85	5	0.671895631	0.75	50
4	26565071	26566345	1274	het_loss	3284	STIM2	N	85	5	0.671895631	0.75	51
4	26565071	26566345	1274	hom_loss	3273	STIM2	N	1	1	0.13732578	13.21	52
4	54838623	54873909	35286	ganancia	3153	PDGFRA	S	0	1	0.005115965	39.43	53
4	90791460	90843887	52427	ganancia	3168		N	0	1	0.005115965	39.43	54
4	90800863	90808258	7395	het_loss	3009		N	0	2	0.005115965	66.59	55
4	90800863	90808258	7395	het_loss	3284		N	0	2	0.005115965	66.59	56
5	45331278	45785151	453873	ganancia	3157	HCN1	S	0	1	0.005115965	39.43	57
5	49771219	49774457	3238	ganancia	3273	MSB	S	0	1	0.005115965	39.43	58
5	66619415	66636116	16701	ganancia	3010		N	0	1	0.005115965	39.43	59
5	78480194	78497296	17102	ganancia	3205		N	0	1	0.005115965	39.43	60
5	78497296	78500551	3255	ganancia	3132		N	0	2	0.005115965	66.59	61
5	78497296	78500551	3255	ganancia	3185		N	0	2	0.005115965	66.59	62
5	78500552	78521408	20856	ganancia	3132		N	0	3	2.49E-05	94.48	63
5	78500552	78521408	20856	ganancia	3185		N	0	3	2.49E-05	94.48	64
5	78500552	78521408	20856	ganancia	3205		N	0	3	2.49E-05	94.48	65
5	78521409	78526637	5228	ganancia	3132		N	0	2	0.005115965	66.59	66
5	78521409	78526637	5228	ganancia	3205		N	0	2	0.005115965	66.59	67
5	78526638	78531091	4453	ganancia	3132		N	0	1	0.005115965	39.43	68
5	83490494	83495169	4675	het_loss	3204	EDIL3	N	0	1	0.005115965	39.43	69
5	133372071	133379727	7656	hom_loss	3153		N	0	1	0.005115965	39.43	70
5	137836466	137843309	6843	hom_loss	3279		N	1	1	0.13732578	13.21	71
5	150159466	150161037	1571	het_loss	3117		N	15	4	0.040487703	3.62	72
5	150159466	150161037	1571	het_loss	3180		N	15	4	0.040487703	3.62	73
5	150159466	150161037	1571	het_loss	3199		N	15	4	0.040487703	3.62	74
5	150159466	150161037	1571	het_loss	3278		N	15	4	0.040487703	3.62	75

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
5	150161038	150181399	20361	het_loss	3117		N	14	4	0.033744017	3.88	76
5	150161038	150181399	20361	het_loss	3180		N	14	4	0.033744017	3.88	77
5	150161038	150181399	20361	het_loss	3199		N	14	4	0.033744017	3.88	78
5	150161038	150181399	20361	het_loss	3278		N	14	4	0.033744017	3.88	79
5	150181400	150185189	3789	het_loss	3117		N	13	4	0.027710312	4.18	80
5	150181400	150185189	3789	het_loss	3180		N	13	4	0.027710312	4.18	81
5	150181400	150185189	3789	het_loss	3199		N	13	4	0.027710312	4.18	82
5	150181400	150185189	3789	het_loss	3278		N	13	4	0.027710312	4.18	83
5	150185190	150191626	6436	hom_loss	3009		N	6	7	8.59E-06	16.65	84
5	150185190	150191626	6436	hom_loss	3143		N	6	7	8.59E-06	16.65	85
5	150185190	150191626	6436	hom_loss	3152		N	6	7	8.59E-06	16.65	86
5	150185190	150191626	6436	hom_loss	3154		N	6	7	8.59E-06	16.65	87
5	150185190	150191626	6436	hom_loss	3193		N	6	7	8.59E-06	16.65	88
5	150185190	150191626	6436	hom_loss	3196		N	6	7	8.59E-06	16.65	89
5	150185190	150191626	6436	hom_loss	3281		N	6	7	8.59E-06	16.65	90
5	150191627	150201145	9518	hom_loss	3009		N	6	7	8.59E-06	16.65	91
5	150191627	150201145	9518	hom_loss	3143		N	6	7	8.59E-06	16.65	92
5	150191627	150201145	9518	hom_loss	3152		N	6	7	8.59E-06	16.65	93
5	150191627	150201145	9518	hom_loss	3154		N	6	7	8.59E-06	16.65	94
5	150191627	150201145	9518	hom_loss	3193		N	6	7	8.59E-06	16.65	95
5	150191627	150201145	9518	hom_loss	3196		N	6	7	8.59E-06	16.65	96
5	150191627	150201145	9518	hom_loss	3281		N	6	7	8.59E-06	16.65	97
5	150201146	150202601	1455	hom_loss	3152		N	1	3	0.001318303	40.7	98
5	150201146	150202601	1455	hom_loss	3154		N	1	3	0.001318303	40.7	99
5	150201146	150202601	1455	hom_loss	3193		N	1	3	0.001318303	40.7	100
5	150202602	150204134	1532	het_loss	3132		N	51	8	0.062987683	2.17	101

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
5	150202602	150204134	1532	het_loss	3180		N	51	8	0.062987683	2.17	102
5	150202602	150204134	1532	het_loss	3196		N	51	8	0.062987683	2.17	103
5	150202602	150204134	1532	het_loss	3273		N	51	8	0.062987683	2.17	104
5	150202602	150204134	1532	het_loss	3277		N	51	8	0.062987683	2.17	105
5	150202602	150204134	1532	het_loss	3278		N	51	8	0.062987683	2.17	106
5	150202602	150204134	1532	het_loss	3280		N	51	8	0.062987683	2.17	107
5	150202602	150204134	1532	het_loss	3282		N	51	8	0.062987683	2.17	108
5	179590681	179626660	35979	het_loss	3172	MAPK9	S	0	1	0.005115965	39.43	109
6	2882577	2947403	64826	het_loss	3196	DKFZP686I1521 7, NQO2,SERPIN B6	S	0	1	0.005115965	39.43	110
6	2964646	2966011	1365	het_loss	3193	HTATSF1P2,NQ O2	S	0	1	0.005115965	39.43	111
6	51766024	51773250	7226	het_loss	3167	PKHD1	N	0	1	0.005115965	39.43	112
6	51952217	51953475	1258	ganancia	3127	PKHD1	N	0	1	0.005115965	39.43	113
6	51953476	51965723	12247	ganancia	3127	PKHD1	N	0	2	0.005115965	66.59	114
6	51953476	51965723	12247	ganancia	3205	PKHD1	N	0	2	0.005115965	66.59	115
6	51965724	51969378	3654	ganancia	3127	PKHD1	N	0	1	0.005115965	39.43	116
6	74396294	74398409	2115	het_loss	3009	SLC17A5	N	0	2	0.005115965	66.59	117
6	74396294	74398409	2115	het_loss	3160	SLC17A5	N	0	2	0.005115965	66.59	118
6	74398410	74404837	6427	het_loss	3009	SLC17A5	S	0	1	0.005115965	39.43	119
6	86416979	86431527	14548	het_loss	3197		N	0	1	0.005115965	39.43	120
6	91131823	91135670	3847	het_loss	3171		N	0	1	0.005115965	39.43	121
6	107882367	107890605	8238	het_loss	3201	PDSS2	S	0	1	0.005115965	39.43	122
6	166418511	166422386	3875	het_loss	3125		N	11	10	3.49E-07	13.49	123
6	166418511	166422386	3875	het_loss	3163		N	11	10	3.49E-07	13.49	124
6	166418511	166422386	3875	het_loss	3192		N	11	10	3.49E-07	13.49	125
6	166418511	166422386	3875	het loss	3193		N	11	10	3.49E-07	13.49	126

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
6	166418511	166422386	3875	het loss	3194		N	11	10	3.49E-07	13.49	127
6	166418511	166422386	3875	het loss	3200		N	11	10	3.49E-07	13.49	128
6	166418511	166422386	3875	het loss	3205		N	11	10	3.49E-07	13.49	129
6	166418511	166422386	3875	het loss	3280		N	11	10	3.49E-07	13.49	130
6	166418511	166422386	3875	het loss	3281		N	11	10	3.49E-07	13.49	131
6	166418511	166422386	3875	het loss	3284		N	11	10	3.49E-07	13.49	132
6	166418511	166422386	3875	hom_loss	3009		N	0	3	2.49E-05	94.48	133
6	166418511	166422386	3875	hom_loss	3152		N	0	3	2.49E-05	94.48	134
6	166418511	166422386	3875	hom_loss	3175		N	0	3	2.49E-05	94.48	135
7	65741238	65768682	27444	ganancia	3152	KCTD7	S	0	2	0.005115965	66.59	136
7	65741238	65768682	27444	ganancia	3202	KCTD7	S	0	2	0.005115965	66.59	137
7	157174966	157177843	2877	het loss	3009	PTPRN2	N	0	1	0.005115965	39.43	138
7	157425841	157496238	70397	ganancia	3189	PTPRN2	N	1	1	0.13732578	13.21	139
7	158000082	158007892	7810	het loss	3279	PTPRN2	N	1	1	0.13732578	13.21	140
7	158007893	158010672	2779	het loss	3279	PTPRN2	N	5	1	0.358539546	2.63	141
7	158010673	158024569	13896	het loss	3279	MIR595,PTPRN2	S	1	1	0.13732578	13.21	142
8	23103186	23125443	22257	het loss	3140	TNFRSF10A	S	0	1	0.005115965	39.43	143
8	39914488	39919594	5106	het loss	3126	IDO2	N	0	1	0.005115965	39.43	144
8	79905654	79910286	4632	het loss	3159		N	0	1	0.005115965	39.43	145
8	99790200	99799839	9639	het loss	3006	STK3	N	0	1	0.005115965	39.43	146
8	102049360	102064431	15071	het loss	3173		N	0	3	2.49E-05	94.48	147
8	102049360	102064431	15071	het loss	3175		N	0	3	2.49E-05	94.48	148
8	102049360	102064431	15071	het loss	3282		N	0	3	2.49E-05	94.48	149
9	571398	580721	9323	het loss	3006	KANK1	N	2	1	0.198831257	6.6	150
9	580722	584647	3925	het loss	3006	KANK1	N	3	3	0.005933668	13.54	151
9	580722	584647	3925	het loss	3200	KANK1	N	3	3	0.005933668	13.54	152

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
9	580722	584647	3925	het loss	3282	KANK1	N	3	3	0.005933668	13.54	153
9	584648	598488	13840	het loss	3200	KANK1	N	2	2	0.02731135	13.37	154
9	584648	598488	13840	het loss	3282	KANK1	N	2	2	0.02731135	13.37	155
9	634039	637589	3550	het loss	3273	KANK1	N	0	2	0.005115965	66.59	156
9	634039	637589	3550	het loss	3282	KANK1	N	0	2	0.005115965	66.59	157
9	74050088	74059447	9359	het loss	3165	GDA	S	0	1	0.005115965	39.43	158
9	93140394	93447826	307432	ganancia	3198	AUH,MIR3163, MIR3910-1,MIR3910-2,NFIL3	S	0	1	0.005115965	39.43	159
9	118564159	118576633	11474	ganancia	3193	ASTN2	N	0	1	0.005115965	39.43	160
9	118657526	118664593	7067	het loss	3144	ASTN2	N	0	1	0.005115965	39.43	161
9	119220847	119233078	12231	ganancia	3005		N	0	1	0.005115965	39.43	162
10	899657	1071401	171744	ganancia	3161	GTPBP4,IDI2,IDI 2-AS 1,LARP4B	S	0	1	0.005115965	39.43	163
10	76217585	76411591	194006	ganancia	3179	KAT6B	S	0	1	0.005115965	39.43	164
10	116000069	116004388	4319	ganancia	3010	VWA2	S	0	1	0.005115965	39.43	165
11	14677012	14689025	12013	het loss	3199	PDE3B	N	0	1	0.005115965	39.43	166
11	34608313	34615878	7565	het loss	3117	EHF	S	0	1	0.005115965	39.43	167
11	62382087	62398462	16375	het loss	3205	SLC3A2	S	0	1	0.005115965	39.43	168
11	76631014	76643625	12611	het loss	3193	GDPD4	S	0	1	0.005115965	39.43	169
12	11616557	12114030	497473	het loss	3126	ETV6	S	0	1	0.005115965	39.43	170
12	12438904	12778142	339238	het_loss	3126	APOLD1,CDKN 1B,CREBL2,DUS P16,GPR19,LOH 12CR1	S	0	1	0.005115965	39.43	171
12	12968705	12971310	2605	ganancia	3127		N	0	1	0.005115965	39.43	172
12	91845527	92201342	355815	het_loss	3126	EEA1,LOC643339	S	0	1	0.005115965	39.43	173
12	92215898	92567120	351222	het_loss	3126	LOC643339,MRP L42,NUDT4,NU DT4P1,SOCS2,S OCS2-AS1,UBE2N	S	0	1	0.005115965	39.43	174

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
12	92568362	93307172	738810	het_loss	3126	CCDC41,CRAD D,PLXNC1	S	0	1	0.005115965	39.43	175
12	111061085	111064486	3401	het_loss	3004	TRAFD1	S	0	1	0.005115965	39.43	176
13	40939924	41026908	86984	ganancia	3140	RGCC	S	0	1	0.005115965	39.43	177
13	75006025	75016304	10279	ganancia	3009	COMD6	S	0	2	0.005115965	66.59	178
13	75006025	75016304	10279	ganancia	3152	COMD6	S	0	2	0.005115965	66.59	179
13	91811087	91814369	3282	het_loss	3143	GPC5	N	1	1	0.13732578	13.21	180
13	91811087	91811118	31	hom_loss	3173	GPC5	N	0	1	0.005115965	39.43	181
13	110754499	110778301	23802	ganancia	3006	ARHGEF7,TEX2 9	S	0	1	0.005115965	39.43	182
14	20021118	20055469	34351	ganancia	3205	RNASE10	S	0	1	0.005115965	39.43	183
14	20426824	20481852	55028	hom_loss	3200	ECRP,RNASE3	S	0	1	0.005115965	39.43	184
14	20430810	20458350	27540	het_loss	3192	ECRP	S	3	1	0.256004559	4.39	185
14	20458351	20481852	23501	het_loss	3192		N	4	1	0.309147091	3.29	186
14	20481853	20490129	8276	het_loss	3192		N	1	1	0.13732578	13.21	187
14	21096689	21105611	8922	het_loss	3125		N	0	5	1.16E-07	152.56	188
14	21096689	21105611	8922	het_loss	3175		N	0	5	1.16E-07	152.56	189
14	21096689	21105611	8922	het_loss	3194		N	0	5	1.16E-07	152.56	190
14	21096689	21105611	8922	het_loss	3204		N	0	5	1.16E-07	152.56	191
14	21096689	21105611	8922	het_loss	3273		N	0	5	1.16E-07	152.56	192
14	21120750	21125513	4763	ganancia	3143		N	1	2	0.014314826	26.77	193
14	21120750	21125513	4763	ganancia	3173		N	1	2	0.014314826	26.77	194
14	60901636	60909492	7856	het_loss	3193	PRKCH	N	0	1	0.005115965	39.43	195
14	60912874	60921269	8395	het_loss	3174	PRKCH	N	0	1	0.005115965	39.43	196
14	63937192	63944459	7267	ganancia	3205	MTHFD1	S	0	1	0.005115965	39.43	197
14	95754535	95759056	4521	het_loss	3009	BDKRB2	N	0	3	2.49E-05	94.48	198
14	95754535	95759056	4521	het_loss	3173	BDKRB2	N	0	3	2.49E-05	94.48	199
14	95754535	95759056	4521	het_loss	3202	BDKRB2	N	0	3	2.49E-05	94.48	200

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
15	66065925	66082418	16493	het_loss	3010		N	0	1	0.005115965	39.43	201
15	70432627	70443017	10390	ganancia	3169	HEXA	S	0	1	0.005115965	39.43	202
15	75096101	75101523	5422	ganancia	3200	PSTPIP1	S	0	1	0.005115965	39.43	203
15	75101524	75105788	4264	ganancia	3132	PSTPIP1	S	0	2	0.005115965	66.59	204
15	75101524	75105788	4264	ganancia	3200	PSTPIP1	S	0	2	0.005115965	66.59	205
15	75105789	75109086	3297	ganancia	3127	PSTPIP1	S	0	5	1.16E-07	152.56	206
15	75105789	75109086	3297	ganancia	3132	PSTPIP1	S	0	5	1.16E-07	152.56	207
15	75105789	75109086	3297	ganancia	3199	PSTPIP1	S	0	5	1.16E-07	152.56	208
15	75105789	75109086	3297	ganancia	3200	PSTPIP1	S	0	5	1.16E-07	152.56	209
15	75105789	75109086	3297	ganancia	3279	PSTPIP1	S	0	5	1.16E-07	152.56	210
15	75109087	75115806	6719	ganancia	3127	PSTPIP1	S	1	5	9.14E-06	69.72	211
15	75109087	75115806	6719	ganancia	3132	PSTPIP1	S	1	5	9.14E-06	69.72	212
15	75109087	75115806	6719	ganancia	3199	PSTPIP1	S	1	5	9.14E-06	69.72	213
15	75109087	75115806	6719	ganancia	3200	PSTPIP1	S	1	5	9.14E-06	69.72	214
15	75109087	75115806	6719	ganancia	3279	PSTPIP1	S	1	5	9.14E-06	69.72	215
15	75115807	75117798	1991	ganancia	3200	PSTPIP1	S	1	1	0.13732578	13.21	216
15	88999998	89016848	16850	het_loss	3172		N	0	1	0.005115965	39.43	217
16	6823677	6884976	61299	het_loss	3126	RBFOX1	N	0	1	0.005115965	39.43	218
16	6886815	6896330	9515	het_loss	3126	RBFOX1	N	0	1	0.005115965	39.43	219
16	6942078	6945539	3461	ganancia	3173	RBFOX1	N	1	3	0.001318303	40.7	220
16	6942078	6945539	3461	ganancia	3175	RBFOX1	N	1	3	0.001318303	40.7	221
16	6942078	6945539	3461	ganancia	3282	RBFOX1	N	1	3	0.001318303	40.7	222
16	23844022	23848772	4750	het_loss	3198	PRKCB	N	7	1	0.447101793	1.88	223
16	23892842	23893968	1126	ganancia	3199	PRKCB	N	0	1	0.005115965	39.43	224
16	23893969	23903495	9526	ganancia	3199	PRKCB	N	0	2	0.005115965	66.59	225
16	23893969	23903495	9526	ganancia	3205	PRKCB	N	0	2	0.005115965	66.59	226

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
16	23903496	23908248	4752	ganancia	3205	PRKCB	S	0	1	0.005115965	39.43	227
16	69047888	69050151	2263	ganancia	3174	FUK	N	0	2	0.005115965	66.59	228
16	69047888	69050151	2263	ganancia	3185	FUK	N	0	2	0.005115965	66.59	229
16	69052450	69081640	29190	het_loss	3197	COG4,FUK	S	0	1	0.005115965	39.43	230
16	70653499	70665447	11948	ganancia	3143	HPR	S	0	6	1.16E-07	182.82	231
16	70653499	70665447	11948	ganancia	3152	HPR	S	0	6	1.16E-07	182.82	232
16	70653499	70665447	11948	ganancia	3192	HPR	S	0	6	1.16E-07	182.82	233
16	70653499	70665447	11948	ganancia	3200	HPR	S	0	6	1.16E-07	182.82	234
16	70653499	70665447	11948	ganancia	3282	HPR	S	0	6	1.16E-07	182.82	235
16	70653499	70665447	11948	ganancia	3284	HPR	S	0	6	1.16E-07	182.82	236
17	69341925	70202523	860598	ganancia	3183	BTBD17,C17orf7 7,CD300A,CD30 0C,CD300E,CD3 00LB,CD300LD,CD300LF,D NAI2,GPR142,GPRC5 C,KIF19,MGC16 275,RAB37,RPL3 8. TTSH2	S	1	1	0.13732578	13.21	237
17	75608151	75611602	3451	het_loss	3144	TBC1D16	N	1	7	5.37E-08	100.4	238
17	75608151	75611602	3451	het_loss	3152	TBC1D16	N	1	7	5.37E-08	100.4	239
17	75608151	75611602	3451	het_loss	3163	TBC1D16	N	1	7	5.37E-08	100.4	240
17	75608151	75611602	3451	het_loss	3192	TBC1D16	N	1	7	5.37E-08	100.4	241
17	75608151	75611602	3451	het_loss	3200	TBC1D16	N	1	7	5.37E-08	100.4	242
17	75608151	75611602	3451	het_loss	3204	TBC1D16	N	1	7	5.37E-08	100.4	243
17	75608151	75611602	3451	het_loss	3284	TBC1D16	N	1	7	5.37E-08	100.4	244
17	75608151	75611602	3451	hom_loss	3009	TBC1D16	N	0	1	0.005115965	39.43	245
17	75611602	75615433	3831	hom_loss	3175	TBC1D16	N	0	1	0.005115965	39.43	246
17	75611603	75615433	3830	het_loss	3144	TBC1D16	N	1	4	0.000112689	55.01	247
17	75611603	75615433	3830	het_loss	3152	TBC1D16	N	1	4	0.000112689	55.01	248

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
17	75611603	75615433	3830	het_loss	3163	TBC1D16	N	1	4	0.000112689	55.01	249
17	75611603	75615433	3830	het_loss	3200	TBC1D16	N	1	4	0.000112689	55.01	250
17	76241510	76247304	5794	ganancia	3205	RPTOR	N	0	1	0.005115965	39.43	251
17	76247305	76265683	18378	ganancia	3127	RPTOR	N	0	2	0.005115965	66.59	252
17	76247305	76265683	18378	ganancia	3205	RPTOR	N	0	2	0.005115965	66.59	253
17	76265684	76267844	2160	ganancia	3205	RPTOR	N	0	1	0.005115965	39.43	254
18	9985530	10125331	139801	ganancia	3175		N	0	1	0.005115965	39.43	255
18	12764095	12781985	17890	ganancia	3191	PTPN2	S	0	1	0.005115965	39.43	256
18	27026203	27029351	3148	het_loss	3125		N	0	3	2.49E-05	94.48	257
18	27026203	27029351	3148	het_loss	3143		N	0	3	2.49E-05	94.48	258
18	27026203	27029351	3148	het_loss	3175		N	0	3	2.49E-05	94.48	259
18	42537949	42663605	125656	ganancia	3125	PIAS2,ST8SIA5	S	0	1	0.005115965	39.43	260
18	46917195	46945018	27823	het_loss	3161		N	0	1	0.005115965	39.43	261
18	59457622	59465699	8077	het_loss	3145	SERPINB4	S	0	1	0.005115965	39.43	262
19	3270755	3291144	20389	ganancia	3205		N	0	1	0.005115965	39.43	263
19	46386511	46388364	1853	hom_loss	3175		N	0	1	0.005115965	39.43	264
19	52496536	52501292	4756	ganancia	3124		N	0	1	0.005115965	39.43	265
19	55247874	55250186	2312	het_loss	3163	FLJ26850	N	4	5	0.000161709	17.38	266
19	55247874	55250186	2312	het_loss	3173	FLJ26850	N	4	5	0.000161709	17.38	267
19	55247874	55250186	2312	het_loss	3192	FLJ26850	N	4	5	0.000161709	17.38	268
19	55247874	55250186	2312	het_loss	3200	FLJ26850	N	4	5	0.000161709	17.38	269
19	55247874	55250186	2312	het_loss	3280	FLJ26850	N	4	5	0.000161709	17.38	270
19	55250187	55252420	2233	het_loss	3163	FLJ26850	N	4	6	1.80E-05	21.15	271
19	55250187	55252420	2233	het_loss	3173	FLJ26850	N	4	6	1.80E-05	21.15	272
19	55250187	55252420	2233	het_loss	3192	FLJ26850	N	4	6	1.80E-05	21.15	273
19	55250187	55252420	2233	het_loss	3194	FLJ26850	N	4	6	1.80E-05	21.15	274

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
19	55250187	55252420	2233	het_loss	3200	FLJ26850	N	4	6	1.80E-05	21.15	275
19	55250187	55252420	2233	het_loss	3280	FLJ26850	N	4	6	1.80E-05	21.15	276
19	55250187	55252420	2233	hom_loss	3175	FLJ26850	N	0	2	0.005115965	96.59	277
19	55250187	55252420	2233	hom_loss	3202	FLJ26850	N	0	2	0.005115965	96.59	278
19	56964168	57308449	344281	ganancia	3155	FPR2,FPR3,ZNF 350,ZNF432,ZNF 577,ZNF613,ZNF 614,ZNF615,ZNF 649,ZNF841	S	3	2	0.043434433	8.91	279
19	56964168	57308449	344281	ganancia	3157	FPR2,FPR3,ZNF 350,ZNF432,ZNF 577,ZNF613,ZNF 614,ZNF615,ZNF 649,ZNF841	S	3	2	0.043434433	8.91	280
19	59016855	59023850	6995	het_loss	3117	NLRP12	S	0	1	0.005115965	39.43	281
19	59249279	59250741	1462	hom_loss	3160	VSTM1	N	37	2	1	0.7	282
19	59249279	59250741	1462	hom_loss	3164	VSTM1	N	37	2	1	0.7	283
19	59250742	59251831	1089	hom_loss	3117	VSTM1	N	38	4	0.533638399	1.39	284
19	59250742	59251831	1089	hom_loss	3160	VSTM1	N	38	4	0.533638399	1.39	285
19	59250742	59251831	1089	hom_loss	3164	VSTM1	N	38	4	0.533638399	1.39	286
19	59250742	59251831	1089	hom_loss	3277	VSTM1	N	38	4	0.533638399	1.39	287
20	17844577	17954650	110073	ganancia	3166	MGME1,OVOL2, SNORD17,SNX5	S	0	1	0.005115965	39.43	288
20	42706630	42711434	4754	het_loss	3125	ADA	N	0	1	0.005115965	39.43	289
21	15237071	15312960	75889	ganancia	3009	NRIP1	S	0	1	0.005115965	39.43	290
21	29643302	29647950	4648	het_loss	3202	BACH1	S	0	1	0.005115965	39.43	291
21	44634707	44637543	2836	ganancia	3200	TRPM2	S	1	3	0.001318303	40.7	292
21	44634707	44637543	2836	ganancia	3205	TRPM2	S	1	3	0.001318303	40.7	293
21	44634707	44637543	2836	ganancia	3279	TRPM2	S	1	3	0.001318303	40.7	294

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
21	44637544	44641658	4114	ganancia	3127	TRPM2	S	1	5	9.14E-06	69.72	295
21	44637544	44641658	4114	ganancia	3185	TRPM2	S	1	5	9.14E-06	69.72	296
21	44637544	44641658	4114	ganancia	3200	TRPM2	S	1	5	9.14E-06	69.72	297
21	44637544	44641658	4114	ganancia	3205	TRPM2	S	1	5	9.14E-06	69.72	298
21	44637544	44641658	4114	ganancia	3279	TRPM2	S	1	5	9.14E-06	69.72	299
21	44641659	44643973	2314	ganancia	3127	TRPM2	S	1	4	0.000112689	55.01	300
21	44641659	44643973	2314	ganancia	3185	TRPM2	S	1	4	0.000112689	55.01	301
21	44641659	44643973	2314	ganancia	3200	TRPM2	S	1	4	0.000112689	55.01	302
21	44641659	44643973	2314	ganancia	3279	TRPM2	S	1	4	0.000112689	55.01	303
21	44643974	44657372	13398	het loss	3161	TRPM2	S	1	1	0.13732578	13.21	304
21	44643975	44657372	13397	ganancia	3127	TRPM2	S	0	5	1.16E-07	152.56	305
21	44643975	44657372	13397	ganancia	3185	TRPM2	S	0	5	1.16E-07	152.56	306
21	44643975	44657372	13397	ganancia	3200	TRPM2	S	0	5	1.16E-07	152.56	307
21	44643975	44657372	13397	ganancia	3205	TRPM2	S	0	5	1.16E-07	152.56	308
21	44643975	44657372	13397	ganancia	3279	TRPM2	S	0	5	1.16E-07	152.56	309
21	44657373	44660198	2825	ganancia	3127	TRPM2	S	0	3	2.49E-05	94.48	310
21	44657373	44660198	2825	ganancia	3200	TRPM2	S	0	3	2.49E-05	94.48	311
21	44657373	44660198	2825	ganancia	3279	TRPM2	S	0	3	2.49E-05	94.48	312
21	44660199	44666832	6633	ganancia	3127	TRPM2	S	0	4	2.49E-05	123.12	313
21	44660199	44666832	6633	ganancia	3200	TRPM2	S	0	4	2.49E-05	123.12	314
21	44660199	44666832	6633	ganancia	3205	TRPM2	S	0	4	2.49E-05	123.12	315
21	44660199	44666832	6633	ganancia	3279	TRPM2	S	0	4	2.49E-05	123.12	316
21	44666833	44669596	2763	ganancia	3127	TRPM2	S	0	3	2.49E-05	94.48	317
21	44666833	44669596	2763	ganancia	3205	TRPM2	S	0	3	2.49E-05	94.48	318
21	44666833	44669596	2763	ganancia	3279	TRPM2	S	0	3	2.49E-05	94.48	319
21	44669597	44671482	1885	ganancia	3205	TRPM2	S	0	2	0.005115965	66.59	320

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
21	44669597	44671482	1885	ganancia	3279	TRPM2	S	0	2	0.005115965	66.59	321
21	44671483	44681194	9711	ganancia	3205	TRPM2	S	0	1	0.005115965	39.43	322
21	45348895	45354820	5925	het loss	3179	ADARB1	N	0	1	0.005115965	39.43	323
22	37689058	37715385	26327	ganancia	3169	APOBEC3A,AP OBEC3A_B,APO BEC3B	S	0	1	0.005115965	39.43	324
22	39257585	39261621	4036	het loss	3005	MKL1	N	0	1	0.005115965	39.43	325
22	40642402	40655210	12808	ganancia	3205	TNFRSF13C	S	0	1	0.005115965	39.43	326
22	40655820	40659632	3812	ganancia	3185		N	0	2	0.005115965	66.59	327
22	40655820	40659632	3812	ganancia	3205		N	0	2	0.005115965	66.59	328
22	40659633	40663049	3416	ganancia	3127		N	0	3	2.49E-05	94.48	329
22	40659633	40663049	3416	ganancia	3185		N	0	3	2.49E-05	94.48	330
22	40659633	40663049	3416	ganancia	3205		N	0	3	2.49E-05	94.48	331
22	40663050	40668079	5029	ganancia	3127	CENPM	S	0	5	1.16E-07	152.56	332
22	40663050	40668079	5029	ganancia	3185	CENPM	S	0	5	1.16E-07	152.56	333
22	40663050	40668079	5029	ganancia	3190	CENPM	S	0	5	1.16E-07	152.56	334
22	40663050	40668079	5029	ganancia	3202	CENPM	S	0	5	1.16E-07	152.56	335
22	40663050	40668079	5029	ganancia	3205	CENPM	S	0	5	1.16E-07	152.56	336
22	40668080	40671866	3786	ganancia	3127	CENPM	S	0	3	2.49E-05	94.48	337
22	40668080	40671866	3786	ganancia	3185	CENPM	S	0	3	2.49E-05	94.48	338
22	40668080	40671866	3786	ganancia	3205	CENPM	S	0	3	2.49E-05	94.48	339
22	40671867	40673250	1383	ganancia	3185	CENPM	S	0	2	0.005115965	66.59	340
22	40671867	40673250	1383	ganancia	3205	CENPM	S	0	2	0.005115965	66.59	341
22	40673251	40675788	2537	ganancia	3205		N	0	1	0.005115965	39.43	342
23	232907	234429	1522	het_loss	3007	PPP2R3B	N	0	1	0.005115965	39.43	343
23	7585301	7769322	184021	ganancia	3172		N	5	1	0.358539546	2.63	344
23	7769323	7773949	4626	ganancia	3172	VCX	S	7	1	0.447101793	1.88	345

Chr	Inicio subregión de CNV	Parada de subregión de CNV	Tamaño de la subregión de CNV	Tipo de CNV	ID de caso de la PML	Símbolo del gen de RefSeq	Solapamiento de exones	Casos de NVE	Casos de PML	FET	OR	No de subregión de CNV (SRN)
23	7773982	7779354	5372	het_loss	3132		N	0	3	2.49E-05	94.48	346
23	7773982	7779354	5372	het_loss	3171		N	0	3	2.49E-05	94.48	347
23	7773982	7779354	5372	het_loss	3204		N	0	3	2.49E-05	94.48	348
23	7773982	7779353	5371	ganancia	3172		N	5	1	0.358539546	2.63	349
23	7779354	7815400	36046	ganancia	3172		N	6	1	0.404443314	2.19	350
23	7779355	8093113	313758	het_loss	3171	MIR651,PNPLA4	S	0	1	0.005115965	39.43	351
23	7815401	7830994	15593	ganancia	3172	PNPLA4	S	7	1	0.447101793	1.88	352
23	48358646	48408854	50208	het_loss	3009		N	0	1	0.005115965	39.43	353
23	64710574	64725828	15254	ganancia	3125		N	0	1	0.005115965	39.43	354
23	73083877	73086192	2315	hom_loss	3193	JPX	N	1	2	0.014314826	26.77	355
23	73083877	73086192	2315	hom_loss	3200	JPX	N	1	2	0.014314826	26.77	356
23	122337025	122340879	3854	hom_loss	3125	GRIA3	N	0	1	0.005115965	39.43	357
23	148452844	148461889	9045	het_loss	3163		N	7	2	0.129983268	3.8	358
23	148452844	148461889	9045	het_loss	3205		N	7	2	0.129983268	3.8	359
23	148459108	148461889	2781	hom_loss	3144		N	0	2	0.005115965	66.59	360
23	148459108	148461889	2781	hom_loss	3193		N	0	2	0.005115965	66.59	361
23	149901706	149902701	995	ganancia	3117	HMGB3	S	0	2	0.005115965	66.59	362
23	149901706	149902701	995	ganancia	3118	HMGB3	S	0	2	0.005115965	66.59	363

La Tabla 2 es idéntica a la Tabla 1, con algunas excepciones. En primer lugar, las coordenadas de las CNV enumeradas se refieren a las subregiones de CNV reales que resultaron ser únicas o significativamente diferentes entre las cohortes de enfermedad y normales, a diferencia de la Tabla 1, que enumera las CNV originales. En segundo lugar, una columna adicional detalla si las subregiones génicas de CNV de interés se solapan con un exón o no. En tercer y cuarto lugar, 2 columnas adicionales detallan el número de casos normales y el número de casos de enfermedad que albergan la subregión de CNV relevante. Por último, 2 columnas informan de la prueba exacta de Fisher de 2 colas (FET) y de la razón de probabilidades (OR). En la Tabla 2 se utiliza la numeración cromosómica estándar utilizada por los expertos en la materia para los cromosomas (1-22) autosómicos, pero, por conveniencia con los métodos de análisis, el cromosoma X se designa como el cromosoma 23 en el presente documento. Todas las coordenadas están en hg18.

Tabla 3: Una lista no redundante de los genes enumerados en la Tabla 2

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
ADARB1	intrónico	104	isoforma 1 editasa 1 específica de ARN de doble cadena	Este gen codifica la enzima responsable de la edición del pre-ARNm de la subunidad B del receptor de glutamato por desaminación específica del sitio de adenosinas. Estudios realizados en ratas descubrieron que esta enzima actuaba sobre sus propias moléculas de pre-ARNm para convertir un dinucleótido AA en un dinucleótido AI, lo que daba lugar a un nuevo sitio de empalme. El empalme alternativo de este gen da lugar a varias variantes del transcripción, algunas de las cuales se han caracterizado por la presencia o ausencia de	2
				una inserción de casete ALU y una región del terminal C corta o larga, [proporcionado por RefSeq, Jul 2008]. Variante de transcripción: Esta variante (1), también conocida como ADAR2a-L1 o DRADA2a, carece de la inserción del casete ALU y contiene la región del terminal C larga, en comparación con la variante 2. La isoforma resultante (1), también conocida como hRED1-Corta, carece de un segmento interno, en comparación con la isoforma 2. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Atributos de RefSeq## sufre edición del ARN :: PMID: 11717408, 12045112 ##FIN de los Atributos de RefSeq## ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AB194370.1, U76420.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025085	

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				[ECO:0000348] ##FIN de los Datos de Evidencia##	
AGBL4	intrónico	84871	carboxipeptidasa 6 citosólica	N/A	3
APOBEC3A	exónico	200315	ADN dC->dU-enzima de edición APOBEC-3A isoforma a	Este gen es un miembro de la familia de genes de la citidina desaminasa. Es uno de los siete genes o pseudogenes relacionados que se encuentran en un grupo, que se cree que es el resultado de la duplicación de genes, en el cromosoma 22. Los miembros del grupo codifican proteínas estructural y funcionalmente relacionadas con la citidina desaminasa editora de ARN de C a U APOBEC1. La proteína codificada por este gen carece de la actividad de unión al zinc de otros miembros de la familia. La proteína desempeña un papel en la inmunidad, al restringir la transmisión de ADN extraño, tal como los virus. Un mecanismo de restricción del ADN extraño es la desaminación de las citidinas hasta las uridinas del ADN de doble cadena extraño, lo que conduce a la degradación del ADN. Sin embargo, se cree que también intervienen otros mecanismos, ya que el efecto antivírico no depende de la actividad de la desaminasa. Para este gen se han encontrado dos variantes de transcripción que codifican isoformas diferentes. [proporcionado por RefSeq, julio de 2012]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (a). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: U03891.2, BC126416.1 [ECO:0000332] Introns de la secuencia de ARN :: una sola muestra respalda todos los introns ERS025081, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	4
APOBEC3AB	intrónico	100913187	probable enzima editora del ADN dC->dU APOBEC-3A	Este gen es un miembro de la familia de genes de la citidina desaminasa. Es uno de los siete genes o pseudogenes relacionados	5

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				que se encuentran en un grupo, que se cree que es el resultado de la duplicación de genes, en el cromosoma 22. Los miembros del grupo codifican proteínas estructural y funcionalmente relacionadas con la citidina desaminasa editora de ARN de C a U APOBEC1. La proteína codificada por este gen carece de la actividad de unión al zinc de otros miembros de la familia. La proteína desempeña un papel en la inmunidad, al restringir la transmisión de ADN extraño, tal como los virus. Un mecanismo de restricción del ADN extraño es la desaminación de las citidinas del ADN de doble cadena extraño a uridinas, lo que conduce a la degradación del ADN. Sin embargo, se cree que también intervienen otros mecanismos, ya que el efecto antivírico no depende de la actividad de la desaminasa. La proteína codificada por este gen es la misma que la codificada por APOBEC3A; sin embargo, este gen es un gen híbrido que resulta de la eliminación de aproximadamente 29.5 kb de secuencia entre el gen APOBEC3A y el gen adyacente APOBEC3B. Los puntos de rotura de la eliminación están dentro de los dos genes, por lo que se predice que el híbrido de eliminación tiene el promotor y la región codificante de APOBEC3A, pero la UTR 3' de APOBEC3B. [proporcionado por RefSeq, julio de 2012]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
APOBEC3B	exónico	9582	ADN dC->dU-enzima de edición APOBEC-3A isoforma a	Este gen es un miembro de la familia de genes de la citidina desaminasa. Es uno de los siete genes o pseudogenes relacionados que se encuentran en un grupo, que se cree que es el resultado de la duplicación de genes, en el cromosoma 22. Los miembros del grupo codifican proteínas	6

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				estructural y funcionalmente relacionadas con la citidina desaminasa editora de ARN de C a U APOBEC1. Se cree que las proteínas pueden ser enzimas de edición del ARN y tener funciones en el crecimiento o el control del ciclo celular. Un gen híbrido resulta de la eliminación de aproximadamente 29.5 kb de secuencia entre este gen, APOBEC3B, y el gen adyacente APOBEC3A. Los puntos de rotura de la eliminación están dentro de los dos genes, por lo que se predice que el alelo de eliminación tiene el promotor y la región codificante de APOBEC3A, pero la UTR 3' de APOBEC3B. Para este gen se han encontrado dos variantes de transcripción que codifican isoformas diferentes. [proporcionado por RefSeq, julio de 2012]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (a). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a publicaciones adicionales.	
				##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AY743217.1 [ECO:0000332] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
APOLD1	exónico	81575	proteína 1 con dominio de apolipoproteína L isoforma 1	APOLD1 es una proteína de respuesta temprana de las células endoteliales que puede desempeñar un papel en la regulación de la señalización de las células endoteliales y la función vascular (Regard et al., 2004 [PubMed 15102925]). [suministrado por OMIM, diciembre de 2008]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (1). Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. La extensión	7

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				de esta transcripción está respaldada por las alineaciones de transcripción. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC042478.1, DR000985.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025086 [ECO:0000348] ##FIN de los Datos de Evidencia##	
ARHGEF7	exónico	8874	factor 7 de intercambio de nucleótidos de guanina rho isoforma a	Las Rho GTPasas desempeñan un papel fundamental en numerosos procesos celulares desencadenados por estímulos extracelulares que actúan a través de receptores acoplados a proteínas G. La proteína codificada pertenece a una familia de proteínas citoplasmáticas que activan la familia Ras de proteínas Rho mediante el intercambio de GDP por GTP. Forma un complejo con la pequeña proteína de unión a GTP Rac1 y recluta a Rac1 a los remolinos de membrana y a las adhesiones focales. Esta proteína puede inducir la alteración de las membranas. Se han descrito múltiples variantes de transcripción empalmadas alternativamente que codifican diferentes isoformas para este gen. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) difiere en la UTR 5', UTR 3', región codificante, y utiliza un codón de inicio secuencia abajo, en comparación con la variante 3. Tanto la variante 1 como la 5 codifican la isoforma a, que tiene un terminal N más corto y un terminal C más largo y distinto, en comparación con la isoforma c. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: D63476.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	8

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
ASTN2	intrónico	23245	precursor de la isoforma a de astrotactina 2	Este gen codifica una proteína que se expresa en el cerebro y puede funcionar en la migración neuronal, basándose en estudios funcionales del gen relacionado astrotactina 1 en humanos y ratones. Una eliminación en este locus se ha asociado a la esquizofrenia. Para este locus se han encontrado múltiples variantes de transcripción que codifican diferentes proteínas. [proporcionado por RefSeq, mayo de 2010]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (a). ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC146756.1, AB014534.1 [ECO:0000332] Introns de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	9
AUH	exónico	549	metilglutaconil-CoA hidratasa, precursor mitocondrial	La proteína mitocondrial metilglutaconil-CoA hidratasa se une al elemento rico en AU (ARE), un elemento común que se encuentra en la UTR 3' de ARNm de rápida descomposición como c-fos, c-myc y el factor estimulante de colonias de granulocitos/macrófagos. Los elementos ARE intervienen en la dirección del ARN hacia la degradación rápida y la desadenilación. La AUH también es homóloga a la enol-CoA hidratasa, una enzima implicada en la degradación de los ácidos grasos, y se ha demostrado que tiene una actividad enzimática de hidratasa intrínseca. Así pues, la AUH es una quimera bifuncional entre la unión al ARN y la actividad enzimática metabólica. Se ha demostrado una posible localización subcelular en las mitocondrias para el homólogo de ratón de esta proteína, que comparte un 92 % de identidad con la proteína humana. Se ha sugerido que AUH puede tener un nuevo papel como proteína de unión a AU localizada en la mitocondria. La AUH humana se expresa como una única especie de ARNm de 1.8 kb, y se traduce como una proteína precursora de 40 kDa que posteriormente se procesa a una forma madura de 32 kDa.	10

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				[proporcionado por RefSeq, mayo de 2010]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Atributos de RefSeq## producto(s) génico(s) localizado(s) en la mitocondria :: reportado por MitoCarta ##FIN de los Atributos de RefSeq## ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: X79888.1, AL533438.3 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084. ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	
BACH1	exónico	571	Dominio BTB y CNC homólogo 1	Este gen codifica un factor de transcripción que pertenece al tipo gorra y cuello de la familia de factores de cremallera de leucina de región básica (CNC-bZip). La proteína codificada contiene amplios dominios complejos, tramtrack, bric-a-brac/poxvirus y dedos de zinc (BTB/POZ), lo que es atípico de los miembros de la familia CNC-bZip. Estos dominios BTB/POZ facilitan las interacciones proteína-proteína y la formación de homooligómeros y/o heterooligómeros. Cuando esta proteína codificada forma un heterodímero con MafK, funciona como represor del elemento de reconocimiento de Maf (MARE) y se reprime la transcripción. Se han identificado múltiples variantes de transcripciones empalmadas alternativamente para este gen. [proporcionado por RefSeq, mayo de 2009].	11
				Variante de transcripción: Esta variante (3), también denominada BACH1t, difiere en la UTR 5', la región codificante 3' y la UTR 3' (en comparación con la variante 1). Esta variante se representa como no codificante porque el uso del codón de inicio traslacional 5' más apoyado, tal como se utiliza en la variante 1, hace que el transcripción a sea candidato a la descomposición del ARNm mediada no sentido (NMD). Este transcripción representa la variante	

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				de empalme reportada por Kanezaki et al. (PMID: 11069897). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025084, ERS025088 [ECO:0000350] ##FIN de los Datos de Evidencia##	
BDKRB2	intrónico	624	Receptor B2 de la bradiquinina	Este gen codifica un receptor para la bradiquinina. El péptido bradiquinina de 9 aa provoca muchas respuestas, tal como vasodilatación, edema, espasmo del músculo liso y estimulación de las fibras del dolor. Este receptor se asocia con proteínas G que estimulan un sistema de segundo mensajero fosfatidilinositol-calcio. Los codones de inicio alternos dan lugar a dos isoformas de la proteína. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: DC369062.1, DC417219.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025090 [ECO:0000348] ##FIN de los Datos de Evidencia##	12
BMPR2	intrónico	659	precursor del receptor de la proteína morfogenética ósea de tipo 2	Este gen codifica un miembro de la familia de receptores de proteínas morfogenéticas óseas (BMP) de serina/treonina quinasas transmembrana. Los ligandos de este receptor son las BMP, miembros de la superfamilia TGF-beta. Las BMP intervienen en la formación del hueso endocondral y en la embriogénesis. Estas proteínas transducen sus señales mediante la formación de complejos heteroméricos de dos tipos diferentes de receptores de serina (treonina) quinasa: receptores de tipo I de aproximadamente 50-55 kD y receptores de tipo II de aproximadamente 70-80 kD. Los	13

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				receptores de tipo II unen ligandos en ausencia de receptores de tipo I, pero requieren sus respectivos receptores de tipo I para la señalización, mientras que los receptores de tipo I requieren sus respectivos receptores de tipo II para la unión del ligando. Las mutaciones en este gen se han asociado con la hipertensión pulmonar primaria, tanto familiar como asociada a la fenfluramina, y con la enfermedad venooclusiva pulmonar. [proporcionado por RefSeq, julio de 2008]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. La extensión de esta transcripción está respaldada por las alineaciones de transcripción. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC052985.2, AK292430.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia###	
BTBD17	exónico	388419	Precursor de la proteína 17 que contiene dominios BTB/POZ	N/A	14
C17orf77	exónico	146723	Precursor de la proteína C17orf77 no caracterizada	N/A	15
CAPZB	intrónico	832	Subunidad beta de la proteína protectora de la actina F isoforma 1	Este gen codifica la subunidad beta de la proteína de unión a la actina de extremo barbado, que pertenece a la familia de proteínas protectora de la actina F. La proteína protectora es una proteína protectora de actina heterodimérica que bloquea el ensamblaje y el desensamblaje de los filamentos de actina en los extremos de los filamentos de crecimiento rápido (barbado) y que interviene en la regulación de la dinámica de los filamentos de actina, así como en la estabilización de la longitud de los filamentos de	16

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				actina en células musculares y no musculares. Un pseudogén de este gen se localiza en el brazo largo del cromosoma 2. Se han encontrado múltiples variantes de transcripción empalmadas alternativamente que codifican diferentes isoformas. [proporcionado por RefSeq, agosto de 2013]. Variante de transcripción: Esta variante (1) codifica la isoforma 1. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC107752.1, BM451686.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	
CCDC41	exónico	51134	proteína centrosomal de 83 kDa	N/A	17
CD300A	exónico	11314	Precursor de la isoforma 1 de la molécula 8 similar a CMRF35	Este gen codifica un miembro de la familia de glicoproteínas CD300, de proteínas de la superficie celular encontradas en los leucocitos que intervienen en las vías de señalización de la respuesta inmunitaria. Este gen está localizado en el cromosoma 17 en un grupo con todos menos uno de los otros miembros de la familia. Se han encontrado múltiples variantes de transcripción que codifican diferentes isoformas para este gen. [proporcionado por RefSeq, febrero de 2012]. Variante de transcripción:	18
				Esta variante (1) representa la transcripción más larga y codifica la proteína más larga (isoforma 1), también denominada IRC1a. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC032352.1, AL531420.3 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025083	

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				[ECO:0000348] ##FIN de los Datos de Evidencia##	
CD300C	exónico	10871	Precursor de la molécula 6 similar a CMRF35	El antígeno CMRF35, que se identificó por reactividad con un anticuerpo monoclonal, está presente en monocitos, neutrófilos y algunos linfocitos T y B (Jackson et al., 1992 [PubMed 1349532]). [suministrado por OMIM, marzo de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC022279.1, BM922826.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025087 [ECO:0000348] ##FIN de los Datos de Evidencia##	19
CD300E	exónico	342510	Precursor de la molécula 2 similar a CMRF35	Este gen codifica un miembro de la familia de glicoproteínas CD300 de proteínas de superficie celular expresadas en células mieloides. La proteína interactúa con la proteína de unión a tirosina quinasa TYRO y se cree que actúa como receptor activador. [proporcionado por RefSeq, nov de 2012]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. Un AUG en el marco se encuentra 41 codones secuencia arriba del sitio de inicio de la traducción anotado, pero no está siendo anotado como un sitio de inicio, ya que no se conserva y se encuentra en un contexto de secuencia Kozak débil. ##INICIO de los Atributos de RefSeq## CDS utiliza AUG secuencia abajo en el marco :: AUG secuencia abajo está asociado con la señal de localización del terminal N ##FIN de los Atributos de RefSeq## ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK303545.1, BX648376.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	20

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
CD300LB	exónico	124599	Precursor de la molécula 7 similar a CMRF35	CD300LB es un receptor activador no clásico de la superfamilia de las inmunoglobulinas (Ig) que se expresa en las células mieloides (Martinez-Barriocanal y Sayos, 2006 [PubMed 16920917]). [suministrado por OMIM, marzo de 2008]. Nota del CCDS: La región de codificación se ha actualizado para acortar el extremo terminal N a uno que esté más respaldado por los datos de conservación disponibles y los miembros de familias parálogas. La actualización tiene un péptido señal del terminal N predicho, que es consistente con el apoyo funcional para la proteína (p. ej., PMID 16920917, 19359216). ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC028091.1, AY359025.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia## ##INICIO de los Atributos de RefSeq## CDS utiliza AUG secuencia abajo en el marco :: AUG secuencia abajo está asociado con la señal de localización del terminal N ##FIN de los Atributos de RefSeq##	21
CD300LD	exónico	100131439	Precursor de la molécula 4 similar a CMRF35	N/A	22
CD300LF	exónico	146722	Precursor de la molécula 1 similar a CMRF35	CD300LF es un receptor inhibidor de la superfamilia Ig expresado en células mieloides. Media señales reguladoras negativas reclutando SHP1 (PTPN6; MIM 176883) o SHIP (INPP5D; MIM 601582) (Sui et al., 2004 [PubMed 15184070]; Alvarez-Errico et al., 2004 [PubMed 15549731]). [suministrado por OMIM, marzo de 2008]. Nota de secuencia: la transcripción y la proteína RefSeq se derivaron de la secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineamientos. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a	23

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF251706.1, AY358545.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
CDKN1B	exónico	1027	inhibidor de la quinasa dependiente de ciclina 1B	Este gen codifica un inhibidor de la quinasa dependiente de ciclina, que comparte una similitud limitada con el inhibidor de CDK CDKN1A/p21. La proteína codificada se une a los complejos ciclina E-CDK2 o ciclina D-CDK4 e impide su activación, controlando así la progresión del ciclo celular en G1. La degradación de esta proteína, desencadenada por su fosforilación dependiente de CDK y su posterior ubiquitinación por complejos SCF, es necesaria para la transición celular de la quiescencia al estado proliferativo. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC001971.1, AY004255.1 [ECO:0000332] Intrones de la secuencia de ARN :: muestra única respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	24
CENPM	exónico	79019	proteína del centrómero M isoforma a	El centrómero es un dominio especializado de la cromatina, presente durante todo el ciclo celular, que actúa como plataforma sobre la que se produce el ensamblaje transitorio del cinetocoro durante la mitosis. Todos los centrómeros activos se caracterizan por la presencia de largas matrices de nucleosomas en las que CENPA (MIM 117139) sustituye a la histona H3 (véase MIM 601128). CENPM es un factor adicional necesario para el ensamblaje del centrómero (Foltz et al., 2006 [PubMed 16622419]). [suministrado por OMIM, marzo de 2008]. Variante de transcripción: Esta variante (1) representa la	25

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				transcripción más larga y codifica la isoforma más larga (a). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC000705.2, BC007495.2 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025085, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	
COG4	exónico	25839	subunidad 4 del complejo de Golgi oligomérico conservado isoforma 1	La proteína codificada por este gen es un componente de un complejo proteico oligomérico implicado en la estructura y función del aparato de Golgi. Los defectos en este gen pueden ser una causa del trastorno congénito de la glicosilación tipo IIj. Para este gen se han encontrado dos variantes de transcripción que codifican isoformas diferentes. [proporcionado por RefSeq, agosto de 2010]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (1). ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC072438.1, AK022874.1 [ECO:0000332] Intrones de la secuencia de ARN :: muestra mixta/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	26
COMMD6	exónico	170622	Proteína 6 que contiene el dominio COMM isoforma a	COMMD6 pertenece a una familia de proteínas inhibidoras de NF-kappa-B (véase RELA; MIM 164014) caracterizadas por la presencia de un dominio COMM (véase COMMD1; MIM 607238) (de Bie et al., 2006 [PubMed 16573520]). [suministrado por OMIM, marzo de 2009]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: HY028175.1, DW440523.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	27

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
CRADD	exónico	8738	proteína que contiene el dominio de muerte CRADD	La proteína codificada por este gen es una proteína que contiene un dominio de muerte (CARD/DD) y se ha demostrado que induce la apoptosis celular. A través de su dominio CARD, esta proteína interactúa con la caspasa 2/ICH1 y la recluta al complejo de transducción de señales de muerte celular que incluye el receptor 1 del factor de necrosis tumoral (TNFR1A), la quinasa RIPK1/RIP y otras proteínas que contienen el dominio CARD. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BX480215.1, BC017042.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025083 [ECO:0000348] ##FIN de los Datos de Evidencia##	28
CREBL2	exónico	1389	proteína similar a la proteína 2 de unión a elementos sensibles al AMPc	La proteína tipo 2 de unión al elemento sensible a cAMP (CREBL2) se identificó en una búsqueda para encontrar genes en una región comúnmente suprimida en el cromosoma 12p13 flanqueada por los genes ETV6 y CDKN1B, frecuentemente asociados a neoplasias hematopoyéticas, así como a cánceres de mama, pulmón de células no pequeñas y ovario. CREBL2 comparte un 41 % de identidad con la proteína de unión a CRE (CREB) en una región de 48 bases de longitud que codifica el dominio bZip de CREB. El dominio bZip consta de unos 30 aminoácidos ricos en residuos básicos implicados en la unión al ADN, seguidos de un motivo de cremallera de leucina implicado en la dimerización de la proteína. Esto sugiere que CREBL2 codifica una proteína con capacidad de unión al ADN. La presencia de eliminación de CREBL2 en neoplasias sugiere que CREBL2 puede actuar como gen supresor de tumores. [proporcionado por RefSeq, julio de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón	29

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				de transcripción :: BC106052.1, AF039081.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
DNAI2	exónico	64446	cadena 2 intermedia de dineína, isoforma axonemal 1	La proteína codificada por este gen es un miembro de la familia de las cadenas intermedias de dineína y forma parte del complejo de dineína de los cilios respiratorios y los flagelos de los espermatozoides. Las mutaciones en este gen están asociadas a la discinesia ciliar primaria de tipo 9. Para este gen se han observado variantes de transcripción empalmadas alternativamente que codifican diferentes isoformas. [proporcionado por RefSeq, marzo de 2010]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga (1). ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF250288.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra soporta todos los intrones ERS025085 [ECO:0000348] ##FIN de los Datos de Evidencia## ##INICIO de los Atributos de RefSeq## NMD candidato :: traducción inferida de la conservación ##FIN de los Atributos de RefSeq##	30
DNER	intrónico	92737	precursor del receptor relacionado con el factor de crecimiento epidérmico tipo delta y Notch	N/A	31
DUSP16	exónico	80824	proteína fosfatasa 16 de doble especificidad	Este gen codifica una fosfatasa de proteína quinasa activada por mitógenos que es miembro de la subfamilia de proteínas fosfatasas de doble especificidad. Estas fosfatasas inactivan sus quinasas diana desfosforilando tanto los residuos de fosfoserina/treonina como los de fosfotirosina. La proteína codificada regula específicamente las vías de la quinasa del terminal amino de c-Jun (JNK) y de la quinasa regulada por señales extracelulares (ERK). [proporcionado por RefSeq, mayo de 2010]. Nota de secuencia: Este registro de RefSeq se creó a partir	32

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF506796.1, AB052156.1 [ECO:0000332] Introns de la secuencia de ARN :: una sola muestra respalda todos los introns ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	
ECRP	exónico	643332	N/A	N/A	33
EDIL3	intrónico	10085	Precursor de la isoforma 1 de la proteína 3 que contiene un dominio similar a la repetición similar a EGF y discoidina I	La proteína codificada por este gen es un ligando de integrina. Desempeña un papel importante en la mediación de la angiogénesis y puede ser importante en la remodelación y el desarrollo de la pared vascular. También influye en el comportamiento de las células endoteliales. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga (1). Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC030828.1, U70312.1 [ECO:0000332] Introns de la secuencia de ARN :: muestra mixta/parcial ERS025081,	34

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
EEA1	exónico	8411	antígeno endosómico temprano 1	N/A	35
EHF	ambos	26298	Precursor de la isoforma 1 del factor homólogo ETS	Este gen codifica una proteína que pertenece a una subfamilia de factores de transcripción ETS caracterizada por una expresión específica epitelial (ESE). La proteína codificada actúa como represor transcripcional y puede estar implicada en la diferenciación epitelial y la carcinogénesis. Para este gen se han encontrado tres variantes de transcripción que codifican diferentes isoformas. [proporcionado por RefSeq, junio de 2011]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga (1). Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK310867.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025086 [ECO:0000348] ##FIN de los Datos de Evidencia##	36
EMB	exónico	133418	precursor de la embigina	Este gen codifica una glicoproteína transmembrana que pertenece a la superfamilia de las inmunoglobulinas. La proteína codificada puede estar implicada en el crecimiento y desarrollo celular al mediar en las interacciones entre la célula y la matriz extracelular. Un pseudogén de este gen se encuentra en el cromosoma 1. [proporcionado por RefSeq, enero de 2009]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC059398.1,	37

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				AK300860.1 [ECO:0000332] Intrones de la secuencia de ARN :: muestra mixta/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
ETV6	exónico	2120	factor de transcripción ETV6	Este gen codifica un factor de transcripción de la familia ETS. El producto de este gen contiene dos dominios funcionales: un dominio del terminal N apuntado (PNT) que participa en las interacciones proteína-proteína consigo mismo y con otras proteínas, y un dominio del terminal C de unión al ADN. Estudios de inactivación génica en ratones sugieren que es necesaria para la hematopoyesis y el mantenimiento de la red vascular en desarrollo. Se sabe que este gen está implicado en un gran número de reordenamientos cromosómicos asociados a la leucemia y al fibrosarcoma congénito. [proporcionado por RefSeq, septiembre de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia##	38
				Combinación de exones de la transcripción :: BC043399.1, U11732.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
FHL2	exónico	2274	proteína 2 de cuatro y medio dominios de LIM	Este gen codifica un miembro de la familia de proteínas de cuatro y medio LIM. Los miembros de la familia contienen dos dominios de dedos de zinc altamente conservados, dispuestos en tándem, con cuatro cisteínas altamente conservadas que unen un átomo de zinc en cada dedo de zinc. Se cree que esta proteína desempeña un papel en el ensamblaje de las membranas extracelulares. Además, este gen se subregula durante la transformación de mioblastos normales en células de rabdomiosarcoma y la proteína codificada puede funcionar como enlace entre la presenilina-2 y una vía de señalización intracelular. Se han identificado múltiples variantes	39

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				empalmadas alternativamente, que codifican la misma proteína. [proporcionado por RefSeq, agosto de 2011]. Variante de transcripción: Esta variante (1) difiere en la UTR 5' en comparación con la variante 2. Las variantes 1, 2, 4 y 5 codifican la misma isoforma. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Atributos de RefSeq## CDS utiliza AUG en marco secuencia abajo :: falta de evidencia para el uso de AUG secuencia arriba ##FIN de los Atributos de RefSeq## ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC093049.1, AL523628.3 [ECO:0000332] Introns de la secuencia de ARN :: una sola muestra respalda todos los introns ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	
FLJ26850	intrónico	400710	N/A	N/A	40
FPR2	exónico	2358	Receptor 2 del péptido N-formilo	N/A	41
FPR3	exónico	2359	Receptor 3 del péptido N-formilo	N/A	42
FUK	ambos	197258	L-fucosa quinasa	La proteína codificada por este gen es un miembro de la familia de las quinasas GHMP (galacto-, homoserina, mevalonato y fosfomevalonato) y cataliza la fosforilación de la L-fucosa para formar beta-L-fucosa 1-fosfato. Esta enzima cataliza el primer paso en la utilización de la L-fucosa libre en la síntesis de glicoproteínas y glicolípidos. La L-fucosa puede ser importante en la mediación de una serie de interacciones célula-célula, tal como el reconocimiento de antígenos del grupo sanguíneo, la inflamación y la metástasis. Aunque pueden existir varias variantes de transcripción para este gen, hasta la fecha sólo se ha descrito la naturaleza completa de una de ellas. [proporcionado por RefSeq, julio de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AJ441184.1, BC032542.1 [ECO:0000332] Introns de la	43

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				secuencia de ARN :: muestra mixta/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
GDA	exónico	9615	guanina desaminasa isoforma a	Este gen codifica una enzima responsable de la desaminación hidrolítica de la guanina. Estudios en el ortólogo de rata sugieren que este gen desempeña un papel en el ensamblaje de microtúbulos. Se han encontrado múltiples variantes de transcripción que codifican diferentes isoformas para este gen. [proporcionado por RefSeq, noviembre de 2011]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga (a). Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. ##INICIO de los Datos de Evidencia## Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025082, ERS025083 [ECO:0000350] ##FIN de los Datos de Evidencia##	44
GDPD4	exónico	220032	proteína 4 que contiene el dominio fosfodiesterasa del glicerofosfo- diéster	N/A	45
GPATCH2	intrónico	55105	Proteína 2 que contiene el dominio de parche G	N/A	46
GPC5	intrónico	2262	precursor de glipicano-5	Los proteoglicanos de sulfato de heparano de la superficie celular están compuestos por un núcleo proteico asociado a la membrana sustituido por un número variable de cadenas de sulfato de heparano. Los miembros de la familia de proteoglicanos de membrana integral relacionados con el glicano (GRIPS) contienen una proteína central anclada a la membrana citoplasmática mediante un enlace glucosil fosfatidilinositol. Estas proteínas pueden desempeñar un papel en el control de la división celular y la regulación del crecimiento. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de	47

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC030584.1, BC039730.1 [ECO:0000332] Intrones de la secuencia de ARN :: respaldo de la muestra mixta/parcial ERS025082, ERS025083 [ECO:0000350] ##FIN de los Datos de Evidencia##	
GPR19	exónico	2842	probable receptor 19 acoplado a proteína G	N/A	48
GPR142	exónico	350383	probable receptor 142 acoplado a proteína G	GPR142 es un miembro de la familia de la rodopsina de los receptores acoplados a proteínas G (GPR) (Fredriksson et al., 2003 [PubMed 14623098]). [suministrado por OMIM, marzo de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AB196530.1, AY288421.1 [ECO:0000332] ##FIN de los Datos de Evidencia##	49
GPRC5C	exónico	55890	Receptor acoplado a proteína G familia C grupo 5 miembro C isoforma a	La proteína codificada por este gen es un miembro de la familia de los receptores acoplados a proteínas G de tipo 3. Los miembros de esta superfamilia se caracterizan por un motivo característico de 7 dominios transmembrana. Se desconoce la función específica de esta proteína; sin embargo, esta proteína puede mediar los efectos celulares del ácido retinoico en la cascada de transducción de señales de la proteína G. Para este gen se han encontrado dos variantes de transcripción que codifican isoformas diferentes. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (a). ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC110848.1, AK131210.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025083 [ECO:0000348] ##FIN de los Datos de Evidencia##	50
GRIA3	intrónico	2892	precursor de la isoforma 1 del	Los receptores de glutamato son los receptores de neurotransmisores	51

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
			receptor de glutamato 3	excitatorios predominantes en el cerebro de los mamíferos y se activan en diversos procesos neurofisiológicos normales. Estos receptores son complejos proteínicos heteroméricos compuestos por múltiples subunidades, dispuestas para formar canales iónicos activados por ligandos. La clasificación de los receptores de glutamato se basa en su activación por diferentes agonistas farmacológicos. La subunidad codificada por este gen pertenece a una familia de receptores de glutamato sensibles a AMPA (propionato de alfa-amino-3-hidroxi-5-metil-4-isoxazol), y está sujeta a la edición del ARN (AGA->GGA; R->G). El empalme alternativo en este locus da lugar a diferentes isoformas, que pueden variar en sus propiedades de transducción de señales. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) codifica la isoforma 1 (también conocida como isoforma alterna). La edición del ARN (AGA->GGA) cambia Arg775Gly. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Atributos de RefSeq## sufre edición del ARN :: PMID: 10688364, 7992055 ##FIN de los Atributos de RefSeq## ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: U10301.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025082, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
GTPBP4	exónico	23560	proteína nucleolar de unión a GTP 1	Las proteínas de unión a GTP son GTPasas y funcionan como interruptores moleculares que pueden alternar entre dos estados: activo, cuando se une GTP, e inactivo, cuando se une GDP. 'Activo' en este contexto suele significar que la molécula actúa como señal para desencadenar otros acontecimientos en la célula. Cuando un ligando extracelular se une a un receptor ligado a una proteína G, el receptor cambia su	52

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				conformación y activa las proteínas G triméricas que se asocian a él haciendo que expulsen su GDP y lo sustituyan por GTP. El interruptor se apaga cuando la proteína G hidroliza su propio GTP unido, convirtiéndolo de nuevo en GDP. Pero antes de que esto ocurra, la proteína activa tiene la oportunidad de difundirse lejos del receptor y transmitir su mensaje durante un periodo prolongado a su objetivo posterior. [proporcionado por RefSeq, julio de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK001552.1, AK222861.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
HCN1	exónico	348980	canal 1 activado por nucleótidos cíclicos activados por hiperpolarización de potasio/sodio	La proteína de membrana codificada por este gen es un canal catiónico activado por hiperpolarización que contribuye a las corrientes marcapasos nativas del corazón y las neuronas. La proteína codificada puede homodimerizarse o heterodimerizarse con otras subunidades formadoras de poros para formar un canal de potasio. Este canal puede actuar como receptor de sabores agrios. [proporcionado por RefSeq, octubre de 2011]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF488549.1, AF064876.1 [ECO:0000332] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025082	53

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				[ECO:0000350] ##FIN de los Datos de Evidencia##	
HEXA	exónico	3073	preproteína de la subunidad alfa de la beta-hexosaminidasa	Este gen codifica la subunidad alfa de la enzima lisosomal beta-hexosaminidasa que, junto con la proteína activadora del cofactor GM2, cataliza la degradación del ganglíosido GM2, y otras moléculas que contienen N-acetil hexosaminas terminales. La beta-hexosaminidasa se compone de dos subunidades, alfa y beta, codificadas por genes distintos. Las subunidades alfa y beta de la beta-hexosaminidasa pertenecen a la familia 20 de las glicosil hidrolasas. Las mutaciones en los genes de las subunidades alfa o beta provocan una acumulación de ganglíosido GM2 en las neuronas y trastornos neurodegenerativos denominados gangliosidosis GM2. Las mutaciones del gen de la subunidad alfa provocan la enfermedad de Tay-Sachs (GM2-gangliosidosis tipo I). [proporcionado por RefSeq, julio de 2009]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica porque no se disponía de una única transcripción para toda la longitud del gen. La extensión de esta transcripción está respaldada por las alineaciones de transcripción. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: M13520.1, CR627386.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	54
HK2	exónico	3099	hexoquinasa-2	Las hexoquinasas fosforilan la glucosa para producir glucosa-6-fosfato, el primer paso en la mayoría de las rutas del metabolismo de la glucosa. Este gen codifica la hexoquinasa 2, la forma predominante que se encuentra en el músculo esquelético. Se localiza en la membrana externa de las mitocondrias. La expresión de este gen responde a la insulina, y los	55

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				estudios realizados en ratas sugieren que está implicado en el aumento de la tasa de glucólisis que se observa en las células cancerosas de crecimiento rápido. [proporcionado por RefSeq, abril de 2009]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC064369.1, AF148513.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025083, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
HMGB3	exónico	3149	proteína del grupo de alta movilidad B3	HMGB3 pertenece a la superfamilia de proteínas del grupo de alta movilidad (HMG). Al igual que HMG1 (MIM 163905) y HMG2 (MIM 163906), HMGB3 contiene dominios de caja HMG de unión al ADN y se clasifica en la subfamilia de caja HMG. Se cree que los miembros de la subfamilia de caja HMG desempeñan un papel fundamental en la replicación del ADN, el ensamblaje de nucleosomas y la transcripción (Wilke et al., 1997 [PubMed 9370291]; Nemeth et al., 2006 [PubMed 16945912]). [suministrado por OMIM, marzo de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: Y10043.1, BG176733.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	56
HPR	exónico	3250	precursor de la proteína relacionada con la haptoglobina	Este gen codifica una proteína relacionada con la haptoglobina que se une a la hemoglobina con la misma eficacia que la haptoglobina. A diferencia de la haptoglobina, la concentración plasmática de esta proteína no se ve afectada en pacientes con anemia falciforme y hemólisis intravascular extensa, lo que sugiere una diferencia en la unión entre los complejos haptoglobina-hemoglobina y proteína relacionada con la	57

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				haptoglobina-hemoglobina a CD163, el receptor depurador de la hemoglobina. Esta proteína también puede ser un predictor clínicamente importante de la recurrencia del cáncer de mama. [proporcionado por RefSeq, octubre de 2011]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: CB147217.1, CB122261.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	
HTATSF1P2	exónico	401233	N/A	N/A	58
IDI2	exónico	91734	delta-isomerasa 2 de difosfato de isopentenilo	N/A	59
IDI2-AS1	exónico	55853	N/A	N/A	60
IDO2	intrónico	169355	indolamina 2,3-dioxigenasa 2	Junto con las enzimas codificadas por los genes INDO (MIM 147435) y TDO2 (MIM 191070), la enzima codificada por el gen INDOL1 metaboliza el triptófano en la vía de la quinurenína (Ball et al., 2007 [PubMed 17499941]). [suministrado por OMIM, febrero de 2011]. Nota de secuencia: La UTR 3' de la transcripción de RefSeq se derivó de la secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas se basaron en alineaciones de transcripciones.	61
IFNLR1	exónico	163702	precursor de la isoforma 1 del receptor 1 de interferón lambda	La proteína codificada por este gen es un miembro de la familia de receptores de citoquinas de clase II. Esta proteína forma un complejo receptor con el receptor de interleuquina 10, beta (IL10RB). Se ha demostrado que el complejo receptor interactúa con tres citoquinas estrechamente relacionadas, tal como la interleuquina 28A (IL28A), la interleuquina 28B (IL28B) y la interleuquina 29 (IL29). La expresión de las tres citoquinas puede ser inducida por una	62

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				infección viral. Las células que sobreexpresan esta proteína se ha	
				encontrado que presentaban una mayor respuesta a la IL28A y la IL29, pero una menor respuesta a la IL28B. Se han descrito tres variantes de la transcripción empalmadas alternativamente que codifican isoformas distintas. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la proteína más larga (isoforma 1). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF439325.1, AK160364.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
IQCB1	exónico	9657	Isoforma a de la proteína 1 que contiene el motivo unión a calmodulina IQ	Este gen codifica una proteína nefrocistina que interactúa con la calmodulina y la proteína reguladora de la GTPasa de la retinosis pigmentaria. La proteína codificada tiene una región central en espiral y dos dominios IQ de unión a calmodulina. Se localiza en los cilios primarios de las células epiteliales renales y en los cilios conectores de las células fotorreceptoras. Se cree que esta proteína desempeña un papel en la función ciliar. Los defectos en este gen dan lugar al síndrome de Senior-Loken tipo 5. El empalme alternativo da lugar a múltiples variantes de la transcripción. [proporcionado por RefSeq, noviembre de 2009]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga (a). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: D25278.1,	63

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				AY714228.1 [ECO:0000332] Intrones de la secuencia de ARN :: muestra mixta/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
JPX	intrónico	554203		JPX es un ARN no codificador de proteínas transcrito a partir de un gen dentro del centro de inactivación del cromosoma X (XIC; MIM 314670) que parece participar en la inactivación del cromosoma X (Tian et al., 2010 [PubMed 21029862]). [suministrado por OMIM, febrero de 2011]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC071776.1 [ECO:0000332] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	64
KANK1	intrónico	23189	Motivo KN e isoforma a de la proteína 1 que contiene el dominio de repetición de anquirina	La proteína codificada por este gen es un miembro de la familia de proteínas Kank, que contienen múltiples dominios de repetición de anquirina. Este miembro de la familia interviene en la formación del citoesqueleto regulando la polimerización de la actina. Este gen es un supresor tumoral candidato para el carcinoma de células renales. Las mutaciones en este gen causan la parálisis cerebral tetrapléjica espástica de tipo 2, un trastorno del desarrollo del sistema nervioso central. Una translocación t(5;9) da lugar a la fusión del gen del receptor beta del factor de crecimiento derivado de plaquetas (PDGFRB) en el cromosoma 5 con este gen en una neoplasia mieloproliferativa que presenta trombocitemia grave. El empalme alternativo de este gen da lugar a múltiples variantes de transcripción. Se ha identificado un pseudogén relacionado en el cromosoma 20. [proporcionado por RefSeq, marzo de 2012]. Variante de transcripción: Esta variante (1) representa la transcripción más corta, pero codifica la isoforma más larga (a, también conocida como Kank-L). Las variantes 1, 3 y 4 codifican todas la isoforma a. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera	65

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AL833161.1, AK292989.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025085 [ECO:0000348] ##FIN de los Datos de Evidencia##	
KAT6B	exónico	23522	histona acetiltransferasa KAT6B isoforma 1	La proteína codificada por este gen es una histona acetiltransferasa y componente del complejo proteico MOZ/MORF. Además de su actividad acetiltransferasa, la proteína codificada tiene actividad de activación transcripcional en su extremo terminal N y actividad de represión transcripcional en su extremo terminal C. Esta proteína es necesaria para la activación transcripcional dependiente de RUNX2 y podría estar implicada en el desarrollo cerebral. Se han encontrado mutaciones en pacientes con síndrome genitopatelar. Una translocación de este gen y del gen CREBBP da lugar a leucemias mieloides agudas. Para este gen se han encontrado tres variantes de transcripción que codifican diferentes isoformas. [proporcionado por RefSeq, marzo de 2012]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (1). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF217500.1, BC150618.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los	66

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				intrones ERS025083, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
KCTD7	exónico	154881	Isoforma 1 de la proteína KCTD7 que contiene el dominio BTB/POZ	Este gen codifica un miembro de la familia de proteínas que contienen dominios de tetramerización de canales de potasio. Los miembros de la familia se identifican sobre una base estructural y contienen un dominio del terminal amino similar al dominio T1 presente en el canal de potasio activado por voltaje. Las mutaciones en este gen se han asociado con la epilepsia mioclónica progresiva 3. El empalme alternativo da lugar a múltiples variantes de transcripción. [proporcionado por RefSeq, enero de 2011]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (1). Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK056631.1, BU902852.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025083 [ECO:0000348] ##FIN de los Datos de Evidencia##	67
KIF19	exónico	124602	proteína similar a la kinesina KIF19	N/A	68
LARP4B	exónico	23185	proteína 4B relacionada con la	Este gen codifica un miembro de una familia de proteínas evolutivamente conservada implicada en el metabolismo y la traducción del ARN. Los miembros de esta familia se caracterizan por la presencia de un motivo La, que a menudo se encuentra adyacente a uno o más motivos de reconocimiento de ARN (RRM). Juntos, los dos motivos constituyen la región funcional de la proteína y permiten su interacción con el sustrato de ARN. Esta familia de proteínas se divide en cinco subfamilias: las auténticas proteínas La y cuatro subfamilias de	69

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				proteínas relacionadas con La (LARP). La proteína codificada por este gen pertenece a la subfamilia LARP 4. Es una proteína citoplasmática que puede desempeñar un papel estimulador de la traducción. [proporcionado por RefSeq, octubre de 2012]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. ##INICIO de los Datos de Evidencia## Combinación de exones CDS :: BC152443.1, D86971.2 [ECO:0000331] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025088 [ECO:0000350] ##FIN de los Datos de Evidencia##	
LOC643339	exónico	643339	N/A	N/A	70
LOH12CR1	exónico	118426	pérdida de heterocigosidad 12 de la proteína de la región cromosómica 1	N/A	71
MALL	exónico	7851	Proteína similar a MAL	Este gen codifica un elemento de la maquinaria para el tráfico mediado por balsas en las células endoteliales. La proteína codificada, miembro de la familia de proteolípidos MAL, se localiza predominantemente en balsas de membrana enriquecidas en glicolípidos y colesterol (GEM). Interactúa con la caveolina-1. [proporcionado por RefSeq, julio de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK125647.1, AK056616.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	72
MAPK9	exónico	5601	proteína quinasa 9 activada por mitógenos isoforma alfa1	La proteína codificada por este gen es un miembro de la familia de las MAP quinasas. Las MAP quinasas actúan como punto de integración de múltiples señales bioquímicas y participan en una amplia variedad de procesos celulares como la proliferación, la diferenciación, la	73

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				regulación de la transcripción y el desarrollo. Esta quinasa se dirige a factores de transcripción específicos, mediando así la expresión génica temprana inmediata en respuesta a diversos estímulos celulares. Está estrechamente relacionada con MAPK8, ambas implicadas en la apoptosis inducida por la radiación UV, que se cree que está relacionada con la vía de la muerte celular mediada por el citocromo c. Este gen y MAPK8 también se conocen como quinasas del terminal N de c-Jun. Esta quinasa bloquea la ubiquitinación del supresor tumoral p53, por lo que aumenta la estabilidad de p53 en células no estresadas. Los estudios del homólogo de este gen en ratones sugieren un papel clave en la diferenciación de las células T. Se han descrito varias variantes de transcripción empalmadas alternativamente que codifican isoformas distintas, [proporcionado por RefSeq, septiembre de 2008]. Variante de transcripción: Esta variante (JNK2-a1) utiliza un sitio de empalme aceptor diferente en el último exón codificante en comparación con la variante de transcripción JNK2-a2, lo que da lugar a un desplazamiento de marco y a una isoforma más corta (JNK2 alfa1) con un terminal C diferente, en comparación con la isoforma JNK2 alfa2. La variante JNK2-a1 difiere de la variante JNK2-b1 en el uso de un exón codificante interno alternativo de la misma longitud. De este modo, la isoforma JNK2 alfa tiene la misma longitud que la isoforma JNK2 beta1, con unas pocas diferencias de aa en un segmento interno de la proteína. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica porque no se disponía de una única transcripción para toda la longitud del gen. La extensión de esta transcripción está respaldada por las alineaciones de transcripción. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones que	
				están disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras	

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exones CDS :: U34821.1 [ECO:0000331] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
MCEE	ambos	84693	metilmalonil-CoA epimerasa, precursor mitocondrial	El producto de este gen cataliza la interconversión de D- y L-metilmalonil-CoA durante la degradación de aminoácidos de cadena ramificada. Ácidos grasos de longitud de cadena impar y otros metabolitos. Las mutaciones en este gen dan lugar a la deficiencia de metilmalonil-CoA epimerasa, que se presenta como aciduria metilmalónica de leve a moderada. [proporcionado por RefSeq, julio de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC020825.1, BG567074.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	74
MGAT5	intrónico	4249	alfa-1,6-manosilglicopro-teína 6-beta-N-acetilglucosaminil transferasa A	La proteína codificada por este gen es un miembro de la familia de las glicosiltransferasas. Cataliza la adición de beta-1,6-N-acetilglucosamina a la manosa ligada alfa de los oligosacáridos ligados a N biantenarios presentes en las glicoproteínas recién sintetizadas. Es una de las enzimas más importantes que intervienen en la regulación de la biosíntesis de los oligosacáridos de las glicoproteínas. Las alteraciones de los oligosacáridos de las glicoproteínas de la superficie celular provocan cambios significativos en el comportamiento adhesivo o migratorio de una célula. El aumento de la actividad de esta enzima se ha correlacionado con la progresión de los tumores malignos invasivos. [proporcionado por RefSeq, octubre de 2011]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones	75

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				se basaron en alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: D17716.1, AF113921.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
MGC16275	exónico	85001	N/A	N/A	76
MGME1	exónico	92667	exonucleasa 1 de mantenimiento del genoma mitocondrial	N/A	77
MIR200A	exónico	406983		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) del tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de	78

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones.	
MIR200B	exónico	406984		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el precursor intermediario de	79
				miARN producido por la escisión de Drosha. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por	

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				favor, consulte el registro de Gene para acceder a otras publicaciones.	
MIR429	exónico	554210		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones.	80
MIR595	exónico	693180		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son	81

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha.	
MIR651	exónico	723779		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidos y poliadenilados (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN	82

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha.	

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
MIR3163	exónico	100423029		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle predicho de microARN. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha.	83
MIR3910-1	exónico	100500821		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III	84

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Es posible que algunas secuencias de los extremos 5' y 3' no estén incluidas en el miARN precursor intermedio producido por la escisión de Drosha.	
MIR3910-2	exónico	100500902		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del	85

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha.	
MIR4267	exónico	100422994		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha.	86

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
MIR4436B1	exónico	100616123		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN previsto de acuerdo con la definición de miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha.	87
MIR4436B2	exónico	100847033		Los microARN (miARN) son ARN no codificantes cortos (20-24 nt) que intervienen en la regulación postranscripcional de la expresión génica en organismos multicelulares al afectar tanto a la estabilidad como a la traducción de los ARNm. Los miARN son transcritos por la ARN polimerasa II como parte de transcripciones primarias protegidas y poliadeniladas (pri-miARN) que pueden ser codificantes o no codificantes de proteínas. La transcripción primaria es escindida por la enzima ribonucleasa III	88

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				Drosha para producir un miARN precursor (pre-miARN) de tallo-bucle de aproximadamente 70 nt, que es escindido a su vez por la ribonucleasa citoplasmática Dicer para generar el miARN maduro y los productos estrella del miARN antisentido (miARN*). El miARN maduro se incorpora a un complejo de silenciamiento inducido por ARN (RISC), que reconoce los ARNm diana a través del emparejamiento imperfecto de bases con el miARN y, en la mayoría de los casos, provoca la inhibición de la traducción o la desestabilización del ARNm diana. El RefSeq representa el tallo-bucle de microARN predicho. [proporcionado por RefSeq, septiembre de 2009]. Nota de secuencia: Este registro representa un tallo-bucle de microARN predicho de acuerdo con lo definido por miRBase. Algunas secuencias de los extremos 5' y 3' pueden no estar incluidas en el miARN precursor intermedio producido por la escisión de Drosha.	
MKL1	intrónico	57591	MKL/proteína similar a la miocardina 1	La proteína codificada por este gen interactúa con el factor de transcripción miocardina, un regulador clave de la diferenciación de las células musculares lisas. La proteína codificada es predominantemente nuclear y puede ayudar a transducir señales del citoesqueleto al núcleo. Este gen está implicado en un evento de translocación específico que crea una fusión de este gen y el gen de la proteína-15 con motivo de unión al ARN. Esta translocación se ha asociado a la leucemia megacariocítica aguda. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AB037859.2, AJ297258.1 [ECO:0000332] Intrones de la secuencia de ARN :: muestra mixta/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	89

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
MRPL42	exónico	28977	proteína ribosómica 39S L42, precursora mitocondrial	Las proteínas ribosómicas mitocondriales de los mamíferos están codificadas por genes nucleares y contribuyen a la síntesis de proteínas en la mitocondria. Los ribosomas mitocondriales (mitoribosomas) constan de una subunidad 28S pequeña y una subunidad 39S grande. Tienen una composición estimada del 75 % de proteínas en relación con el ARNr, en comparación con los ribosomas procariotas, en los que esta proporción se invierte. Otra diferencia entre los mitoribosomas de los mamíferos y los ribosomas procariotas es que estos últimos contienen un ARNr 5S. Entre las distintas especies, las proteínas que componen el mitorribosoma difieren mucho en secuencia y, a veces, en propiedades bioquímicas, lo que impide reconocerlas fácilmente por homología de secuencia. Este gen codifica una proteína identificada como perteneciente tanto a la subunidad 28S como a la 39S. El empalme alternativo da lugar a múltiples variantes del transcripción. Los pseudogenes correspondientes a este gen se encuentran en los cromosomas 4q, 6p, 6q, 7p y 15q. [proporcionado por RefSeq, mayo de 2011]. Variante de transcripción: Esta variante (1) codifica la proteína soportada. Las variantes 1 y 2 codifican la misma proteína. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. ##INICIO de los Atributos de RefSeq## producto(s) génico(s) localizado(s) en la mitocondria. :: reportado por MitoCarta ##FIN de los Atributos de RefSeq## ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK000285.1, AF151038.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	90

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
MTHFD1	exónico	4522	C-1-tetrahidrofolato sintasa, citoplasmática	Este gen codifica una proteína que posee tres actividades enzimáticas distintas, la 5,10-metilen tetrahidrofolato deshidrogenasa, la 5,10-metiltetrahidrofolato ciclohidrolasa y la 10-formil tetrahidrofolato sintetasa. Cada una de estas actividades cataliza una de las tres reacciones secuenciales en la interconversión de derivados de 1 carbono del tetrahidrofolato, que son sustratos para las síntesis de metionina, timidilato y purina de novo. Las actividades enzimáticas trifuncionales son conferidas por dos dominios principales, una porción aminoterminal que contiene las actividades deshidrogenasa y ciclohidrolasa y un dominio sintetasa de mayor tamaño. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Atributos de RefSeq## CDS utiliza AUG en marco secuencia abajo :: evidencia experimental (PMID:3053686) ##FIN de los Atributos de RefSeq## ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC050420.1, J04031.1	91
				[ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
NFIL3	exónico	4783	proteína regulada por el factor nuclear interleucina-3	La expresión de interleucina-3 (IL3; MIM 147740) está restringida a células T activadas, células asesinas naturales (NK) y mastocitos. El inicio de la transcripción depende de la capacidad activadora de factores proteicos específicos, tales como NFIL3, que se unen a regiones reguladoras del gen, normalmente secuencia arriba del sitio de inicio de la transcripción (Zhang et al., 1995 [PubMed 7565758]). [suministrado por OMIM, febrero de 2009]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por	92

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: S79880.1, U26173.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
NLRP12	exónico	91662	Proteína 12 que contiene los dominios NACHT, LRR y PYD isoforma 2	Este gen codifica un miembro de la familia CATERPILLER de proteínas citoplasmáticas. La proteína codificada, que contiene un dominio terminal N de pirina, un dominio NACHT, un dominio asociado a NACHT y una región del terminal C de repetición rica en leucina, funciona como factor atenuante de la inflamación al suprimir las respuestas inflamatorias en monocitos activados. Las mutaciones en este gen causan el síndrome autoinflamatorio familiar por frío de tipo 2. El empalme alternativo da lugar a múltiples variantes de la transcripción. [proporcionado por RefSeq, marzo de 2013]. Variante de transcripción: Esta variante (2) utiliza un sitio de empalme alternativo en la región codificante central, en comparación con la variante 3, que da como resultado una isoforma (2) que es 1 aa más corta que la isoforma 3. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AY095146.1, BC028069.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025089 [ECO:0000348] ##FIN de los Datos de Evidencia##	93
NQO2	exónico	4835	Ribosildihidronico-tinamida deshidrogenasa [quinona]	La NQO2 (EC 1.10.99.2) es una flavoproteína que cataliza la reducción de 2 electrones de diversas quinonas, colorantes redox y la menadiona de la vitamina K. La NQO2 utiliza predominantemente ribósido dihidronicotinamida (NRH) como donante de electrones (resumen de Wu et al., 1997	94

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				[PubMed 9367528]). [suministrado por OMIM, julio de 2010]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: J02888.1, AK311746.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
NRIP1	exónico	8204	proteína 1 de interacción con receptores nucleares	La proteína 1 de interacción con receptores nucleares (NRIP1) es una proteína nuclear que interactúa específicamente con el dominio de activación dependiente de hormonas AF2 de los receptores nucleares. También conocida como RIP140, esta proteína modula la actividad transcripcional del receptor de estrógenos. [proporcionado por RefSeq, julio de 2008]. Nota de secuencia: la transcripción y la proteína RefSeq se derivaron de la transcripción y la secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. La extensión de esta transcripción de RefSeq está respaldada por alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK289786.1, DA230125.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025098 [ECO:0000348] ##FIN de los Datos de Evidencia##	95
NUDT4	exónico	11163	difosfoinositol polifosfato fosfohidrolasa 2 isoforma alfa	La proteína codificada por este gen regula el recambio de los polifosfatos de difosfoinositol. El recambio de estos polifosfatos de difosfoinositol de alta energía representa una actividad de cambio molecular con importantes	96

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				consecuencias reguladoras. La conmutación molecular por polifosfatos de difosfoinositol puede contribuir a regular el tráfico intracelular. Se han descrito varias variantes de la transcripción empalmadas alternativamente, pero no se ha determinado la naturaleza de la longitud completa de algunas variantes. Las isoformas DIPP2alfa y DIPP2beta se distinguen entre sí únicamente porque la DIPP2beta posee un aminoácido adicional debido al deslizamiento del límite del intrón en el empalme alternativo. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) codifica la isoforma predominante (alfa). ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF191651.1, AF191650.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
NUDT4P1	exónico	440672	N/A	N/A	97
OVOL2	exónico	58495	factor de transcripción similar a Ovo 2	N/A	98
PDE3B	intrónico	5140	fosfodiesterasa B cíclica 3',5' inhibida por cGMP	N/A	99
PDGFRA	exónico	5156	precursor del receptor alfa del factor de crecimiento derivado de plaquetas	Este gen codifica un receptor tirosina quinasa de superficie celular para miembros de la familia del factor de crecimiento derivado de plaquetas. Estos factores de crecimiento son mitógenos para las células de origen mesenquimal. La identidad del factor de crecimiento unido a un monómero del receptor determina si el receptor funcional es un homodímero o un heterodímero, compuesto por los polipéptidos alfa y beta del receptor del factor de crecimiento derivado de plaquetas. Los estudios sugieren que este gen desempeña un papel en el desarrollo de órganos, la cicatrización de heridas y la progresión tumoral. Las mutaciones en este gen se han asociado con el síndrome hipereosinofílico idiopático, a tumores somáticos y familiares del estroma gastrointestinal y a otros tipos de	100

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				cáncer. [proporcionado por RefSeq, marzo de 2012]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica porque no se disponía de una única transcripción para toda la longitud del gen. La extensión de esta transcripción está respaldada por alineaciones de transcripciones y datos ortólogos. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: M21574.1, M22734.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
PDSS2	exónico	57107	decaprenil-difosfato sintasa subunidad 2	La proteína codificada por este gen es una enzima que sintetiza la cadena lateral de prenilo de la coenzima Q, o ubiquinona, uno de los elementos clave de la cadena respiratoria. El producto génico cataliza la formación de todos los pirofosfatos de transpoliprenilo a partir de difosfato de isopentilo en el ensamblaje de cadenas laterales poliisoprenoides, el primer paso en la biosíntesis de la coenzima Q. Los defectos en este gen son una de las causas de la deficiencia de coenzima Q10. [proporcionado por RefSeq, octubre de 2009]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC039906.1, AF254956.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	101
PHACTR4	exónico	65979	fosfatasa y regulador de actina 4 isoforma 1	Este gen codifica un miembro de la familia de fosfatasas y reguladores de actina (PHACTR). Se ha demostrado que otros miembros de	102

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				la familia PHACTR inhiben la actividad de la proteína fosfatasa 1 (PP1), y que el homólogo de este gen en el ratón interactúa con la actina y la PP1. Se han encontrado múltiples variantes de transcripción que codifican diferentes isoformas para este gen. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) representa la transcripción más larga, pero codifica la isoforma más corta (1). Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: CR749449.1, BC029266.1 [ECO:0000332] Introns de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025087 [ECO:0000348] ##FIN de los Datos de Evidencia##	
PIAS2	exónico	9063	E3 SUMO-proteína ligasa PIAS2 isoforma alfa	Este gen codifica un miembro de la familia de la proteína inhibidora de STAT activada (PIAS). Las proteínas PIAS funcionan como ligasas SUMO E3 y desempeñan un papel importante en muchos procesos celulares al mediar en la sumoilación de proteínas diana. Para este gen se han observado variantes de transcripción empalmadas alternativamente que codifican múltiples isoformas. Las isoformas de la proteína codificada potencian la sumoilación de proteínas diana específicas, entre ellas la proteína supresora de tumores p53, c-Jun y el receptor de andrógenos. Un pseudogén de este gen se localiza en el brazo corto del cromosoma 4. El símbolo MIZ1 también se ha asociado con ZBTB17 que es un gen diferente localizado en el cromosoma 1. [proporcionado por RefSeq, agosto de 2011]. Variante de transcripción: Esta variante (alfa) utiliza un exón codificante 3' alternativo, en comparación con la variante beta, que da como resultado una	103

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				isoforma más corta (alfa) que tiene un terminal C único en comparación con la isoforma beta. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC015190.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	
PIK3CD	exónico	5293	subunidad catalítica 4,5-bisfosfato 3-quinasa de fosfatidilinositol isoforma delta	Las fosfoinositida 3 quinasas (PI3K) fosforilan los lípidos de inositol y participan en la respuesta inmunitaria. La proteína codificada por este gen es una PI3K de clase I que se encuentra principalmente en los leucocitos. Al igual que otras PI3K de clase I (p110-alfa, p110-beta y p110-gamma), la proteína codificada se une a proteínas adaptadoras p85 y a RAS unido a GTP. Sin embargo, a diferencia de las otras PI3K de clase I, esta proteína se fosforila a sí misma, no a la proteína p85. [proporcionado por RefSeq, julio de 2010]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: U86453.1, Y10055.2 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025089 [ECO:0000348] ##FIN de los Datos de Evidencia##	104
PKHD1	intrónico	5314	precursor de la isoforma 1 de la fibrocistina	Se predice que la proteína codificada por este gen tiene un único dominio transmembrana (TM)	105

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				y múltiples copias de un dominio de factor de transcripción de plexina similar a la inmunoglobulina. El empalme alternativo da lugar a dos variantes de transcripción que codifican isoformas diferentes. Se han descrito otras transcripciones empalmadas alternativamente, pero no se han determinado las secuencias de longitud completa. Se prevé que varios de estas transcripciones codifican productos truncados que carecen de TM y pueden ser secretados. Las mutaciones en este gen causan la enfermedad renal poliquística autosómica recesiva, también conocida como riñón poliquístico y enfermedad hepática 1. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga de esta proteína. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AY074797.1, AF480064.1 [ECO:0000332] Introns de la secuencia de ARN :: muestra mixta/parcial ERS025084, ERS025085 [ECO:0000350] ##FIN de los Datos de Evidencia##	
PLXNC1	exónico	10154	Plexina C1	Este gen codifica un miembro de la familia de la plexina. Las plexinas son receptores transmembrana de las semaforinas, una gran familia de proteínas que regulan la orientación de los axones, la motilidad y migración celular y la respuesta inmunitaria. La proteína codificada y su ligando regulan la adhesión de los melanocitos, y las semaforinas virales pueden modular la respuesta inmunitaria uniéndose a este receptor. La proteína codificada puede ser una proteína supresora de tumores para el melanoma. Se han observado variantes de transcripción empalmadas alternativamente para este gen. [proporcionado por RefSeq, enero de 2011]. Variante de transcripción: Esta variante (2) carece de múltiples exones 5' pero contiene un exón 5' alternativo, en comparación con la variante 1. Esta variante se	106

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				representa como no codificante debido a la presencia de un ORF secuencia arriba que se predice que interfiere con la traducción del ORF más largo dentro del marco. La traducción del ORF secuencia arriba hace que la transcripción sea candidata a la descomposición del ARNm mediada no sentido (NMD). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	
PNPLA4	exónico	8228	precursor de la isoforma 1 de la proteína 4 que contiene un dominio de fosfolipasa similar a la patatina	Este gen codifica un miembro de la familia de las fosfolipasas similares a la patatina. La enzima codificada tiene actividades tanto de triacilglicerol lipasa como de transacilasa y puede estar implicada en la homeostasis de los triglicéridos en los adipocitos. El empalme alternativo da lugar a múltiples variantes de transcripción. Un pseudogén de este gen se encuentra en el cromosoma Y. [proporcionado por RefSeq, febrero de 2010]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (1). Las variantes 1 y 2 codifican la misma isoforma (1). Nota de secuencia: la transcripción y la proteína RefSeq se derivaron del transcripción y la secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. La extensión de esta transcripción está respaldada por las alineaciones de transcripción. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: U03886.1, AK289888.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025091, ERS025098 [ECO:0000348] ##FIN de los Datos de Evidencia##	107
PNPT1	ambos	87178	polirribonucleótido nucleotidiltransfera-	La proteína codificada por este gen es un miembro de la familia de las polinucleótido fosforilasas,	108

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
			sa 1, precursor mitocondrial	conservada evolutivamente y compuesta por exorribonucleasas dependientes de fosfatos 3'-a-5' implicadas en el procesamiento y la degradación del ARN. Esta enzima se localiza predominantemente en el espacio intermembrana mitocondrial y participa en la importación de ARN a las mitocondrias. Las mutaciones en este gen se han asociado a la deficiencia combinada de fosforilación oxidativa-13 y a la sordera autosómica recesiva no sindrómica-70. Los pseudogenes relacionados se encuentran en los cromosomas 3 y 7. [proporcionado por RefSeq, diciembre de 2012]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC053660.1, AJ458465.1 [ECO:0000332] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia## ##INICIO de los Atributos de RefSeq## producto(s) génico(s) localizado(s) en la mitocondria :: PMID: 12798676; comunicado por MitoCarta ##FIN de los Atributos de RefSeq##	
PPP2R3B	intrónico	28227	serina/treonina-proteína fosfatasa 2A subunidad reguladora B'' subunidad beta	La proteína fosfatasa 2 (antes denominada tipo 2A) es una de las cuatro fosfatasas Ser/Thr principales y está implicada en el control negativo del crecimiento y la división celular. Las holoenzimas de la proteína fosfatasa 2 son proteínas heterotriméricas compuestas por una subunidad estructural A, una subunidad catalítica C y una subunidad reguladora B. La subunidad reguladora está codificada por un conjunto diverso de genes que se han agrupado en las familias B/PR55, B'/PR61 y B''/PR72. Estas diferentes subunidades reguladoras confieren distintas especificidades enzimáticas y localizaciones intracelulares a la holoenzima. El producto de este gen es un miembro de la familia B''. La familia "B" se ha dividido a su vez en subfamilias. El	109

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				producto de este gen pertenece a la subfamilia beta de la subunidad reguladora B". [proporcionado por RefSeq, abril de 2010]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BK000521.1, BC063429.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
PRKCB	ambos	5579	proteína quinasa C tipo beta isoforma 1	La proteína quinasa C (PKC) es una familia de proteínas quinasas específicas de serina y treonina que pueden ser activadas por el calcio y el segundo mensajero, el diacilglicerol. Los miembros de la familia PKC fosforilan una amplia variedad de proteínas diana y se sabe que participan en diversas vías de señalización celular. Los miembros de la familia PKC también actúan como receptores principales de los ésteres de forbol, una clase de promotores tumorales. Cada miembro de la familia PKC tiene un perfil de expresión específico y se cree que desempeña una función distinta en las células. La proteína codificada por este gen es uno de los miembros de la familia PKC. Esta proteína quinasa interviene en numerosas funciones celulares, tal como la activación de los linfocitos B, la inducción de la apoptosis, la proliferación de las células endoteliales y la absorción intestinal de azúcares. Los estudios en ratones también sugieren que esta quinasa también puede regular las funciones neuronales y correlacionar el comportamiento conflictivo inducido por el miedo tras el estrés. Se han descrito variantes de transcripción empalmadas alternativamente que codifican isoformas distintas. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) utiliza un empalme alternativo en el	110

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				extremo 5' del último exón en comparación con la variante 2. La isoforma resultante (1) tiene un terminal C distinto y más corto en comparación con la isoforma 2. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica porque no se disponía de una única transcripción para toda la longitud del gen. La extensión de esta transcripción está respaldada por las alineaciones de transcripción. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: X06318.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025082, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
PRKCH	intrónico	5583	proteína quinasa C tipo eta	La proteína quinasa C (PKC) es una familia de proteínas quinasas específicas de serina y treonina que pueden ser activadas por el calcio y el segundo mensajero, el diacilglicerol. Los miembros de la familia PKC fosforilan una amplia variedad de proteínas diana y se sabe que participan en diversas vías de señalización celular. Los miembros de la familia PKC también actúan como receptores principales de los ésteres de forbol, una clase de promotores tumorales. Cada miembro de la familia PKC tiene un perfil de expresión específico y se cree que desempeña una función distinta en las células. La proteína codificada por este gen es uno de los miembros de la familia PKC. Es una proteína quinasa independiente del calcio y dependiente de los fosfolípidos. Se expresa predominantemente en los tejidos epiteliales y se ha demostrado que reside específicamente en el núcleo celular. Esta proteína quinasa puede regular la diferenciación de los queratinocitos mediante la activación de la cascada de la proteína quinasa activada por la MAP quinasa MAPK13 (p38delta) que se dirige a CCAAT/proteína alfa	111

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				de unión al potenciador (CEBPA). También se ha descubierto que media en la activación de la transcripción del gen de la transglutaminasa 1 (TGM1). [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC037268.1, AK290183.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025083 [ECO:0000348] ##FIN de los Datos de Evidencia##	
PSTPIP1	exónico	9051	proteína 1 que interactúa con la fosfatasa prolina-serina-treonina	La proteína codificada por este gen se une a la cola citoplasmática de CD2, un efector de la activación y adhesión de células T, afectando negativamente a la activación de células T desencadenada por CD2. La proteína codificada parece ser una proteína andamio y un regulador del citoesqueleto de actina. También se ha demostrado que se une a ABL1, PTPN18, WAS, CD2AP y PTPN12. Las mutaciones en este gen son causa del síndrome PAPA, [proporcionado por RefSeq, Jul 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC008602.1, U94778.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	112
PTPN2	exónico	5771	tirosina-proteína fosfatasa no receptora tipo 2 isoforma 1	La proteína codificada por este gen es un miembro de la familia de las proteínas tirosina fosfatasas (PTP). Los miembros de la familia PTP comparten un motivo catalítico altamente conservado, que es esencial para la actividad catalítica. Se sabe que las PTP son moléculas de señalización que regulan diversos procesos celulares, tal	113

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				como el crecimiento celular, la diferenciación, el ciclo mitótico y la transformación oncogénica. El receptor del factor de crecimiento epidérmico y la proteína adaptadora She son sustratos de esta PTP, lo que sugiere su papel en la señalización celular mediada por el factor de crecimiento. Se han encontrado múltiples variantes de transcripción empalmadas alternativamente que codifican diferentes isoformas. Se ha informado de la existencia de dos pseudogenes altamente relacionados, pero procesados de forma distinta, que se localizan en los cromosomas 1 y 13, respectivamente. [proporcionado por RefSeq, mayo de 2011]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga (1). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: M25393.1, AK292570.1 [ECO:0000332] Intrones de la secuencia de ARN :: muestra mixta/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
PTPRN2	intrónico	5799	precursor de la isoforma 1 de la fosfatasa tirosina-proteína de tipo receptor N2	La proteína codificada por este gen es un miembro de la familia de las proteínas tirosina fosfatasas (PTP). Se sabe que las PTP son moléculas de señalización que regulan diversos procesos celulares, tal como el crecimiento celular, la diferenciación, el ciclo mitótico y la transformación oncogénica. Esta PTP posee una región extracelular, una única región transmembrana y un único dominio catalítico intracelular, por lo que representa una PTP de tipo receptor. El dominio catalítico de esta PTP está estrechamente relacionado con la PTPRN/IA-2beta. Se ha descubierto que tanto el PTP como el PTPRN son autoantígenos importantes asociados a la diabetes mellitus insulinodependiente. Se han descrito tres variantes de transcripción de este gen empalmadas alternativamente, que codifican proteínas distintas.	114

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				[proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga (1). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: U66702.1, AF007555.1 [ECO:0000332] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
RAB37	exónico	326624	proteína relacionada con ras Rab-37 isoforma 2	Las proteínas Rab son GTPasas de baja masa molecular que regulan de forma decisiva el tráfico de vesículas. Para más información sobre las proteínas Rab, véase MIM 179508. [suministrado por OMIM, abril de 2006]. Variante de transcripción: Esta variante (2) representa el uso de un promotor alternativo, UTR 5' y un codón de inicio alternativo, e incluye un exón codificante alternativo, en comparación con la variante 3. La isoforma resultante (2) tiene un terminal N distinto y más largo, en comparación con la isoforma 3. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK098068.1, BX332255.2 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084, ERS025088 [ECO:0000348] ##FIN de los Datos de Evidencia##	115
RBFOX1	intrónico	54715	Proteína de unión a ARN fox-1 homóloga 1 isoforma 1	La familia Fox-1 de proteínas de unión a ARN está conservada evolutivamente y regula el empalme alternativo específico de tejido en los metazoos. Fox-1 reconoce un tramo (U)GCAUG en exones regulados o en intrones flanqueantes. La proteína se une al terminal C de la ataxina-2 y puede contribuir a la patología restringida de la ataxia espinocerebelosa tipo 2 (SCA2). La ataxina-2 es el producto del gen SCA2, causante de enfermedades neurodegenerativas familiares. Fox-1 y ataxina-2 se localizan en la red trans-Golgi. Para este gen se han encontrado varias	116

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				variantes de transcripción empalmadas alternativamente que codifican diferentes isoformas. [proporcionado por RefSeq, noviembre de 2011]. Variante de transcripción: Esta variante (1), también conocida como gamma, codifica la isoforma más larga (1). Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica porque no se disponía de una única transcripción para toda la longitud del gen. La extensión de esta transcripción está respaldada por las alineaciones de transcripción. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF229057.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025083, ERS025099 [ECO:0000348] ##FIN de los Datos de Evidencia##	
CCR1	exónico	1104	N/A	N/A	117
RGCC	exónico	28984	regulador del ciclo celular RGCC	Se cree que este gen regula la progresión del ciclo celular. Es inducida por p53 en respuesta a daños en el ADN, o por niveles sublíticos de las proteínas del sistema del complemento que dan lugar a la activación del ciclo celular. La proteína codificada se localiza en el citoplasma durante la interfase y en los centrosomas durante la mitosis. La proteína forma un complejo con la quinasa 1 tipo polo. La proteína también se transloca al núcleo en respuesta al tratamiento con proteínas del sistema del complemento, y puede asociarse con la proteína 2 del ciclo de división celular y aumentar su actividad quinasa. En diferentes ensayos y tipos celulares, se ha demostrado que la sobreexpresión de esta proteína activa o suprime la progresión del ciclo celular. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por	118

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC066334.1, BG037019.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025092 [ECO:0000348] ##FIN de los Datos de Evidencia##	
RHOQ	intrónico	23433	proteína de unión a GTP relacionada con rho precursor de RhoQ	Este gen codifica un miembro de la familia Rho de pequeñas GTPasas, que alternan entre estados inactivos unidos a GDP y estados activos unidos a GTP y funcionan como interruptores moleculares en cascadas de transducción de señales. Las proteínas Rho promueven la reorganización del citoesqueleto de actina y regulan la forma, la adhesión y la motilidad celular. La proteína codificada es una proteína de señalización importante para el ensamblaje del sarcómero y se ha demostrado que desempeña un papel significativo en la exocitosis de la familia 2 transportadora de solutos, el miembro 4 del transportador facilitado de glucosa y otras proteínas, actuando posiblemente como la señal que pone en marcha la maquinaria de fusión de membranas. Se han identificado tres pseudogenes relacionados en los cromosomas 2 y 14. [proporcionado por RefSeq, agosto de 2011]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BX428852.2, BC013135.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	119
RNASA3	exónico	6037	precursor de la proteína catiónica de los eosinófilos	N/A	120
RNASA10	exónico	338879	precursor inactivo de la proteína 10 similar a la ribonucleasa	N/A	121

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
RPL38	exónico	6169	proteína ribosómica 60S L38	Los ribosomas, los orgánulos que catalizan la síntesis de proteínas, constan de una subunidad 40S pequeña y una subunidad 60S grande. En conjunto, estas subunidades están compuestas por 4 especies de ARN y aproximadamente 80 proteínas estructuralmente distintas. Este gen codifica una proteína ribosómica que es un componente de la subunidad 60S. La proteína pertenece a la familia L38E de proteínas ribosómicas. Se encuentra en el citoplasma. Se han identificado variantes de empalme alternativas, ambas codifican la misma proteína. Como es típico de los genes que codifican proteínas ribosómicas, existen múltiples pseudogenes procesados de este gen dispersos por el genoma, incluido uno localizado en la región promotora del gen del receptor de angiotensina II tipo 1. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) es la transcripción más larga y predominante. Las variantes 1 y 2 codifican la misma proteína. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BQ276548.1, BU569438.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	122
RPTOR	intrónico	57521	proteína reguladora asociada a mTOR isoforma 1	Este gen codifica un componente de una vía de señalización que regula el crecimiento celular en respuesta a los niveles de nutrientes e insulina. La proteína codificada forma un complejo estequiométrico con la quinasa mTOR, y también se asocia con la proteína-1 de unión al factor de iniciación eucariótico 4E y la proteína ribosómica S6 quinasa. La proteína regula positivamente la proteína ribosomal efectora S6 quinasa secuencia abajo, y regula negativamente la quinasa mTOR. Se han encontrado múltiples variantes de transcripción que codifican diferentes isoformas para este gen. [proporcionado por RefSeq, septiembre de 2009]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la	123

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				isoforma más larga (1). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AY090663.1, BC136652.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025083, ERS025085 [ECO:0000348] ##FIN de los Datos de Evidencia###	
SERPINB4	exónico	6318	serpina B4	N/A	124

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
SERPINB6	exónico	5269	serpina B6 isoforma a	La proteína codificada por este gen pertenece a la superfamilia de las serpinas (inhibidores de la serina proteinasa) y a la subfamilia de las serpinas de la ovoalbúmina (ov). Se descubrió originalmente como inhibidor de la trombina placentaria. Se descubrió que el homólogo del ratón se expresa en las células ciliadas del oído interno. Las mutaciones en este gen están asociadas a la pérdida de audición progresiva no sindrómica, lo que sugiere que esta serpina desempeña un papel importante en el oído interno en la protección contra la fuga de contenido lisosomal durante el estrés, y que la pérdida de esta protección provoca la muerte celular y la pérdida de audición neurosensorial. Se han encontrado variantes de transcripción empalmadas alternativamente para este gen. [proporcionado por RefSeq, septiembre de 2010]. Variante de transcripción: Esta variante (1) representa la transcripción predominante. Las variantes 1, 5 y 6 codifican la misma isoforma (a). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK314578.1, BC098564.1 [ECO:0000332] Intrones de la secuencia de ARN :: muestra mixta/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	125
SLC3A2	ambos	6520	cadena pesada del antígeno de superficie celular 4F2 isoforma b	Este gen es un miembro de la familia transportadora de soluto y codifica una proteína transmembrana de la superficie celular. La proteína existe como la cadena pesada de un heterodímero, unido covalentemente a través de enlaces di-sulfuro a una de varias cadenas ligeras posibles. El transportador codificado desempeña un papel en la regulación de los niveles de calcio intracelular y transporta aminoácidos de tipo L. Se han caracterizado variantes de la transcripción empalmadas alternativamente, que codifican	126

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				diferentes isoformas. [proporcionado por RefSeq, noviembre de 2010]. Variante de transcripción: Esta variante (2) representa la transcripción más larga y codifica la isoforma más larga (b). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK025584.1 [ECO:0000332] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025082, ERS025084 [ECO:0000350] ##FIN de los Datos de Evidencia##	
SLC17A5	ambos	26503	sialina	Este gen codifica un transportador de membrana que exporta ácidos siálicos libres que han sido escindidos de lípidos y proteínas de la superficie celular desde los lisosomas. Las mutaciones en este gen causan enfermedades por almacenamiento de ácido siálico, incluido el trastorno por almacenamiento de ácido siálico infantil y la enfermedad de Salla, una forma adulta. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC020961.2, AJ387747.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	127
SNHG3	exónico	8420	N/A	N/A	128
SNORD17	exónico	692086	N/A	N/A	129
SNX5	exónico	27131	Clasificación de nexina-5 isoforma a	Este gen codifica un miembro de la clasificación de la familia de nexina. Los miembros de esta familia contienen un dominio phox (PX), que es un dominio de unión a fosfoinositida, y participan en el tráfico intracelular. Esta proteína interviene en la clasificación endosomal, la vía de señalización	130

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				de la fosfoinositida y macropinocitosis. Este gen puede desempeñar un papel en la tumorigénesis del carcinoma papilar de tiroides. El empalme alternativo da lugar a múltiples variantes de transcripción que codifican diferentes isoformas. [proporcionado por RefSeq, septiembre de 2013]. Variante de transcripción: Esta variante (1) difiere en la UTR 5', en comparación con la variante 2. Las variantes 1 y 2 codifican la misma proteína (isoforma a). ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC000100.3, AF121855.1 [ECO:0000332] Introns de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	
SOCS2	exónico	8835	supresor de la señalización de citoquinas 2	Este gen codifica un miembro de la familia de supresores de la señalización de citoquinas (SOCS). Los miembros de la familia SOCS son reguladores negativos inducibles por citoquinas de la señalización de los receptores de citoquinas a través de la vía Janus quinasa/transductor de señales y activación de la transcripción (la vía JAK/STAT). Las proteínas de la familia SOCS interactúan con las principales moléculas de los complejos de señalización para bloquear la transducción ulterior de la señal, en parte, mediante el agotamiento proteosomal de los receptores o de las proteínas transductoras de señales a través de la ubiquitinación. La expresión de este gen puede ser inducida por un subconjunto de citoquinas, entre ellas la eritropoyetina, el GM-CSF, la IL10, el interferón (IFN)-gamma y por receptores de citoquinas como el receptor de la hormona de crecimiento. La proteína codificada por este gen interactúa con el dominio citoplasmático del receptor del factor de crecimiento 1 similar a la insulina (IGF1R) y se cree que está implicada en la regulación de la señalización celular mediada por IGF1R. Este gen tiene pseudogenes en los cromosomas 20 y 22. El empalme alternativo da lugar a múltiples variantes de transcripción. [proporcionado por	131

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				RefSeq, julio de 2012]. Variante de transcripción: Esta variante (1) difiere en la UTR 5', en comparación con la variante 5. Las variantes 1-6 codifican la misma proteína. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK313165.1. AL522912.3 [ECO:0000332]	
				Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025082 [ECO:0000348] ##FIN de los Datos de Evidencia##	
SOCS2-AS1	exónico	144481	N/A	N/A	132
ST8SIA5	exónico	29906	alfa-2,8-sialiltransferasa 8E	La proteína codificada por este gen es una proteína de membrana de tipo II que puede estar presente en el aparato de Golgi. La proteína codificada, que es miembro de la familia 29 de las glicosiltransferasas, puede estar implicada en la síntesis de los gangliósidos GD1c, GT1a, GQ1b y GT3 a partir de GD1a, GT1b, GM1b y GD3, respectivamente. [proporcionado por RefSeq, julio de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK056270.1, BC108910.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025082, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	133
STIM2	intrónico	57620	Precursor de la isoforma 1 de la molécula de interacción estromal 2	Este gen es un miembro de la familia de las moléculas de interacción estromal (STIM) y probablemente surgió, junto con el miembro relacionado de la familia STIM1, de un gen ancestral común. La proteína codificada funciona para regular las concentraciones de calcio en el citosol y el retículo endoplásmico, y está implicada en la activación de los canales de entrada de Ca(2+) Orai de la membrana plasmática. Este gen inicia la traducción a partir de un sitio de inicio no-AUG (UUG). Se escinde un péptido señal de la proteína resultante. Múltiples	134

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				variantes de la transcripción son el resultado de empalme alternativo. [proporcionado por RefSeq, diciembre de 2009]. Variante de transcripción: Esta variante (1) codifica la isoforma más larga (1). Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC136449.1, AK096846.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia## ##INICIO de los Atributos de RefSeq## CDS utiliza secuencia abajo en el marco AUG :: evidencia experimental (PMID:11463338) codón de iniciación no AUG :: PMID: 11463338 ##FIN de los Atributos de RefSeq##	
TBC1D16	intrónico	125058	Miembro 16 de la familia del dominio TBC1 isoforma a	N/A	136
TEX29	exónico	121793	proteína de la secuencia 29 expresada en testículo	N/A	137
TNFRSF10A	exónico	8797	miembro 10A de la superfamilia de receptores del factor de necrosis tumoral	La proteína codificada por este gen pertenece a la superfamilia de los receptores del TNF. Este receptor es activado por el ligando inductor de la apoptosis relacionado con el factor de necrosis tumoral (TNFSF10/TRAIL), por lo que transduce la señal de muerte celular e induce la apoptosis celular. Los estudios con ratones deficientes en FADD sugirieron que la FADD, una proteína adaptadora que contiene un dominio de muerte, es necesaria para la apoptosis mediada por esta proteína. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC012866.1,	138

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				AK291299.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025081, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
TNFRSF13C	exónico	115650	miembro 13C de la superfamilia de receptores del factor de necrosis tumoral	El factor activador de células B (BAFF) aumenta la supervivencia de las células B *in vitro* y es un regulador de la población periférica de células B. La sobreexpresión de Baff en ratones provoca hiperplasia de células B maduras y síntomas de lupus eritematoso sistémico (SLE). Además, algunos pacientes con SLE presentan niveles elevados de BAFF en suero. Por lo tanto, se ha propuesto que los niveles anormalmente altos de BAFF pueden contribuir a la patogénesis de las enfermedades autoinmunes al aumentar la supervivencia de las células B autorreactivas. La proteína codificada por este gen es un receptor para BAFF y es una proteína transmembrana de tipo III que contiene un único dominio extracelular rico en cisteína. Se cree que este receptor es el receptor principal necesario para la supervivencia de las células B maduras mediada por BAFF. [proporcionado por RefSeq, julio de 2008]. Nota de secuencia: la transcripción y la proteína RefSeq se derivaron de la secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineamientos. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF373846.1, BC112030.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025089 [ECO:0000348] ##FIN de los Datos de Evidencia##	139

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
TNFRSF18	exónico	8784	miembro 18 de la superfamilia de receptores del factor de necrosis tumoral precursor de la isoforma 1	Este gen codifica un miembro de la superfamilia de receptores de TNF. Se ha demostrado que el receptor codificado tiene una mayor expresión tras la activación de las células T, y se cree que desempeña un papel clave en la autotolerancia inmunológica dominante mantenida por las células T reguladoras CD25(+)CD4(+). Estudios de inactivación en ratones también sugieren que el papel de este receptor es la regulación de la activación de las células T impulsada por CD3 y la muerte celular programada. Se han descrito tres variantes de la transcripción de este gen empalmadas alternativamente que codifican isoformas distintas. [proporcionado por RefSeq, febrero de 2011]. Variante de transcripción: Esta variante (1) representa la transcripción más larga. Contiene un segmento de codificación adicional, lo que provoca un desplazamiento de marco, en comparación con la variante 2. La proteína resultante (isoforma 1) contiene un del terminal C distinto y más corto, en comparación con la isoforma 2. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas para el registro de transcripciones se basaron en alineaciones de transcripciones. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AY358877.1, AF125304.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025089, ERS025093 [ECO:0000348] ##FIN de los Datos de Evidencia##	140
TRAFD1	exónico	10906	Proteína 1 con dominio de dedos de zinc de tipo TRAF	El sistema inmunitario innato confiere al huésped defensa frente a infecciones virales y microbianas, y TRAFD1 es un regulador de	141

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				retroalimentación negativa que controla las respuestas inmunitarias excesivas (Sanada et al., 2008 [PubMed 18849341]). [suministrado por OMIM, diciembre de 2009]. Variante de transcripción: Esta variante (1) representa la transcripción más larga. Las variantes 1 y 2 codifican ambas la misma proteína. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AK122620.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025083, ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	
TRPM2	exónico	7226	Canal de cationes potenciales receptores transitorios Subfamilia M Miembro 2	La proteína codificada por este gen es un canal catiónico permeable al calcio que está regulado por la ADP-ribosa intracelular libre. La proteína codificada se activa por estrés oxidativo y confiere susceptibilidad a la muerte celular. Se han descrito varias variantes del transcripción de este gen empalmadas alternativamente, pero se desconoce su naturaleza completa. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (2) utiliza una unión de empalme alternativa en el marco en el extremo 5' de un exón en comparación con la variante 1. Esto da lugar a la introducción de un codón de parada prematuro y convierte al transcripción en un candidato a la descomposición del ARNm mediada no sentido (NMD). Por lo tanto, no se cree que esta transcripción codifique proteínas. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025082 [ECO:0000350] ##FIN de los Datos de Evidencia##	142
TTLL10	exónico	254173	poliglicilasa inactiva TTLL10 isoforma 1	N/A	143
TTYH2	exónico	94015	proteína tweety homóloga 2 isoforma 1	Este gen codifica un miembro de la familia de proteínas tweety. Los miembros de esta familia funcionan como canales aniónicos de cloruro.	144

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				La proteína codificada funciona como un canal de cloruro(-) de gran conductancia activado por calcio(2+), y puede desempeñar un papel en la tumorigénesis renal. Para este gen se han identificado dos variantes de transcripción que codifican isoformas distintas. [proporcionado por RefSeq, julio de 2008]. Variante de transcripción: Esta variante (1) representa la transcripción más larga y codifica la isoforma más larga (1). ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF319952.1, BC107492.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025082, ERS025083 [ECO:0000348] ##FIN de los Datos de Evidencia##	
UBE2N	exónico	7334	enzima conjugadora de ubiquitina E2 N	La modificación de las proteínas con ubiquitina es un importante mecanismo celular para destinar a la degradación las proteínas anormales o de vida corta. En la ubiquitinación intervienen al menos tres clases de enzimas: las enzimas activadoras de ubiquitina, o E1, las enzimas conjugadoras de ubiquitina, o E2, y las ubiquitina-proteína ligasas, o E3. Este gen codifica un miembro de la familia de enzimas conjugadoras de ubiquitina E2. Estudios realizados en ratones sugieren que esta proteína desempeña un papel en la reparación del ADN tras la replicación. [proporcionado por RefSeq, julio de 2008]. Nota de publicación: Este registro de RefSeq incluye un subconjunto de las publicaciones disponibles para este gen. Por favor, consulte el registro de Gene para acceder a otras publicaciones. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: BC000396.2, D83004.1 [ECO:0000332] Intrones de la secuencia de ARN :: una sola muestra respalda todos los intrones ERS025084 [ECO:0000348] ##FIN de los Datos de Evidencia##	145
VCX	exónico	26609	proteína 1 ligada al cromosoma X de carga variable	Este gen es un miembro de la familia de genes VCX/Y, que tiene múltiples miembros tanto en el cromosoma X como en el Y, y todos se expresan exclusivamente en las	146

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				células germinales masculinas. Los miembros ligados al cromosoma X se agrupan en el cromosoma Xp22 y los miembros ligados al cromosoma Y son dos copias idénticas del gen dentro de una región palindrómica en Y q 11. Los miembros de la familia comparten un alto grado de identidad de secuencia, con la excepción de que una unidad de 30 pb se repite en tándem en los miembros ligados al cromosoma X, pero	
				sólo se produce una vez en los miembros ligados a Y. El grupo de genes VCX es polimórfico en términos de número de copias; diferentes individuos pueden tener un número diferente de genes VCX. Los genes VCX/Y codifican proteínas pequeñas y muy cargadas de función desconocida. La presencia de una señal de localización nuclear bipartita putativa sugiere que los miembros VCX/Y son proteínas nucleares. Este gen contiene 10 repeticiones de la unidad de 30 pb. [proporcionado por RefSeq, julio de 2008]. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AF167081.2 [ECO:0000332] ##FIN de los Datos de Evidencia##	
VSTM1	intrónico	284415	Precursor de la proteína 1 que contiene dominios del V-set y transmembrana	N/A	147
VWA2	exónico	340706	precursor de la proteína 2 que contiene el dominio A del factor de von Willebrand	Este gen codifica un miembro de la superfamilia de proteínas del dominio similar al factor A de von Willebrand. La proteína codificada se localiza en la matriz extracelular y puede servir como componente estructural en las membranas basales o en estructuras de anclaje en andamios de colágeno VII o fibrilina. Este gen se ha relacionado con la diabetes de tipo 1A y es un marcador serológico candidato para el cáncer de colon. [proporcionado por RefSeq, enero de 2013]. Nota de secuencia: Este registro de RefSeq se creó a partir de datos de transcripción y secuencia genómica para que la secuencia fuera coherente con el ensamblaje del genoma de referencia. Las coordenadas genómicas utilizadas	148

Símbolo del gen de RefSeq	Solapamiento de exones	ID del gen de NCBI	Descripción del gen	Resumen de RefSeq	Gen # (GN)
				para el registro de transcripciones se basaron en alineaciones de transcripciones. Nota del CCDS: La región codificante se ha actualizado para representar un patrón de empalme 3' alternativo que está más respaldado por los datos disponibles sobre transcripción y proteínas. ##INICIO de los Datos de Evidencia## Combinación de exón de transcripción :: AY572972.1, AJ536328.2 [ECO:0000332] Intrones de la secuencia de ARN :: soporte de muestra mixto/parcial ERS025081, ERS025084 [ECO:0000350] ##FIN de los Datos de Evidencia###	
ZNF350	exónico	59348	proteína dedo de zinc 350	N/A	149
ZNF432	exónico	9668	proteína dedo de zinc 432	N/A	150
ZNF577	exónico	84765	N/A	N/A	151
ZNF613	exónico	79898	proteína dedo de zinc 613 isoforma 1	N/A	152
ZNF614	exónico	80110	proteína dedo de zinc 614	N/A	153
ZNF615	exónico	284370	proteína dedo de zinc 615 isoforma 1	N/A	154
ZNF649	exónico	65251	proteína dedo de zinc 649	N/A	155
ZNF841	exónico	284371	proteína dedo de zinc 841	N/A	156

Para todos los genes listados en la Tabla 2 (es decir, aquellos relevantes para las subregiones de CNV de interés), la Tabla 3 representa una lista no redundante.

Tabla 4: Una lista no redundante de variantes de transcripción que corresponden a los genes de la Tabla 3

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
MIR200B	exónico	NR_029639	microARN 200b (MIR200B) de Homo sapiens, microARN.	173
MIR200A	exónico	NR_029834	microARN 200a (MIR200A) de Homo sapiens, microARN.	174
MIR429	exónico	NR_029957	microARN 429 (MIR429) de Homo sapiens, microARN.	175
TTLL10	exónico	NM_001130045	Familia similar a la tirosina ligasa de tubulina de Homo sapiens, miembro 10 (TTLL10), variante de transcripción 1, ARNm.	176
TTLL10	exónico	NM_153254	Familia similar a la tirosina ligasa de tubulina de Homo sapiens, miembro 10 (TTLL10), variante de transcripción 2, ARNm.	177

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
TNFRSF18	exónico	NM_004195	Superfamilia de receptores del factor de necrosis tumoral de Homo sapiens, miembro 18 (TNFRSF18), variante de transcripción 1, ARNm.	178
TNFRSF18	exónico	NM_148901	Superfamilia de receptores del factor de necrosis tumoral de Homo sapiens, miembro 18 (TNFRSF18), variante de transcripción 2, ARNm.	179
TNFRSF18	exónico	NM_148902	Superfamilia de receptores del factor de necrosis tumoral de Homo sapiens, miembro 18 (TNFRSF18), variante de transcripción 3, ARNm.	180
PIK3CD	exónico	NM_005026	Fosfatidilinositol-4,5-bisfosfato 3-quinasa de Homo sapiens, subunidad catalítica delta (PIK3CD), ARNm.	181
CAPZB	intrónico	NR_038125	Proteína protectora de Homo sapiens (filamento de actina) línea Z de músculo, beta (CAPZB), variante de transcripción 4, ARN no codificante.	182
CAPZB	intrónico	NM_001206540	Proteína protectora de Homo sapiens (filamento de actina) línea Z de músculo, beta (CAPZB), variante de transcripción 2, ARNm.	183
CAPZB	intrónico	NM_004930	Proteína protectora de Homo sapiens (filamento de actina) línea Z de músculo, beta (CAPZB), variante de transcripción 1, ARNm.	184
IFNLR1	exónico	NM_170743	Interferón de Homo sapiens, receptor lambda 1 (IFNLR1), variante de transcripción 1, ARNm.	185
IFNLR1	exónico	NM_173064	Interferón de Homo sapiens, receptor lambda 1 (IFNLR1), variante de transcripción 2, ARNm.	186
IFNLR1	exónico	NM_173065	Interferón de Homo sapiens, receptor lambda 1 (IFNLR1), variante de transcripción 3, ARNm.	187
PHACTR4	exónico	NM_001048183	Fosfatasa de Homo sapiens y regulador de actina 4 (PHACTR4), variante de transcripción 1, ARNm.	188
PHACTR4	exónico	NM_023923	Fosfatasa de Homo sapiens y regulador de actina 4 (PHACTR4), variante de transcripción 2, ARNm.	189
SNHG3	exónico	NR_002909	Gen 3 huésped de ARN nucleolar pequeño de Homo sapiens (no codificador de proteína) (SNHG3), variante de transcripción 2, ARN no codificante.	190
SNHG3	exónico	NR_036473	Gen 3 huésped de ARN nucleolar pequeño de Homo sapiens (no codificador de proteína) (SNHG3), variante de transcripción 1, ARN no codificante.	191
CCR1	exónico	NM_001048199	Regulador de condensación 1 del cromosoma de Homo sapiens (RCC1), variante de transcripción 4, ARNm.	192
CCR1	exónico	NR_030725	Regulador de condensación 1 del cromosoma de Homo sapiens (RCC1), variante de transcripción 5, ARN no codificante.	193
CCR1	exónico	NR_030726	Regulador de condensación 1 del cromosoma de Homo sapiens (RCC1), variante de transcripción 6, ARN no codificante.	194
CCR1	exónico	NM_001048194	Regulador de condensación 1 del cromosoma de Homo sapiens (RCC1), variante de transcripción 1, ARNm.	195

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
CCR1	exónico	NM_001048195	Regulador de condensación 1 del cromosoma de Homo sapiens (RCC1), variante de transcripción 2, ARNm.	196
CCR1	exónico	NM_001269	Regulador de condensación 1 del cromosoma de Homo sapiens (RCC1), variante de transcripción 3, ARNm.	197
AGBL4	intrónico	NM_032785	Proteína tipo 4 de unión a ATP/GTP de Homo sapiens (AGBL4), ARNm.	198
GPATCH2	intrónico	NM_018040	Contiene dos dominios parche G de Homo sapiens (GPATCH2), ARNm.	199
RHOQ	intrónico	NM_012249	Miembro Q de la familia homóloga ras de Homo sapiens (RHOQ), ARNm.	200
PNPT1	ambos	NM_033109	Polirribonucleótido nucleotidiltransferasa 1 (PNPT1) de Homo sapiens, ARNm.	201
MCEE	ambos	NM_032601	Metilmalonil CoA epimerasa (MCEE) de Homo sapiens, ARNm.	202
HK2	exónico	NM_000189	Hexoquinasa 2 (HK2) de Homo sapiens, ARNm.	203
FHL2	exónico	NM_201557	Cuatro y medio dominios LIM 2 (FHL2) de Homo sapiens, variante de transcripción 4, ARNm.	204
FHL2	intrónico	NM_001039492	Cuatro y medio dominios LIM 2 (FHL2) de Homo sapiens, variante de transcripción 5, ARNm.	205
FHL2	intrónico	NM_001450	Cuatro y medio dominios LIM 2 (FHL2) de Homo sapiens, variante de transcripción 1, ARNm.	206
FHL2	intrónico	NM_201555	Cuatro y medio dominios LIM 2 (FHL2) de Homo sapiens, variante de transcripción 2, ARNm.	207
MIR4267	exónico	NR_036225	microARN 4267 (MIR4267) de Homo sapiens, microARN.	208
MALL	exónico	NM_005434	Similar a proteína de diferenciación de células T de Homo sapiens mal (MALL), ARNm.	209
MIR4436B1	exónico	NR_039941	microARN 4436b-1 de Homo sapiens (MIR4436B1), microARN.	210
MIR4436B2	exónico	NR_049830	microARN 4436b-2 de Homo sapiens (MIR4436B2), microARN.	211
MGAT5	intrónico	NM_002410	Mannosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa (MGATS) de Homo sapiens, ARNm.	212
BMPR2	intrónico	NM_001204	Receptor de la proteína morfogenética ósea de Homo sapiens, tipo II (serina/treonina quinasa) (BMPR2), ARNm.	213
DNER	intrónico	NM_139072	Contiene la repetición de EGF tipo delta/notch de Homo sapiens (DNER), ARNm.	214
IQCB1	exónico	NM_001023570	Motivo IQ que contiene B1 (IOCB1) de Homo sapiens, variante de transcripción 1, ARNm.	215
IQCB1	exónico	NM_001023571	Motivo IQ que contiene B1 (IQCB1) de Homo sapiens, variante de transcripción 3, ARNm.	216
STIM2	intrónico	NM_001169117	Molécula de interacción estromal 2 (STIM2) de Homo sapiens, variante de transcripción 3, ARNm.	217
STIM2	intrónico	NM_001169118	Molécula de interacción estromal 2 (STIM2) de Homo sapiens, variante de transcripción 1, ARNm.	218

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
STIM2	intrónico	NM_020860	Molécula de interacción estromal 2 (STIM2) de Homo sapiens variante de transcripción, ARNm.	219
PDGFRA	exónico	NM_006206	Receptor del factor de crecimiento derivado de plaquetas de Homo sapiens, polipéptido alfa (PDGFRA), ARNm.	220
HCN1	exónico	NM_021072	Canal 1 de potasio activado por nucleótido cíclico activado por hiperpolarización (HCN1) de Homo sapiens, ARNm.	221
MSB	exónico	NM_198449	Embigina (EMB) de Homo sapiens, ARNm.	222
EDIL3	intrónico	NM_005711	Repeticiones similares a EGF y 3 dominios similares a discoidina 1 (EDIL3) de Homo sapiens, variante de transcripción 1, ARNm.	223
MAPK9	exónico	NM_002752	Proteína quinasa 9 activada por mitógenos (MAPK9) de Homo sapiens, variante de transcripción JNK2-a2, ARNm.	224
MAPK9	exónico	NM_139068	Proteína quinasa 9 activada por mitógenos (MAPK9) de Homo sapiens, variante de transcripción JNK2-a1, ARNm.	225
MAPK9	exónico	NM_139069	Proteína quinasa 9 activada por mitógenos (MAPK9) de Homo sapiens, variante de transcripción JNK2-b1, ARNm.	226
MAPK9	exónico	NM_139070	Proteína quinasa 9 activada por mitógenos (MAPK9) de Homo sapiens, variante de transcripción JNK2-b2, ARNm.	227
MAPK9	exónico	NM_001135044	Proteína quinasa 9 activada por mitógenos (MAPK9) de Homo sapiens, variante de transcripción JNK2-g, ARNm.	228
SERPINB6	exónico	NM_001271825	Inhibidor de serpina peptidasa de Homo sapiens, clado B (ovoalbúmina), miembro 6 (SERPINB6), variante de transcripción 6, ARNm.	229
SERPINB6	exónico	NM_001271823	Inhibidor de serpina peptidasa de Homo sapiens, clado B (ovoalbúmina), miembro 6 (SERPINB6), variante de transcripción 4, ARNm.	230
SERPINB6	exónico	NM_001271822	Inhibidor de serpina peptidasa de Homo sapiens, clado B (ovoalbúmina), miembro 6 (SERPINB6), variante de transcripción 3, ARNm.	231
SERPINB6	exónico	NM_001195291	Inhibidor de serpina peptidasa de Homo sapiens, clado B (ovoalbúmina), miembro 6 (SERPINB6), variante de transcripción 2, ARNm.	232
SERPINB6	exónico	NM_001271824	Inhibidor de serpina peptidasa de Homo sapiens, clado B (ovoalbúmina), miembro 6 (SERPINB6), variante de transcripción 5, ARNm.	233
SERPINB6	exónico	NM_004568	Inhibidor de serpina peptidasa de Homo sapiens, clado B (ovoalbúmina), miembro 6 (SERPINB6), variante de transcripción 1, ARNm.	234
DKFZP686I15217	exónico	NR_026855	ARN 1011 no codificante de proteínas intergénico largo (LINC01011) de Homo sapiens, variante de transcripción 1, ARN no codificante.	235
DKFZP686I15217	exónico	NR_026856	ARN 1011 no codificante de proteínas intergénico largo (LINC01011) de Homo sapiens, variante de transcripción 2, ARN no codificante.	236

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
NQO2	exónico	NM_000904	NAD(P)H deshidrogenasa de Homo sapiens, quinona 2 (NQO2). ARNm.	237
HTATSF1P2	exónico	NR_033884	Pseudogén 2 del factor 1 específico de Tat de VIH-1 (HTATSF1P2) de Homo sapiens, ARN no codificante.	238
PKHD1	intrónico	NM_138694	Enfermedad 1 renal y hepática poliquística (autosómica recesiva) (PKHD1) de Homo sapiens, variante de transcripción 1, ARNm.	239
PKHD1	intrónico	NM_170724	Enfermedad 1 renal y hepática poliquística (autosómica recesiva) (PKHD1) de Homo sapiens, variante de transcripción 2, ARNm.	240
SLC17A5	ambos	NM_012434	Familia 17 transportadora de soluto (transportadora de azúcares ácidos), miembro 5 (SLC17A5), ARNm.	241
PDSS2	exónico	NM_020381	Prenil (decaprenil) difosfato sintasa, subunidad 2 (PDSS2) de Homo sapiens, ARNm.	242
KCTD7	exónico	NM_001167961	7 que contiene dominios de tetramerización del canal de potasio (KCTD7) de Homo sapiens, variante de transcripción 2, ARNm.	243
KCTD7	exónico	NM_153033	7 que contiene dominios de tetramerización del canal de potasio (KCTD7) de Homo sapiens, variante de transcripción 1, ARNm.	244
PTPRN2	intrónico	NM_002847	Proteína tirosina fosfatasa, tipo receptor, polipéptido N 2 (PTPRN2) de Homo sapiens, variante de transcripción 1, ARNm.	245
PTPRN2	intrónico	NM_130842	Proteína tirosina fosfatasa, tipo receptor, polipéptido N 2 (PTPRN2) de Homo sapiens, variante de transcripción 2, ARNm.	246
PTPRN2	intrónico	NM_130843	Proteína tirosina fosfatasa, tipo receptor, polipéptido N 2 (PTPRN2) de Homo sapiens, variante de transcripción 3, ARNm.	247
MIR595	exónico	NR_030325	microARN 595 (MIR595) de Homo sapiens, microARN.	248
TNFRSF10A	exónico	NM_003844	Superfamilia de receptores del factor de necrosis tumoral, miembro 10a (TNFRSF10A) de Homo sapiens, ARNm.	249
IDO2	intrónico	NM_194294	Indolamina 2,3-dioxigenasa 2 (IDO2) de Homo sapiens, ARNm.	250
STK3	intrónico	NM_001256313	Serina/treonina quinasa 3 (STK3) de Homo sapiens, variante de transcripción 3, ARNm.	251
STK3	intrónico	NM_006281	Serina/treonina quinasa 3 (STK3) de Homo sapiens, variante de transcripción 1, ARNm.	252
STK3	intrónico	NM_001256312	Serina/treonina quinasa 3 (STK3) de Homo sapiens, variante de transcripción 2, ARNm.	253
KANK1	intrónico	NM_001256876	Dominios 1 del motivo KN y repetición de anquirina (KANK1) de Homo sapiens, variante de transcripción 3, ARNm.	254
KANK1	intrónico	NM_001256877	Dominios 1 del motivo KN y repetición de anquirina (KANK1) de Homo sapiens, variante de transcripción 4, ARNm.	255

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
KANK1	intrónico	NM_015158	Dominios 1 del motivo KN y repetición de anquirina (KANK1) de Homo sapiens, variante de transcripción 1, ARNm.	256
KANK1	intrónico	NM_153186	Dominios 1 del motivo KN y repetición de anquirina (KANK1) de Homo sapiens, variante de transcripción 2, ARNm.	257
GDA	exónico	NM_001242507	Guanina desaminasa (GDA) de Homo sapiens, variante de transcripción 4, ARNm.	258
GDA	exónico	NM_001242505	Guanina desaminasa (GDA) de Homo sapiens, variante de transcripción 1, ARNm.	259
GDA	exónico	NM_001242506	Guanina desaminasa (GDA) de Homo sapiens, variante de transcripción 3, ARNm.	260
GDA	exónico	NM_004293	Guanina desaminasa (GDA) de Homo sapiens, variante de transcripción 2, ARNm.	261
AUH	exónico	NM_001698	Proteína/enoil-CoA hidratasa (AUH) de unión a ARN AU de Homo sapiens, ARNm.	262
MIR3163	exónico	NR_036121	microARN 3163 (MIR3163) de Homo sapiens, microARN.	263
NFIL3	exónico	NM_005384	Factor nuclear de Homo sapiens, regulado por interleuquina 3 (NFIL3), ARNm.	264
MIR3910-1	exónico	NR_037472	microARN 3910-1 (MIR3910-1) de Homo sapiens, microARN.	265
MIR3910-2	exónico	NR_037489	microARN 3910-2 (MIR3910-2)de Homo sapiens, microARN.	266
ASTN2	intrónico	NM_014010	Astrotactina 2 (ASTN2) de Homo sapiens, variante de transcripción 1, ARNm.	267
ASTN2	intrónico	NM_198186	Astrotactina 2 (ASTN2) de Homo sapiens, variante de transcripción 2, ARNm.	268
ASTN2	intrónico	NM_001184734	Astrotactina 2 (ASTN2) de Homo sapiens, variante de transcripción 5, ARNm.	269
ASTN2	intrónico	NM_198187	Astrotactina 2 (ASTN2) de Homo sapiens, variante de transcripción 3, ARNm.	270
ASTN2	intrónico	NM_198188	Astrotactina 2 (ASTN2) de Homo sapiens, variante de transcripción 4, ARNm.	271
ASTN2	intrónico	NM_001184735	Astrotactina 2 (ASTN2) de Homo sapiens, variante de transcripción 6, ARNm.	272
LARP4B	exónico	NM_015155	Familia del dominio de ribonucleoproteína La, miembro 4B (LARP4B) de Homo sapiens, ARNm.	273
GTPBP4	exónico	NM_012341	Proteína 4 de unión a GTP (GTPBP4) de Homo sapiens, ARNm.	274
IDI2	exónico	NM_033261	Difosfato de isopentenilo delta isomerasa 2 (IDI2) de Homo sapiens, ARNm.	275
IDI2-AS1	exónico	NR_024628	ARN 1antisentido IDI2 (IDI2-AS1) de Homo sapiens, variante de transcripción 1, ARN no codificante.	276
IDI2-AS1	exónico	NR_024629	ARN 1antisentido IDI2 (IDI2-AS1) de Homo sapiens, variante de transcripción 2, ARN no codificante.	277
IDI2-AS1	exónico	NR_027708	ARN 1antisentido IDI2 (IDI2-AS1) de Homo sapiens, variante de transcripción 3, ARN no codificante.	278

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
IDI2-AS1	exónico	NR_027709	ARN 1antisentido IDI2 (IDI2-AS1) de Homo sapiens, variante de transcripción 4, ARN no codificante.	279
KAT6B	exónico	NM_001256468	K(lisina) acetiltransferasa 6B (KAT6B) de Homo sapiens, variante de transcripción 2, ARNm.	280
KAT6B	exónico	NM_001256469	K(lisina) acetiltransferasa 6B (KAT6B) de Homo sapiens, variante de transcripción 3, ARNm.	281
KAT6B	exónico	NM_012330	K(lisina) acetiltransferasa 6B (KAT6B) de Homo sapiens, variante de transcripción 1, ARNm.	282
VWA2	exónico	NM_001272046	2 que contiene dominios del factor A de von Willebrand (VWA2) de Homo sapiens, ARNm.	283
PDE3B	intrónico	NM_000922	Fosfodiesterasa 3B, inhibida por cGMP (PDE3B) de Homo sapiens, ARNm.	284
EHF	intrónico	NM_001206615	Factor homólogo ets (EHF) de Homo sapiens, variante de transcripción 3, ARNm.	285
EHF	intrónico	NM_012153	Factor homólogo ets (EHF) de Homo sapiens, variante de transcripción 2, ARNm.	286
EHF	exónico	NM_001206616	Factor homólogo ets (EHF) de Homo sapiens, variante de transcripción 1, ARNm.	287
SLC3A2	exónico	NM_001012662	Familia 3 del transportador de soluto (cadena pesada transportadora de aminoácidos), miembro 2 (SLC3A2) de Homo sapiens, variante de transcripción 2, ARNm.	288
SLC3A2	intrónico	NM_001012664	Familia 3 del transportador de soluto (cadena pesada transportadora de aminoácidos), miembro 2 (SLC3A2) de Homo sapiens, variante de transcripción 5, ARNm.	289
SLC3A2	exónico	NM_002394	Familia 3 del transportador de soluto (cadena pesada transportadora de aminoácidos), miembro 2 (SLC3A2) de Homo sapiens, variante de transcripción 3, ARNm.	290
SLC3A2	intrónico	NM_001013251	Familia 3 del transportador de soluto (cadena pesada transportadora de aminoácidos), miembro 2 (SLC3A2) de Homo sapiens, variante de transcripción 6, ARNm.	291
SLC3A2	intrónico	NR_037193	Familia 3 del transportador de soluto (cadena pesada transportadora de aminoácidos), miembro 2 (SLC3A2) de Homo sapiens, variante de transcripción 7, ARNm.	292
GDPD4	exónico	NM_182833	4 que contiene dominios de glicerofosfodiéster fosfodiesterasa (GDPD4) de Homo sapiens, ARNm.	293
ETV6	exónico	NM_001987	Variante 6 est (ETV6) de Homo sapiens, ARNm.	294
LOH12CR1	exónico	NM_058169	Pérdida de heterocigosidad, 12, región cromosómica 1 (LOH12CR1) de Homo sapiens, ARNm.	295
DUSP16	exónico	NM_030640	Fosfatasa 16 de doble especificidad (DUSP16) de Homo sapiens, ARNm.	296
CREBL2	exónico	NM_001310	Similar a la proteína 2 de unión al elemento sensible a cAMP (CREBL2) de Homo sapiens, ARNm.	297
GPR19	exónico	NM_006143	Receptor 19 acoplado a proteína G (GPR19) de Homo sapiens, ARNm.	298

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
CDKN1B	exónico	NM_004064	Inhibidor 1B de la quinasa dependiente de ciclina (p27, Kip1) (CDKN1B) de Homo sapiens, ARNm.	299
APOLD1	exónico	NM_001130415	1 que contiene dominio de apolipoproteína L (APOLD1) de Homo sapiens, variante de transcripción 1, ARNm.	300
APOLD1	intrónico	NM_030817	1 que contiene dominio de apolipoproteína L (APOLD1) de Homo sapiens, variante de transcripción 2, ARNm.	301
EEE1	exónico	NM_003566	Antígeno 1 endosomal temprano (EEA1) de Homo sapiens, ARNm.	302
LOC643339	exónico	NR_040096	LOC643339 no caracterizado (LOC643339) de Homo sapiens, ARN no codificante.	303
NUDT4	exónico	NM_019094	Motivo 4 de tipo (nucleósido difosfato enlazado a la fracción X) (NUDT4) de Homo sapiens nudix, variante de transcripción 1, ARNm.	304
NUDT4	exónico	NM_199040	Motivo 4 de tipo (nucleósido difosfato enlazado a la fracción X) (NUDT4) de Homo sapiens nudix, variante de transcripción 2, ARNm.	305
NUDT4P1	exónico	NR_002212	Pseudogén 1 del motivo 4 tipo (nucleósido difosfato enlazado a la fracción X) (NUDT4P1) de Homo sapiens nudix, ARN no codificante.	306
UBE2N	exónico	NM_003348	Enzima conjugadora de ubiquitina E2N (UBE2N) de Homo sapiens, ARNm.	307
MRPL42	exónico	NM_014050	Proteína ribosómica mitocondrial L42 (MRPL42) de Homo sapiens, variante de transcripción 1, ARNm.	308
MRPL42	exónico	NM_172177	Proteína ribosómica mitocondrial L42 (MRPL42) de Homo sapiens, variante de transcripción 2, ARNm.	309
MRPL42	exónico	NR_038159	Proteína ribosómica mitocondrial L42 (MRPL42) de Homo sapiens, variante de transcripción 3, ARN no codificante.	310
MRPL42	exónico	NR_038160	Proteína ribosómica mitocondrial L42 (MRPL42) de Homo sapiens, variante de transcripción 4, ARN no codificante.	311
MRPL42	exónico	NR_038161	Proteína ribosómica mitocondrial L42 (MRPL42) de Homo sapiens, variante de transcripción 5, ARN no codificante.	312
SOCS2-AS1	exónico	NR_038263	ARN 1 antisentido SOCS2 (SOCS2-AS1) de Homo sapiens, ARN no codificante.	313
SOCS2	exónico	NM_003877	Supresor de la señalización de citoquinas 2 (SOCS2) de Homo sapiens, variante de transcripción 1, ARNm.	314
SOCS2	exónico	NM_001270467	Supresor de la señalización de citoquinas 2 (SOCS2) de Homo sapiens, variante de transcripción 2, ARNm.	315
SOCS2	exónico	NM_001270468	Supresor de la señalización de citoquinas 2 (SOCS2) de Homo sapiens, variante de transcripción 3, ARNm.	316
SOCS2	exónico	NM_001270469	Supresor de la señalización de citoquinas 2 (SOCS2) de Homo sapiens, variante de transcripción 3, ARNm.	317

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
SOCS2	exónico	NM_001270470	Supresor de la señalización de citoquinas 2 (SOCS2) de Homo sapiens, variante de transcripción 5, ARNm.	318
SOCS2	exónico	NM_001270471	Supresor de la señalización de citoquinas 2 (SOCS2) de Homo sapiens, variante de transcripción 6, ARNm.	319
CRADD	exónico	NM_003805	Dominio CASP2 y RIPK1 que contiene adaptador con dominio de muerte (CRADD) de Homo sapiens, ARNm.	320
PLXNC1	exónico	NM_005761	Plexina C1 (PLXNC1) de Homo sapiens, variante de transcripción 1, ARNm.	321
PLXNC1	exónico	NR_037687	Plexina C1 (PLXNC1) de Homo sapiens, variante de transcripción 2, ARNm no codificante.	322
CCDC41	exónico	NM_001042399	41 que contiene un dominio de bobina enrollada (CCDC41) de Homo sapiens, variante de transcripción 2, ARNm.	323
CCDC41	exónico	NM_016122	41 que contiene un dominio de bobina enrollada (CCDC41) de Homo sapiens, variante de transcripción 1, ARNm.	324
TRAFD1	exónico	NM_001143906	1 que contiene dominio dedo de zinc tipo TRAF (TRAFD1) de Homo sapiens, variante de transcripción 1, ARNm.	325
TRAFD1	exónico	NM_006700	1 que contiene dominio dedo de zinc tipo TRAF (TRAFD1) de Homo sapiens, variante de transcripción 2, ARNm.	326
RGCC	exónico	NM_014059	Regulador del ciclo celular (RGCC) de Homo sapiens, ARNm.	327
COMD6	exónico	NM_203495	6 que contiene dominio COMM (COMMD6) de Homo sapiens, variante de transcripción 2, ARNm.	328
COMD6	exónico	NM_203497	6 que contiene dominio COMM (COMMD6) de Homo sapiens, variante de transcripción 1, ARNm.	329
GPC5	intrónico	NM_004466	Glipicano 5 (GPC5) de Homo sapiens, ARNm.	330
ARHGEF7	exónico	NM_003899	Factor 7 de intercambio de nucleótidos de Rho guanina (GEF) 7 (ARHGEF7) de Homo sapiens, variante de transcripción 1, ARNm.	331
ARHGEF7	exónico	NM_001113513	Factor 7 de intercambio de nucleótidos de Rho guanina (GEF) 7 (ARHGEF7) de Homo sapiens, variante de transcripción 5, ARNm.	332
TEX29	exónico	NM_152324	Expresado en testículo 29 (TEX29) de Homo sapiens, ARNm.	333
ARHGEF7	intrónico	NM_001113511	Factor 7 de intercambio de nucleótidos de Rho guanina (GEF) 7 (ARHGEF7) de Homo sapiens, variante de transcripción 3, ARNm.	334
ARHGEF7	intrónico	NM_001113512	Factor 7 de intercambio de nucleótidos de Rho guanina (GEF) 7 (ARHGEF7) de Homo sapiens, variante de transcripción 4, ARNm.	335
ARHGEF7	intrónico	NM_145735	Factor 7 de intercambio de nucleótidos de Rho guanina (GEF) 7 (ARHGEF7) de Homo sapiens, variante de transcripción 2, ARNm.	336

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
RNASE10	exónico	NM_001012975	Ribonucleasa, familia de RNasa A, 10 (no activa) (RNASA10) de Homo sapiens, ARNm.	337
RNASE3	exónico	NM_002935	Ribonucleasa, familia de RNasa A, 3 (no activa) (RNASA3) de Homo sapiens, ARNm.	338
ECRP	exónico	NR_033909	ribonucleasa, familia de RNasa A, 2 (hígado, neurotoxina derivada de eosinófilos) pseudogén (ECRP) de Homo sapiens, ARN no codificante.	339
PRKCH	intrónico	NM_006255	Proteína quinasa C, eta (PRKCH) de Homo sapiens, ARNm.	340
MTHFD1	exónico	NM_005956	Metilentetrahidrofolato deshidrogenasa (dependiente de NADP+) 1, meteniltetrahidrofolato ciclohidrolasa, formiltetrahidrofolato sintetasa (MTHFD1) de Homo sapiens, ARNm.	341
BDKRB2	intrónico	NM_000623	Receptor B2 de la bradiquinina (BDKRB2) de Homo sapiens, ARNm.	342
HEXA	exónico	NM_000520	Hexosaminidasa A (polipéptido alfa) (HEXA) de Homo sapiens, ARNm.	343
PSTPIP1	exónico	NM_003978	Proteína 1 que interactúa con prolina-serina-treonina-fosfatasa (PSTPIP1) de Homo sapiens, ARNm.	344
RBFOX1	intrónico	NM_001142333	Proteína de unión al ARN, homólogo de fox-1 (C. elegans) 1 (RBFOX1) de Homo sapiens, variante de transcripción 5, ARNm.	345
RBFOX1	intrónico	NM_018723	Proteína de unión al ARN, homólogo de fox-1 (C. elegans) 1 (RBFOX1) de Homo sapiens, variante de transcripción 4, ARNm.	346
RBFOX1	intrónico	NM_001142334	Proteína de unión al ARN, homólogo de fox-1 (C. elegans) 1 (RBFOX1) de Homo sapiens, variante de transcripción 6, ARNm.	347
RBFOX1	intrónico	NM_145891	Proteína de unión al ARN, homólogo de fox-1 (C. elegans) 1 (RBFOX1) de Homo sapiens, variante de transcripción 1, ARNm.	348
RBFOX1	intrónico	NM_145892	Proteína de unión al ARN, homólogo de fox-1 (C. elegans) 1 (RBFOX1) de Homo sapiens, variante de transcripción 2, ARNm.	349
RBFOX1	intrónico	NM_145893	Proteína de unión al ARN, homólogo de fox-1 (C. elegans) 1 (RBFOX1) de Homo sapiens, variante de transcripción 3, ARNm.	350
PRKCB	ambos	NM_002738	Proteína quinasa C, beta (PRKCB) de Homo sapiens, variante de transcripción 2, ARNm.	351
PRKCB	ambos	NM_212535	Proteína quinasa C, beta (PRKCB) de Homo sapiens, variante de transcripción 1, ARNm.	352
FUK	ambos	NM_145059	Fucoquinasa (FUK) de Homo sapiens, ARNm.	353
COG4	exónico	NM_001195139	Componente del complejo oligomérico de golgi 4 (COG4) de Homo sapiens, variante de transcripción 2, ARNm.	354
COG4	exónico	NM_015386	Componente del complejo oligomérico de golgi 4 (COG4) de Homo sapiens, variante de transcripción 1, ARNm.	355

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
HPR	exónico	NM_020995	Proteína relacionada con la haptoglobina (HPR) de Homo sapiens, ARNm.	356
RPL38	exónico	NM_000999	Proteína ribosomal L38 (RPL38) de Homo sapiens, variante de transcripción 1, ARNm.	357
RPL38	exónico	NM_001035258	Proteína ribosomal L38 (RPL38) de Homo sapiens, variante de transcripción 2, ARNm.	358
MGC 16275	exónico	NR_026914	Proteína MGC16275 no caracterizada (MGC16275) de Homo sapiens, ARN no codificante.	359
TTYH2	exónico	NM_032646	Miembro 2 de la familia tweety (TTYH2) de Homo sapiens, variante de transcripción 1, ARNm.	360
TTYH2	exónico	NM_052869	Miembro 2 de la familia tweety (TTYH2) de Homo sapiens, variante de transcripción 2, ARNm.	361
DNAI2	exónico	NM_001172810	Dineína, axonemal, cadena intermedia 2 (DNAI2) de Homo sapiens, variante de transcripción 2, ARNm.	362
DNAI2	exónico	NM_023036	Dineína, axonemal, cadena intermedia 2 (DNAI2) de Homo sapiens, variante de transcripción 1, ARNm.	363
KIF19	exónico	NM_153209	Miembro 19 de la familia de la kinesina (KIF19) de Homo sapiens, ARNm.	364
BTBD17	exónico	NM_001080466	17 que contiene el dominio BTB (POZ) (BTBD17) de Homo sapiens, ARNm.	365
GPR142	exónico	NM_181790	Receptor 142 acoplado a proteína G (GPR142) de Homo sapiens, ARNm.	366
GPRC5C	exónico	NM_022036	Receptor acoplado a proteína G, familia C, grupo 5, miembro C (GPRC5C) de Homo sapiens, variante de transcripción 1, ARNm.	367
GPRC5C	exónico	NM_018653	Receptor acoplado a proteína G, familia C, grupo 5, miembro C (GPRC5C) de Homo sapiens, variante de transcripción 2, ARNm.	368
CD300A	exónico	NM_001256841	Molécula CD300a (CD300A) de Homo sapiens, variante de transcripción 2, ARNm.	369
CD300A	exónico	NM_007261	Molécula CD300a (CD300A) de Homo sapiens, variante de transcripción 1, ARNm.	370
CD300LB	exónico	NM_174892	Miembro b de la familia similar a la molécula CD300 (CD300LB) de Homo sapiens, ARNm.	371
CD300C	exónico	NM_006678	Molécula CD300c (CD300C) de Homo sapiens, ARNm.	372
CD300LD	exónico	NM_001115152	Miembro d de la familia similar a la molécula CD300 (CD300LD) de Homo sapiens, ARNm.	373
C17orf77	exónico	NM_152460	Marco de lectura abierto 77 del cromosoma 17 (C17orf77) de Homo sapiens, ARNm.	374
CD300E	exónico	NM_181449	Molécula CD300e (CD300E) de Homo sapiens, ARNm.	375
RAB37	exónico	NM_175738	RAB37, miembro de la familia del oncogén RAS (RAB37) de Homo sapiens, variante de transcripción 3, ARNm.	376
CD300LF	exónico	NM_139018	Miembro f de la familia similar a la molécula CD300 (CD300LF) de Homo sapiens, ARNm.	377

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
RAB37	intrónico	NM_001163989	RAB37, miembro de la familia del oncogén RAS (RAB37) de Homo sapiens, variante de transcripción 4, ARNm.	378
RAB37	intrónico	NM_001006638	RAB37, miembro de la familia del oncogén RAS (RAB37) de Homo sapiens, variante de transcripción 2, ARNm.	379
RAB37	intrónico	NM_001163990	RAB37, miembro de la familia del oncogén RAS (RAB37) de Homo sapiens, variante de transcripción 5, ARNm.	380
TBC1D16	intrónico	NM_019020	Familia del dominio de TBC1, miembro 16 (TBC1D16) de Homo sapiens, variante de transcripción 1, ARNm.	381
TBC1D16	intrónico	NM_001271844	Familia del dominio de TBC1, miembro 16 (TBC1D16) de Homo sapiens, variante de transcripción 2, ARNm.	382
TBC1D16	intrónico	NM_001271845	Familia del dominio de TBC1, miembro 16 (TBC1D16) de Homo sapiens, variante de transcripción 3, ARNm.	383
TBC1D16	intrónico	NM_001271846	Familia del dominio de TBC1, miembro 16 (TBC1D16) de Homo sapiens, variante de transcripción 4, ARNm.	384
RPTOR	intrónico	NM_001163034	Proteína reguladora asociada de MTOR, complejo 1 (RPTOR) de Homo sapiens, variante de transcripción 2, ARNm.	385
RPTOR	intrónico	NM_020761	Proteína reguladora asociada de MTOR, complejo 1 (RPTOR) de Homo sapiens, variante de transcripción 1, ARNm.	386
PTPN2	exónico	NM_001207013	Proteína tirosina fosfatasa, no receptora tipo 2 (PTPN2) de Homo sapiens, variante de transcripción 4, ARNm.	387
PTPN2	exónico	NM_080422	Proteína tirosina fosfatasa, no receptora tipo 2 (PTPN2) de Homo sapiens, variante de transcripción 2, ARNm.	388
PTPN2	exónico	NM_080423	Proteína tirosina fosfatasa, no receptora tipo 2 (PTPN2) de Homo sapiens, variante de transcripción 3, ARNm.	389
PTPN2	intrónico	NM_002828	Proteína tirosina fosfatasa, no receptora tipo 2 (PTPN2) de Homo sapiens, variante de transcripción 1, ARNm.	390
ST8SIA5	exónico	NM_013305	ST8 alfa-N-acetil-neuraminida alfa-2,8-sialiltrans-ferasa 5 (ST8SIA5) de Homo sapiens, ARNm.	391
PIAS2	exónico	NM_004671	Inhibidor proteico de STAT activado, 2 (PIAS2) de Homo sapiens, variante de transcripción beta, ARNm.	392
PIAS2	exónico	NM_173206	Inhibidor proteico de STAT activado, 2 (PIAS2) de Homo sapiens, variante de transcripción alfa, ARNm.	393
SERPINB4	exónico	NM_002974	Inhibidor de la serpina peptidasa, clado B (ovoalbúmina), miembro 4 (SERPINB4) de Homo sapiens, ARNm.	394
FLJ26850	intrónico	NR_027257	Proteína FLJ26850 (FLJ26850) de Homo sapiens, ARN no codificante.	395

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
FPR2	exónico	NM_001005738	Receptor 2 del péptido formilo (FPR2) de Homo sapiens, variante de transcripción 2, ARNm.	396
FPR2	exónico	NM_001462	Receptor 2 del péptido formilo (FPR2) de Homo sapiens, variante de transcripción 1, ARNm.	397
FPR3	exónico	NM_002030	Receptor 3 del péptido formilo (FPR3) de Homo sapiens, ARNm.	398
ZNF577	exónico	NR_024181	Proteína 577 dedo de zinc (ZNF577) de Homo sapiens, variante de transcripción 3, ARN no codificante.	399
ZNF577	exónico	NM_001135590	Proteína 577 dedo de zinc (ZNF577) de Homo sapiens, variante de transcripción 2, ARN no codificante.	400
ZNF577	exónico	NM_032679	Proteína 577 dedo de zinc (ZNF577) de Homo sapiens, variante de transcripción 1, ARN no codificante.	401
ZNF649	exónico	NM_023074	Proteína 649 dedo de zinc (ZNF649) de Homo sapiens, ARNm.	402
ZNF613	exónico	NM_001031721	Proteína 613 dedo de zinc (ZNF613) de Homo sapiens, variante de transcripción 1, ARNm.	403
ZNF613	exónico	NM_024840	Proteína 613 dedo de zinc (ZNF613) de Homo sapiens, variante de transcripción 2, ARNm.	404
ZNF350	exónico	NM_021632	Proteína 350 dedo de zinc (ZNF350) de Homo sapiens, ARNm.	405
ZNF615	exónico	NM_001199324	Proteína 615 dedo de zinc (ZNF615) de Homo sapiens, variante de transcripción 1, ARNm.	406
ZNF615	exónico	NM_198480	Proteína 615 dedo de zinc (ZNF615) de Homo sapiens, variante de transcripción 2, ARNm.	407
ZNF614	exónico	NM_025040	Proteína 614 dedo de zinc (ZNF614) de Homo sapiens, ARNm.	408
ZNF432	exónico	NM_014650	Proteína 432 dedo de zinc (ZNF432) de Homo sapiens, ARNm.	409
ZNF841	exónico	NM_001136499	Proteína 841 dedo de zinc (ZNF841) de Homo sapiens, ARNm.	410
NLRP12	exónico	NM_001277126	Familia de NLR, 12 que contiene el dominio de pirina (NLRP12) de Homo sapiens, variante de transcripción 3, ARNm.	411
NLRP12	exónico	NM_001277129	Familia NLR, 12 que contiene el dominio de pirina (NLRP12) de Homo sapiens, variante de transcripción 4, ARNm.	412
NLRP12	exónico	NM_144687	Familia NLR, 12 que contiene el dominio de pirina (NLRP12) de Homo sapiens, variante de transcripción 2, ARNm.	413
VSTM1	intrónico	NM_198481	1 que contiene el dominio de V-set y transmembrana (VSTM1) de Homo sapiens, ARNm.	414
SNX5	exónico	NM_014426	Nexina clasificadora 5 (SNX5) de Homo sapiens, variante de transcripción 2, ARNm.	415
SNX5	exónico	NM_152227	Nexina clasificadora 5 (SNX5) de Homo sapiens, variante de transcripción 1, ARNm.	416

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
SNORD17	exónico	NR_003045	ARN nucleolar pequeño, C/D caja 17 (SNORD17) de Homo sapiens, ARN nucleolar pequeño.	417
MGME1	exónico	NM_052865	Exonucleasa 1 de mantenimiento del genoma mitocondrial (MGME1) de Homo sapiens, ARNm.	418
OVOL2	exónico	NM_021220	Similar a ovo 2 (Drosophila) (OVOL2) de Homo sapiens, ARNm.	419
ADA	intrónico	NM_000022	Adenosina desaminasa (ADA) de Homo sapiens, ARNm.	420
NRIP1	exónico	NM_003489	Proteína 1 de interacción con receptores nucleares (NRIP1) de Homo sapiens, ARNm.	421
BACH1	exónico	NR_027655	BTB y CNC de homología 1, factor 1 de transcripción de cremallera de leucina básica (BACH1) de Homo sapiens, variante de transcripción 3, ARN no codificante.	422
BACH1	intrónico	NM_001186	BTB y CNC de homología 1, factor 1 de transcripción de cremallera de leucina básica (BACH1) de Homo sapiens, variante de transcripción 2, ARN no codificante.	423
BACH1	intrónico	NM_206866	BTB y CNC de homología 1, factor 1 de transcripción de cremallera de leucina básica (BACH1) de Homo sapiens, variante de transcripción 1, ARN no codificante.	424
TRPM2	exónico	NM_003307	Canal catiónico de potencial receptor transitorio, subfamilia M, miembro 2 (TRPM2) de Homo sapiens, variante de transcripción 1, ARNm.	425
TRPM2	exónico	NR_038257	Canal catiónico de potencial receptor transitorio, subfamilia M, miembro 2 (TRPM2) de Homo sapiens, variante de transcripción 2, ARN no codificante.	426
ADARB1	intrónico	NM_001112	Adenosina desaminasa, específica de ARN, B1 (ADARB1) de Homo sapiens, variante de transcripción 1, ARNm.	427
ADARB1	intrónico	NM_001160230	Adenosina desaminasa, específica de ARN, B1 (ADARB1) de Homo sapiens, variante de transcripción 7, ARNm.	428
ADARB1	intrónico	NM_015833	Adenosina desaminasa, específica de ARN, B1 (ADARB1) de Homo sapiens, variante de transcripción 2, ARNm.	429
ADARB1	intrónico	NM_015834	Adenosina desaminasa, específica de ARN, B1 (ADARB1) de Homo sapiens, variante de transcripción 3, ARNm.	430
ADARB1	intrónico	NR_027672	Adenosina desaminasa, específica de ARN, B1 (ADARB1) de Homo sapiens, variante de transcripción 5, ARNm.	431
ADARB1	intrónico	NR_027673	Adenosina desaminasa, específica de ARN, B1 (ADARB1) de Homo sapiens, variante de transcripción 4, ARNm.	432
ADARB1	intrónico	NR_027674	Adenosina desaminasa, específica de ARN, B1 (ADARB1) de Homo sapiens, variante de transcripción 6, ARNm.	433

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
ADARB1	intrónico	NR_073200	Adenosina desaminasa, específica de ARN, B1 (ADARB1) de Homo sapiens, variante de transcripción 8, ARNm.	434
APOBEC3A	exónico	NM_001270406	Enzima editora del ARNm de la apolipoproteína B, similar al polipéptido catalítico 3A (APOBEC3A) de Homo sapiens, variante de transcripción 3, ARNm.	435
APOBEC3A	exónico	NM_145699	Enzima editora del ARNm de la apolipoproteína B, similar al polipéptido catalítico 3A (APOBEC3A) de Homo sapiens, variante de transcripción 1, ARNm.	436
APOBEC3A B	intrónico	NM_001193289	Híbrido de eliminación APOBEC3A y APOBEC3B (APOBEC3A B) de Homo sapiens, ARNm.	437
APOBEC3B	exónico	NM_001270411	Enzima editora del ARNm de la apolipoproteína B, similar al polipéptido catalítico 3B (APOBEC3B) de Homo sapiens, variante de transcripción 2, ARNm.	438
APOBEC3B	exónico	NM_004900	Enzima editora del ARNm de la apolipoproteína B, similar al polipéptido catalítico 3B (APOBEC3B) de Homo sapiens, variante de transcripción 1, ARNm.	439
MKL1	intrónico	NM_020831	Leucemia megacarioblástica (translocación) 1 (MKL1) de Homo sapiens, ARNm.	440
TNFRSF13C	exónico	NM_052945	Superfamilia de receptores del factor de necrosis tumoral, miembro 13C (TNFRSF13C) de Homo sapiens, ARNm.	441
CENPM	exónico	NM_001110215	Proteína M del centrómero (CENPM) de Homo sapiens, variante de transcripción 3, ARNm.	442
CENPM	exónico	NM_001002876	Proteína M del centrómero (CENPM) de Homo sapiens, variante de transcripción 2, ARNm.	443
CENPM	exónico	NM_024053	Proteína M del centrómero (CENPM) de Homo sapiens, variante de transcripción 1, ARNm.	444
PPP2R3B	intrónico	NM_013239	Proteína fosfatasa 2, subunidad reguladora B'', beta (PPP2R3B) de Homo sapiens, ARNm.	445
VCX	exónico	NM_013452	carga variable, ligada al cromosoma X (VCX) de Homo sapiens, ARNm.	446
PNPLA4	exónico	NM_004650	4 que contiene el dominio de fosfolipasa similar a patatina (PNPLA4) de Homo sapiens, variante de transcripción 1, ARNm.	447
PNPLA4	exónico	NM_001142389	4 que contiene el dominio de fosfolipasa similar a patatina (PNPLA4) de Homo sapiens, variante de transcripción 2, ARNm.	448
PNPLA4	exónico	NM_001172672	4 que contiene el dominio de fosfolipasa similar a patatina (PNPLA4) de Homo sapiens, variante de transcripción 3, ARNm.	449
MIR651	exónico	NR_030380	microARN 651 (MIR651) de Homo sapiens, microARN.	450
JPX	intrónico	NR_024582	Transcripción JPX, activador XIST (no codificante de proteínas) (JPX) de Homo sapiens, ARN no codificante.	451
GRIA3	intrónico	NM_000828	Receptor de glutamato, ionotrópico, AMPA 3 (GRIA3) de Homo sapiens, variante de transcripción 2, ARNm.	452

Símbolo del gen de RefSeq	Solapamiento de exones	Número de acceso RefSeq	Descripción del ARNm	SEQ ID
GRIA3	intrónico	NM_007325	Receptor de glutamato, ionotrópico, AMPA 3 (GRIA3) de Homo sapiens, variante de transcripción 1, ARNm.	453
GRIA3	intrónico	NM_001256743	Receptor de glutamato, ionotrópico, AMPA 3 (GRIA3) de Homo sapiens, variante de transcripción 3, ARNm.	454
HMGB3	exónico	NM_005342	Caja 3 del grupo de alta movilidad (HMGB3) de Homo sapiens, ARNm.	455

Para todos los genes listados en la Tabla 2 (es decir, aquellos relevantes para las subregiones de CNV de interés), la Tabla 4 representa una lista no redundante.

Tabla 5: El conjunto de SNV reportados en las Tablas 7-10, 14, o 15 que se encontraron en los 70 casos de PML en este estudio para los que se generaron datos de WES.

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
1	9777599	C	G	1000
1	12172008	T	C	1001
1	24486004	G	T	1002
1	33476435	C	A	1003
1	33478900	T	A	1004
1	33487007	C	T	1005
1	36932047	C	T	1006
1	36933715	A	G	1007
1	42047208	C	G	1008
1	59248085	G	C	1009
1	59248339	T	C	1010
1	92941660	C	T	1011
1	92946625	G	C	1012
1	92946625	G	C	1013
1	150053494	C	T	1014
1	155317682	C	T	1015
1	155449630	T	G	1016
1	155450331	C	T	1017
1	182554557	C	T	1018
1	198717250	T	G	1019
1	198717272	A	T	1020
1	206945738	C	T	1021
1	207641950	C	T	1022
1	235840495	G	T	1023
1	235897907	C	T	1024
1	235909815	A	T	1025
2	24431184	C	T	1026
2	24432937	C	T	1027

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
2	24435599	G	A	1028
2	47205921	C	T	1029
2	47273468	A	G	1030
2	47277182	T	C	1031
2	55910961	T	C	1032
2	71337203	C	T	1033
2	98351032	C	G	1034
2	98351066	C	T	1035
2	98351081	C	T	1036
2	113589000	C	T	1037
2	163124051	C	T	1038
2	163133194	T	C	1039
2	163134203	G	T	1040
2	163136505	C	G	1041
2	163139025	C	T	1042
2	163139085	A	T	1043
2	163144899	G	A	1044
2	163174327	C	A	1045
2	163174328	T	G	1046
2	219942026	T	A	1047
2	220023045	C	T	1048
2	230231632	C	T	1049
2	230450646	T	A	1050
3	38181899	G	T	1051
3	39323163	A	C	1052
3	53213691	G	C	1053
3	53221390	T	C	1054
3	121415370	T	C	1055
3	128204761	C	T	1056
3	128205808	C	T	1057
3	142272098	A	G	1058
3	142274880	G	C	1059
3	142281353	C	G	1060
3	142286928	C	T	1061
3	196199032	A	T	1062
3	196199204	G	T	1063
3	196210704	G	A	1064
3	196210764	T	C	1065
3	196214320	C	T	1066
4	27019452	C	T	1067

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
4	27024170	A	G	1068
4	103522068	A	G	1069
4	103522150	G	A	1070
4	103528328	C	T	1071
4	151199080	G	A	1072
4	151520216	G	A	1073
4	187003729	C	G	1074
4	187004074	C	T	1075
4	187005854	A	C	1076
5	67591018	A	C	1077
5	77311370	C	T	1078
5	77311370	C	T	1079
5	77334964	T	C	1080
5	77334964	T	C	1081
5	77335015	G	T	1082
5	77335015	G	T	1083
5	77437092	G	C	1084
5	77437092	G	C	1085
5	78596018	G	C	1086
5	138856923	C	T	1087
5	156593120	C	T	1088
5	169081453	G	C	1089
6	3077139	T	C	1090
6	12121113	C	T	1091
6	12122102	T	G	1092
6	12123538	G	T	1093
6	12124215	C	T	1094
6	12125232	C	T	1095
6	12162068	C	T	1096
6	12163657	C	T	1097
6	31928306	A	G	1098
6	31935750	G	A	1099
6	31936679	C	T	1100
6	32797809	C	T	1101
6	32810794	T	A	1102
6	32811752	C	T	1103
6	51483961	T	C	1104
6	51484077	G	C	1105
6	51491885	G	A	1106
6	51497503	C	A	1107

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
6	51524339	C	G	1108
6	51524409	G	T	1109
6	51612746	G	A	1110
6	51712759	T	C	1111
6	51747943	T	A	1112
6	51798908	C	T	1113
6	52101833	C	T	1114
6	83884161	C	G	1115
6	143081232	T	C	1116
6	143092151	T	C	1117
6	143092673	G	A	1118
6	144508353	G	A	1119
6	144508563	G	A	1120
7	2959240	C	T	1121
7	2962933	C	T	1122
7	2983958	T	C	1123
8	39840234	A	G	1124
8	39862881	C	T	1125
8	39862893	T	A	1126
8	42176189	G	A	1127
8	48690299	A	G	1128
8	48773526	T	C	1129
8	48798507	T	C	1130
8	48826575	C	G	1131
8	61654298	T	A	1132
8	61732632	A	G	1133
8	61757805	C	T	1134
8	61769428	A	G	1135
8	61777914	C	G	1136
8	61777922	C	G	1137
8	90990521	T	C	1138
8	100205255	G	A	1139
8	100791158	G	A	1140
8	100865941	G	A	1141
8	145154222	G	A	1142
8	145154222	G	A	1143
8	145154257	C	G	1144
8	145154824	A	C	1145
9	286491	G	A	1146
9	286593	C	A	1147

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
9	304628	G	A	1148
9	312134	G	A	1149
9	328047	T	A	1150
9	334277	G	A	1151
9	368128	C	T	1152
9	399233	A	G	1153
9	446401	A	G	1154
9	711359	C	T	1155
9	713132	G	T	1156
9	32526077	C	T	1157
9	32526077	C	T	1158
9	120466814	A	G	1159
9	120475302	A	G	1160
9	120475602	C	T	1161
9	120476568	A	G	1162
9	120476816	C	T	1163
10	1060218	G	A	1164
10	14974905	T	C	1165
10	14976727	G	C	1166
10	14977469	C	A.T	1167
10	72358167	G	A	1168
10	76602923	G	T	1169
10	76748831	C	T	1170
10	89720659	G	T	1171
10	90771767	G	A	1172
10	116045796	G	A	1173
11	4104626	C	A	1174
11	4112582	C	T	1175
11	9598696	G	A	1176
11	9608330	G	A	1177
11	36595321	C	T	1178
11	36596528	G	C	1179
11	36596863	C	T	1180
11	36597513	G	A	1181
11	36614561	G	T	1182
11	36615033	C	T	1183
11	67814983	G	A	1184
11	67818269	G	A	1185
11	76954833	G	A	1186
11	76979511	A	G	1187

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
11	108117787	C	T	1188
11	108119823	T	C	1189
11	108123551	C	T	1190
11	108138003	T	C	1191
11	108143456	C	G	1192
11	108175462	G	A	1193
11	108181014	A	G	1194
11	108186610	G	A	1195
11	108186631	A	G	1196
11	108198384	C	G	1197
11	108202772	G	T	1198
12	12673965	G	A	1199
12	12870798	G	A	1200
12	44166753	A	G	1201
12	44167821	A	T	1202
12	64878241	G	A	1203
12	64879775	C	T	1204
12	88900891	C	A	1205
12	93196332	C	T	1206
12	93205148	T	G	1207
12	112583447	A	C	1208
12	122064788	G	GT	1209
12	133201381	T	A	1210
12	133202816	C	T	1211
12	133209020	G	C	1212
12	133220526	T	C	1213
12	133220544	C	T	1214
12	133237658	T	G	1215
12	133245026	G	A	1216
12	133252406	C	A	1217
12	133253971	C	T	1218
12	133253995	G	A	1219
13	47466549	T	C	1220
13	108861092	G	T	1221
13	108863591	G	A	1222
14	21992397	T	C	1223
14	21993359	G	A	1224
14	22004996	G	T	1225
14	24805463	G	T	1226
14	24806303	G	A	1227

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
14	61924007	C	G	1228
14	103369593	G	A	1229
15	41011016	G	A	1230
15	68378781	A	C	1231
15	77329479	C	T	1232
15	91306241	G	A	1233
15	91310209	A	G	1234
15	91326099	C	T	1235
15	91328219	G	T	1236
15	91328310	A	G	1237
15	91341543	A	C	1238
16	1498408	G	A	1239
16	1510535	C	T	1240
16	1524855	C	G	1241
16	7568296	C	T	1242
16	7703891	A	G	1243
16	7714909	C	T	1244
16	7759119	G	A	1245
16	7759496	C	T	1246
16	24124365	A	G	1247
16	27460020	G	A	1248
16	30133233	T	C	1249
16	30134529	A	C	1250
16	50733536	T	C	1251
16	50741791	C	T	1252
16	50741791	C	T	1253
16	50744688	A	G	1254
16	50745021	C	T	1255
16	50753867	G	T	1256
16	70503095	A	G	1257
16	81819605	C	T	1258
16	81902826	C	T	1259
16	81904539	C	T	1260
16	81939089	T	C	1261
16	81942028	C	G	1262
16	81942175	A	G	1263
16	81946278	A	G	1264
16	81960772	C	A	1265
17	7577069	C	T	1266
17	16852187	A	G	1267

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
17	77926526	C	T	1268
18	43445580	C	T	1269
18	43445601	T	G	1270
18	43456296	C	T	1271
18	43458306	G	A	1272
18	43460105	C	A	1273
18	43464763	C	T	1274
18	43479473	T	C	1275
18	43488030	T	C	1276
18	43496370	G	A	1277
18	43496539	G	A	1278
18	43497710	A	G	1279
18	43523240	C	T	1280
18	43529551	C	T	1281
18	43531186	C	T	1282
18	44392443	T	C	1283
18	48584504	C	T	1284
18	56401523	C	T	1285
18	60036429	G	A	1286
18	60052034	A	C	1287
19	4817657	C	T	1288
19	4817852	G	A	1289
19	7705818	C	T	1290
19	7712287	G	C	1291
19	48631258	G	A	1292
19	48639022	T	C	1293
20	3843027	C	A	1294
20	3846397	C	T	1295
20	31383307	G	A	1296
20	31384614	G	T	1297
20	62305450	C	T	1298
20	62309621	T	C	1299
20	62326964	C	G	1300
21	16338814	T	C	1301
21	16339852	T	C	1302
21	30698953	T	G	1303
21	34809232	C	T	1304
21	45786650	C	T	1305
21	45795833	G	T	1306
21	45795877	G	T	1307

Cromosoma	Posición_hg19	REF	ALT	SEQ_ID
21	45811411	G	T	1308
21	45811438	C	T	1309
21	45815307	T	C	1310
21	45815331	G	A	1311
21	45815343	A	G	1312
21	45815425	C	G	1313
21	45820196	C	T	1314
21	45826486	G	A	1315
21	45826616	C	T	1316
21	45838333	C	T	1317
21	45844780	C	T	1318
21	45845528	G	A	1319
21	45845661	A	G	1320
21	45845699	G	A	1321
21	45855099	C	T	1322
22	21235389	A	G	1323
22	23915583	T	C	1324
22	23915745	G	A	1325
22	23917192	G	T	1326
22	36661354	C	T	1327
X	24759574	G	T	1328
X	24759574	G	T	1329

La Tabla 5 enumera, en orden de coordenadas genómicas, todas las variantes de un solo nucleótido (SNV) que son relevantes para el presente estudio, ya sea como soluciones a nivel de caso (Tablas 7, 8) o soluciones potenciales (Tablas 9, 10), o a nivel de análisis de carga de variantes (Tablas 14, 15). Todas las coordenadas del genoma se basan en hg19.

Tabla 6: Lista no redundante de 419 genes implicados en el sistema inmunitario y/o vinculados a PML a través de una CNV

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
ACADM	AR	Base de datos pública	MySql	157
ACKR1	AD	Base de datos pública	MySql	158
ACP5	AR	Base de datos pública	PMID: 26052098, 27260006, 27821552	159
ADAR	ADAR	Base de datos pública	PMID: 26052098, 27260006, 27821552	160
ADARB1	desconocido	PBio	PMID: 16227093, 17376196, 19482597, 20220308, 21682836, 21809195, 22001568, 22085847, 22110393, 24586166, 24725957, 24760760, 25826567	2
ADK	AR	PBio	PMID: 17205396, 23592612, 25654762, 25720338, 25979489, 26341619, 26642971	161
AGBL4	desconocido	PBio	PMID: 17244818, 21074048, 23085998, 25416787, 25332286, 26502776	3
AICDA	AD_AR	Base de datos pública	MySql; PMID: 23765059	162
AK2	AR	Base de datos pública	PMID: 19043417, 19782549, 20068220, 23765059, 24135998, 24753205, 26454313	163
ALG12	AR	Base de datos pública	MySql	164
ALPL	AD_AR	PBio	PMID: 18821074, 20049532, 20977932, 21191615, 21289095, 23091474, 23454488, 23860646, 26219705, 26219711, 26219717	165
AP3B1	AR	Base de datos pública	MySql; PMID: 11590544, 19782549, 24302998, 24753205, 24916509, 25980904, 27889060	166
AP3B2	AR	Base de datos pública	PMID: 26377319, 27889060	167
AP3D1	AR	Base de datos pública	PMID: 26744459, 27889060	168
APOBEC3A	desconocido	PBio	PMID: 16720547, 17303427, 20062055, 20615867, 22896897, 23344558, 23540892, 25262471, 25576866, 26416327, 26469788, 26678087	4
APOBEC3B	desconocido	PBio	PMID: 16720547, 17303427, 20062055, 20615867, 22896897, 23344558, 23540892, 25262471, 25576866, 26416327, 26469788, 26678087	6
APOL1	asociación	Base de datos pública	PMID: 27042582	169
ARHGEF7	desconocido	PBio	PMID: 11160719, 16983070, 18378701, 19861492, 25284783, 25500533	8
ASH1L	desconocido	PBio	PMID: 17923682, 17981149, 22541069, 24012418, 24244179, 25866973, 26002201, 27154821, 27229316, 27434206	170
ASTN2	desconocido	PBio	PMID: 22230861,86025532,20573900,24357807,24381304,25146927,25410587,265146622,265 86575	9
ATL2	desconocido	PBio	PMID: 18270207, 19665976, 25773277	171

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
ATM	AR	Base de datos pública	MySql; PMID: 19903823, 20301790, 23765059, 24799566, 25692705, 27042682, 27484032, 27884168, 27895165	172
ATR	ADAR	Base de datos pública	PMID: 17564965,17151099, 19903823, 20506465, 21615334, 24799566, 25910481	173
AUH	AR	PBio	PMID: 12434311, 12655555, 17130438, 20855850, 25280001, 25597510	10
BACH1	desconocido	PBio	PMID: 15068237, 18555605, 22024395, 22791292, 23456643, 23562577, 24752012, 25344725, 25391381, 24752012, 26045540, 26894991	11
BACH2	desconocido	PBio	PMID: 17262715, 17991429, 18769450, 22791292, 23728300, 24367030, 24608439, 24681888, 24694524, 25123280, 25344725, 25665584, 25686607, 26444573, 26620562, 26731475, 26894991, 26981933	174
BCL10	AR	Base de datos pública	MySql	176
BDKRB2	desconocido	PBio	PMID: 7787759, 18930543, 22047990, 22095814, 24925394	12
BLM	AR	PBio	PMID: 15137905, 15493327, 17210642, 17321898, 19109166, 19709744, 2032252, 23572515.24606147	177
BLNK	AR	Base de datos pública	PMID: 23765059	178
BLOC1S6	AR	Base de datos pública	MySql	179
BMPR2	AD	PBio	PMID: 15877825, 19191909, 23733693, 24334027	13
C11orf65	desconocido	Base de datos pública	MySql	181
C1QA	AR	Base de datos pública	PMID: 27821552	182
C1QB	AR	Base de datos pública	PMID: 27821552	183
C1QC	AR	Base de datos pública	PMID: 27821552	184
C5AR1	desconocido	PBio	PMID: 1847994, 22964232, 25041699, 25174320, 25455139, 25539817, 25769922, 26059553.26283482.26537334	185
CAPZB	desconocido	PBio	PMID: 99354614, 19806181, 22493691, 22706086, 22710966, 22918941, 23178720, 26758871	186
TARJETA11	AD_AR	Base de datos pública	MySql; PMID: 23765059, 25645939, 26525107	187
TARJETA9	AR	Base de datos pública	PMID: 27222657	188
CASP8	AR	Base de datos pública	PMID: 22365665, 26454313, 27873163, 27999438	189
CCL11	AD	Base de datos pública	MySql	190

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
CCL2	asociación	Base de datos pública	MySql	191
CCL5	asociación	Base de datos pública	MySql	192
CCR2	asociación	Base de datos pública	MySql	193
CCR5	asociación	Base de datos pública	MySql	194
CD180	desconocido	PBio	PMID: 9763566, 10880523, 21918197, 21959264, 22484241, 23103284, 23483427, 24019553, 25749095, 26371254, 26384474, 26482097, 26555723, 26371254	195
CD19	AR	Base de datos pública	MySql; PMID: 23765059, 26453379	196
CD209	asociación	Base de datos pública	MySql	197
CD247	AR	Base de datos pública	PMID: 26454313	198
CD27	AR	Base de datos pública	MySql; PMID: 23765059	199
CD27-AS1	desconocido	Base de datos pública	MySql	200
CD300LF	desconocido	PBio	PMID: 15184070, 15549731, 17202342, 18688020, 19592130, 22288587, 23072861, 23293083,24035150	23
CD34	desconocido	Base de datos pública	PMID: 27042682	201
CD3D	AR	Base de datos pública	PMID: 23765059, 26454313	202
CD3E	AR	Base de datos pública	PMID: 23765059, 26454313	203
CD3G	AR	Base de datos pública	PMID: 23765059, 26454313	204
CD40	AR	Base de datos pública	MySql; PMID: 23765059, 26453379	205
CD55	desconocido	PBio	PMID: 12417446, 1385527, 16406700, 16503113, 17678954, 18424707, 19660813, 21143144, 22795896, 24588829, 24639397, 25156074, 25954012, 26423932	207
CD59	AR	Base de datos pública	MySql	208
CD79A	AR	Base de datos pública	PMID: 23765059	209
CD79B	AR	Base de datos pública	PMID: 23765059	210
CD81	AR	Base de datos pública	MySql	211
CD8A	AR	Base de datos pública	PMID: 26454313	212
CDCA7	AR	Base de datos pública	PMID: 26216346	213

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
CDKN1B	AD	PBio	PMID: 10799578, 10825149, 10916090, 11123298, 11123306, 15557280, 16410832, 17273559, 20854895, 21078910, 22454463, 24317118, 25213837	24
CEBPB	desconocido	Base de datos pública	PMID: 27042682	214
CENPM	desconocido	PBio	PMID: 15183305,16391015,19711193,25006165	25
CHD7	AD	Base de datos pública	PMID: 18505430, 18976358, 26454313, 27484032	215
CHEK1	desconocido	Base de datos pública	PMID: 19903823, 27042682	216
CIITA	AR	Base de datos pública	PMID: 23765059, 26454313, 27484032	217
CLCN7	AD		PMID: 21107136, 25992615	218
COG4	AR	PBio	PMID: 18086915, 18256213, 20065092, 20143049, 21421995, 23462996, 23865579, 24784932, 26125015	26
COG6	AR	Base de datos pública	MySql	219
COMD6	desconocido	PBio	PMID: 14685242, 15799966, 16573520, 20126548, 25355947, 27441653	27
CORO1A	AR	Base de datos pública	PMID: 23887241, 26454313	220
CR2	AR	Base de datos pública	MySql	221
CRADD	AR	PBio	PMID: 11573962, 21242994, 22323537, 24958727, 26190521, 27135977	28
CRTC3	desconocido	PBio	PMID: 15466468, 2032252, 21536665, 23033494, 23241891, 25114223, 25316186, 25351958,26937622	222
CSF3R	AR	Base de datos pública	PMID: 24753537, 26324699, 27789332	223
CTLA4	AD	Base de datos pública	PMID: 18219311, 25213377, 25329329	224
CTPS1	AR	Base de datos pública	MySql	225
CTSC	AR	Base de datos pública	PMID: 27222657	226
CX3CR1	asociación	Base de datos pública	MySql	227
CXCL12	asociación	Base de datos pública	MySql	228
CXCL9	desconocido	Base de datos pública	PMID: 27042682	229
CXCR1	asociación	Base de datos pública	MySql	230
CXCR4	AD	Base de datos pública	PMID: 19782549, 19950235, 23765059, 24753205, 25645939, 28009639	231

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
CXorf40A	desconocido	PBio	PMID: 15541360, 24916366, 26881174	232
CYBB	XLR	Base de datos pública	PMID: 27222657	233
CYP2S1	desconocido	PBio	PMID: 15681441, 23933117	234
DCLRE1C	AR	Base de datos pública	PMID: 26454313, 26476407, 27484032	235
DDX1	desconocido	Base de datos pública	PMID: 27042682	236
DDX58	AD	Base de datos pública	PMID: 23592984, 25692705, 25794939, 26052098, 26748340, 26848516, 26987611, 27260006, 27821552	237
DHX58	asociación	Base de datos pública	PMID: 25794939, 26748340, 26848516	238
DKC1	XLR	Base de datos pública	MySql; PMID: 23765059	239
DNER	desconocido	PBio	PMID: 15965470, 16298139, 16997755, 17765022, 18474614, 20058045, 20367751, 22447725, 23041955, 23328254, 24248099, 24935874, 26869529	31
MUELLE2	AR	Base de datos pública	MySql	241
DSC1	desconocido	PBio	PMID: 16048752, 20222919, 22692770, 24460202, 24680560, 25078507, 25244249, 26043694,26758100	243
DUSP16	desconocido	PBio	PMID: 15284860, 21613215, 24311790, 25716993, 26381291, 27162525	32
ECRP	desconocido	PBio	PMID: 9826755, 12855582, 25271100, 26184157	33
EDIL3	desconocido	PBio	PMID: 22601780, 23518061, 24060278, 24504014, 25385367, 26038125	34
EEE1	desconocido	PBio	PMID: 16670179, 22591512, 24491918, 24561067, 26478006, 26909655, 27077111	35
EGR1	desconocido	PBio	PMID: 15308739, 19050264, 19812322, 20414733, 21368226, 21622185, 22554935, 25613134, 26052046, 26980486, 11910893, 14647476, 18203138, 24627779, 25368162,27192563	244
EHF	desconocido	PBio	PMID: 16380452, 17027647, 19801549, 20879862, 21617703, 24219556, 25217163	36
ELANE	AD	Base de datos pública	PMID: 20008220, 24145314, 27222657	245
MSB	desconocido	PBio	PMID: 8432389, 15917240, 18209069, 19164284, 25773908	37
EPG5	AR	Base de datos pública	MySql: PMID: 21965116, 23222957, 23838600, 26917586, 26927810, 27588602	246
ETF1	desconocido	PBio	PMID: 20418372, 25606970, 26384426, 26833392, 27387891	247

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
ETV6	AD	PBio	PMID: 19264918, 20350684, 21714648, 22438058, 25581430, 25807284, 26718572, 27365488	38
F9	XLR	Base de datos pública	MySql	248
FAS	AD	Base de datos pública	PMID: 27222657	249
FASLG	AD	Base de datos pública	PMID: 27222657	250
FCGR2A	AD_AR	Base de datos pública	MySql	251
FCGR3A	AR	Base de datos pública	MySql	252
FCN3	AR	Base de datos pública	MySql	253
FEZ1	desconocido	Base de datos pública	PMID: 27042682	254
FHL2	desconocido	PBio	PMID: 16389449, 20592280, 22417706, 22633286, 23212909	39
FOS	desconocido	Base de datos pública	PMID: 27042682	255
FOXH1	desconocido	Base de datos pública	PMID: 27042682	256
FOXN1	AR	Base de datos pública	MySql	257
FOXP3	XLR	Base de datos pública	PMID: 26454313	258
FPR1	desconocido	PBio	PMID: 8994115, 10229829, 10611407, 17084101, 22934745, 23230437, 25605714, 25826286.26101324.26701131. 27034344, 27100350, 27131862, 27154726	259
FPR2	desconocido	PBio	PMID: 8994115, 10229829, 10611407, 17084101, 22934745, 23230437, 25605714, 25826286.26101324.26701131. 27034344, 27100350, 27131862, 27154726	41
FPR3	desconocido	PBio	PMID: 8994115, 10229829, 10611407, 17084101, 22934745, 23230437, 25605714, 25826286, 26101324, 26701131, 27034344, 27100350, 27131862, 27154726	42
FUK	desconocido	PBio	PMID: 11753075, 12651883, 15774760, 19394435, 19647987, 20383321, 22134107, 22203233, 22276660, 22461019, 24239607	43
G6PC3	AR	Base de datos pública	PMID: 20008220, 24145314, 25879134, 26479985	260
GATA2	AD	Base de datos pública	PMID: 23765059, 23887241	261
GDA	desconocido	PBio	PMID: 10595517, 18600524, 20826431, 23838888, 24834013	44
GDPD4	desconocido	PBio	PMID: 24373430, 24977479, 24977485, 24977489, 25528375, 25596343	45
GFI1	AD	Base de datos pública	PMID: 20008220, 24145314	262

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
GOLGB1	desconocido	PBio	PMID: 17475246, 21217069, 22034594, 23555793, 24046448	263
GPATCH2	desconocido	PBio	PMID: 19432882, 25353171, 25376275	46
GPC5	desconocido	PBio	PMID: 24130709, 24943672, 25354479, 26224662, 26349477	47
GPRC5A	desconocido	PBio	PMID: 19593893, 20959490, 22239913, 25621293, 25714996, 26165721	264
GRAP2	desconocido	Base de datos pública	PMID: 25452106, 25636200, 26246585	265
GRIA3	XLR	PBio	PMID: 10441169, 12682273, 17202328, 18590483, 25904555, 26648591	51
GTPBP4	desconocido	PBio	PMID: 17785438, 26015807	52
HAX1	AR	Base de datos pública	PMID: 20008220, 24145314	266
HCN1	AD	PBio	PMID: 9405696, 9630217, 9634236, 9921901, 11133998, 23042740, 23077068, 23319474, 24403084, 24747641, 24756635, 25580535, 26578877	53
HELLS	AR	Base de datos pública	PMID: 26216346	267
HEXA	AR	PBio	PMID: 20301397, 21997228, 23727835, 24445368	54
HIVEP1	asociación	Base de datos pública	MySql; PMID 20226436, 26117544	268
HIVEP2	AD	Base de datos pública	MySql; PMID: 21475200, 21936769, 23389689, 24366360, 26153216, 26483320, 27003583	269
HIVEP3	desconocido	Base de datos pública	MySql	270
HK2	desconocido	PBio	PMID: 2749583, 4265132, 19519254, 24966891, 23874603, 25525876, 25602755	55
HMGB3	desconocido	PBio	PMID: 12714519, 15259015, 15358624, 16945912, 22014684, 23994280, 26553261	56
HNRNPLL	desconocido	PBio	PMID: 18669861, 18719244, 19100700, 20505149, 22073166, 23934048, 24476532, 25825742	271
HP	desconocido	PBio	PMID: 16046400, 19380867, 19795414, 26445729	272
HPCAL1	desconocido	PBio	PMID: 12445467, 24699524, 25519916, 26659654, 26729710	273
HPR	desconocido	PBio	PMID: 16046400, 19380867, 19795414, 26445729	57
HTR2A	asociación	Base de datos pública	PMID: 19204164, 24089568, 25078361, 26056932, 27003757, 27042682	274
ICOS	AR	Base de datos pública	MySql; PMID: 19380800, 23765059, 27250108	275
IDI1	desconocido	PBio	PMID: 14629038, 17202134, 19454010, 20955688, 22579571, 23585482, 25950736	276
IDI2	desconocido	PBio	PMID: 14629038, 17202134, 19454010, 20955688, 22579571, 23585482, 25950736	59

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
IDI2-AS1	desconocido	PBio	Véase IDI2	60
IDO2	desconocido	PBio	PMID: 18219311, 18364004, 19487973, 19799997, 20197554, 20484729, 20693847, 21084489, 21406395, 21990421, 22754762, 24391212, 24402311, 24844751, 25477879, 25478733, 25541666, 25949913, 26308414, 27183624	61
IFIH1	AD	Base de datos pública	PMID: 21156324, 24686847, 24995871, 25794939, 26052098, 26748340, 27260006, 27821552	277
IFNAR1	asociación	Base de datos pública	PMID: 27821552	278
IFNAR2	AR	Base de datos pública	PMID: 26424569, 27821552	279
IFNG	asociación	Base de datos pública	MySql	280
IFNGR1	AD AR	Base de datos pública	MySql	281
IFNGR2	ADAR	Base de datos pública	MySql; PMID: 15356149, 23161749	282
IFNLR1	desconocido	PBio	PMID: 12469119, 12483210, 15166220, 22386267, 22891284, 25634147, 25904743, 25941255	62
IGLL1	AR	Base de datos pública	PMID: 25502423	283
IKBKB	AD_AR	Base de datos pública	MySql: PMID: 17047224, 17072332, 25764117, 25930993, 26117626, 26525107	284
IKBKG	XLD XLR	Base de datos pública	MySql: PMID: 17047224, 21455173, 21455181, 23765059, 25764117, 25886387, 25930993, 26117626, 26525107	285
IKZF1	AD	Base de datos pública	PMID: 26454313, 26981933	286
IL10	AR	Base de datos pública	PMID: 23887241	287
IL10RA	AR	Base de datos pública	PMID: 23887241	288
IL10RB	AR	Base de datos pública	PMID: 23887241	289
IL12B	AR	Base de datos pública	MySql	290
IL12RB1	AR	Base de datos pública	MySql	291
IL17F	AD	Base de datos pública	PMID: 22264928, 23887241, 24240291, 24690400, 25890879, 27144517	292
IL17RA	AR	Base de datos pública	PMID: 23887241	293
IL1B	AD	Base de datos pública	PMID: 15327898, 20543597, 24248593, 26525107, 27730320, 27873163, 27999438	294
IL21	AR	Base de datos pública	MySql	295

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
IL21R	AD AR	Base de datos pública	PMID: 23765059, 23887241	296
IL2RA	AR	Base de datos pública	MySql	297
IL2RG	XLR	Base de datos pública	PMID: 23765059, 26454313, 27484032	298
IL4R	asociación	Base de datos pública	MySql	299
IL7	desconocido	PBio	PMID: 21508983, 22288682, 24507157, 24979548, 25130296, 25214510, 25411246, 25734144, 26537673, 26675348, 26908786	300
IL7R	AR	Base de datos pública	PMID: 26454313, 27484032	301
IRAK4	AD_AR	Base de datos pública	PMID: 23766853, 25232776, 25344726, 25764117, 25886387, 25930993, 26785681, 27845762	302
IRF3	AD	Base de datos pública	PMID: 23388631, 26513235, 26748340	303
IRF7	AR	Base de datos pública	MySql; PMID: 26748340	304
IRF8	AD AR	Base de datos pública	PMID: 23887241	305
IRGM	asociación	PBio	PMID: 14707092, 17911638, 22174682, 22722598, 23084913, 23335927	306
ISG15	AR	Base de datos pública	MySql; PMID: 26052098, 27260006, 27821552	307
ITSN2	desconocido	PBio	PMID: 11748279, 15020715, 17696400, 17696407, 22558309, 22975684, 23986746, 24097067, 24284073, 25797047, 26479042	309
JAGN1	AR	Base de datos pública	PMID: 25129144	310
JAK3	AR	Base de datos pública	PMID: 23765059, 26454313	311
JMY	desconocido	PBio	PMID: 19287377, 20573979, 20574148, 20888769, 21965285, 23291261, 26015719, 26223951, 26305109	312
JPX	asociación	PBio	PMID: 21029862, 23791181, 23943155	64
JUN	desconocido	Base de datos pública	PMID: 16928756, 27042682	313
KANK1	asociación	PBio	PMID: 18458160, 20164854, 21685469, 24399197, 25961457, 26656975	65
KAT6B	AD	PBio	PMID: 17460191, 17694082, 22715153, 25920810	66
KCTD7	AR	PBio	PMID: 17455289, 20301601, 21710140, 22606975, 22638565, 22748208, 25060828, 27629772, 27742667	67
KITLG	AD	Base de datos pública	PMID: 27042682	314

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
LAMTOR2	AR	Base de datos pública	MySql; PMID 19782549, 20008220, 24753205	315
LARP4B	desconocido	PBio	PMID: 20573744, 23815932, 25534202, 26001795, 26501340, 26644407	69
LCP2	desconocido	Base de datos pública	PMID: 12874226, 18219311, 19056831, 23494777, 26246585	317
LIG1	AR	Base de datos pública	MySql	318
LIG4	AR	Base de datos pública	MySql	319
LOC 102724297	desconocido	Base de datos pública	MySql	320
LOC400710	desconocido	PBio	ARNnc, información génica limitada; véase familia génica SNAR (locus adyacente)	321
LRBA	AR	Base de datos pública	MySql; PMID: 23765059, 27873163, 27192563	322
LYST	AR	Base de datos pública	PMID: 19302049, 19782549, 20008220, 24753205, 24916509, 26454313, 27881733	323
MAGEA9	desconocido	PBio	PMID: 15222021, 15900605, 21093980, 21791470, 25315972, 25445503, 25755744	324
MAGEA9B	desconocido	PBio	PMID: 15222021, 15900605, 21093980, 21791470, 25315972, 25445503, 25755744	325
MAGT1	XLR	Base de datos pública	PMID: 23887241, 27873163, 25504528	326
MALL	desconocido	PBio	PMID: 11294831, 19064697, 24101378, 24746959, 26109641, 26622604, 26641089, 26772392, 27583248, 27846891	72
MALT1	AR	Base de datos pública	MySql; PMID: 26525107	327
MAP3K2	desconocido	PBio	PMID: 11032806, 11278622, 12138187, 14734742, 16430878, 21333552, 2437584, 24847879, 25012295, 26056008	328
MAPK1	desconocido	Base de datos pública	PMID: 14671106, 27042682	329
MAPK3	desconocido	Base de datos pública	PMID: 14671106, 27042682	330
MAPK9	desconocido	PBio	PMID: 15023353, 23685277, 24673683, 25762148, 26141991	73
MAVS	asociación	Base de datos pública	PMID: 23582325, 26513235, 26987611	331
MCEE	AR	PBio	PMID: 17846917, 20301409, 21365456, 23726524, 24532006, 25763508, 26725562	74
MECP2	XLD XLR	Base de datos pública	PMID: 27042682	332
MEX3C	desconocido	PBio	PMID: 18779327, 22357625, 22658931, 22863774, 22927639, 23140835, 23446422, 23999169, 24706898, 24741071	333
MGAT5	desconocido	PBio	PMID: 12417426, 15585841, 18292539, 20089585, 20117844, 25768892, 26972830	75

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
MKL1	AR	PBio	PMID: 12944485, 22626970, 26098208, 26098211, 26221020, 26241940, 26405212, 26224645,26554816	89
MRE11A	AR	Base de datos pública	PMID: 23388631, 23765059	334
MS4A1	AR	Base de datos pública	MySql; PMID: 23765059	335
MSN	desconocido	PBio	PMID: 9070665, 10444190, 11777944, 12445265, 14758359, 16368573, 17110458, 18025306, 18725395, 21486194, 23526587, 23613524, 24250818, 24358210, 24760896,25746045	336
MTHFD1	AR	Ambos	PMID: 26454313	337
MYD88	AD_AR	Base de datos pública	PMID: 23766853, 25344726, 25764117, 25886387, 25930993, 26371186, 27435819	338
NBN	AD_AR	Base de datos pública	MySql; PMID: 23765059	339
NFIC	desconocido	PBio	PMID: 11559801, 15327898, 16928756, 18474555, 19058033, 22205750	340
NFIL3	desconocido	PBio	PMID: 20080759, 20697558, 22075207, 23453631, 24070385, 24277151, 24280221, 24442434, 24909887, 25092873, 25113970, 25310240, 25611557, 25614966, 25801035, 25993115, 26153760, 26379372, 26806130, 26880402	92
NFKB1	AD	Base de datos pública	PMID: 22081022, 26279205	341
NFKB2	AD	Base de datos pública	MySql; PMID: 25764117	342
NFKBIA	AD	Base de datos pública	MySql; PMID: 23765059, 25645939, 25764117	343
NHEJ1	AR	Base de datos pública	MySql: PMID: 23765059	344
NLRP12	AD	PBio	PMID: 17947705, 18230725, 20861596, 21978668, 23318142, 23970817, 24282415, 24347638, 25249449, 25620184, 25902475, 26083549, 26343520, 26386126, 26521018	93
NLRP3	AD	Base de datos pública	PMID: 16724804, 19302049, 23592984, 26848516, 27999438	345
NOD2	AD	Base de datos pública	PMID: 16724804, 19302049, 23584365, 26509073, 26848516, 26953272	346
NQO2	desconocido	PBio	PMID: 16253210, 16905546, 17720881, 18552348, 26046590	94
NRIP1	desconocido	PBio	PMID: 18267075, 23241901, 24969109, 25066731, 25697398, 25879677, 26937622	95
ORAI1	AD_AR	Base de datos pública	PMID: 19075015, 20004786, 21790973, 22144678, 23765059, 26454313, 26469693	347
OSTM1	AR	Base de datos pública	PMID: 16813530, 19507210, 21107136, 23685543	348
OVOL2	AD	PBio	PMID: 16423343, 25267199, 26619963, 26749309	98

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
PDE3B	desconocido	PBio	PMID: 17220874, 23276671, 25816736, 26203135, 26297890, 26374610	99
PDGFRA	asociación	PBio	PMID: 12660384, 18634583, 18701889, 19246520, 19839938, 20032375, 20569695, 21123594, 21975205, 22449623, 22523564, 23771592, 25319708, 25940087	100
PDSS2	AR	PBio	PMID: 17186472, 18437205, 18764258, 21567994, 21871565, 21983691, 23150520	101
PGM3	AR	Base de datos pública	MySql; PMID: 25502423	349
PHACTR4	desconocido	PBio	PMID: 15107502, 17609112, 22215804, 22215812, 22766235, 23076051, 23203801, 23319639, 24748504, 26850007	102
PIAS1	desconocido	PBio	PMID: 10805787, 10858346, 14644436, 15297606, 15311277, 17065208, 17540171, 18056374, 19857525, 20966256, 22369086, 22982248, 23299081, 24036127	103
PIAS2	desconocido	PBio	PMID: 9724754, 11117529, 12077349, 12764129, 14514699, 15582666, 16460827, 19549844, 21156324, 21779164, 22210188, 22982248, 24344134, 25484205, 25434787, 26223632	350
PIK3CD	AD	Ambos	MySql; PMID: 24165795, 25133419, 25645939, 26437962, 26453379, 27379089, 27426521, 27873163, 14647476, 27192563	104
PIK3R1	AD AR	Base de datos pública	PMID: 23765059, 23867241, 25645939, 26246585, 26453379, 27076228, 14647476, 27192563	351
PKHD1	AR	PBio	PMID: 8178467, 15052665, 17450421, 23423256, 24964219, 24984783, 25186187, 26502924	105
PLCG2	AD	Base de datos pública	PMID: 19056831, 23000145, 23765059, 23867241, 25452106, 25536200, 25645939, 26246585,27192563	352
PMS2	AR	Base de datos pública	MySql; PMID: 23765059	353
PNPLA4	desconocido	PBio	PMID: 22289388, 23741432, 26017929, 26184793, 26713677, 26741492, 26968210	107
PNPT1	AR	PBio	PMID: 14563561, 15492272, 16410805, 16687933, 17983748, 19580345, 23084291, 23221631, 24143183, 24729470, 25457163	108
POLA1	XLR	Base de datos pública	PMID: 27019227, 27821552	355
POLE	AR	Base de datos pública	MySql; PMID: 23230001, 23765059, 25948378	356
PPP2R3B	desconocido	PBio	PMID: 9847399, 10629059, 11593413, 18353419, 20485545, 26983421	109
PRF1	AD AR	Base de datos pública	PMID: 17311987, 19302049, 21881043, 24316509, 25776844, 26454313, 26864340, 27391055	357

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
PRKCB	desconocido	PBio	PMID: 10872892, 15488737, 16935002, 17060474, 17395590, 19907441, 21997316, 22994860,23959874,24550541,25548371,24550541,25808972,26509731, 26510741	110
PRKCD	AR	Base de datos pública	MySql: PMID: 23319571. 27250108, 27873163	358
PRKCH	desconocido	PBio	PMID: 15327898, 16571806, 18353419, 22114277, 22155788, 22892130, 23868949, 24705298. 25617472. 25889880	111
PRKDC	ADAR	Base de datos pública	PMID: 12847277, 23722905, 26454313, 26838362, 27980111	359
PROC	AD AR	PBio	PMID: 2437584, 18751723, 21114396, 22447930, 24162617	360
PSMB8	AR	Base de datos pública	PMID: 26052098, 27260006, 27821552	361
PSTPIP1	AD	PBio	PMID: 9488710, 11313252, 12530983, 14707117, 16724804, 19290936, 19302049, 24421327, 25040622, 25645939, 25814341, 26386126, 26919742	112
PTEN	AD	Base de datos pública	PMID: 26246517, 27426521	362
PTPN2	desconocido	PBio	PMID: 11909529, 12359225, 12847239, 19290937, 19825843, 19930043, 20473312, 20564182, 20848498, 21220691, 22080861, 22080863, 22671594, 24442435, 24445916. 24608439. 24849651. 24997008. 25548153. 25581833	113
PTPRC	AR	Base de datos pública	PMID: 26454313	363
PTPRN2	desconocido	PBio	PMID: 9714834, 10426369, 11086001, 11086294, 11793386, 15114673, 19361477, 23595248, 24988487, 26141787, 26609326	114
PURA	AD	Base de datos pública	PMID: 27042682	364
RAB27A	AR	Base de datos pública	PMID: 19302049, 20008220, 21881043, 23810987	365
RAB37	desconocido	PBio	PMID: 21805469, 22899725, 26931073, 27798165	115
RAB7A	AD		PMID: 25992615, 27588602	366
RABGEF1	desconocido	PBio	PMID: 12505986, 15143060, 15235600, 16499958, 16533754, 16605131, 17341663, 20829437, 22846990, 23552075, 24569883, 24957337, 25427001, 26567216, 26588713, 27791468	367
RAC2	AD	Base de datos pública	MySql	368
RAD51	AD	Base de datos pública	PMID: 25310191, 27042682	369
RAG2	AR	Base de datos pública	PMID: 23765059, 23887241, 26454313, 27808398	371

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
RBCK1	AR	Base de datos pública	MySql; PMID: 21455173, 21455181, 23765059, 23969028, 24958845, 25764117, 25930993, 26008899, 26525107, 27810922	372
RBFOX1	desconocido	PBio	PMID: 23350840, 24039908, 25043849, 26500751, 26687839	116
CCR1	desconocido	PBio	PMID: 1961752, 18442486, 19060893, 20347844, 23536659, 25452301, 26864624	117
RFX5	AR	Base de datos pública	PMID: 23765059, 26454313	373
RFXANK	AR	Base de datos pública	PMID: 23765059, 26454313	374
RFXAP	AR	Base de datos pública	PMID: 23765059, 26454313	375
RGCC	desconocido	PBio	PMID: 19158077, 19652095, 23000427, 24973210, 25770350, 26134570	118
RHOQ	desconocido	PBio	PMID: 10490598, 12456725, 14734537, 16246732, 17016434, 19258391, 22916134, 24223996, 24297911, 24663214, 24667291	119
RIPK1	asociación	Base de datos pública	PMID: 21455173, 27999438	376
RIPK3	asociación	Base de datos pública	PMID: 22365665, 27999438	377
RMRP	AR	Base de datos pública	MySql; PMID 19782549, 20008220, 24753205	378
RNASE3	desconocido	PBio	PMID: 19515815,26184157	120
RNASEH2A	AR	Base de datos pública	PMID: 26052098, 27260006, 27821552	379
RNASEH2B	AR	Base de datos pública	PMID: 26052098, 27260006, 27821552	380
RNASEH2C	AR	Base de datos pública	PMID: 26052098, 27260006, 27821552	381
RNASEL	asociación	Base de datos pública	PMID: 24995003, 27525044	382
RNF168	AR	Base de datos pública	MySql; PMID: 23765059	383
RNF31	AR	Base de datos pública	PMID: 21455173, 21455181, 23969028, 24958845, 26008899, 26525107, 27810922	384
RNU4ATAC	AR	Base de datos pública	PMID: 27222657	385
RPTOR	desconocido	PBio	PMID: 16959881, 22810227, 23349361, 23812589, 24287405, 24303063, 24671993, 24948799, 26678875	123
RTEL1	AR	Base de datos pública	MySql; PMID: 23329068, 23765059, 24009516, 25607374, 26810774	386
RTEL 1-TNFRSF6B	desconocido	Base de datos pública	MySql; PMID: 23329068, 23765059, 25607374	387

Simbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
SALL2	AR	PBio	PMID: 11734654, 15082782, 18818376, 19076363, 19131967, 21362508, 21689070, 21791360, 22074632, 22978642, 23029531, 24040083, 24412933, 24903482, 25360671, 25580951, 25608837, 26181197	388
SAMHD1	AR	Base de datos pública	PMID: 26052098, 27260006, 27821552	389
SBDS	AR	Base de datos pública	PMID: 20008220, 21062271, 27418648, 27658964	390
SERPINB4	desconocido	PBio	PMID: 15203215, 19070595, 21857942, 22451727, 22808225, 24560885, 24635038, 25111616, 25133778, 25213322	124
SERPINB6	AR	PBio	PMID: 14670919, 20451170, 24172014, 24359430	125
SH2D1A	XLR	Base de datos pública	MySql; PMID: 19302049, 23765059, 25744037	391
SHARPIN	desconocido	Base de datos pública	PMID: 21455181, 22901541, 23969028, 24958845, 26525107, 26848516, 27810922, 27892465	392
SKIV2L	AR	Base de datos pública	PMID: 27260006, 27821552	393
SLC17A5	AR	PBio	PMID: 14742248, 15006695, 15172005, 16575519, 18399798, 20007460, 20951965, 21628664, 22778404, 23760462, 23889254, 25494612, 25855729, 25879139	127
SLC37A4	AR	Base de datos pública	PMID: 20008220, 20301489	394
SLC3A2	desconocido	PBio	PMID: 22588539, 22624878, 23297381, 24491544, 25002078, 26172215, 26439699, 26444422	126
SLC46A1	AR	Base de datos pública	PMID: 26454313	395
SLC8A1	desconocido	PBio	PMID: 23224883, 23224887, 23224890, 23224891, 26045217, 26418956, 26775040, 26859825, 26924806	396
SMAD2	desconocido	Base de datos pública	PMID: 27042682	397
SMAD3	AD	Base de datos pública	PMID: 27042682	398
SMAD4	AD	Ambos	PMID: 12202226, 14987161, 16800882, 19420158, 25637015, 25705527, 26454313, 27042682	399
SNAP29	AR	Base de datos pública	PMID: 15968592, 21073448, 27588602	400
SNAR-A1	desconocido	PBio	PMID: 25327818, 25447144	401
SNAR-A10	desconocido	PBio	PMID: 25327818, 25447144	402
SNAR-A11	desconocido	PBio	PMID: 25327818, 25447144	403

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
SNAR-A12	desconocido	PBio	PMID: 25327818, 25447144	404
SNAR-A13	desconocido	PBio	PMID: 25327818, 25447144	405
SNAR-A14	desconocido	PBio	PMID: 25327818, 25447144	406
SNAR-A2	desconocido	PBio	PMID: 25327818, 25447144	407
SNAR-A3	desconocido	PBio	PMID: 25327818, 25447144	408
SNAR-A4	desconocido	PBio	PMID: 25327818, 25447144	409
SNAR-A5	desconocido	PBio	PMID: 25327818, 25447144	410
SNAR-A6	desconocido	PBio	PMID: 25327818, 25447144	411
SNAR-A7	desconocido	PBio	PMID: 25327818, 25447144	412
SNAR-A8	desconocido	PBio	PMID: 25327818, 25447144	413
SNAR-A9	desconocido	PBio	PMID: 25327818, 25447144	414
SNAR-B1	desconocido	PBio	PMID: 25327818, 25447144	415
SNAR-B2	desconocido	PBio	PMID: 25327818, 25447144	416
SNAR-C1	desconocido	PBio	PMID: 25327818, 25447144	417
SNAR-C2	desconocido	PBio	PMID: 25327818, 25447144	418
SNAR-C3	desconocido	PBio	PMID: 25327818, 25447144	419
SNAR-C4	desconocido	PBio	PMID: 25327818, 25447144	420
SNAR-C5	desconocido	PBio	PMID: 25327818, 25447144	421
SNAR-D	desconocido	PBio	PMID: 25327818, 25447144	422
SNAR-E	desconocido	PBio	PMID: 25327818, 25447144	423
SNAR-F	desconocido	PBio	PMID: 25327818, 25447144	424
SNAR-G1	desconocido	PBio	PMID: 25327818, 25447144	425
SNAR-G2	desconocido	PBio	PMID: 25327818, 25447144	426
SNAR-H	desconocido	PBio	PMID: 25327818, 25447144	427
SNAR-I	desconocido	PBio	PMID: 25327818, 25447144	428

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
SNCA	AD	PBio	PMID: 12406186, 14648159, 16953112, 19115126, 19432400, 19652146, 22209147, 23378275, 23771222, 24586351, 24593806, 25092570, 25450953, 25522431, 25635231, 25866630, 26087293, 26272943, 26342897, 26646749	429
SNHG3	desconocido	PBio	PMID: 22308462, 22843687, 26373735	128
SNX10	AR	Base de datos pública	PMID: 22499339, 23123320	430
SNX5	desconocido	PBio	PMID: 10600472, 11128621, 14499622, 15133132, 15561769, 16857196, 18596235, 21725319, 21903422, 21943487, 23213485, 24820351, 26220253	130
SOCS2	desconocido	PBio	PMID: 19279332, 21403007, 22693634, 22795647, 23455506, 24400794, 26216515, 26709655, 26765997, 27071013, 27158906, 27330188, 27338192	131
SP110	AR	Base de datos pública	MySql	431
SP140	desconocido	Base de datos pública	MySql	432
SPINK5	AR	Base de datos pública	PMID: 19683336, 26865388, 27222657, 27905021	433
SOSTM1	AD AR	Base de datos pública	PMID: 19229298, 27715390	434
SRSF1	desconocido	Base de datos pública	PMID: 27042682	435
ST8SIA5	desconocido	PBio	PMID: 11089916, 15829700	133
STAT2	AR	Base de datos pública	PMID: 23391734, 26122121, 27821552	437
STAT5B	AR	Base de datos pública	MySql	439
STIM1	AD_AR	Base de datos pública	PMID: 20004786, 21790973, 23765059, 26454313, 26469693	440
STIM2	desconocido	PBio	PMID: 20004786, 21790973, 21880262, 22129055, 22477146, 22914293, 25157823, 26109647,26469693	134
STK4	AR	Base de datos pública	PMID: 19782549, 23765059, 23887241, 24753205, 26029204	441
STX11	AR	Base de datos pública	PMID: 19302049, 21881043, 24916509, 26454313	442
STXBP2	AD_AR	Base de datos pública	PMID: 21881043, 24916509, 25564401, 26454313	443
SYNCRIP	desconocido	PBio	PMID: 10734137, 18045242, 19331829, 19232660, 22493061, 22935615, 23679954, 23700384, 24844655, 25100733, 26641092	444
T	AD AR	PBio	PMID: 11897834, 17438107, 23064415, 23662285, 24253444, 24556085, 25186612, 26210634, 26919728	445
TAP1	AR	Base de datos pública	PMID: 26454313	446

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
TAP2	AR	Base de datos pública	PMID: 26454313	447
TAPBP	desconocido	Base de datos pública	PMID: 26454313	448
TAZ	XLR	Base de datos pública	PMID: 20008220	449
TBC1D16	desconocido	PBio	PMID: 16923123, 19077034, 21250943, 23019362, 23485563, 23812537, 24513270, 26030178	136
TBK1	AD	Base de datos pública	PMID: 23887241, 25930993, 26513235, 28049150	450
TBX1	AD	Base de datos pública	PMID: 26454313	451
TCIRG1	AD AR	Base de datos pública	MySql; PMID 19507210, 19782549, 24753205, 27233968	452
TICAM1	ADAR	Base de datos pública	PMID: 22105173, 23887241, 25764117, 25930993, 26513235, 28049150	453
TLR3	AD	Base de datos pública	PMID: 23592984, 23887241, 25930993, 26513235, 27810922, 27873163, 27881733	454
TLR4	asociación	Ambos	PMID: 12124407, 17893200, 18946062, 19843948, 20521908, 21677132, 22474023, 22962435, 23055527, 23890253, 25365308, 25454804, 25930993, 26189680, 26453379, 27881733	455
TMMS173	AD	Base de datos pública	PMID: 23388631, 25645939, 25692705, 26052098, 27260006, 27801882, 27821552	456
TNF	asociación	Base de datos pública	MySql; PMID: 27042682	457
TNFAIP3	AD	Base de datos pública	PMID: 23969028, 26642243, 27845235	458
TNFRSF10A	desconocido	PBio	PMID: 10889508, 11602752, 11704827, 11777550, 11844843, 12390973, 12694389, 14975593, 15007095, 16394652, 16554480, 17671142, 19690337, 20921531	138
TNFRSF11A	AD AR	Base de datos pública	PMID: 17088646, 17360404, 18281276, 18606301, 19380800, 19507210, 25102334, 25393853,27003757,27016605	459
TNFRSF11B	AR	Base de datos pública	PMID: 19507210, 25102334, 25393853, 27003757	460
TNFRSF13B	AD AR	Base de datos pública	MySql; PMID: 17467261, 17492055, 18978466, 18981294, 19629655,20889194, 21458042, 22697072, 23765059, 25454804, 25930993, 26727773, 27123465	461
TNFRSF13C	AR	Ambos	MySql; PMID: 16769579, 17785824, 18784835, 18813230, 19136305, 19406831, 20547827, 20547828, 20817206, 21897850, 22028296, 22030463, 23684423, 24101550, 24953530, 25454804, 25637018, 25724205, 26419927, 26453379, 26600308, 26888554	139
TNFRSF18	desconocido	PBio	PMID: 16439533, 19162554, 19363449, 22017440, 23432692, 24484736, 25738498	140
TNFRSF4	AR	Base de datos pública	MySql	462

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
TNFRSF8	desconocido	PBio	PMID: 10921351, 15990453, 16472805, 18852356, 20141444, 20378007, 21933041, 23115213, 23307550, 23654079, 24809535, 25999451	463
TNFSF11	AR	Base de datos pública	PMID: 17088646, 17360404, 18281276, 18606301, 19507210, 25992615, 27003757	464
TNFSF12	asociación	Base de datos pública	PMID: 23765059	465
TP53	AD AR	Base de datos pública	MySql; PMID: 11048806, 11079782, 12009037, 19282432, 26870672	466
TRAF3	AD	Base de datos pública	PMID: 20832341, 23887241, 25764117, 25930993, 28049150	467
TRAF6	desconocido	Base de datos pública	PMID: 10215628, 10421844, 25200954, 27808398, 27999438	468
TRAFD1	desconocido	PBio	PMID: 16221674, 18849341, 23913580, 25909814, 25992615, 26283173	141
TREX1	ADAR	Base de datos pública	PMID: 26052098, 27260006, 27821552	469
TRNT1	AR	Base de datos pública	MySql; PMID: 25193871	470
TRPM2	desconocido	PBio	PMID: 9806837, 16585058, 18569867, 19411837, 20107186, 25012489, 25049394, 25088676,26300888,26558786,26679996,26942016,26969190,27405665, 27872485	142
TTC7A	AR	Base de datos pública	MySql; PMID: 27873163	471
UBE2N	desconocido	PBio	PMID: 21512573, 23159053, 24906799, 25343992, 25503582, 25548215, 26085214, 26150489, 26212332, 26518362	145
UNC119	AD	Base de datos pública	MySql	472
UNC13D	AR	Base de datos pública	PMID: 19302049, 21881043, 24916509, 25564401, 25980904, 26454313	473
UNC93B1	asociación	Base de datos pública	PMID: 23810987, 23887241, 25930993, 27873163	474
UNG	AR	Base de datos pública	MySql: PMID: 23765059	475
USP18	AR	Base de datos pública	PMID: 27016605, 27325888, 27801882, 27821552	476
USP20	desconocido	Base de datos pública	PMID: 27801882	477
VAPA	desconocido	PBio	PMID: 9657962, 10523508, 10655491, 11511104, 12931207, 18713837, 23536298, 24076421, 24569996, 25015719	478
VCP	AD	Base de datos pública	PMID: 24246593, 27730320	479
VDAC1	desconocido	PBio	PMID: 10620603, 25874870, 26322231, 26542804, 26616244, 26758954, 26878172	480
VPS13B	AR	Base de datos pública	PMID: 20008220, 20301655	481
VPS45	AR	Base de datos pública	MySql: PMID: 23738510, 24145314, 24164830, 26358756	482

Símbolo del gen de RefSeq	Modelo de enfermedad	Fuente del gen	Anotación de la fuente	Número de gen (GN)
VSTM1	desconocido	PBio	PMID: 22960280, 23436183, 24205237, 25351446, 25887911, 26760041	147
VWA2	desconocido	PBio	PMID: 14506275, 18434322, 21385852, 23443151, 23960233, 26121272	148
WEE1	desconocido	Base de datos pública	PMID: 19903823, 25088202, 26598692, 26881506, 27042682	484
WIPF1	AR	Base de datos pública	PMID: 23765059, 26029204, 26453379	485
XIAP	XLD XLR	Base de datos pública	MySql: PMID: 22365665, 25744037, 26953272	486
YBX1	desconocido	Base de datos pública	PMID: 27042682	487
YWHAZ	desconocido	Ambos	PMID: 25894827, 27042682	488
ZAP70	AD_AR	Base de datos pública	PMID: 18219311, 23494777, 23765059, 24164480, 26454313	489
ZBTB24	AR	Base de datos pública	MySql: PMID: 23486536, 23765059, 26851945, 27098601	490

La Tabla 6 es una lista completa de 419 genes ejemplares (denominados en el presente documento como "genes de PML-419" o "lista de genes de PML-419") interrogados en el presente estudio, junto con información relacionada con el patrón de herencia asumido para el análisis y la razón para la inclusión del gen. Fuentes de genes para la Tabla 6 (encabezado de columna "Fuente del gen"): 1) nominados sobre la base de estar relacionados con la inmunodeficiencia, tal como curados de bases de datos públicas (indicadas por 'Base de datos pública') como PubMed y ClinVar, 2) genes identificados por CNV de PBio ('PBio', véase la Tabla 6 encabezado de columna 'Fuente del gen') de un estudio de descubrimiento de genes CGH de arreglo de genoma completo de 71 casos de PML, o 3) curados de bases de datos públicas e identificados en el estudio de descubrimiento de genes PML de PBio (indicados por 'Ambos'). Se propuso una predisposición genética a PML sobre la base del genoma del huésped; es decir, la(s) variante(s) genética(s) de la línea germinal en el genoma del paciente con PML, en lugar de las variantes genéticas que están presentes en el virus de JC, son la causa de la PML del paciente (Hatchwell, Front Immunol., 6: 216 (2015). Detalles sobre la fuente de los genes en la lista de genes de PML-419 pueden encontrarse en las siguientes fuentes de genes relacionados con inmunodeficiencia e inmunidad: Durandy et al., Nat Rev Immunol., 13(7): 519-33 (2013); Milner et al., Nat Rev Immunol., 13(9): 635-48 (2013); Paciolla et al., Genes Immun., 16(4): 239-46 (2015); Hatchwell, Front Immunol., 6: 216 (2015); Thijssen et al., Nat Commun., 6: 7870 (2015); Chinn et al., Immunol Allergy Clin North Am., 35(4): 671-94 (2015); Zhou et al., Nat Genet., 48(1): 67-73 (2015); Navabi et al., Allergy Asthma Clin Immunol., 12: 27 (2016); y Tsujita et al., J Allergy Clin Immunol. (2016). Los genes de 'MySql' proceden de la base de datos ClinVar. Se realizaron búsquedas en ClinVar utilizando los términos "inmuno deficiencia" e "inmunodeficiencia". Se excluyeron las entradas que describían grandes reordenamientos genómicos, que contenían múltiples genes. Se compiló una lista no redundante de 125 genes combinando el resultado de las dos búsquedas y se depositó en una base de datos MySQL. NOTA: Un subconjunto de estos genes no están marcados como 'MySql' si aparecieron en uno o más de los artículos de revisión de genes inmunes señalados anteriormente. van der Kolk et al., Ann Clin Transl Neurol.; 3(3): 226-32 (2016) fue la fuente del conocido gen de PML BAG3 (véase abajo) y de 28 genes de PML candidatos en base a su conexión con el JCV. Van der Kolk et al., citan un método como el siguiente: "esto último se realizó buscando JCV en NCBI, y seleccionando genes en humanos". Esto produjo 30 genes humanos, 5 de los cuales se solapaban con la lista de genes de PML y 2 genes (HLA-DQB1, HLA-DRB1) fueron excluidos porque los loci HLA son difíciles de interpretar. Los genes ADA, BAG3, BTK, CD40LG, DOCK8, STAT1, WAS y WIPF1 se derivaron de Hatchwell, Front Immunol., 6: 216 (2015) (véase la Tabla 1 para las referencias primarias); van der Kolk et al., Ann Clin Transl Neurol., 3(3): 226-32 (2016); y Zerbe et al., Clin Infect Dis., 62(8): 986-94 (2016). Los genes de PBio se basan en estudios de CNV y un subconjunto se solapa con las listas de genes de la revisión inmunológica (anotados como 'Ambos' en el encabezado de la columna 'Fuente del gen'). Los genes de nivel 1 se utilizaron como posibles soluciones para los casos de PML. La determinación del modelo de enfermedad autosómica dominante (AD), autosómica recesiva (AR), dominante ligada al cromosoma X (XLD) o ligada al cromosoma X (XLR) para cada gen se derivó de los artículos de revisión de inmunodeficiencias y/o de las anotaciones de OMIM. Las entradas marcadas con 'asociación' indican que las variantes se asociaron con una enfermedad relacionada con el sistema inmunitario; 'desconocido' indica que no hay pruebas en la bibliografía de un modelo de AD o AR.

Tabla 7: Causa potencial de la PML en cada paciente en el estudio

ID de la muestra	Etnia	Género	Enfermedad primaria	Símbolo del gen de RefSeq	Tipo de variante	Detalles de la frecuencia variante (específicos de cada etnia)	Frecuencia (Recíproca) (Específica de cada etnia)	SEQ ID
MVGS1116-8a	EUR	F	MS (NZ Rx)	MUELLE8	SNV hom:SNV het	0.499:0.00447	1 de cada 1,792	1147: 1148
MVGS1359	EUR	F	MS (NZ Rx)	IL17F	SNV het	0.00024	1 de cada 4,170	1114
MVGS1368	EUR	F	MS (NZ Rx)	IDO2	SNV hom:SNV het	0.508:0.065	1 de cada 121	1125: 1126
MVGS540-374b	EUR	M	MS (NZ Rx)	SHARPIN	SNV hom	0.00217	1 de cada 461	1142
MVGS540-393b	EUR	F	MS (NZ Rx)	MUELLE8	SNV hom:SNV het:SNV het	0.499:0.00153:0.19 4	1 de cada 5,246	1147: 1154: 11 52
MVGS694-6a	EUR	F	Otra	CHD7	SNV het	0.00028	1 de cada 3,528	1135
MVGS811-13a	EUR	M	VIH	PIK3CD PIK3 CD-AS1	CNV hom	novedosos	0	2
MVGS995-4a	EUR	M	MS (NZ Rx)	EPG5	SNV hom:SNV het	0.495:0.000251	1 de cada 32,224	1279: 1273
PML01	EUR	F	VIH	ITSN2	SNV hom	0.00183	1 de cada 547	1028
PML02	EUR	M	Otra	IKBKB	SNV het	novedosos	0	1127
PML03	EUR	F	MS (NZ Rx)	FPR2	CNV hom	2.23E-06	1 de cada 448,833	140
PML04	EUR	M	VIH	*sin resolver*	n/a	n/a	n/a	n/a
PML05	LAT	M	VIH	TBK1	SNV het	novedosos	0	1203
PML06	AFR	M	VIH	TICAM1	SNV het	0.000777	1 de cada 1,287	1289
PML09	EUR	M	VIH	LIG4	SNV:SNV comp het	0.00399:0.287	1 de cada 3497	1221: 1222
PML 10	EUR	F	VIH	TNFRSF11A	SNV het	novedosos	0	1287
PML 12	LAT	F	VIH	BLM	SNV hom	0.000874	1 de cada 1,144	1235
PML 13	AFR	M	VIH	PLCG2	SNV:SNV comp het	0.00167:0.0187	1 de cada 128,105	1261: 1263
PML 14	EUR	M	VIH	PLCG2	SNV:SNV comp het	0.00998:0.0159	1 de cada 25,259	1261: 1263
PML15	LAT	M	VIH	NOD2	SNV het	novedosos	0	1255
PML 16	AFR	F	VIH	TNFRSF11A	SNV het	novedosos	0	1287

ID de la muestra	Etnia	Género	Enfermedad primaria	Símbolo del gen de RefSeq	Tipo de variante	Detalles de la frecuencia variante (específicos de cada etnia)	Frecuencia (Recíproca) (Específica de cada etnia)	SEQ ID
PML 17	EUR	M	VIH	ZAP70	SNV het	0.00009	1 de cada 11,110	1035
PML18	EUR	M	VIH	*sin resolver*	n/a	n/a	n/a	n/a
PML 19	AFR	M	VIH	ATM	SNV:SNV comp het	0.0479:novedosos	0	1193: 1194
PML20	AFR	M	VIH	NFKB1	SNV het	0.00173	1 de cada 577	1069
PML21	EUR	M	VIH	ZAP70	SNV het	0.0000602	1 de cada 16,623	1034
PML22	EUR	M	VIH	*sin resolver*	n/a	n/a	n/a	n/a
PML23	EUR	F	VIH	DCLRE1C	SNV hom	novedosos	0	1167
PML25	EUR	F	VIH	PLCG2	SNV het	0.000150	1 en 6,672	1259
PML26	EUR	M	VIH	TRAFD1	SNV hom	0.000689	1 de cada 1,451	1208
PML27	EUR	M	VIH	TAP2	SNV hom	0.00837	1 de cada 120	1101
PML28	EUR	F	MS (NZ Rx)	TRPM2	SNV hom	novedosos	0	1311
PML29	AFR	M	VIH	KCTD7 RAB GEF1	CNV hom	0.000387	1 de cada 2,584	65
PML30	EUR	M	VIH	TNFRSF11A	SNV het	novedosos	0	1287
PML31	AFR	F	VIH	DDX58	SNV het	0.000779	1 en 1,283	1157
PML32	EUR	M	VIH	*sin resolver*	n/a	n/a	n/a	n/a
PML33	EUR	M	VIH	TNFRSF11A	SNV het	novedosos	0	1287
PML35	EUR	F	VIH	TNFRSF11A	SNV het	novedosos	0	1287
PML36	AFR	F	VIH	TCIRG1	SNV het	0.002134	1 de cada 469	1184
PML37	AFR	M	VIH	GATA2	SNV het	novedosos	0	1056
PML38	EUR	M	VIH	MALL	CNV hom	3.95E-06	1 de cada 253,036	26
PML39	AFR	M	VIH	*sin resolver*	n/a	n/a	n/a	n/a
PML40	LAT	F	VIH	PNPT1	SNV hom	novedosos	0	1032
PML41	AFR	M	VIH	ZAP70	SNV het	novedosos	0	1036
PML43	EUR	M	VIH	PTPRC	SNV hom	novedosos	0	1020
PML44	EUR	M	VIH	TNFRSF11A	SNV het	novedosos	0	1287
PML45	EUR	F	Otra	TARJETA11	SNV het	0.0024	1 de cada 417	1123

ID de la muestra	Etnia	Género	Enfermedad primaria	Símbolo del gen de RefSeq	Tipo de variante	Detalles de la frecuencia variante (específicos de cada etnia)	Frecuencia (Recíproca) (Específica de cada etnia)	SEQ ID
PML46	LAT	M	VIH	EPG5	SNV:SNV comp het	0.0123:0.436	1 de cada 745	1278: 1279
PML48	EUR	M	VIH	SMAD4	SNV het	0.000901	1 de cada 11,100	1284
PML49	EUR	M	VIH	STIM1	SNV het	novedosos	0	1174
PML50	AFR	M	VIH	NOD2	SNV het	novedosos	0	1256
PML51	EUR	M	VIH	TICAM1	SNV het	0.00265	1 de cada 377	1289
PML52	EUR	F	Otra	*sin resolver*	n/a	n/a	n/a	n/a
PML53	EUR	M	Otra	GFI1	SNV het	0.00003	1 de cada 32,635	1011
PML54	EUR	F	VIH	TNFRSF11A	SNV het	novedosos	0	1287
PML55	EUR	F	VIH	RTEL1	SNV het	0.00326	1 de cada 307	1299
PML56	EUR	M	VIH	TNFRSF11A	SNV het	novedosos	0	1287
PML57	EUR	F	Otra	TRAF3	SNV het	0.00093	1 de cada 1,075	1229
PML58	AFR	M	VIH	MUELLE8	SNV:SNV comp het	0.0575:0.478	1 de cada 146	1146: 1147
PML59	AFR	M	VIH	IFIH1	SNV het	0.00281	1 de cada 356	1040
PML60	EUR	M	VIH	*sin resolver*	n/a	n/a	n/a	n/a
PML61	AFR	F	VIH	TNFRSF11A	SNV het	novedosos	0	1287
PML62	AFR	F	VIH	*sin resolver*	n/a	n/a	n/a	n/a
PML63	AFR	M	VIH	PLCG2	SNV het	0.00195	1 de cada 514	1260
PML64	AFR	M	VIH	PIK3R1	SNV het	novedosos	0	1077
PML65	AFR	M	VIH	ITSN2	CNV hom	0.00313	1 de cada 319	14
PML66	AFR	M	VIH	*sin resolver*	n/a	n/a	n/a	n/a
PML67	EUR	F	MS (NZ Rx)	*sin resolver* (datos de CNV, sin datos de WES)	n/a	n/a	n/a	n/a
PML68	EUR	F	MS (NZ Rx)	LRBA	SNV hom	0.00162	1 de cada 618	1073
PML69	EUR	M	Otra	EGR1 ETF1	CNV hom	0.001	1 de cada 1,005	45
PML72	AFR	F	VIH	NOD2	SNV het	0.004036	1 de cada 248	1252

La Tabla 7 contiene una única solución/explicación genética que es la causa potencial de PML en cada paciente del estudio (71 casos fueron evaluados con CGH de arreglo de genoma completo y 71 también fueron evaluados por secuenciación de exoma completo), con la excepción de 19 casos "no resueltos". Las soluciones se basan en una combinación de variantes CNV y SNV, conectadas por las SEQ ID a las tablas 1, 4 y 5. Para las soluciones de variantes homocigotas o heterocigotas compuestas, las frecuencias poblacionales esperadas se calcularon del siguiente modo:

Frecuencia de población esperada para la variante a (frecuencia p) y la variante b (frecuencia q) = pq/4

Por ejemplo, PML09 tiene 2 variantes, SEQ ID 1221 y 1222, con frecuencias individuales en la población normal de 0.00399, 0.287. La frecuencia esperada en una población normal emparejada étnicamente para esta combinación es (0.00399*0.287)*0.25 = 0.000286283 = 1/3,497.

Los identificadores de Enfermedad primaria de la Tabla 7 son: VIH, infección por el virus de la inmunodeficiencia humana; MS (NZ Rx), esclerosis múltiple tratada con natalizumab; Otras, que incluye una variedad de trastornos/afecciones (MVGS694-6a tenía anemia aplásica, PML02 y PML52 tenían linfoma, PML45 y PML 57 tenían leucemia linfocítica crónica, PML53 tenía sarcoidosis y PML69 es un paciente con trasplante de riñón que recibía belatacept).

Las soluciones se consideraron con base en la presencia de variantes raras (CNV y/o SNV) en o cerca de genes que se enumeran en la Tabla 6. Los modelos de enfermedad autosómica recesiva (AR) y autosómica dominante (AD) comprenden este conjunto de soluciones, basadas en la búsqueda de SNV homocigóticas, CNV homocigóticas, SNV heterocigóticas compuestas o SNV heterocigóticas. Nueve casos de PML de la Tabla 7 se consideraron 'no resueltos' sobre la base de analizar tanto los datos de CNV como los de SNV, y un caso (PML67) se evaluó sólo para CNV ya que los datos de WES no estaban disponibles. En algunos casos, se consideró que un caso no tenía solución óptima (Tabla 7), pero en el Tabla 8 se indicaron soluciones alternativas (véase más abajo).

Para los casos de PML que tenían más de una solución potencial. En estos casos, la 'mejor' solución (Tabla 7) se determinó en función de la rareza de la(s) variante(s) genética(s) y de la fuerza relativa de la biología de los genes de PML-419 (Tabla 6). En la Tabla 8 se presentan soluciones alternativas. Por ejemplo, para el caso de PML MVGS 1116-8a, se encontraron tres soluciones, que afectaban a los genes DOCK8, HIVEP2 y RNF168. En este ejemplo, las SNV heterocigotas compuestas DOCK8 (Tabla 7, SNV hom:SNV het) fueron seleccionados como la mejor solución porque DOCK8 es un gen de PML conocido. En otro ejemplo, el caso de PML MVGS 1359 tiene IL17F (het SNV) listado como la mejor solución en la Tabla 7 porque es más raro que las soluciones alternativas para los genes ATR y STXBP2.

Mientras que algunos pacientes con PML pueden tener múltiples genes/variantes causando y/o contribuyendo a su PML, en muchos pacientes con PML sólo un único gen será la causa primaria análoga a los pacientes diagnosticados con trastornos de inmunodeficiencia primaria. Además de las soluciones alternativas de la Tabla 8, que se basan únicamente en los resultados genéticos de SNV, en la Tabla 1 se presentan otras soluciones alternativas basadas en los resultados genéticos de CNV.

Tabla 8: Soluciones/explicaciones genéticas alternativas como posible causa de la PML en el estudio

ID de la muestra	Símbolo del gen de RefSeq	Tipo de variante	Detalles de la frecuencia de variantes (específicas de cada etnia)	Frecuencia (Reciproca) (Específica de la etnia)	SEQ ID
MVGS1116-8a	HIVEP2	SNV het	novedosos	0	1118
MVGS1116-8a	RNF168	SNV hom:SNV het	0.469:0.00818	1 en 1.041	1063: 1066
MVGS1359	ATR	SNV het	0.00393	1 de cada 254	1058
MVGS1359	STXBP2	SNV het	0.00501	1 de cada 199	1291
MVGS540-374b	MKL1	CNV hom	3.99E-08	1 entre 25.081.515	157
MVGS540-393b	PRKDC	SNV het	0.00097	1 de cada 1.031	1130
MVGS811-13a	CLCN7	SNV het	0.00028	1 de cada 3.571	1239
MVGS995-4a	KAT6B	SNV het	0.00003	1 en 33.357	1169
MVGS995-4a	PRF1	SNV het	0.00243	1 de cada 412	1168

ID de la muestra	Símbolo del gen de RefSeq	Tipo de variante	Detalles de la frecuencia de variantes (específicas de cada etnia)	Frecuencia (Reciproca) (Específica de la etnia)	SEQ ID
PML03	CDKN1B	SNV het	0.00003	1 de cada 32.209	1200
PML05	ATR	SNV het	novedosos	0	1061
PML05	NFKB1	SNV het	0.00501	1 de cada 200	1070
PML06	CHD7	SNV het	0.00797	1 de cada 125	1136
PML06	DOCK8	SNV hom:SNV het	0.478:0.0313	1 de cada 267	1147: 1152
PML09	RIPK3	SNV het	0.00398	1 de cada 251	1227
PML 10	JUN	SNV het	0.00103	1 de cada 968	1009
PML 10	RAG1	SNV het	0.00039	1 de cada 2.566	1179
PML 12	CARD11	SNV het	novedosos	0	1122
PML 12	PRKDC	SNV het	novedosos	0	1128
PML 13	DOCK8	SNV hom:SNV het	0.478:0.0313	1 de cada 267	1147: 1152
PML 13	IRAK4	SNV het	novedosos	0	1202
PML 13	PIK3CD	SNV het	0.00679	1 de cada 147	1000
PML 14	NBN	SNV het	0.0039	1 de cada 256	1138
PML 14	NFKB1	SNV het	novedosos	0	1071
PML 15	ASH1L	SNV:SNV comp het	novedosos:0.0019	0	1016: 1017
PML 15	CHD7	SNV het	0.00176	1 de cada 568	1133
PML 15	HIVEP2	SNV het	novedosos	0	1116
PML 15	STIM1	SNV het	0.00587	1 de cada 170	1175
PML 16	TBK1	SNV het	novedosos	0	1204
PML 16	TLR3	SNV het	0.00136	1 de cada 738	1076
PML 17	APOL1	SNV het	0.0021	1 de cada 475	1327
PML18	PKHD1	SNV hom:SNV het	0.498:0.0471	1 de cada 171	1104: 1107
PML 19	DOCK8	SNV:SNV comp het	0.0575:0.478	1 de cada 146	1146: 1147
PML 19	IFIH1	SNV het	0.00444	1 de cada 225	1041
PML20	JUN	SNV het	0.00535	1 de cada 187	1010
PML21	PRKCH	SNV het	novedosos	0	1228
PML21	PSTPIP1	SNV het	0.00093	1 de cada 1.074	1232
PML21	RAG2	SNV het	novedosos	0	1182
PML22	RIPK3	SNV hom	0.00309	1 de cada 324	1226
PML22	VPS45	SNV het	0.00114	1 de cada 878	1014
PML23	NOD2	SNV het	novedosos	0	1251
PML23	RAG1	SNV het	0.00003	1 de cada 33.317	1180
PML28	PKHD1	SNV hom:SNV het	0.498:0.0471	1 de cada 171	1104: 1107
PML28	TNFRSF13B	SNV het	0.00929	1 de cada 108	1267

ID de la muestra	Símbolo del gen de RefSeq	Tipo de variante	Detalles de la frecuencia de variantes (específicas de cada etnia)	Frecuencia (Reciproca) (Específica de la etnia)	SEQ ID
PML30	RTEL1	SNV het	0.000124	1 de cada 8.068	1300
PML31	AP3B1	SNV het	novedosos	0	1084
PML31	PRKDC	SNV het	novedosos	0	1129
PML33	STIM2	SNV het	0.00003	1 de cada 32.688	1068
PML33	TLR3	SNV hom:SNV het	0.413:0.00435	1 de cada 2.227	1075: 1074
PML33	TLR4	SNV hom:SNV hom	0.00283:0.00285	1 de cada 354	1161: 1160
PML35	PRKCB	SNV het	0.00276	1 de cada 362	1247
PML36	NOD2	SNV het	0.00871	1 de cada 115	1254
PML36	PIK3CD	SNV het	0.00679	1 de cada 147	1000
PML37	AP3B1	SNV het	novedosos	0	1080
PML37	ATR	SNV het	0.00038	1 de cada 2.601	1059
PML37	WEE1	SNV het	0.00825	1 de cada 121	1177
PML38	MYD88	SNV het	novedosos	0	1051
PML40	MCEE	SNV hom	0.01	1 de cada 100	1033
PML41	AP3B1	SNV het	0.00173	1 de cada 577	1082
PML41	CHD7	SNV het	novedosos	0	1137
PML41	DOCK8	SNV:SNV comp het	0.0575:0.478	1 de cada 146	1146: 1147
PML41	POLE	SNV hom	0.00019	1 de cada 5.203	1219
PML41	RNF 168	SNV:SNV comp het	0.412:novedosos	0	1063: 1062
PML43	DOCK8	SNV hom:SNV het	0.499:novedosos	0	1147: 1150
PML44	DCLRE1C	SNV hom:SNV hom	0.0287:0.00575	1 de cada 174	1166: 1165
PML44	GFI1	SNV het	0.00708	1 de cada 141	1012
PML45	POLA1	SNV het	novedosos	0	1328
PML46	AP3B1	SNV het	0.00587	1 de cada 170	1082
PML46	IL21R	SNV het	0.00573	1 de cada 175	1248
PML46	PRKDC	SNV het	0.00017	1 de cada 5.781	1131
PML48	TNFRSF11A	SNV het	0.00233	1 de cada 429	1286
PML49	DCLRE1C	SNV hom:SNV hom	0.00575:0.0287	1 de cada 174	1166: 1165
PML49	PTEN	SNV het	novedosos	0	1171
PML49	RIPK1	SNV het	0.00090	1 de cada 1.112	1090
PML50	AP3B1	SNV het	0.00387	1 de cada 259	1078
PML50	PIAS2	SNV het	0.00357	1 de cada 280	1283
PML50	STXBP2	SNV het	0.00038	1 de cada 2.598	1290
PML52	GFI1	SNV het	0.00708	1 de cada 141	1012
PML53	IL1B	SNV het	novedosos	0	1037

ID de la muestra	Símbolo del gen de RefSeq	Tipo de variante	Detalles de la frecuencia de variantes (específicas de cada etnia)	Frecuencia (Reciproca) (Específica de la etnia)	SEQ ID
PML53	STXBP2	SNV het	0.00501	1 de cada 199	1291
PML54	EPG5	SNV:SNV comp het	0.0638:0.495	1 de cada 127	1278: 1279
PML54	IFNGR2	SNV het	0.00009	1 en 11.096	1304
PML54	RAG1	SNV het	0.00003	1 de cada 33.352	1178
PML54	RAG2	SNV het	novedosos	0	1183
PML57	PIAS1	SNV het	novedosos	0	1231
PML57	PKHD1	SNV hom:SNV het	0.498:0.0471	1 de cada 171	1104: 1107
PML57	SKIV2L	SNV hom:SNV hom:SNV het	0.157:0.214:0.0471	1 de cada 538	1098: 1100: 1099
PML58	GFI1	SNV het	0.00144	1 de cada 693	1012
PML59	IFNLR1	SNV het	novedosos	0	1002
PML59	NOD2	SNV het	0.00404	1 de cada 248	1252
PML59	NRIP1	SNV hom	0.00711	1 de cada 141	1301
PML59	RAD51	SNV het	0.00865	1 de cada 116	1230
PML60	MAPK3	SNV het	novedosos	0	1250
PML60	TP53	SNV het	0.00048	1 de cada 2.085	1266
PML61	GATA2	SNV het	0.00024	1 de cada 4.139	1057
PML61	PTPRC	SNV hom	novedosos	0	1019
PML61	TNFRSF8	SNV het	novedosos	0	1001
PML62	PRKCD	SNV het	novedosos	0	1054
PML63	HTR2A	SNV hom	0.00519	1 de cada 193	1220
PML63	MAPK3	SNV het	0.00193	1 de cada 518	1249
PML64	PLCG2	SNV het	0.00044	1 de cada 2.276	1264
PML64	WEE1	SNV het	novedosos	0	1176
PML65	IRAK4	SNV het	0.00118	1 de cada 850	1201
PML66	PIK3CD	SNV het	0.00679	1 de cada 147	1000
PML68	RAG1	SNV het	0.00586	1 de cada 171	1181
PML72	CARD11	SNV het	0.00242	1 de cada 413	1121
PML72	HIVEP1	SNV hom	0.00164	1 de cada 610	1092
PML72	IFIH1	SNV het	0.00843	1 de cada 119	1043

La Tabla 8 contiene información análoga a la de la Tabla 7, con la excepción de que Etnia, Sexo y Enfermedad primaria no se repiten. La Tabla 8 contiene soluciones/explicaciones genéticas alternativas como causa potencial de la PML para los pacientes del estudio (71 casos se evaluaron con CGH de arreglo de genoma completo y 70 se evaluaron también mediante secuenciación del exoma completo). Las soluciones de la Tabla 8 también se refieren a casos concretos y representan soluciones alternativas secundarias para los casos enumerados (utilizando los mismos criterios empleados para identificar posibles soluciones en la Tabla 7). En otras palabras, para algunos individuos se identificó más de una solución razonable y, aunque las de la Tabla 7 se consideran las más probables, las de la Tabla 8 también son soluciones potenciales. Los expertos en la materia pueden apreciar que los datos adicionales sobre nuevos casos de PML, pacientes con trastornos de inmunodeficiencia de base genética o estudios funcionales sobre un gen determinado pueden respaldar la

selección de una solución de la Tabla 8 como la 'mejor' solución única (es decir, una solución actual de la Tabla 7 podría considerarse en su lugar como una solución de la Tabla 8, y viceversa).

Tabla 9: Pares de SNV que afectan al mismo gen

Muestra ID	Símbolo del gen de RefSeq	Tipo de variante	Cromosoma	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
MVGS1359	TTC7A	SNV het	2	47273468	A	G	K252R	0.00684	1030
MVGS1359	TTC7A	SNV het	2	47277182	T	C	S318P	0.00683	1031
MVGS1368	RNF 168	SNV het	3	196199204	G	T	P401Q	0.46947	1063
MVGS1368	RNF 168	SNV het	3	196210764	T	C	n/a	0.00003	1065
MVGS1368	TLR4	SNV het	9	120475302	A	G	D259G	0.10251	1160
MVGS1368	TLR4	SNV het	9	120475602	C	T	T359I	0.10560	1161
MVGS811-13a	HIVEP1	SNV het	6	12121113	C	T	P362L	0.00024	1091
MVGS811-13a	HIVEP1	SNV het	6	12123538	G	T	K1170N	0.08730	1093
MVGS995-4a	EEE1	SNV het	12	93196332	C	T	E840K	0.01949	1206
MVGS995-4a	EEE1	SNV het	12	93205148	T	G	E702D	0.00003	1207
PML02	RBFOX1	SNV het	16	7759119	G	A	G326S	0.00504	1245
PML02	RBFOX1	SNV het	16	7759496	C	T	P401S	novedosos	1246
PML04	POLE	SNV het	12	133220526	T	C	N1369S	0.22363	1213
PML04	POLE	SNV het	12	133237658	T	G	Q766P	novedosos	1215
PML05	TLR4	SNV het	9	120475302	A	G	D259G	0.04628	1160
PML05	TLR4	SNV het	9	120475602	C	T	T359I	0.04180	1161
PML05	POLE	SNV het	12	133220526	T	C	N1369S	0.12669	1213
PML05	POLE	SNV het	12	133252406	C	A	A121S	novedosos	1217
PML10	TLR4	SNV het	9	120475302	A	G	D259G	0.10251	1160
PML10	TLR4	SNV het	9	120475602	C	T	T359I	0.10560	1161
PML12	IDO2	SNV het	8	39840234	A	G	I127V	0.38971	1124
PML12	IDO2	SNV het	8	39862881	C	T	R23 5 W	0.50282	1125
PML12	IDO2	SNV het	8	39862893	T	A	S239T	0.02384	1126
PML13	STX11	SNV het	6	144508353	G	A	V197M	novedosos	1119
PML13	STX11	SNV het	6	144508563	G	A	V267M	0.00202	1120
PML13	DCLRE1C	SNV het	10	14974905	T	C	H123R	0.16298	1165

Muestra ID	Símbolo del gen de RefSeq	Tipo de variante	Cromosoma	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
PML13	DCLRE1C	SNV het	10	14976727	G	C	P171R	0.22295	1166
PML13	EPG5	SNV het	18	43497710	A	G	V1058A	0.42740	1279
PML13	EPG5	SNV het	18	43531186	C	T	S424N	0.00600	1282
PML14	ATM	SNV het	11	108117787	C	T	S333F	0.00280	1188
PML14	ATM	SNV het	11	108175462	G	A	D1853N	0.24654	1193
PML14	TRPM2	SNV het	21	45815425	C	G	I621M	novedosos	1313
PML14	TRPM2	SNV het	21	45845699	G	A	V1242M	0.00537	1321
PML16	TLR3	SNV het	4	187004074	C	T	L135F	0.12378	1075
PML16	TLR3	SNV het	4	187005854	A	C	I571L	0.00136	1076
PML16	HIVEP1	SNV het	6	12121113	C	T	P362L	0.07856	1091
PML16	HIVEP1	SNV het	6	12162068	C	T	S160F	0.01979	1096
PML16	PKHD1	SNV het	6	51483961	T	C	Q4048R	0.50029	1104
PML16	PKHD1	SNV het	6	51747943	T	A	D2433V	0.07153	1112
PML16	POLE	SNV het	12	133209020	G	C	Q2044E	novedosos	1212
PML16	POLE	SNV het	12	133220526	T	C	N1369S	0.24889	1213
PML17	RNF 168	SNV het	3	196199204	G	T	P401Q	0.46947	1063
PML17	RNF 168	SNV het	3	196210704	G	A	P206L	0.00003	1064
PML17	HIVEP1	SNV het	6	12123538	G	T	K1170N	0.08730	1093
PML17	HIVEP1	SNV het	6	12125232	C	T	S1735F	0.00027	1095
PML17	PKHD1	SNV het	6	51483961	T	C	Q4048R	0.49837	1104
PML17	PKHD1	SNV het	6	51497503	C	A	R3842L	0.04707	1107
PM17	DCLRE1C	SNV het	10	14974905	T	C	H123R	0.27332	1165
PML17	DCLRE1C	SNV het	10	14976727	G	C	P171R	0.13896	1166
PML17	ATM	SNV het	11	108119823	T	C	V410A	0.00643	1189
PML17	ATM	SNV het	11	108175462	G	A	D1853N	0.24654	1193
PML17	EPG5	SNV het	18	43464763	C	T	G1708D	0.00013	1274
PML17	EPG5	SNV het	18	43497710	A	G	V1058A	0.49513	1279

Muestra ID	Símbolo del gen de RefSeq	Tipo de variante	Cromosoma	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
PML18	TLR4	SNV het	9	120475302	A	G	D259G	0.10251	1160
PML18	TLR4	SNV het	9	120475602	C	T	T359I	0.10560	1161
PML20	AK2	SNV het	1	33476435	C	A	n/a	novedosos	1003
PML20	AK2	SNV het	1	33478900	T	A	Y159F	0.04954	1004
PML20	HIVEP1	SNV het	6	12124215	C	T	P1396L	0.06774	1094
PML20	HIVEP1	SNV het	6	12163657	C	T	P2374S	0.06733	1097
PML20	KANK1	SNV het	9	711359	C	T	S198F	0.11985	1155
PML20	KANK1	SNV het	9	713132	G	T	G631V	0.00136	1156
PML21	DOCK8	SNV het	9	286593	C	A	P29T	0.49889	1147
PML21	DOCK8	SNV het	9	286593	C	A	P29T	0.49889	1147
PML21	DOCK8	SNV het	9	312134	G	A	E169K	0.06358	1149
PML21	DOCK8	SNV het	9	312134	G	A	E169K	0.06358	1149
PML21	TLR4	SNV het	9	120475302	A	G	D259G	0.10251	1160
PML21	TLR4	SNV het	9	120475302	A	G	D259G	0.10251	1160
PML21	TLR4	SNV het	9	120475602	C	T	T359I	0.10560	1161
PML21	TLR4	SNV het	9	120475602	C	T	T359I	0.10560	1161
PML21	ATM	SNV het	11	108138003	T	C	F858L	0.02864	1191
PML21	ATM	SNV het	11	108138003	T	C	F858L	0.02864	1191
PML21	ATM	SNV het	11	108143456	C	G	P1054R	0.05069	1192
PML21	ATM	SNV het	11	108143456	C	G	P1054R	0.05069	1192
PML21	TRPM2	SNV het	21	45786650	C	T	S146F	0.00072	1305
PML21	TRPM2	SNV het	21	45786650	C	T	S146F	0.00072	1305
PML21	TRPM2	SNV het	21	45820196	C	T	R735C	0.10374	1314
PML21	TRPM2	SNV het	21	45820196	C	T	R735C	0.10374	1314
PML22	SKIV2L	SNV het	6	31928306	A	G	Q151R	0.15759	1098
PML22	SKIV2L	SNV het	6	31935750	G	A	V724M	0.04718	1099
PML22	SKIV2L	SNV het	6	31936679	C	T	A1071V	0.21419	1100

Muestra ID	Símbolo del gen de RefSeq	Tipo de variante	Cromosoma	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
PML22	DOCK8	SNV het	9	286593	C	A	P29T	0.49889	1147
PML22	DOCK8	SNV het	9	304628	G	A	R151Q	0.00447	1148
PML22	GDPD4	SNV het	11	76954833	G	A	H383Y	0.44867	1186
PML22	GDPD4	SNV het	11	76979511	A	G	I233T	0.00504	1187
PML22	ATM	SNV het	11	108117787	C	T	S333F	0.00280	1188
PML22	ATM	SNV het	11	108175462	G	A	D1853N	0.24654	1193
PML22	BLM	SNV het	15	91306241	G	A	R643H	0.00799	1233
PML22	BLM	SNV het	15	91341543	A	C	N1112H	novedosos	1238
PML23	PKHD1	SNV het	6	51483961	T	C	Q4048R	0.49837	1104
PML23	PKHD1	SNV het	6	51497503	C	A	R3842L	0.04707	1107
PML23	SHARPIN	SNV het	8	145154222	G	A	P294S	0.08789	1142
PML23	SHARPIN	SNV het	8	145154257	C	G	S282T	0.14880	1144
PML23	DOCK8	SNV het	9	286491	G	A	D63N	0.27362	1146
PML23	DOCK8	SNV het	9	334277	G	A	R325H	0.00015	1151
PML25	SKIV2L	SNV het	6	31928306	A	G	Q151R	0.15759	1098
PML25	SKIV2L	SNV het	6	31935750	G	A	V724M	0.04718	1099
PML25	SKIV2L	SNV het	6	31936679	C	T	A1071V	0.21419	1100
PML25	PKHD1	SNV het	6	51483961	T	C	Q4048R	0.49837	1104
PML25	PKHD1	SNV het	6	51524409	G	T	S3505R	0.02049	1109
PML25	EPG5	SNV het	18	43445601	T	G	I174L	novedosos	1270
PML25	EPG5	SNV het	18	43531186	C	T	S424N	0.02391	1282
PML27	LYST	SNV het	1	235897907	C	T	G2804D	0.00114	1024
PML27	LYST	SNV het	1	235909815	A	T	F165Y	0.00102	1025
PML27	EPG5	SNV het	18	43445601	T	G	I174L	novedosos	1270
PML27	EPG5	SNV het	18	43497710	A	G	V1058A	0.49513	1279
PML29	LIG1	SNV het	19	48631258	G	A	T546I	0.07515	1292
PML29	LIG1	SNV het	19	48639022	T	C	M412V	0.05385	1293

Muestra ID	Símbolo del gen de RefSeq	Tipo de variante	Cromosoma	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
PML30	DCLRE1C	SNV het	10	14974905	T	C	H123R	0.27332	1165
PML30	DCLRE1C	SNV het	10	14976727	G	C	P171R	0.13896	1166
PML30	ATM	SNV het	11	108138003	T	C	F858L	0.02864	1191
PML30	ATM	SNV het	11	108143456	C	G	P1054R	0.05069	1192
PML30	ATM	SNV het	11	108186610	G	A	G2023R	0.00465	1195
PML31	LYST	SNV het	1	235897907	C	T	G2804D	0.23000	1024
PML31	LYST	SNV het	1	235909815	A	T	F165Y	0.15155	1025
PML31	PKHD1	SNV het	6	51483961	T	C	Q4048R	0.50029	1104
PML31	PKHD1	SNV het	6	51524339	C	G	E3529Q	0.07244	1108
PML31	PKHD1	SNV het	6	51747943	T	A	D2433V	0.07153	1112
PML31	PKHD1	SNV het	6	51798908	C	T	G2041S	0.00173	1113
PML32	PKHD1	SNV het	6	51483961	T	C	Q4048R	0.49837	1104
PML32	PKHD1	SNV het	6	51491885	G	A	Q3899*	novedosos	1106
PML32	EPG5	SNV het	18	43496539	G	A	S1083L	0.06375	1278
PML32	EPG5	SNV het	18	43497710	A	G	V1058A	0.49513	1279
PML32	EPG5	SNV het	18	43529551	C	T	V466M	0.00006	1281
PML33	AK2	SNV het	1	33476435	C	A	n/a	novedosos	1003
PML33	AK2	SNV het	1	33487007	C	T	S129N	0.01100	1005
PML33	EPG5	SNV het	18	43497710	A	G	V1058A	0.49513	1279
PML33	EPG5	SNV het	18	43523240	C	T	M610I	0.00066	1280
PML35	RNF 168	SNV het	3	196199204	G	T	P401Q	0.46947	1063
PML35	RNF 168	SNV het	3	196214320	C	T	E170K	0.00818	1066
PML36	IDO2	SNV het	8	39862881	C	T	R23 5 W	0.46108	1125
PML36	IDO2	SNV het	8	39862893	T	A	S239T	0.01135	1126
PML39	LYST	SNV het	1	235897907	C	T	G2804D	0.23000	1024
PML39	LYST	SNV het	1	235909815	A	T	F165Y	0.15155	1025
PML39	NHEJ1	SNV het	2	219942026	T	A	Q181L	0.06324	1047

Muestra ID	Símbolo del gen de RefSeq	Tipo de variante	Cromosoma	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
PML39	NHEJ1	SNV het	2	220023045	C	T	A14T	0.23543	1048
PML40	ATM	SNV het	11	108186631	A	G	I2030V	0.00173	1196
PML40	ATM	SNV het	11	108186631	A	G	I2030V	0.03446	1196
PML40	ATM	SNV het	11	108198384	C	G	L2330V	0.00035	1197
PML40	ATM	SNV het	11	108198384	C	G	L2330V	0.00491	1197
PML41	PKHD1	SNV het	6	51483961	T	C	Q4048R	0.50029	1104
PML41	PKHD1	SNV het	6	51497503	C	A	R3842L	0.00654	1107
PML41	IDO2	SNV het	8	39840234	A	G	I127V	0.06350	1124
PML41	IDO2	SNV het	8	39862881	C	T	R23 5 W	0.46108	1125
PML45	VPS13B	SNV het	8	100791158	G	A	E2560K	0.00964	1140
PML45	VPS13B	SNV het	8	100865941	G	A	A3442T	novedosos	1141
PML48	EPG5	SNV het	18	43497710	A	G	V1058A	0.49513	1279
PML48	EPG5	SNV het	18	43531186	C	T	S424N	0.02391	1282
PML51	TRPM2	SNV het	21	45826486	G	A	V914I	novedosos	1315
PML51	TRPM2	SNV het	21	45855099	C	T	R1300W	0.00021	1322
PML53	EPG5	SNV het	18	43445580	C	T	D181N	novedosos	1269
PML53	EPG5	SNV het	18	43497710	A	G	V1058A	0.49513	1279
PML56	TLR4	SNV het	9	120475302	A	G	D259G	0.10251	1160
PML56	TLR4	SNV het	9	120475602	C	T	T359I	0.10560	1161
PML56	DCLRE1C	SNV het	10	14974905	T	C	H123R	0.27332	1165
PML56	DCLRE1C	SNV het	10	14976727	G	C	P171R	0.13896	1166
PML57	CSF3R	SNV het	1	36932047	C	T	E359K	0.01706	1006
PML57	CSF3R	SNV het	1	36933715	A	G	Y113H	0.00087	1007
PML57	TLR4	SNV het	9	120475302	A	G	D259G	0.10251	1160
PML57	TLR4	SNV het	9	120475602	C	T	T359I	0.10560	1161
PML57	ATM	SNV het	11	108138003	T	C	F858L	0.02864	1191
PML57	ATM	SNV het	11	108143456	C	G	P1054R	0.05069	1192

Muestra ID	Símbolo del gen de RefSeq	Tipo de variante	Cromosoma	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
PML57	ATM	SNV het	11	108175462	G	A	D1853N	0.24654	1193
PML58	DOCK8	SNV het	9	399233	A	G	N1002D	0.19737	1153
PML58	DCLRE1C	SNV het	10	14974905	T	C	H123R	0.16298	1165
PML58	DCLRE1C	SNV het	10	14976727	G	C	P171R	0.22295	1166
PML58	DNMT3B	SNV het	20	31383307	G	A	G311S	0.00192	1296
PML58	DNMT3B	SNV het	20	31384614	G	T	G343V	novedosos	1297
PML59	LYST	SNV het	1	235897907	C	T	G2804D	0.23000	1024
PML59	LYST	SNV het	1	235897907	C	T	G2804D	0.23000	1024
PML59	LYST	SNV het	1	235909815	A	T	F165Y	0.15155	1025
PML59	LYST	SNV het	1	235909815	A	T	F165Y	0.15155	1025
PML59	LIG1	SNV het	19	48631258	G	A	T546I	0.07515	1292
PML59	LIG1	SNV het	19	48631258	G	A	T546I	0.07515	1292
PML59	LIG1	SNV het	19	48639022	T	C	M412V	0.05385	1293
PML59	LIG1	SNV het	19	48639022	T	C	M412V	0.05385	1293
PML60	DCLRE1C	SNV het	10	14974905	T	C	H123R	0.27332	1165
PML60	DCLRE1C	SNV het	10	14976727	G	C	P171R	0.13896	1166
PML60	POLE	SNV het	12	133202816	C	T	E2113K	0.04686	1211
PML60	POLE	SNV het	12	133220526	T	C	N1369S	0.22363	1213
PML62	TLR4	SNV het	9	120475302	A	G	D259G	0.13066	1160
PML62	TLR4	SNV het	9	120475602	C	T	T359I	0.02672	1161
PML63	HIVEP1	SNV het	6	12124215	C	T	P1396L	0.06774	1094
PML63	HIVEP1	SNV het	6	12163657	C	T	P2374S	0.06733	1097
PML63	PLCG2	SNV het	16	81942175	A	G	N571S	0.01870	1263
PML63	TRPM2	SNV het	21	45795833	G	T	V297L	0.00097	1306
PML63	TRPM2	SNV het	21	45815307	T	C	V582A	0.00724	1310
PML64	DNER	SNV het	2	230231632	C	T	D687N	0.00058	1049
PML64	DNER	SNV het	2	230450646	T	A	T259S	0.00692	1050

Muestra ID	Símbolo del gen de RefSeq	Tipo de variante	Cromosoma	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
PML64	IDO2	SNV het	8	39862881	C	T	R235W	0.46108	1125
PML64	IDO2	SNV het	8	39862893	T	A	S239T	0.01135	1126
PML65	POLE	SNV het	12	133201381	T	A	I2228F	0.00232	1210
PML65	POLE	SNV het	12	133253971	C	T	R233Q	0.02037	1218
PML66	PKHD1	SNV het	6	51483961	T	C	Q4048R	0.50029	1104
PML66	PKHD1	SNV het	6	51612746	G	A	S3223L	0.00000	1110
PML66	PKHD1	SNV het	6	51712769	T	C	T2641A	0.04812	1111
PML66	EPG5	SNV het	18	43456296	C	T	R1985Q	0.07733	1271
PML66	EPG5	SNV het	18	43497710	A	G	V1058A	0.42740	1279
PML68	DCLRE1C	SNV het	10	14974905	T	C	H123R	0.27332	1165
PML68	DCLRE1C	SNV het	10	14976727	G	C	P171R	0.13896	1166
PML72	PSMB8	SNV het	6	32810794	T	A	T70S	0.04224	1102
PML72	PSMB8	SNV het	6	32811752	C	T	G8R	0.04845	1103
PML72	POLE	SNV het	12	133220526	T	C	N1369S	0.24889	1213
PML72	POLE	SNV het	12	133245026	G	A	P477S	0.02332	1216
PML72	RBFOX1	SNV het	16	7568296	C	T	P102S	0.00692	1242
PML72	RBFOX1	SNV het	16	7703891	A	G	T235A	novedosos	1243

La Tabla 9 enumera, para cada caso (en múltiples filas), las variantes para las que no fue posible, utilizando los datos de secuenciación del exoma completo (WES) disponibles, determinar la fase (es decir, si dos variantes están en cis - en el mismo cromosoma - o trans - en cromosomas opuestos). Determinar la fase es una consideración importante cuando se trata de trastornos que se evalúan sobre una base autosómica recesiva (AR). Si se sabe que están presentes dos variantes, pero es imposible determinar si están en cis o en trans, entonces es imposible concluir que ambas copias del gen están afectadas, a diferencia de sólo una (aunque con 2 variantes). Este problema no se plantea en el caso de las variantes homocigotas, para las que es obvio que las variantes deben estar en trans (es decir, sólo es un problema para las variantes no idénticas). Todas las coordenadas del genoma se basan en la construcción hg19.

En resumen, la Tabla 9 enumera todas las soluciones de SNV heterocigotas compuestas a nivel de caso sin desfasar, que podrían representar otras soluciones a nivel de caso, si desfasar hubiera sido posible. Además, puede ser apreciado por los expertos en la materia que las soluciones sin desfasar reportadas en la Tabla 9 (2 SNV het por gen) o la Tabla 10 (véase abajo, que reporta SNV het en pacientes que también tienen una CNV reportada en la Tabla 1) pueden potencialmente causar o contribuir a la PML del paciente si el análisis genético de seguimiento revela que el par de variantes están en diferentes alelos (es decir, cada copia del gen impactada por una variante). Las variantes reportadas en las Tablas 1, 9 o 10 también pueden resultar ser significativamente deletéreas por sí solas (p. ej., en estudios funcionales en células derivadas de pacientes, modelos animales, etc.) y constituir así una solución de modelo de AD (es decir, los genes actualmente enumerados como modelo 'AR' en la Tabla 6) pueden ser causales o contribuir a la enfermedad a través de un modelo AD o AR, al igual que varios genes ya conocidos como AD o AR (Tabla 6, modelo de enfermedad 'AD AR').

Tabla 10: SNV encontradas en genes sospechosos de estar afectados por CNV adquiridas

ID de la muestra	Símbolo del gen de RefSeq	Tipo de variante	Chr	Posición	Alelo de referencia	Alelo Alt	Cambio de aminoácidos	Detalles de la frecuencia de variantes (específicas de cada etnia)	SEQ ID NO
MVGS811-13a	NRIP1	SNV het	21	16338814	T	C	N567S	0.00060	1301
MVGS995-4a	VWA2	SNV het	10	116045796	G	A	V366M	0.02392	1173
PML01	PKHD1	SNV het	6	51497503	C	A	R3842L	0.04707	1107
PML01	PKHD1	SNV het	6	51483961	T	C	Q4046R	0.49637	1104
PML02	DUSP16	SNV het	12	12673965	G	A	T23M	0.00015	1189
PML39	SALL2	SNV het	14	22004996	G	T	S13R	0.00231	1225
PML51	JMY	SNV het	5	78596018	G	C	D524H	novedosos	1086
PML65	SALL2	SNV het	14	21992397	T	C	S347G	0.07709	1223

NOTA: Se trata de SNV het potencialmente heterocigotos compuestos con una CNV en el alelo. Véase la descripción en el texto.
La SNV de DUSP16 (chr12:12673965) estaba en trans con una eliminación de chr12 de DUSP16 en este paciente (PML02), cuyo diagnóstico primario fue linfoma.

La Tabla 10 es una lista de todas las SNV heterocigotas que son potencialmente heterocigotas compuestas con una CNV en el alelo. Véase el texto para una explicación más completa. Todas las coordenadas del genoma se basan en la construcción hg19.

Tabla 11: Clave que relaciona la ID de la muestra para los casos de PML con los números de ID del caso de PML

ID de la muestra	ID del caso de PML
MVGS1116-8a	3006
MVGS1359	3117
MVGS1368	3118
MVGS540-374b	3005
MVGS540-393b	3004
MVGS694-6a	3007
MVGS811-13a	3009
MVGS995-4a	3010
PML01	3127
PML02	3126
PML03	3155
PML04	3156
PML05	3125
PML06	3124
PML09	3132
PML10	3157
PML12	3159
PML13	3160
PML14	3161
PML15	3194
PML16	3163
PML17	3140
PML18	3141
PML19	3164
PML20	3143
PML21	3144
PML22	3145
PML23	3165
PML25	3166
PML26	3167
PML27	3168
PML28	3151
PML29	3152
PML30	3153
PML31	3154
PML32	3169
PML33	3170

ID de la muestra	ID del caso de PML
PML35	3171
PML36	3172
PML37	3173
PML38	3174
PML39	3175
PML40	3273
PML41	3177
PML43	3178
PML44	3179
PML45	3180
PML46	3196
PML48	3197
PML49	3183
PML50	3198
PML51	3185
PML52	3186
PML53	3187
PML54	3188
PML55	3189
PML56	3190
PML57	3191
PML58	3192
PML59	3193
PML60	3199
PML61	3200
PML62	3201
PML63	3202
PML64	3203
PML65	3204
PML66	3205
PML67	3277
PML68	3278
PML69	3279
PML72	3282
PML70_control	3280
PML71_control	3281
PML73_control	3283
PML74_control	3284
PML75_control	3285
PML76_control	3286

La Tabla 11 provee la ID de la muestra y la ID del caso de PML (ID experimental para datos de CGH) para 77 'casos de PML' (incluye 6 casos de VIH sin PML listados como controles).

Tabla 12: Lista no redundante de variantes de transcripción que corresponden al conjunto de genes de los que no se han reportado 'soluciones' de CNV en los 71 casos de PML

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
ACADM	NM_000016	Acil-CoA deshidrogenasa, cadena recta de C-4 a C-12 (ACADM) de Homo sapiens, variante de transcripción 1, ARNm.	1500
ACADM	NM_001127328	Receptor atípico de quimioquinas 1 (grupo sanguíneo Duffy) (ACKR1) de Homo sapiens, variante de transcripción 1, ARNm.	1501
ACKR1	NM_002036	Receptor atípico de quimioquinas 1 (grupo sanguíneo Duffy) (ACKR1) de Homo sapiens, variante de transcripción 2, ARNm.	1502
ACKR1	NM_001122951	Receptor atípico de quimioquinas 1 de Homo sapiens isoforma a	1503
ACP5	NM_001611	Fosfatasa ácida 5, resistente al tartrato (ACP5) de Homo sapiens, variante de transcripción 4, ARNm.	1504
ACP5	NM_001111034	Fosfatasa ácida 5, resistente al tartrato (ACP5) de Homo sapiens, variante de transcripción 2, ARNm.	1505
ACP5	NM_001111035	Fosfatasa ácida 5, resistente al tartrato (ACP5) de Homo sapiens, variante de transcripción 1, ARNm.	1506
ACP5	NM_001111036	Fosfatasa ácida 5, resistente al tartrato (ACP5) de Homo sapiens, variante de transcripción 3, ARNm.	1507
ADAR	NM_001111	Adenosina desaminasa, específica de ARN (ADAR) de Homo sapiens, variante de transcripción 1, ARNm.	1508
ADAR	NM_015840	Adenosina desaminasa, específica de ARN (ADAR) de Homo sapiens, variante de transcripción 2, ARNm.	1509
ADAR	NM_015841	Adenosina desaminasa, específica de ARN (ADAR) de Homo sapiens, variante de transcripción 3, ARNm.	1510
ADAR	NM_001025107	Adenosina desaminasa, específica de ARN (ADAR) de Homo sapiens, variante de transcripción 4, ARNm.	1511
ADAR	NM_001193495	Adenosina desaminasa, específica de ARN (ADAR) de Homo sapiens, variante de transcripción 5, ARNm.	1512
ADK	NM_001202450	Adenosina quinasa (ADK) de Homo sapiens, variante de transcripción 4, ARNm.	1513
ADK	NM_006721	Adenosina quinasa (ADK) de Homo sapiens, variante de transcripción 2, ARNm.	1514
ADK	NM_001123	Adenosina quinasa (ADK) de Homo sapiens, variante de transcripción 1, ARNm.	1515
ADK	NM_001202449	Adenosina quinasa (ADK) de Homo sapiens, variante de transcripción 3, ARNm.	1516
AICDA	NM_020661	Citidina desaminasa inducida por activación (AICDA) de Homo sapiens, ARNm.	1517
AK2	NM_001199199	Adenilato quinasa 2 (AK2) de Homo sapiens, variante de transcripción 3, ARNm.	1518
AK2	NM_013411	Adenilato quinasa 2 (AK2) de Homo sapiens, variante de transcripción 2, ARNm.	1519
AK2	NM_001625	Adenilato quinasa 2 (AK2) de Homo sapiens, variante de transcripción 1, ARNm.	1520
ALG12	NM_024105	ALG12, alfa-1,6-manosiltransferasa (ALG12) de Homo sapiens, ARNm.	1521

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
ALPL	NM_000478	Fosfatasa alcalina, hígado/hueso/riñón (ALPL) de Homo sapiens, variante de transcripción 1, ARNm.	1522
ALPL	NM_001127501	Fosfatasa alcalina, hígado/hueso/riñón (ALPL) de Homo sapiens, variante de transcripción 2, ARNm.	1523
ALPL	NM_001177520	Fosfatasa alcalina, hígado/hueso/riñón (ALPL) de Homo sapiens, variante de transcripción 3, ARNm.	1524
AP3B1	NM_001271769	Complejo proteico 3 relacionado con el adaptador subunidad beta 1 (AP3B1) de Homo sapiens, variante de transcripción 2, ARNm.	1525
AP3B1	NM_003664	Complejo proteico 3 relacionado con el adaptador subunidad beta 1 (AP3B1) de Homo sapiens, variante de transcripción 1, ARNm.	1526
AP3B2	NM_004644	Complejo proteico 3 relacionado con el adaptador subunidad beta 1 (AP3B2) de Homo sapiens, variante de transcripción 2, ARNm.	1527
AP3D1	NM_003938	Complejo proteico 3 relacionado con el adaptador subunidad beta 1 (AP3D1) de Homo sapiens, variante de transcripción 2, ARNm.	1528
AP3D1	NM_001261826	Complejo proteico 3 relacionado con el adaptador subunidad beta 1 (AP3D1) de Homo sapiens, variante de transcripción 3, ARNm.	1529
APOL1	NM_001136540	Apolipoproteína L1 (APOL1) de Homo sapiens, variante de transcripción 3, ARNm.	1530
APOL1	NM_001136541	Apolipoproteína L1 (APOL1) de Homo sapiens, variante de transcripción 4, ARNm.	1531
APOL1	NM_003661	Apolipoproteína L1 (APOL1) de Homo sapiens, variante de transcripción 1, ARNm.	1532
APOL1	NM_145343	Apolipoproteína L1 (APOL1) de Homo sapiens, variante de transcripción 2, ARNm.	1533
ASH1L	NM_018489	ASH1 como histona lisina metiltransferasa (ASH1L) de Homo sapiens, ARNm.	1534
ATL2	NM_001135673	Atlastina GTPasa 2 (ATL2) de Homo sapiens, variante de transcripción 2, ARNm.	1535
ATL2	NM_022374	Atlastina GTPasa 2 (ATL2) de Homo sapiens, variante de transcripción 1, ARNm.	1536
ATL2	NR_024191	Atlastina GTPasa 2 (ATL2) de Homo sapiens, variante de transcripción 3, ARN no codificante.	1537
ATM	NM_000051	ATM serina/treonina quinasa (ATM) de Homo sapiens, ARNm.	1538
ATR	NM_001184	ATR serina/treonina quinasa (ATR) de Homo sapiens, ARNm.	1539
BACH2	NM_001170794	Dominio BTB y CNC homólogo 2 (BACH2) de Homo sapiens, variante de transcripción 2, ARNm.	1540
BACH2	NM_021813	Dominio BTB y CNC homólogo 2 (BACH2) de Homo sapiens, variante de transcripción 1, ARNm.	1541
BAG3	NM_004281	Atanogen 3 asociado a BCL2 (BAG3) de Homo sapiens, ARNm.	1542
BCL10	NM_003921	CLL/linfoma de células B 10 (BCL10) de Homo sapiens, variante de transcripción 1, ARNm.	1543
BLM	NM_000057	Síndrome de Bloom similar a la helicasa RecQ (BLM) de Homo sapiens, variante de transcripción 1, ARNm.	1544
BLNK	NM_001114094	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 2, ARNm.	1545
BLNK	NM_001258440	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 3, ARNm.	1546

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
BLNK	NM_001258441	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 4, ARNm.	1547
BLNK	NM_001258442	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 5, ARNm.	1548
BLNK	NM_013314	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 1, ARNm.	1549
BLNK	NR_047680	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 6, ARNm.	1550
BLNK	NR_047681	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 7, ARNm.	1551
BLNK	NR_047682	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 8, ARNm.	1552
BLNK	NR_047683	Enlazador de células B (BLNK) de Homo sapiens, variante de transcripción 9, ARNm.	1553
BLOC1S6	NM_012388	Biogénesis del complejo de orgánulos lisosomales 1 subunidad 6 (BLOC1S6) de Homo sapiens, variante de transcripción 2, ARNm.	1554
BTK	NM_000061	Tirosina quinasa Bruton (BTK) de Homo sapiens, variante de transcripción 1, ARNm.	1555
C11orf65	NM_152587	Marco de lectura abierto 65 del cromosoma 11 (C11orf65) de Homo sapiens, ARNm.	1556
C1QA	NM_015991	Componente 1 del complemento, subcomponente q, cadena A (C1QA) de Homo sapiens, ARNm.	1557
C1QB	NM_000491	Componente 1 del complemento, subcomponente q, cadena B (C1QB) de Homo sapiens, ARNm.	1558
C1QC	NM_001114101	Componente 1 del complemento, subcomponente q, cadena C (C1QC) de Homo sapiens, variante de transcripción 1, ARNm.	1559
C1QC	NM_172369	Componente 1 del complemento, subcomponente q, cadena B (C1QB) de Homo sapiens, variante de transcripción 2, ARNm.	1560
C5AR1	NM_001736	Receptor 1 del componente 5a del complemento (C5AR1) de Homo sapiens, ARNm.	1561
CARD 11	NM_032415	Miembro 11 de la familia del dominio de reclutamiento de caspasas (CARD11) de Homo sapiens, variante de transcripción 2, ARNm.	1562
CARD9	NM_052813	Familia del dominio de reclutamiento de caspasas, miembro 9 (CARD9) de Homo sapiens, variante de transcripción 1, ARNm.	1563
CARD9	NM_052814	Familia del dominio de reclutamiento de caspasas, miembro 9 (CARD9) de Homo sapiens, variante de transcripción 2, ARNm.	1564
CASP8	NM 001080124	Caspasa 8 (CASP8) de Homo sapiens, variante de transcripción F, ARNm.	1565
CASP8	NM_001228	Caspasa 8 (CASP8) de Homo sapiens, variante de transcripción A, ARNm.	1566
CASP8	NM_033355	Caspasa 8 (CASP8) de Homo sapiens, variante de transcripción B, ARNm.	1567
CASP8	NM_033358	Caspasa 8 (CASP8) de Homo sapiens, variante de transcripción E, ARNm.	1568
CASP8	NM_001080125	Caspasa 8 (CASP8) de Homo sapiens, variante de transcripción G, ARNm.	1569
CASP8	NM_-033356	Caspasa 8 (CASP8) de Homo sapiens, variante de transcripción C, ARNm.	1570

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
CCL11	NM_002986	Ligando 11 de quimioquinas con motivo C-C (CCL11) de Homo sapiens, ARNm.	1571
CCL2	NM_002982	Ligando 2 de quimioquinas con motivo C-C (CCL2) de Homo sapiens, ARNm.	1572
CCL5	NM_002985	Ligando 5 de quimioquinas con motivo C-C (CCLS) de Homo sapiens, variante de transcripción 1, ARNm.	1573
CCR2	NM_001123041	Receptor 2 de quimioquinas con motivo C-C (CCR2) de Homo sapiens, variante de transcripción A, ARNm.	1574
CCR2	NM_001123396	Receptor 2 de quimioquinas con motivo C-C (CCR2) de Homo sapiens, variante de transcripción B, ARNm.	1575
CCR5	NM_000579	Receptor 5 de quimioquinas con motivo C-C (gen/pseudogén) (CCRS) de Homo sapiens, variante de transcripción A, ARNm.	1576
CCR5	NM_-001100168	Receptor 5 de quimioquinas con motivo C-C (gen/pseudogén) (CCR5) de Homo sapiens, variante de transcripción B, ARNm.	1577
CD180	NM_005582	Molécula CD180 (CD180) de Homo sapiens, ARNm.	1578
CD19	NM_001178098	Molécula CD19 (CD19) de Homo sapiens, variante de transcripción 1, ARNm.	1579
CD19	NM_001770	Molécula CD19 (CD19) de Homo sapiens, variante de transcripción 2, ARNm.	1580
CD209	NM_001144893	Molécula CD209 (CD209) de Homo sapiens, variante de transcripción 5, ARNm.	1581
CD209	NM_001144894	Molécula CD209 (CD209) de Homo sapiens, variante de transcripción 6, ARNm.	1582
CD209	NM_001144895	Molécula CD209 (CD209) de Homo sapiens, variante de transcripción 7, ARNm.	1583
CD209	NM_001144896	Molécula CD209 (CD209) de Homo sapiens, variante de transcripción 3, ARNm.	1584
CD209	NM_001144897	Molécula CD209 (CD209) de Homo sapiens, variante de transcripción 4, ARNm.	1585
CD209	NM_001144899	Molécula CD209 (CD209) de Homo sapiens, variante de transcripción 8, ARNm.	1586
CD209	NM_021155	Molécula CD209 (CD209) de Homo sapiens, variante de transcripción 1, ARNm.	1587
CD209	NR_026692	Molécula CD209 (CD209) de Homo sapiens, variante de transcripción 2, ARN no codificante.	1588
CD247	NM_000734	Molécula CD247 (CD247) de Homo sapiens, variante de transcripción 2, ARNm.	1589
CD247	NM_198053	Molécula CD247 (CD247) de Homo sapiens, variante de transcripción 1, ARNm.	1590
CD27	NM_001242	Molécula CD27 (CD27) de Homo sapiens, ARNm.	1591
CD27-AS1	NR_015382	ARN1 antisentido de CD27 (CD27-AS1) de Homo sapiens, ARN no codificante largo.	1592
CD34	NM 001025109	Molécula CD34 (CD34) de Homo sapiens, variante de transcripción 1, ARNm.	1593
CD34	NM_001773	Molécula CD34 (CD34) de Homo sapiens, variante de transcripción 2, ARNm.	1594
CD3D	NM_000732	Molécula CD3d (CD3D) de Homo sapiens, variante de transcripción 1, ARNm.	1595

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
CD3D	NM_001040651	Molécula CD3d (CD3D) de Homo sapiens, variante de transcripción 2, ARNm.	1596
CD3E	NM_000733	Molécula CD3e (CD3E) de Homo sapiens, ARNm.	1597
CD3G	NM_000073	Molécula CD3g (CD3G) de Homo sapiens, ARNm.	1598
CD40	NM_001250	Molécula CD40 (CD40) de Homo sapiens, variante de transcripción 1, ARNm.	1599
CD40	NM_152854	Molécula CD40 (CD40) de Homo sapiens, variante de transcripción 2, ARNm.	1600
CD40LG	NM_000074	Ligando CD40 (CD40LG) de Homo sapiens, ARNm.	1601
CD55	NM_000574	Molécula CD55 (grupo sanguíneo Cromer) (CD55) de Homo sapiens, variante de transcripción 1, ARNm.	1602
CD55	NM_001114752	Molécula CD55 (grupo sanguíneo Cromer) (CD55) de Homo sapiens, variante de transcripción 2, ARNm.	1603
CD59	NM_000611	Molécula CD59 (CD59) de Homo sapiens, variante de transcripción 2, ARNm.	1604
CD59	NM_001127223	Molécula CD59 (CD59) de Homo sapiens, variante de transcripción 5, ARNm.	1605
CD59	NM_001127225	Molécula CD59 (CD59) de Homo sapiens, variante de transcripción 6, ARNm.	1606
CD59	NM_001127226	Molécula CD59 (CD59) de Homo sapiens, variante de transcripción 7, ARNm.	1607
CD59	NM_001127227	Molécula CD59 (CD59) de Homo sapiens, variante de transcripción 8, ARNm.	1608
CD59	NM_203329	Molécula CD59 (CD59) de Homo sapiens, variante de transcripción 3, ARNm.	1609
CD59	NM_203330	Molécula CD59 (CD59) de Homo sapiens, variante de transcripción 1, ARNm.	1610
CD59	NM_203331	Molécula CD59 (CD59) de Homo sapiens, variante de transcripción 4, ARNm.	1611
CD79A	NM 001783	Molécula CD79a (CD79A) de Homo sapiens, variante de transcripción 1, ARNm.	1612
CD79A	NM_021601	Molécula CD79a (CD79A) de Homo sapiens, variante de transcripción 2, ARNm.	1613
CD79B	NM_000626	Molécula CD79b (CD79B) de Homo sapiens, variante de transcripción 1, ARNm.	1614
CD79B	NM_001039933	Molécula CD79b (CD79B) de Homo sapiens, variante de transcripción 3, ARNm.	1615
CD79B	NM_021602	Molécula CD79b (CD79B) de Homo sapiens, variante de transcripción 2, ARNm.	1616
CD81	NM_004356	Molécula CD81 (CD81) de Homo sapiens, variante de transcripción 1, ARNm.	1617
CD8A	NM_001145873	Molécula CD8a (CD8A) de Homo sapiens, variante de transcripción 3, ARNm.	1618
CD8A	NM_001768	Molécula CD8a (CD8A) de Homo sapiens, variante de transcripción 1, ARNm.	1619
CD8A	NM_171827	Molécula CD8a (CD8A) de Homo sapiens, variante de transcripción 2, ARNm.	1620

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
CD8A	NR_027353	Molécula CD8a (CD8A) de Homo sapiens, variante de transcripción 4, ARN no codificante.	1621
CDCA7	NM_031942	Ciclo de división celular asociado a 7 (CDCA7) de Homo sapiens, variante de transcripción 1, ARNm.	1622
CDCA7	NM_145810	Ciclo de división celular asociado a 7 (CDCA7) de Homo sapiens, variante de transcripción 2, ARNm.	1623
CEBPB	NM_005194	CCAAT/proteína beta de unión a potenciadores (CEBPB) de Homo sapiens, variante de transcripción 1, ARNm.	1624
CHD7	NM_017780	Proteína 7 de unión al ADN de la helicasa de cromodominio (CHD7) de Homo sapiens, variante de transcripción 1, ARNm.	1625
CHEK1	NM_001114121	Quinasa 1 de punto de control (CHEK1) de Homo sapiens, variante de transcripción 2, ARNm.	1626
CHEK1	NM_001114122	Quinasa 1 de punto de control (CHEK1) de Homo sapiens, variante de transcripción 1, ARNm.	1627
CHEK1	NM_001244846	Quinasa 1 de punto de control (CHEK1) de Homo sapiens, variante de transcripción 4, ARNm.	1628
CHEK1	NR_045204	Quinasa 1 de punto de control (CHEK1) de Homo sapiens, variante de transcripción 5, ARNm.	1629
CHEK1	NR_045205	Quinasa 1 de punto de control (CHEK1) de Homo sapiens, variante de transcripción 6, ARNm.	1630
CHEK1	NM_001274	Quinasa 1 de punto de control (CHEK1) de Homo sapiens, variante de transcripción 3, ARNm.	1631
CIITA	NM_000246	Transactivador del complejo principal de histocompatibilidad de clase II (CIITA) de Homo sapiens, variante de transcripción 2, ARNm.	1632
CLCN7	NM_001114331	Canal de cloruro, sensible al voltaje 7 (CLCN7) de Homo sapiens, variante de transcripción 2, ARNm.	1633
CLCN7	NM_001287	Canal de cloruro, sensible al voltaje 7 (CLCN7) de Homo sapiens, variante de transcripción 1, ARNm.	1634
COG6	NM_001145079	Componente del complejo oligomérico de golgi 6 (COG6) de Homo sapiens, variante de transcripción 2, ARNm.	1635
COG6	NM_020751	Componente del complejo oligomérico de golgi 6 (COG6) de Homo sapiens, variante de transcripción 1, ARNm.	1636
COG6	NR_026745	Componente del complejo oligomérico de golgi 6 (COG6) de Homo sapiens, variante de transcripción 3, ARNm.	1637
CORO 1A	NM_001193333	Coronina 1A (CORO1A) de Homo sapiens, variante de transcripción 1, ARNm.	1638
CORO 1A	NM_007074	Coronina 1A (CORO1A) de Homo sapiens, variante de transcripción 2, ARNm.	1639
CR2	NM_001006658	Receptor 2 del componente 3d del complemento (CR2) de Homo sapiens, variante de transcripción 1, ARNm.	1640
CR2	NM_001877	Receptor 2 del componente 3d del complemento (CR2) de Homo sapiens, variante de transcripción 2, ARNm.	1641
CRTC3	NM_001042574	Coactivador 3 de transcripción regulado por CREB (CRTC3), de Homo sapiens, variante de transcripción 2, ARNm.	1642
CRTC3	NM_022769	Coactivador 3 de transcripción regulado por CREB (CRTC3), de Homo sapiens, variante de transcripción 1, ARNm.	1643

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
CSF3R	NM_000760	Receptor del factor 3 estimulante de colonias (granulocitos) (CSF3R) de Homo sapiens, variante de transcripción 1, ARNm.	1644
CSF3R	NM_156039	Receptor del factor 3 estimulante de colonias (granulocito) (CSF3R) de Homo sapiens, variante de transcripción 3, ARNm.	1645
CSF3R	NM_172313	Receptor del factor 3 estimulante de colonias (granulocito) (CSF3R) de Homo sapiens, variante de transcripción 4, ARNm.	1646
CTLA4	NM_005214	Proteína 4 asociada a linfocitos T citotóxicos (CTLA4) de Homo sapiens, variante de transcripción 1, ARNm.	1647
CTLA4	NM_001037631	Proteína 4 asociada a linfocitos T citotóxicos (CTLA4) de Homo sapiens, variante de transcripción 2, ARNm.	1648
CTPS1	NM_001905	CTP sintasa 1 (CTPS1) de Homo sapiens, variante de transcripción 1, ARNm.	1649
CTSC	NM_148170	Catepsina C (CTSC) de Homo sapiens, variante de transcripción 2, ARNm.	1650
CTSC	NM_001114173	Catepsina C (CTSC) de Homo sapiens, variante de transcripción 3, ARNm.	1651
CTSC	NM_001814	Catepsina C (CTSC) de Homo sapiens, variante de transcripción 1, ARNm.	1652
CX3CR1	NM_001171171	Receptor 1 de quimiocina del motivo C-X3-C (CX3CR1) de Homo sapiens, variante de transcripción 2, ARNm.	1653
CX3CR1	NM_001171172	Receptor 1 de quimiocina del motivo C-X3-C (CX3CR1) de Homo sapiens, variante de transcripción 3, ARNm.	1654
CX3CR1	NM_001171174	Receptor 1 de quimiocina del motivo C-X3-C (CX3CR1) de Homo sapiens, variante de transcripción 1, ARNm.	1655
CX3CR1	NM_001337	Receptor 1 de quimiocina del motivo C-X3-C (CX3CR1) de Homo sapiens, variante de transcripción 4, ARNm.	1656
CXCL 12	NM_000609	Ligando 12 de quimiocina del motivo C-X-C (CXCL12) de Homo sapiens, variante de transcripción 2, ARNm.	1657
CXCL 12	NM_001033886	Ligando 12 de quimiocina del motivo C-X-C (CXCL12) de Homo sapiens, variante de transcripción 3, ARNm.	1658
CXCL 12	NM_001178134	Ligando 12 de quimiocina del motivo C-X-C (CXCL12) de Homo sapiens, variante de transcripción 4, ARNm.	1659
CXCL 12	NM_199168	Ligando 12 de quimiocina del motivo C-X-C (CXCL12) de Homo sapiens, variante de transcripción 1, ARNm.	1660
CXCL9	NM_002416	Ligando 9 de quimiocina del motivo C-X-C (CXCL9) de Homo sapiens, ARNm.	1661
CXCR1	NM_000634	Receptor 1 de quimiocina del motivo C-X-C (CXCR1) de Homo sapiens, ARNm.	1662
CXCR4	NM_001008540	Receptor 4 de quimiocina del motivo C-X-C (CXCR4) de Homo sapiens, variante de transcripción 1, ARNm.	1663
CXCR4	NM_003467	Receptor 4 de quimiocina del motivo C-X-C (CXCR4) de Homo sapiens, variante de transcripción 2, ARNm.	1664
CXorf40A	NM_001171907	Marco de lectura abierta 40A del cromosoma X (CXorf40A) de Homo sapiens, variante de transcripción 2, ARNm.	1665
CXorf40A	NM_001171908	Marco de lectura abierta 40A del cromosoma X (CXorf40A) de Homo sapiens, variante de transcripción 3, ARNm.	1666
CXorf40A	NM_178124	Marco de lectura abierta 40A del cromosoma X (CXorf40A) de Homo sapiens, variante de transcripción 1, ARNm.	1667

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
CXorf40A	NM_001171909	Marco de lectura abierta 40A del cromosoma X (CXorf40A) de Homo sapiens, variante de transcripción 4, ARNm.	1668
CYBB	NM_000397	Citocromo b-245, polipéptido beta (CYBB) de Homo sapiens, ARNm.	1669
CYP2S1	NM 030622	Citocromo P450 familia 2 subfamilia S miembro 1 (CYP2S1) de Homo sapiens. ARNm.	1670
DCLRE1C	NM_001033855	Reparación 1C de enlaces cruzados de ADN (DCLRE1C), variante de transcripción a, ARNm.	1671
DCLRE1C	NM 001033857	Reparación 1C de enlaces cruzados de ADN (DCLRE1C), variante de transcripción d, ARNm.	1672
DCLRE1C	NM_001033858	Reparación 1C de enlaces cruzados de ADN (DCLRE1C), variante de transcripción c, ARNm.	1673
DCLRE1C	NM_022487	Reparación 1C de enlaces cruzados de ADN (DCLRE1C), variante de transcripción b, ARNm.	1674
DDX1	NM_004939	DEAD/H-box helicasa 1 (DDX1) de Homo sapiens, ARNm.	1675
DDX58	NM_014314	Polipéptido DEAD (Asp-Glu-Ala-Asp) box 58 (DDX58) de Homo sapiens, ARNm.	1676
DHX58	NM_024119	Polipéptido DEXH (Asp-Glu-X-His) box 58 (DHX58) de Homo sapiens, ARNm.	1677
DKC1	NM_001142463	Disquerina pseudouridina sintasa 1 (DKC1) de Homo sapiens, variante de transcripción 2, ARNm.	1678
DKC1	NM_001363	Disquerina pseudouridina sintasa 1 (DKC1) de Homo sapiens, variante de transcripción 1, ARNm.	1679
DNMT3B	NM_001207055	ADN metiltransferasa 3 beta (DNMT3B) de Homo sapiens, variante de transcripción 7, ARNm.	1680
DNMT3B	NM_001207056	ADN metiltransferasa 3 beta (DNMT3B) de Homo sapiens, variante de transcripción 8, ARNm.	1681
DNMT3B	NM_006892	ADN metiltransferasa 3 beta (DNMT3B) de Homo sapiens, variante de transcripción 1, ARNm.	1682
DNMT3B	NM_175848	ADN metiltransferasa 3 beta (DNMT3B) de Homo sapiens, variante de transcripción 2, ARNm.	1683
DNMT3B	NM_175849	ADN metiltransferasa 3 beta (DNMT3B) de Homo sapiens, variante de transcripción 3, ARNm.	1684
DNMT3B	NM_175850	ADN metiltransferasa 3 beta (DNMT3B) de Homo sapiens, variante de transcripción 6, ARNm.	1685
DOCK2	NM_004946	Dedicador de la citocinesis 2 (DOCK2) de Homo sapiens, ARNm.	1686
DOCK8	NM_203447	Dedicador de la citocinesis 8 (DOCK8) de Homo sapiens, variante de transcripción 1, ARNm.	1687
DOCK8	NM_001190458	Dedicador de la citocinesis 8 (DOCK8) de Homo sapiens, variante de transcripción 2, ARNm.	1688
DOCK8	NM_001193536	Dedicador de la citocinesis 8 (DOCK8) de Homo sapiens, variante de transcripción 3, ARNm.	1689
DSC1	NM_004948	Desmocolina 1 (DSC1) de Homo sapiens, variante de transcripción Dsc1b, ARNm.	1690
DSC1	NM_024421	Desmocolina 1 (DSC1) de Homo sapiens, variante de transcripción Dsc1a, ARNm.	1691
EGR1	NM_001964	Respuesta 1 de crecimiento temprano (EGR1) de Homo sapiens, ARNm.	1692

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
ELANE	NM_001972	Elastasa, expresada por neutrófilos (ELANE) de Homo sapiens, ARNm.	1693
EPG5	NM_020964	Homólogo de la proteína 5 de autofagia de los gránulos P ectópicos (EPGS) de Homo sapiens, ARNm.	1694
ETF1	NM_004730	Factor 1 de terminación de la traducción eucariótica (ETF1) de Homo sapiens, variante de transcripción 1, ARNm.	1695
F9	NM_000133	Factor IX de coagulación (F9) de Homo sapiens, variante de transcripción 1, ARNm.	1696
FAS	NM_000043	Receptor de muerte de superficie celular Fas (FAS) de Homo sapiens, variante de transcripción 1, ARNm.	1697
FAS	NM_152871	Receptor de muerte de superficie celular Fas (FAS) de Homo sapiens, variante de transcripción 2, ARNm.	1698
FAS	NM_152872	Receptor de muerte de superficie celular Fas (FAS) de Homo sapiens, variante de transcripción 3, ARNm.	1699
FAS	NR_028033	Receptor de muerte de superficie celular Fas (FAS) de Homo sapiens, variante de transcripción 4, ARN no codificante.	1700
FAS	NR_028034	Receptor de muerte de superficie celular Fas (FAS) de Homo sapiens, variante de transcripción 5, ARN no codificante.	1701
FAS	NR_028035	Receptor de muerte de superficie celular Fas (FAS) de Homo sapiens, variante de transcripción 6, ARN no codificante.	1702
FAS	NR_028036	Receptor de muerte de superficie celular Fas (FAS) de Homo sapiens, variante de transcripción 7, ARN no codificante.	1703
FASLG	NM_000639	Ligando Fas (superfamilia TNF, miembro 6) (FASLG) de Homo sapiens, ARNm.	1704
FCGR2A	NM_001136219	Fragmento Fc del receptor IIa de IgG (FCGR2A) de Homo sapiens, variante de transcripción 1, ARNm.	1705
FCGR2A	NM_021642	Fragmento Fc del receptor IIa de IgG (FCGR2A) de Homo sapiens, variante de transcripción 2, ARNm.	1706
FCGR3A	NM_000569	Fragmento Fc del receptor IIIa de IgG (FCGR3A) de Homo sapiens, variante de transcripción 1, ARNm.	1707
FCGR3A	NM_001127592	Fragmento Fc del receptor IIIa de IgG (FCGR3A) de Homo sapiens, variante de transcripción 2, ARNm.	1708
FCGR3A	NM_001127593	Fragmento Fc del receptor IIIa de IgG (FCGR3A) de Homo sapiens, variante de transcripción 3, ARNm.	1709
FCGR3A	NM_001127595	Fragmento Fc del receptor IIIa de IgG (FCGR3A) de Homo sapiens, variante de transcripción 4, ARNm.	1710
FCGR3A	NM_001127596	Fragmento Fc del receptor IIIa de IgG (FCGR3A) de Homo sapiens, variante de transcripción 5, ARNm.	1711
FCN3	NM_003665	Ficolina 3 (FCN3) de Homo sapiens, variante de transcripción 1, ARNm.	1712
FCN3	NM_173452	Ficolina 3 (FCN3) de Homo sapiens, variante de transcripción 2, ARNm.	1713
FEZ1	NM_005103	Fasciculación y elongación de la proteína zeta 1 (FEZ1) de Homo sapiens, variante de transcripción 1, ARNm.	1714
FEZ1	NM_022549	Fasciculación y elongación de la proteína zeta 1 (FEZ1) de Homo sapiens, variante de transcripción 2, ARNm.	1715
FOS	NM_005252	Fos proto-oncogén, subunidad del factor de transcripción AP-1 (FOS) de Homo sapiens, ARNm.	1716

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
FOXH1	NM_003923	Forkhead box H1 (FOXH1) de Homo sapiens, ARNm.	1717
FOXN1	NM_003593	Forkhead box N1 (FOXN1) de Homo sapiens, ARNm.	1718
FOXP3	NM_001114377	Forkhead box P3 (FOXP3) de Homo sapiens, variante de transcripción 2, ARNm.	1719
FOXP3	NM_014009	Forkhead box P3 (FOXP3) de Homo sapiens, variante de transcripción 1, ARNm.	1720
FPR1	NM_001193306	Receptor 1 del péptido formilo (FPR1) de Homo sapiens, variante de transcripción 1, ARNm.	1721
FPR1	NM_002029	Receptor del péptido formilo (FPR1) de Homo sapiens, variante de transcripción 2, ARNm.	1722
G6PC3	NM 138387	Glucosa 6 fosfatasa, catalítica, 3 (G6PC3) de Homo sapiens, variante de transcripción 1, ARNm.	1723
G6PC3	NR_028582	Glucosa 6 fosfatasa, catalítica, 3 (G6PC3) de Homo sapiens, variante de transcripción 2, ARN no codificante.	1724
G6PC3	NR_028581	Glucosa 6 fosfatasa, catalítica, 3 (G6PC3) de Homo sapiens, variante de transcripción 3, ARN no codificante.	1725
GATA2	NM_001145661	Proteína 2 de unión a GATA (GATA2) de Homo sapiens, variante de transcripción 1, ARNm.	1726
GATA2	NM_001145662	Proteína 2 de unión a GATA (GATA2) de Homo sapiens, variante de transcripción 3, ARNm.	1727
GATA2	NM_032638	Proteína 2 de unión a GATA (GATA2) de Homo sapiens, variante de transcripción 2, ARNm.	1728
GFI1	NM_005263	Factor de crecimiento independiente 1 represor de la transcripción (GFI1) de Homo sapiens, variante de transcripción 1, ARNm.	1729
GFI1	NM_001127215	Factor de crecimiento independiente 1 represor de la transcripción (GFI1) de Homo sapiens, variante de transcripción 2, ARNm.	1730
GFI1	NM_001127216	Factor de crecimiento independiente 1 represor de la transcripción (GFI1) de Homo sapiens, variante de transcripción 3, ARNm.	1731
GOLGB1	NM_001256486	Golgina B1 (GOLGB1) de Homo sapiens, variante de transcripción 1, ARNm.	1732
GOLGB1	NM 001256487	Golgina B1 (GOLGB1) de Homo sapiens, variante de transcripción 3, ARNm.	1733
GOLGB1	NM_001256488	Golgina B1 (GOLGB1) de Homo sapiens, variante de transcripción 4, ARNm.	1734
GOLGB1	NM_004487	Golgina B1 (GOLGB1) de Homo sapiens, variante de transcripción 2, ARNm.	1735
GPRC5A	NM_003979	Receptor acoplado a proteína G clase C grupo 5 miembro A (GPRC5A) de Homo sapiens, ARNm.	1736
GRAP2	NM_004810	Proteína 2 adaptadora relacionada con GRB2 (GRAP2) de Homo sapiens, ARNm.	1737
HAX1	NM_006118	Proteína X-1 asociada a HCLS1 (HAX1) de Homo sapiens, variante de transcripción 1, ARNm.	1738
HAX1	NM_001018837	Proteína X-1 asociada a HCLS1 (HAX1) de Homo sapiens, variante de transcripción 2, ARNm.	1739
HELLS	NM_018063	Helicasa, específica de linfoides (HELLS) de Homo sapiens, variante de transcripción 1, ARNm.	1740

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
HIVEP1	NM_002114	Proteína 1 de unión al potenciador del virus de la inmunodeficiencia humana tipo I (HIVEP1) de Homo sapiens, ARNm.	1741
HIVEP2	NM_006734	Proteína 2 de unión al potenciador del virus de la inmunodeficiencia humana tipo I (HIVEP2), de Homo sapiens, ARNm.	1742
HIVEP3	NM_001127714	Proteína 3 de unión al potenciador del virus de la inmunodeficiencia humana tipo I (HIVEP3), de Homo sapiens, variante de transcripción 2, ARNm.	1743
HIVEP3	NM_024503	Proteína 3 de unión al potenciador del virus de la inmunodeficiencia humana tipo I (HIVEP3), de Homo sapiens, variante de transcripción 1, ARNm.	1744
HIVEP3	NR_038260	Proteína 3 de unión al potenciador del virus de la inmunodeficiencia humana tipo I (HIVEP3), de Homo sapiens, variante de transcripción 3, ARN no codificante.	1745
HIVEP3	NR_038261	Proteína 3 de unión al potenciador del virus de la inmunodeficiencia humana tipo I (HIVEP3), de Homo sapiens, variante de transcripción 4, ARN no codificante.	1746
HP	NM_001126102	Haptoglobina (HP) de Homo sapiens, variante de transcripción 2, ARNm.	1747
HP	NM_005143	Haptoglobina (HP) de Homo sapiens, variante de transcripción 1, ARNm.	1748
HPCAL1	NM_002149	Hipocalcina tipo 1 (HPCAL1) de Homo sapiens, variante de transcripción 1, ARNm.	1749
HPCAL1	NM_134421	Hipocalcina tipo 1 (HPCAL1) de Homo sapiens, variante de transcripción 2, ARNm.	1750
HPCAL1	NM_001258357	Hipocalcina tipo 1 (HPCAL1) de Homo sapiens, variante de transcripción 3, ARNm.	1751
HPCAL1	NM_001258358	Hipocalcina tipo 1 (HPCAL1) de Homo sapiens, variante de transcripción 4, ARNm.	1752
HPCAL1	NM_001258359	Hipocalcina tipo 1 (HPCAL1) de Homo sapiens, variante de transcripción 5, ARNm.	1753
HTR2A	NM_000621	Receptor 2A de la 5-hidroxitriptamina (HTR2A) de Homo sapiens, variante de transcripción 1, ARNm.	1754
HTR2A	NM_001165947	Receptor 2A de 5-hidroxitriptamina (serotonina), acoplado a proteína G (HTR2A) de Homo sapiens, variante de transcripción 2, ARNm.	1755
ICOS	NM_012092	Coestimulador inducible de células T (ICOS) de Homo sapiens, ARNm.	1756
IDI1	NM_004508	Delta isomerasa 1 de difosfato de isopentenilo (IDI1) de Homo sapiens, variante de transcripción 1, ARNm.	1757
IFIH1	NM_022168	Interferón inducido con dominio 1 de helicasa C (IFIH1) de Homo sapiens, ARNm.	1758
IFNAR1	NM_000629	Receptor 1 del interferón (alfa, beta y omega) (IFNAR1) de Homo sapiens, ARNm.	1759
IFNAR2	NM-_-207584	Receptor 2 del interferón (alfa, beta y omega) (IFNAR2) de Homo sapiens, variante de transcripción 3, ARNm.	1760
IFNAR2	NM_207585	Receptor 2 del interferón (alfa, beta y omega) (IFNAR2) de Homo sapiens, variante de transcripción 1, ARNm.	1761

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
IFNAR2	NM_000874	Receptor 2 del interferón (alfa, beta y omega) (IFNAR2) de Homo sapiens, variante de transcripción 2, ARNm.	1762
IFNG	NM_000619	Interferón gamma (IFNG) de Homo sapiens, ARNm.	1763
IFNGR1	NM 000416	Receptor 1 de interferón gamma (IFNGR1) de Homo sapiens, ARNm.	1764
IFNGR2	NM_005534	Receptor 2 de interferón gamma (transductor 1 de interferón gamma) (IFNGR2) de Homo sapiens, variante de transcripción 2, ARNm.	1765
IGLL1	NM_020070	Polipéptido 1 tipo inmunoglobulina lambda 1 (IGLL1) de Homo sapiens, variante de transcripción 1, ARNm.	1766
IGLL1	NM_152855	Polipéptido 1 tipo inmunoglobulina lambda 1 (IGLL1) de Homo sapiens, variante de transcripción 2, ARNm.	1767
IKBKB	NM_001190720	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa beta (IKBKB) de Homo sapiens, variante de transcripción 2, ARNm.	1768
IKBKB	NM_001242778	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa beta (IKBKB) de Homo sapiens, variante de transcripción 7, ARNm.	1769
IKBKB	NM_001556	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa beta (IKBKB) de Homo sapiens, variante de transcripción 1, ARNm.	1770
IKBKB	NR_033818	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa beta (IKBKB) de Homo sapiens, variante de transcripción 5, ARN no codificante.	1771
IKBKB	NR_033819	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa beta (IKBKB) de Homo sapiens, variante de transcripción 6, ARN no codificante.	1772
IKBKB	NR_040009	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa beta (IKBKB) de Homo sapiens, variante de transcripción 8, ARN no codificante.	1773
IKBKG	NM 001099856	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa gamma (IKBKB) de Homo sapiens, variante de transcripción 2, ARNm.	1774
IKBKG	NM_001099857	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa gamma (IKBKB) de Homo sapiens, variante de transcripción 1, ARNm.	1775
IKBKG	NM_001145255	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa gamma (IKBKB) de Homo sapiens, variante de transcripción 4, ARNm.	1776
IKBKG	NM_003639	Inhibidor del potenciador del gen del polipéptido ligero kappa en células B, quinasa gamma (IKBKB) de Homo sapiens, variante de transcripción 3, ARNm.	1777
IKZF1	NM_001220765	Dedo de zinc 1 de la familia IKAROS (IKZF1) de Homo sapiens, variante de transcripción 2, ARNm.	1778
IKZF1	NM_001220767	Dedo de zinc 1 de la familia IKAROS (IKZF1) de Homo sapiens, variante de transcripción 4, ARNm.	1779
IKZF1	NM_001220768	Dedo de zinc 1 de la familia IKAROS (IKZF1) de Homo sapiens, variante de transcripción 5, ARNm.	1780
IKZF1	NM_001220770	Dedo de zinc 1 de la familia IKAROS (IKZF1) de Homo sapiens, variante de transcripción 7, ARNm.	1781

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
IKZF1	NM_001220771	Dedo de zinc 1 de la familia IKAROS (IKZF1) de Homo sapiens, variante de transcripción 8, ARNm.	1782
IKZF1	NM_006060	Dedo de zinc 1 de la familia IKAROS (IKZF1) de Homo sapiens, variante de transcripción 1, ARNm.	1783
IL10	NM_000572	Interleuquina 10 (IL10) de Homo sapiens, ARNm.	1784
IL10RA	NM_001558	Receptor de interleuquina 10 subunidad alfa (IL10RA) de Homo sapiens, variante de transcripción 1, ARNm.	1785
IL10RA	NR_026691	Receptor de interleuquina 10 subunidad alfa (IL10RA) de Homo sapiens, variante de transcripción 2, ARN no codificante.	1786
IL10RB	NM_000628	Receptor de interleuquina 10 subunidad beta (IL10RB) de Homo sapiens, ARNm.	1787
IL12B	NM_002187	Interleuquina 12B (IL12B) de Homo sapiens, ARNm.	1788
IL12RB1	NM_005535	Receptor de interleuquina 12 subunidad beta 1 (IL12RB1) de Homo sapiens, variante de transcripción 1, ARNm.	1789
IL12RB1	NM_153701	Receptor de interleuquina 12 subunidad beta 1 (IL12RB1) de Homo sapiens, variante de transcripción 2, ARNm.	1790
IL17F	NM_052872	Interleuquina 17F (IL17F) de Homo sapiens, ARNm.	1791
IL 17RA	NM_014339	Receptor A de interleuquina 17 (IL17RA) de Homo sapiens, variante de transcripción 1, ARNm.	1792
IL1B	NM_000576	interleuquina 1, beta (IL1B) de Homo sapiens, ARNm.	1793
IL21	NM_001207006	Interleuquina 21 (IL21) de Homo sapiens, variante de transcripción 2, ARNm.	1794
IL21	NM_-021803	Interleuquina 21 (IL21) de Homo sapiens, variante de transcripción 1, ARNm.	1795
IL21R	NM_181078	Receptor de interleuquina 21 (IL21R) de Homo sapiens, variante de transcripción 2, ARNm.	1796
IL21R	NM_181079	Receptor de interleuquina 21 (IL21R) de Homo sapiens, variante de transcripción 3, ARNm.	1797
IL21R	NM_021798	Receptor de interleuquina 21 (IL21R) de Homo sapiens, variante de transcripción 1, ARNm.	1798
IL2RA	NM_000417	Receptor de interleuquina 2, alfa (IL2RA) de Homo sapiens, variante de transcripción 1, ARNm.	1799
IL2RG	NM_000206	Subunidad gamma del receptor de interleuquina 2 (IL2RG) de Homo sapiens, ARNm.	1800
IL4R	NM_000418	Receptor de interleuquina 4 (IL4R) de Homo sapiens, variante de transcripción 1, ARNm.	1801
IL4R	NM_001257406	Receptor de interleuquina 4 (IL4R) de Homo sapiens, variante de transcripción 3, ARNm.	1802
IL4R	NM_001257407	Receptor de interleuquina 4 (IL4R) de Homo sapiens, variante de transcripción 4, ARNm.	1803
IL4R	NM_001257997	Receptor de interleuquina 4 (IL4R) de Homo sapiens, variante de transcripción 5, ARNm.	1804
IL7	NM_000880	Interleuquina 7 (IL7) de Homo sapiens, variante de transcripción 1, ARNm.	1805
IL7	NM 001199886	Interleuquina 7 (IL7) de Homo sapiens, variante de transcripción 2, ARNm.	1806

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
IL7	NM_001199887	Interleuquina 7 (IL7) de Homo sapiens, variante de transcripción 3, ARNm.	1807
IL7	NM_001199888	Interleuquina 7 (IL7) de Homo sapiens, variante de transcripción 4, ARNm.	1808
IL7R	NM_002185	Receptor de interleuquina 7 (IL7R) de Homo sapiens, variante de transcripción 1, ARNm.	1809
IRAK4	NM_001114182	Receptor de interleuquina 1 asociado a quinasa 4 (IRAK4) de Homo sapiens, variante de transcripción 1, ARNm.	1810
IRAK4	NM_001145256	Receptor de interleuquina 1 asociado a quinasa 4 (IRAK4) de Homo sapiens, variante de transcripción 3, ARNm.	1811
IRAK4	NM_001145257	Receptor de interleuquina 1 asociado a quinasa 4 (IRAK4) de Homo sapiens, variante de transcripción 4, ARNm.	1812
IRAK4	NM_001145258	Receptor de interleuquina 1 asociado a quinasa 4 (IRAK4) de Homo sapiens, variante de transcripción 5, ARNm.	1813
IRAK4	NM_016123	Receptor de interleuquina 1 asociado a quinasa 4 (IRAK4) de Homo sapiens, variante de transcripción 2, ARNm.	1814
IRF3	NM_001197122	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 2, ARNm.	1815
IRF3	NM 001197123	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 3, ARNm.	1816
IRF3	NM_001197124	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 4, ARNm.	1817
IRF3	NM_001197125	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 5, ARNm.	1818
IRF3	NM_001197126	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 6, ARNm.	1819
IRF3	NM_001197127	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 7, ARNm.	1820
IRF3	NM_001197128	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 8, ARNm.	1821
IRF3	NM_001571	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 1, ARNm.	1822
IRF3	NR_045568	Factor 3 regulador de interferón (IRF3) de Homo sapiens, variante de transcripción 9, ARN no codificante.	1823
IRF7	NM_001572	Factor 7 regulador de interferón (IRF7) de Homo sapiens, variante de transcripción a, ARNm.	1824
IRF7	NM_004029	Factor 7 regulador de interferón (IRF7) de Homo sapiens, variante de transcripción b, ARNm.	1825
IRF7	NM_004031	Factor 7 regulador de interferón (IRF7) de Homo sapiens, variante de transcripción d, ARNm.	1826
IRF8	NM_002163	Factor 8 regulador del interferón (IRF8) de Homo sapiens, ARNm.	1827
IRGM	NM_001145805	GTPasa M relacionada con la inmunidad (IRGM) de Homo sapiens, ARNm.	1828
ISG15	NM_005101	Modificador similar a la ubiquitina ISG15 (ISG15) de Homo sapiens, ARNm.	1829
ITK	NM_005546	Quinasa de células T inducible por IL2 (ITK) de Homo sapiens, ARNm.	1830

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
ITSN2	NM_006277	Intersectina 2 (ITSN2) de Homo sapiens, variante de transcripción 1, ARNm.	1831
ITSN2	NM_019595	Intersectina 2 (ITSN2) de Homo sapiens, variante de transcripción 3, ARNm.	1832
ITSN2	NM_147152	Intersectina 2 (ITSN2) de Homo sapiens, variante de transcripción 2, ARNm.	1833
JAGN1	NM_032492	Homólogo 1 de jagunal (Drosophila) (JAGN1) de Homo sapiens, ARNm.	1834
JAK3	NM_000215	Janus quinasa 3 (JAK3) de Homo sapiens, ARNm.	1835
JMY	NM_152405	Proteína mediadora y reguladora de la unión, cofactor de p53 (JMY) de Homo sapiens, ARNm.	1836
JUN	NM_002228	Protooncogén Jun, subunidad del factor de transcripción AP-1 (JUN) de Homo sapiens, ARNm.	1837
KITLG	NM_000899	Ligando KIT (KITLG) de Homo sapiens, variante de transcripción b, ARNm.	1838
KITLG	NM_003994	Ligando KIT (KITLG) de Homo sapiens, variante de transcripción a, ARNm.	1839
LAMTOR2	NM_001145264	Adaptador endosomal/lisosomal tardío, activador 2 de MAPK y MTOR (LAMTOR2) de Homo sapiens, variante de transcripción 2, ARNm.	1840
LAMTOR2	NM_014017	Adaptador endosomal/lisosomal tardío, activador 2 de MAPK y MTOR 2 (LAMTOR2) de Homo sapiens, variante de transcripción 1, ARNm.	1841
LCK	NM_005356	Protooncogén LCK, tirosina quinasa de la familia Src (LCK) de Homo sapiens, variante de transcripción 2, ARNm.	1842
LCK	NM_-001042771	Protooncogén LCK, tirosina quinasa de la familia Src (LCK) de Homo sapiens, variante de transcripción 1, ARNm.	1843
LCP2	NM_005565	Proteína 2 citosólica linfocitaria (dominio SH2 que contiene proteína leucocitaria de 76 kDa) (LCP2) de Homo sapiens, ARNm.	1844
LIG1	NM_000234	ADN ligasa 1 (LIG1) de Homo sapiens, variante de transcripción 1, ARNm.	1845
LIG4	NM_001098268	ADN ligasa 4 (LIG4) de Homo sapiens, variante de transcripción 3, ARNm.	1846
LIG4	NM_002312	ADN ligasa 4 (LIG4) de Homo sapiens, variante de transcripción 1, ARNm.	1847
LIG4	NM_206937	ADN ligasa 4 (LIG4) de Homo sapiens, variante de transcripción 2, ARNm.	1848
LRBA	NM_001199282	Proteína de anclaje de tipo beige sensible a LPS (LRBA) de Homo sapiens, variante de transcripción 1, ARNm.	1849
LRBA	NM_006726	Proteína de anclaje de tipo beige sensible a LPS (LRBA) de Homo sapiens, variante de transcripción 2, ARNm.	1850
LYST	NM_000081	Regulador del tráfico lisosomal (LYST) de Homo sapiens, variante de transcripción 1, ARNm.	1851
MAGEA9	NM_005365	Miembro A9 de la familia MAGE (MAGEA9) de Homo sapiens, ARNm.	1852
MAGEA9B	NM_001080790	Miembro A9B de la familia MAGE (MAGEA9B) de Homo sapiens, ARNm.	1853
MAGT1	NM_032121	Transportador 1 de magnesio (MAGT1) de Homo sapiens, ARNm.	1854

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
MALT1	NM_006785	Paracaspasa MALT1 (MALT1) de Homo sapiens, variante de transcripción 1, ARNm.	1855
MALT1	NM_173844	Paracaspasa MALT1 (MALT1) de Homo sapiens, variante de transcripción 2, ARNm.	1856
MAP3K2	NM_006609	Proteína quinasa quinasa quinasa 2 activada por mitógenos (MAP3K2) de Homo sapiens, ARNm.	1857
MAPK1	NM_002745	Proteína quinasa 1 activada por mitógenos (MAPK1) de Homo sapiens, variante de transcripción 1, ARNm.	1858
MAPK1	NM_138957	Proteína quinasa 1 activada por mitógenos (MAPK1) de Homo sapiens, variante de transcripción 2, ARNm.	1859
MAPK3	NM_001040056	Proteína quinasa 3 activada por mitógenos (MAPK3) de Homo sapiens, variante de transcripción 2, ARNm.	1860
MAPK3	NM_001109891	Proteína quinasa 3 activada por mitógenos (MAPK3) de Homo sapiens, variante de transcripción 3, ARNm.	1861
MAPK3	NM_002746	Proteína quinasa 3 activada por mitógenos (MAPK3) de Homo sapiens, variante de transcripción 1, ARNm.	1862
MAVS	NM_020746	Proteína de señalización antiviral mitocondrial (MAVS) de Homo sapiens, variante de transcripción 1, ARNm.	1863
MAVS	NM_001206491	Proteína de señalización antiviral mitocondrial (MAVS) de Homo sapiens, variante de transcripción 3, ARNm.	1864
MAVS	NR_037921	Proteína de señalización antiviral mitocondrial (MAVS) de Homo sapiens, variante de transcripción 2, ARN no codificante.	1865
MECP2	NM_004992	Proteína 2 de unión a metil-CpG (MECP2) de Homo sapiens, variante de transcripción 1, ARNm.	1866
MECP2	NM_001110792	Proteína 2 de unión a metil-CpG (MECP2) de Homo sapiens, variante de transcripción 2, ARNm.	1867
MEX3C	NM_016626	Miembro C de la familia de unión al ARN de mex-3 (MEX3C) de Homo sapiens, ARNm.	1868
MRE11A	NM_005590	Homólogo A de MRE11, nucleasa de reparación de rotura de doble cadena (MRE11A) de Homo sapiens, variante de transcripción 2, ARNm.	1869
MRE11A	NM_005591	Homólogo A de MRE11, nucleasa de reparación de rotura de doble cadena (MRE11A) de Homo sapiens, variante de transcripción 1, ARNm.	1870
MS4A1	NM_021950	Membrana que abarca 4 dominios A1 (MS4A1) de Homo sapiens, variante de transcripción 3, ARNm.	1871
MS4A1	NM_152866	Membrana que abarca 4 dominios A1 (MS4A1) de Homo sapiens, variante de transcripción 1, ARNm.	1872
MSN	NM_-002444	Moesina (MSN) de Homo sapiens, ARNm.	1873
MYD88	NM_001172566	Respuesta 88 primaria de diferenciación mieloide (MYD88) de Homo sapiens, variante de transcripción 5, ARNm.	1874
MYD88	NM_001172567	Respuesta 88 primaria de diferenciación mieloide (MYD88) de Homo sapiens, variante de transcripción 1, ARNm.	1875
MYD88	NM_001172568	Respuesta 88 primaria de diferenciación mieloide (MYD88) de Homo sapiens, variante de transcripción 3, ARNm.	1876
MYD88	NM_001172569	Respuesta 88 primaria de diferenciación mieloide (MYD88) de Homo sapiens, variante de transcripción 4, ARNm.	1877

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
MYD88	NM_002468	Respuesta 88 primaria de diferenciación mieloide (MYD88) de Homo sapiens, variante de transcripción 2, ARNm.	1878
NBN	NM_002485	Nibrina (NBN) de Homo sapiens, ARNm.	1879
NFIC	NM_001245005	Factor nuclear I C (NFIC) de Homo sapiens, variante de transcripción 4, ARNm.	1880
NFIC	NM_205843	Factor nuclear I C (NFIC) de Homo sapiens, variante de transcripción 2, ARNm.	1881
NFIC	NM_001245002	Factor nuclear I C (NFIC) de Homo sapiens, variante de transcripción 1, ARNm.	1882
NFIC	NM_001245004	Factor nuclear I C (NFIC) de Homo sapiens, variante de transcripción 3, ARNm.	1883
NFIC	NM_005597	Factor nuclear I C (NFIC) de Homo sapiens, variante de transcripción 5, ARNm.	1884
NFKB1	NM_003998	Factor nuclear del potenciador del gen del polipéptido ligero kappa en células B 1 (NFKB1) de Homo sapiens, variante de transcripción 1, ARNm.	1885
NFKB1	NM_001165412	Factor nuclear del potenciador del gen del polipéptido ligero kappa en células B 1 (NFKB1) de Homo sapiens, variante de transcripción 2, ARNm.	1886
NFKB2	NM_001077494	Factor nuclear del potenciador del gen del polipéptido ligero kappa en células B 2 (NFKB2) de Homo sapiens, variante de transcripción 1, ARNm.	1887
NFKB2	NM_002502	Factor nuclear del potenciador del gen del polipéptido ligero kappa en células B 2 (NFKB2) de Homo sapiens, variante de transcripción 2, ARNm.	1888
NFKB2	NM_001261403	Factor nuclear del potenciador del gen del polipéptido ligero kappa en células B 2 (NFKB2) de Homo sapiens, variante de transcripción 4, ARNm.	1889
NFKBIA	NM_020529	Inhibidor alfa de NFKB (NFKBIA) de Homo sapiens, ARNm.	1890
NHEJ1	NM_024782	Factor 1 de unión de extremo no homólogo (NHEJ1) de Homo sapiens, ARNm.	1891
NLRP3	NM_183395	Familia NLR, 3 que contiene dominio de pirina (NLRP3) de Homo sapiens, variante de transcripción 2, ARNm.	1892
NLRP3	NM_004895	Familia NLR, 3 que contiene dominio de pirina (NLRP3) de Homo sapiens,, variante de transcripción 1, ARNm.	1893
NLRP3	NM_001127462	Familia NLR, 3 que contiene dominio de pirina (NLRP3) de Homo sapiens, variante de transcripción 5, ARNm.	1894
NLRP3	NM_001127461	Familia NLR, 3 que contiene dominio de pirina (NLRP3) de Homo sapiens, variante de transcripción 4, ARNm.	1895
NLRP3	NM_001079821	Familia NLR, 3 que contiene dominio de pirina (NLRP3) de Homo sapiens, variante de transcripción 3, ARNm.	1896
NLRP3	NM_001243133	Familia NLR, 3 que contiene dominio de pirina (NLRP3) de Homo sapiens, variante de transcripción 6, ARNm.	1897
NOD2	NM_022162	2 que contiene dominio de oligomerización de unión a nucleótidos (NOD2) de Homo sapiens, ARNm.	1898
ORAI1	NM_032790	Modulador 1 de calcio activado por liberación de calcio ORAI (ORAI1) de Homo sapiens, ARNm.	1899
OSTM1	NM_014028	Proteína 1 transmembrana asociada a osteopetrosis (OSTM1) de Homo sapiens, ARNm.	1900

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
PGM3	NM_001199917	Fosfoglucomutasa 3 (PGM3) de Homo sapiens, variante de transcripción 1, ARNm.	1901
PGM3	NM_001199918	Fosfoglucomutasa 3 (PGM3) de Homo sapiens, variante de transcripción 3, ARNm.	1902
PGM3	NM_015599	Fosfoglucomutasa 3 (PGM3) de Homo sapiens, variante de transcripción 2, ARNm.	1903
PGM3	NM_001199919	Fosfoglucomutasa 3 (PGM3) de Homo sapiens, variante de transcripción 4, ARNm.	1904
PIAS1	NM_016166	Inhibidor de proteína de STAT 1 activada (PIAS1) de Homo sapiens, variante de transcripción 2, ARNm.	1905
PIK3R1	NM_181523	Subunidad 1 reguladora de fosfoinositido-3-quinasa (PIK3R1) de Homo sapiens, variante de transcripción 1, ARNm.	1906
PIK3R1	NM_181524	Subunidad 1 reguladora de fosfoinositido-3-quinasa (PIK3R1) de Homo sapiens, variante de transcripción 3, ARNm.	1907
PIK3R1	NM_181504	Subunidad 1 reguladora de fosfoinositido-3-quinasa (PIK3R1) de Homo sapiens, variante de transcripción 2, ARNm.	1908
PIK3R1	NM_001242466	Subunidad 1 reguladora de fosfoinositido-3-quinasa (PIK3R1) de Homo sapiens, variante de transcripción 4, ARNm.	1909
PLCG2	NM_002661	Fosfolipasa C gamma 2 (PLCG2) de Homo sapiens, ARNm.	1910
PMS2	NM_000535	Homólogo 2 de PMS1, componente del sistema de reparación de emparejamientos erróneos (PMS2) de Homo sapiens, variante de transcripción 1, ARNm.	1911
PNP	NM_000270	Purina nucleósido fosforilasa (PNP) de Homo sapiens, ARNm.	1912
POLA1	NM_016937	Polimerasa (dirigida al ADN), alfa 1, subunidad catalítica (POLA1) de Homo sapiens, ARNm.	1913
POLE	NM_006231	ADN polimerasa épsilon, subunidad catalítica (POLE) de Homo sapiens, ARNm.	1914
PRF1	NM_001083116	Perforina 1 (PRF1) de Homo sapiens, variante de transcripción 2, ARNm.	1915
PRF1	NM_005041	Perforina 1 (PRF1) de Homo sapiens, variante de transcripción 1, ARNm.	1916
PRKCD	NM_006254	Proteína quinasa C delta (PRKCD) de Homo sapiens, variante de transcripción 1, ARNm.	1917
PRKCD	NM_212539	Proteína quinasa C delta (PRKCD) de Homo sapiens, variante de transcripción 2, ARNm.	1918
PRKDC	NM_001081640	Proteína quinasa, activada por ADN, polipéptido catalítico (PRKDC) de Homo sapiens, variante de transcripción 2, ARNm.	1919
PRKDC	NM_006904	Proteína quinasa, activada por ADN, polipéptido catalítico (PRKDC) de Homo sapiens, variante de transcripción 1, ARNm.	1920
PROC	NM_000312	Proteína C, inactivadora de los factores de coagulación Va y Villa (PROC) de Homo sapiens, ARNm.	1921
PSMB8	NM_004159	Proteasoma (prosoma, macropaina) subunidad, tipo beta, 8 (PSMB8) de Homo sapiens, variante de transcripción 1, ARNm.	1922
PSMB8	NM_148919	Proteasoma (prosoma, macropaina) subunidad, tipo beta, 8 (PSMB8) de Homo sapiens, variante de transcripción 2, ARNm.	1923
PTEN	NM_000314	Fosfatasa y homólogo de tensina (PTEN) de Homo sapiens, variante de transcripción 1, ARNm.	1924

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
PTPRC	NM_001267798	Proteína tirosina fosfatasa, receptor tipo C (PTPRC) de Homo sapiens, variante de transcripción 5, ARNm.	1925
PTPRC	NM_002838	Proteína tirosina fosfatasa, receptor tipo C (PTPRC) de Homo sapiens, variante de transcripción 1, ARNm.	1926
PTPRC	NM_080921	Proteína tirosina fosfatasa, receptor tipo C (PTPRC), variante de transcripción 2, ARNm.	1927
PTPRC	NR_052021	Proteína tirosina fosfatasa, receptor tipo C (PTPRC) de Homo sapiens, variante de transcripción 4, ARN no codificante.	1928
PURA	NM_005859	Proteína A de unión a elementos ricos en purinas (PURA) de Homo sapiens, ARNm.	1929
RAB27A	NM_183235	RAB27A, miembro de la familia de oncogenes RAS (RAB27A) de Homo sapiens, variante de transcripción 3, ARNm.	1930
RAB27A	NM_183236	RAB27A, miembro de la familia de oncogenes RAS (RAB27A) de Homo sapiens, variante de transcripción 4, ARNm.	1931
RAB27A	NM_004580	RAB27A, miembro de la familia de oncogenes RAS (RAB27A) de Homo sapiens, variante de transcripción 1, ARNm.	1932
RAB27A	NM_183234	RAB27A, miembro de la familia de oncogenes RAS (RAB27A) de Homo sapiens, variante de transcripción 2, ARNm.	1933
RAB7A	NM_004637	RAB7A, miembro de la familia oncogénica RAS (RAB7A) de Homo sapiens, ARNm.	1934
RABGEF1	NM_014504	Factor de intercambio de nucleótidos de guanina RAB (GEF) 1 (RABGEF1) de Homo sapiens, variante de transcripción 4, ARNm.	1935
RAC2	NM_002872	Sustrato 2 de toxina botulínica C3 relacionado con ras (familia rho, proteína pequeña de unión a GTP Rac2) (RAC2) de Homo sapiens, ARNm.	1936
RAD51	NM_001164270	RAD51 recombinasa (RAD51) de Homo sapiens, variante de transcripción 3, ARNm.	1937
RAD51	NM_002875	RAD51 recombinasa (RAD51) de Homo sapiens, variante de transcripción 1, ARNm.	1938
RAD51	NM_133487	RAD51 recombinasa (RAD51) de Homo sapiens, variante de transcripción 2, ARNm.	1939
RAD51	NM_001164269	RAD51 recombinasa (RAD51) de Homo sapiens, variante de transcripción 4, ARNm.	1940
RAG1	NM_000448	Gen 1 activador de la recombinación (RAG1) de Homo sapiens, ARNm.	1941
RAG2	NM_000536	Gen 2 activador de la recombinación (RAG2) de Homo sapiens, variante de transcripción 1, ARNm.	1942
RAG2	NM_001243785	Gen 2 activador de la recombinación (RAG2) de Homo sapiens, variante de transcripción 3, ARNm.	1943
RAG2	NM_001243786	Gen 2 activador de la recombinación (RAG2) de Homo sapiens, variante de transcripción 4, ARNm.	1944
RBCK1	NM_006462	Dedo de zinc tipo RANBP2 y tipo C3HC4 que contiene 1 (RBCK1) de Homo sapiens, variante de transcripción 1, ARNm.	1945
RBCK1	NM_031229	Dedo de zinc tipo RANBP2 y tipo C3HC4 que contiene 1 (RBCK1) de Homo sapiens, variante de transcripción 2, ARNm.	1946
RFX5	NM_000449	Factor regulador X5 (RFX5) de Homo sapiens, variante de transcripción 1, ARNm.	1947

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
RFX5	NM_001025603	Factor regulador X5 (RFXS) de Homo sapiens, variante de transcripción 2, ARNm.	1948
RFXANK	NM_003721	Proteína que contiene anquirina asociada al factor regulador X (RFXANK) de Homo sapiens, variante de transcripción 1, ARNm.	1949
RFXANK	NM_134440	Proteína que contiene anquirina asociada al factor regulador X (RFXANK) de Homo sapiens, variante de transcripción 2, ARNm.	1950
RFXAP	NM_000538	Proteína asociada al factor regulador X (RFXAP) de Homo sapiens, ARNm.	1951
RIPK1	NM_003804	Serina-treonina quinasa 1 que interactúa con el receptor (TNFRSF) (RIPK1) de Homo sapiens, ARNm.	1952
RIPK3	NM_006871	Serina-treonina quinasa 3 que interactúa con el receptor (RIPK3) de Homo sapiens, ARNm.	1953
RMRP	NR_003051	Componente del ARN de la endorribonucleasa de procesamiento del ARN mitocondrial (RMRP) de Homo sapiens, ARN RNasa MRP.	1954
RNASEH2A	NM_006397	Ribonucleasa H2, subunidad A (RNASAH2A) de Homo sapiens, ARNm.	1955
RNASEH2B	NM_001142279	Ribonucleasa H2, subunidad B (RNASAH2B) de Homo sapiens, variante de transcripción 2, ARNm.	1956
RNASEH2B	NM_024570	Ribonucleasa H2, subunidad B (RNASAH2B) de Homo sapiens, variante de transcripción 1, ARNm.	1957
RNASEH2C	NM_032193	Ribonucleasa H2, subunidad C (RNASAH2C) de Homo sapiens, ARNm.	1958
RNASEL	NM_021133	Ribonucleasa L (dependiente de 2',5'-oligoisoadenilato sintetasa) (RNASAL) de Homo sapiens, ARNm.	1959
RNF168	NM_152617	Proteína dedo anular 168 (RNF168) de Homo sapiens, ARNm.	1960
RNF31	NM_017999	Proteína dedo anular 31 (RNF31) de Homo sapiens, ARNm.	1961
RNU4ATAC	NR_023343	ARN, nuclear pequeño U4atac (empalme dependiente de U12) (RNU4ATAC) de Homo sapiens, ARN nuclear pequeño.	1962
RTEL1	NM_016434	Regulador del alargamiento del telómero helicasa 1 (RTEL1) de Homo sapiens, variante de transcripción 1, ARNm.	1963
RTEL1	NM_032957	Regulador del alargamiento del telómero helicasa 1 (RTEL1) de Homo sapiens, variante de transcripción 2, ARNm.	1964
RTEL1-TNFRSF6B	NR_037882	Lectura completa RTEL1-TNFRSF6B (candidato a NMD) (RTEL1-TNFRSF6B) de Homo sapiens, ARN no codificante largo.	1965
SALL2	NM_005407	Factor 2 de transcripción similar a spalt (SALL2) de Homo sapiens, variante de transcripción 1, ARNm.	1966
SAMHD1	NM_015474	Dominio SAM y dominio 1 HD (SAMHD1) de Homo sapiens, ARNm.	1967
SBDS	NM_016038	Síndrome de Shwachman-Bodian-Diamond (SBDS) de Homo sapiens, ARNm.	1968
SH2D1A	NM_001114937	Dominio SH2 que contiene 1A (SH2D1A) de Homo sapiens, variante de transcripción 2, ARNm.	1969
SH2D1A	NM_002351	Dominio SH2 que contiene 1A (SH2D1A) de Homo sapiens, variante de transcripción 1, ARNm.	1970
SHARPIN	NM_030974	Interactuador del dominio RH asociado a SHANK (SHARPIN) de Homo sapiens, variante de transcripción 1, ARNm.	1971

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
SHARPIN	NR_038270	Interactuador del dominio RH asociado a SHANK (SHARPIN) de Homo sapiens, variante de transcripción 2, ARN no codificante.	1972
SKIV2L	NM_006929	Similar a la actividad 2 viralicida superasesina (S. cerevisiae) (SKIV2L) de Homo sapiens, ARNm.	1973
SLC37A4	NM_001164277	Familia 37 de transportadores de soluto (transportador de glucosa-6-fosfato), miembro 4 (SLC37A4) de Homo sapiens, variante de transcripción 1, ARNm.	1974
SLC37A4	NM_001164278	Familia 37 de transportadores de soluto (transportador de glucosa-6-fosfato), miembro 4 (SLC37A4) de Homo sapiens, variante de transcripción 2, ARNm.	1975
SLC37A4	NM_001164279	Familia 37 de transportadores de soluto (transportador de glucosa-6-fosfato), miembro 4 (SLC37A4) de Homo sapiens, variante de transcripción 3, ARNm.	1976
SLC37A4	NM_001467	Familia 37 de transportadores de soluto (transportador de glucosa-6-fosfato), miembro 4 (SLC37A4) de Homo sapiens, variante de transcripción 4, ARNm.	1977
SLC37A4	NM_001164280	Familia 37 de transportadores de soluto (transportador de glucosa-6-fosfato), miembro 4 (SLC37A4) de Homo sapiens, variante de transcripción 5, ARNm.	1978
SLC46A1	NM_001242366	Familia 46 de transportadores de soluto miembro 1 (SLC46A1) de Homo sapiens, variante de transcripción 2, ARNm.	1979
SLC46A1	NM_080669	Familia 46 de transportadores de soluto miembro 1 (SLC46A1) de Homo sapiens, variante de transcripción 1, ARNm.	1980
SLC8A1	NM_001112800	Familia 8 de transportadores de soluto miembro A1 (SLC8A1) de Homo sapiens, variante de transcripción B, ARNm.	1981
SLC8A1	NM_001112801	Familia 8 de transportadores de soluto miembro A1 (SLC8A1) de Homo sapiens, variante de transcripción C, ARNm.	1982
SLC8A1	NM_001112802	Familia 8 de transportadores de soluto miembro A1 (SLC8A1) de Homo sapiens, variante de transcripción D, ARNm.	1983
SLC8A1	NM_001252624	Familia 8 de transportadores de soluto miembro A1 (SLC8A1) de Homo sapiens, variante de transcripción E, ARNm.	1984
SLC8A1	NM_021097	Familia 8 de transportadores de soluto miembro A1 (SLC8A1) de Homo sapiens, variante de transcripción A, ARNm.	1985
SMAD2	NM_001003652	Familia SMAD miembro 2 (SMAD2) de Homo sapiens, variante de transcripción 2, ARNm.	1986
SMAD2	NM_001135937	Familia SMAD miembro 2 (SMAD2) de Homo sapiens, variante de transcripción 3, ARNm.	1987
SMAD2	NM_005901	Familia SMAD miembro 2 (SMAD2) de Homo sapiens, variante de transcripción 1, ARNm.	1988
SMAD3	NM_005902	Familia SMAD miembro 3 (SMAD3) de Homo sapiens, variante de transcripción 1, ARNm.	1989
SMAD3	NM_001145102	Familia SMAD miembro 3 (SMAD3) de Homo sapiens, variante de transcripción 2, ARNm.	1990
SMAD3	NM_001145103	Familia SMAD miembro 3 (SMAD3) de Homo sapiens, variante de transcripción 3, ARNm.	1991
SMAD3	NM_001145104	Familia SMAD miembro 3 (SMAD3) de Homo sapiens, variante de transcripción 4, ARNm.	1992
SMAD4	NM_005359	Familia SMAD miembro 4 (SMAD4) de Homo sapiens, ARNm.	1993

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
SNAP29	NM_004782	Proteína asociada a sinaptosomas, 29 kDa (SNAP29) de Homo sapiens, ARNm.	1994
SNAR-A1	NR_004435	ARN A1 pequeño asociado a ILF3/NF90 (SNAR-A1) de Homo sapiens, ARN nuclear pequeño.	1995
SNAR-A10	NR_024229	ARN A10 pequeño asociado a ILF3/NF90 (SNAR-A10) de Homo sapiens, ARN nuclear pequeño.	1996
SNAR-A11	NR_024225	ARN A11 pequeño asociado a ILF3/NF90 (SNAR-A11) de Homo sapiens, ARN nuclear pequeño.	1997
SNAR-A12	NR_004437	ARN A12 pequeño asociado a ILF3/NF90 (SNAR-A12) de Homo sapiens, ARN nuclear pequeño.	1998
SNAR-A13	NR_024216	ARN A13 pequeño asociado a ILF3/NF90 (SNAR-A13) de Homo sapiens, ARN nuclear pequeño.	1999
SNAR-A14	NR_024242	ARN A14 pequeño asociado a ILF3/NF90 (SNAR-A14) de Homo sapiens, ARN nuclear pequeño.	2000
SNAR-A2	NR_004436	ARN A2 pequeño asociado a ILF3/NF90 (SNAR-A2) de Homo sapiens, ARN nuclear pequeño.	2001
SNAR-A3	NR_024214	ARN A3 pequeño asociado a ILF3/NF90 (SNAR-A3) de Homo sapiens, ARN nuclear pequeño.	2002
SNAR-A4	NR_024215	ARN A4 pequeño asociado a ILF3/NF90 (SNAR-A4) de Homo sapiens, ARN nuclear pequeño.	2003
SNAR-A5	NR_024223	ARN A5 pequeño asociado a ILF3/NF90 (SNAR-A5) de Homo sapiens, ARN nuclear pequeño.	2004
SNAR-A6	NR_024227	ARN A6 pequeño asociado a ILF3/NF90 (SNAR-A6) de Homo sapiens, ARN nuclear pequeño.	2005
SNAR-A7	NR_024224	ARN A7 pequeño asociado a ILF3/NF90 (SNAR-A7) de Homo sapiens, ARN nuclear pequeño.	2006
SNAR-A8	NR_024228	ARN A8 pequeño asociado a ILF3/NF90 (SNAR-A8) de Homo sapiens, ARN nuclear pequeño.	2007
SNAR-A9	NR_024226	ARN A9 pequeño asociado a ILF3/NF90 (SNAR-A9) de Homo sapiens, ARN nuclear pequeño.	2008
SNAR-B1	NR_024231	ARN B1 pequeño asociado a ILF3/NF90 (SNAR-B1) de Homo sapiens, ARN nuclear pequeño.	2009
SNAR-B2	NR_024230	ARN B2 pequeño asociado a ILF3/NF90 (SNAR-B2) de Homo sapiens, ARN nuclear pequeño.	2010
SNAR-C1	NR_024220	ARN C1 pequeño asociado a ILF3/NF90 (SNAR-C1) de Homo sapiens, ARN nuclear pequeño.	2011
SNAR-C2	NR_024217	ARN C2 pequeño asociado a ILF3/NF90 (SNAR-C2) de Homo sapiens, ARN nuclear pequeño.	2012
SNAR-C3	NR_024221	ARN C3 pequeño asociado a ILF3/NF90 (SNAR-C3) de Homo sapiens, ARN nuclear pequeño.	2013
SNAR-C4	NR_024218	ARN C4 pequeño asociado a ILF3/NF90 (SNAR-C4) de Homo sapiens, ARN nuclear pequeño.	2014
SNAR-C5	NR_024219	ARN C5 pequeño asociado a ILF3/NF90 (SNAR-CS) de Homo sapiens, ARN nuclear pequeño.	2015
SNAR-D	NR_024243	ARN D pequeño asociado a ILF3/NF90 (SNAR-D) de Homo sapiens, ARN nuclear pequeño.	2016
SNAR-E	NR_024258	ARN E pequeño asociado a ILF3/NF90 (SNAR-E) de Homo sapiens, ARN nuclear pequeño.	2017

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
SNAR-F	NR_004384	ARN F pequeño asociado a ILF3/NF90 (SNAR-F) de Homo sapiens, ARN nuclear pequeño.	2018
SNAR-G1	NR_004383	ARN G1 pequeño asociado a ILF3/NF90 (SNAR-G1) de Homo sapiens, ARN nuclear pequeño.	2019
SNAR-G2	NR_024244	ARN G2 pequeño asociado a ILF3/NF90 (SNAR-G2) de Homo sapiens, ARN nuclear pequeño.	2020
SNAR-H	NR_024342	ARN H pequeño asociado a ILF3/NF90 (SNAR-H) de Homo sapiens, ARN nuclear pequeño.	2021
SNAR-I	NR_024343	ARN I pequeño asociado a ILF3/NF90 (SNAR-I) de Homo sapiens, ARN nuclear pequeño.	2022
SNCA	NM_000345	Sinucleína, alfa (componente no A4 del precursor amiloide) (SNCA) de Homo sapiens, variante de transcripción 1, ARNm.	2023
SNCA	NM_001146054	Sinucleína, alfa (componente no A4 del precursor amiloide) (SNCA) de Homo sapiens, variante de transcripción 2, ARNm.	2024
SNCA	NM_001146055	Sinucleína, alfa (componente no A4 del precursor amiloide) (SNCA) de Homo sapiens, variante de transcripción 3, ARNm.	2025
SNCA	NM_007308	Sinucleína, alfa (componente no A4 del precursor amiloide) (SNCA) de Homo sapiens, variante de transcripción 4, ARNm.	2026
SNX10	NM_013322	Clasificación de nexina 10 (SNX10) de Homo sapiens, variante de transcripción 2, ARNm.	2027
SNX10	NM_001199835	Clasificación de nexina 10 (SNX10) de Homo sapiens, variante de transcripción 1, ARNm.	2028
SNX10	NM_001199837	Clasificación de nexina 10 (SNX10) de Homo sapiens, variante de transcripción 3, ARNm.	2029
SNX10	NM_001199838	Clasificación de nexina 10 (SNX10) de Homo sapiens, variante de transcripción 4, ARNm.	2030
SNX10	NR_037670	Clasificación de nexina 10 (SNX10) de Homo sapiens, variante de transcripción 5, ARN no codificante.	2031
SP110	NM_004509	Proteína del cuerpo nuclear SP110 (SP110) de Homo sapiens, variante de transcripción a, ARNm.	2032
SP110	NM_080424	Proteína del cuerpo nuclear SP110 (SP110) de Homo sapiens, variante de transcripción c, ARNm.	2033
SP110	NM_001185015	Proteína del cuerpo nuclear SP110 (SP110) de Homo sapiens, variante de transcripción d, ARNm.	2034
SP110	NM_004510	Proteína del cuerpo nuclear SP110 (SP110) de Homo sapiens, variante de transcripción b, ARNm.	2035
SP140	NM_001005176	Proteína del cuerpo nuclear SP140 (SP140) de Homo sapiens, variante de transcripción 2, ARNm.	2036
SP140	NM_007237	Proteína del cuerpo nuclear SP140 (SP140) de Homo sapiens, variante de transcripción 1, ARNm.	2037
SPINK5	NM_001127698	Inhibidor de serina peptidasa, Kazal tipo 5 (SPINK5) de Homo sapiens, variante de transcripción 1, ARNm.	2038
SPINK5	NM_006846	Inhibidor de la serina peptidasa, Kazal tipo 5 (SPINK5) de Homo sapiens, variante de transcripción 2, ARNm.	2039
SPINK5	NM_001127699	Inhibidor de la serina peptidasa, Kazal tipo 5 (SPINK5) de Homo sapiens, variante de transcripción 3, ARNm.	2040
SQSTM1	NM_003900	Sequestosoma 1 (SQSTM1) de Homo sapiens, variante de transcripción 1, ARNm.	2041

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
SQSTM1	NM_001142298	Sequestosoma 1 (SQSTM1) de Homo sapiens, variante de transcripción 2, ARNm.	2042
SQSTM1	NM_001142299	Sequestosoma 1 (SQSTM1) de Homo sapiens, variante de transcripción 3, ARNm.	2043
SRSF1	NM_001078166	Factor 1 de empalme rico en serina y arginina (SRSF1) de Homo sapiens, variante de transcripción 2, ARNm.	2044
SRSF1	NM_006924	Factor de empalme 1 rico en serina y arginina (SRSF1) de Homo sapiens, variante de transcripción 1, ARNm.	2045
SRSF1	NR_034041	Factor de empalme 1 rico en serina y arginina (SRSF1) de Homo sapiens, variante de transcripción 3, ARN no codificante.	2046
STAT1	NM_007315	Transductor de señales y activador de la transcripción 1 (STAT1) de Homo sapiens, variante de transcripción alfa, ARNm.	2047
STAT1	NM_139266	Transductor de señales y activador de la transcripción 1 (STAT1) de Homo sapiens, variante de transcripción beta, ARNm.	2048
STAT2	NM_005419	Transductor de señales y activador de la transcripción 2, 113 kDa (STAT2) de Homo sapiens, variante de transcripción 1, ARNm.	2049
STAT2	NM_198332	Transductor de señales y activador de la transcripción 2, 113 kDa (STAT2) de Homo sapiens, variante de transcripción 2, ARNm.	2050
STAT3	NM_003150	Transductor de señales y activador de la transcripción 3 (STAT3) de Homo sapiens, variante de transcripción 2, ARNm.	2051
STAT3	NM_139276	Transductor de señales y activador de la transcripción 3 (STAT3) de Homo sapiens, variante de transcripción 1, ARNm.	2052
STAT3	NM_213662	Transductor de señales y activador de la transcripción 3 (STAT3) de Homo sapiens, variante de transcripción 3, ARNm.	2053
STAT5B	NM_012448	transductor de señales y activador de la transcripción 5B (STAT5B) de Homo sapiens, ARNm.	2054
STIM1	NM_003156	Molécula de interacción estromal 1 (STIM1) de Homo sapiens, variante de transcripción 2, ARNm.	2055
STK4	NM_006282	Serina/treonina quinasa 4 (STK4) de Homo sapiens, ARNm.	2056
STX11	NM_003764	Sintaxina 11 (STX11) de Homo sapiens, ARNm.	2057
STXBP2	NM_001127396	Proteína 2 de unión a sintaxina (STXBP2) de Homo sapiens, variante de transcripción 2, ARNm.	2058
STXBP2	NM_001272034	Proteína 2 de unión a sintaxina (STXBP2) de Homo sapiens, variante de transcripción 3, ARNm.	2059
STXBP2	NM_006949	Proteína 2 de unión a sintaxina (STXBP2) de Homo sapiens, variante de transcripción 1, ARNm.	2060
STXBP2	NR_073560	Proteína 2 de unión a sintaxina (STXBP2) de Homo sapiens, variante de transcripción 4, ARN no codificante.	2061
SYNCRIP	NM_001159673	Proteína que interactúa con el ARN citoplasmático de unión a sinaptotagmina (SYNCRIP) de Homo sapiens, variante de transcripción 2, ARNm.	2062
SYNCRIP	NM_001159674	Proteína que interactúa con el ARN citoplasmático de unión a sinaptotagmina (SYNCRIP) de Homo sapiens, variante de transcripción 3, ARNm.	2063
SYNCRIP	NM_001159676	Proteína que interactúa con el ARN citoplasmático de unión a sinaptotagmina (SYNCRIP) de Homo sapiens, variante de transcripción 5, ARNm.	2064

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
SYNCRIP	NM_001159677	Proteína que interactúa con el ARN citoplasmático de unión a sinaptotagmina (SYNCRIP) de Homo sapiens, variante de transcripción 6, ARNm.	2065
SYNCRIP	NM_001253771	Proteína que interactúa con el ARN citoplasmático de unión a sinaptotagmina (SYNCRIP) de Homo sapiens, variante de transcripción 7, ARNm.	2066
SYNCRIP	NM_001159675	Proteína que interactúa con el ARN citoplasmático de unión a sinaptotagmina (SYNCRIP) de Homo sapiens, variante de transcripción 4, ARNm.	2067
SYNCRIP	NM_006372	Proteína que interactúa con el ARN citoplasmático de unión a sinaptotagmina (SYNCRIP) de Homo sapiens, variante de transcripción 1, ARNm.	2068
T	NM_001270484	Factor de transcripción de braquiuria T (T) de Homo sapiens, variante de transcripción 2, ARNm.	2069
T	NM_003181	Factor de transcripción de braquiuria T (T) de Homo sapiens, variante de transcripción 1, ARNm.	2070
TAP1	NM_000593	Transportador 1, miembro de la subfamilia B de casetes de unión a ATP (TAP1) de Homo sapiens, variante de transcripción 1, ARNm.	2071
TAP2	NM_018833	Transportador 2, miembro de la subfamilia B de casetes de unión a ATP (TAP2) de Homo sapiens, variante de transcripción 2, ARNm.	2072
TAP2	NM_000544	Transportador 2, miembro de la subfamilia B de casetes de unión a ATP (TAP2) de Homo sapiens, variante de transcripción 1, ARNm.	2073
TAPBP	NM_003190	Proteína de unión a TAP (tapasina) (TAPBP) de Homo sapiens, variante de transcripción 1, ARNm.	2074
TAPBP	NM_172209	Proteína de unión a TAP (tapasina) (TAPBP) de Homo sapiens, variante de transcripción 3, ARNm.	2075
TAPBP	NM_172208	Proteína de unión a TAP (tapasina) (TAPBP) de Homo sapiens, variante de transcripción 2, ARNm.	2076
TAZ	NM_000116	Tafazina (TAZ) de Homo sapiens, variante de transcripción 1, ARNm.	2077
TAZ	NM_181312	Tafazina (TAZ) de Homo sapiens, variante de transcripción 3, ARNm.	2078
TAZ	NM_181311	Tafazina (TAZ) de Homo sapiens, variante de transcripción 2, ARNm.	2079
TAZ	NM_181313	Tafazina (TAZ) de Homo sapiens, variante de transcripción 4, ARNm.	2080
TAZ	NR_024048	Tafazina (TAZ) de Homo sapiens, variante de transcripción 5, ARN no codificante.	2081
TBK1	NM_013254	Quinasa 1 de unión a TANK (TBK1) de Homo sapiens, ARNm.	2082
TBX1	NM_005992	T-box 1 (TBX1) de Homo sapiens, variante de transcripción B, ARNm.	2083
TBX1	NM_080646	T-box 1 (TBX1) de Homo sapiens, variante de transcripción A, ARNm.	2084
TBX1	NM_080647	T-box 1 (TBX1) de Homo sapiens, variante de transcripción C, ARNm.	2085

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
TCIRG1	NM_006019	Regulador 1 inmunitario de células T, ATPasa H+ que transporta la subunidad a3 de V0 (TCIRG1) de Homo sapiens, variante de transcripción 1, ARNm.	2086
TCIRG1	NM_006053	Regulador 1 inmunitario de células T, ATPasa H+ que transporta la subunidad a3 de V0 (TCIRG1) de Homo sapiens, variante de transcripción 2, ARNm.	2087
TICAM1	NM_182919	Molécula 1 adaptadora del receptor tipo Toll (TICAM1) de Homo sapiens, ARNm.	2088
TLR3	NM_003265	Receptor 3 tipo toll (TLR3) de Homo sapiens, ARNm.	2089
TLR4	NM_003266	Receptor 4 tipo toll (TLR4) de Homo sapiens, variante de transcripción 3, ARNm.	2090
TLR4	NM_138554	Receptor 4 tipo toll (TLR4) de Homo sapiens, variante de transcripción 1, ARNm.	2091
TLR4	NM_138557	Receptor 4 tipo Toll (TLR4) de Homo sapiens, variante de transcripción 4, ARNm.	2092
TMMS173	NM_198282	Proteína 173 transmembrana (TMMS173) de Homo sapiens, ARNm.	2093
TNF	NM_000594	Factor de necrosis tumoral (TNF) de Homo sapiens, ARNm.	2094
TNFAIP3	NM_001270507	Proteína 3 inducida por TNF alfa (TNFAIP3) de Homo sapiens, variante de transcripción 2, ARNm.	2095
TNFAIP3	NM_001270508	Proteína 3 inducida por TNF alfa (TNFAIP3) de Homo sapiens, variante de transcripción 1, ARNm.	2096
TNFAIP3	NM_006290	Proteína 3 inducida por TNF alfa (TNFAIP3) de Homo sapiens, variante de transcripción 3, ARNm.	2097
TNFRSF11A	NM_003839	Superfamilia de receptores del factor de necrosis tumoral, miembro 11a, activador de NFKB (TNFRSF11A) de Homo sapiens, variante de transcripción 1, ARNm.	2098
TNFRSF11A	NM_001270949	Superfamilia de receptores del factor de necrosis tumoral, miembro 11a, activador de NFKB (TNFRSF11A) de Homo sapiens, variante de transcripción 2, ARNm.	2099
TNFRSF11A	NM_001270950	Superfamilia de receptores del factor de necrosis tumoral, miembro 11a, activador de NFKB (TNFRSF11A) de Homo sapiens, variante de transcripción 3, ARNm.	2100
TNFRSF11A	NM_001270951	Superfamilia de receptores del factor de necrosis tumoral, miembro 11a, activador de NFKB (TNFRSF11A) de Homo sapiens, variante de transcripción 4, ARNm.	2101
TNFRSF11B	NM_002546	Superfamilia de receptores del factor de necrosis tumoral, miembro 11b (TNFRSF11B) de Homo sapiens, ARNm.	2102
TNFRSF13B	NM_012452	Superfamilia de receptores de TNF miembro 13B (TNFRSF13B) de Homo sapiens, ARNm.	2103
TNFRSF4	NM_003327	Superfamilia de receptores de TNF miembro 4 (TNFRSF4) de Homo sapiens, ARNm.	2104
TNFRSF8	NM_001243	Superfamilia de receptores de TNF miembro 8 (TNFRSF8) de Homo sapiens, variante de transcripción 1, ARNm.	2105
TNFSF11	NM_003701	Superfamilia del factor de necrosis tumoral (ligando), miembro 11 (TNFSF11) de Homo sapiens, variante de transcripción 1, ARNm.	2106
TNFSF11	NM_033012	Superfamilia del factor de necrosis tumoral (ligando), miembro 11 (TNFSF11) de Homo sapiens, variante de transcripción 2, ARNm.	2107

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
TNFSF12	NM_003809	Superfamilia del factor de necrosis tumoral miembro 12 (TNFSF12) de Homo sapiens, variante de transcripción 1, ARNm.	2108
TNFSF12	NR_037146	Superfamilia del factor de necrosis tumoral miembro 12 (TNFSF12) de Homo sapiens, variante de transcripción 2, ARN no codificante.	2109
TP53	NM_000546	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 1, ARNm.	2110
TP53	NM_001126112	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 2, ARNm.	2111
TP53	NM_001126113	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 4, ARNm.	2112
TP53	NM_001126114	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 3, ARNm.	2113
TP53	NM_001126115	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 5, ARNm.	2114
TP53	NM_001126116	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 6, ARNm.	2115
TP53	NM_001126117	Homo sapiens proteína tumoral p53 (TP53), variante de transcripción 7, ARNm.	2116
TP53	NM_001126118	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 8, ARNm.	2117
TP53	NM_001276695	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 4, ARNm.	2118
TP53	NM_001276696	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 3, ARNm.	2119
TP53	NM_001276697	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 5, ARNm.	2120
TP53	NM_001276698	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 6, ARNm.	2121
TP53	NM_001276699	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 7, ARNm.	2122
TP53	NM_001276760	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 1, ARNm.	2123
TP53	NM_001276761	Proteína tumoral p53 (TP53) de Homo sapiens, variante de transcripción 2, ARNm.	2124
TRAF3	NM_001199427	Factor 3 asociado al receptor de TNF (TRAF3) de Homo sapiens, variante de transcripción 4, ARNm.	2125
TRAF3	NM_003300	Factor 3 asociado al receptor de TNF (TRAF3) de Homo sapiens, variante de transcripción 3, ARNm.	2126
TRAF3	NM_145725	Factor 3 asociado al receptor de TNF (TRAF3) de Homo sapiens, variante de transcripción 1, ARNm.	2127
TRAF3	NM_145726	Factor 3 asociado al receptor de TNF (TRAF3) de Homo sapiens, variante de transcripción 2, ARNm.	2128
TRAF6	NM_004620	Factor 6 asociado al receptor de TNF, ubiquitina proteína ligasa E3 (TRAF6) de Homo sapiens, variante de transcripción 2, ARNm.	2129
TRAF6	NM_145803	Factor 6 asociado al receptor de TNF, ubiquitina proteína ligasa E3 (TRAF6) de Homo sapiens, variante de transcripción 1, ARNm.	2130
TREX1	NM_007248	Exonucleasa 1 de reparación principal 3 (TREX1) de Homo sapiens, variante de transcripción 5, ARNm.	2131

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
TREX1	NM_033629	Exonucleasa 1 de reparación principal 3 (TREX1) de Homo sapiens, variante de transcripción 4, ARNm.	2132
TREX1	NM_016381	Exonucleasa 1 de reparación principal 3 (TREX1) de Homo sapiens, variante de transcripción 1, ARNm.	2133
TRNT1	NM_182916	ARNt nucleotidil transferasa 1 (TRNT1) de Homo sapiens, variante de transcripción 1, ARNm.	2134
TTC7A	NM_020458	Dominio 7A de repetición de tetratricopéptido (TTC7A) de Homo sapiens, variante de transcripción 2, ARNm.	2135
TYK2	NM_003331	Tirosina quinasa 2 (TYK2) de Homo sapiens, ARNm.	2136
UNC119	NM_005148	Chaperona de unión a lípidos unc-119 (UNC119) de Homo sapiens, variante de transcripción 1, ARNm.	2137
UNC119	NM_054035	Chaperona de unión a lípidos unc-119 (UNC119) de Homo sapiens, variante de transcripción 2, ARNm.	2138
UNC13D	NM_199242	unc-13 homólogo D (UNC13D) de Homo sapiens, ARNm.	2139
UNC93B1	NM_030930	unc-93 homólogo B1 (C. elegans) (UNC93B1) de Homo sapiens, ARNm.	2140
UNG	NM_080911	Uracilo ADN glicosilasa (UNG) de Homo sapiens, variante de transcripción 2, ARNm.	2141
UNG	NM_003362	Uracilo ADN glicosilasa (UNG) de Homo sapiens, variante de transcripción 1, ARNm.	2142
USP18	NM_017414	Peptidasa 18 específica de ubiquitina (USP18) de Homo sapiens, ARNm.	2143
USP20	NM_006676	Peptidasa 20 específica de ubiquitina (USP20) de Homo sapiens, variante de transcripción 1, ARNm.	2144
USP20	NM_001008563	Peptidasa 20 específica de ubiquitina (USP20) de Homo sapiens, variante de transcripción 2, ARNm.	2145
USP20	NM_001110303	Peptidasa 20 específica de ubiquitina (USP20) de Homo sapiens, variante de transcripción 3, ARNm.	2146
VAPA	NM_003574	Proteína A asociada a VAMP (VAPA) de Homo sapiens, variante de transcripción 1, ARNm.	2147
VAPA	NM_194434	Proteína A asociada a VAMP (VAPA) de Homo sapiens, variante de transcripción 2, ARNm.	2148
VCP	NM_007126	Proteína que contiene valosina (VCP) de Homo sapiens, ARNm.	2149
VDAC1	NM_003374	Canal 1 aniónico dependiente de voltaje (VDAC1) de Homo sapiens, variante de transcripción 1, ARNm.	2150
VDAC1	NR_036624	Canal 1 aniónico dependiente de voltaje (VDAC1) de Homo sapiens, variante de transcripción 3, ARN no codificante.	2151
VDAC1	NR_036625	Canal 1 aniónico dependiente de voltaje (VDAC1) de Homo sapiens, variante de transcripción 2, ARN no codificante.	2152
VPS13B	NM_017890	Clasificación 13 de proteína vacuolar homólogo B (levadura) (VPS13B) de Homo sapiens, variante de transcripción 5, ARNm.	2153
VPS13B	NM_181661	Clasificación 13 de proteína vacuolar homólogo B (levadura) (VPS13B) de Homo sapiens, variante de transcripción 4, ARNm.	2154
VPS13B	NM_015243	Clasificación 13 de proteína vacuolar homólogo B (levadura) (VPS13B) de Homo sapiens, variante de transcripción 3, ARNm.	2155
VPS13B	NR_047582	Clasificación 13 de proteína vacuolar homólogo B (levadura) (VPS13B) de Homo sapiens, variante de transcripción 6, ARN no codificante.	2156

Símbolo del gen de RefSeq	Número de acceso de RefSeq	Descripción del ARNm	SEQ ID
VPS13B	NM_152564	Clasificación 13 de proteína vacuolar homólogo B (levadura) (VPS13B) de Homo sapiens, variante de transcripción 1, ARNm.	2157
VPS45	NM_007259	Clasificación 45 de proteína vacuolar homólogo (VPS45) de Homo sapiens, variante de transcripción 1, ARNm.	2158
WAS	NM_000377	Síndrome de Wiskott-Aldrich (WAS) de Homo sapiens, ARNm.	2159
WEE1	NM_003390	Quinasa de punto de control WEE1 G2 (WEE1) de Homo sapiens, variante de transcripción 1, ARNm.	2160
WEE1	NM_001143976	Quinasa de punto de control WEE1 G2 (WEE1) de Homo sapiens, variante de transcripción 2, ARNm.	2161
WIPF1	NM_001077269	Familia de la proteína que interactúa con WAS/WASL miembro 1 (WIPF1) de Homo sapiens, variante de transcripción 2, ARNm.	2162
WIPF1	NM_003387	Familia de la proteína que interactúa con WAS/WASL miembro 1 (WIPF1) de Homo sapiens, variante de transcripción 1, ARNm.	2163
XIAP	NM_001204401	Inhibidor de la apoptosis ligado al cromosoma X, ubiquitina proteína ligasa E3 (XIAP) de Homo sapiens, variante de transcripción 2, ARNm.	2164
XIAP	NM_001167	Inhibidor de la apoptosis ligado al cromosoma X, ubiquitina proteína ligasa E3 (XIAP) de Homo sapiens, variante de transcripción 1, ARNm.	2165
XIAP	NR_037916	Inhibidor de la apoptosis ligado al cromosoma X, ubiquitina proteína ligasa E3 (XIAP) de Homo sapiens, variante de transcripción 3, ARN no codificante.	2166
YBX1	NM_004559	proteína 1 de unión a Y-box (YBX1) de Homo sapiens, variante de transcripción 1, ARNm.	2167
YWHAZ	NM_001135699	Proteína zeta activadora de la tirosina 3-monooxigenasa/triptófano 5-monooxigenasa (YWHAZ) de Homo sapiens, variante de transcripción 3, ARNm.	2168
YWHAZ	NM_001135700	Proteína zeta activadora de la tirosina 3-monooxigenasa/triptófano 5-monooxigenasa (YWHAZ) de Homo sapiens, variante de transcripción 4, ARNm	2169
YWHAZ	NM_001135701	Proteína zeta activadora de la tirosina 3-monooxigenasa/triptófano 5-monooxigenasa (YWHAZ) de Homo sapiens, variante de transcripción 5, ARNm	2170
YWHAZ	NM_001135702	Proteína zeta activadora de la tirosina 3-monooxigenasa/triptófano 5-monooxigenasa (YWHAZ) de Homo sapiens, variante de transcripción 6, ARNm	2171
YWHAZ	NM_003406	Proteína zeta activadora de la tirosina 3-monooxigenasa/triptófano 5-monooxigenasa (YWHAZ) de Homo sapiens, variante de transcripción 1, ARNm	2172
YWHAZ	NM_145690	Proteína zeta activadora de la tirosina 3-monooxigenasa/triptófano 5-monooxigenasa (YWHAZ) de Homo sapiens, variante de transcripción 2, ARNm	2173
ZAP70	NM_001079	Cadena z de proteína quinasa 70 asociada al receptor de células T (ZAP70) de Homo sapiens, variante de transcripción 1, ARNm.	2174
ZAP70	NM_207519	Cadena z de proteína quinasa 70 asociada al receptor de células T (ZAP70) de Homo sapiens, variante de transcripción 2, ARNm.	2175
ZBTB24	NM_014797	Dedo de zinc y dominio BTB que contiene 24 (ZBTB24) de Homo sapiens, variante de transcripción 1, ARNm.	2176
ZBTB24	NM_001164313	Dedo de zinc y dominio BTB que contiene 24 (ZBTB24) de Homo sapiens, variante de transcripción 2, ARNm.	2177

La Tabla 12 enumera todas las variantes de transcripción para los genes de la Tabla 6 que no fueron 'descubiertas' por PBio sobre la base de aCGH (genes identificados por CNV). Las SEQ ID NOs corresponden a variantes de transcripción (a menudo más de una por gen).

Tabla 13: Genes para los que la carga total de variantes heterocigotas dañinas resultó ser estadísticamente mayor en los casos de PML frente a los controles del ExAC

Gen	Casos promedio	Casos promedio del ExAC	Muestras promedio del ExAC	FET promedio	FET promedio corregida (419)	OR promedio	Etnia	Solapamiento
PLCG2	17	1,806	31,277	1.43E-10	6.21E-08	10.27	EUR	EUR + AFR
RBCK1	6	187	29,324	4.27E-07	1.86E-04	24.60	EUR	
EPG5	9	764	32,835	7.11E-07	3.09E-04	10.79	EUR	
IL17F	4	61	33,346	1.67E-06	7.28E-04	54.57	EUR	
SHARPIN	8	646	32,162	2.58E-06	1.12E-03	10.84	EUR	
PRF1	8	715	33,027	4.44E-06	1.93E-03	10.04	EUR	
JAGN1	5	163	27,768	6.80E-06	2.96E-03	21.71	EUR	
TAP1	5	203	28,125	1.80E-05	7.82E-03	17.63	EUR	
POLE	11	1,660	29,108	2.84E-05	1.23E-02	5.51	EUR	EUR + AFR
LRBA	11	1,876	32,136	3.47E-05	1.51E-02	5.38	EUR	
EHF	3	49	32,588	4.83E-05	2.10E-02	48.59	EUR	
IL12B	3	58	33,112	7.44E-05	3.23E-02	41.70	EUR	
ATL2	8	31	5,041	4.03E-11	1.75E-08	90.11	AFR	
NHEJ1	6	27	4,384	5.48E-09	2.39E-06	64.56	AFR	
LYST	11	291	4,748	1.09E-08	4.76E-06	16.85	AFR	
HIVEP1	9	150	4,432	7.41E-08	3.22E-05	23.83	AFR	
AP3B1	5	46	4,937	1.69E-06	7.36E-04	33.23	AFR	
TNFRSF10A	7	149	4,626	3.28E-06	1.43E-03	15.03	AFR	
PIK3CD	7	148	4,549	3.52E-06	1.53E-03	14.87	AFR	
PLCG2	8	256	4,410	1.47E-05	6.41E-03	9.99	AFR	EUR + AFR
PNP	3	11	5,189	2.00E-05	8.69E-03	78.45	AFR	
POLE	8	297	4,752	2.48E-05	1.08E-02	9.23	AFR	EUR + AFR
MCEE	3	13	5,164	3.10E-05	1.35E-02	66.04	AFR	
DOCK2	6	173	5,023	6.39E-05	2.78E-02	11.21	AFR	
ALG12	4	43	4,252	6.73E-05	2.93E-02	23.03	AFR	

La Tabla 13 enumera los genes para los cuales la carga total de variantes heterocigotas dañinas fue encontrada estadísticamente mayor en los casos de PML versus los controles del ExAC. El análisis de la carga génica se realizó tal y como se describe a continuación con valores de corte de la frecuencia alélica menor (MAF) de 0.01, 0.02, 0.03, 0.04 y 0.05. No todos los genes sobrevivieron al análisis estadístico en todos los cortes de MAF. Para cada gen que sobrevivió a múltiples cortes de MAF, se calcularon los promedios de la prueba exacta de Fisher (FET), nominal y corregida, así como las demás métricas relevantes. Dos genes se solaparon entre los análisis de AFR y EUR. Las FET se corrigieron para pruebas múltiples con el número de genes utilizado en este estudio (419). Sólo se consideraron para su inclusión los genes para los que FET corregida era < 0.05 y en los que las variantes afectaban a > 10 % de los casos dentro de la etnia dada (>2 para AFR, >4 para EUR).

Tabla 14: Nivel superior de variantes consideradas significativas de acuerdo con el análisis de la carga de variantes

Gen	Variante (hg19)	Tipo Genómico	PML EUR 44	PML AFR 21	PML LAT 6	ExAC EUR	ExAC AFR	ExAC LAT	PML EUR OR	PML EUR FET	PML AFR OR	PML AFR FET	PML ALL OR	PML ALL FET
PLCG2	cr16:81942175, A>G	het	2	5	0	512/32261	68/4707	116/5546	2.95	0.154755	16.40	0.0000	6.49	0.0002
IFIH1	cr2:163136505, C>G	het	6	1	0	611/33155	23/5182	119/5671	8.41	0.000156	11.22	0.0927	6.38	0.0002
TCIRG1	cr11:67818269, G>A	het	0	4	0	103/33193	200/5170	60/5770	NA	NA	5.85	0.0082	7.31	0.0028
IGLL1	cr22:23917192, G>T	het	4	3	1	751/33348	603/5183	236/5762	4.34	0.017218	1.27	0.7286	3.47	0.0036
MAVS	cr20:3846397, C>T	hom	4	4	0	800/32122	684/4982	52/5681	3.82	0.023868	1.48	0.5191	3.47	0.0036
SHARPIN	cr8:145154222, G>A	het	0	4	0	2916/33177	59/4865	171/5760	2.31	0.053526	19.17	0.0001	2.68	0.0040
CHD7	cr8:6165429298, T>A	het	5	0	0	1103/33106	29/4840	93/5725	3.72	0.015268	NA	NA	2.64	0.0485
CX3CR1	cr3:39323163, A>C	hom	11	4	0	4723/31219	193/4376	1357/5491	1.67	0.088087	5.10	0.0128	1.51	0.1606
LRBA	cr4:151199060, G>A	hom	3	3	0	2260/33328	20/5195	54/5785	1.01	1.000000	43.13	0.0001	1.69	0.2736
HIVEP3	cr1:42047208, C>G	het	5	3	1	3383/32494	123/5061	902/5756	1.10	0.803620	6.69	0.0143	1.30	0.4283
IFIH1	cr2:163124051, C>T	hom	20	3	1	12107/33356	184/5199	1076/5776	1.46	0.212471	4.54	0.0374	1.21	0.4372
RNASAL	cr1:182554557, C>T	hom	7	2	0	4543/33356	78/5202	167/5785	1.20	0.658473	6.91	0.0403	1.22	0.5616

La Tabla 14 enumera el nivel superior de variantes que se consideraron significativas sobre la base del análisis de carga de variantes, tal como se describe a continuación. Para cada variante (las coordenadas del genoma se basan en UCSC hg19), se presenta información detallada del número de casos EUR y AFR que portan la variante, junto con las métricas estadísticas agregadas y específicas de cada etnia.

Tabla 16. Nivel superior de variantes encontradas con base en el análisis de la carga de variantes

Gen	Variante (hg19)	Tipo Genómico	PML EUR 44	PML AFR 21	PML LAT 5	EXAC EUR	EXAC AFR	EXAC LAT	PML EUR OR	PML EUR FET	PML AFR OR	PML AFR FET	PML ALL OR	PML ALL FET
SHARPIN	chr8:144164324, A>C	het	3	0	0	2/30,670	0/4,471	0/5,302	1122.60	0.000000	NA	NA	905.40	0.0000
RTEL1	chr20:62305450, C>T	het	0	2	0	1/32,552	0/4,838	0/5,737	NA	NA	1240.64	0.0000	1368.41	0.0000
IGLL1	chr22:23916745, G>A	het	2	0	1	19/33,348	74/5,184	9/5,783	83.53	0.000351	NA	NA	19.41	0.0006
PGM3	chr6:83884161, C>G	het	0	2	0	0/33,080	26/5,167	3/5,748	NA	NA	20.81	0.0055	44.58	0.0011
ATM	chr11:108202772, G>T	het	3	0	0	170/32,707	3/5,099	7/5,713	14.00	0.001636	NA	NA	10.78	0.0032
TMEM173	chr5:138356923, C>T	het	2	2	0	109/32,327	204/4,842	58/5,770	14.31	0.009663	2.39	0.2226	6.97	0.0033
CLCN7	chr16:1510535, C>T	het	0	2	0	1/32,896	66/5,119	0/5,732	NA	NA	8.06	0.0308	19.18	0.0055
MAVS	chr20:3843027, C>A	hom	4	2	0	80/33,206	167/5,171	46/5,779	4.04	0.021706	3.15	0.1480	3.98	0.0056
ORAI1	chr12:1220064786, G>GT	het	4	0	0	371/26,708	5/3,555	16/5,354	7.64	0.002662	NA	NA	5.76	0.0064
RBFOX1	chr16:7714909, C>T	het	0	2	0	1/33,367	69/4,802	4/5,782	NA	NA	7.37	0.0361	17.46	0.0086
MALT1	chr18:56401523, C>T	het	4	0	0	468/33,239	8/5,179	40/5,760	7.03	0.003411	NA	NA	5.14	0.0083
GFI1	chr1:92945825, G>C	het	2	1	0	236/29,111	6/4,156	39/5,114	6.68	0.039381	24.58	0.0347	6.80	0.0113
MUELLE2	chr5:169081453, G>C	het	0	2	0	48/33,350	43/5,201	27/5,786	NA	NA	12.62	0.0137	11.02	0.0155
ATM	chr11:108117787, C>T	het	2	0	0	93/33,256	2/5,151	28/5,758	16.98	0.007947	NA	NA	10.53	0.0189
SNAP29	chr22:21233389, A>G	het	3	0	0	283/32,917	21/5,149	32/5,740	8.44	0.006384	NA	NA	5.79	0.0171
TICAM1	chr19:4817657, C>T	het	0	2	0	32/31,437	7/4,814	19/5,687	NA	NA	7.03	0.0392	10.08	0.0163

Gen	Variante (hg19)	Tipo Genómico	PML EUR 44	PML AFR 21	PML LAT 5	EXAC EUR	EXAC AFR	EXAC LAT	PML EUR OR	PML EUR FET	PML AFR OR	PML AFR FET	PML ALL OR	PML ALL FET
GTPBP4	cr10:1060 218, G>A	hom	3	0	0	334/33,367	20/5,202	21/5,786	7.24	0.009825	NA	NA	5.25	0.0220
BACH1	cr21:3089 6953, T>G	het	2	0	0	134/33,122	4/5,103	6/5,778	11.72	0.014110	NA	NA	6.96	0.0227
MUELLE8	cr9:30462 8, G>A	het	2	0	0	149/33,298	2/5,161	5/5,762	10.59	0.017020	NA	NA	6.31	0.0261
STXBP2	cr19:7712 287, G>C	het	2	0	0	161/32,104	4/4,626	11/5,696	9.45	0.021028	NA	NA	7.08	0.0350
FAS	cr10:9077 1767, G>A	het	2	0	0	175/33,264	3/5,162	10/5,731	9.01	0.022902	NA	NA	6.89	0.0365
GOLGB1	cr3:12141 5370, T>C	het	3	2	0	1,111/33,34 9	26/5,179	84/5,779	2.12	0.180743	20.86	0.0055	2.71	0.0443
FUK	cr16:7050 3095, A>G	het	4	0	0	741/33,341	23/4,899	73/5,787	4.40	0.018488	NA	NA	3.13	0.0449
IL10	cr1:20684 5738, C>T	het	2	0	0	206/33,343	2/5,198	6/5,787	7.66	0.030787	NA	NA	6.06	0.0458
ITK	cr5:15659 3120, C>T	het	2	0	0	206/33,353	5/5,203	5/5,789	7.66	0.030770	NA	NA	6.01	0.0466
STIM2	cr4:27019 452, C>T	het	2	0	0	219/33,369	5/5,202	9/5,789	7.21	0.034346	NA	NA	5.57	0.0532
ASH1L	cr1:15531 7682, C>T	het	2	0	0	218/33,367	2/5,203	17/5,789	7.24	0.034067	NA	NA	5.48	0.0548
TBC1D16	cr17:7792 6526, C>T	het	3	0	0	496/31,905	15/4,845	21/5,707	4.63	0.031269	NA	NA	3.53	0.0584
LYST	cr1:23584 0495, G>T	het	3	0	0	517/33,239	10/5,156	35/5,756	4.63	0.031299	NA	NA	3.47	0.0606
SALL2	cr14:2199 3359, G>A	het	3	0	0	519/31,729	14/4,520	17/5,718	4.40	0.035528	NA	NA	3.37	0.0650
CHD7	cr6:61757 805, C>T	het	3	0	0	517/32,980	14/4,872	41/5,765	4.58	0.032169	NA	NA	3.36	0.0654
BLM	cr15:9130 6241, G>A	het	2	0	0	266/33,277	10/5,061	17/5,756	5.91	0.048875	NA	NA	4.40	0.0799
NOD2	cr16:5074 1791, C>T	het	0	2	0	265/33,369	21/5,203	2/5,789	NA	NA	25.97	0.0037	4.21	0.0860
IGLL1	cr22:2381 6583, T>C	het	2	0	0	265/33,334	21/5,183	26/5,787	5.94	0.048403	NA	NA	4.15	0.0881

Gen	Variante (hg19)	Tipo Genómico	PML EUR 44	PML AFR 21	PML LAT 5	ExAC EUR	ExAC AFR	ExAC LAT	PML EUR OR	PML EUR FET	PML AFR OR	PML AFR FET	PML ALL OR	PML ALL FET
TTC7A	chr2:47205921, C>T	het	3	0	0	589/33,202	13/5,173	61/5,759	4.05	0.043427	NA	NA	2.94	0.0891
KITLG	chr12:88900891, C>A	het	4	0	0	1,023/33,226	31/5,158	40/5,760	3.15	0.046242	NA	NA	2.36	0.0964
ATR	chr3:142281353, C>G	het	4	0	0	1,037/33,343	14/5,130	69/5,785	3.12	0.047671	NA	NA	2.33	0.1021
ATM	chr11:108123551, C>T	het	0	2	0	217/29,921	66/4,955	40/5,435	NA	NA	7.80	0.0327	3.64	0.1093
CR2	chr1:207641950, C>T	het	0	2	0	391/33,363	10/5,203	8/5,754	NA	NA	28.72	0.0031	3.09	0.1422
HIVEP2	chr6:143092151, T>C	het	3	2	0	1,718/33,370	50/4,901	209/5,769	1.35	0.494339	10.21	0.0202	1.64	0.2466
ITSN2	chr2:24431164, C>T	hom	3	2	0	2,019/33,338	17/5,186	55/5,764	1.14	0.748301	32.01	0.0025	1.55	0.3862
ITSN2	chr2:24432937, C>T	hom	3	2	0	2,026/32,472	17/4,958	56/5,672	1.10	0.753075	30.59	0.0028	1.50	0.3937
DOCK8	chr9:312134, G>A	het	3	2	0	2,114/33,251	79/5,180	16/15,768	1.06	0.757661	6.80	0.0415	1.37	0.4238
VPS13B	chr8:100205255, G>A	het	0	2	0	811/33,348	19/5,192	100/5,778	NA	NA	28.66	0.0031	1.37	0.6600
NRIP1	chr21:16339852, T>C	het	0	2	0	901/33,355	19/5,203	64/5,780	NA	NA	28.72	0.0031	1.30	0.6899

La Tabla 15 enumera el segundo nivel de variantes que se encontraron sobre la base del análisis de carga de variantes, tal como se describe a continuación. Para cada variante (las coordenadas del genoma son UCSC hg19), se presenta información detallada del número de casos EUR y AFR que portan la variante, junto con las métricas estadísticas agregadas y específicas de la etnia.

Tabla 16: Posible escenario de pruebas, basado en los principales resultados de carga de variantes

				Información al paciente		
Gen/Variante	Casos resueltos	Proporción de la cohorte (n = 70)	Método de ensayo	Enfermedad primaria	Etnia	Género
Los 4 SNV	28	40 %	genotipado	M, H, O	A, E	ambos
SNV SHARPIN, IFIH1, PLCG2	24	34 %	genotipado	M, H, O	A, E	ambos
SNV IFIH1, PLCG2	13	19 %	genotipado	M, H, O	A, E	ambos
SNV SHARPIN	13	19 %	genotipado	M, H	A, E	ambos
SNV IFIH1	7	10 %	genotipado	M, H, O	A, E	ambos
SNV PLCG2	7	10 %	genotipado	M, H	A, E	ambos
SNV CHD7	5	7 %	genotipado	M, H. O	E	ambos

La Tabla 16 enumera un escenario de prueba potencial, basado en los resultados de carga de variantes principales (reportados en la Tabla 14). El análisis es meramente ilustrativo, reconociéndose que pueden obtenerse mayores rendimientos diagnósticos ensayando un mayor número de variantes, incluidas las enumeradas en la Tabla 15. Se dan ejemplos de rendimiento diagnóstico utilizando variantes únicas, así como diversas combinaciones, incluido el uso de las 4 variantes principales. Para este conjunto de variantes, el método de prueba se describe como genotipado, a diferencia de la secuenciación de genes completos (es decir, la determinación del estado en cada una de las bases, que arroja un resultado binario, a diferencia de la identificación de variantes en otras partes de los genes relevantes).

Tabla 17: Posible escenario de pruebas utilizando genes identificados con una mayor carga de variantes heterocigotas perjudiciales en la cohorte de PML

Gen	Casos promedio	Etnia	Rendimiento específico por etnia (%)	Rendimiento global (EUR + AFR) (%)	Método de ensayo
PLCG2	17/44	EUR	38	38	Secuenciación de genes
PLCG2	8/21	AFR	38		Secuenciación de genes
POLE	8/21	AFR	38		Secuenciación de genes
POLE	11/44	EUR	25	29	Secuenciación de genes
LRBA	11/44	EUR	25		Secuenciación de genes
EPG5	9/44	EUR	20		Secuenciación de genes
SHARPIN	8/44	EUR	18		Secuenciación de genes

La Tabla 17 enumera un escenario de prueba potencial usando genes identificados por tener una mayor carga de variantes dañinas heterocigotas en la cohorte de PML (véase la Tabla 13). La naturaleza del método de ensayo es la 'secuenciación de genes', ya que las variantes no se conocen de antemano - en un ensayo de este tipo deben tenerse en cuenta todas y cada una de las variantes potencialmente perjudiciales.

Tabla 18: Resumen de genes que sobreviven los análisis de casos, carga génica y/o carga de variantes

Gen	Nivel de caso	Carga de variantes	Carga de genes
PLCG2	Sí	Sí	Sí
CHD7	Sí	Sí	
IFIH1	Sí	Sí	

Gen	Nivel de caso	Carga de variantes	Carga de genes
AP3B1	Sí		Sí
EPG5	Sí		Sí
PIK3CD	Sí		Sí
LRBA		Sí	Sí
SHARPIN		Sí	Sí

La Tabla 18 representa un resumen de los genes que sobreviven a los análisis a nivel de caso (2 o más ejemplos en las Tablas 7, 8), de carga génica y/o de carga de variantes (basados en las Tablas 13 y 14). Cabe destacar que PLCG2 satisface los 3 criterios (2 o más ejemplos, en la Tabla 8, presencia en las Tablas 13, 14). Este resumen demuestra que muchos genes han sido identificados como significativos sobre la base de métodos de análisis independientes.

Ejemplo 11 - Figuras referenciadas en este estudio

Las Figuras 1-12 representan ejemplos de datos de CNV del estudio de descubrimiento del gen de PML (71 casos de PML, (véase la Tabla 7 para información de pacientes) usando arreglo de CGH (métodos descritos en el presente documento). En cada figura/dibujo: 1) las coordenadas del genoma se enumeran en la parte superior (ensamblaje hg18, número de cromosoma y posición representada); 2) la pista de datos 1 (marcada como 'Genes') representa la ubicación de los genes de RefSeq (los exones son porciones de color gris oscuro de las barras, los intrones son porciones de color gris claro de las barras); 3) la pista de datos 2 (marcada como 'Cohorte Normal') representa el tamaño y la ubicación de las CNV encontradas en la cohorte de NVE (base de datos de control propiedad de PBio que consiste en hallazgos de CNV en sujetos aparentemente sanos - es decir, normales, ver métodos en el presente documento) con el eje y correspondiente al número de sujetos de NVE que tienen la CNV; y 4) las pistas de datos restantes son datos de CNV encontrados en pacientes individuales de PML donde el eje y corresponde a la relación log2 (véanse los métodos en el presente documento), los puntos representan sondas individuales en el microarreglo, y los segmentos de línea están desplazados positivamente (ganancia de número de copias) o negativamente (pérdida de número de copias) basado en la salida de DNAcopy, el algoritmo de llamada de CNV. Las relaciones log2 típicas para ganancias y pérdidas en la micromatriz Agilent 1M (véanse los métodos en el presente documento) y nuestros protocolos experimentales son: 0.6 para las duplicaciones, 1.0 para las triplicaciones (o duplicaciones homocigotas), -1.0 para las eliminaciones heterocigotas y <-2 (a menudo de -4.0 a -6.0) para las eliminaciones homocigotas. Los genes relevantes se marcan en la pista de datos 'Genes'.

La Figura 1 representa un ejemplo de un gen afectado por CNV de línea germinal y adquiridas. Las CNV de línea germinal que afectan al gen PRKCB incluyen al paciente PML50 con una pérdida heterocigota intrónica de 4.8 Kb (también encontrada en 7 sujetos normales) y al paciente PML11 con una ganancia intrónica de 7.3 Kb (también encontrada en 1 sujeto normal). Se encontraron CNV adquiridas en 6 pacientes con PML, una serie de ganancias en ~23.9 Mb con relaciones log2 variables, sugestivas de una población celular mixta (los experimentos de arreglo CGH se realizaron en ADN genómico derivado de la sangre).

La Figura 2 representa un ejemplo de genes potencialmente relevantes para PML (TNFRSF13C y CENPM) afectados por CNV adquiridas. Se encontraron CNV adquiridas en 9 pacientes con PML, una serie de ganancias en ~40.6 Mb con relaciones log2 variables, sugestivas de una población celular mixta (los experimentos de arreglo CGH se realizaron en ADN genómico derivado de la sangre). Los 9 pacientes con PML (véase la Tabla 7 para la información sobre los pacientes) tenían un diagnóstico primario de VIH y eran de género (3 mujeres y 6 hombres) y etnia mixtos (4 de ascendencia africana y 5 de ascendencia europea).

La Figura 3 representa un ejemplo de un gen afectado por CNV. En el paciente PML26 se detectó una pérdida heterocigota intrónica de 7.2 Kb (no encontrada en sujetos normales, pero se encuentra una pérdida adyacente en 8 sujetos normales) que afecta al gen PKHD1. Se encontraron CNV en 3 pacientes con PML, una serie de ganancias en ~51.9 Mb con relaciones log2 variables, sugestivas de una población celular mixta (los experimentos de arreglo CGH se realizaron en ADN genómico derivado de la sangre).

La Figura 4 representa un ejemplo de un gen afectado por una pérdida de CNV. La eliminación intrónica de 14.7 Kb afecta al gen BMPR2. Se detectaron eliminaciones heterocigotas en los pacientes PML58 y MVGS811-13a (también encontradas en 2 sujetos normales), y se detectó una eliminación homocigótica en el paciente PML29 (ninguna encontrada en sujetos normales). Los tres pacientes con PML son varones y su enfermedad primaria es el VIH (véase la Tabla 7).

La Figura 5 representa un ejemplo de un gen afectado por una ganancia de CNV. La ganancia exónica de 10.2 Kb interrumpe el gen COMMD6. Dos pacientes con PML, PML29 y MVGS811-13a, tienen una duplicación homocigota (relación log2 comparable a triplicaciones) basada en la observación de que los sujetos de 1000

genomas se informa que tienen esta ganancia (véase la variante esv3632749 de DGV del ensamblaje hg19, que informa que 148 de 2504 sujetos tienen esta ganancia; no se encontraron normales en la base de datos de NVE de PBio). Ambos pacientes con PML son varones y su enfermedad primaria es el VIH (véase la Tabla 7).

La Figura 6 representa un ejemplo de un gen afectado por una ganancia de CNV. La ganancia exónica de 27.4 Kb interrumpe el gen KCTD7 y el punto de rotura derecho está 16-90 Kb secuencia arriba de las variantes de transcripción RABGEF1 (RefSeq: NM_001287060, NR_104676, NM_014504, NM_001287062, NM_001287061). El paciente PML29 tiene una duplicación homocigota (relación log2 comparable a triplicaciones) basado en la observación de que los sujetos de 1000 genomas se informa que tienen esta ganancia (véase la variante esv3606925 de DGV del ensamblaje hg19, que informa que 28 de 2504 sujetos tienen esta ganancia; no se encontraron normales en la base de datos de NVE de PBio). El paciente PML63 tiene una duplicación. Ambos pacientes con PML son varones de ascendencia africana y su enfermedad primaria es el VIH (véase la Tabla 7).

La Figura 7 representa un ejemplo de un gen afectado por una ganancia de CNV. La ganancia exónica de 344 Kb interrumpe los genes FPR2 y ZNF616 (a través de los puntos de ruptura izquierdo y derecho) y los genes adicionales totalmente englobados por esta CNV son: FPR3, ZNF350, ZNF350-AS1, ZNF432, ZNF577, ZNF613, ZNF614, ZNF615, ZNF649, ZNF649-AS1, ZNF841. El paciente PML03 tiene una duplicación homocigota (relación log2 comparable a triplicaciones) basada en la observación de que 3 sujetos normales (base de datos de NVE de PBio) se encuentran con una duplicación de esta región, junto con el paciente PML10. Ambos pacientes con PML son mujeres de ascendencia europea y sus enfermedades primarias son el VIH y la MS (véase la Tabla 7).

La Figura 8 representa un ejemplo de un gen afectado por una pérdida de CNV. La eliminación exónica de 1.1 Kb afecta a los genes PIK3CD y PIK3CD-AS1 (el símbolo del gen anterior era C1orf200). Se detectó una eliminación homocigota en el paciente MVGS811-13a y esta pérdida (heterocigota u homocigota) no se encontró en sujetos normales ni en la base de datos pública de CNV de DGV. El paciente con PML es un varón y su enfermedad primaria es el VIH (véase la Tabla 7). Se presume que es de ascendencia EUR (no se disponía de etnias para las muestras de MVGS).

La Figura 9 representa un ejemplo de un gen afectado por una ganancia de CNV intergénica. La ganancia intergénica de 16.7 Kb tiene un punto de rotura a la izquierda que se encuentra 105 Kb secuencia arriba del gen CD180 (variante de transcripción de RefSeq NM_005582). El paciente MVGS995-4a tiene una duplicación homocigota (relación log2 comparable a triplicaciones) basado en la observación de que los sujetos de 1000 genomas se informa que tienen esta ganancia (véase la variante esv3605336 de DGV del ensamblaje hg19, que informa que 2 de 2504 sujetos tienen esta ganancia; no se encontraron normales en base de datos de NVE de PBio). El paciente con PML es un varón de ascendencia europea y su enfermedad primaria es la MS (véase la Tabla 7).

La Figura 10 representa un ejemplo de un gen afectado por una pérdida de CNV intergénica. La eliminación homocigota intergénica de 7.7 Kb tiene un punto de rotura a la izquierda que se encuentra 3-4 Kb secuencia arriba de las variantes de transcripción VDAC1 (RefSeq: NM_003374, NR_036625, NR_036624). Esta pérdida (heterocigota u homocigota) no se encontró en sujetos normales ni en la base de datos pública de CNV de DGV. El paciente PML30 es un varón de ascendencia europea y su enfermedad primaria es el VIH (véase la Tabla 7).

La Figura 11 representa un ejemplo de un gen afectado por una pérdida intergénica de CNV. La eliminación homocigota intergénica de 6.8 Kb tiene un punto de rotura a la izquierda que está 4 Kb secuencia abajo de la variante de transcripción EGR1 (RefSeq: NM_001964) y 26 Kb secuencia abajo de las variantes de transcripción ETF1 (RefSeq: NM_001256302, NM_004730, NM_001282185, NM_001291975, NM_001291974). Se encontró que esta pérdida era homocigota en 1 sujeto normal y también se informó de la pérdida en la base de datos pública de CNV de DGV (véase la variante esv3606925 de DGV del ensamblaje hg19, que informa de que 33 de 2504 sujetos tienen esta pérdida, se desconoce si los sujetos son homocigotos o heterocigotos). El paciente PML69 es un varón de ascendencia europea y su enfermedad (afección) principal es el trasplante de riñón (véase la Tabla 7, reportado como 'Otros'). El paciente PML69 fue tratado con CTLA4-Ig (belatacept, un bloqueador de la coestimulación de CD28-B7 e inductor de anergia de las células T). La vía CD28 incluye vínculos con el hallazgo genético del paciente (p. ej., eliminación homocigota adyacente al gen EGR1) y varios otros genes que pueden estar relacionados con la inmunodeficiencia (p. ej., CD40LG, ITK, LCK, LRBA, PIK3CD, PIK3R1, PLCG2, WAS y ZAP70) (Dekeyser M et al., Foro Abierto de Enfermedades Infecciosas, 2016 (Anergia de células T refractaria y leucoencefalopatía multifocal progresiva rápidamente mortal tras terapia prolongada con CTLA4).

La Figura 12 representa un ejemplo de un gen afectado por una pérdida de CNV intergénica. La eliminación homocigota intergénica de 5.6 Kb tiene un punto de rotura a la izquierda que se encuentra 20 Kb secuencia arriba de las variantes de transcripción ITSN2 (RefSeq: NM _019595, NM_006277, NM_147152). Se encontraron pérdidas heterocigotas en 50 sujetos normales y también se informó de la pérdida en la base de

datos pública de CNV de DGV (véase la variante esv3590068 de DGV del ensamblaje hg19, que informa de que 222 de 2504 sujetos tienen esta pérdida, se desconocen los sujetos homocigóticos frente a los heterocigóticos). El paciente PML65 es un varón de ascendencia africana y su enfermedad principal es el VIH (véase la Tabla 7).

La Figura 13 representa un ejemplo de interacciones proteicas conocidas y/o predichas utilizando la base de datos String (string-db.org; véase Szklarczyk *et al.*, (2015) y las referencias citadas allí). Una lista no redundante de todos los genes reportados en la Tabla 7 (43 genes, que incluyen aquellos cuya expresión fue inferida para ser impactada por una CNV intergénica cercana) como mejores soluciones/explicaciones para 61 de 71 casos de PML (11 casos de PML son reportados como 'no resueltos', incluyendo, 1 caso para el cual sólo se obtuvieron datos de CGH) fue evaluada usando la base de datos String. La 'puntuación de interacción mínima requerida' se fijó en 'confianza alta (0.7)' y no se añadieron 'interactuadores' adicionales. De los 43 genes de entrada, se encontró que 21 tenían interacciones de alta confianza, tal como se muestra en la figura, junto con la anotación del número de casos de PML que tenían cada uno de estos genes como solución/explicación (p. ej., se encontró que 3 casos de PML en la Tabla 7 tenían una solución PLCG2).

Ejemplo 12 - Análisis de la carga génica

El análisis de la carga génica se realizó de la siguiente manera. Utilizando una serie de scripts propios y datos descargados del ExAC (exac.broadinstitute.org), se realizó un recuento de todas las variantes presentes en cada uno de los 419 genes enumerados en la Tabla 6. Cada variante se clasificó en función de si se consideraba perjudicial (sobre la base de al menos uno de los algoritmos de predicción SIFT, PolyPhen2 o MutationTaster) o no perjudicial, heterocigota u homocigota. Esto se realizó en paralelo para las variantes de PML y las encontradas en el ExAC. No se utilizaron los datos del ExAC cuya calidad/cobertura era < 80 % de lo esperado y, por lo tanto, no se pudo realizar el análisis de la carga génica.

A continuación, se realizó una comparación específica de cada etnia (EUR o AFR solamente, había muy pocos casos LAT para este tipo de análisis) para cada una de las 4 categorías:

- Homocigoto perjudicial
- Homocigoto no perjudicial
- Heterocigoto perjudicial
- Heterocigoto no perjudicial

Para las 4 categorías, variantes con cortes de frecuencia alélica menor (MAF) de 0.01, 0.02. 0.03. 0.04, 0.05 y 0.1.

Para cada comparación, razón de probabilidades (OR) y prueba exacta de Fisher (FET) fueron calculados para la comparación de números de casos de PML con al menos una variante del tipo en consideración y aquellos en el ExAC. La corrección por pruebas múltiples se realizó multiplicando la FET por el número de genes considerados (419). Sólo los genes para los que la FET corregida fue < 0.05 se incluyeron en la Tabla 13, que contiene datos sobre los valores promedio para un gen dado en todos los MAF que pasaron la corrección de FET. En la práctica, sólo la categoría de heterocigotos dañinos produjo genes significativos.

Ejemplo 13 - Análisis de carga de variantes

Para cada variante identificada en al menos un caso de PML, se realizó un recuento para obtener la frecuencia de la misma variante en el conjunto de la cohorte. Estos datos agregados se compararon con los recuentos de la misma variante notificados en el ExAC. Los datos del ExAC se filtraron por calidad/cobertura y el análisis de la carga de variantes no se realizó si la cobertura del ExAC era < 80 % de la esperada.

El análisis de la carga de variantes se realizó por separado para las cohortes EUR (n = 44 casos) y AFR (n = 21 casos) (la cohorte LAT era demasiado pequeña) y se calcularon los valores de OR y FET. A partir de este análisis, sólo se consideraron las variantes con OR>1 (es decir, potencialmente indicativas de un mayor riesgo de PML) para ambas etnias (AFR y EUR) y para las que la frecuencia del ExAC de la variante era <5 %. Además, sólo se consideraron de primer nivel aquellas variantes cuya frecuencia en la cohorte específica de la etnia era >10 % (5 o más casos EUR, 3 o más casos AFR) (Tabla 14), aunque otras variantes se han tabulado en la Tabla 15.

Ejemplo 14 - Pruebas ejemplares de predicción del riesgo de PML

La Tabla 16 proporciona marcadores ejemplares para crear una prueba de predicción de riesgo de PML de bajo coste y simple (SNV específicas de genotipo). Se podrían idear otros ejemplos similares a partir de otras SNV reportadas en las Tablas 14 y 15. Se podrían utilizar diferentes combinaciones de SNV de las Tablas 14, 15 en pruebas de complejidad variable, para desarrollar una prueba que produjera mayores rendimientos de diagnóstico que el ejemplo superior de la Tabla 16 (es decir, 40 %).

La Tabla 17 proporciona genes ejemplares que podrían incluirse en una prueba de secuenciación de panel de genes para la predicción del riesgo de PML. Otros ejemplos podrían elaborarse de forma similar a partir de los genes descritos en la Tabla 13, o de otras tablas divulgadas en el presente documento.

La Tabla 9 contiene variantes de 'ejemplo' que pueden ser consideradas como causas de 'AD' de inmunodeficiencia (es decir, la presencia de sólo 1 de las 2 SNV het reportadas en un paciente dado puede estar causando inmunodeficiencia), lo que puede aumentar el riesgo de PML. Por ejemplo, esto puede ser un escenario más probable para SNV het que son 'novedosas' en la base de datos del ExAC (es decir, no se encuentran en la población general), y aún más probable si tales SNV novedosas se encuentran en >=2 casos de PML (independientemente del modelo de enfermedad invocado). Ejemplos de ello son los 3 genes siguientes:

- AK2, 2 casos (Tabla 9)

chr1:33476435, C>A, novedosos en el ExAC
PML20 y PML33, AFR y EUR, ambos con VIH

EPG5, 2 casos (Tabla 9)

cr18: 43445601, T>G, novedosos en el ExAC
PML25 y PML27, ambos EUR, ambos con VIH

- TNFRSF11A, 9 casos (Tabla 7)

cr18: 60052034, A>C, novedosos en el ExAC
véase la Tabla 7 para los ID de los casos, 2 AFR y 7 EUR, todos con VIH

Puede ser apreciado por aquellos expertos en la técnica que los genes de inmunodeficiencia actualmente conocidos por causar la enfermedad AR pueden potencialmente causar la enfermedad AD. Numerosos ejemplos han sido reportados en la bibliografía, incluyendo, varios de los genes listados en la Tabla 6 (p. ej., Modelo de enfermedad se indica como AD_AR para 32 genes, tal como ADAR y TICAM1).

Ejemplo 15 - Pruebas ejemplares de predicción del riesgo de PML con un panel de 96 genes

La Tabla 19 contiene un panel ejemplar de 96 genes basado en genes que fueron encontrados con al menos un caso de PML de las Tablas 7 y 8. Las columnas "Genes" y "Soluciones a nivel de caso" mostraban los genes y el número total de casos de PML (con al menos una solución 'a nivel de caso') indicados en las Tablas 7 y 8. Además, los 7 genes principales (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) de la Tabla 14 con SNV basadas en valores de 'PML_ALL_FET' <0.05 (columna O) también se incluyeron en la Tabla 19. Entre estos 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) con SNV se basaron en valores 'PML_ALL_FET' <0.05 (columna O) de la Tabla 15.

Tabla 19: panel ejemplar de 96 genes

Genes	Soluciones a nivel de caso
AP3B1	5
APOL1	1
ASH1L	1
ATM	1
ATR	3
BLM	1
CARD11	3
CDKN1B	1
CHD7	4
CLCN7	1
DCLRE1C	3
DDX58	1
DOCK8	8
EGR1	1

Genes	Soluciones a nivel de caso
EPGS	3
ETF1	1
FPR2	1
GATA2	2
GFI1	4
HIVEP1	1
HIVEP2	2
HTR2A	1
IDO2	1
IFIH1	3
IFNGR2	1
IFNLR1	1
IGLL1	0
IKBKB	1
IL17F	1
IL1B	1
IL21R	1
IRAK4	2
ITSN2	2
JUN	2
KAT6B	1
KCTD7	1
LIG4	1
LRBA	1
MALL	1
MAPK3	2
MAVS	0
MCEE	1
MKL1	1
MYD88	1
NBN	1
NFKB1	3
NOD2	6
NRIP1	1
PIAS1	1
PIAS2	1
PIK3CD	4
PIK3CD-AS1	1
PIK3R1	1
PKHD1	3

Genes	Soluciones a nivel de caso
PLCG2	5
PNPT1	1
POLA1	1
POLE	1
PRF1	1
PRKCB	1
PRKCD	1
PRKCH	1
PRKDC	4
PSTPIP1	1
PTEN	1
PTPRC	2
RABGEF1	1
RAD51	1
RAG1	4
RAG2	2
RIPK1	1
RIPK3	2
RNF168	2
RTEL1	2
SHARPIN	1
SKIV2L	1
SMAD4	1
STIM1	2
STIM2	1
STXBP2	3
TAP2	1
TBK1	2
TCIRG1	1
TICAM1	2
TLR3	2
TLR4	1
TNFRSF11A	10
TNFRSF13B	1
TNFRSF8	1
TP53	1
TRAF3	1
TRAFD1	1
TRPM2	1
VPS45	1

Genes	Soluciones a nivel de caso
WEE1	2
ZAP70	3
TOTAL (96 genes)	172
Casos no redundantes	67
Rendimiento diagnóstico para la cohorte de PML (n = 70)	95.7 %

El número no redundante de casos de PML y el rendimiento diagnóstico están listados en las 2 últimas filas de la Tabla 19. En concreto, una prueba que incluyera los 96 genes tuvo un rendimiento diagnóstico del 95.7 % basado en los hallazgos genéticos de los 70 casos de PML utilizados en el presente estudio.

Ejemplo 16 - Pruebas ejemplares de predicción del riesgo de PML con un panel de 39 genes

La Tabla 20 contiene un panel ejemplar de 39 genes basado en genes que fueron encontrados con múltiples casos de PML de las Tablas 7 y 8. Las columnas "Genes" y "Soluciones a nivel de caso" mostraban los genes y el número total de casos de PML (con al menos dos soluciones 'a nivel de caso') reportados en las Tablas 7 y 8. Además, los 7 genes principales (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) de la Tabla 14 con SNV basadas en valores 'PML_ALL_FET' <0.05 (columna O) también se incluyeron en la Tabla 20. Entre estos 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) con SNV se basaron en valores 'PML_ALL_FET' <0.05 (columna O) de la Tabla 15.

Tabla 20: panel ejemplar de 39 genes

Genes	Soluciones a nivel de caso
AP3B1	5
ATR	3
CARD11	3
CHD7	4
DCLRE1C	3
DOCK8	8
EPGS	3
GATA2	2
GFI1	4
HIVEP2	2
IFIH1	3
IGLL1	0
IRAK4	2
ITSN2	2
JUN	2
MAPK3	2
MAVS	0
NFKB1	3
NOD2	6
PIK3CD	4
PKHD1	3
PLCG2	5
PRKDC	4
PTPRC	2

Genes	Soluciones a nivel de caso
RAG1	4
RAG2	2
RIPK3	2
RNF168	2
RTEL1	2
SHARPIN	1
STIM1	2
STXBP2	3
TBK1	2
TCIRG1	1
TICAM1	2
TLR3	2
TNFRSF11A	10
WEE1	2
ZAP70	3
TOTAL (39 genes)	115
Casos no redundantes	57
Rendimiento diagnóstico para la cohorte de PML (n = 70)	81.4 %

El número no redundante de casos de PML y el rendimiento diagnóstico están listados en las 2 últimas filas de la Tabla 20. En concreto, una prueba que incluyera los 39 genes tuvo un rendimiento diagnóstico del 81.4 % basado en los hallazgos genéticos de los 70 casos de PML utilizados en el presente estudio.

Ejemplo 17 - Pruebas ejemplares de predicción del riesgo de PML con un panel de 23 genes

La Tabla 21 contiene un panel ejemplar de 23 genes basado en genes que fueron encontrados con múltiples casos de PML de las Tablas 7 y 8. Las columnas "Genes" y "Soluciones a nivel de caso" mostraban los genes y el número total de casos de PML (con al menos tres soluciones 'a nivel de caso') de los que se informaba en las Tablas 7 y 8. Además, los 7 genes principales (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) de la Tabla 14 con SNV basadas en valores 'PML_ALL_FET' <0.05 (columna O) también se incluyeron en la Tabla 21. Entre estos 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) con SNV se basaron en valores 'PML_ALL_FET' <0.05 (columna O) de la Tabla 15.

Tabla 21: panel ejemplar de 23 genes

Genes	Soluciones a nivel de caso
AP3B1	5
ATR	3
CARD11	3
CHD7	4
DCLRE1C	3
DOCK8	8
EPGS	3
GFI1	4
IFIH1	3
IGLL1	0
MAVS	0
NFKB1	3

Genes	Soluciones a nivel de caso
NOD2	6
PIK3CD	4
PKHD1	3
PLCG2	5
PRKDC	4
RAG1	4
SHARPIN	1
STXBP2	3
TCIRG1	1
TNFRSF11A	10
ZAP70	3
TOTAL (23 genes)	83
Casos no redundantes	50
Rendimiento diagnóstico para la cohorte de PML (n = 70)	71.4 %

El número no redundante de casos de PML y el rendimiento diagnóstico están listados en las 2 últimas filas de la Tabla 21. En concreto, una prueba que incluyera los 23 genes tuvo un rendimiento diagnóstico del 71.4 % basado en los hallazgos genéticos de los 70 casos de PML utilizados en el presente estudio.

Ejemplo 18 - Pruebas ejemplares de predicción del riesgo de PML con un panel de 15 genes

La Tabla 22 contiene un panel ejemplar de 15 genes basado en genes que fueron encontrados con múltiples casos de PML de las Tablas 7 y 8. Las columnas "Genes" y "Soluciones a nivel de caso" mostraban los genes y el número total de casos de PML (con al menos cuatro soluciones 'a nivel de caso') de los que se informaba en las Tablas 7 y 8. Además, los 7 genes principales (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) de la Tabla 14 con SNV basadas en valores 'PML_ALL_FET' <0.05 (columna O) también se incluyeron en la Tabla 22. Entre estos 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) con SNV se basaron en valores 'PML_ALL_FET' <0.05 (columna O) de la Tabla 15.

Tabla 22: panel ejemplar de 15 genes

Genes	Soluciones a nivel de caso
AP3B1	5
CHD7	4
DOCK8	8
GFI1	4
IFIH1	3
IGLL1	0
MAVS	0
NOD2	6
PIK3CD	4
PLCG2	5
PRKDC	4
RAG1	4
SHARPIN	1
TCIRG1	1
TNFRSF11A	10

Genes	Soluciones a nivel de caso
TOTAL (15 genes)	59
Casos no redundantes	39
Rendimiento diagnóstico para la cohorte de PML (n = 70)	55.7 %

El número no redundante de casos de PML y el rendimiento diagnóstico están listados en las 2 últimas filas de la Tabla 22. En concreto, una prueba que incluyera los 15 genes tuvo un rendimiento diagnóstico del 55.7 % basado en los hallazgos genéticos de los 70 casos de PML utilizados en el presente estudio.

Ejemplo 19 - Pruebas ejemplares de predicción del riesgo de PML con un panel de 11 genes

La Tabla 23 contiene un panel ejemplar de 11 genes basado en genes que fueron encontrados con múltiples casos de PML de las Tablas 7 y 8. Las columnas "Genes" y "Soluciones a nivel de caso" mostraban los genes y el número total de casos de PML (con al menos cinco soluciones 'a nivel de caso') reportados en las Tablas 7 y 8. Además, los 7 genes principales (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) de la Tabla 14 con SNV basados en valores 'PML_ALL_FET' <0.05 (columna O) también se incluyeron en la Tabla 23. Entre estos 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) con SNV se basaron en valores 'PML_ALL_FET' <0.05 (columna O) de la Tabla 15.

Tabla 23: panel ejemplar de 11 genes

Genes	Soluciones a nivel de caso
AP3B1	5
CHD7	4
DOCK8	8
IFIH1	3
IGLL1	0
MAVS	0
NOD2	6
PLCG2	5
SHARPIN	1
TCIRG1	1
TNFRSF11A	10
TOTAL (11 genes)	43
Casos no redundantes	33
Rendimiento diagnóstico para la cohorte de PML (n = 70)	47.1 %

El número no redundante de casos de PML y el rendimiento diagnóstico están listados en las 2 últimas filas de la Tabla 23. En concreto, una prueba que incluyera los 11 genes tuvo un rendimiento diagnóstico del 47.1 % basado en los hallazgos genéticos de los 70 casos de PML utilizados en el presente estudio.

Ejemplo 20 - Pruebas ejemplares de predicción del riesgo de PML con un panel de 10 genes

La Tabla 24 contiene un panel ejemplar de 10 SNV basado en los 7 SNV principales de la Tabla 14 y 3 SNV de la Tabla 15 (basado en genes solapados entre 14 y 15: IGLL1, MAVS, SHARPIN). En concreto, utilizando los 10 SNV principales (7 de la Tabla 14, junto con 3 de la Tabla 15, que residen en genes ya seleccionados de la Tabla 14), se realizó un recuento aditivo (columna "Total aditivo de casos (no redundante)") para determinar cuántos casos de PML tenían al menos una de las variantes cuando éstas se consideraban en orden (p. ej., columna "Orden (FET)": '1', primero, seguido de '1' + '2', seguido de '1' + '2' + '3', etc.). Dado que algunos individuos albergan más de una variante, el recuento aditivo no es igual a la suma simple de los números de casos de PML para cada variante (columna "Total de casos por SNV"). Todas las coordenadas del genoma se basan en la construcción hg19.

Se realizó un conteo aditivo para los sujetos del ExAC (columna "sujetos del ExAC totales aditivos (redundante)"), de la siguiente manera: i) Se calculó el tamaño promedio de la cohorte para el ExAC para todas las variantes; ii) Cada conteo total de sujetos (todas las etnias) fue normalizado a este tamaño promedio de cohorte. El recuento aditivo del ExAC representa una adición simple: marcado como "redundante" en la

columna " sujetos del ExAC totales aditivos (redundante)", porque no se dispone de información relativa a la posible presencia de múltiples variantes en el mismo individuo; iii) se calcularon las razones de probabilidades (OR) y los valores de la prueba exacta de Fisher (FET) (columnas "PML ALL OR aditiva" y "PML ALL FET aditiva").

Tabla 24: panel ejemplar de 10 genes

Orden (FET)[1]	Fuente de la tabla	Gen	Variante (hg19)	Genotipo	Total de casos por SNV	Caso total aditivo (no redundante)[2]	Rendimiento de diagnóstico (no redundante)	Total aditivo de sujetos ExAC (redundante)[3]	PML ALL OR aditivo	PML ALL FET aditivo
1	14	PLCG2	chr16: 81942175, A>G	het	7	7	10 %	730	6.50	2.00E -04
2	14	IFIH1	cr2: 163136505, C>G	het	7	13	19 %	1,473	6.49	6.37E -07
3	14	TCIRG1	chr11: 67818269, G>A	het	4	16	23 %	1,830	6.73	2.94E -08
4	14	IGLL1	chr22: 23917192, G>T	het	8	22	31 %	3,368	5.42	9.41E -09
5	14	MAVS	cr20: 3846397, C>T	hom	8	26	37 %	4,947	4.60	2.13E -08
6	14	SHARPIN	chr8: 145154222, G>A	het	12	33	47 %	8,064	3.91	5.10E -08
7	14	CHD7	chr8: 61654298, T>A	het	5	36	51 %	9,292	3.89	3.26E -08
8	15	SHARPIN	chr8: 145154824, A>C	het	3	37	53 %	9,294	4.12	8.10E -09
9	15	IGLL1	chr22: 23915745, G>A	het	3	38	54 %	9,394	4.30	2.59E -09
10	15	MAVS	cr20: 3843027, C>A	hom	6	38	54 %	10,393	3.77	5.26E -08

[1] Orden de SNV basado en el valor FET más bajo reportado en las Tablas 14 y 15 para las etnias combinadas
[2] Total de casos de PML = 70
[3] Sujetos totales ExAC = 43,419 (promedio de las 10 SNV)

Puede ser apreciado por aquellos expertos en la técnica que los paneles de genes arriba mencionados fueron seleccionados basados en los presentes hallazgos genéticos en 70 casos de PML. Además, puede añadirse al panel de genes un gen no seleccionado actualmente para ninguno de estos paneles de genes ejemplares. Por ejemplo, los genes en los que sólo se encontró 1 caso de PML con variantes que cumplían los criterios pueden añadirse al panel de genes si la validación genética en casos adicionales de PML muestra que un gen "n=1 caso" está afectado por más de 1 caso de PML cuando los datos se examinan para un nuevo conjunto de casos de PML. En algunos casos, pueden añadirse genes adicionales (p. ej., genes ligados a PML como DOCK8, BAG3, STAT1) al panel de genes.

Ejemplo 21 - Identificar variaciones genéticas adicionales

Los métodos y protocolos descritos en los ejemplos anteriores pueden utilizarse para identificar cualquier posible variación genética. Se pueden generar datos comparando las variaciones genéticas identificadas en 2 cohortes: 1) cohorte no enferma que incluye 1000 individuos no enfermos (p. ej., individuos sin PML); y 2) cohorte enferma que incluye 100 individuos enfermos (p. ej., individuos con PML). Los individuos de las cohortes pueden estar emparejados por sexo y/o etnia. Las variaciones genéticas presentes en la cohorte no enferma y en la cohorte enferma pueden identificarse utilizando el análisis de CNV (p. ej., descrito en el Ejemplo 2) o la secuenciación del exoma completo (p. ej., descrita en el Ejemplo 3). Se identifican dos nuevas variaciones genéticas, CNV_1 (localizada en el gen #1) y CNV_2 (localizada en el gen #2), p. ej., comparando los datos de la secuencia con una secuencia de referencia (p. ej., UCSC hg19). Los datos de una base de datos de CNV creada utilizando datos de CNV de todo el genoma en sujetos sanos (o individuos sin PML) como el motor de variación normal (NVE) descrito en el presente documento se pueden utilizar para determinar si una CNV encontrada en una cohorte de PML se produce con mayor frecuencia o no en comparación con el NVE. Del mismo modo, las SNV identificadas en una cohorte de PML pueden ser interpretadas por la relevancia potencial para PML utilizando los recursos disponibles públicamente del Exome Aggregation Consortium (ExAC) o la base de datos de agregación del genoma (gnomAD) [Lek M et al., Nature. Agosto de 2016 18; 536(7616):285-91]; es decir, obtener datos de frecuencia (p. ej., frecuencia específica de una etnia) para las variantes consideradas. En otros casos, se pueden crear bases de datos de CNV para la evaluación de CNV para una variedad de etnias (p. ej., sujetos de ascendencia africana y latina) para determinar la relevancia de una CNV en una cohorte de PML en comparación con una base de datos de CNV emparejada étnicamente en individuos sin PML y/o utilizando datos específicos étnicos de bases de datos de CNV disponibles públicamente como la Base de Datos de Variantes Genómicas [MacDonald J et al., Nucleic Acids Res. Enero de 2014; 42 (número de base de datos):D986-92]. En un ejemplo, 20 de los 100 individuos enfermos tienen CNV_1, y sólo 10 de los 1000 individuos no enfermos tienen CNV_1. En otro ejemplo, 10 de los 100 individuos enfermos tienen CNV_2, y sólo 5 de los 1000 individuos no enfermos tienen CNV_2. El valor p puede calcularse utilizando pruebas estándar, tal como la prueba exacta de Fisher (FET), y los datos pueden seleccionarse utilizando un determinado valor de significancia. Por ejemplo, se incluyen las variaciones genéticas con un valor p inferior a 0.05. Además, la frecuencia y la razón de probabilidades de las dos variaciones genéticas pueden calcularse y resumirse en una tabla ejemplar:

Variación genética	Genes	Frecuencia en la cohorte enferma	Frecuencia en la cohorte no enferma	OR
CNV_1	Gen #1	20/100 = 20 %	10/1000 = 1 %	(20/80) / (10/990) = 24.75
CNV_2	Gen #2	10/100 = 10 %	5/1000 = 0.5 %	(10/90) / (5/995) = 22.11

Un sujeto sin CNV_1, CNV_2, o ambos, puede tener un riesgo disminuido de PML debido a una infección del cerebro por el virus John Cunningham (JCV), y por lo tanto se le puede administrar un medicamento inmunosupresor, tal como natalizumab.

Otras variaciones genéticas como SNV pueden ser identificadas usando el mismo método descrito anteriormente.

Aunque se han mostrado y descrito en el presente documento realizaciones preferidas de la presente invención, será obvio para los expertos en la materia que dichas realizaciones se proporcionan únicamente a modo de ejemplo. La invención puede modificarse dentro del ámbito de las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un medicamento inmunosupresor para su uso en el tratamiento de una afección en un sujeto que necesita terapia con medicamentos inmunosupresores,

en el que el sujeto tiene un menor riesgo de leucoencefalopatía multifocal progresiva (PML) debida a una infección del cerebro por el virus John Cunningham (JCV),

en el que la disminución del riesgo del sujeto se debe a la ausencia de una variación genómica seleccionada del grupo que consiste en: chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A, y
en la que la posición cromosómica de la variación genómica se define con respecto a UCSC hg19.

2. Un medicamento inmunosupresor para uso de acuerdo con la reivindicación 1, en el que la afección es un cáncer, un trasplante de órganos o una enfermedad autoinmune;
opcionalmente, en el que el sujeto es un sujeto humano.

3. Un medicamento inmunosupresor para uso de acuerdo con la reivindicación 2, en el que la afección es una enfermedad autoinmune,

que se selecciona del grupo que consiste en enfermedad de Addison, encefalitis por receptor anti-NMDA, síndrome anti sintetasa, anemia aplásica, anemias autoinmunes, anemia hemolítica autoinmune, pancreatitis autoinmune, enfermedad de Behcet, trastornos ampollosos de la piel, enfermedad celíaca - esprúe (enteropatía sensible al gluten), síndrome de fatiga crónica, polineuropatía desmielinizante inflamatoria crónica, leucemia linfocítica crónica, enfermedad de Crohn, dermatomiositis, enfermedad de Devic, eritroblastopenia, síndrome de Evans, glomeruloesclerosis segmentaria focal, granulomatosis con poliangitis, enfermedad de Graves, oftalmopatía de Graves, síndrome de Guillain-Barré, tiroiditis de Hashimoto, púrpura trombocitopénica idiopática (ITP), nefropatía por IgA, enfermedades autoinmunes mediadas por IgA, enfermedad relacionada con IgG4, enfermedad inflamatoria intestinal, artritis idiopática juvenil, esclerosis múltiple, miastenia gravis, mieloma, linfoma no Hodgkin, síndrome opsoclono mioclono (OMS), penfigoide, penfigo, penfigo vulgar, anemia perniciosa, polimiositis, psoriasis, aplasia pura de eritrocitos, artritis reactiva, artritis reumatoide, sarcoidosis, esclerodermia, síndrome de Sjögren, lupus eritematoso sistémico, púrpura trombocitopénica, púrpura trombocitopénica trombótica, diabetes de tipo I, colitis ulcerosa, vasculitis, vitíligo y combinaciones de los mismos;
opcionalmente, en la que la enfermedad autoinmune es la esclerosis múltiple o la enfermedad de Crohn; u
opcionalmente, en donde la esclerosis múltiple es una forma recidivante de esclerosis múltiple.

4. Un medicamento inmunosupresor para uso de acuerdo con una cualquiera de las reivindicaciones 1 a 3,
en el que el medicamento inmunosupresor comprende un glucocorticoide, un citostático, un anticuerpo, un fármaco que actúa sobre las inmunofilinas, un interferón, un opioide, una proteína de unión al TNF, un micofenolato, un agente biológico pequeño, una molécula pequeña, un compuesto orgánico o cualquier combinación de los mismos,

opcionalmente, el medicamento inmunosupresor comprende abatacept, adalimumab, alefacept, alemtuzumab, anakinra, azatioprina, belimumab, bendamustina, bevacizumab, bortezomib, brentuximab vedotina, capecitabina, carboplatino, cetuximab, clorambucilo, cladribina, ciclofosfamida, ciclosporina, daclizumab, doxorrubicina, eculizumab, efalizumab, etanercept, etopósido, everolimus, fludarabina, fluorouracilo, ácido fumárico, gemcitabina, ibritumomab tiuxetano, imatinib, infliximab, leflunomida, lenalidomida, metilprednisolona, metotrexato, micofenolato mofetilo, natalizumab, oxaliplatino, rituximab, tocilizumab, tofacitinib, ustekinumab, vedolizumab, vincristina, belatacept, quimioterapia citotóxica, corticosteroides, Ig antitimocítica, basiliximab, muromonab-CD3, ácido micofenólico, prednisona/prednisolona, sirolimus, tacrolimus, fumarato de dimetilo, fingolimod, ruxolitinib, interferón beta-la, interferón beta-1b, acetato de glatiramero, peginterferón beta-la, teriflunomida, mitoxantrona, ocrelizumab, asparaginasa, bleomicina, busulfano, carmustina, certolizumab, ibrutinib, idarrubicina, idelalisib, hidrocortisona, ifosfamida, levamisol, mercaptopurina, mizoribina, obinutuzumab, ofatumumab, tegafur/gimeracilo/oteracilo, tiotepa, vinblastina, o cualquier combinación de los mismos,
opcionalmente, en el que el medicamento inmunosupresor comprende interferón beta-la, interferón beta-1b, acetato de glatiramero, peginterferón beta-la, teriflunomida, fingolimod, fumarato de dimetilo, alemtuzumab, mitoxantrona, natalizumab, daclizumab, ocrelizumab, o cualquier combinación de los mismos; u
opcionalmente, en el que el medicamento inmunosupresor comprende natalizumab.

5. Un medicamento inmunosupresor para uso de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que el medicamento inmunosupresor comprende una molécula de anticuerpo o un fragmento de la misma,

opcionalmente en el que:

(i) la molécula de anticuerpo o un fragmento de la misma es una molécula de anticuerpo recombinante humanizada o un fragmento de la misma,
(ii) la molécula de anticuerpo o un fragmento de la misma es una molécula de anticuerpo monoclonal recombinante humanizada IgG4κ o un fragmento de la misma;
(iii) la molécula de anticuerpo o un fragmento de la misma se produce en células de mieloma murino;
(iv) la molécula de anticuerpo o un fragmento de la misma se une a una integrina, en la que la integrina se expresa opcionalmente en la superficie de un leucocito, que opcionalmente no es un neutrófilo;
(v) la molécula de anticuerpo o un fragmento de la misma se une a la integrina α4β1, a la integrina α4β7 o a ambas;
(vi) la molécula de anticuerpo o un fragmento de la misma se une a la subunidad α4 de la integrina α4β1, a la integrina α4β7, o a ambas;
(vii) la molécula de anticuerpo o un fragmento de la misma inhibe la adhesión mediada por α4 de un leucocito a su receptor;
(viii) la molécula de anticuerpo o un fragmento de la misma comprende una secuencia que tiene al menos un 60 %, 70 %, 80 %, 90 %, 95 % o 100 % de identidad de secuencia con la SEQ ID NO. 2178;
(ix) la molécula de anticuerpo o un fragmento de la misma comprende una secuencia que tiene al menos un 60 %, 70 %, 80 %, 90 %, 95 % o 100 % de identidad de secuencia con la SEQ ID NO. 2179; o
(x) la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena pesada de anticuerpo, o un fragmento de la misma de unión a α4, que comprende CDR no humanas en las posiciones 31-35 (CDR1), 50-65 (CDR2) y 95-102 (CDR3) (numeración de Kabat) de un anticuerpo anti-α4 de ratón y que tiene residuos no humanos en las posiciones marco 27-30 (numeración de Kabat), en las que dichas posiciones 27-30 tienen la secuencia de aminoácidos Phe 27, Asn 28, Ile 29 y Lys 30,

opcionalmente en el que:

(a) la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena ligera de anticuerpo, o un fragmento de la misma de unión a α4, que comprende: una CDR1 de cadena ligera (LC) con una secuencia de aminoácidos de la SEQ ID NO: 2180 (KTSQDINKYMA), una CDR2 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2181 (YTSALQP), y una CDR3 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2182 (LQYDNLWT),
(b) la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena ligera de anticuerpo, o un fragmento de la misma de unión a α4, que comprende: una CDR1 de cadena ligera (LC) con una secuencia de aminoácidos de la SEQ ID NO: 2183 (QASQDIIKYLN), una CDR2 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2184 (EASNLQA), y una CDR3 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2185 (QQYQSLPYT),
(c) la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena ligera de anticuerpo, o un fragmento de la misma de unión a α4, que comprende: una CDR1 de cadena ligera (LC) con una secuencia de aminoácidos de la SEQ ID NO: 2186 (KASQSVTNDVA), una CDR2 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2187 (YASNRYT), y una CDR3 de LC con una secuencia de aminoácidos de la SEQ ID NO: 2188 (QQDYSSPYT),
(d) en el que la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena pesada de anticuerpo, o un fragmento de la misma de unión a α4, que comprende: una CDR1 de cadena pesada (HC) con una secuencia de aminoácidos de la SEQ ID NO: 2189 (DTYIH), una CDR2 de HC con una secuencia de aminoácidos de la SEQ ID NO: 2190 (RIDPANGYTKYDPKFQG), y una CDR3 de HC con una secuencia de aminoácidos de la SEQ ID NO: 2191 (EGYYGNYGVYAMDY),
(e) en el que la molécula de anticuerpo o fragmento de la misma comprende al menos una cadena pesada de anticuerpo, o un fragmento de unión a α4 de la misma, que comprende: una CDR1 de cadena pesada (HC) con una secuencia de aminoácidos de la SEQ ID NO: 2192 (DTYMH), una CDR2 de HC con una secuencia de aminoácidos de la SEQ ID NO: 2193 (RIDPASGDTKYDPKFQV), y una CDR3 de HC con una secuencia de aminoácidos de la SEQ ID NO: 2194 (DGMWVSTGYALDF),
(f) en el que la molécula de anticuerpo o fragmento de la misma comprende una cadena pesada humanizada, o un fragmento de la misma de unión a a4, que comprende: una región variable de cadena pesada seleccionada del grupo que consiste en:

SEQ ID NO.: 2195

(MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVRQPP
GRGLEWIGRIDPASGDTKYDPKFQVKATITADTSSNQFSLRLSSVTAADTAVYYCADGMWV
STGYALDFWGQGTTVTVSSGES).

SEQ ID NO.: 2196

(QVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVRQPPGRGLEWIGRIDPASGDTKYDP
KFQVRVTMLVDTSSNQFSLRLSSVTSEDTAVYYCADGMWVSTGYALDFWGQGTTVTVSSG
ES),

SEQ ID NO.: 2197

(MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVKQRP
GRGLEWIGRIDPASGDTKYDPKFQVRVTMLVDTSSNQFSLRLSSVTAADTAVYYCADGMWV
STGYALDFWGQGTTVTVSSGES),

SEQ ID NO.: 2198

(MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTASGFNIKDTYMHWVRQPP
GRGLEWIGRIDPASGDTKYDPKFQVRVTMLVDTSSNQFSLRLSSVTAADTAVYYCADGMWV
STGYALDFWGQGTTVTVSSGES),

y

SEQ ID NO.: 2199

(QVQLVQSGAEVKKPGASVKVSCKASGFNIKDTYIHWVRQAPGQRLEWMGRIDPANGYTKY
DPKFQGRVTITADTSASTAYMELSSLRSEDTAVYYCAREGYYGNYGVYAMDYWGQGTLVT
VSS),

o

(g) en el que la molécula de anticuerpo o fragmento del mismo comprende una cadena ligera humanizada, o un fragmento de la misma de unión a a4, que comprende una región de cadena ligera variable seleccionada del grupo que consiste en:

SEQ ID NO.: 2200

(MGWSCIILFLVATATGVHSDIQLTQSPSSLSASVGDRVTITCKASQSVTNDVAWYQQKPGKA
PKLLIYYASNRYTGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQDYSSPYTFGQGTKVEIKR
K),

SEQ ID NO.: 2201

(MGWSCIILFLVATATGVHSSIVMTQSPSSLSASVGDRVTITCKASQSVTNDVAWYQQKPGKA
PKLLIYYASNRYTGVPDRFSGSGYGTDFTFTISSLQPEDIATYYCQQDYSSPYTFGQGTKVEIK
RK),

SEQ ID NO.: 2202

(MGWSCIILFLVATATGVHSDIQMTQSPSSLSASVGDRVTITCKASQSVTNDVAWYQQKPGK
APKLLIYYASNRYTGVPDRFSGSGYGTDFTFTISSLQPEDIATYYCQQDYSSPYTFGQGTKVEI
KRK),

y

SEQ ID NO.: 2203

(DIQMTQSPSSLSASVGDRVTITCKTSQDINKYMAWYQQTPGKAPRLLIHYTSALQPGIPSRFS
GSGSGRDYTFTISSLQPEDIATYYCLQYDNLWTFGQGTKVEIKRTV),

opcionalmente en el que el medicamento inmunosupresor comprende natalizumab, que se administra mediante infusión intravenosa,
opcionalmente en el que:

(i) se administrarán de 100 mg a 500 mg de natalizumab,
(ii) se administrarán de 100 mg a 500 mg de natalizumab en cuatro semanas, o
(ii) se administrarán de 100 mg a 500 mg de natalizumab mediante infusión intravenosa en cuatro semanas.

6. Un medicamento inmunosupresor para uso de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en el que la ausencia de las variaciones genómicas se establece mediante una prueba genética para detectar las variaciones genómicas en una muestra de ácido polinucleico del sujeto,

opcionalmente, en el que el ácido polinucleico se ensaya mediante una prueba genética seleccionada entre análisis de microarreglos, PCR, secuenciación, hibridación de ácidos nucleicos o cualquier combinación de los mismos,
opcionalmente en el que:

(i) la prueba genética comprende un análisis de microarreglos seleccionado del grupo formado por un análisis de arreglos de hibridación genómica comparativa (CGH) y un análisis de arreglos de SNP, o
(ii) la prueba genética comprende secuenciación, en la que la secuenciación se selecciona del grupo que consiste en secuenciación de firmas masivas en paralelo (MPSS), secuenciación de Polony, pirosecuenciación 454, secuenciación de Illumina, secuenciación de Illumina (Solexa) utilizando la preparación de bibliotecas Genomics 10X, secuenciación SOLiD, secuenciación por semiconductores iónicos, secuenciación por nanobolas de ADN, secuenciación heliscópica de moléculas individuales, secuenciación de moléculas individuales en tiempo real (SMRT), secuenciación RNAP, secuenciación de ADN Nanopore, secuenciación por hibridación y secuenciación microfluídica Sanger,
(iii) la prueba genética comprende el análisis de todo el genoma del sujeto,
(iv) la prueba genética comprende el análisis de un exoma completo del sujeto, o
(v) la prueba genética comprende el análisis de la información sobre ácidos nucleicos ya obtenida para un genoma completo o un exoma completo del sujeto,
opcionalmente, en la que la información sobre el ácido nucleico se obtiene a partir de un análisis *in silico*; o
(vi) la muestra de ácido polinucleico comprende un ácido polinucleico de sangre, saliva, orina, suero, lágrimas, piel, tejido o pelo del sujeto.

7. Un medicamento inmunosupresor para su uso de acuerdo con cualquiera de las reivindicaciones 1 a 6, que debe administrarse con un agente que reduce una carga viral en el sujeto,
opcionalmente en el que:

(i) el medicamento inmunosupresor debe administrarse una vez reducida la carga viral;
(ii) la carga viral es una carga viral del JCV; o
(iii) el agente que reduce la carga viral es un agente dirigido contra el JCV.

8. Un medicamento inmunosupresor para su uso de acuerdo con una cualquiera de las reivindicaciones 1 a 7, en el que se va a analizar la presencia del JCV en una muestra biológica del sujeto,

opcionalmente, en el que el análisis para detectar la presencia del JCV comprende poner en contacto un reactivo de detección del JCV con la muestra biológica, u
opcionalmente, en el que el reactivo de detección del JCV se selecciona del grupo que consiste en un anticuerpo anti-JCV, un cebador específico del JCV y combinaciones de los mismos.

9. Un método de detección de un biomarcador de PML que comprende

(a) detección de las muestras biológicas obtenidas de sujetos con PML para obtener información sobre ácidos nucleicos;
(b) detectar una variación genómica en un gen de la Tabla 19 en la información del ácido nucleico;
(c) determinar una frecuencia de la variación genómica en al menos 100 sujetos humanos sin PML; y
(d) identificar la variación genómica como un biomarcador de PML si la variación genómica está presente a un nivel estadísticamente significativo en la información del ácido nucleico en comparación con la frecuencia de la variación genómica en los al menos 100 sujetos humanos sin PML,

que comprende opcionalmente:

(i) detectar la variación genómica en una muestra de ácido polinucleico de un sujeto sospechoso de padecer PML y utilizar dicha detección para predecir una respuesta adversa del sujeto sospechoso de padecer PML a la terapia, en la que la terapia es una terapia inmunosupresora que comprende uno o más medicamentos inmunosupresores;
(ii) confirmar que cada muestra biológica no es un duplicado de ninguna otra muestra biológica basándose en la información sobre el ácido nucleico;
(iii) determinar un genotipo de sexo para cada muestra biológica basándose en la información del ácido nucleico; o
(iv) confirmar que el genotipo del sexo de cada muestra coincide con el fenotipo del sexo del sujeto de los sujetos con PML.

10. El medicamento inmunosupresor para uso de acuerdo con una cualquiera de las reivindicaciones 1 a 8 o un método de la reivindicación 9, que comprende además analizar al sujeto con una prueba de anticuerpos contra el JCV, una prueba de CD62L o una prueba de bandas oligoclonales de IgM del CSF,

opcionalmente comprende además probar al sujeto con la prueba de anticuerpos contra el JCV, en el que la prueba de anticuerpos contra el JCV no detecta la presencia del JCV,
opcionalmente en el que la prueba de anticuerpos contra el JCV comprende el contacto de un reactivo de detección del JCV con una muestra biológica del sujeto,
opcionalmente, en el que el reactivo de detección del JCV se selecciona del grupo que consiste en un anticuerpo anti-JCV, un cebador específico del JCV y combinaciones de los mismos.

11. Un método para identificar a un sujeto con riesgo de desarrollar leucoencefalopatía multifocal progresiva (PML) que comprende

(a) analizar una muestra de ácido polinucleico del sujeto para detectar una variación genética, y establecer que la variación genética está presente en la muestra de ácido polinucleico;
(b) identificar al sujeto con riesgo de desarrollar PML;

en el que el sujeto está inmunodeprimido, y
en el que el riesgo del sujeto se debe a la presencia de una variación genómica seleccionada del grupo que consiste en chr19:7712287, G>C; chr16:81942175, A>G; chr2:163136505, C>G; chr1:92946625, G>C; chr9:446401, A>G; chr8:61732632, A>G; chr16:81939089, T>C; chr20:62309621, T>C; chr22:36661354, C>T; chr6:52101833, C>T; chr19:7705818, C>T; chr16:81902826, C>T; chr16:81946278, A>G; chr2:98351032, C>G; chr2:98351066, C>T; chr2:98351081, C>T y chr4:103522150, G>A, y
en el que la posición en el cromosoma de la variación genómica se define con respecto a UCSC hg19,
opcionalmente, en el que el sujeto tiene una infección por VIH, o padece una enfermedad seleccionada entre un cáncer, una neoplasia hematológica, una enfermedad autoinmune o una linfocitopenia CD4+ idiopática, o ha recibido un trasplante de órganos.

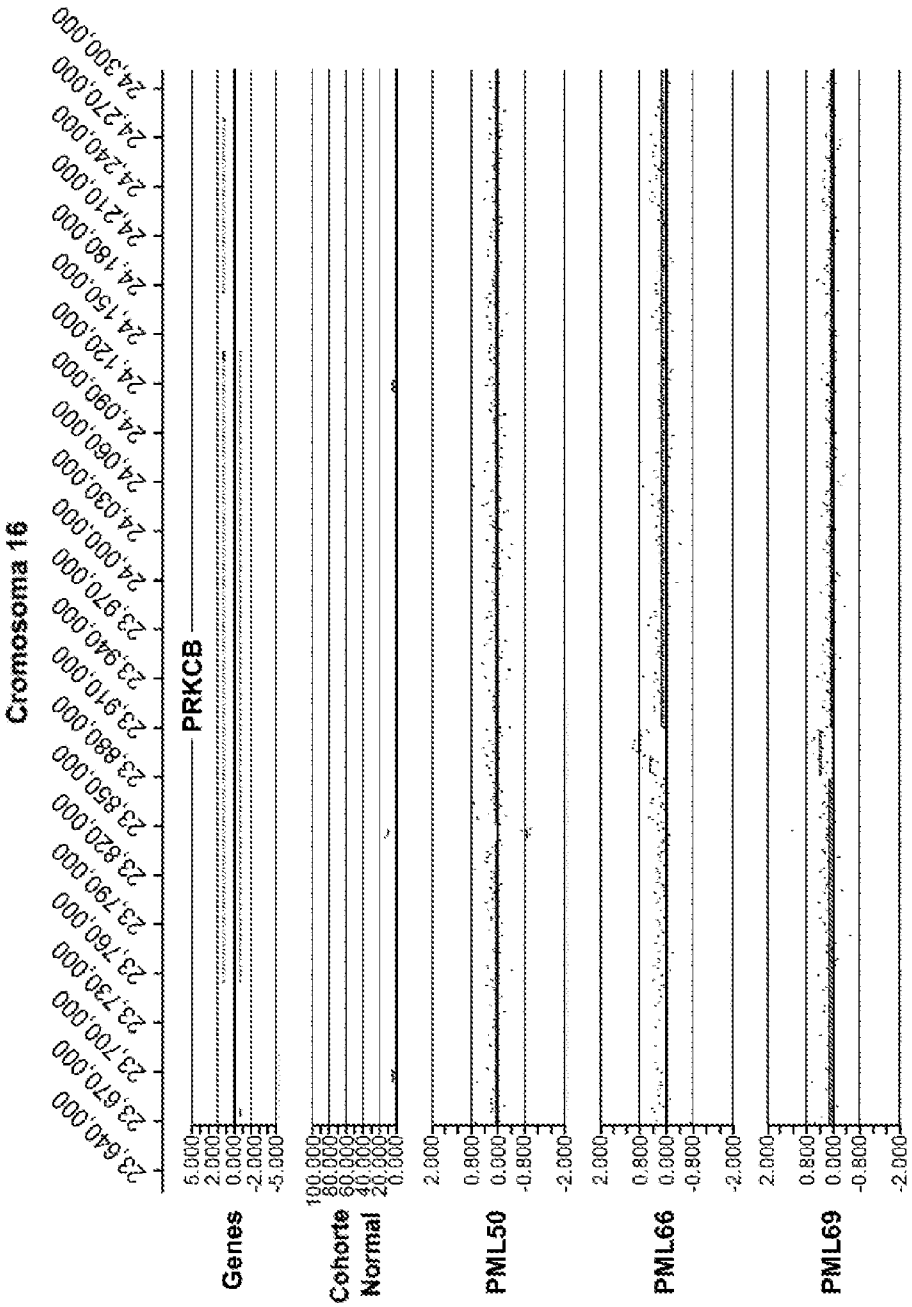

Cromosoma 16
23,640,000
23,670,000
23,700,000
23,730,000
23,760,000
23,790,000
23,820,000
23,850,000
23,880,000
23,910,000
23,940,000
23,970,000
24,000,000
24,030,000
24,060,000
24,090,000
24,120,000
24,150,000
24,180,000
24,210,000
24,240,000
24,270,000
24,300,000
Genes
PRKCB
5.000
2.000
0.000
-2.000
-5.000
Cohorte
Normal
100.000
80.000
60.000
40.000
20.000
0.000
PML50
2.000
0.800
0.000
-0.800
-2.000
PML66
2.000
0.800
0.000
-0.800
-2.000
PML69
2.000
0.800
0.000
-0.800
-2.000

FIG. 1

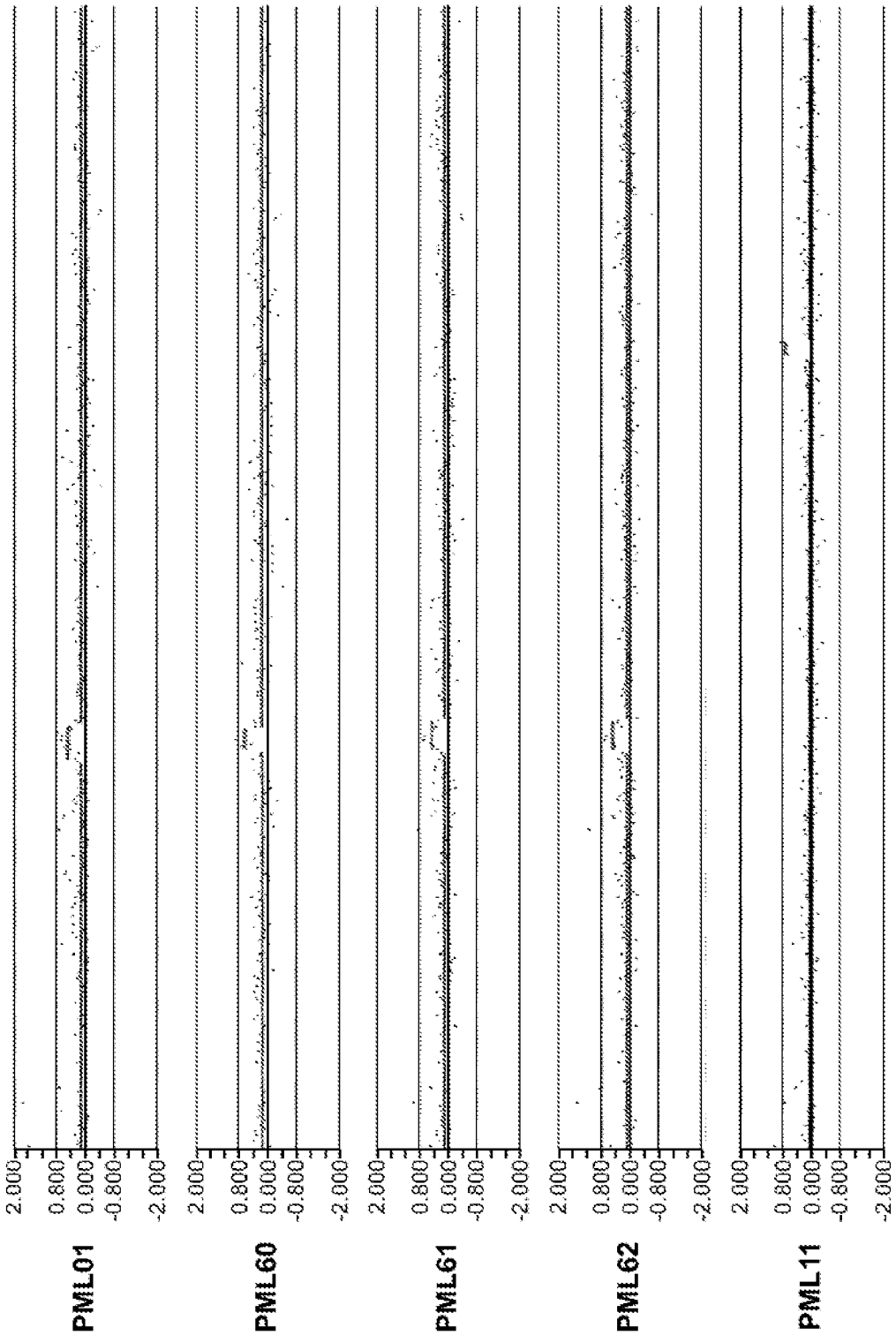


FIG. 1 (Cont.)

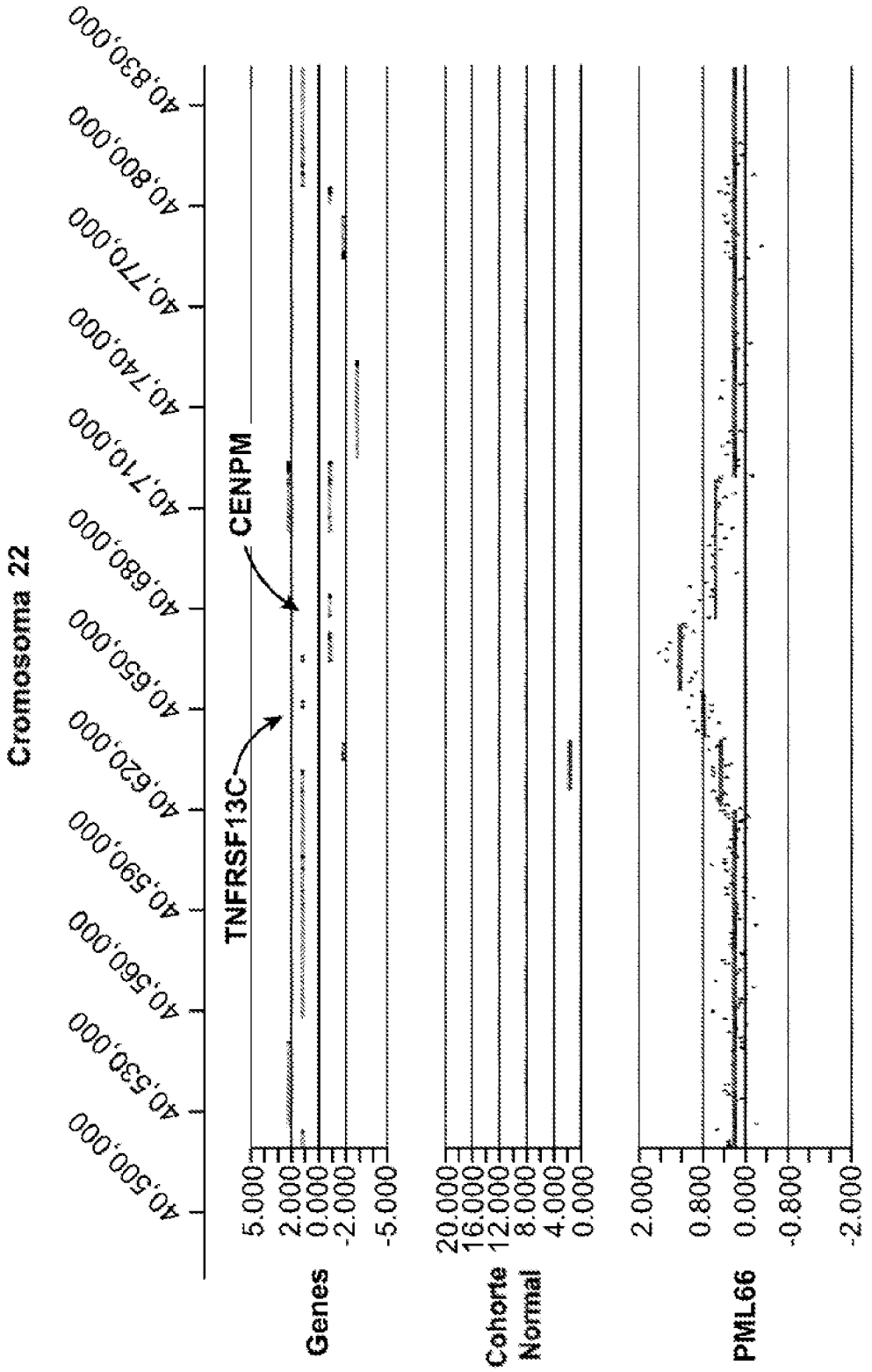

Cromosoma 22
40,500,000
40,530,000
40,560,000
40,590,000
40,620,000
40,650,000
40,680,000
40,710,000
40,740,000
40,770,000
40,800,000
40,830,000
TNFRSF13C
CENPM
Genes
5.000
2.000
0.000
-2.000
-5.000
Cohorte Normal
20.000
16.000
12.000
8.000
4.000
0.000
PML66
2.000
0.800
0.000
-0.800
-2.000

FIG. 2 (Cont. 1)

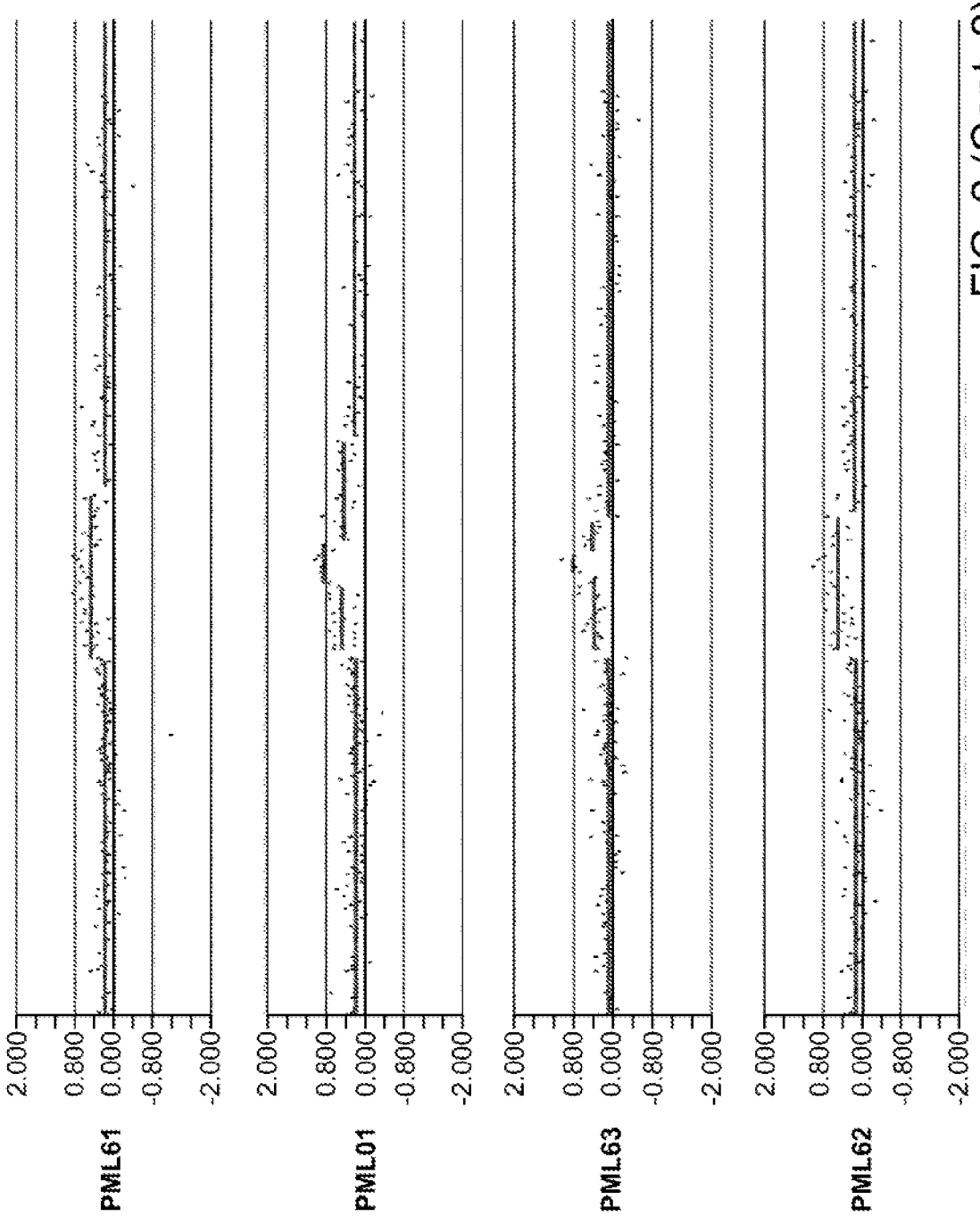


FIG. 2 (Cont. 2)

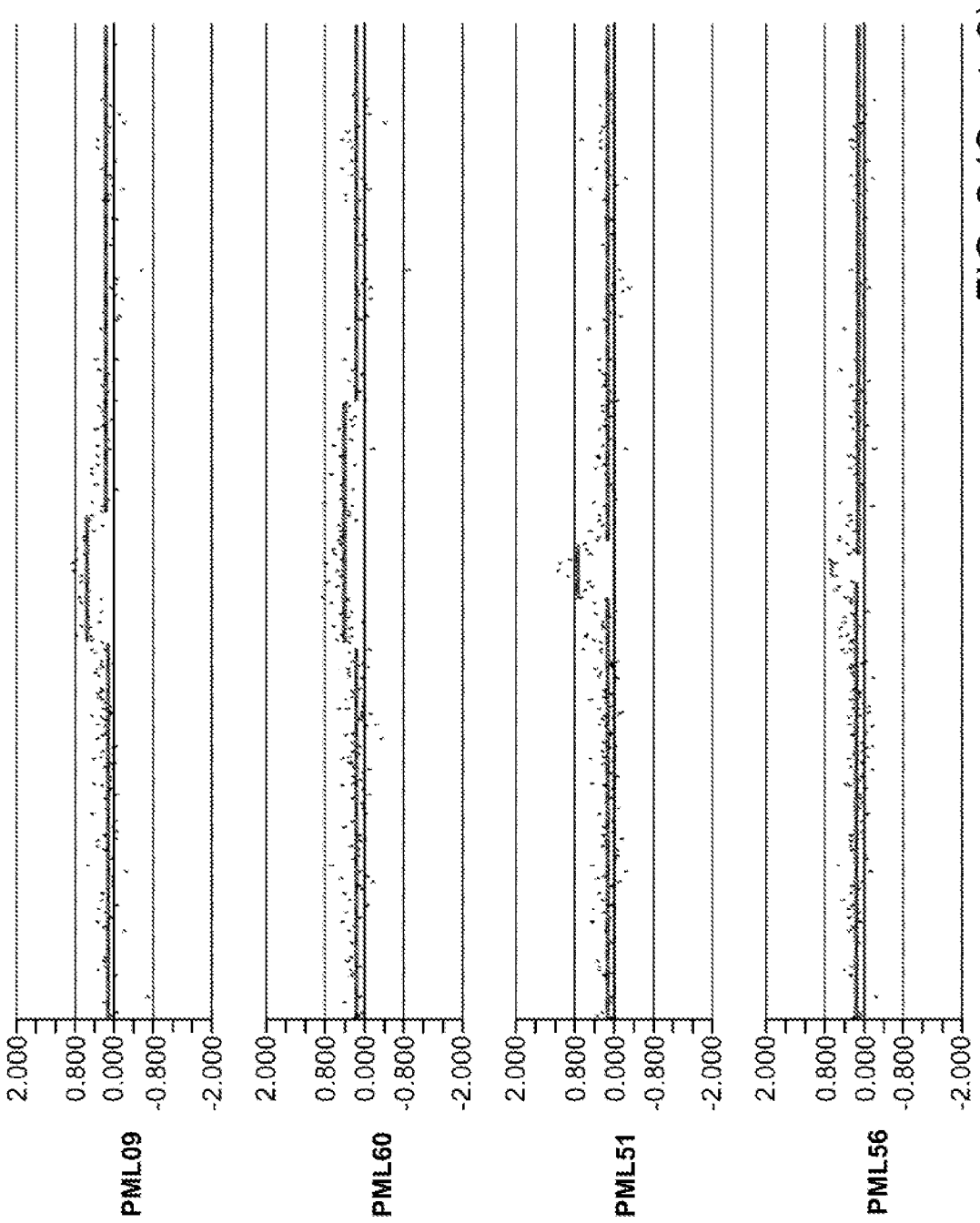


FIG. 2 (Cont. 3)

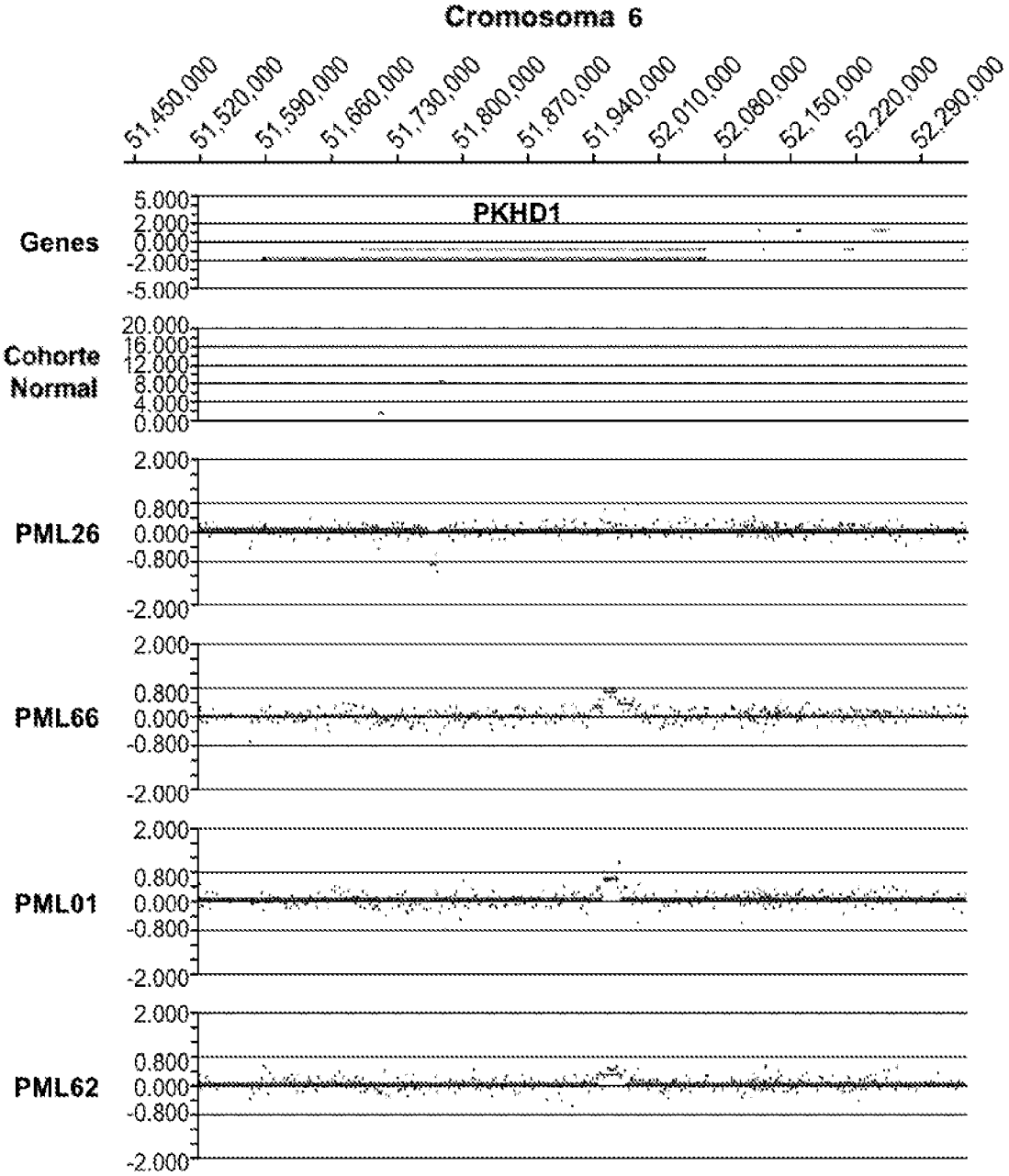

Cromosoma 6
51,450,000
51,520,000
51,590,000
51,660,000
51,730,000
51,800,000
51,870,000
51,940,000
52,010,000
52,080,000
52,150,000
52,220,000
52,290,000
Genes
PKHD1
5.000
2.000
0.000
-2.000
-5.000
Cohorte Normal
20.000
16.000
12.000
8.000
4.000
0.000
PML26
2.000
0.800
0.000
-0.800
-2.000
PML66
2.000
0.800
0.000
-0.800
-2.000
PML01
2.000
0.800
0.000
-0.800
-2.000
PML62
2.000
0.800
0.000
-0.800
-2.000

FIG. 3

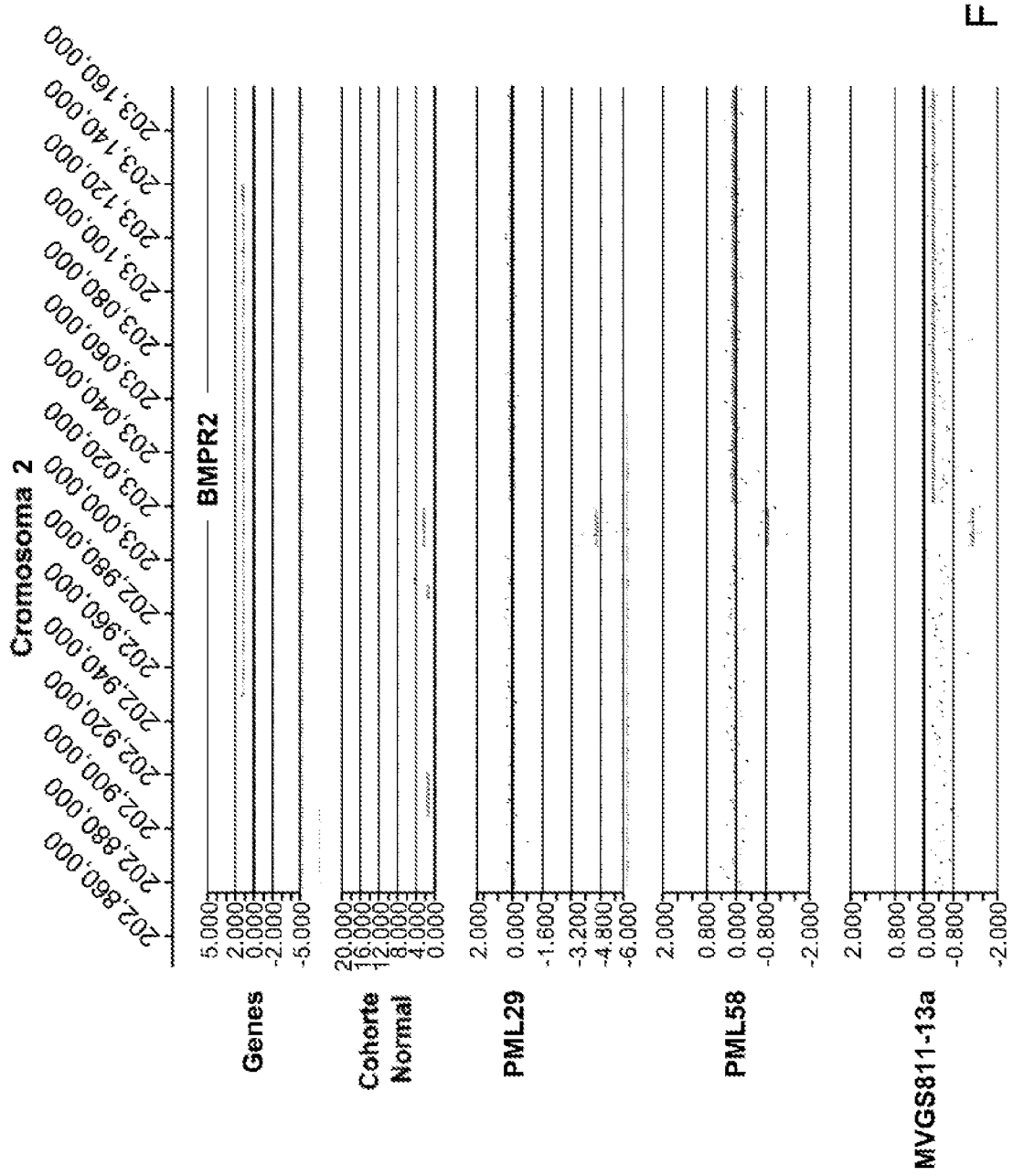

Cromosoma 2
202.860.000
202.880.000
202.900.000
202.920.000
202.940.000
202.960.000
202.980.000
203.000.000
203.020.000
203.040.000
203.060.000
203.080.000
203.100.000
203.120.000
203.140.000
203.160.000
Genes
BMPR2
5.000
2.000
0.000
-2.000
-5.000
Cohorte
Normal
20.000
16.000
12.000
8.000
4.000
0.000
PML29
2.000
0.000
-1.600
-3.200
-4.800
-6.000
PML58
2.000
0.800
0.000
-0.800
-2.000
MVGS811-13a
2.000
0.800
0.000
-0.800
-2.000

FIG. 4

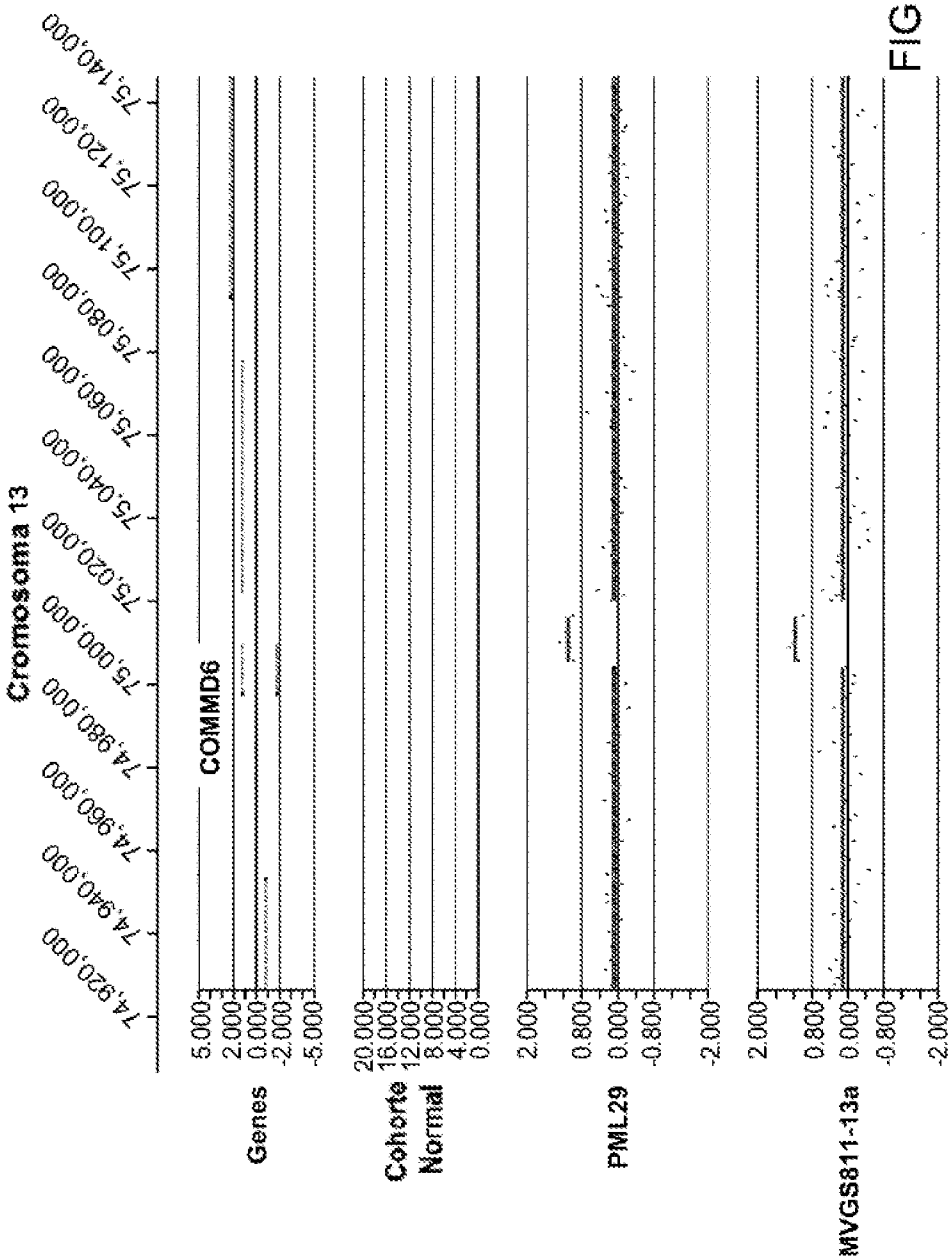

Cromosoma 13
74,920,000
74,940,000
74,960,000
74,980,000
75,000,000
75,020,000
75,040,000
75,060,000
75,080,000
75,100,000
75,120,000
75,140,000
COMMD6
Genes
5.000
2.000
0.000
-2.000
-5.000
Cohorte
Normal
20.000
16.000
12.000
8.000
4.000
0.000
PML29
2.000
0.800
0.000
-0.800
-2.000
MVGS811-13a
2.000
0.800
0.000
-0.800
-2.000

FIG. 5

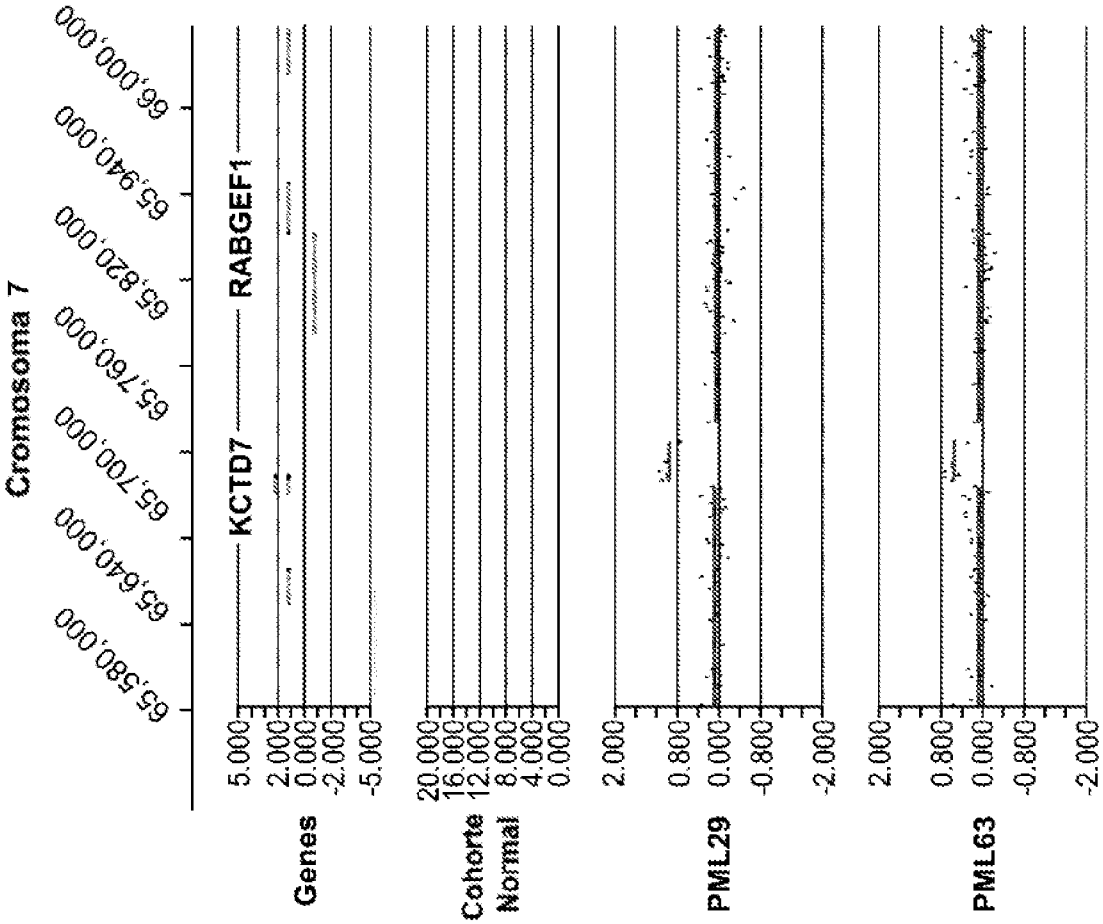

Cromosoma 7
65,580,000
65,640,000
65,700,000
65,760,000
65,820,000
65,940,000
66,000,000
KCTD7
RABGEF1
Genes
5.000
2.000
0.000
-2.000
-5.000
Cohorte
Normal
20.000
16.000
12.000
8.000
4.000
0.000
PML29
2.000
0.800
0.000
-0.800
-2.000
PML63
2.000
0.800
0.000
-0.800
-2.000

FIG. 6

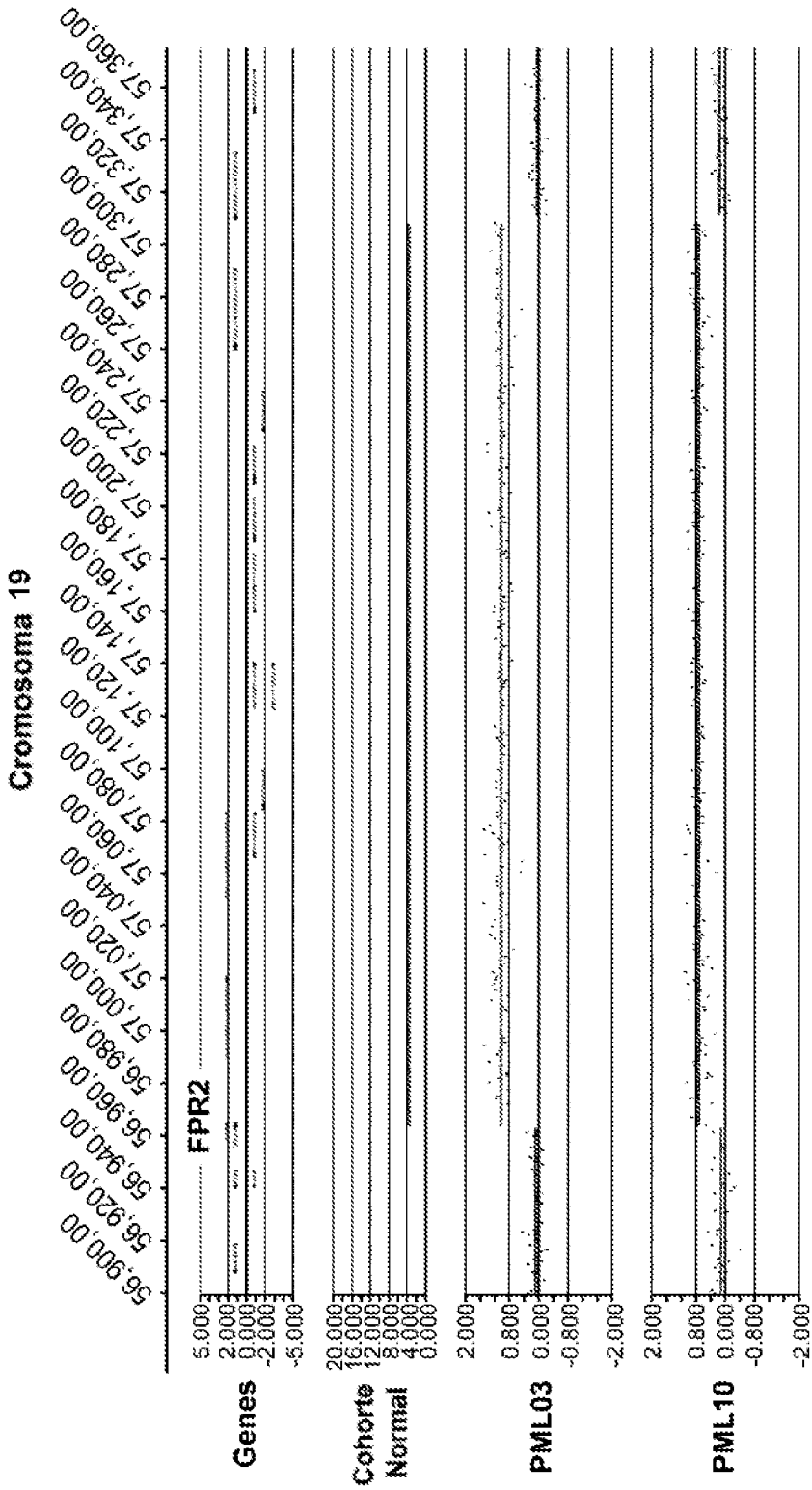

Cromosoma 19
56.900,00
56.920,00
56.940,00
56.960,00
56.980,00
57.000,00
57.020,00
57.040,00
57.060,00
57.080,00
57.100,00
57.120,00
57.140,00
57.160,00
57.180,00
57.200,00
57.220,00
57.240,00
57.260,00
57.280,00
57.300,00
57.320,00
57.340,00
57.360,00
Genes
FPR2
5.000
2.000
0.000
-2.000
-5.000
Cohorte Normal
20.000
16.000
12.000
8.000
4.000
0.000
PML03
2.000
0.800
0.000
-0.800
-2.000
PML10
2.000
0.800
0.000
-0.800
-2.000

FIG. 7

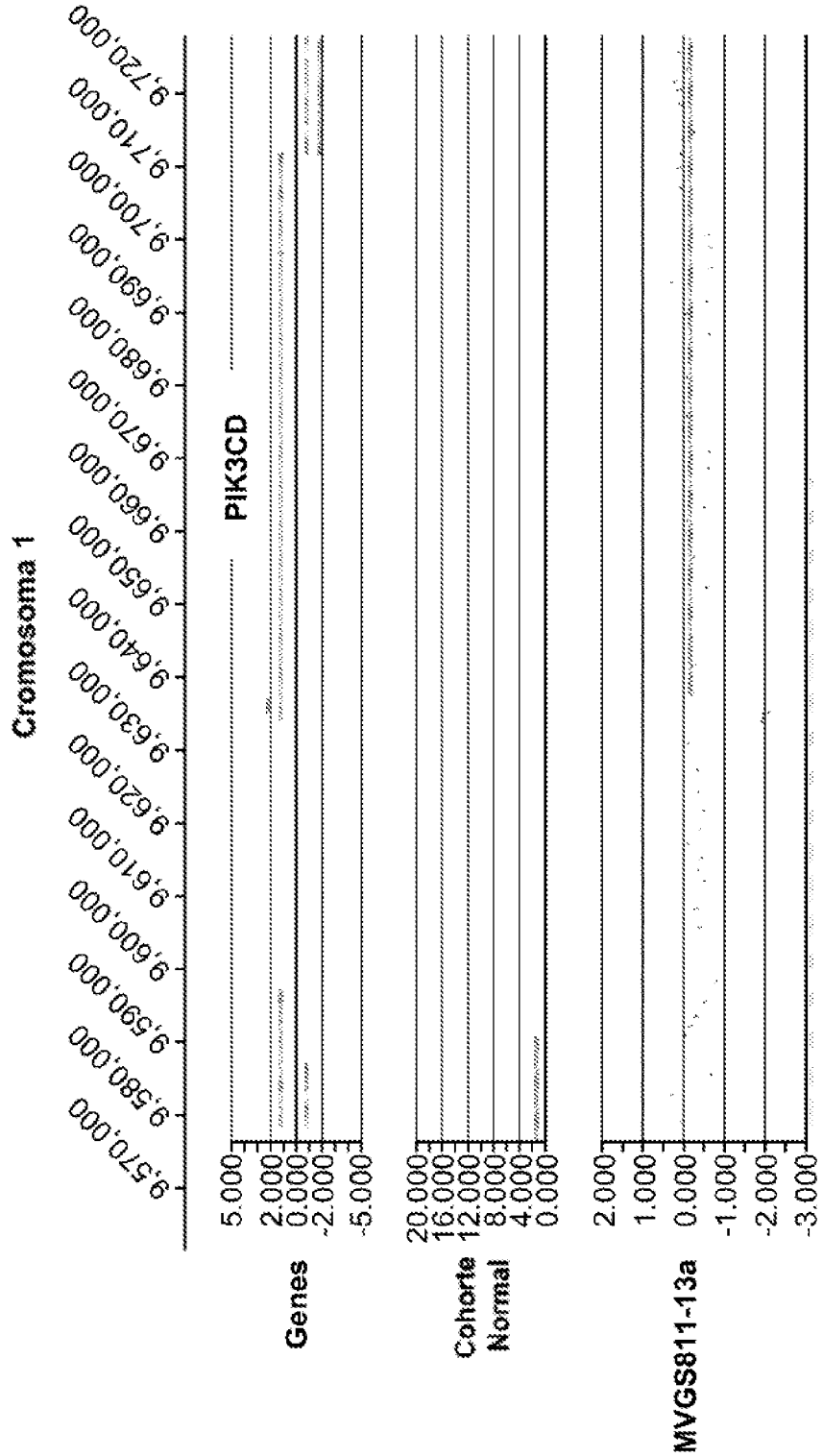

Cromosoma 1
9,570,000
9,580,000
9,590,000
9,600,000
9,610,000
9,620,000
9,630,000
9,640,000
9,650,000
9,660,000
9,670,000
9,680,000
9,690,000
9,700,000
9,710,000
9,720,000
PIK3CD
Genes
5.000
2.000
0.000
-2.000
-5.000
Cohorte Normal
20.000
16.000
12.000
8.000
4.000
0.000
MVGS811-13a
2.000
1.000
0.000
-1.000
-2.000
-3.000

FIG. 8

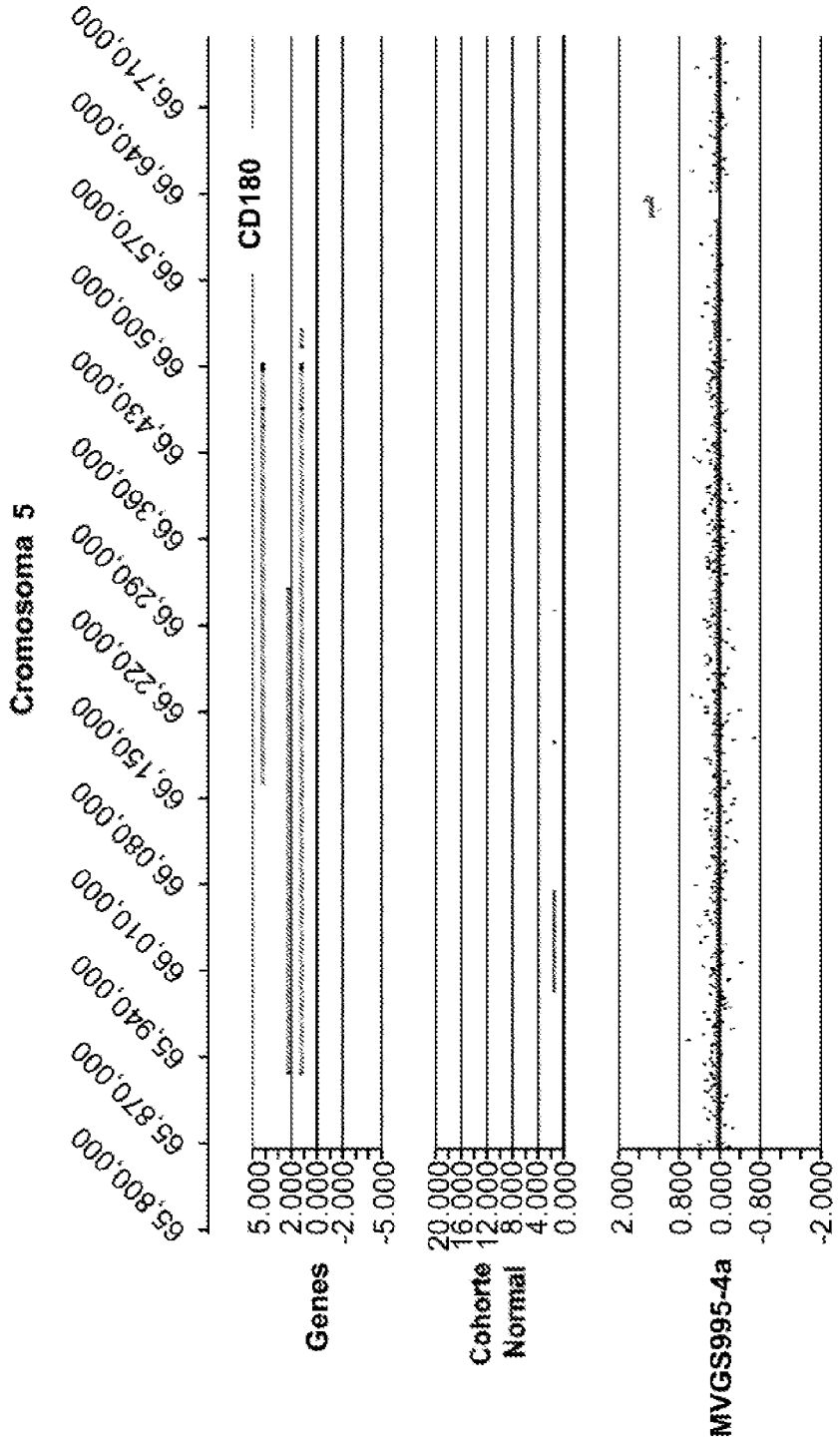

Cromosoma 5
65.800.000
65.870.000
65.940.000
66.010.000
66.080.000
66.150.000
66.220.000
66.290.000
66.360.000
66.430.000
66.500.000
66.570.000
66.640.000
66.710.000
CD180
Genes
5.000
2.000
0.000
-2.000
-5.000
Cohorte
Normal
20.000
16.000
12.000
8.000
4.000
0.000
MVGS995-4a
2.000
0.800
0.000
-0.800
-2.000

FIG. 9

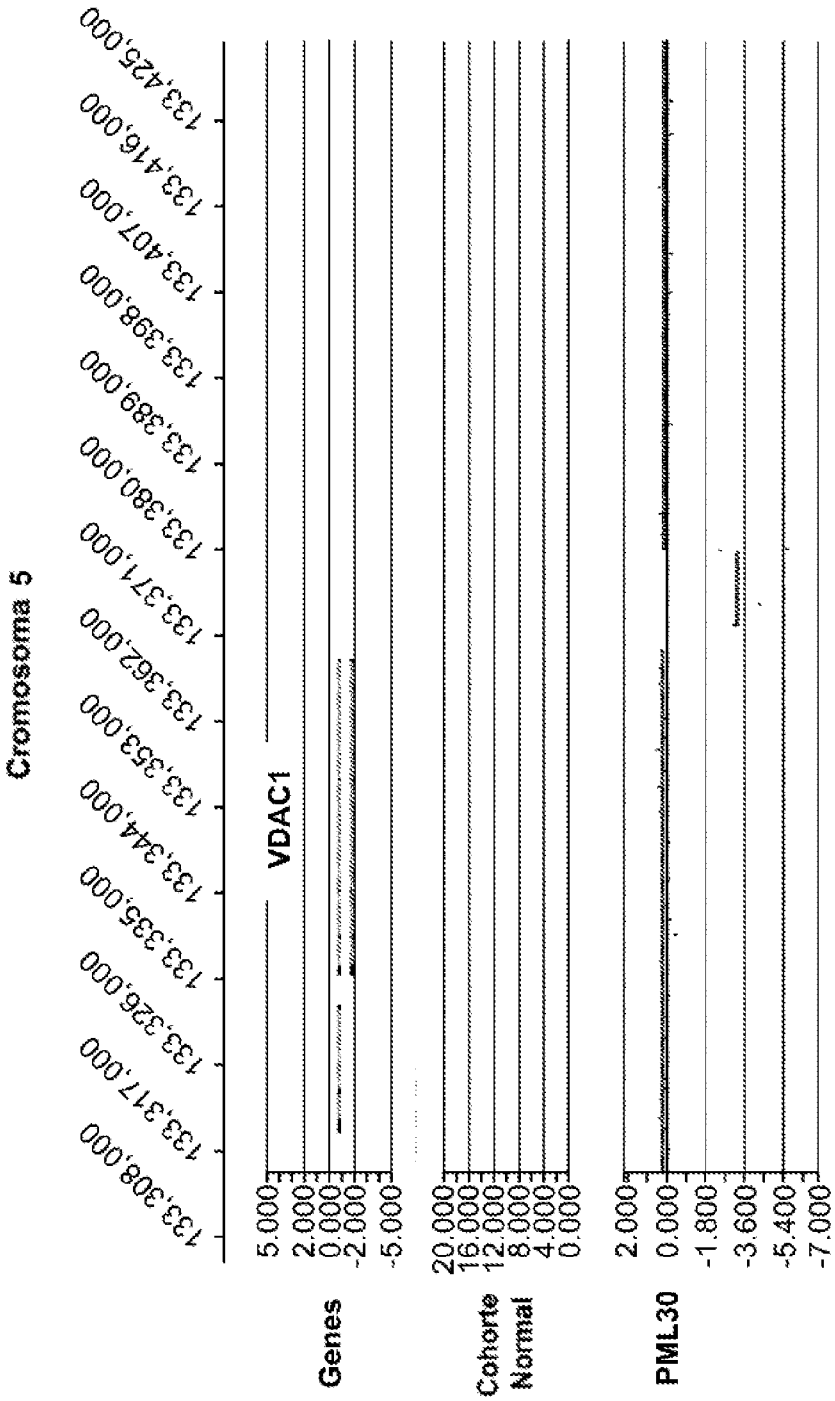

Cromosoma 5
133,308,000
133,317,000
133,326,000
133,335,000
133,344,000
133,353,000
133,362,000
133,371,000
133,380,000
133,389,000
133,398,000
133,407,000
133,416,000
133,425,000
Genes
VDAC1
5.000
2.000
0.000
-2.000
-5.000
Cohorte Normal
20.000
16.000
12.000
8.000
4.000
0.000
PML30
2.000
0.000
-1.800
-3.600
-5.400
-7.000

FIG. 10

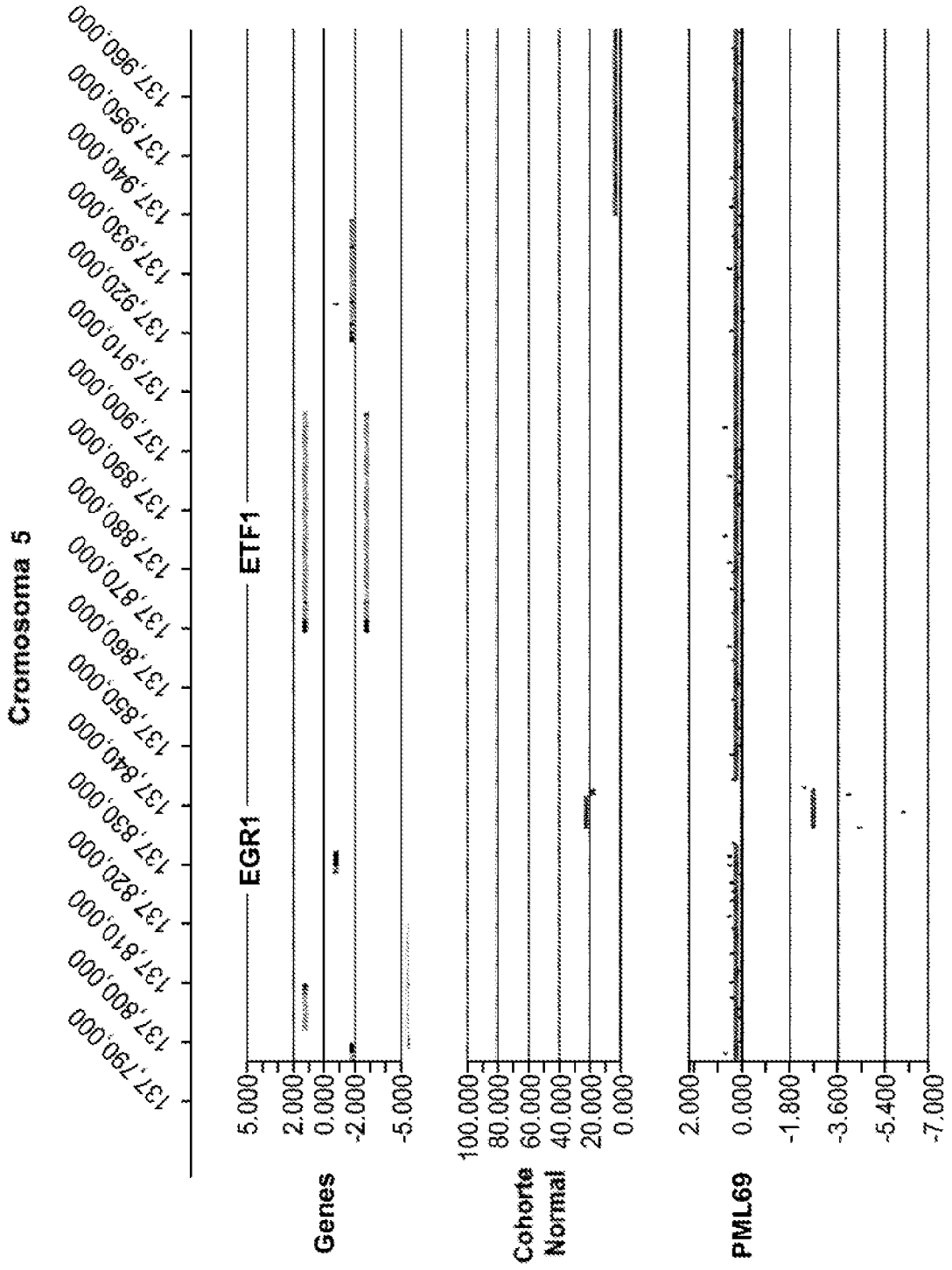

Cromosoma 5
137,790,000
137,800,000
137,810,000
137,820,000
137,830,000
137,840,000
137,850,000
137,860,000
137,870,000
137,880,000
137,890,000
137,900,000
137,910,000
137,920,000
137,930,000
137,940,000
137,950,000
137,960,000
Genes
EGR1
ETF1
5.000
2.000
0.000
-2.000
-5.000
Cohorte Normal
100.000
80.000
60.000
40.000
20.000
0.000
PML69
2.000
0.000
-1.800
-3.600
-5.400
-7.000

FIG. 11

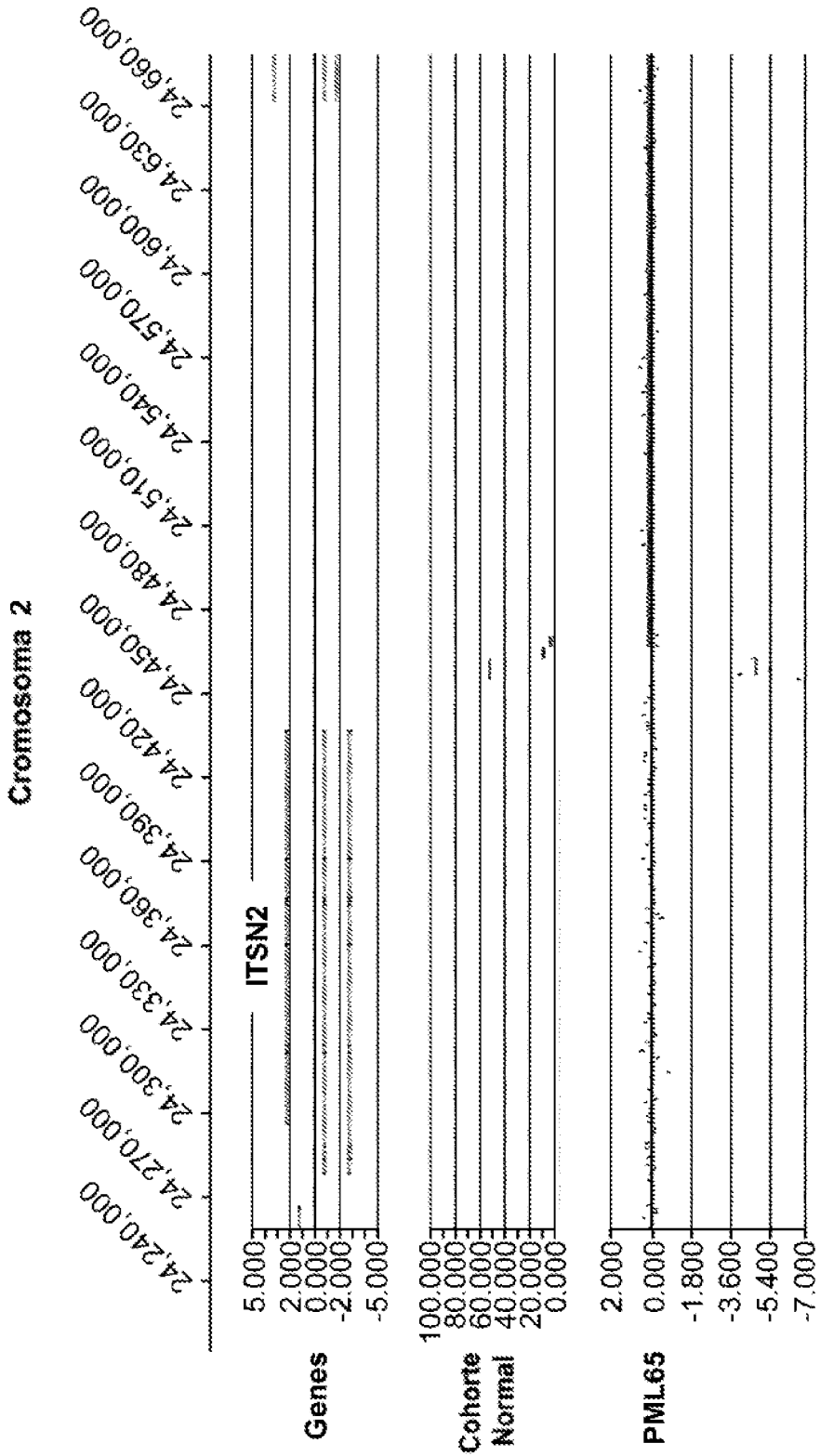

Cromosoma 2
24,240,000
24,270,000
24,300,000
24,330,000
24,360,000
24,390,000
24,420,000
24,450,000
24,480,000
24,510,000
24,540,000
24,570,000
24,600,000
24,630,000
24,660,000
ITSN2
Genes
5.000
2.000
0.000
-2.000
-5.000
Cohorte Normal
100.000
80.000
60.000
40.000
20.000
0.000
PML65
2.000
0.000
-1.800
-3.600
-5.400
-7.000

FIG. 12

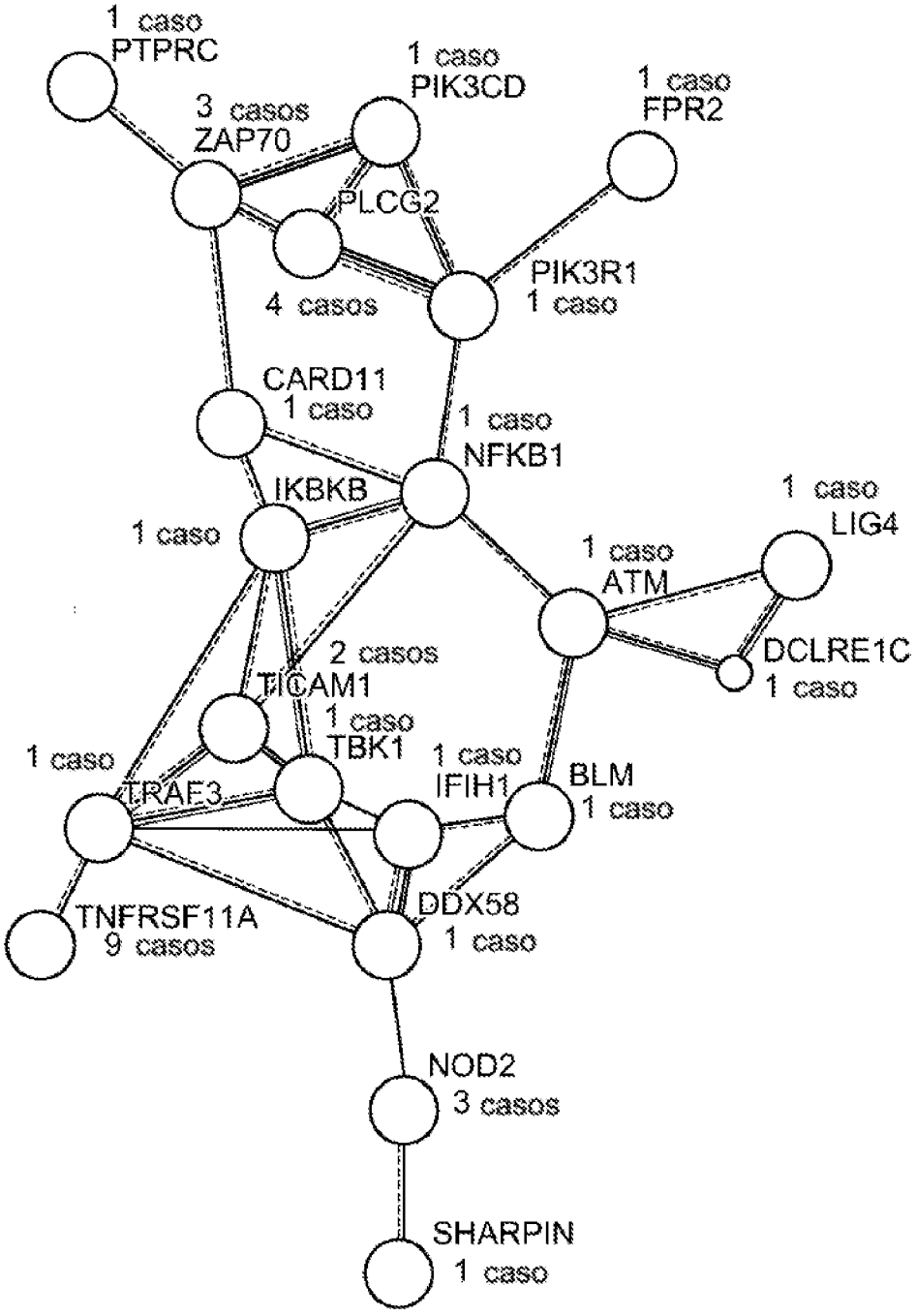

1 caso
PTPRC
1 caso
PIK3CD
1 caso
FPR2
3 casos
ZAP70
PLCG2
4 casos
PIK3R1
1 caso
CARD11
1 caso
1 caso
NFKB1
IKBKB
1 caso
1 caso
LIG4
1 caso
ATM
2 casos
TICAM1
DCLRE1C
1 caso
1 caso
TBK1
1 caso
1 caso
IFIH1
BLM
1 caso
TRAF3
TNFRSF11A
9 casos
DDX58
1 caso
NOD2
3 casos
SHARPIN
1 caso

FIG. 13