(19) DANMARK

(10) **DK/EP 4177356 T3**

Patent- og Varemærkestyrelsen

(12) Oversættelse af europæisk patentskrift

(51) Int.Cl.: ***C 12 Q 1/6883 (2018.01)***

(45) Oversættelsen bekendtgjort den: **2024-08-19**

(80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2024-05-08**

(86) Europæisk ansøgning nr.: **22199371.0**

(86) Europæisk indleveringsdag: **2019-08-08**

(87) Den europæiske ansøgnings publiceringsdag: **2023-05-10**

(30) Prioritet: **2018-08-08 US 201862716072 P**
2018-08-08 US 201862716183 P
2018-08-08 US 201861716183 P

(62) Stamansøgningsnr: **19765344.7**

(84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(73) Patenthaver: **PML Screening, LLC, 23 Corporate Plaza, Suite 150, Newport Beach, CA 92660, USA**
Université Paris-Saclay, Bâtiment Breguet , 3 Rue Joliot Curie, 91190 Gif sur Yvette, Frankrig
The Assistance Publique-Hopitaux de Paris (APHP), 149 rue de Sèvres, 75015 Paris, Frankrig
The Institut National de la Santé et de la Recherche Médicale (INSERM), 101 rue Tolbiac, 75654 Paris Cedex 13, Frankrig

(72) Opfinder: **HATCHWELL, Eli, , New York, NY 10036, USA**
EIS, Peggy S., , New York, NY 10036, USA
SMITH, Edward B., III, , New York, NY 10036, USA
TAOUFIK, Yassine, , 75014 Paris, Frankrig

(74) Fuldmægtig i Danmark: **NORDIC PATENT SERVICE A/S, Bredgade 30, 1260 København K, Danmark**

(54) Benævnelse: **FREMGANGSMÅDER TIL VURDERING AF RISIKO FOR UDVIKLING AF EN VIRUSSYGDOM VED HJÆLP AF EN GENETISK TEST**

(56) Fremdragne publikationer:
WO-A2-2018/158632
US-A1- 2017 016 919
ELIZABETH A. MILLS ET AL: "Understanding Progressive Multifocal Leukoencephalopathy Risk in Multiple Sclerosis Patients Treated with Immunomodulatory Therapies: A Bird's Eye View", FRONTIERS IN IMMUNOLOGY, vol. 9, 2 February 2018 (2018-02-02), XP055647863, DOI: 10.3389/fimmu.2018.00138
GUERINI ET AL: "Analysis of CCR5, CCR2, SDF1 and RANTES gene polymorphisms in subjects with HIV-related PML and not determined leukoencephalopathy", BIOMEDICINE AND PHARMACOTHERAPY, ELSEVIER, FR, vol. 62, no. 1, 22 May 2007 (2007-05-22), pages 26-30, XP022418147, ISSN: 0753-3322, DOI: 10.1016/J.BIOPHA.2007.04.005

DESCRIPTION

Description

BACKGROUND OF THE DISCLOSURE

[0001] Progressive multifocal leukoencephalopathy (PML) is a rare and potentially fatal opportunistic infection of the central nervous system that is caused by a ubiquitous polyomavirus, the JC virus (JCV). While JCV is present at very high rates in the general population, PML remains a rare disorder, albeit an important one because of the poor survival and the severe neurological sequelae, and the recently demonstrated association with a variety of useful therapies, for example, natalizumab in multiple sclerosis (MS). A number of risk factors for PML have been described but these are better viewed as necessary but not sufficient. While these risk factors are highly relevant, they do not, on their own, predict who will develop PML, since the vast majority of individuals with these risk factors will not develop the disorder. Other factors need to be considered and there is growing evidence for the role of host genetic factors in susceptibility to PML US2017/016919 to Plavina et al. describes methods of assessing a patient's risk of developing PML. Mills et al. in "Understanding Progressive Multifocal Leukoencephalopathy Risk in Multiple Sclerosis Patients Treated with immunomodulatory Therapies: A Bird's eye view", A Bird's Eye View"; Front. Immunol. 9:138 describe metrics used to predict risk for PML.

[0002] The ability to more accurately predict who is at risk of developing PML will be of enormous benefit in the context of drug treatment with compounds that are highly effective in their disease context (natalizumab in MS, for example) but carry a risk of a devastating disorder. There is a need to develop a companion diagnostic testing, in order to effectively exclude those that were at risk of PML, in the process reassuring those with negative tests about their dramatically reduced risk of developing PML.

BRIEF DESCRIPTION OF THE DRAWINGS

[0003] The novel features of the disclosure are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present disclosure will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the disclosure are utilized, and the accompanying drawings.

Figure 1 represents an example of a gene (PRKCB) impacted by germline and acquired CNVs.

Figure 2 represents an example of genes (TNFRSF13C and CENPM) impacted by acquired CNVs.

Figure 3 represents an example of a gene (PKHD1) impacted by germline and acquired CNVs.

Figure 4 represents an example of a gene (BMPR2) impacted by a CNV (homozygous and heterozygous losses).

Figure 5 represents an example of a gene (COMMD6) impacted by a CNV (*e.g.*, homozygous duplication).

Figure 6 represents an example of genes (KCTD7, RABGEF1) directly and potentially impacted by a CNV (*e.g.*, homozygous duplication).

Figure 7 represents an example of a gene (FPR2) impacted by a CNV (*e.g.*, homozygous duplication).

Figure 8 represents an example of a gene (PIK3CD) impacted by a CNV (*e.g.*, homozygous loss).

Figure 9 represents an example of a gene (CD180) potentially impacted by an intergenic CNV gain (*e.g.*, homozygous duplication).

Figure 10 represents an example of a gene (VDAC1) potentially impacted by an intergenic CNV (homozygous loss).

Figure 11 represents an example of genes (EGR1 and ETF1) potentially impacted by an intergenic CNV (homozygous loss).

Figure 12 represents an example of a gene (ITSN2) potentially impacted by an intergenic CNV (homozygous loss).

Figure 13 represents an example of known and/or predicted protein interactions using the String database for 21 of 43 genes (non-redundant list) reported in **Table 7.** The number of PML cases found to harbor variants impacting a given gene is indicated next to each gene.

Figure 14 represents an example gene set analysis of protein-protein interactions using the String database described herein. The input gene list was 74 genes (see **Table 42**) and the largest network from the String database analysis output, a 24-gene network, is depicted. The genes are color-coded based on the GO pathway ID with the largest number of genes (26) that was in the top 5 GO results: GO:0006955, dark gray colored genes.

SUMMARY OF THE INVENTION

[0004] In a first aspect the present invention provides an immunosuppressive medicament for use as defined in claim 1. In a second aspect the present invention provides a method as defined in claim 9. In a third aspect the present invention provides a method as defined in claim 10. Further features are defined in the dependent claims.

[0005] The immune deficiency can be X-linked agammaglobulinemia (XLA), common variable immunodeficiency (CVID), severe combined immunodeficiency (SCID), acquired immune deficiency syndrome (AIDS), cancers of the immune system (e.g., leukemia), immune-complex diseases (e.g., viral hepatitis), or multiple myeloma.

[0006] In some embodiments, the multiple sclerosis is relapsing-remitting multiple sclerosis (RRMS). In some embodiments, the multiple sclerosis is primary progressive multiple sclerosis (PPMS). In some embodiments, the multiple sclerosis is secondary progressive multiple sclerosis (SPMS).

[0007] In some embodiments, the immunosuppressive medicament comprises interferon beta-1a, interferon beta-1b, glatiramer acetate, peginterferon beta-1a,teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, mitoxantrone, rituximab, natalizumab, daclizumab, ocrelizumab, diroximel fumarate, siponimod or any combination thereof.

[0008] In some embodiments, the subject has not taken the immunosuppressive medicament. In some embodiments, the subject has taken the immunosuppressive medicament. In some embodiments, the subject is taking the immunosuppressive medicament.

[0009] In some embodiments, the immunosuppressive medicament comprises natalizumab (e.g., TYSABRI). In some embodiments, at least about 10 mg of the natalizumab is administered, for example, at least about 10 mg, at least about 15 mg, at least about 20 mg, at least about 30 mg, at least about 40 mg, at least about 50 mg, at least about 60 mg, at least about 70 mg, at least about 80 mg, at least about 90 mg, at least about 100 mg, at least about 150 mg, at least about 200 mg, at least about 250 mg, or at least about 300 mg of the natalizumab is administered. In some embodiments, at least about 10 mg of the natalizumab is administered via intravenous infusion. In some embodiments, at least about 10 mg of the natalizumab is administered via intravenous infusion in four weeks. In some embodiments, about 100 mg to about 500 mg of the natalizumab is administered, for example, about 100 mg to about 200 mg, about 100 mg to about 300 mg, about 100 mg to about 400 mg, about 100 mg to about 500 mg, about 200 mg to about 300 mg, about 200 mg to about 400 mg, about 200 mg to about 500 mg, about 300 mg to about 400 mg, about 300 mg to about 500 mg, or about 400 mg to about 500 mg of the natalizumab is administered. In some embodiments, about 100 mg to about 500 mg of the natalizumab is administered via intravenous infusion. In some embodiments, about 100 mg to about 500 mg of the natalizumab is administered via intravenous infusion in four weeks. In some embodiments, about 300 mg of the natalizumab is administered. In some embodiments, about 300 mg of the natalizumab is administered via intravenous infusion. In some embodiments, about 300 mg of the natalizumab is administered via intravenous infusion in four weeks. In some embodiments, at least about 10 mg of the natalizumab is administered via intravenous infusion in six weeks. In some embodiments, at least about 10 mg of the natalizumab is administered via intravenous infusion in eight weeks. In some embodiments, about 100 mg to about 500 mg of the natalizumab is administered via intravenous infusion in six weeks. In some embodiments, about 100 mg to about 500 mg of the natalizumab is administered via intravenous infusion in eight weeks. In some embodiments, about 300 mg of the natalizumab is administered via intravenous infusion in six weeks. In some embodiments, about 300 mg of the natalizumab is administered via intravenous infusion in eight weeks.

[0010] In some embodiments, the immunosuppressive medicament comprises dimethyl fumarate. In some embodiments, about 100 mg to about 500 mg of the dimethyl fumarate is administered, for example, about 100 mg to about 200 mg, about 100 mg to about 300 mg, about 100 mg to about 400 mg, about 100 mg to about 500 mg, about 200 mg to about 300 mg, about 200 mg to about 400 mg, about 200 mg to about 500 mg, about 300 mg to about 400 mg, about 300 mg to about 500 mg, or about 400 mg to about 500 mg of the dimethyl fumarate is administered. In some embodiments, about 120 mg of the dimethyl fumarate is administered. In some embodiments, about 240 mg of the dimethyl fumarate is administered.

[0011] In some embodiments, the immunosuppressive medicament comprises diroximel fumarate. In some embodiments, the immunosuppressive medicament comprise diroximel fumarate. In some embodiments, about 100 mg to about 500 mg of the diroximel fumarate is administered, for example, about 100 mg to about 200 mg, about 100 mg to about 300 mg, about 100 mg to about 400 mg, about 100 mg to about 500 mg, about 200 mg to about 300 mg, about 200 mg to about 400 mg, about 200 mg to about 500 mg, about 300 mg to about 400 mg, about 300 mg to about 500 mg, or about 400 mg to about 500 mg of the diroximel fumarate is administered. In some embodiments, about 400, 410, 420, 430, 440, 450, 460, 462, 470, 480, 490 or 500 mg of the diroximel fumarate is administered.

[0012] In some embodiments, the immunosuppressive medicament comprises fingolimod. In some embodiments, about 0.01 mg to about 5 mg of the fingolimod is administered, for example, about 0.01 mg to about 2 mg, about 0.01 mg to about 3 mg, about 0.01 mg to about 4 mg, about 0.01 mg to about 5 mg, about 0.1 mg to about 2 mg, about 0.1 mg to about 3 mg, about 0.1 mg to about 4 mg, about 0.1 mg to about 5 mg, about 0.2 mg to about 3 mg, about 0.2 mg to about 4 mg, about 0.2 mg to about 5 mg, about 0.3 mg to about 4 mg, about 0.3 mg to about 5 mg, about 0.4 mg to about 5 mg, about 0.1 mg to about 0.2 mg, about 0.1 mg to about 0.3 mg, about 0.1 mg to about 0.4 mg, about 0.1 mg to about 0.5 mg, about 0.2 mg to about 0.3 mg, about 0.2 mg to about 0.4 mg, about 0.2 mg to about 0.5 mg, about 0.3 mg to about 0.4 mg, about 0.3 mg to about 0.5 mg, about 0.4 mg to about 0.5 mg, or about 0.4 mg to about 0.6 mg of the

fingolimod is administered. In some embodiments, about 0.25 mg or 0.5 mg of the fingolimod is administered.

[0013] In some embodiments, the immunosuppressive medicament comprises rituximab. In some embodiments, about 100 mg to about 1000 mg of the rituximab is administered, for example, about 100 mg to about 200 mg, about 100 mg to about 300 mg, about 100 mg to about 400 mg, about 100 mg to about 500 mg, about 100 mg to about 600 mg, about 100 mg to about 700 mg, about 100 mg to about 800 mg, about 100 mg to about 900 mg of the rituximab is administered. The dose may be by weight or a fixed dose. In some embodiments, about 250 mg/m^2, 375 mg/m^2, 500 mg/m^2, 500 mg, or 1000 mg of the rituximab is administered. In some embodiments, about 250 mg/m^2, 375 mg/m^2, 500 mg/m^2, 500 mg, or 1000 mg of the rituximab is administered every week, every 2 weeks, every 4 weeks, every 8 weeks, or every 6 months. In some embodiments, about 250 mg/m^2, 375 mg/m^2, 500 mg/m^2, 500 mg, or 1000 mg of the rituximab is administered every 8 weeks or every 6 months for treating MS. The total dose cab be from about 50 and 4000 mg, for example, from about 75 and 3000 mg, from about 100 and 2000 mg, from about 100 and 1000 mg, from about 150 and 1000 mg, or from about 200 and 1000 mg, including doses of about 200, 300, 400, 500, 600, 700, 800, 900, 1000 mg, and 2000 mg. These doses may be given as a single dose or as multiple doses, for example, two to four doses. Such doses may be done by infusions, for example.

[0014] In some embodiments, the immunosuppressive medicament comprises siponimod. In some embodiments, about 0.1 mg to about 5 mg of the siponimod is administered. In some embodiments, about 1 mg or about 2 mg of the siponimod is administered. In some embodiments, about 1 mg or about 2 mg of the siponimod is administered to a subject with a CYP2C9* 1/*3 or CYP2C9*2/*3 genotype.

[0015] In some embodiments, the subject does not have one or more genetic variations associated with a risk of developing PML. In some embodiments, the subject does not have one or more genetic variations associated with a high risk of developing PML.

[0016] In some embodiments, the genetic test comprises detecting the genomic variations of the claims in a polynucleic acid sample from the subject. In some embodiments, the genetic test comprises detecting one or more genomic variations of the claims in a polynucleic acid sample from the subject.

[0017] The subject may be in need of natalizumab, interferon beta-1a, interferon beta-1b, glatiramer acetate, peginterferon beta-1a, teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, mitoxantrone, rituximab, daclizumab, ocrelizumab, diroximel fumarate or siponimod therapy, and the immunosuppressive medicament may be natalizumab, interferon beta-1a, interferon beta-1b, glatiramer acetate, peginterferon beta-1a, teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, mitoxantrone, rituximab, daclizumab, ocrelizumab, diroximel fumarate or siponimod.

[0018] In an aspect, the invention provides natalizumab, interferon beta-1a, interferon beta-1b, glatiramer acetate, peginterferon beta-1a, teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, mitoxantrone, rituximab, daclizumab, ocrelizumab, diroximel fumarate or siponimod for use in treating multiple sclerosis, wherein the subject is identified as not having a genomic variation of the claims.

[0019] In an aspect, the invention provides natalizumab, interferon beta-1a, interferon beta-1b, glatiramer acetate, peginterferon beta-1a, teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, mitoxantrone, rituximab, daclizumab, ocrelizumab, diroximel fumarate or siponimod for use in treating Crohn's disease, wherein the subject is identified as not having a genomic variations of the claims.

[0020] Provided herein is a method of reducing a risk of a subject developing progressive multifocal leukoencephalopathy (PML) comprising testing a subject for the presence of a genomic variation of the claims, determining that the subject has the genomic variation, and advising against administering natalizumab to the subject that was determined to have at least one of the genomic variations.

[0021] In some embodiments, the subject is identified as not having a risk of developing progressive multifocal leukoencephalopathy (PML) by a genetic test. In some embodiments, the subject is identified as not having a high risk of developing progressive multifocal leukoencephalopathy (PML) by a genetic test.

[0022] In some embodiments, the immunosuppressive medicament for use or the methods of the invention further comprise analyzing for a presence of JCV in a biological sample from the subject. In some embodiments, the analyzing comprises a JCV-antibody test. In some embodiments, the JCV-antibody test has a negative result. In some embodiments, the JCV-antibody test does not detect a presence of JCV in the biological sample from the subject. In some embodiments, the JCV-antibody test detects a presence of JCV in the biological sample from the subject.

[0023] In some embodiments, the subject is on an immunosuppressive therapy.

[0024] In some embodiments, the analyzing for a presence of JCV comprises a JCV-antibody test, a CD62L test, or a CSF IgM oligoclonal bands test. In some embodiments, the analyzing for a presence of JCV is performed prior to the genetic test. In some embodiments, the analyzing for a presence of JCV is performed after the genetic test. In some embodiments, the analyzing for a presence of JCV is performed concurrently with the genetic test. In some embodiments, the analyzing for a presence of JCV identifies the subject as having JCV. In some embodiments, the analyzing for a presence of JCV identifies the subject as not having JCV. In some embodiments, the genetic test result identifies the subject as having a risk or an increased risk of developing PML. In some embodiments, the genetic test result identifies the subject as not having a risk or not having an increased risk of developing PML.

[0025] In some embodiments, the subject is immunosuppressed. In some embodiments, the subject has HIV. In some embodiments, the subject has HIV infection. In some embodiments, the subject is at risk of HIV infection.

[0026] In some embodiments, the condition is a cancer, a hematologic malignancy, an organ transplant, or an autoimmune disease. In some embodiments, the condition is idiopathic CD4+ lymphocytopenia (ICL).

[0027] In some embodiments, the condition is an autoimmune disease.

[0028] In some embodiments, the autoimmune disease is selected from the group consisting of Addison disease, Behcet's Disease, Inflammatory bowel disease, Celiac disease - sprue (gluten-sensitive enteropathy), Crohn's disease, Dermatomyositis, Focal segmental glomerulosclerosis, Graves disease, Hashimoto thyroiditis, Multiple sclerosis, Myasthenia gravis, Pemphigus, Pemphigoid, Aplastic anemia, Pernicious anemia, Autoimmune hemolytic anemia, Erythroblastopenia, Thrombocytopenic purpura, Evans syndrome, Vasculitis, Granulomatosis with polyangiitis, Chronic inflammatory demyelinating polyneuropathy, Guillain-Barre syndrome, Anti-NMDA receptor encephalitis, Devic's disease, Autoimmune pancreatitis, Opsoclonus myoclonus syndrome, IgG4-related disease, Psoriasis, Reactive arthritis, Rheumatoid arthritis, Juvenile idiopathic arthritis, Sarcoidosis, Sjögren syndrome, Systemic lupus erythematosus, Type I diabetes, Vitiligo, or Ulcerative colitis.

[0029] In some embodiments, the immunosuppressive medicament comprises an interferon beta-1a, interferon beta-1b, glatiramer acetate, peginterferon beta-1a, teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, mitoxantrone, rituximab, natalizumab, daclizumab, ocrelizumab, siponimod or any combination thereof.

[0030] In an aspect the invention provides natalizumab, interferon beta-1a, interferon beta-1b, glatiramer acetate, peginterferon beta-1a, teriflunomide, fingolimod, dimethyl fumarate, alemtuzumab, mitoxantrone, rituximab, daclizumab, ocrelizumab, diroximel fumarate or siponimod for use in treating a subject in need thereof, wherein the subject has a decreased risk of progressive multifocal leukoencephalopathy (PML) due to an infection of the brain by John Cunningham virus (JCV), wherein the subject's decreased risk is due to the absence of a genomic variation of the claims.

[0031] In an aspect the invention provides an immunosuppressive medicament for use in treating a condition in a subject in need of immunosuppressive medicament therapy, wherein the subject has a decreased risk of progressive multifocal leukoencephalopathy (PML) due to an infection of the brain by John Cunningham virus (JCV), wherein the subject's decreased risk is due to the absence of a genomic variation of the claims.

[0032] In some embodiments, the immunosuppressive medicament comprises an antibody molecule or a fragment thereof. In some embodiments, the antibody molecule or a fragment thereof is a recombinant antibody molecule or a fragment thereof. In some embodiments, the antibody molecule or a fragment thereof is a humanized antibody molecule or a fragment thereof. In some embodiments, the antibody molecule or fragment thereof is a humanized recombinant antibody molecule or fragment thereof. In some embodiments, the antibody molecule or fragment thereof is a humanized recombinant IgG4κ monoclonal antibody molecule or fragment thereof. In some embodiments, the antibody molecule or fragment thereof comprises a sequence in CAS Registry Number: 189261-10-7. In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody heavy chain. In some embodiments, the antibody molecule or fragment thereof comprises two antibody heavy chains. In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody light chain. In some embodiments, the antibody molecule or fragment thereof comprises two antibody light chains. In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody heavy chain and at least one antibody light chain.

[0033] In some embodiments, the antibody molecule or fragment thereof is produced in myeloma cells. In some embodiments, the antibody molecule or fragment thereof is produced in rabbit hybridoma cells.

[0034] In some embodiments, the antibody molecule or fragment thereof binds a receptor. In some embodiments, the antibody molecule or fragment thereof binds an integrin. In some embodiments, the integrin is expressed on surface of a leukocyte. In some embodiments, the leukocyte is a neutrophil. In some embodiments, the leukocyte is not a neutrophil. In some embodiments, the antibody molecule or a fragment thereof binds α4β1 integrin, α4β7 integrin, or both. In some embodiments, the antibody molecule or a fragment thereof binds α4-subunit of α4β1 integrin, α4β7 integrin, or both. In some embodiments, the antibody molecule or a fragment thereof inhibits α4-mediated adhesion of a leukocyte to its receptor.

[0035] In some embodiments, the immunosuppressive medicament comprises an antibody or a fragment thereof, which comprises a sequence that has at least about 50%, 60%, 70%, 80%, 90%, 95%, or 100% sequence identity to SEQ ID NO. 3527 (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). In some embodiments, the antibody or fragment thereof comprises a sequence that has about 50%-100% identity, for example, about 50%-60%, about 50%-70%, about 60%-70%, about 60%-80%, about 70%-80%, about 70%-90%, about 80%-90%, about 80%-95%, about 90%-95%, about 90%-99%, about 90%-100%, about 95%-99%, or about 99%-100% sequence identity to SEQ ID NO. 3275.

[0036] In some embodiments, the immunosuppressive medicament comprises an antibody or a fragment thereof, which comprises a sequence that has at least about 50%, 60%, 70%, 80%, 90%, 95%, or 100% sequence identity to SEQ ID NO. 3528 (DIQMTQSPSS LSASVGDRVT ITCKTSQDIN KYMAWYQQTP GKAPRLLIHY TSALQPGIPS RFSGSGSGRD YTFTISSLQP EDIATYYCLQ YDNLWTFGQG TKVEIKRTVA APSVFIFPPS DEQLKSGTAS VVCLLNNFYP REAKVQWKVD NALQSGNSQE SVTEQDSKDS TYSLSSTLTL SKADYEKHKV YACEVTHQGL SSPVTKSFNRGEC). In some embodiments, the antibody or fragment thereof comprises a sequence that has about 50%-100% identity, for example, about 50%-60%, about 50%-70%, about 60%-70%, about 60%-80%, about 70%-80%, about 70%-90%, about 80%-90%, about 80%-95%, about 90%-95%, about 90%-99%, about 90%-100%, about 95%-99%, or about 99%-100% sequence identity to SEQ ID NO. 3276.

[0037] In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody heavy chain, or an α4-binding fragment thereof, comprising non-human CDRs at positions 31-35 (CDR1), 50-65 (CDR2) and 95-102 (CDR3) (Kabat numbering) from a mouse anti-α4 antibody and having non-human residues at framework positions 27-30 (Kabat numbering), wherein the positions 27-30 have the amino acid sequence Phe 27, Asn 28, Ile 29 and Lys 30.

[0038] In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody light chain, or an α4-binding fragment thereof, comprising: a light chain (LC) CDR1 with an amino acid sequence of SEQ ID NO.: 3529 (KTSQDINKYMA), a LC CDR2 with an amino acid sequence of SEQ ID NO.: 3530 (YTSALQP), and a LC CDR3 with an amino acid sequence of SEQ ID NO.: 3531 (LQYDNLWT).

[0039] In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody light chain, or an α4-binding fragment thereof, comprising: a light chain (LC) CDR1 with an amino acid sequence of SEQ ID NO.: 3532 (QASQDIIKYLN), a LC CDR2 with an amino acid sequence of SEQ ID NO.: 3533 (EASNLQA), and a LC CDR3 with an amino acid sequence of SEQ ID NO.: 3534 (QQYQSLPYT).

[0040] In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody light chain, or an α4-binding fragment thereof, comprising: a light chain (LC) CDR1 with an amino acid sequence of SEQ ID NO.: 3535 (KASQSVTNDVA), a LC CDR2 with an amino acid sequence of SEQ ID NO.: 3536 (YASNRYT), and a LC CDR3 with an amino acid sequence of SEQ ID NO.: 3537 (QQDYSSPYT).

[0041] In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody heavy chain, or an α4-binding fragment thereof, comprising: a heavy chain (HC) CDR1 with an amino acid sequence of SEQ ID NO.: 3538 (DTYIH), a HC CDR2 with an amino acid sequence of SEQ ID NO.: 3539 (RIDPANGYTKYDPKFQG), and a HC CDR3 with an amino acid sequence of SEQ ID NO.: 3540 (EGYYGNYGVYAMDY).

[0042] In some embodiments, the antibody molecule or fragment thereof comprises at least one antibody heavy chain, or an α4-binding fragment thereof, comprising: a heavy chain (HC) CDR1 with an amino acid sequence of SEQ ID NO.: 3541 (DTYMH), a HC CDR2 with an amino acid sequence of SEQ ID NO.: 3542 (RIDPASGDTKYDPKFQV), and a HC CDR3 with an amino acid sequence of SEQ ID NO.: 3543 (DGMWVSTGYALDF).

[0043] In some embodiments, the antibody molecule or fragment thereof comprises a humanized heavy chain, or an a4-binding fragment thereof, comprising: a variable heavy chain region selected from the group consisting of SEQ ID NO.: 3544

(MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVRQPPGR GLEWIGRIDPASGDTKYDPKFQVKATITADTSSNQFSLRLSSVTAADTAVYYCADGMWVSTGY ALDFWGQGTTVTVSSGES), SEQ ID NO.: 3545

(QVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVRQPPGRGLEWIGRIDPASGDTKYDPKF QVRVTMLVDTSSNQFSLRLSSVTSEDTAVYYCADGMWVSTGYALDFWGQGTTVTVSSGES), SEQ ID NO.: 3546

(MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTVSGFNIKDTYMHWVKQRPGR GLEWIGRIDPASGDTKYDPKFQVRVTMLVDTSSNQFSLRLSSVTAADTAVYYCADGMWVSTGY ALDFWGQGTTVTVSSGES), SEQ ID NO.: 3547

(MDWTWRVFCLLAVAPGAHSQVQLQESGPGLVRPSQTLSLTCTASGFNIKDTYMHWVRQPPGR GLEWIGRIDPASGDTKYDPKFQVRVTMLVDTSSNQFSLRLSSVTAADTAVYYCADGMWVSTGY ALDFWGQGTTVTVSSGES), and SEQ ID NO.: 3548

(QVQLVQSGAEVKKPGASVKVSCKASGFNIKDTYIHWVRQAPGQRLEWMGRIDPANGYTKYDP KFQGRVTITADTSASTAYMELSSLRSEDTAVYYCAREGYYGNYGVYAMDYWGQGTLVTVSS).

[0044] In some embodiments, the antibody molecule or fragment thereof comprises a humanized light chain, or an a4-binding fragment thereof, comprising a variable light chain region selected from the group consisting of SEQ ID NO.: 3549

(MGWSCIILFLVATATGVHSDIQLTQSPSSLSASVGDRVTITCKASQSVTNDVAWYQQKPGKAPK
LLIYYASNRYTGVPSRFSGSGSGTDFTFTISSLQPEDIATYYCQQDYSSPYTFGQGTKVEIKRK), SEQ ID NO.: 3550

(MGWSCIILFLVATATGVHSSIVMTQSPSSLSASVGDRVTITCKASQSVTNDVAWYQQKPGKAPK
LLIYYASNRYTGVPDRFSGSGYGTDFTFTISSLQPEDIATYYCQQDYSSPYTFGQGTKVEIKRK), SEQ ID NO.: 3551

(MGWSCIILFLVATATGVHSDIQMTQSPSSLSASVGDRVTITCKASQSVTNDVAWYQQKPGKAP
KLLIYYASNRYTGVPDRFSGSGYGTDFTFTISSLQPEDIATYYCQQDYSSPYTFGQGTKVEIKRK), and SEQ ID NO.: 3552

(DIQMTQSPSSLSASVGDRVTITCKTSQDINKYMAWYQQTPGKAPRLLIHYTSALQPGIPSRFSGS
GSGRDYTFTISSLQPEDIATYYCLQYDNLWTFGQGTKVEIKRTV).

[0045] In some embodiments, a biological product can be a regulatory agency-approved biological product. For example, the biological product can be approved by the U.S. Food and Drug Administration (FDA) and/or the European medicines Agency (EMA). In some embodiments, the biological product can be a reference product. In some cases, the biological product can be a biosimilar product. In some embodiments, the biological product can be an interchangeable product.

[0046] In some embodiments, a biosimilar product can be similar to a reference product (see, e.g. Table 67). In some embodiments, a biosimilar product can have no clinically meaningful differences in terms of safety and effectiveness from the reference product. In some embodiments, a biosimilar product can have the same clinically inactive components. In some embodiments, a biosimilar product can have different clinically inactive components. In some embodiments, a biosimilar product specifically interacts with a substrate and the reference product specifically interacts with the same substrate. In some embodiments, a response rate of human subjects administered the biosimilar product can be 50%-150% of the response rate of human subjects administered the reference product. For example, the response rate of human subjects administered the biosimilar product can be 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%, or 140%-150% of the response rate of human subjects administered the reference product. In some embodiments, a biosimilar product and a reference product can utilize the same mechanism or mechanisms of action for the condition or conditions of use prescribed, recommended, or suggested in the proposed labeling, but only to extent the mechanism or mechanisms are known for the reference product.

[0047] In some embodiments, an interchangeable product can be a biosimilar product that meets additional standards for interchangeability. In some embodiments, an interchangeable product can produce the same clinical result as a reference product in all of the reference product's licensed conditions of use. In some embodiments, an interchangeable product can be substituted for the reference product by a pharmacist without the intervention of the health care provider who prescribed the reference product. In some embodiments, when administered more than once to an individual, the risk in terms of safety or diminished efficacy of alternating or switching between use of the biological product and the reference product is not greater than the risk of using the reference product without such alternation or switch. In some embodiments, an interchangeable product can be a regulatory agency approved product. In some embodiments, a response rate of human subjects administered the interchangeable product can be 80%-120% of the response rate of human subjects administered the reference product. For example, the response rate of human subjects administered the interchangeable product can be 80%-100%, 80%-110%, 80%-120%, 90%-100%, 90%-110%, 90%-120%, 100%-110%, 100%-120%, or 110%-120 of the response rate of human subjects administered the reference product.

[0048] In some embodiments, the natalizumab is administered via intravenous infusion. In some embodiments, about 100 mg to about 500 mg of the natalizumab is administered.

[0049] In some embodiments, the method comprises testing the subject for a genetic predisposition for PML with a genetic assay. In some embodiments, the genetic assay has a diagnostic yield of at least 5%. In some cases, the genetic assay has a diagnostic yield of at least about 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99%. In some cases, the genetic assay has a diagnostic yield of about 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%, or 99%-100%. In some embodiments, the genetic assay has a diagnostic yield of at least 20%.

[0050] In an aspect the invention provides a method of identifying a subject as not having a risk of developing PML, comprising: (a) analyzing a polynucleic acid sample from the subject for a genomic variation of the claims, wherein the genomic variation is not present in the polynucleic acid sample; and (b) identifying the subject as not having a risk of developing PML.

DETAILED DESCRIPTION OF THE DISCLOSURE

[0051] The details of one or more inventive embodiments are set forth in the accompanying drawings, the claims, and in the description

herein. Other features, objects, and advantages of inventive embodiments disclosed and contemplated herein will be apparent from the description and drawings, and from the claims.

[0052] As used herein, unless otherwise indicated, the article "a" means one or more unless explicitly otherwise provided for.

[0053] As used herein, unless otherwise indicated, terms such as "contain," "containing," "include," "including," and the like mean "comprising."

[0054] As used herein, unless otherwise indicated, the term "or" can be conjunctive or disjunctive. As used herein, unless otherwise indicated, any embodiment can be combined with any other embodiment.

[0055] As used herein, unless otherwise indicated, some inventive embodiments herein contemplate numerical ranges. When ranges are present, the ranges include the range endpoints. Additionally, every subrange and value within the range is present as if explicitly written out.

[0056] As used herein, unless otherwise indicated, the term "about" in relation to a reference numerical value and its grammatical equivalents include a range of values plus or minus 10% from that value, such as a range of values plus or minus 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1% from that value. For example, the amount "about 10" includes amounts from 9 to 11.

[0057] As used herein, unless otherwise indicated, the term "biological product" refers to a virus, therapeutic serum, toxin, antitoxin, vaccine, blood, blood component or derivative, allergenic product, protein (any alpha amino acid polymer with a specific defined sequence that is greater than 40 amino acids in size), or analogous product, or arsphenamine or derivative of arsphenamine (or any trivalent organic arsenic compound), applicable to the prevention, treatment, or cure of a disease or condition of human beings.

[0058] As used herein, unless otherwise indicated, the term "biosimilar product" refers to 1) a biological product having an amino acid sequence that is identical to a reference product; 2) a biological product having a different amino acid sequence (e.g., N- or C-terminal truncations) from a reference product; or 3) a biological product having a different posttranslational modification (e.g., glycosylation or phosphorylation) from a reference product, wherein the biosimilar product and the reference product utilize the same mechanism or mechanisms of action for the prevention, treatment, or cure of a disease or condition.

[0059] As used herein, "mechanism of action" refers to an interaction or activity through which a drug product (e.g., a biological product) produces a pharmacological effect.

[0060] As used herein, unless otherwise indicated, the term "interchangeable product" refers to a biosimilar product, wherein a response rate of a human subject administered the interchangeable product is from 80% to 120% of the response rate of the human subject administered the reference product.

[0061] As used herein, unless otherwise indicated, the term "reference product" refers to 1) a biological product having an amino acid sequence that is identical to a biosimilar product; 2) a biological product having a different amino acid sequence (e.g., N- or C-terminal truncations) from a biosimilar product; or 3) a biological product having a different posttranslational modification (e.g., glycosylation or phosphorylation) from a biosimilar product, wherein the reference product and the biosimilar product utilize the same mechanism or mechanisms of action for the prevention, treatment, or cure of a disease or condition.

[0062] As used herein, unless otherwise indicated, any nonproprietary or generic name of a biological product includes the biological product and any biosimilar product thereof. For example, the nonproprietary name, filgrastim, refers to the biological product sold under the trade name NEUPOGEN; it also includes the biosimilar product, filgrastim-sndz, sold under the trade name ZARXIO. In another example, the nonproprietary name, natalizumab, refers to the biological product sold under the trade name TYSABRI; it also includes any biosimilar product of the biological product.

[0063] All drug molecules and compounds provided herein include all salts, polymorphs, prodrugs, tautomers, zwitterionic forms, etc. thereof.

Progressive Multifocal Leukoencephalopathy (PML)

[0064] Progressive multifocal leukoencephalopathy (PML) is a rare and usually fatal viral disease characterized by progressive damage or inflammation of the white matter of the brain at multiple locations. The cause of PML can be a type of polyomavirus called the John Cunningham (JC) virus (or JCV), which can be harmless except in cases of weakened immune systems. While JCV is present at very high rates in the general population, PML remains a rare disorder, albeit an important one because of the clinical sequelae.

[0065] PML can occur in patients with severe immune deficiency, which allows reactivation of the JC virus, such as: 1) most commonly among patients with acquired immune deficiency syndrome (AIDS) that results from infection with human immunodeficiency virus (HIV), 2) patients on immunosuppressive medicaments like corticosteroids for organ transplant (e.g., renal, liver, lung, and heart) and in people with cancer (e.g., Hodgkin's disease, leukemia, or lymphoma, and myeloproliferative neoplasms such as myelofibrosis), and 3) individuals

with autoimmune diseases (e.g., multiple sclerosis, rheumatoid arthritis, psoriasis, and systemic lupus erythematosus) with therapies that depress the immune response. Several immunosuppressive drugs have been reported in the context of drug-induced PML or drug-associated PML. For example, see: Melis et al. CNS Drugs. 2015;29(10):879-91); Maas et al. J Neurol. 2016 Oct;263(10):2004-21; Colin et al. Fundam Clin Pharmacol. 2016 Oct 13. Immunosuppressive medicaments can include, but are not limited to, a glucocorticoid, cytostatic, antibody, drug acting on immunophilins, interferon, opioid, TNF binding protein, mycophenolate, small biological agent, small molecule, organic compound, A2aR antagonist, Akt inhibitor, anti CD20, Anti-amyloidotic (AA) Agent, anti-CD37 protein therapeutic, anti-CTLA4 mAb, Anti-CXCR4, anti-huCD40 mAb, anti-LAG3 mAb, anti-PD-1 mAb, anti-PD-L1 agent, anti-PD-L1 agent, anti-PD-L1 mAb, anti-TGFb mAb, anti-TIGIT mAb, anti-TIM-3 mAb, Aurora kinase inhibitor, Bcl-2 Inhibitor, bifunctional fusion protein targeting TGFb and PD-L1, bispecific anti-PD-1 and anti-LAG3 mAb, CD1d ligand, CD40 agonist, Complement C5a inhibitor, CSF1R inhibitor, EZH2 inhibitor, FGFR3 inhibitor, FGFR4 inhibitor, FGFrR3 inhibitor, glucocorticoid-induced tumor necrosis factor receptor-related gene [GITR] agonist, glutaminase inhibitor, Human monoclonal antibody against IL-12, ICOS agonist, IDO1 inhibitor, IL2 mutein, IL2 receptor agonist, MEK inhibitor, multitargeted receptor tyrosine kinase inhibitor, neutrophil elastase inhibitor, Notch Inhibitor, p38 MAPK inhibitor, PD-1 inhibitor, recombinant human Flt3L, ROCK inhibitor, selective sphingosine-1-phosphate receptor modulator, Src kinase inhibitor, TLR4 agonist, TLR9 agonist, abatacept (e.g. ORENCIA), abrilumab, acalabrutinib, adalimumab, adrenocorticotropic hormone, agatolimod sodium, AJM300, aldesleukin, alefacept, alemtuzumab, alisertib, alvespimycin hydrochloride, alvocidib, ambrisentan (e.g. LETAIRIS), aminocamptothecin, amiselimod, anakinra, andecaliximab, andrographolides (a botanical medicinal herb also known as IB-MS), anifrolumab, antithymocyte Ig, apatinib, apelisib, asparaginase, atacicept, atezolizumab, avelumab, azacitidine, azathioprine, bafetinib, baminercept, baricitinib, basiliximab, becatecarin, begelomab, belatacept, belimumab, bemcentinib, bendamustine, bendamustine (e.g. bendamustine hydrochloride), betalutin with lilotomab, bevacizumab, BIIB033, BIIB059, BIIB061, bimekizumab, binimetinib, bleomycin, blinatumomab, BNZ-1, bortezomib (e.g. VELCADE), brentuximab vedotin, bryostatin 1, bucillamine, buparlisib, busulfan, canakinumab, capecitabine, carboplatin, carfilzomib, carmustine, cediranib maleate, cemiplimab, ceralifimod, cerdulatinib, certolizumab (e.g. certolizumab pegol), cetuximab, chidamide, chlorambucil, CHS-131, cilengitide, cirmtuzumab, cisplatin, cladribine, clazakizumab, clemastine, clioquinol, corticosteroids, cyclophosphamide, cyclosporine, cytarabine, cytotoxic chemotherapy, daclizumab, dalfampridine (e.g. AMPYRA), daprolizumab pegol, daratumumab, dasatinib, defactinib, defibrotide, denosumab, dexamethasone, diacerein, dimethyl fumarate, dinaciclib, diroximel fumarate (e.g. VUMERITY), doxorubicin, doxorubicin (e.g. doxorubicin hydrochloride), durvalumab, duvelisib, duvortuxizumab, eculizumab (e.g. SOLIRIS), efalizumab, eftilagimod alpha, EK-12 (a neuropeptide combination of metenkefalin and tridecactide), elezanumab, elotuzumab (e.g. EMPLICITI), encorafenib, enfuvirtida (e.g. FUZEON), entinostat, entospletinib, enzastaurin, epacadostat, epirubicin, epratuzumab, eritoran tetrasodium, etanercept, etoposide, etrolizumab, everolimus, evobrutinib, filgotinib, fingolimod (e.g. fingolimod hydrochloride), firategrast, fludarabine, fluorouracil, fontolizumab, forodesine hydrochloride, fostamatinib, galunisertib, ganetespib, ganitumab, gemcitabine, gemtuzumab ozogamicin, gerilimzumab, glasdegib, glassia, glatiramer acetate, glembatumumab vedotin, glesatinib, golimumab (e.g. SIMPONI), guadecitabine, hydrocortisone, hydroxychloroquine sulfate, hydroxyurea, ibritumomab tiuxetan, ibrutinib, ibudilast, idarubicin, idebenone, idelalisib, ifosfamide, iguratimod, imatinib, imexon, IMU-838, infliximab, inotuzumab ozogamicin, interferon alfa-2, interferon beta-1a, interferon beta-1b, interferon gamma-1, ipilimumab, irofulven, isatuximab, ispinesib, itacitinib, ixazomib, lapatinib, laquinimod, laromustine, ld-aminopterin, leflunomide, lenalidomide, lenvatinib, letrozole (e.g. FEMARA), levamisole, levocabastine, lipoic acid, lirilumab, lonafamib, lumiliximab, maraviroc (e.g. SELZENTRY), masitinib, mavrilimumab, melphalan, mercaptopurine, methotrexate, methoxsalen, methylprednisone, milatuzumab, mitoxantrone, mizoribine, mocetinostat, monalizumab, mosunetuzumab, motesanib diphosphate, moxetumomab pasudotox, muromonab-CD3, mycophenolate mofetil (e.g. mycophenolate mofetil hydrochloride), mycophenolic acid, namilumab, natalizumab, navitoclax, neihulizumab, nerispirdine, neurovax, niraparib, nivolumab, obatoclax mesylate, obinutuzumab, oblimersen sodium, ocrelizumab, ofatumumab, olokizumab, opicinumab, oprelvekin, osimertinib, otelixizumab, oxaliplatin, oxcarbazepine, ozanimod, paclitaxel, pacritinib, palifermin, panobinostat, pazopanib, peficitinib, pegfilgrastim (e.g. NEULASTA), peginterferon beta-1a, pegsunercept (peg stnf-ri), pembrolizumab, pemetrexed, penclomedine, pentostatin, perifosine, pevonedistat, pexidartinib, picoplatin, pidilizumab, pivanex, pixantrone, pleneva, plovamer acetate, polatuzumab vedotin, pomalidomide, ponatinib, ponesimod, prednisone/prednisolone, pyroxamide, R-411, ravulizimab-cwvz (e.g. (ULTOMIRIS), recombinant il-12, relatlimab, rhigf-1, rhigm22, rigosertib, rilonacept, ritonavir (e.g. NORVIR), rituximab, ruxolitinib, SAR442168/PRN2246, sarilumab, secukinumab, selumetinib, simvastatin, sintilimab, siplizumab, siponimod (e.g. MAYZENT), sirolimus (rapamycin), sirukumab, sitravatinib, sonidegib, sorafenib, sotrastaurin acetate, sunitinib, sunphenon epigallocatechin-gallate, tabalumab, tacrolimus (e.g. tacrolimus anhydrous), talabostat mesylate, talacotuzumab, tanespimycin, tegafur/gimeracil/oteracil, temozolomide, temsirolimus, tenalisib, terameprocol, teriflunomide, thalidomide, thiarabine, thiotepa, tipifarnib, tirabrutinib, tislelizumab, tivozanib, tocilizumab, tofacitinib, TR-14035, tregalizumab, tremelimumab, treosulfan, ublituximab, umbralisib, upadacitinib, urelumab, ustekinumab, varlilumab, vatelizumab, vedolizumab, veliparib, veltuzumab, venetoclax, vinblastine, vincristine, vinorelbine ditartrate, visilizumab, vismodegib, vistusertib, voriconazole (e.g. VFEND), vorinostat, vosaroxin, ziv-aflibercept, 2B3-201, 3PRGD2, 4SC-202, 506U78, 6,8-bis(benzylthio)octanoic acid, 68Ga-BNOTA-PRGD2, 852A, 89Zr-DFO-CZP, ABBV-257, ABL001, ABP 501, ABP 710, ABP 798, ABT-122, ABT-199, ABT-263, ABT-348, ABT-494, ABT-555, ABT-874, ABX-1431 HCl, ACP-196, ACP-319, ACT-128800, ACY-1215, AD 452, Ad-P53, ADCT-301, ADCT-402, ADL5859, ADS-5102, AFX-2, AGEN1884, AGEN2034, AGS67E, AIN457, AK106-001616, ALD518, ALKS 8700, ALT-803, ALT-803, ALX-0061, ALXN1007, ALXN6000, AMD3100, AMG 108, AMG 319, AMG 357, AMG 570, AMG 592, AMG 714, AMG 719, AMG 827, AMP-110, AP1903, APL A12, APO866, APX005M, AQ4N, AR-42, ARN-6039, ARQ 531, ARRY-371797, ARRY-382, ARRY-438162, ART-I02, ART621, ASK8007, ASN002, ASP015K, ASP1707, ASP2408, ASP2409, ASP5094, AT-101, AT7519M, AT9283, ATA188, ATN-103, ATX-MS-1467, AVL-292, AVP-923, AZD4573, AZD5672, AZD5991, AZD6244, AZD6738, AZD9056, AZD9150, AZD9567, AZD9668, B-701, BAF312, BAY1830839, BBI608, BCD-054, BCD-055, BCD-063, BCD-089, BCD-100, BCD-132, BCD-145, BEZ235, BG00012, BG9924, BGB-3111, BGB-A333, BGG492, BHT-3009, BI 655064, BI 695500, BI 695501, BI 836826, BI-1206, BIBR 796 BS, BIIB017, BIIB023, BIIB057, BIIB061, BIIL 284 BS, BLZ945, BMMNC, BMN 673, BMS-247550, BMS-582949, BMS-817399, BMS-936558, BMS-936564, BMS-945429, BMS-986104, BMS-986142, BMS-986156, BMS-986195, BMS-986205, BMS-986213, BMS-986226, BMS-986251, BNC105P, BOW015, BP1001, BT061, BTT-1023, C105, CAL-101, CAM-3001, CAT-8015, CB-839, CBL0137, CC-1088, CC-115, CC-122, CC-292, CC100, CCI-779, CCX 354-C, CDKI AT7519, CDP323, CDP6038, CDP870, CDX-1127, CDX-301, CE-224535,

CF101, CFZ533, CGP 77116, CH-1504, CH-4051, CHR-5154, CHS-0214, CK-2017357, CLAG-M, CLR 131, CMAB008, CMP-001, CNF2024 (BIIB021), CNM-Au8, CNTO 1275, CNTO 136, CNTO 148, CNTO 6785, CP-195543, CP-461, CpG 7909, CPI-1205, CR6086, CRx-102, CS-0777, CS1002, CT-011, CT-1530, CT-P10, CV301, CX-3543, DAC-HYP, DCDT2980S, DI-B4, DPA-714 FDG, DS-3032b, DT2219ARL, DTRM-505, DTRM-555, DTRMWXHS-12, DWP422, E6011, E7449, EK-12, ELND002, ENIA11, EOC202, ETBX-011, F8IL10, FBTA05, FEDAA1106 (BAY85-8101), FGF401, FKB327, FPA008, FR104, FS118, FTY720, G100, GCS-100, GDC-0199, GDC-0853, GEH120714, GLPG0259, GLPG0634, GNbAC1, GNKG168, GP2013, GP2015, GRN163L, GS-1101, GS-5745, GS-9219, GS-9820, GS-9876, GS-9901, GSK1223249, GSK1827771, GSK2018682, GSK21110183, GSK239512, GSK2618960, GSK2831781, GSK2982772, GSK3117391, GSK3152314A, GSK3196165, GSK3358699, GSK706769, GW-1000-02, GW274150, GW406381, GW856553, GZ402668, HCD122, HE3286, HL2351, HL237, hLL1-DOX (IMMU-115), HLX01, HM71224, HMPL-523, HSC835, HZT-501, ICP-022, IDEC-C2B8, ILV-094, IMGN529, IMMU-114, IMO-2125, INCAGN02385, INCB018424, INCB028050, INCB039110, INCB047986, INCMGA00012, INNO-406, INT131, INT230-6, INVAC-1, IPI-145, IPX056, ISF35, ISIS 104838, ITF2357, JCARH125, JHL1101, JNJ 38518168, JNJ-39758979, JNJ-40346527, JNJ-63723283, JS001, JTE-051, JTX-2011, KB003, KD025, KPT-330, KW-2449, KW-2478, KX2-391, L-778123, LAG525, LAM-002A, LBEC0101, LBH589, LFB-R603, LMB-2, LX3305, LY2127399, LY2189102, LY2439821, LY3009104, LY3090106, LY3300054, LY3321367, LY3337641, M2951, M7824, M923, MBG453, MBP8298, MBS2320, MD1003, MDG013, MDV9300, MDX-1100, MDX-1342, MDX-1411, ME-401, MEDI-522, MEDI-538, MEDI-551, MEDI4920, MGA012, MGCD0103, MGD007, MIS416, MK-0873, MK-4280, MK-4827, MK-8457, MK-8808, MK0359, MK0457, MK0752, MK0782, MK0812, MK2206, MLN1202, MLTA3698A, MM-093, MN-122, MN-166, monoclonal antibody M-T412, monoclonal antibody mono-dgA-RFB4, MOR00208, MOR103, MORAb-022, MP-435, MP470, MRC375, MRG-106, MS-533, MSB11022, MSC2490484A, MT-1303, MT-3724, MTIG7192A, MTRX1011A, NBI-5788, NC-503, NI-0101, NI-071, NIS793, NKTR-214, NNC 0141-0000-0100, NNC 0151-0000-0000, NNC0109-0012, NNC0114-0000-0005, NNC0114-0006, NNC0142-0002, NNC0215-0384, NNC109-0012, NOX-A12, NT-KO-003, NU100, OMB157, OMP-313M32, ON01910Na, ONO-2506PO, ONO-4641, ONTAK, OPB 31121, OSI-461, OTS167IV, P1446A-05, PBF-509, PBR06, PCI 32765, PCI-24781, PD 0360324, PDA001, PDR001, PF-04171327, PF-04236921, PF-04308515, PF-04629991, PF-05280586, PF-06342674, PF-06410293, PF-06438179, PF-06650833, PF-06651600, PF-06835375, PG-760564, PH-797804, PLA-695, PLX3397, PLX5622, POL6326, PROB 1921, PRO283698, PRTX-100, PS-341, PTL201, R(+)XK469, R788, RAD001, RC18, REGN1979, REGN3767, REGN2810, REGN4659, RFT5-SMPT-dgA, RG2077, RGB-03, RGI-2001, RHB-104, RNS60, RO5045337, RO7123520, Rob 803, RPC1063, RWJ-445380, S 55746, SAIT101, SAN-300, SAR245409, SB-681323, SB683699, SBI-087, SC12267 (4SC-101), SCH 727965, SCIO-469, SD-101, SG2000, SGN-40, SHC014748M, SHR-1210, SHR0302, SHR1020, SJG-136, SKI-O-703, SMP-114, SNS-032, SNS-062, SNX-5422, SPARC1103 I, SPC2996, SSR150106, STA 5326 mesylate, Sunpharma1505, SyB L-0501, Sym022, Sym023, SYN060, T-614, T0001, TA-650, TAB08, TAK-715, TAK-783, TAK-901, TGR-1202, TH-302, TL011, TMI-005, TMP001, TNFa Kinoid, TP-0903, TRU-015, TRU-016, TSR-022, TSR-033, TSR-042, TXA127, VAY736, VP-16, VSN16R, VX-509, VX-702, VX-745, VX15/2503, XCEL-MC-ALPHA, XL228, XL844, XmAb13676, XmAb5574, XOMA 052, YRA-1909, Z102, ZEN003365 or any combination thereof.

[0066] Exemplary small molecule immunosuppressive medicaments include dimethyl fumarate, fingolimod, diroximel fumarate, and ruxolitinib. In some embodiments, an immunosuppressive therapy is classified as a Class 1 (high risk) therapeutic agent, such as efalizumab and natalizumab as reported in Calabrese L. H. et al., Nat Rev Rheumatol. (2015).

[0067] In some cases, the immunosuppressive medicaments can be DNA and/or RNA crosslinking agents, including alkylating agents, nitrogen mustard alkylating agents, topoisomerase inhibitors, anthracyclines, and platinum-based anticancer drugs. In some cases, the immunosuppressive medicaments can be kinase inhibitors, including phosphoinositide-3-kinase, cyclin-dependent kinase (e.g., CDK9), Aurora kinase, ROCK, Akt, or PKC. In some cases, the immunosuppressive medicaments can be tyrosine kinase inhibitors, including inhibitors of the fusion protein breakpoint cluster region-Abelson murine leukemia viral oncogene homolog 1 (BCR-ABL), Bruton's tyrosine kinase (BTK), epidermal growth factor receptor (EGFR), Janus kinase (JAK), Syk, Lyn, MEK, FAK, BRAF, AXL, or vascular endothelial growth factor (VEGF). In some cases, the immunosuppressive medicaments can be monoclonal antibodies and/or antibody-drug conjugates directed at proteins including cluster of differentiation (CD) proteins, such as CD2, CD3, CD11a, CD20, CD30, CD52, CD-19, CD-38, CD-26, CD-37, CD-22, CD-33, CD-23, CD-74, CD-162, CD-79, CD-123, CD-4, CD-137, CD-27, CD-36, CD-39, CD-73, CD-226, CD-155, CD-40; interleukins (IL), such as IL-1, IL-2, IL-6, IL-12, IL-23; tumor necrosis factor (TNF) family proteins, such as TNFα; and integrins, such as integrin α4, $\alpha_v\beta_3$, $\alpha_v\beta_5$, $\alpha_v\beta_3$, or α_2. In some cases, the immunosuppressive medicaments can be monoclonal antibodies and/or antibody-drug conjugates directed at Programmed cell death receptor 1 (PD-1), Programmed cell death ligand 1 (PD-L1), Cytotoxic T-lymphocyte associated protein 4 (CTLA-4), Lymphocyte activation gene 3 (LAG-3), T-cell immunoglobulin and mucin-domain containing-3 (TIM-3), T-cell immunoreceptor with Ig and ITIM domains (TIGIT), also known as WUCAM or Vstm3, B and T lymphocyte attenuator (BTLA), Glucocorticoid-induced TNFR family related gene (GITR), OX40, HSP90, killer-cell immunoglobulin-like receptor (KIR), Toll-like receptor 9 (TLR9), Toll-like receptor 4 (TLR4), Matrix metallopeptidase 9 (MMP), Interferon receptor, Interferon gamma, Transforming growth factor 1b (TGF1β), Insulin growth factor 1 receptor (IGF1 R), Fibroblast growh factor receptor (FGFrR3, FGFR4), Neuromedin B, Granulocyte-macrophage colony stimulating factor receptor (GM-CSF R), Natural killer cell receptor (NKG-2a), Leucine rich repeat and Immunoglobin-like domain-containing protein 1 (LINGO1), B-cell activating factor (BAFF), Inducible T-cell co-stimulator (ICOS). In some cases, the monoclonal antibody/antibody-drug conjugate can activate the target.

[0068] In some cases, the monoclonal antibody/antibody-drug conjugate can inhibit the target. In some cases, the immunosuppressive medicaments can be inhibitors of RANKL (receptor activator of nuclear factor kappa-B ligand). In some cases, the immunosuppressive medicaments can be inhibitors of histone deacetylase (HDAC). In some cases, the immunosuppressive medicaments can be inhibitors of heat shock protein 90 (HSP90). In some cases, the immunosuppressive medicaments can be inhibitors of cytidine deaminase (CDA). In some cases, the immunosuppressive medicaments can be inhibitors of Hedgehog signaling pathway (including Sonic hedgehog and Smoothened). In some cases, the immunosuppressive medicaments can be inhibitors of alpha-1-proteinase. In some cases, the immunosuppressive medicaments can be inhibitors of cyclooxygenase 2 (COX2). In some cases, the immunosuppressive medicaments

can be inhibitors of complement (C5a). In some cases, the immunosuppressive medicaments can be inhibitors of colony stimulating factor 1 receptor (CSF1R). In some cases, the immunosuppressive medicaments can be inhibitors of Notch. In some cases, the immunosuppressive medicaments can be inhibitors of kinesin. In some cases, the immunosuppressive medicaments can be inhibitors of famesyltransferase. In some cases, the immunosuppressive medicaments can be inhibitors of poly(ADP-ribose) polymerase (PARP). In some cases, the immunosuppressive medicaments can be inhibitors of Neural Precursor Cell Expressed, Developmentally Down-Regulated (NEDD8). In some cases, the immunosuppressive medicaments can be inhibitors of dipeptidyl peptidase IV (DPP-IV). In some cases, the immunosuppressive medicaments can be inhibitors of leucine-rich repeat kinase 2 (LRRK2). In some cases, the immunosuppressive medicaments can be inhibitors of immune checkpoint proteins. In some cases, the immunosuppressive medicaments can be inhibitors of indoleamine 2,3-dioxygenase-1 (IDO1). In some cases, the immunosuppressive medicaments can be inhibitors of chemokine receptors (CCR4, CCR5, CCR7). In some cases, the immunosuppressive medicaments can be immunosuppression-inducing therapies such as T-cells or regulatory T-cells modified with a chimeric antigen receptor (CAR-T, CAR-Tregs). In some cases, the immunosuppressive medicaments can be structured lipids. In some cases, the immunosuppressive medicaments can be Ras mimetic. In some cases, the immunosuppressive medicaments can be inhibitors of NOD-like receptor pyrin domain-containing protein 3 (NLRP3). In some cases, the immunosuppressive medicaments can be mTOR and/or calcineurin inhibitors. In some cases, the immunosuppressive medicaments can be complement inhibitors. In some cases, the immunosuppressive medicaments can be immunosuppressive antimetabolites, nucleoside metabolic inhibitors, imidazole nucleosides, nucleotide analogs, nucleoside synthesis inhibitors, purine synthesis inhibitors, pyrimidine synthesis inhibitors, or pyrimidine synthase inhibitors. In some cases, the immunosuppressive medicaments can be recombinant proteins, such as recombinant interferon beta, IL-2, IL-11, Lymphotoxin B fusion protein, Therapeutic T cell receptor peptide vaccine, Keratinocyte growth factor, or Tumor necrosis factor (TNF) receptor.

[0069] In some cases, the immunosuppressive medicaments can be sphingosine-1-phosphate receptor and/or nicotinic acetylcholine receptor modulators. For example, siponimod (BAF312) can be used for the treatment of secondary progressive MS (Kappos L et al. 2018, PMID 29576505). Another medicament, ibudilast (MN-122), can be used for the treatment of progressive MS (Fox R et al. 2016, PMID 27521810). In some cases, the immunosuppressive medicaments can be therapeutic antibodies, including Immunoglobulin G. In some cases, the immunosuppressive medicaments can be asparaginase inhibitors. In some cases, the immunosuppressive medicaments can be B-lymphocyte stimulator (BLyS)-specific inhibitor. In some cases, the immunosuppressive medicaments can be T-cell costimulation modulators. In some cases, the immunosuppressive medicaments can be cyclic polypeptide immunosuppressants and/or synthetic polypeptides that modify immune processes. In some cases, the immunosuppressive medicaments can be corticosteroids. In some cases, the immunosuppressive medicaments can be cytotoxic chemotherapy drugs. In some cases, the immunosuppressive medicaments can be cytotoxic glycopeptide antibiotics and/or mixtures thereof. In some cases, the immunosuppressive medicaments can be molecules that inhibit pro-inflammatory cytokine production. In some cases, the immunosuppressive medicaments can be thalidomide analogues.

[0070] In some cases, the immunosuppressive medicament can be a Complement C5a inhibitor. In some cases the immunosuppressive medicament can be a CD40 agonist. In some cases, the immunosuppressive medicament can be a p38 inhibitor. In some cases, the immunosuppressive medicament can be a CSF1R inhibitor. In some cases, the immunosuppressive medicament can be a MEK inhibitor. In some cases, the immunosuppressive medicament can be a neutrophil elastase inhibitor. In some cases, the immunosuppressive medicament can be FGFrR3 inhibitor. In some cases, the immunosuppressive medicament can be anti-LAG3 mAb, Anti-CXCR, glucocorticoid-induced tumor necrosis factor receptor-related gene [GITR] agonist, IDO1 inhibitor, ICOS agonist, glutaminase inhibitor, recombinant human Flt3L, TLR9 agonist, EZH2 inhibitor, anti-CTLA4 mAb, PD-1 inhibitor, PD-L1 inhibitor, anti-PD-L1 mAb, FGFR4 inhibitor, bispecific anti-PD-1 and anti-LAG3 mAb, TLR4 agonist, Bcl-2 Inhibitor, or anti-LAG3 mAb. In some cases, the immunosuppressive medicaments can be inhibitors of cell degradation pathways, such as proteasome inhibitors. In some cases, the immunosupressive medicament can be selected from A2aR antagonist, Akt inhibitor, anti CD20, Anti-amyloidotic (AA) Agent, anti-CD37 protein therapeutic, anti-CTLA4 mAb, Anti-CXCR4, anti-huCD40 mAb, anti-LAG3 mAb, anti-PD-1 mAb, anti-PD-L1 agent, anti-PD-L1 agent, anti-PD-L1 mAb, anti-TGFb mAb, anti-TIGIT mAb, anti-TIM-3 mAb, Aurora kinase inhibitor, Bcl-2 Inhibitor, bifunctional fusion protein targeting TGFb and PD-L1, bispecific anti-PD-1 and anti-LAG3 mAb, CD1d ligand, CD40 agonist, Complement C5a inhibitor, CSF1R inhibitor, EZH2 inhibitor, FGFR3 inhibitor, FGFR4 inhibitor, FGFrR3 inhibitor, glucocorticoid-induced tumor necrosis factor receptor-related gene [GITR] agonist, glutaminase inhibitor, Human monoclonal antibody against IL-12, ICOS agonist, IDO1 inhibitor, IL2 mutein, IL2 receptor agonist, MEK inhibitor, multitargeted receptor tyrosine kinase inhibitor, neutrophil elastase inhibitor, Notch Inhibitor, p38 MAPK inhibitor, PD-1 inhibitor, recombinant human Flt3L, ROCK inhibitor, selective sphingosine-1-phosphate receptor modulator, Src kinase inhibitor, TLR4 agonist, TLR9 agonist.

[0071] In some cases, the immunosupressive medicament can be selected from 2B3-201, 3PRGD2, 4SC-202, 506U78, 6,8-bis(benzylthio)octanoic acid, 68Ga-BNOTA-PRGD2, 852A, 89Zr-DFO-CZP, ABBV-257, ABL001, ABP 501, ABP 710, ABP 798, ABT-122, ABT-199, ABT-263, ABT-348, ABT-494, ABT-555, ABT-874, ABX-1431 HCl, ACP-196, ACP-319, ACT-128800, ACY-1215, AD 452, Ad-P53, ADCT-301, ADCT-402, ADL5859, ADS-5102, AFX-2, AGEN1884, AGEN2034, AGS67E, AIN457, AK106-001616, ALD518, ALKS 8700, ALT-803, ALT-803, ALX-0061, ALXN1007, ALXN6000, AMD3100, AMG 108, AMG 319, AMG 357, AMG 570, AMG 592, AMG 714, AMG 719, AMG 827, AMP-110, AP1903, APL A12, APO866, APX005M, AQ4N, AR-42, ARN-6039, ARQ 531, ARRY-371797, ARRY-382, ARRY-438162, ART-I02, ART621, ASK8007, ASN002, ASP015K, ASP1707, ASP2408, ASP2409, ASP5094, AT-101, AT7519M, AT9283, ATA188, ATN-103, ATX-MS-1467, AVL-292, AVP-923, AZD4573, AZD5672, AZD5991, AZD6244, AZD6738, AZD9056, AZD9150, AZD9567, AZD9668, B-701, BAF312, BAY1830839, BBI608, BCD-054, BCD-055, BCD-063, BCD-089, BCD-100, BCD-132, BCD-145, BEZ235, BG00012, BG9924, BGB-3111, BGB-A333, BGG492, BHT-3009, BI 655064, BI 695500, BI 695501, BI 836826, BI-1206, BIBR 796 BS, BIIB017, BIIB023, BIIB057, BIIB061, BIIL 284 BS, BLZ945, BMMNC, BMN 673, BMS-247550, BMS-582949, BMS-817399, BMS-936558, BMS-936564, BMS-945429, BMS-986104, BMS-986142, BMS-986156, BMS-986195, BMS-986205, BMS-986213, BMS-986226,

BMS-986251, BNC105P, BOW015, BP1001, BT061, BTT-1023, C105, CAL-101, CAM-3001, CAT-8015, CB-839, CBL0137, CC-1088, CC-115, CC-122, CC-292, CC100, CCI-779, CCX 354-C, CDKI AT7519, CDP323, CDP6038, CDP870, CDX-1127, CDX-301, CE-224535, CF101, CFZ533, CGP 77116, CH-1504, CH-4051, CHR-5154, CHS-0214, CK-2017357, CLAG-M, CLR 131, CMAB008, CMP-001, CNF2024 (BIIB021), CNM-Au8, CNTO 1275, CNTO 136, CNTO 148, CNTO 6785, CP-195543, CP-461, CpG 7909, CPI-1205, CR6086, CRx-102, CS-0777, CS1002, CT-011, CT-1530, CT-P10, CV301, CX-3543, DAC-HYP, DCDT2980S, DI-B4, DPA-714 FDG, DS-3032b, DT2219ARL, DTRM-505, DTRM-555, DTRMWXHS-12, DWP422, E6011, E7449, EK-12, ELND002, ENIA11, EOC202, ETBX-011, F8IL10, FBTA05, FEDAA1106 (BAY85-8101), FGF401, FKB327, FPA008, FR104, FS118, FTY720, G100, GCS-100, GDC-0199, GDC-0853, GEH120714, GLPG0259, GLPG0634, GNbAC1, GNKG168, GP2013, GP2015, GRN163L, GS-1101, GS-5745, GS-9219, GS-9820, GS-9876, GS-9901, GSK1223249, GSK1827771, GSK2018682, GSK21110183, GSK239512, GSK2618960, GSK2831781, GSK2982772, GSK3117391, GSK3152314A, GSK3196165, GSK3358699, GSK706769, GW-1000-02, GW274150, GW406381, GW856553, GZ402668, HCD122, HE3286, HL2351, HL237, hLL1-DOX (IMMU-115), HLX01, HM71224, HMPL-523, HSC835, HZT-501, ICP-022, IDEC-C2B8, ILV-094, IMGN529, IMMU-114, IMO-2125, INCAGN02385, INCB018424, INCB028050, INCB039110, INCB047986, INCMGA00012, INNO-406, INT131, INT230-6, INVAC-1, IPI-145, IPX056, ISF35, ISIS 104838, ITF2357, JCARH125, JHL1101, JNJ 38518168, JNJ-39758979, JNJ-40346527, JNJ-63723283, JS001, JTE-051, JTX-2011, KB003, KD025, KPT-330, KW-2449, KW-2478, KX2-391, L-778123, LAG525, LAM-002A, LBEC0101, LBH589, LFB-R603, LMB-2, LX3305, LY2127399, LY2189102, LY2439821, LY3009104, LY3090106, LY3300054, LY3321367, LY3337641, M2951, M7824, M923, MBG453, MBP8298, MBS2320, MD1003, MDG013, MDV9300, MDX-1100, MDX-1342, MDX-1411, ME-401, MEDI-522, MEDI-538, MEDI-551, MEDI4920, MGA012, MGCD0103, MGD007, MIS416, MK-0873, MK-4280, MK-4827, MK-8457, MK-8808, MK0359, MK0457, MK0752, MK0782, MK0812, MK2206, MLN1202, MLTA3698A, MM-093, MN-122, MN-166, monoclonal antibody M-T412, monoclonal antibody mono-dgA-RFB4, MOR00208, MOR103, MORAb-022, MP-435, MP470, MRC375, MRG-106, MS-533, MSB11022, MSC2490484A, MT-1303, MT-3724, MTIG7192A, MTRX1011A, NBI-5788, NC-503, NI-0101, NI-071, NIS793, NKTR-214, NNC 0141-0000-0100, NNC 0151-0000-0000, NNC0109-0012, NNC0114-0000-0005, NNC0114-0006, NNC0142-0002, NNC0215-0384, NNC109-0012, NOX-A12, NT-KO-003, NU100, OMB157, OMP-313M32, ON01910 Na, ONO-2506PO, ONO-4641, ONTAK, OPB 31121, OSI-461, OTS167IV, P1446A-05, PBF-509, PBR06, PCI 32765, PCI-24781, PD 0360324, PDA001, PDR001, PF-04171327, PF-04236921, PF-04308515, PF-04629991, PF-05280586, PF-06342674, PF-06410293, PF-06438179, PF-06650833, PF-06651600, PF-06835375, PG-760564, PH-797804, PLA-695, PLX3397, PLX5622, POL6326, PROB 1921, PRO283698, PRTX-100, PS-341, PTL201, R(+)XK469, R788, RAD001, RC18, REGN1979, REGN3767, REGN2810, REGN4659, RFT5-SMPT-dgA, RG2077, RGB-03, RGI-2001, RHB-104, RNS60, RO5045337, RO7123520, Rob 803, RPC1063, RWJ-445380, S 55746, SAIT101, SAN-300, SAR245409, SB-681323, SB683699, SBI-087, SC12267 (4SC-101), SCH 727965, SCIO-469, SD-101, SG2000, SGN-40, SHC014748M, SHR-1210, SHR0302, SHR1020, SJG-136, SKI-O-703, SMP-114, SNS-032, SNS-062, SNX-5422, SPARC1103 I, SPC2996, SSR150106, STA 5326 mesylate, Sunpharma1505, SyB L-0501, Sym022, Sym023, SYN060, T-614, T0001, TA-650, TAB08, TAK-715, TAK-783, TAK-901, TGR-1202, TH-302, TL011, TMI-005, TMP001, TNFa Kinoid, TP-0903, TRU-015, TRU-016, TSR-022, TSR-033, TSR-042, TXA127, VAY736, VP-16, VSN16R, VX-509, VX-702, VX-745, VX15/2503, XCEL-MC-ALPHA, XL228, XL844, XmAb13676, XmAb5574, XOMA 052, YRA-1909, Z102, ZEN003365.

[0072] PML can be diagnosed in a patient with a progressive course of the disease, finding JC virus DNA in spinal fluid together with consistent white matter lesions on brain magnetic resonance imaging (MRI); alternatively, a brain biopsy can be diagnostic when the typical histopathology of demyelination, bizarre astrocytes, and enlarged oligodendroglial nuclei are present, coupled with techniques showing the presence of JC virus. Characteristic evidence of PML on brain CT scan images can be multifocal, non-contrast enhancing hypodense lesions without mass effect, but MRI can be more sensitive than CT. The most common area of involvement can be the cortical white matter of frontal and parieto-occipital lobes, but lesions may occur anywhere in the brain, like the basal ganglia, external capsule, and posterior cranial fossa structures like the brainstem and cerebellum.

[0073] In general, treatment of PML aims at reversing the immune deficiency to slow or stop the disease progress. Patients on an immunosuppression regime can stop taking the immunosuppressive medicament or plasma exchange (PLEX) can be used to accelerate the removal of the immunosuppressive medicament that put the person at risk for PML. HIV-infected patients can start highly active antiretroviral therapy (HAART). Occurrence of PML can also occur in the context of immune reconstitution inflammatory syndrome (IRIS), wherein onset of PML can occur or PML symptoms may get worse after cessation of immunosuppression (e.g., as reviewed by Pavlovic et al. Ther Adv Neurol Disord. 2015 Nov;8(6):255-73 and Bowen et al. Nat Rev Neurol. 2016 Oct 27;12(11):662-674). For example, in MS patients that develop PML during treatment with natalizumab, IRIS often results when treatment is stopped and PLEX is used to remove natalizumab from the patient's circulation. Treatment of IRIS in PML patients can include administration of corticosteroids. Other potential treatments of PML can include cidofovir, cytarabine, anti-malaria drug mefloquine, interleukin-2, and 1-O-hexadecyloxypropyl-cidofovir (CMX001, aka brincidofovir). As reviewed by Pavlovic (Ther Adv Neurol Disord. 2015 Nov;8(6):255-73), potential treatments for PML include antiviral agents (e.g., chlorpromazine, citalopram, mirtazapine, risperidone, ziprasidone, retro-2cycl, brefeldin A, cidofovir, brincidofovir, cytarabine, ganciclovir, leflunomide, topotecan, mefloquine, 3-aminobenzamide, imatinib, and Ag122), immune response modulators (e.g., IFN-alpha, IL-2, IL-7, maraviroc, and glucocorticoids), and immunization (e.g., recombinant human anti-JCV VP-1 monoclonal antibodies, JCV-specific cytotoxic T lymphocyte therapy, IL-7 plus JCV VP1 vaccine, and JCV oral vaccine).

[0074] The term "diagnostic yield" as used herein refers to the percentage of cases that would identify the presence of one or more genetic variations (*e.g.*, CNV, SNV) in a PML cohort using an assay. For example, if 40 cases would identify the presence of one or more genetic variations (*e.g.,* CNV, SNV) in a cohort of 100 PML patients, the diagnostic yield of the assay is 40%. In some cases, the patients in the PML cohort are clinically diagnosed with PML. In some cases, a patient is clinically diagnosed with PML when JC virus DNA is present in spinal fluid and consistent white matter lesions is present on brain magnetic resonance imaging (MRI). In some cases, a patient is clinically diagnosed with PML when typical histopathology of demyelination, bizarre astrocytes, and enlarged oligodendroglial nuclei are present in a brain biopsy, coupled with the presence of JC virus. In some cases, the PML cohort has at least 5 PML cases, for

example, at least 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900, or 1000 PML cases. In some cases, the PML cohort is a cohort listed herein. For example, the PML cohort is the PML patient cohort listed in **Table 7.** In some cases, the assay is JCV-antibody assay. In some cases, the assay is not JCV-antibody assay. In some cases, the assay is a genetic assay. In some cases, the genetic assay tests the genetic predisposition for PML.

[0075] The genetic assay can comprise any method disclosed herein. In some cases, the genetic assay has a diagnostic yield of at least about 5%, 6%, 7%, 8%, 9%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99%. In some cases, the genetic assay has a diagnostic yield of about 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%, or 99%-100%.

Genetic Variations Associated with PML

[0076] The present invention involves a genomic variation selected from chr9: 137779251, G>A and chr1:160769595, AG>A. The following provides a general explanation of genetic variations, and describes examples of variations not forming subject matter being claimed herein, insofar as it relates to variations other than chr9:137779251, G>A and chr1:160769595, AG>A. Described herein, are methods that can be used to detect genetic variations. Detecting specific genetic variations, for example polymorphic markers and/or haplotypes, copy number, absence or presence of an allele, or genotype associated with a condition (e.g., disease or disorder) as described herein, can be accomplished by methods known in the art for analyzing nucleic acids and/or detecting sequences at polymorphic or genetically variable sites, for example, amplification techniques, hybridization techniques, sequencing, microarrays/arrays, or any combination thereof. Thus, by use of these methods disclosed herein or other methods available to the person skilled in the art, one or more alleles at polymorphic markers, including microsatellites, single nucleotide polymorphisms (SNPs), single nucleotide variations (SNVs), insertions/deletions (indels), copy number variations (CNVs), or other types of genetic variations, can be identified in a sample obtained from a subject.

[0077] Genomic sequences within populations exhibit variability between individuals at many locations in the genome. For example, the human genome exhibits sequence variations that occur on average every 500 base pairs. Such genetic variations in polynucleic acid sequences are commonly referred to as polymorphisms or polymorphic sites. As used herein, a polymorphism, e.g., genetic variation, includes a variation in the sequence of the genome amongst a population, such as allelic variations and other variations that arise or are observed. Thus, a polymorphism refers to the occurrence of two or more genetically determined alternative sequences or alleles in a population. These differences can occur in coding (e.g., exonic) and non-coding (e.g., intronic or intergenic) portions of the genome, and can be manifested or detected as differences in polynucleic acid sequences, gene expression, including, for example transcription, processing, translation, transport, protein processing, trafficking, DNA synthesis; expressed proteins, other gene products or products of biochemical pathways or in post-translational modifications and any other differences manifested amongst members of a population. Polymorphisms that arise as the result of a single base change, such as single nucleotide polymorphisms (SNPs) or single nucleotide variations (SNVs), can include an insertion, deletion or change in one nucleotide. A polymorphic marker or site is the locus at which divergence occurs. Such sites can be as small as one base pair (an SNP or SNV). Polymorphic markers include, but are not limited to, restriction fragment length polymorphisms (RFLPs), variable number of tandem repeats (VNTRs), hypervariable regions, minisatellites, dinucleotide repeats, trinucleotide repeats, tetranucleotide repeats and other repeating patterns, simple sequence repeats and insertional elements, such as Alu. Polymorphic forms also are manifested as different mendelian alleles for a gene. Polymorphisms can be observed by differences in proteins, protein modifications, RNA expression modification, DNA and RNA methylation, regulatory factors that alter gene expression and DNA replication, and any other manifestation of alterations in genomic polynucleic acid or organelle polynucleic acids. Those skilled in the art can appreciate that polymorphisms are sometimes considered to be a subclass of variations, defined on the basis of a particular frequency cutoff in a population. For example, polymorphisms are considered to genetic variants/variations that occur at >1%, or >5%, frequency in the population.

[0078] These genetic variations can be found to be associated with one or more disorders and/or diseases using the methods disclosed herein. These genetic variations can be found to be associated with absence of one or more disorders and/or diseases (e.g. the one or more variants are protective against development of the disorder and/or diseases) using the methods disclosed herein.

[0079] These genetic variations can comprise point mutations, polymorphisms, single nucleotide polymorphisms (SNPs), single nucleotide variations (SNVs), translocations, insertions, deletions, amplifications, inversions, interstitial deletions, copy number variations (CNVs), structural variation (SV), loss of heterozygosity, or any combination thereof. As genetic variation includes any deletion, insertion or base substitution of the genomic DNA of one or more individuals in a first portion of a total population which thereby results in a difference at the site of the deletion, insertion or base substitution relative to one or more individuals in a second portion of the total population. Thus, the term "genetic variation" encompasses "wild type" or the most frequently occurring variation, and also includes "mutant," or the less frequently occurring variation. A wild type allele may be referred to as an ancestral allele.

[0080] As used herein, a target molecule that is "associated with" or "correlates with" a particular genetic variation is a molecule that can be functionally distinguished in its structure, activity, concentration, compartmentalization, degradation, secretion, and the like, as a result of such genetic variation. Polymorphisms (e.g., polymorphic markers, genetic variations, or genetic variants) can comprise any nucleotide position at which two or more sequences are possible in a subject population. Each version of a nucleotide sequence, with respect to the polymorphism/variation, can represent a specific allele of the polymorphism/variation. Genomic DNA from a subject can contain two

alleles for any given polymorphic marker, representative of each copy of the marker on each chromosome. An allele can be a nucleotide sequence of a given location on a chromosome. Polymorphisms/variations can comprise any number of specific alleles. A polymorphism/variation can be characterized by the presence of two or more alleles in a population. The polymorphism/variation can be characterized by the presence of three or more alleles. The polymorphism/variation can be characterized by four or more alleles, five or more alleles, six or more alleles, seven or more alleles, nine or more alleles, or ten or more alleles. An allele can be associated with one or more diseases or disorders, for example, a PML risk allele can be an allele that is associated with increased or decreased risk of developing PML. Genetic variations and alleles can be used to associate an inherited phenotype with a responsible genotype. A PML risk allele can be a variant allele that is statistically associated with a screening of PML. Genetic variations can be of any measurable frequency in the population, for example, a frequency higher than 10%, a frequency from 5-10%, a frequency from 1-5%, a frequency from 0.1-1%, or a frequency below 0.1%. As used herein, variant alleles can be alleles that differ from a reference allele. As used herein, a variant can be a segment of DNA that differs from the reference DNA, such as a genetic variation. Genetic variations can be used to track the inheritance of a gene that has not yet been identified, but whose approximate location is known.

[0081] As used herein, a "haplotype" can be information regarding the presence or absence of one or more genetic markers in a given chromosomal region in a subject. A haplotype can be a segment of DNA characterized by one or more alleles arranged along the segment, for example, a haplotype can comprise one member of the pair of alleles for each genetic variation or locus. The haplotype can comprise two or more alleles, three or more alleles, four or more alleles, five or more alleles, or any combination thereof, wherein, each allele can comprise one or more genetic variations along the segment.

[0082] A genetic variation can be a functional aberration that can alter gene function, gene expression, polypeptide expression, polypeptide function, or any combination thereof. A genetic variation can be a loss-of-function mutation, gain-of-function mutation, dominant negative mutation, or reversion. A genetic variation can be part of a gene's coding region or regulatory region. Regulatory regions can control gene expression and thus polypeptide expression. A regulatory region can be a segment of DNA wherein regulatory polypeptides, for example, transcription or splicing factors, can bind. A regulatory region can be positioned near the gene being regulated, for example, positions upstream or downstream of the gene being regulated. A regulatory region (e.g., enhancer element) can be several thousands of base pairs upstream or downstream of a gene.

[0083] Variants can include changes that affect a polypeptide, such as a change in expression level, sequence, function, localization, binding partners, or any combination thereof. A genetic variation can be a frameshift mutation, nonsense mutation, missense mutation, neutral mutation, or silent mutation. For example, sequence differences, when compared to a reference nucleotide sequence, can include the insertion or deletion of a single nucleotide, or of more than one nucleotide, resulting in a frame shift; the change of at least one nucleotide, resulting in a change in the encoded amino acid; the change of at least one nucleotide, resulting in the generation of a premature stop codon; the deletion of several nucleotides, resulting in a deletion of one or more amino acids encoded by the nucleotides; the insertion of one or several nucleotides, such as by unequal recombination or gene conversion, resulting in an interruption of the coding sequence of a reading frame; duplication of all or a part of a sequence; transposition; or a rearrangement of a nucleotide sequence. Such sequence changes can alter the polypeptide encoded by the nucleic acid, for example, if the change in the nucleic acid sequence causes a frame shift, the frame shift can result in a change in the encoded amino acids, and/or can result in the generation of a premature stop codon, causing generation of a truncated polypeptide. A genetic variation associated with PML can be a synonymous change in one or more nucleotides, for example, a change that does not result in a change in the amino acid sequence. Such a polymorphism can, for example, alter splice sites, affect the stability or transport of mRNA, or otherwise affect the transcription or translation of an encoded polypeptide. A synonymous mutation can result in the polypeptide product having an altered structure due to rare codon usage that impacts polypeptide folding during translation, which in some cases may alter its function and/or drug binding properties if it is a drug target. The changes that can alter DNA increase the possibility that structural changes, such as amplifications or deletions, occur at the somatic level. A polypeptide encoded by the reference nucleotide sequence can be a reference polypeptide with a particular reference amino acid sequence, and polypeptides encoded by variant nucleotide sequences can be variant polypeptides with variant amino acid sequences.

[0084] The most common sequence variants comprise base variations at a single base position in the genome, and such sequence variants, or polymorphisms, are commonly called single nucleotide polymorphisms (SNPs) or single nucleotide variants (SNVs). A SNP may represent a genetic variant present at greater than or equal to 1% occurrence in a population and a SNP or an SNV can represent a genetic variant present at any frequency level in a population. A SNP can be a nucleotide sequence variation occurring when a single nucleotide at a location in the genome differs between members of a species or between paired chromosomes in a subject. SNPs can include variants of a single nucleotide, for example, at a given nucleotide position, some subjects can have a 'G', while others can have a 'C'. SNPs can occur in a single mutational event, and therefore there can be two possible alleles possible at each SNP site; the original allele and the mutated allele. SNPs that are found to have two different bases in a single nucleotide position are referred to as biallelic SNPs, those with three are referred to as triallelic, and those with all four bases represented in the population are quadallelic. SNPs can be considered neutral. SNPs can affect susceptibility to a condition (e.g., PML). SNP polymorphisms can have two alleles, for example, a subject can be homozygous for one allele of the polymorphism wherein both chromosomal copies of the individual have the same nucleotide at the SNP location, or a subject can be heterozygous wherein the two sister chromosomes of the subject contain different nucleotides. The SNP nomenclature as reported herein is the official Reference SNP (rs) ID identification tag as assigned to each unique SNP by the National Center for Biotechnological Information (NCBI).

[0085] Another genetic variation can be copy number variations (CNVs). As used herein, "CNVs" include alterations of the DNA of a genome that results in an abnormal number of copies of one or more sections of DNA. A CNV can comprise a CNV-subregion. As used

herein, a "CNV-subregion" includes a continuous nucleotide sequence within a CNV. The nucleotide sequence of a CNV-subregion can be shorter than the nucleotide sequence of the CNV, and CNV-subregion can be equivalent to the CNV (e.g., such as for some CNVs). CNVs can be inherited or caused by de novo mutation and can be responsible for a substantial amount of human phenotypic variability, behavioral traits, and disease susceptibility. CNVs can be associated with susceptibility to one or more conditions, for example, PML. CNVs can include a single gene or include a contiguous set of genes. CNVs can be caused by structural rearrangements of the genome, for example, unbalanced translocations or inversions, insertions, deletions, amplifications, and interstitial deletions. These structural rearrangements occur on one or more chromosomes. Low copy repeats (LCRs), which are region-specific repeat sequences (also known as segmental duplications), can be susceptible to these structural rearrangements, resulting in CNVs. Factors such as size, orientation, percentage similarity and the distance between the copies can influence the susceptibility of LCRs to genomic rearrangement. In addition, rearrangements may be mediated by the presence of high copy number repeats, such as long interspersed elements (LINEs) and short interspersed elements (SINEs), often via non-homologous recombination. For example, chromosomal rearrangements can arise from non-allelic homologous recombination during meiosis or via a replication-based mechanism such as fork stalling and template switching (FoSTeS) (Zhang F. et al., Nat. Genet. (2009) or microhomology-mediated break-induced repair (MMBIR) (Hastings P. J. et al., PLoS Genetics (2009). CNVs can be referred to as structural variants, which are a broader class of variant that also includes copy number neutral alterations such as balanced inversions and balanced translocations.

[0086] CNVs can account for genetic variation affecting a substantial proportion of the human genome, for example, known CNVs can cover over 15% of the human genome sequence (Estivill and Armengol, PLoS Genetics (2007)). CNVs can affect gene expression, phenotypic variation and adaptation by disrupting or impairing gene dosage, and can cause disease, for example, microdeletion and microduplication disorders, and can confer susceptibility to diseases and disorders. Updated information about the location, type, and size of known CNVs can be found in one or more databases, for example, the Database of Genomic Variants (*See,* MacDonald JR et al., Nucleic Acids Res., 42, D986-92 (2014), which currently contains data for over 500,000 CNVs (as of May, 2016).

[0087] Other types of sequence variants can be found in the human genome and can be associated with a disease or disorder, including but not limited to, microsatellites. Microsatellite markers are stable, polymorphic, easily analyzed, and can occur regularly throughout the genome, making them especially suitable for genetic analysis. A polymorphic microsatellite can comprise multiple small repeats of bases, for example, CA repeats, at a particular site wherein the number of repeat lengths varies in a population. Microsatellites, for example, variable number of tandem repeats (VNTRs), can be short segments of DNA that have one or more repeated sequences, for example, about 2 to 5 nucleotides long, that can occur in non-coding DNA. Changes in microsatellites can occur during genetic recombination of sexual reproduction, increasing or decreasing the number of repeats found at an allele, or changing allele length.

[0088] The genetic variations disclosed herein can be associated with a risk of developing PML in a subject. In some cases, the subject can have a decreased risk due to the absence of one or more genetic variations that disrupt or modulate a corresponding gene according to Tables 1 to 26. For example, the subject can have a decreased risk due to the absence of one or more genetic variations that disrupt or modulate a corresponding gene according to Tables 3 and 6. In some cases, the subject can have an increased risk due to the presence of one or more genetic variations that disrupt or modulate a corresponding gene according to Tables 1 to 26. For example, the subject can have an increased risk due to the presence of one or more genetic variations that disrupt or modulate a corresponding gene according to Tables 3 and 6. In some cases, one or more genes listed in Tables 25A, 25B, and 26 can be removed from any one of the Tables 1-24. In some cases, one or more genes listed in Tables 25A, 25B, and 26 can be added to any one of the Tables 1-24.

Table 25A: Exemplary 8-gene panel

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
BAG3	AR	Public db	PMID: 19229298, 19282432, 22984599, 27042682	175
BTK	XLR	Public db	PMID: 18281276, 23765059, 25930993, 26029204	180
CD40LG	XLR	Public db	PMID: 17360404, 21455173, 23765059, 26008899, 26029204	206
DOCK8	AR	Public db	PMID: 23765059, 23887241, 26029204, 26454313	242
MAGT1	XLR	Public db	PMID: 23887241, 25504528, 27873163	326
RAG1	ADAR	Public db	PMID: 23122631, 23765059, 23887241, 25976673, 26029204,26454313,27484032,27808398	370
STAT1	ADAR	Public db	PMID: 23887241, 25645939, 26029204, 26513235, 26743090, 27821552, 27873163	436
WAS	XLR	Both	PMID: 12874226, 14647476, 19782549, 20008220, 24753205,26029204,26371186	483

Table 25B: Exemplary 16-gene panel

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
ADA	AR	Both	PMID: 23765059, 24135998, 25930993, 26029204, 26454313	1
BAG3	AR	Public db	PMID: 19229298, 19282432, 22984599, 27042682	175
BTK	XLR	Public db	PMID: 18281276, 23765059, 25930993, 26029204	180
CD40LG	XLR	Public db	PMID: 14647476, 17360404, 21455173, 23765059, 26008899,	206

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
			26029204	
DNMT3B	AR	Public db	PMID: 23486536, 23765059, 26029204, 26851945	240
DOCK8	AR	Public db	PMID: 23765059, 23887241, 26029204, 26454313	242
ITK	AR	Public db	PMID: 14647476, 23765059, 26029204, 26454313	308
LCK	AR	Public db	PMID: 14647476, 23765059, 26029204, 26454313	316
PNP	AR	Both	PMID: 26029204, 26454313	354
RAG1	ADAR	Public db	PMID: 23122631, 23765059, 23887241, 25976673, 26029204,26454313,27484032,27808398	370
STAT1	ADAR	Public db	PMID: 23887241, 25645939, 26029204, 26513235, 26743090,27821552,27873163	436
STAT3	AD	Public db	PMID: 23765059, 23887241, 25645939, 25930993, 26029204, 27658964, 27873163	438
STK3	unknow n	Both	PMID: 26029204	135
TYK2	AR	Public db	PMID: 26029204, 26513235, 27821552	144
WAS	XLR	Both	PMID: 12874226, 19782549, 20008220, 24753205, 26029204, 26371186	483
WIPF 1	AR	Public db	PMID: 23765059, 26029204, 26453379	485

Table 26: Exemplary 2-gene panel

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
ADA	intronic	100	adenosine deaminase	This gene encodes an enzyme that catalyzes the hydrolysis of adenosine to inosine. Various mutations have been described for this gene and have been linked to human diseases. Deficiency in this enzyme causes a form of severe combined immunodeficiency disease (SCID), in which there is dysfunction of both B and T lymphocytes with impaired cellular immunity and decreased production of immunoglobulins, whereas elevated levels of this enzyme have been associated with congenital hemolytic anemia. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC040226.1, X02994.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	1
STK3	intronic	6788	serine/threonine-protein kinase 3 isoform 1	This gene encodes a serine/threonine protein kinase activated by proapoptotic molecules indicating the encoded protein functions as a growth suppressor. Cleavage of the protein product by caspase removes the inhibitory C-terminal portion. The N-terminal portion is transported to the nucleus where it homodimerizes to form the active kinase which promotes the condensation of chromatin during apoptosis. Multiple transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Jan 2012]. Transcript Variant: This variant (1) encodes isoform 1. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: U26424.1, BC010640.2 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	135

Subjects

[0089] A "subject", as used herein, can be an individual of any age or sex from whom a sample containing polynucleotides is obtained for analysis by one or more methods described herein so as to obtain polynucleic acid information; for example, a male or female adult, child, newborn, or fetus. In some embodiments, a subject can be any target of therapeutic administration. In some embodiments, a subject can

be a test subject or a reference subject.

[0090] As used herein, a "cohort" can represent an ethnic group, a patient group, a particular age group, a group not associated with a particular condition (e.g., disease or disorder), a group associated with a particular condition (e.g., disease or disorder), a group of asymptomatic subjects, a group of symptomatic subjects, or a group or subgroup of subjects associated with a particular response to a treatment regimen or enrolled in a clinical trial. In some embodiments, a patient can be a subject afflicted with a condition (e.g., disease or disorder). In some embodiments, a patient can be a subject not afflicted with a condition (e.g., disease or disorder) and is considered apparently healthy, or a normal or control subject. In some embodiments, a subject can be a test subject, a patient or a candidate for a therapeutic, wherein genomic DNA from the subject, patient, or candidate is obtained for analysis by one or more methods of the present disclosure herein, so as to obtain genetic variation information of the subject, patient or candidate.

[0091] In some embodiments, the polynucleic acid sample can be obtained prenatally from a fetus or embryo or from the mother, for example, from fetal or embryonic cells in the maternal circulation. In some embodiments, the polynucleic acid sample can be obtained with the assistance of a health care provider, for example, to draw blood. In some embodiments, the polynucleic acid sample can be obtained without the assistance of a health care provider, for example, where the polynucleic acid sample is obtained non-invasively, such as a saliva sample, or a sample comprising buccal cells that is obtained using a buccal swab or brush, or a mouthwash sample.

[0092] Genetic variations may be assessed in subjects who are members of a target population. Such a target population maybe a population or group of subjects at risk of developing the condition (e.g., disease or disorder), based on, for example, other genetic factors, biomarkers, biophysical parameters, diagnostic testing such as magnetic resonance imaging (MRI), family history of the condition, previous screening or medical history, or any combination thereof.

[0093] The genetic variations of the present disclosure found to be associated with a condition (e.g., disease or disorder) can show similar association in other human populations. Particular embodiments comprising subject human populations are thus also contemplated and within the scope of the disclosure. Such embodiments relate to human subjects that are from one or more human populations including, but not limited to, Caucasian, Ashkenazi Jewish, Sephardi Jewish, European, American, Eurasian, Asian, Central/South Asian, East Asian, Middle Eastern, African, Hispanic, Caribbean, and Oceanic populations. European populations include, but are not limited to, Swedish, Norwegian, Finnish, Russian, Danish, Icelandic, Irish, Celt, English, Scottish, Dutch, Belgian, French, German, Spanish, Portuguese, Italian, Polish, Bulgarian, Slavic, Serbian, Bosnian, Czech, Greek and Turkish populations. The ethnic contribution in subjects can also be determined by genetic analysis, for example, genetic analysis of ancestry can be carried out using unlinked microsatellite markers or single nucleotide polymorphisms (SNPs) such as those set out in Smith *et al.,* (Smith M. W. et al., Am. J. Hum. Genet., 74:1001 (2004)).

[0094] Certain genetic variations can have different population frequencies in different populations, or are polymorphic in one population but not in another. The methods available and as thought herein can be applied to practice the present disclosure in any given human population. This can include assessment of genetic variations of the present disclosure, so as to identify those markers that give strongest association within the specific population. Thus, the at-risk variants of the present disclosure can reside on different haplotype background and in different frequencies in various human populations.

Conditions and Immunosuppressive Medicaments

[0095] In some embodiments, a subject can be diagnosed or undiagnosed with a condition (e.g., disease or disorder), can be asymptomatic or symptomatic, can have increased or decreased susceptibility to a condition (e.g., disease or disorder), can be currently under or previously under or not under a treatment for a condition (e.g., disease or disorder), or any combination thereof. In some embodiments, the condition can be AIDS, cancer, organ transplant, or an autoimmune disease. In some embodiments, the condition is PML.

[0096] In some embodiments, a subject can be diagnosed or undiagnosed with PML, can be asymptomatic or symptomatic, can have increased or decreased susceptibility to PML, can be currently under or previously under or not under a treatment for PML, or any combination thereof. In some embodiments, a subject can be diagnosed or undiagnosed with AIDS (e.g., individuals infected with HIV), can be asymptomatic or symptomatic, can have increased or decreased susceptibility to AIDS, can be currently under or previously under or not under a treatment for AIDS, or any combination thereof. In some embodiments, a subject can be diagnosed or undiagnosed with cancer (e.g., Hodgkin's disease, leukemia, lymphoma, or myelofibrosis), can be asymptomatic or symptomatic, can have increased or decreased susceptibility to cancer, can be currently under or previously under or not under a treatment for cancer, or any combination thereof. In some embodiments, a subject can be currently diagnosed or previously diagnosed or undiagnosed with an autoimmune disease (e.g., multiple sclerosis, rheumatoid arthritis, psoriasis, systemic lupus erythematosus), can be asymptomatic or symptomatic, can have increased or decreased susceptibility to an autoimmune disease, can be currently under or previously under or not under a treatment for an autoimmune disease, or any combination thereof.

[0097] The term "cancer" is meant to include all types of cancerous growths or oncogenic processes, metastatic tissues or malignantly transformed cells, tissues, or organs, irrespective of histopathologic type or stage of invasiveness. A metastatic tumor can arise from a multitude of primary tumor types, including but not limited to those of breast, lung, liver, colon and ovarian origin. Examples of cancers include, but are not limited to, a fibrosarcoma, myosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma,

angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, gastric cancer, esophageal cancer, rectal cancer, pancreatic cancer, ovarian cancer, prostate cancer, uterine cancer, cancer of the head and neck, skin cancer, brain cancer, squamous cell carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinoma, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilm's tumor, cervical cancer, testicular cancer, small cell lung carcinoma, non-small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, meningioma, melanoma, neuroblastoma, retinoblastoma, leukemia, lymphoma, myelofibrosis, or Kaposi sarcoma.

[0098] The term "autoimmune disease" is meant to include all types of pathological states arising from abnormal immune responses of the body to substances and tissues that are normally present in the body. Examples of autoimmune diseases include, but are not limited to, Addison disease, Anti-NMDA receptor encephalitis, antisynthetase syndrome, Aplastic anemia, autoimmune anemias, Autoimmune hemolytic anemia, Autoimmune pancreatitis, Behcet's Disease, bullous skin disorders, Celiac disease - sprue (gluten-sensitive enteropathy), chronic fatigue syndrome, Chronic inflammatory demyelinating polyneuropathy, chronic lymphocytic leukemia, Crohn's disease, Dermatomyositis, Devic's disease, Erythroblastopenia, Evans syndrome, Focal segmental glomerulosclerosis, Granulomatosis with polyangiitis, Graves disease, Graves' ophthalmopathy, Guillain-Barre syndrome, Hashimoto thyroiditis, idiopathic thrombocytopenic purpura (ITP), IgA nephropathy, IgA-mediated autoimmune diseases, IgG4-related disease, Inflammatory bowel disease, Juvenile idiopathic arthritis, Multiple sclerosis, Myasthenia gravis, myeloma, non-Hodgkin's lymphoma, Opsoclonus myoclonus syndrome (OMS), Pemphigoid, Pemphigus, pemphigus vulgaris, Pernicious anemia, polymyositis, Psoriasis, pure red cell aplasia, Reactive arthritis, Rheumatoid arthritis, Sarcoidosis, scleroderma, Sjögren syndrome, Systemic lupus erythematosus, Thrombocytopenic purpura, Thrombotic thrombocytopenic purpura, Type I diabetes, Ulcerative colitis, Vasculitis (e.g., vasculitis associated with anti-neutrophil cytoplasmic antibody) and Vitiligo.

[0099] In some embodiments, a subject can be currently treated with an immunosuppressive medicament. In some embodiments, a subject can be previously treated with an immunosuppressive medicament. In some embodiments, a subject can be not yet treated with an immunosuppressive medicament. The immunosuppressive medicament can include but not limited to glucocorticoids, cytostatics, antibodies, drugs acting on immunophilins, interferons, opioids, TNF binding proteins, mycophenolate, or other small biological agents. For example, glucocorticoids can include but not limited to cortisol (hydrocortisone), cortisone, prednisone, prednisolone, methylprednisolone, dexamethasone, betamethasone, triamcinolone, beclometasone, fludrocortisone acetate, deoxycorticosterone acetate (DOCA), or aldosterone. Cytostatics can include but not limited to nitrogen mustards (e.g., cyclophosphamide), nitrosoureas, platinum compounds, folic acid analogues such as methotrexate, purine analogues such as azathioprine and mercaptopurine, pyrimidine analogues such as fluorouracil, protein synthesis inhibitors, cytotoxic antibiotics such as dactinomycin, anthracyclines, mitomycin C, bleomycin, or mithramycin. Antibodies can include but not limited to polyclonal antibodies such as atgam and thymoglobuline, monoclonal antibodies such as CD25- and CD3-directed antibodies, muromonab-CD3, basiliximab (e.g., SIMULECT), and daclizumab (e.g., ZENAPAX). Drugs acting on immunophilins can include but not limited to ciclosporin, tacrolimus, sirolimus, or everolimus. TNF binding proteins can include but not limited to infliximab (e.g., REMICADE), etanercept (e.g., ENBREL), or adalimumab (e.g., HUMIRA). Other small biological agents can include but not limited to fingolimod, myriocin, and rituximab (e.g., RITUXAN).

[0100] In some embodiments, the immunosuppressive medicament can be drugs for treating multiple sclerosis include but not limited to interferon beta-1a (e.g., AVONEX, REBIF), interferon beta-1b (e.g., BETASERON, EXTAVIA), glatiramer acetate (e.g., COPAXONE, GLATOPA), peginterferon beta-1a (e.g., PLEGRIDY), teriflunomide (e.g., AUBAGIO), fingolimod (e.g., GILENYA), dimethyl fumarate (e.g., TECFIDERA), alemtuzumab (e.g., LEMTRADA), mitoxantrone (e.g., NOVANTRONE), natalizumab (e.g., TYSABRI), daclizumab (e.g., ZINBRYTA), or ocrelizumab (e.g., OCREVUS).

[0101] In some embodiments, the immunosuppressive medicament can be adalimumab (e.g., HUMIRA), alemtuzumab (e.g., LEMTRADA), alemtuzumab (e.g., CAMPATH), azathioprine (e.g., IMURAN), belimumab (e.g., BENLYSTA), bevacizumab (e.g., AVASTIN), bortezomib (e.g., VELCADE), eculizumab (e.g., SOLIRIS), leflunomide, brentuximab vedotin (e.g., ADCETRIS), cetuximab (e.g., ERBITUX), cyclophosphamid, dimethyl fumarate (e.g., TECFIDERA), efalizumab (e.g., RAPTIVA), fingolimod (e.g., GILENYA), fludarabine (e.g., FLUDARA), fumaric acid, imatinib (e.g., GLEEVEC, GLIVEC), infliximab (e.g., REMICADE), methotrexate (e.g., TREXALL, RHEUMATREX), mycophenolate mofetil (e.g., CELLCEPT), natalizumab (e.g., TYSABRI), daclizumab (e.g., ZINBRYTA), rituximab (e.g., RITUXAN), vedolizumab (e.g., ENTYVIO), ruxolitinib (e.g., JAKAFI, JAKAVI), or ocrelizumab (e.g., Ocrevus). For example, rituximab can be used to treat MS patients (e.g., off-label), both relapsing-remitting (RRMS) and progressive (PMS) forms; for instance, as reported by Memon A et al. 2018 (PMID 29309416), Alcala C et al. 2018 (PMID 29785523), and Berntsson S et al. 2018 (PMID 29797711).

Samples

[0102] Samples that are suitable for use in the methods described herein can be polynucleic acid samples from a subject. A "polynucleic acid sample" as used herein can include RNA or DNA, or a combination thereof. In another embodiment, a "polypeptide sample" (e.g., peptides or proteins, or fragments therefrom) can be used to ascertain information that an amino acid change has occurred, which is the result of a genetic variant. Polynucleic acids and polypeptides can be extracted from one or more samples including but not limited to, blood, saliva, urine, mucosal scrapings of the lining of the mouth, expectorant, serum, tears, skin, tissue, or hair. A polynucleic acid sample can be assayed for polynucleic acid information. "Polynucleic acid information," as used herein, includes a polynucleic acid

sequence itself, the presence/absence of genetic variation in the polynucleic acid sequence, a physical property which varies depending on the polynucleic acid sequence (e.g., Tm), and the amount of the polynucleic acid (e.g., number of mRNA copies). A "polynucleic acid" means any one of DNA, RNA, DNA including artificial nucleotides, or RNA including artificial nucleotides. As used herein, a "purified polynucleic acid" includes cDNAs, fragments of genomic polynucleic acids, polynucleic acids produced using the polymerase chain reaction (PCR), polynucleic acids formed by restriction enzyme treatment of genomic polynucleic acids, recombinant polynucleic acids, and chemically synthesized polynucleic acid molecules. A "recombinant" polynucleic acid molecule includes a polynucleic acid molecule made by an artificial combination of two otherwise separated segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of polynucleic acids by genetic engineering techniques. As used herein, a "polypeptide" includes proteins, fragments of proteins, and peptides, whether isolated from natural sources, produced by recombinant techniques, or chemically synthesized. A polypeptide may have one or more modifications, such as a post-translational modification (e.g., glycosylation, phosphorylation, etc.) or any other modification (e.g., pegylation, etc.). The polypeptide may contain one or more non-naturally-occurring amino acids (e.g., such as an amino acid with a side chain modification).

[0103] In some embodiments, the polynucleic acid sample can comprise cells or tissue, for example, cell lines. Exemplary cell types from which nucleic acids can be obtained using the methods described herein include, but are not limited to, the following: a blood cell such as a B lymphocyte, T lymphocyte, leukocyte, erythrocyte, macrophage, or neutrophil; a muscle cell such as a skeletal cell, smooth muscle cell or cardiac muscle cell; a germ cell, such as a sperm or egg; an epithelial cell; a connective tissue cell, such as an adipocyte, chondrocyte; fibroblast or osteoblast; a neuron; an astrocyte; a stromal cell; an organ specific cell, such as a kidney cell, pancreatic cell, liver cell, or a keratinocyte; a stem cell; or any cell that develops therefrom. A cell from which nucleic acids can be obtained can be a blood cell or a particular type of blood cell including, for example, a hematopoietic stem cell or a cell that arises from a hematopoietic stem cell such as a red blood cell, B lymphocyte, T lymphocyte, natural killer cell, neutrophil, basophil, eosinophil, monocyte, macrophage, or platelet. Generally, any type of stem cell can be used including, without limitation, an embryonic stem cell, adult stem cell, or pluripotent stem cell.

[0104] In some embodiments, a polynucleic acid sample can be processed for RNA or DNA isolation, for example, RNA or DNA in a cell or tissue sample can be separated from other components of the polynucleic acid sample. Cells can be harvested from a polynucleic acid sample using standard techniques, for example, by centrifuging a cell sample and resuspending the pelleted cells, for example, in a buffered solution, for example, phosphate-buffered saline (PBS). In some embodiments, after centrifuging the cell suspension to obtain a cell pellet, the cells can be lysed to extract DNA. In some embodiments, the nucleic acid sample can be concentrated and/or purified to isolate DNA. All nucleic acid samples obtained from a subject, including those subjected to any sort of further processing, are considered to be obtained from the subject. In some embodiments, standard techniques and kits known in the art can be used to extract RNA or DNA from a nucleic acid sample, including, for example, phenol extraction, a QIAAMP® Tissue Kit (Qiagen, Chatsworth, Calif.), a WIZARD® Genomic DNA purification kit (Promega), or a Qiagen Autopure method using Puregene chemistry, which can enable purification of highly stable DNA well-suited for archiving.

[0105] In some embodiments, determining the identity of an allele or determining copy number can, but need not, include obtaining a polynucleic acid sample comprising RNA and/or DNA from a subject, and/or assessing the identity, copy number, presence or absence of one or more genetic variations and their chromosomal locations within the genomic DNA (e.g. subject's genome) derived from the polynucleic acid sample.

[0106] The individual or organization that performs the determination need not actually carry out the physical analysis of a nucleic acid sample from a subject. In some embodiments, the methods can include using information obtained by analysis of the polynucleic acid sample by a third party. In some embodiments, the methods can include steps that occur at more than one site. For example, a polynucleic acid sample can be obtained from a subject at a first site, such as at a health care provider or at the subject's home in the case of a self-testing kit. The polynucleic acid sample can be analyzed at the same or a second site, for example, at a laboratory or other testing facility.

Nucleic Acids

[0107] The nucleic acids and polypeptides described herein can be used in methods and kits of the present disclosure. In some embodiments, aptamers that specifically bind the nucleic acids and polypeptides described herein can be used in methods and kits of the present disclosure. As used herein, a nucleic acid can comprise a deoxyribonucleotide (DNA) or ribonucleotide (RNA), whether singular or in polymers, naturally occurring or non-naturally occurring, double-stranded or single-stranded, coding, for example a translated gene, or non-coding, for example a regulatory region, or any fragments, derivatives, mimetics or complements thereof. In some embodiments, nucleic acids can comprise oligonucleotides, nucleotides, polynucleotides, nucleic acid sequences, genomic sequences, complementary DNA (cDNA), antisense nucleic acids, DNA regions, probes, primers, genes, regulatory regions, introns, exons, open-reading frames, binding sites, target nucleic acids and allele-specific nucleic acids.

[0108] A "probe," as used herein, includes a nucleic acid fragment for examining a nucleic acid in a specimen using the hybridization reaction based on the complementarity of nucleic acid.

[0109] A "hybrid" as used herein, includes a double strand formed between any one of the abovementioned nucleic acid, within the same

type, or across different types, including DNA-DNA, DNA-RNA, RNA-RNA or the like.

[0110] "Isolated" nucleic acids, as used herein, are separated from nucleic acids that normally flank the gene or nucleotide sequence (as in genomic sequences) and/or has been completely or partially purified from other transcribed sequences (e.g., as in an RNA library). For example, isolated nucleic acids of the disclosure can be substantially isolated with respect to the complex cellular milieu in which it naturally occurs, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized. In some instances, the isolated material can form part of a composition, for example, a crude extract containing other substances, buffer system or reagent mix. In some embodiments, the material can be purified to essential homogeneity using methods known in the art, for example, by polyacrylamide gel electrophoresis (PAGE) or column chromatography (e.g., HPLC). With regard to genomic DNA (gDNA), the term "isolated" also can refer to nucleic acids that are separated from the chromosome with which the genomic DNA is naturally associated. For example, the isolated nucleic acid molecule can contain less than about 250 kb, 200 kb, 150 kb, 100 kb, 75 kb, 50 kb, 25 kb, 10 kb, 5 kb, 4 kb, 3 kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of the nucleotides that flank the nucleic acid molecule in the gDNA of the cell from which the nucleic acid molecule is derived.

[0111] Nucleic acids can be fused to other coding or regulatory sequences can be considered isolated. For example, recombinant DNA contained in a vector is included in the definition of "isolated" as used herein. In some embodiments, isolated nucleic acids can include recombinant DNA molecules in heterologous host cells or heterologous organisms, as well as partially or substantially purified DNA molecules in solution. Isolated nucleic acids also encompass in vivo and in vitro RNA transcripts of the DNA molecules of the present disclosure. An isolated nucleic acid molecule or nucleotide sequence can be synthesized chemically or by recombinant means. Such isolated nucleotide sequences can be useful, for example, in the manufacture of the encoded polypeptide, as probes for isolating homologous sequences (e.g., from other mammalian species), for gene mapping (e.g., by in situ hybridization with chromosomes), or for detecting expression of the gene, in tissue (e.g., human tissue), such as by Northern blot analysis or other hybridization techniques disclosed herein. The disclosure also pertains to nucleic acid sequences that hybridize under high stringency hybridization conditions, such as for selective hybridization, to a nucleotide sequence described herein Such nucleic acid sequences can be detected and/or isolated by allele- or sequence-specific hybridization (e.g., under high stringency conditions). Stringency conditions and methods for nucleic acid hybridizations are well known to the skilled person (see, e.g., Current Protocols in Molecular Biology, Ausubel, F. et al., John Wiley & Sons, (1998), and Kraus, M. and Aaronson, S., Methods Enzymol., 200:546-556 (1991).

[0112] Calculations of "identity" or "percent identity" between two or more nucleotide or amino acid sequences can be determined by aligning the sequences for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first sequence). The nucleotides at corresponding positions are then compared, and the percent identity between the two sequences is a function of the number of identical positions shared by the sequences (e.g. % identity = # of identical positions/total # of positions x 100). For example, a position in the first sequence is occupied by the same nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

[0113] In some embodiments, the length of a sequence aligned for comparison purposes is at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%, of the length of the reference sequence. The actual comparison of the two sequences can be accomplished by well-known methods, for example, using a mathematical algorithm. A non-limiting example of such a mathematical algorithm is described in Karlin, S. and Altschul, S., Proc. Natl. Acad. Sci. USA, 90- 5873-5877 (1993). Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0), as described in Altschul, S. et al., Nucleic Acids Res., 25:3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, any relevant parameters of the respective programs (e.g., NBLAST) can be used. For example, parameters for sequence comparison can be set at score= 100, word length= 12, or can be varied (e.g., W=5 or W=20). Other examples include the algorithm of Myers and Miller, CABIOS (1989), ADVANCE, ADAM, BLAT, and FASTA. In some embodiments, the percent identity between two amino acid sequences can be accomplished using, for example, the GAP program in the GCG software package (Accelrys, Cambridge, UK).

[0114] "Probes" or "primers" can be oligonucleotides that hybridize in a base-specific manner to a complementary strand of a nucleic acid molecule. Probes can include primers, which can be a single-stranded oligonucleotide probe that can act as a point of initiation of template-directed DNA synthesis using methods including but not limited to, polymerase chain reaction (PCR) and ligase chain reaction (LCR) for amplification of a target sequence. Oligonucleotides, as described herein, can include segments or fragments of nucleic acid sequences, or their complements. In some embodiments, DNA segments can be between 5 and 10,000 contiguous bases, and can range from 5, 10, 12, 15, 20, or 25 nucleotides to 10, 15, 20, 25, 30, 40, 50, 100, 200, 500, 1000 or 10,000 nucleotides. In addition to DNA and RNA, probes and primers can include polypeptide nucleic acids (PNA), as described in Nielsen, P. et al., Science 254: 1497-1500 (1991). A probe or primer can comprise a region of nucleotide sequence that hybridizes to at least about 15, typically about 20-25, and in certain embodiments about 40, 50, 60 or 75, consecutive nucleotides of a nucleic acid molecule.

[0115] The present disclosure also provides isolated nucleic acids, for example, probes or primers, that contain a fragment or portion that can selectively hybridize to a nucleic acid that comprises, or consists of, a nucleotide sequence, wherein the nucleotide sequence can comprise at least one polymorphism or polymorphic allele contained in the genetic variations described herein or the wild-type nucleotide that is located at the same position, or the complements thereof. In some embodiments, the probe or primer can be at least 70% identical, at least 80% identical, at least 85% identical, at least 90% identical, or at least 95% identical, to the contiguous nucleotide sequence or to the complement of the contiguous nucleotide sequence.

[0116] In some embodiments, a nucleic acid probe can be an oligonucleotide capable of hybridizing with a complementary region of a gene associated with a condition (e.g., PML) containing a genetic variation described herein. The nucleic acid fragments of the disclosure can be used as probes or primers in assays such as those described herein.

[0117] The nucleic acids of the disclosure, such as those described above, can be identified and isolated using standard molecular biology techniques well known to the skilled person. In some embodiments, DNA can be amplified and/or can be labeled (e.g., radiolabeled, fluorescently labeled) and used as a probe for screening, for example, a cDNA library derived from an organism. cDNA can be derived from mRNA and can be contained in a suitable vector. For example, corresponding clones can be isolated, DNA obtained fallowing in vivo excision, and the cloned insert can be sequenced in either or both orientations by art-recognized methods to identify the correct reading frame encoding a polypeptide of the appropriate molecular weight. Using these or similar methods, the polypeptide and the DNA encoding the polypeptide can be isolated, sequenced and further characterized.

[0118] In some embodiments, nucleic acid can comprise one or more polymorphisms, variations, or mutations, for example, single nucleotide polymorphisms (SNPs), single nucleotide variations (SNVs), copy number variations (CNVs), for example, insertions, deletions, inversions, and translocations. In some embodiments, nucleic acids can comprise analogs, for example, phosphorothioates, phosphoramidates, methyl phosphonate, chiralmethyl phosphonates, 2-0-methyl ribonucleotides, or modified nucleic acids, for example, modified backbone residues or linkages, or nucleic acids combined with carbohydrates, lipids, polypeptide or other materials, or peptide nucleic acids (PNAs), for example, chromatin, ribosomes, and transcriptosomes. In some embodiments nucleic acids can comprise nucleic acids in various structures, for example, A DNA, B DNA, Z-form DNA, siRNA, tRNA, and ribozymes. In some embodiments, the nucleic acid may be naturally or non-naturally polymorphic, for example, having one or more sequence differences, for example, additions, deletions and/or substitutions, as compared to a reference sequence. In some embodiments, a reference sequence can be based on publicly available information, for example, the U.C. Santa Cruz Human Genome Browser Gateway (genome.ucsc.edu/cgi-bin/hgGateway) or the NCBI website (www.ncbi.nlm.nih.gov). In some embodiments, a reference sequence can be determined by a practitioner of the present disclosure using methods well known in the art, for example, by sequencing a reference nucleic acid.

[0119] In some embodiments, a probe can hybridize to an allele, SNP, SNV, or CNV as described herein. In some embodiments, the probe can bind to another marker sequence associated with PML as described herein.

[0120] One of skill in the art would know how to design a probe so that sequence specific hybridization can occur only if a particular allele is present in a genomic sequence from a test nucleic acid sample. The disclosure can also be reduced to practice using any convenient genotyping method, including commercially available technologies and methods for genotyping particular genetic variations

[0121] Control probes can also be used, for example, a probe that binds a less variable sequence, for example, a repetitive DNA associated with a centromere of a chromosome, can be used as a control. In some embodiments, probes can be obtained from commercial sources. In some embodiments, probes can be synthesized, for example, chemically or in vitro, or made from chromosomal or genomic DNA through standard techniques. In some embodiments sources of DNA that can be used include genomic DNA, cloned DNA sequences, somatic cell hybrids that contain one, or a part of one, human chromosome along with the normal chromosome complement of the host, and chromosomes purified by flow cytometry or microdissection. The region of interest can be isolated through cloning, or by site-specific amplification using PCR.

[0122] One or more nucleic acids for example, a probe or primer, can also be labeled, for example, by direct labeling, to comprise a detectable label. A detectable label can comprise any label capable of detection by a physical, chemical, or a biological process for example, a radioactive label, such as ^{32}P or ^{3}H, a fluorescent label, such as FITC, a chromophore label, an affinity-ligand label, an enzyme label, such as alkaline phosphatase, horseradish peroxidase, or I2 galactosidase, an enzyme cofactor label, a hapten conjugate label, such as digoxigenin or dinitrophenyl, a Raman signal generating label, a magnetic label, a spin label, an epitope label, such as the FLAG or HA epitope, a luminescent label, a heavy atom label, a nanoparticle label, an electrochemical label, a light scattering label, a spherical shell label, semiconductor nanocrystal label, such as quantum dots (described in U.S. Pat. No. 6,207,392), and probes labeled with any other signal generating label known to those of skill in the art, wherein a label can allow the probe to be visualized with or without a secondary detection molecule. A nucleotide can be directly incorporated into a probe with standard techniques, for example, nick translation, random priming, and PCR labeling. A "signal," as used herein, include a signal suitably detectable and measurable by appropriate means, including fluorescence, radioactivity, chemiluminescence, and the like.

[0123] Non-limiting examples of label moieties useful for detection include, without limitation, suitable enzymes such as horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; members of a binding pair that are capable of forming complexes such as streptavidin/biotin, avidin/biotin or an antigen/antibody complex including, for example, rabbit IgG and anti-rabbit IgG; fluorophores such as umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, tetramethyl rhodamine, eosin, green fluorescent protein, erythrosin, coumarin, methyl coumarin, pyrene, malachite green, stilbene, lucifer yellow, Cascade Blue, Texas Red, dichlorotriazinylamine fluorescein, dansyl chloride, phycoerythrin, fluorescent lanthanide complexes such as those including Europium and Terbium, cyanine dye family members, such as Cy3 and Cy5, molecular beacons and fluorescent derivatives thereof, as well as others known in the art as described, for example, in Principles of Fluorescence Spectroscopy, Joseph R. Lakowicz (Editor), Plenum Pub Corp, 2nd edition (July 1999) and the 6th Edition of the Molecular Probes Handbook by Richard P. Hoagland; a luminescent material such as luminol; light scattering or plasmon resonant materials such as gold or silver particles or quantum dots; or radioactive material include ^{14}C, ^{121}I, ^{124}I, ^{125}I, Tc99m, ^{32}P, ^{33}P, ^{35}S or ^{3}H.

[0124] Other labels can also be used in the methods of the present disclosure, for example, backbone labels. Backbone labels comprise nucleic acid stains that bind nucleic acids in a sequence independent manner. Non-limiting examples include intercalating dyes such as phenanthridines and acridines (e.g., ethidium bromide, propidium iodide, hexidium iodide, dihydroethidium, ethidium homodimer-1 and -2, ethidium monoazide, and ACMA); some minor grove binders such as indoles and imidazoles (e.g., Hoechst 33258, Hoechst 33342, Hoechst 34580 and DAPI); and miscellaneous nucleic acid stains such as acridine orange (also capable of intercalating), 7-AAD, actinomycin D, LDS751, and hydroxystilbamidine. All of the aforementioned nucleic acid stains are commercially available from suppliers such as Molecular Probes, Inc. Still other examples of nucleic acid stains include the following dyes from Molecular Probes: cyanine dyes such as 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, SYBRDX, SYTO-40, -41, -42, -43, -44, -45 (blue), SYTO-13, -16, -24, -21, -23, -12, -11, -20, -22, - 15, -14, -25 (green), SYTO-81, -80, -82, -83, -84, -85 (orange), SYTO-64, -17, -59, -61, -62, -60, -63 (red).

[0125] In some embodiments, fluorophores of different colors can be chosen, for example, 7-amino-4-methylcoumarin-3-acetic acid (AMCA), 5-(and-6)-carboxy-X-rhodamine, lissamine rhodamine B, 5-(and-6)-carboxyfluorescein, fluorescein-5-isothiocyanate (FITC), 7-diethylaminocoumarin-3-carboxylic acid, tetramethylrhodamine-5-(and-6)-isothiocyanate, 5-(and-6)-carboxytetramethylrhodamine, 7-hydroxycoumarin-3-carboxylic acid, 6-[fluorescein 5-(and-6)-carboxamido]hexanoic acid, N-(4,4-difluoro-5,7-dimethyl-4-bora-3a,4a diaza-3-indacenepropionic acid, eosin-5-isothiocyanate, erythrosin-5-isothiocyanate, TRITC, rhodamine, tetramethylrhodamine, R-phycoerythrin, Cy-3, Cy-5, Cy-7, Texas Red, Phar-Red, allophycocyanin (APC),and CASCADETM blue acetylazide, such that each probe in or not in a set can be distinctly visualized. In some embodiments, fluorescently labeled probes can be viewed with a fluorescence microscope and an appropriate filter for each fluorophore, or by using dual or triple band-pass filter sets to observe multiple fluorophores. In some embodiments, techniques such as flow cytometry can be used to examine the hybridization pattern of the probes.

[0126] In other embodiments, the probes can be indirectly labeled, for example, with biotin or digoxygenin, or labeled with radioactive isotopes such as ^{32}P and/or ^{3}H. As a non-limiting example, a probe indirectly labeled with biotin can be detected by avidin conjugated to a detectable marker. For example, avidin can be conjugated to an enzymatic marker such as alkaline phosphatase or horseradish peroxidase. In some embodiments, enzymatic markers can be detected using colorimetric reactions using a substrate and/or a catalyst for the enzyme. In some embodiments, catalysts for alkaline phosphatase can be used, for example, 5-bromo-4-chloro-3-indolylphosphate and nitro blue tetrazolium. In some embodiments, a catalyst can be used for horseradish peroxidase, for example, diaminobenzoate.

[0127] One or more genes disclosed herein can be in conditions or molecular pathways related to various aspects of immune function including, but not limited to, Type I interferon response (*e.g.*, PMID 26052098), B cell receptor pathway (*e.g.*, Wikipathways WP23; PMID 22566564), RANKL/RANK signaling pathway (*e.g.,* Wikipathways WP2018), TCR signaling pathway (*e.g.*, Wikipathways WP69), NF-kB signaling (*e.g.,* PMID 28362430), JAK-STAT pathway (*e.g.,* PMID 28255960), post-translational modification biology such as ubiquitination via LUBAC (*e.g.*, PMID 23104095, 24958845, 25086647, 26085218, 26111062, 26525107, 26848516, 26877205, 27178468, 27786304, 27892465), Aicardi-Goutieres syndrome (*e.g.,* PMID 26052098), eosinophilia (*e.g.,* PMID 27222657), congenital neutropenia (*e.g.,* PMID 24753205), T cell receptor defects (*e.g.,* PMID 25452106, 25636200, 26246585, 26379669, 26453379, 28400082), and autophagy defects (*e.g.*, 19229298, 22984599, 23222957, 26917586, 26953272, 27588602). In some embodiments, one or more genes disclosed herein can be related to JC virus biology (*e.g.*, PMID 15327898, 19282432, 19903823, 22984599, 25910481). In some embodiments, one or more genes disclosed herein can be antibiral immune response genes.

Table 27: Exemplary pathways and biology for PML risk genes (96-gene panel)

Genes	Autoinflammatory diseases	Autophagy defects	B cell defects	B cell receptor pathway	Deubiquitinase pathway	Eosinophilia-associated immunodeficiency	JC virus biology	Osteopetrosis	PI3K signaling	RANKL/RANK pathway	T cell defects	T cell receptor pathway	TLR signaling	Type I interferon pathway
AP3B1								x						
APOL1							x							
ASH1L														
ATM			x			x	x							x
ATR							x							x
BLM														
CARD11			x	x		x						x	x	
CDKN1B														
CHD7						x								
CLCN7								x						
DCLRE1C														
DDX58														x
DOCK8			x			x				x				
EGR1														
EPG5		x												
ETF1														

FPR2														
GATA2			x											
GFI1								x						
HIVEP1														
HIVEP2														
HTR2A							x							
IDO2														
IFIH1														x
IFNGR2														
IFNLR1														
IGLL1			x	x										
IKBKB	x		x	x	x			x		x		x	x	x
IL17F														
IL1B					x		x					x	x	x
IL21R			x											
IRAK4			x	x	x								x	x
ITSN2														
JUN				x			x							x
KAT6B														
KCTD7														
LIG4						x								
LRBA		x	x								x			
MALL														
MAPK3							x			x		x	x	
MAVS														x
MCEE														
MKL1														x
MYD88			x		x								x	x
NBN														
NFKB1	x		x	x	x			x	x	x		x	x	x
NOD2	x	x			x	x								x
NRIP1														
PIAS1														x
PIAS2														x
PIK3CD									x					x
PIK3CD-AS1									x					
PIK3R1			x	x					x	x		x		x
PKHD1														
PLCG2			x	x					x	x	x	x	x	
PNP1														

POLA1														x
POLE			x											
PRF1														
PRKCB				x										
PRKCD				x								x		x
PRKCH														
PRKDC														x
PSTPIP1	x													
PTEN									x					
PTPRC				x					x					x
RABGEF1														
RAD51							x							
RAG1			x			x			x					
RAG2			x			x			x					
RIPK1	x				x								x	x
RIPK3					x									x
RNF168														
RTEL1														
SHARPIN	x		x		x				x				x	x
SKIV2L														x
SMAD4							x							
STIM1			x											x
STIM2														
STXBP2														
TAP2														
TBK1													x	x
TCIRG1									x					
TICAM1					x								x	x
TLR3	x				x								x	x
TLR4	x				x								x	x
TNFRSF11A	x				x			x		x			x	x
TNFRSF13B			x										x	
TNFRSF8														
TP53														

TRAF3								x		x			x	x
TRAFD1								x					x	x
TRPM2														
VPS45														
WEE1							x							
ZAP70			x			x			x			x		

[0128] Table 27 contains a set of exemplary pathways and biology for PML risk genes based on the 96-gene panel listed in **Table 19.** The genes disclosed herein, such as the genes in the 96-gene panel, can be grouped based on the pathway or biological processes they are involved in.

Methods of Screening

[0129] The present invention involves determining the absence or presence of a genomic variation selected from chr9:137779251, G>A and chr1:160769595, AG>A. The following provides a general explanation of screening methods that may be used to determine the presence or absence of the genomic variations of the claims, and includes examples not forming subject matter being claimed herein, insofar as they relate to variations other than chr9: 137779251, G>A and chr1:160769595, AG>A. As used herein, screening a subject comprises diagnosing or determining, theranosing, or determining the susceptibility to developing (prognosing) a condition, for example, PML. The presence of, or a susceptibility to, PML, may be determined by detecting at least one genetic variation in a sample from a subject as described herein. Detection of particular alleles, markers, variations, or haplotypes may be indicative of a presence or susceptibility to a condition (e.g., PML).

[0130] While means for screening PML using a JCV antibody test exist, PML risk is not adequately assessed by the JCV antibody test alone. Thus there exists a need for an improved screening test for assessing the risk of developing PML. Described herein are methods of screening an individual for a risk of developing PML, including but not limited to, determining the identity and location of genetic variations, such as variations in nucleotide sequence and copy number, and the presence or absence of alleles or genotypes in one or more samples from one or more subjects using any of the methods described herein. Determining an association to having or developing PML can be performed by detecting particular variations that appear more frequently in test subjects compared to reference subjects and analyzing the molecular and physiological pathways these variations can affect.

[0131] Within any given population, there can be an absolute susceptibility of developing a disease or trait, defined as the chance of a person developing the specific disease or trait over a specified time-period. Susceptibility (e.g., being at-risk) is typically measured by looking at very large numbers of people, rather than at a particular individual. As described herein, certain copy number variations (genetic variations) and/or single nucleotide variations are found to be useful for susceptibility assessment of PML. Susceptibility assessment can involve detecting particular genetic variations in the genome of individuals undergoing assessment. Particular genetic variations are found more frequently in individuals with PML, than in individuals without PML. Therefore, these genetic variations have predictive value for detecting PML, or a susceptibility to PML, in an individual. Without intending to be limited by theory, it is believed that the genetic variations described herein to be associated with susceptibility of PML represent functional variants predisposing to the disease. A genetic variation can confer a susceptibility of the condition, for example carriers of the genetic variation are at a different risk of the condition than non-carriers. The presence of such a genetic variation may be indicative of increased susceptibility to PML.

[0132] Screening can be performed using any of the methods disclosed, alone or in combination. Screening can be performed using Polymerase Chain Reaction (PCR). Screening can be performed using Array Comparative Genomic Hybridization (aCGH) to detect CNVs. Screening can be performed using exome sequencing to detect SNVs, indels, and in some cases CNVs using appropriate analysis algorithms. Screening may be performed using high-throughput (also known as next generation) whole genome sequencing methods and appropriate algorithms to detect all or nearly all genetic variations present in a genomic DNA sample. The genetic variation information as it relates to the current disclosure can be used in conjunction with any of the above mentioned symptomatic screening tests to screen a subject for PML, for example, using a combination of aCGH and/or sequencing with a JCV screening test, such as the JCV antibody test, CD62L test, or CSF IgM oligoclonal band test. The L-selectin (CD62L) expressed by $CD3^{+}CD4^{+}$ T cells in, for example, cryopreserved peripheral blood mononuclear cells (PBMCs), can be a biomarker for JCV screening. A CD62L expression can be correlated with the risk of PML.

[0133] In the present context, the term screening comprises diagnosis, prognosis, and theranosis. Screening can refer to any available screening method, including those mentioned herein. As used herein, susceptibility can be proneness of a subject towards the development of PML, or towards being less able to resist PML than one or more control subjects. Susceptibility can encompass increased susceptibility. For example, particular nucleic acid variations of the disclosure as described herein can be characteristic of increased susceptibility to PML. particular nucleic acid variations can confer decreased susceptibility, for example particular nucleic variations of the disclosure as described herein can be characteristic of decreased susceptibility to development of PML.

[0134] A genetic variant associated with PML can be used to predict the susceptibility of the disease for a given genotype. For any genetic variation, there can be one or more possible genotypes, for example, homozygote for the at-risk variant (e.g., in autosomal recessive disorders), heterozygote, and non-carrier of the at-risk variant. Autosomal recessive disorders can also result from two distinct

genetic variants impacting the same gene such that the individual is a compound heterozygote (e.g., the maternal allele contains a different mutation than the paternal allele). Compound heterozygosity may result from two different SNVs, two different CNVs, an SNV and a CNV, or any combination of two different genetic variants but each present on a different allele for the gene. For X-linked genes, males who possess one copy of a variant-containing gene may be affected, while carrier females, who also possess a wild-type gene, may remain unaffected. Susceptibility associated with variants at multiple loci can be used to estimate overall susceptibility. For multiple genetic variants, there can be k (k = 3^n * 2^P) possible genotypes; wherein n can be the number of autosomal loci and p can be the number of gonosomal (sex chromosomal) loci. Overall susceptibility assessment calculations can assume that the relative susceptibilities of different genetic variants multiply, for example, the overall susceptibility associated with a particular genotype combination can be the product of the susceptibility values for the genotype at each locus. If the susceptibility presented is the relative susceptibility for a person, or a specific genotype for a person, compared to a reference population, then the combined susceptibility can be the product of the locus specific susceptibility values and can correspond to an overall susceptibility estimate compared with a population. If the susceptibility for a person is based on a comparison to non-carriers of the at-risk allele, then the combined susceptibility can correspond to an estimate that compares the person with a given combination of genotypes at all loci to a group of individuals who do not carry at-risk variants at any of those loci. The group of non-carriers of any at-risk variant can have the lowest estimated susceptibility and can have a combined susceptibility, compared with itself, for example, non-carriers, of 1.0, but can have an overall susceptibility, compared with the population, of less than 1.0.

[0135] Overall risk for multiple risk variants can be performed using standard methodology. Genetic variations described herein can form the basis of risk analysis that combines other genetic variations known to increase risk of PML, or other genetic risk variants for PML. A plurality of variants (genetic variations, variant alleles, and/or haplotypes) can be used for overall risk assessment. These variants may be selected from the genetic variations as disclosed herein. The variants of the present disclosure may be useful in combination with other variants known to be useful for screening a susceptibility to PML. Here, the genotype status of a plurality of genetic variations, markers and/or haplotypes is determined in an individual, and the status of the individual compared with the population frequency of the associated variants, or the frequency of the variants in clinically healthy subjects, such as age-matched and sex-matched subjects.

[0136] Multivariate analyses or joint risk analyses, including the use of multiplicative model for overall risk assessment, can subsequently be used to determine the overall risk conferred based on the genotype status at the multiple loci. Use of a multiplicative model, for example, assuming that the risk of individual risk variants multiply to establish the overall effect, allows for a straight-forward calculation of the overall risk for multiple markers. The multiplicative model is a parsimonious model that usually fits the data of complex traits reasonably well. Deviations from multiplicity have been rarely described in the context of common variants for common diseases, and if reported are usually only suggestive since very large sample sizes can be required to be able to demonstrate statistical interactions between loci. Assessment of risk based on such analysis can subsequently be used in the methods, uses and kits of the disclosure, as described herein.

[0137] The significance of increased or decreased susceptibility can be measured by a percentage. A significant increased susceptibility can be measured as a relative susceptibility of at least 1.2, including but not limited to: at least 1.3, at least 1.4, at least 1.5, at least 1.6, at least 1.7, at least 1.8, at least 1.9, at least 2.0, at least 2.5, at least 3.0, at least 4.0, at least 5.0, at least 6.0, at least 7.0, at least 8.0, at least 9.0, at least 10.0, and at least 15.0. A relative susceptibility of at least 2.0, at least 3.0, at least 4.0, at least, 5.0, at least 6.0, or at least 10.0 may be significant. Other values for significant susceptibility are also contemplated, for example, at least 2.5, 3.5, 4.5, 5.5, or any suitable other numerical values, wherein the values are also within scope of the present disclosure. A significant increase in susceptibility may be at least about 20%, including but not limited to about 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%, and 1500%. A significant increase in susceptibility may be at least 100%. A significant increase in susceptibility may be at least 200%, at least 300%, at least 400%, at least 500%, at least 700%, at least 800%, at least 900% and at least 1000%. Other cutoffs or ranges as deemed suitable by the person skilled in the art to characterize the disclosure are also contemplated, and those are also within scope of the present disclosure. A significant increase in susceptibility may be characterized by a p-value, such as a p-value of less than 0.5, less than 0.4, less than 0.3, less than 0.2, less than 0.1, less than 0.05, less than 0.01, less than 0.001, less than 0.0001, less than 0.00001, less than 0.000001, less than 0.0000001, less than 0.00000001, or less than 0.000000001.

[0138] The significance of increased or decreased susceptibility can be determined according to the ratio of measurements from a test subject to a reference subject. Losses or gains of one or more CNVs can be determined according to a threshold $\log_2$ ratio determined by these measurements. A $\log_2$ ratio value greater than 0.35, or 0.5, may be indicative of a gain of one or more CNVs. A $\log_2$ ratio value less than - 0.35, or -0.5, is indicative of a loss of one or more CNVs. The ratio of measurements from a test subject to a reference subject may be inverted such that the log2 ratios of copy number gains are negative and the log2 ratios of copy number losses are positive.
the combined or overall susceptibility associated with a plurality of variants associated with PML can also be assessed; for example, the genetic variations described herein to be associated with susceptibility to PML can be combined with other common genetic risk factors. Combined risk for such genetic variants can be estimated in an analogous fashion to the methods described herein.

[0139] Calculating risk conferred by a particular genotype for the individual can be based on comparing the genotype of the individual to previously determined risk expressed, for example, as a relative risk (RR) or an odds ratio (OR), for the genotype, for example, for a heterozygous carrier of an at-risk variant for PML. An odds ratio can be a statistical measure used as a metric of causality. For example, in genetic disease research it can be used to convey the significance of a variant in a disease cohort relative to an unaffected/normal cohort. The calculated risk for the individual can be the relative risk for a subject, or for a specific genotype of a subject, compared to the average population. The average population risk can be expressed as a weighted average of the risks of different genotypes, using

results from a reference population, and the appropriate calculations to calculate the risk of a genotype group relative to the population can then be performed. Alternatively, the risk for an individual can be based on a comparison of particular genotypes, for example, heterozygous and/or homozygous carriers of an at-risk allele of a marker compared with non-carriers of the at-risk allele (or pair of alleles in the instance of compound heterozygous variants, wherein one variant impacts the maternally inherited allele and the other impacts the paternally inherited allele). Using the population average can be more convenient, since it provides a measure that can be easy to interpret for the user, for example, a measure that gives the risk for the individual, based on his/her genotype, compared with the average in the population.

[0140] The OR value can be calculated as follows: OR = (A/(N1-A))/(U/(N2-U)), where A = number of affected cases with variant, N1 = total number of affected cases, U = number of unaffected cases with variant and N2 = total number of unaffected cases. In circumstances where U = 0, it is conventional to set U=1, so as to avoid infinities. The OR can be calculated essentially as above, except that where U or A = 0, 0.5 is added to all of A, N1, U, N2. A Fisher's Exact Test (FET) can be calculated using standard methods. The p-values can be corrected for false discovery rate (FDR) using the Benjamini-Hochberg method (Benjamini Y. and Hochberg Y., J. Royal Statistical Society 57:289 (1995); Osborne J. A. and Barker C. A. (2007)).

[0141] A genetic variation may be correlated to PML by referencing genetic variation data to a look-up table that comprises correlations between the genetic variation and PML. The genetic variation may comprise at least one indication of the genetic variation. The table may comprise a correlation for one genetic variation. The table may comprise a correlation for a plurality of genetic variations in both scenarios, by referencing to a look-up table that gives an indication of a correlation between a genetic variation and PML, a risk for PML, or a susceptibility to PML, can be identified in the individual from whom the nucleic acid sample is derived.

[0142] The present disclosure also pertains to methods of clinical screening, for example, diagnosis, prognosis, or theranosis of a subject performed by a medical professional using the methods disclosed herein. The methods of screening may be performed by a layman. The layman can be a customer of a genotyping, microarray, exome sequencing, or whole genome sequencing service provider. The layman can also be a genotype, microarray, exome sequencing, or whole genome sequencing service provider, who performs genetic analysis on a DNA sample from an individual, in order to provide service related to genetic risk factors for particular traits or diseases, based on the genotype status of the subject obtained from use of the methods described herein. The resulting genotype or genetic information can be made available to the individual and can be compared to information about PML or risk of developing PML associated with one or various genetic variations, including but not limited to, information from public or private genetic variation databases or literature and scientific publications. The screening applications of PML-associated genetic variations, as described herein, can, for example, be performed by an individual, a health professional, or a third party, for example a service provider who interprets genotype information from the subject. The genetic analysis may be performed in a CLIA-certified laboratory (e.g. the federal regulatory standards the U.S. that are specified in the Clinical Laboratory Improvement Amendments, administered by the Centers for Medicare and Medicaid Services) or equivalent laboratories in Europe and elsewhere in the world.

[0143] The information derived from analyzing sequence data can be communicated to any particular body, including the individual from which the nucleic acid sample or sequence data is derived, a guardian or representative of the individual, clinician, research professional, medical professional, service provider, and medical insurer or insurance company. Medical professionals can be, for example, doctors, nurses, medical laboratory technologists, and pharmacists. Research professionals can be, for example, principle investigators, research technicians, postdoctoral trainees, and graduate students.

[0144] A medical professional can initiate or modify treatment after receiving information regarding a subject's screening for PML, for example. A medical professional can recommend a change in therapy or exclude a therapy. A medical professional can enroll a subject in a clinical trial for, by way of example, detecting correlations between a haplotype as described herein and any measurable or quantifiable parameter relating to the outcome of the treatment as described above.

[0145] Results of these tests, and optionally interpretive information, can be returned to the subject, the health care provider or to a third party. The results can be communicated to the tested subject, for example, with a prognosis and optionally interpretive materials that can help the subject understand the test results and prognosis; used by a health care provider, for example, to determine whether to administer a specific drug, or whether a subject should be assigned to a specific category, for example, a category associated with a specific disease endophenotype, or with drug response or non-response; used by a third party such as a healthcare payer, for example, an insurance company or HMO, or other agency, to determine whether or not to reimburse a health care provider for services to the subject, or whether to approve the provision of services to the subject. For example, the healthcare payer can decide to reimburse a health care provider for treatments for PML if the subject has PML or has an increased risk of developing PML.

Methods of Screening using Variations in RNA and/or Polypeptides

[0146] The present invention involves determining the absence or presence of a genomic variation selected from chr9:137779251, G>A and chr1:160769595, AG>A. The following provides a general explanation of screening methods that may be used to determine the presence or absence of the genomic variations of the claims, and includes examples not forming subject matter being claimed herein, insofar as they relate to variations other than chr9:137779251, G>A and chr1:160769595, AG>A. Screening of PML can be made by examining or comparing changes in expression, localization, binding partners, and composition of a polypeptide encoded by a nucleic

acid variant associated with PML, for example, in those instances where the genetic variations of the present disclosure results in a change in the composition or expression of the polypeptide and/or RNA, for example, mRNAs, microRNAs (miRNAs), and other noncoding RNAs (ncRNAs). Thus, screening of PML can be made by examining expression and/or composition of one of these polypeptides and/or RNA, or another polypeptide and/or RNA encoded by a nucleic acid associated with PML, in those instances where the genetic variation of the present disclosure results in a change in the expression, localization, binding partners, and/or composition of the polypeptide and/or RNA. Screening can comprise diagnosing a subject. Screening can comprise determining a prognosis of a subject, for example determining the susceptibility of developing PML. Screening can comprise theranosing a subject.

[0147] The genetic variations described herein that show association to PML can play a role through their effect on one or more of these genes, either by directly impacting one or more genes or influencing the expression of one or more nearby genes. For example, while not intending to be limited by theory, it is generally expected that a deletion of a chromosomal segment comprising a particular gene, or a fragment of a gene, can either result in an altered composition or expression, or both, of the encoded polypeptide and/or mRNA. Likewise, duplications, or high number copy number variations, are in general expected to result in increased expression of encoded polypeptide and/or RNA if the gene they are expressed from is fully encompassed within the duplicated (or triplicated, or even higher copy number gains) genomic segment, or conversely can result in decreased expression or a disrupted RNA or polypeptide if one or both breakpoints of the copy number gain disrupt a given gene. Other possible mechanisms affecting genes within a genetic variation region include, for example, effects on transcription, effects on RNA splicing, alterations in relative amounts of alternative splice forms of mRNA, effects on RNA stability, effects on transport from the nucleus to cytoplasm, and effects on the efficiency and accuracy of translation. Thus, DNA variations can be detected directly, using the subjects unamplified or amplified genomic DNA, or indirectly, using RNA or DNA obtained from the subject's tissue(s) that are present in an aberrant form or expression level as a result of the genetic variations of the disclosure showing association to PML. DNA variations can be detected indirectly using a polypeptide or protein obtained from the subject's tissue(s) that is present in an aberrant form or expression level as a result of genetic variations of the disclosure showing association to the PML. An aberrant form or expression level of a polypeptide or protein that results from one or more genetic variations of the disclosure showing association to PML can be detected indirectly via another polypeptide or protein present in the same biological/cellular pathway that is modulated or interacts with said polypeptide or protein that results from one or more genetic variations of the disclosure. The genetic variations of the disclosure showing association to PML can affect the expression of a gene within the genetic variation region. A genetic variation affecting an exonic region of a gene can affect, disrupt, or modulate the expression of the gene. A genetic variation affecting an intronic or intergenic region of a gene can affect, disrupt, or modulate the expression of the gene.

[0148] Certain genetic variation regions can have flanking duplicated segments, and genes within such segments can have altered expression and/or composition as a result of such genomic alterations. Regulatory elements affecting gene expression can be located far away, even as far as tens or hundreds of kilobases away, from the gene that is regulated by said regulatory elements. Thus, regulatory elements for genes that are located outside the gene (e.g., upstream or downstream of the gene) can be located within the genetic variation, and thus be affected by the genetic variation. It is thus contemplated that the detection of the genetic variations described herein, can be used for assessing expression for one or more of associated genes not directly impacted by the genetic variations. A genetic variation affecting an intergenic region of a gene can affect, disrupt, or modulate the expression of a gene located elsewhere in the genome, such as described above. For example, a genetic variation affecting an intergenic region of a gene can affect, disrupt, or modulate the expression of a transcription factor, located elsewhere in the genome, which regulates the gene. Regulatory elements can also be located within a gene, such as within intronic regions, and similarly impact the expression level of the gene and ultimately the protein expression level without changing the structure of the protein. The effects of genetic variants on regulatory elements can manifest in a tissue-specific manner; for example, one or more transcription factors that bind to the regulatory element that is impacted by one or more genetic variations may be expressed at higher concentration in neurons as compared to skin cells (e.g., the impact of the one or more genetic variations may be primarily evident in neuronal cells).

[0149] Genetic variations of the disclosure showing association to PML can affect protein expression at the translational level. It can be appreciated by those skilled in the art that this can occur by increased or decreased expression of one or more microRNAs (miRNAs) that regulates expression of a protein known to be important, or implicated, in the cause, onset, or progression of PML. Increased or decreased expression of the one or more miRNAs can result from gain or loss of the whole miRNA gene, disruption or impairment of a portion of the gene (e.g., by an indel or CNV), or even a single base change (SNP or SNV) that produces an altered, non-functional or aberrant functioning miRNA sequence. It can also be appreciated by those skilled in the art that the expression of protein, for example, one known to cause PML by increased or decreased expression, can result due to a genetic variation that results in alteration of an existing miRNA binding site within the polypeptide's mRNA transcript, or even creates a new miRNA binding site that leads to aberrant polypeptide expression.

[0150] A variety of methods can be used for detecting polypeptide composition and/or expression levels, including but not limited to enzyme linked immunosorbent assays (ELISA), Western blots, spectroscopy, mass spectrometry, peptide arrays, colorimetry, electrophoresis, isoelectric focusing, immunoprecipitations, immunoassays, and immunofluorescence and other methods well-known in the art. A test nucleic acid sample from a subject can be assessed for the presence of an alteration in the expression and/or an alteration in composition of the polypeptide encoded by a nucleic acid associated with PML. An "alteration" in the polypeptide expression or composition, as used herein, refers to an alteration in expression or composition in a test nucleic acid sample, as compared to the expression or composition of the polypeptide in a control nucleic acid sample. Such alteration can, for example, be an alteration in the quantitative polypeptide expression or can be an alteration in the qualitative polypeptide expression, for example, expression of a mutant polypeptide or of a different splicing variant, or a combination thereof. Screening of PML can be made by detecting a particular splicing variant encoded by a nucleic acid associated with PML, or a particular pattern of splicing variants.

[0151] Antibodies can be polyclonal or monoclonal and can be labeled or unlabeled. An intact antibody or a fragment thereof can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled as previously described herein. Other non-limiting examples of indirect labeling include detection of a primary antibody using a labeled secondary antibody, for example, a fluorescently-labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently-labeled streptavidin.

Methods of Detecting Genetic Variations

[0152] The present invention involves determining the absence or presence of a genomic variation selected from chr9:137779251, G>A and chr1:160769595, AG>A. The following provides a general explanation of methods that may be used to determine the presence or absence of the genomic variations of the claims, and includes examples not forming subject matter being claimed herein, insofar as they relate to variations other than chr9: 137779251, G>A and chr1:160769595, AG>A. Standard techniques for genotyping for the presence genetic variations, for example, amplification, can be used. Amplification of nucleic acids can be accomplished using methods known in the art. Generally, sequence information from the region of interest can be used to design oligonucleotide primers that can be identical or similar in sequence to opposite strands of a template to be amplified. Amplification methods can include but are not limited to, fluorescence-based techniques utilizing PCR, for example, ligase chain reaction (LCR), Nested PCR, transcription amplification, self-sustained sequence replication, nucleic acid based sequence amplification (NASBA), and multiplex ligation-dependent probe amplification (MLPA). Guidelines for selecting primers for PCR amplification are well known in the art. A computer program can be used to design primers, for example, Oligo (National Biosciences, Inc, Plymouth Minn), MacVector (Kodak/IBI), and GCG suite of sequence analysis programs.

[0153] Commercial methodologies available for genotyping, for example, SNP genotyping, can be used, but are not limited to, TaqMan genotyping assays (Applied Biosystems), SNPlex platforms (Applied Biosystems), gel electrophoresis, capillary electrophoresis, size exclusion chromatography, mass spectrometry, for example, MassARRAY system (Sequenom), minisequencing methods, real-time Polymerase Chain Reaction (PCR), Bio-Plex system (BioRad), CEQ and SNPstream systems (Beckman), array hybridization technology, for example, Affymetrix GeneChip (Perlegen), BeadArray Technologies, for example, Illumina GoldenGate and Infinium assays, array tag technology, Multiplex Ligation-dependent Probe Amplification (MLPA), and endonuclease-based fluorescence hybridization technology (Invader assay, either using unamplified or amplified genomic DNA, or unamplified total RNA, or unamplified or amplified cDNA; Third Wave/Hologic). PCR can be a procedure in which target nucleic acid is amplified in a manner similar to that described in U.S. Pat. No. 4,683,195 and subsequent modifications of the procedure described therein. PCR can include a three phase temperature cycle of denaturation of DNA into single strands, annealing of primers to the denatured strands, and extension of the primers by a thermostable DNA polymerase enzyme. This cycle can be repeated so that there are enough copies to be detected and analyzed. Real-time quantitative PCR can be used to determine genetic variations, wherein quantitative PCR can permit both detection and quantification of a DNA sequence in a nucleic acid sample, for example, as an absolute number of copies or as a relative amount when normalized to DNA input or other normalizing genes. Methods of quantification can include the use of fluorescent dyes that can intercalate with double-stranded DNA, and modified DNA oligonucleotide probes that can fluoresce when hybridized with a complementary DNA.

[0154] A nucleic acid sample obtained from the subject can be collected and PCR can be used to amplify a fragment of nucleic acid that comprises one or more genetic variations that can be indicative of a susceptibility to PML. Detection of genetic variations can be accomplished by expression analysis, for example, by using quantitative PCR. This technique can assess the presence or absence of a genetic alteration in the expression or composition of one or more polypeptides or splicing variants encoded by a nucleic acid associated with PML.

[0155] The nucleic acid sample from a subject containing a SNP can be amplified by PCR prior to detection with a probe. The amplified DNA may serve as the template for a detection probe and/or an enhancer probe. The detection probe, the enhancer probe, and/or the primers used for amplification of the template by PCR can comprise the use of modified bases, for example, modified A, T, C, G, and U, wherein the use of modified bases can be useful for adjusting the melting temperature of the nucleotide probe and/or primer to the template DNA, Modified bases may be used in the design of the detection nucleotide probe. Any modified base known to the skilled person can be selected in these methods, and the selection of suitable bases is well within the scope of the skilled person based on the teachings herein and known bases available from commercial sources as known to the skilled person.

[0156] Identification of genetic variations can be accomplished using hybridization methods. The presence of a specific marker allele or a particular genomic segment comprising a genetic variation, or representative of a genetic variation, can be indicated by sequence-specific hybridization of a nucleic acid probe specific for the particular allele or the genetic variation in a nucleic acid sample that has or has not been amplified but methods described herein. The presence of more than one specific marker allele or several genetic variations can be indicated by using two or more sequence-specific nucleic acid probes, wherein each is specific for a particular allele and/or genetic variation.

[0157] Hybridization can be performed by methods well known to the person skilled in the art, for example, hybridization techniques such as fluorescent in situ hybridization (FISH), Southern analysis, Northern analysis, or in situ hybridization. Hybridization may refer to specific hybridization, wherein hybridization can be performed with no mismatches. Specific hybridization, if present, can be using standard

methods. If specific hybridization occurs between a nucleic acid probe and the nucleic acid in the nucleic acid sample, the nucleic acid sample can contain a sequence that can be complementary to a nucleotide present in the nucleic acid probe. If a nucleic acid probe can contain a particular allele of a polymorphic marker, or particular alleles for a plurality of markers, specific hybridization is indicative of the nucleic acid being completely complementary to the nucleic acid probe, including the particular alleles at polymorphic markers within the probe. A probe can contain more than one marker alleles of a particular haplotype, for example, a probe can contain alleles complementary to 2, 3, 4, 5 or all of the markers that make up a particular haplotype. Detection of one or more particular markers of the haplotype in the nucleic acid sample is indicative that the source of the nucleic acid sample has the particular haplotype.

[0158] PCR conditions and primers can be developed that amplify a product only when the variant allele is present or only when the wild type allele is present, for example, allele-specific PCR. When carrying out allele-specific PCR, a method utilizing a detection oligonucleotide probe comprising a fluorescent moiety or group at its 3' terminus and a quencher at its 5' terminus, and an enhancer oligonucleotide, can be employed (*see e.g.,* Kutyavin et al., Nucleic Acid Res. 34:e128 (2006)).

[0159] An allele-specific primer/probe can be an oligonucleotide that is specific for particular a polymorphism can be prepared using standard methods. Allele-specific oligonucleotide probes can specifically hybridize to a nucleic acid region that contains a genetic variation. Hybridization conditions can be selected such that a nucleic acid probe can specifically bind to the sequence of interest, for example, the variant nucleic acid sequence.

[0160] Allele-specific restriction digest analysis can be used to detect the existence of a polymorphic variant of a polymorphism, if alternate polymorphic variants of the polymorphism can result in the creation or elimination of a restriction site. Allele-specific restriction digests can be performed, for example, with the particular restriction enzyme that can differentiate the alleles. PCR can be used to amplify a region comprising the polymorphic site, and restriction fragment length polymorphism analysis can be conducted. For sequence variants that do not alter a common restriction site, mutagenic primers can be designed that can introduce one or more restriction sites when the variant allele is present or when the wild type allele is present.

[0161] Fluorescence polarization template-directed dye-terminator incorporation (FP-TDI) can be used to determine which of multiple polymorphic variants of a polymorphism can be present in a subject. Unlike the use of allele-specific probes or primers, this method can employ primers that can terminate adjacent to a polymorphic site, so that extension of the primer by a single nucleotide can result in incorporation of a nucleotide complementary to the polymorphic variant at the polymorphic site.

[0162] DNA containing an amplified portion can be dot-blotted, using standard methods and the blot contacted with the oligonucleotide probe. The presence of specific hybridization of the probe to the DNA can then be detected. The methods can include determining the genotype of a subject with respect to both copies of the polymorphic site present in the genome, wherein if multiple polymorphic variants exist at a site, this can be appropriately indicated by specifying which variants are present in a subject. Any of the detection means described herein can be used to determine the genotype of a subject with respect to one or both copies of the polymorphism present in the subject's genome.

[0163] A peptide nucleic acid (PNA) probe can be used in addition to, or instead of, a nucleic acid probe in the methods described herein. A PNA can be a DNA mimic having a peptide-like, inorganic backbone, for example, N-(2-aminoethyl) glycine units with an organic base (A, G, C, T or U) attached to the glycine nitrogen via a methylene carbonyl linker.

[0164] Nucleic acid sequence analysis can also be used to detect genetic variations, for example, genetic variations can be detected by sequencing exons, introns, 5' untranslated sequences, or 3' untranslated sequences. One or more methods of nucleic acid analysis that are available to those skilled in the art can be used to detect genetic variations, including but not limited to, direct manual sequencing, automated fluorescent sequencing, single-stranded conformation polymorphism assays (SSCP); clamped denaturing gel electrophoresis (CDGE); denaturing gradient gel electrophoresis (DGGE), two-dimensional gel electrophoresis (2DGE or TDGE); conformational sensitive gel electrophoresis (CSGE); denaturing high performance liquid chromatography (DHPLC), infrared matrix-assisted laser desorption/ionization (IR-MALDI) mass spectrometry, mobility shift analysis, quantitative real-time PCR, restriction enzyme analysis, heteroduplex analysis; chemical mismatch cleavage (CMC), RNase protection assays, use of polypeptides that recognize nucleotide mismatches, allele-specific PCR, real-time pyrophosphate DNA sequencing, PCR amplification in combination with denaturing high performance liquid chromatography (dHPLC), and combinations of such methods.

[0165] Sequencing can be accomplished through classic Sanger sequencing methods, which are known in the art. Sequencing can be performed using high-throughput sequencing methods some of which allow detection of a sequenced nucleotide immediately after or upon its incorporation into a growing strand, for example, detection of sequence in substantially real time or real time. In some cases, high throughput sequencing generates at least 1,000, at least 5,000, at least 10,000, at least 20,000, at least 30,000, at least 40,000, at least 50,000, at least 100,000 or at least 500,000 sequence reads per hour; with each read being at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 120 or at least 150 bases per read (or 500 - 1,000 bases per read for 454).

[0166] High-throughput sequencing methods can include but are not limited to, Massively Parallel Signature Sequencing (MPSS, Lynx Therapeutics), Polony sequencing, 454 pyrosequencing, Illumina (Solexa) sequencing, Illumina (Solexa) sequencing using 10X Genomics library preparation, SOLiD sequencing, on semiconductor sequencing, DNA nanoball sequencing, Helioscope™ single molecule sequencing, Single Molecule SMRT™ sequencing, Single Molecule real time (RNAP) sequencing, Nanopore DNA sequencing, and/or

sequencing by hybridization, for example, a non-enzymatic method that uses a DNA microarray, or microfluidic Sanger sequencing.

[0167] High-throughput sequencing can involve the use of technology available by Helicos BioSciences Corporation (Cambridge, Mass.) such as the Single Molecule Sequencing by Synthesis (SMSS) method. SMSS is unique because it allows for sequencing the entire human genome in up to 24 hours. This fast sequencing method also allows for detection of a SNP/nucleotide in a sequence in substantially real time or real time. Finally, SMSS is powerful because, like the MIP technology, it does not use a pre-amplification step prior to hybridization. SMSS does not use any amplification. SMSS is described in US Publication Application Nos. 20060024711; 20060024678; 20060012793; 20060012784; and 20050100932. High-throughput sequencing can involve the use of technology available by 454 Life Sciences, Inc. (a Roche company, Branford, Conn.) such as the PicoTiterPlate device which includes a fiber optic plate that transmits chemiluminescent signal generated by the sequencing reaction to be recorded by a CCD camera in the instrument. This use of fiber optics allows for the detection of a minimum of 20 million base pairs in 4.5 hours.

[0168] PCR-amplified single-strand nucleic acid can be hybridized to a primer and incubated with a polymerase, ATP sulfurylase, luciferase, apyrase, and the substrates luciferin and adenosine 5' phosphosulfate. Next, deoxynucleotide triphosphates corresponding to the bases A, C, G, and T (U) can be added sequentially. A base incorporation can be accompanied by release of pyrophosphate, which can be converted to ATP by sulfurylase, which can drive synthesis of oxyluciferin and the release of visible light. Since pyrophosphate release can be equimolar with the number of incorporated bases, the light given off can be proportional to the number of nucleotides adding in any one step. The process can repeat until the entire sequence can be determined. Pyrosequencing can be utilized to analyze amplicons to determine whether breakpoints are present. Pyrosequencing can map surrounding sequences as an internal quality control.

[0169] Pyrosequencing analysis methods are known in the art. Sequence analysis can include a four-color sequencing by ligation scheme (degenerate ligation), which involves hybridizing an anchor primer to one of four positions. Then an enzymatic ligation reaction of the anchor primer to a population of degenerate nonamers that are labeled with fluorescent dyes can be performed. At any given cycle, the population of nonamers that is used can be structured such that the identity of one of its positions can be correlated with the identity of the fluorophore attached to that nonamer. To the extent that the ligase discriminates for complementarily at that queried position, the fluorescent signal can allow the inference of the identity of the base. After performing the ligation and four-color imaging, the anchor primer: nonamer complexes can be stripped and a new cycle begins. Methods to image sequence information after performing ligation are known in the art.

[0170] Analysis by restriction enzyme digestion can be used to detect a particular genetic variation if the genetic variation results in creation or elimination of one or more restriction sites relative to a reference sequence. Restriction fragment length polymorphism (RFLP) analysis can be conducted, wherein the digestion pattern of the relevant DNA fragment indicates the presence or absence of the particular genetic variation in the nucleic acid sample.

[0171] Arrays of oligonucleotide probes that can be complementary to target nucleic acid sequence segments from a subject can be used to identify genetic variations. An array of oligonucleotide probes can comprise an oligonucleotide array, for example, a microarray. Arrays that include a substrate having a plurality of addressable areas can be used. At least one area of the plurality includes a nucleic acid probe that binds specifically to a sequence comprising a genetic variation, and can be used to detect the absence or presence of the genetic variation, for example, one or more SNPs, microsatellites, or CNVs, as described herein, to determine or identify an allele or genotype. For example, the array can include one or more nucleic acid probes that can be used to detect a genetic variation associated with a gene and/or gene product. The array can further comprise at least one area that includes a nucleic acid probe that can be used to specifically detect another marker associated with PML as described herein.

[0172] Microarray hybridization can be performed by hybridizing a nucleic acid of interest, for example, a nucleic acid encompassing a genetic variation, with the array and detecting hybridization using nucleic acid probes. The nucleic acid of interest can be amplified prior to hybridization. Hybridization and detecting can be carried out according to standard methods described in Published PCT Applications: WO 92/10092 and WO 95/11995, and U.S. Pat. No. 5,424,186. For example, an array can be scanned to determine the position on the array to which the nucleic acid hybridizes. The hybridization data obtained from the scan can be, for example, in the form of fluorescence intensities as a function of location on the array.

[0173] Arrays can be formed on substrates fabricated with materials such as paper; glass; plastic, for example, polypropylene, nylon, or polystyrene; polyacrylamide; nitrocellulose; silicon; optical fiber; or any other suitable solid or semisolid support; and can be configured in a planar, for example, glass plates or silicon chips); or three dimensional, for example, pins, fibers, beads, particles, microtiter wells, and capillaries, configuration.

[0174] Methods for generating arrays are known in the art and can include for example; photolithographic methods (U.S. Pat. Nos. 5,143,854, 5,510,270 and 5,527,681); mechanical methods, for example, directed-flow methods (U.S. Pat. No. 5,384,261); pin-based methods (U.S. Pat. No. 5;288;514); bead-based techniques (PCT US/93/04145); solid phase oligonucleotide synthesis methods; or by other methods known to a person skilled in the art (*see*, *e.g.*, 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), and references cited therein). Many additional descriptions of the preparation and use of oligonucleotide arrays for detection of polymorphisms can be found, for example, in 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, and EP 373 203. Methods for array production, hybridization, and analysis are also described in Snijders et al., Nat. Genetics

29:263-264 (2001); Klein et al., Proc. Natl. Acad. Sci. USA 96:4494-4499 (1999); Albertson et al., Breast Cancer Research and Treatment 78:289-298 (2003); and Snijders et al., "BAC microarray based comparative genomic hybridization," in: Zhao et al., (eds), Bacterial Artificial Chromosomes: Methods and Protocols, Methods in Molecular Biology, Humana Press (2002).

[0175] Oligonucleotide probes forming an array can be attached to a substrate by any number of techniques, including, but not limited to, in situ synthesis, for example, high-density oligonucleotide arrays, using photolithographic techniques; spotting/printing a medium to low density on glass, nylon, or nitrocellulose; by masking; and by dot-blotting on a nylon or nitrocellulose hybridization membrane. Oligonucleotides can be immobilized via a linker, including but not limited to, by covalent, ionic, or physical linkage. Linkers for immobilizing nucleic acids and polypeptides, including reversible or cleavable linkers, are known in the art (U.S. Pat. No. 5,451,683 and WO98/20019). Oligonucleotides can be non-covalently immobilized on a substrate by hybridization to anchors, by means of magnetic beads, or in a fluid phase, for example, in wells or capillaries.

[0176] An array can comprise oligonucleotide hybridization probes capable of specifically hybridizing to different genetic variations. Oligonucleotide arrays can comprise a plurality of different oligonucleotide probes coupled to a surface of a substrate in different known locations. Oligonucleotide probes can exhibit differential or selective binding to polymorphic sites, and can be readily designed by one of ordinary skill in the art, for example, an oligonucleotide that is perfectly complementary to a sequence that encompasses a polymorphic site, for example, a sequence that includes the polymorphic site, within it, or at one end, can hybridize preferentially to a nucleic acid comprising that sequence, as opposed to a nucleic acid comprising an alternate polymorphic variant.

[0177] Arrays can include multiple detection blocks, for example, multiple groups of probes designed for detection of particular polymorphisms. These arrays can be used to analyze multiple different polymorphisms. Detection blocks can be grouped within a single array or in multiple, separate arrays, wherein varying conditions, for example, conditions optimized for particular polymorphisms, can be used during hybridization. General descriptions of using oligonucleotide arrays for detection of polymorphisms can be found, for example, in U.S. Pat. Nos. 5,858,659 and 5,837,832. In addition to oligonucleotide arrays, cDNA arrays can be used.

[0178] Example methods can include but are not limited to providing an array as described herein; contacting the array with a nucleic acid sample, and detecting binding of a nucleic acid from the nucleic acid sample to the array. Such a method can comprise amplifying nucleic acid from the nucleic acid sample, for example, a region associated with PML or a region that includes another region associated with PML. The methods described herein can include using an array that can identify differential expression patterns or copy numbers of one or more genes in nucleic acid samples from control and affected individuals. For example, arrays of probes to a marker described herein can be used to identify genetic variations between DNA from an affected subject, and control DNA obtained from an individual that does not have PML. Since the nucleotides on the array can contain sequence tags, their positions on the array can be accurately known relative to the genomic sequence.

[0179] It can be desirable to employ methods that can detect the presence of multiple genetic variations, for example, polymorphic variants at a plurality of polymorphic sites, in parallel or substantially simultaneously. These methods can comprise oligonucleotide arrays and other methods, including methods in which reactions, for example, amplification and hybridization, can be performed in individual vessels, for example, within individual wells of a multi-well plate or other vessel.

[0180] Determining the identity of a genetic variation can also include or consist of reviewing a subject's medical history, where the medical history includes information regarding the identity, copy number, presence or absence of one or more alleles or SNPs in the subject, e.g., results of a genetic test.

[0181] Extended runs of homozygosity (ROH) may be useful to map recessive disease genes in outbred populations. Furthermore, even in complex disorders, a high number of affected individuals may have the same haplotype in the region surrounding a disease mutation. Therefore, a rare pathogenic variant and surrounding haplotype can be enriched in frequency in a group of affected individuals compared with the haplotype frequency in a cohort of unaffected controls. Homozygous haplotypes (HH) that are shared by multiple affected individuals can be important for the discovery of recessive disease genes in a condition such as PML. The traditional homozygosity mapping method can be extended by analyzing the haplotype within shared ROH regions to identify homozygous segments of identical haplotype that are present uniquely or at a higher frequency in PML probands compared to parental controls. Such regions are termed risk homozygous haplotypes (rHH), which may contain low-frequency recessive variants that contribute to PML risk in a subset of PML patients.

[0182] Genetic variations can also be identified using any of a number of methods well known in the art. For example, genetic variations available in public databases, which can be searched using methods and custom algorithms or algorithms known in the art, can be used. A reference sequence can be from, for example, the human draft genome sequence, publicly available in various databases, or a sequence deposited in a database such as GenBank.

[0183] A comparison of one or more genomes relative to one or more other genomes with array CGH, or a variety of other genetic variation detection methods, can reveal the set of genetic variations between two genomes, between one genome in comparison to multiple genomes, or between one set of genomes in comparison to another set of genomes. An array CGH experiment can be performed by hybridizing a single test genome against a pooled nucleic acid sample of two or more genomes, which can result in minimizing the detection of higher frequency variants in the experiment. A test genome can be hybridized alone (e.g., one-color detection) to a microarray, for example, using array CGH or SNP genotyping methods, and the comparison step to one or more reference genomes

can be performed *in silico* to reveal the set of genetic variations in the test genome relative to the one or more reference genomes. A single test genome can be compared to a single reference genome in a 2-color experiment wherein both genomes are cohybridized to the microarray. The whole genome or whole exome from one or more subjects can be analyzed. Nucleic acid information may have already been obtained for the whole genome or whole exome from one or more individuals and the nucleic acid information is obtained from *in silico* analysis.

[0184] Any of the polynucleotides described, including polynucleotides comprising a genetic variation, can be made synthetically using methods known in the art.

Immunosuppressive Medicament For Use In Therapy

[0185] In an aspect the invention provides an immunosuppressive medicament for use in the treatment of a condition in a subject in need of immunosuppressive medicament therapy. The immunosuppressive medicament for use of the present invention involves determining the absence of a genomic variation selected from chr9:137779251, G>A and chr1:160769595, AG>A. The following provides a general explanation of an immunosuppressive medicament for use, and includes examples not forming subject matter being claimed herein, insofar as they relate to variations other than chr9: 137779251, G>A and chr1:160769595, AG>A. The condition can be HIV/AIDS, cancer, or an autoimmune disease. In some embodiments, the condition can be PML. For example, the condition can be multiple sclerosis.. The one or more immunosuppressive medicaments can be adalimumab (e.g., HUMIRA), alemtuzumab (e.g., LEMTRADA), alemtuzumab (e.g., CAMPATH), azathioprine (e.g., IMURAN), belimumab (e.g., BENLYSTA), bevacizumab (e.g., AVASTIN), bortezomib (e.g., VELCADE), eculizumab (e.g., SOLIRIS), leflunomide, brentuximab vedotin (e.g., ADCETRIS), cetuximab (e.g., ERBITUX), cyclophosphamid, dimethyl fumarate (e.g., TECFIDERA), efalizumab (e.g., RAPTIVA), fingolimod (e.g., GILENYA), fludarabine (e.g., FLUDARA), fumaric acid, imatinib (e.g., GLEEVEC, GLIVEC), infliximab (e.g., REMICADE), methotrexate (e.g., TREXALL, RHEUMATREX), mycophenolate mofetil (e.g., CELLCEPT), natalizumab (e.g., TYSABRI), rituximab (e.g., RITUXAN), daclizumab (e.g., ZINBRYTA), vedolizumab (ENTYVIO), ruxolitinib (e.g., JAKAFI, JAKAVI), ocrelizumab (e.g., OCREVUS), or any combinations thereof. The term "animal subject" as used herein includes humans as well as other mammals. The term "treating" as used herein includes achieving a therapeutic benefit and/or a prophylactic benefit. By therapeutic benefit is meant eradication or amelioration of the underlying viral infection (e.g., HIV), cancer, or autoimmune disease.

[0186] In some embodiments, the immunosuppressive medicament for use is for administration to a subject can be currently treated with an antiretroviral medicament. In some embodiments, the immunosuppressive medicament for use is for administration to a subject who has previously been treated with an antiretroviral medicament. In some embodiments, the immunosuppressive medicament for use is for administration to a subject who has not yet been treated with an antiretroviral medicament. The antiretroviral medicament can include but not limited to Nucleoside Reverse Transcriptase Inhibitors (NRTIs), Non-Nucleoside Reverse Transcriptase Inhibitors (NNRTIs), Protease Inhibitors (PIs), Fusion Inhibitors, Entry Inhibitors, Integrase Inhibitors, Pharmacokinetic Enhancers, and Combination HIV Medicines. In some cases, the Nucleoside Reverse Transcriptase Inhibitors can include but not limited to abacavir, didanosine, emtricitabine, lamivudine, stavudine, tenofovir disoproxil fumarate, and zidovudine. In some cases, the Non-Nucleoside Reverse Transcriptase Inhibitors can include but not limited to efavirenz, etravirine, nevirapine, and rilpivirine. In some cases, the Protease Inhibitors can include but not limited to atazanavir, darunavir, fosamprenavir, indinavir, nelfinavir, ritonavir, saquinavir, and tipranavir. In some cases, the Fusion Inhibitors can include but not limited to enfuvirtide. In some cases, the Entry Inhibitors can include but not limited to maraviroc. In some cases, the Integrase Inhibitors can include but not limited to dolutegravir, elvitegravir, and raltegravir. In some cases, the Pharmacokinetic Enhancers can include but not limited to cobicistat. In some cases, the Combination HIV Medicines can include but not limited to abacavir and lamivudine, abacavir, dolutegravir, and lamivudine, abacavir, lamivudine, and zidovudine, atazanavir and cobicistat, darunavir and cobicistat, efavirenz, emtricitabine, and tenofovir disoproxil fumarate, elvitegravir, cobicistat, emtricitabine, and tenofovir alafenamide fumarate, elvitegravir, cobicistat, emtricitabine, and tenofovir disoproxil fumarate, emtricitabine, rilpivirine, and tenofovir alafenamide, emtricitabine, rilpivirine, and tenofovir disoproxil fumarate, emtricitabine and tenofovir alafenamide, emtricitabine and tenofovir disoproxil fumarate, lamivudine and zidovudine, lopinavir and ritonavir, and any combination of antiretroviral medicaments listed above.

[0187] When a subject is identified as having at least one of the genetic variants described herein, the immunosuppressive medicament for use is for administration to the subject with an agent targeting the JC Virus. In some embodiments, the immunosuppressive medicament for use is for administration to a subject with a medicament that prevents PML from developing, or it can reduce, lessen, shorten and/or otherwise ameliorate the progression of PML, or symptoms that develop. such a medicament can modulate or target JC Virus. In some embodiments, the immunosuppressive medicament for use is for administration to a subject identified as having PML with an agent that reduces a viral load in the subject. In some embodiments, the immunosuppressive medicament for use is for administration prior to, or in conjunction with, an agent that reduces a viral load in the subject. In some embodiments, the immunosuppressive medicament for use is for administration to a subject identified as having a risk of developing PML with an agent that prevents an increase in a viral load in the subject. In some embodiments, the immunosuppressive medicament for use is for administration to a subject identified as having a high risk of developing PML with an agent that prevents an increase in a viral load in the subject. In some embodiments, the immunosuppressive medicament for use is for administration prior to, or in conjunction with, an agent that prevents an increase in a viral load in the subject. The agent that reduces a viral load in the subject or that prevents an increase in a viral load in the subject can be, for example, an agent that targets JC Virus. Exemplary agents include antibodies, such as broadly neutralizing JCV antibodies. For example, an agent can be a broadly neutralizing human monoclonal JC polyomavirus VP-1 specific antibody (See, e.g., Jelcic et al., Science Translational Medicine, Vol. 7, Issue 306, pp. 306ra150 (2015) and Ray et al., Science Translational Medicine, Vol. 7, Issue 306, pp 306ra151 (2015)). Additional exemplary agents include antiretroviral agents, cidofovir, hexadecyloxypropyl-cidofovir (a lipid-

ester derivative), cytarabine (e.g., cytosine arabinoside), agents that block the 5HT2a receptor (e.g., olanzapine, zisprasidone, mirtazapine, cyproheptadine, and risperidone), topoisomerase inhibitors (e.g., topotecan), and mefloquine.

[0188] In some embodiments, the immunosuppressive medicament for use is for administration to a subject at risk of developing PML, or to a subject reporting one or more of the physiological symptoms of PML, even though a screening of the condition cannot have been made.

[0189] The present disclosure also includes kits that can be used to treat a condition in animal subjects. These kits comprise one or more immunosuppressive medicaments and optionally instructions teaching the use of the kit according to the various methods and approaches described herein. Such kits can also include information, such as scientific literature references, package insert materials, clinical trial results, and/or summaries of these and the like, which indicate or establish the activities and/or advantages (or risks and/or disadvantages) of the agent. Such information can be based on the results of various studies, for example, studies using experimental animals involving in vivo models and studies based on human clinical trials. Kits described herein can be provided, marketed and/or promoted to health providers, including physicians, nurses, pharmacists, formulary officials, and the like.

Formulations, Routes of Administration, and Effective Doses

[0190] In an aspect the invention provides an immunosuppressive medicament for use in the treatment of a condition in a subject in need of immunosuppressive medicament therapy. The following provides a general explanation of compositions, formulations, routes of administration, and effective doses useful for such an immunosuppressive medicament for use, but not forming subject matter being claimed herein. Further, The immunosuppressive medicament for use of the present invention involves determining the absence of a genomic variation selected from chr9: 137779251, G>A and chr1:160769595, AG>A. The following provides a general explanation of variations, and thus includes examples not forming subject matter being claimed herein, insofar as they relate to variations other than chr9: 137779251, G>A and chr1:160769595, AG>A. Yet another aspect of the present disclosure relates to formulations, routes of administration and effective doses for pharmaceutical compositions comprising an agent or combination of agents of the instant disclosure. Such pharmaceutical compositions can be used to treat a condition (e.g., multiple sclerosis) as described above.

[0191] Compounds of the disclosure can be administered as pharmaceutical formulations including those suitable for oral (including buccal and sub-lingual), rectal, nasal, topical, transdermal patch, pulmonary, vaginal, suppository, or parenteral (including intramuscular, intraarterial, intrathecal, intradermal, intraperitoneal, subcutaneous and intravenous) administration or in a form suitable for administration by aerosolization, inhalation or insufflation. General information on drug delivery systems can be found in Ansel et al., Pharmaceutical Dosage Forms and Drug Delivery Systems (Lippencott Williams & Wilkins, Baltimore Md. (1999).

[0192] The pharmaceutical composition may include carriers and excipients (including but not limited to buffers, carbohydrates, mannitol, polypeptides, amino acids, antioxidants, bacteriostats, chelating agents, suspending agents, thickening agents and/or preservatives), water, oils including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like, saline solutions, aqueous dextrose and glycerol solutions, flavoring agents, coloring agents, detackifiers and other acceptable additives, adjuvants, or binders, other pharmaceutically acceptable auxiliary substances to approximate physiological conditions, such as pH buffering agents, tonicity adjusting agents, emulsifying agents, wetting agents and the like. Examples of excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The pharmaceutical preparation may be substantially free of preservatives. The pharmaceutical preparation can contain at least one preservative. General methodology on pharmaceutical dosage forms is found in Ansel et al., Pharmaceutical Dosage Forms and Drug Delivery Systems (Lippencott, Williams, & Wilkins, Baltimore Md. (1999)). It can be recognized that, while any suitable carrier known to those of ordinary skill in the art can be employed to administer the compositions of this disclosure, the type of carrier can vary depending on the mode of administration.

[0193] Compounds can also be encapsulated within liposomes using well-known technology. Biodegradable microspheres can also be employed as carriers for the pharmaceutical compositions of this disclosure. Suitable biodegradable microspheres are disclosed, for example, in U.S. Pat. Nos. 4,897,268, 5,075,109, 5,928,647, 5,811,128, 5,820,883, 5,853,763, 5,814,344 and 5,942,252.

[0194] The compound can be administered in liposomes or microspheres (or microparticles). Methods for preparing liposomes and microspheres for administration to a subject are well known to those of skill in the art. U.S. Pat. No. 4,789,734, describes methods for encapsulating biological materials in liposomes. Essentially, the material is dissolved in an aqueous solution, the appropriate phospholipids and lipids added, and along with surfactants if required, and the material dialyzed or sonicated, as necessary. A review of known methods is provided by G. Gregoriadis, Chapter 14, "Liposomes," Drug Carriers in Biology and Medicine, pp. 2.sup.87-341 (Academic Press, 1979).

[0195] Microspheres formed of polymers or polypeptides are well known to those skilled in the art, and can be tailored for passage through the gastrointestinal tract directly into the blood stream. Alternatively, the compound can be incorporated and the microspheres, or composite of microspheres, implanted for slow release over a period of time ranging from days to months. See, for example, U.S. Pat. Nos. 4,906,474, 4,925,673 and 3,625,214, and Jein, TIPS 19:155-157 (1998).

[0196] The concentration of drug can be adjusted, the pH of the solution buffered and the isotonicity adjusted to be compatible with intravenous injection, as is well known in the art.

[0197] The compounds of the disclosure can be formulated as a sterile solution or suspension, in suitable vehicles, well known in the art. The pharmaceutical compositions can be sterilized by conventional, well-known sterilization techniques, or can be sterile filtered. The resulting aqueous solutions can be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. Suitable formulations and additional carriers are described in Remington "The Science and Practice of Pharmacy" (20th Ed., Lippincott Williams & Wilkins, Baltimore MD).

[0198] The agents or their pharmaceutically acceptable salts can be provided alone or in combination with one or more other agents or with one or more other forms. For example, a formulation can comprise one or more agents in particular proportions, depending on the relative potencies of each agent and the intended indication. For example, in compositions for targeting two different host targets, and where potencies are similar, about a 1:1 ratio of agents can be used. The two forms can be formulated together, in the same dosage unit e.g., in one cream, suppository, tablet, capsule, aerosol spray, or packet of powder to be dissolved in a beverage; or each form can be formulated in a separate unit, e.g., two creams, two suppositories, two tablets, two capsules, a tablet and a liquid for dissolving the tablet, two aerosol sprays, or a packet of powder and a liquid for dissolving the powder, etc.

[0199] The term "pharmaceutically acceptable salt" means those salts which retain the biological effectiveness and properties of the agents used in the present disclosure, and which are not biologically or otherwise undesirable.

[0200] Typical salts are those of the inorganic ions, such as, for example, sodium, potassium, calcium, magnesium ions, and the like. Such salts include salts with inorganic or organic acids, such as hydrochloric acid, hydrobromic acid, phosphoric acid, nitric acid, sulfuric acid, methanesulfonic acid, p toluenesulfonic acid, acetic acid, fumaric acid, succinic acid, lactic acid, mandelic acid, malic acid, citric acid, tartaric acid or maleic acid. In addition, if the agent(s) contain a carboxyl group or other acidic group, it can be converted into a pharmaceutically acceptable addition salt with inorganic or organic bases. Examples of suitable bases include sodium hydroxide, potassium hydroxide, ammonia, cyclohexylamine, dicyclohexyl-amine, ethanolamine, diethanolamine, triethanolamine, and the like.

[0201] A pharmaceutically acceptable ester or amide refers to those which retain biological effectiveness and properties of the agents used in the present disclosure, and which are not biologically or otherwise undesirable. Typical esters include ethyl, methyl, isobutyl, ethylene glycol, and the like. Typical amides include unsubstituted amides, alkyl amides, dialkyl amides, and the like.

[0202] An agent can be administered in combination with one or more other compounds, forms, and/or agents, e.g., as described above. Pharmaceutical compositions with one or more other active agents can be formulated to comprise certain molar ratios. For example, molar ratios of about 99:1 to about 1:99 of a first active agent to the other active agent can be used. The range of molar ratios of a first active agent: other active agents may be selected from about 80:20 to about 20:80; about 75:25 to about 25:75, about 70:30 to about 30:70, about 66:33 to about 33:66, about 60:40 to about 40:60; about 50:50; and about 90:10 to about 10:90. The molar ratio of a first active: other active agents can be about 1:9, or can be about 1:1. The two agents, forms and/or compounds can be formulated together, in the same dosage unit e.g., in one cream, suppository, tablet, capsule, or packet of powder to be dissolved in a beverage; or each agent, form, and/or compound can be formulated in separate units, e.g., two creams, suppositories, tablets, two capsules, a tablet and a liquid for dissolving the tablet, an aerosol spray a packet of powder and a liquid for dissolving the powder, etc.

[0203] If necessary or desirable, the agents and/or combinations of agents can be administered with still other agents. The choice of agents that can be co-administered with the agents and/or combinations of agents of the instant disclosure can depend, at least in part, on the condition being treated. Agents of particular use in the formulations of the present disclosure include, for example, any agent having a therapeutic effect for a viral infection, including, e.g., drugs used to treat inflammatory conditions. For example, in treatments for influenza, formulations of the instant disclosure can additionally contain one or more conventional anti-inflammatory drugs, such as an NSAID, e.g., ibuprofen, naproxen, acetaminophen, ketoprofen, or aspirin. Or, for the treatment of influenza formulations of the instant disclosure can additionally contain one or more conventional influenza antiviral agents, such as amantadine, rimantadine, zanamivir, and oseltamivir. In treatments for retroviral infections, such as HIV, formulations of the instant disclosure can additionally contain one or more conventional antiviral drug, such as protease inhibitors (lopinavir/ritonavir {e.g., KALETRA}, indinavir {e.g., CRIXIVAN}, ritonavir {e.g., NORVIR}, nelfinavir {e.g., VIRACEPT}, saquinavir hard gel capsules {e.g., INVIRASE}, atazanavir {e.g., REYATAZ}, amprenavir {e.g., AGENERASE}, fosamprenavir {e.g., TELZIR}, tipranavir{e.g., APTIVUS}), reverse transcriptase inhibitors, including non-nucleoside and nucleoside/nucleotide inhibitors (AZT {zidovudine, e.g., Retrovir}, ddI {didanosine, e.g., VIDEX}, 3TC {lamivudine, e.g., EPIVIR}, d4T {stavudine, e.g., ZERIT}, abacavir {e.g., ZIAGEN}, FTC {emtricitabine, e.g., EMTRIVA}, tenofovir {e.g., VIREAD}, efavirenz {e.g., SUSTIVA} and nevirapine {e.g., VIRAMUNE}), fusion inhibitors T20 {enfuvirtide, e.g., FUZEON}, integrase inhibitors (Raltegravir, e.g., ISENTRESS, MK-0518; and elvitegravir, e.g., VITEKTA, GS-9137), and maturation inhibitors (bevirimat {PA-457}). As another example, formulations can additionally contain one or more supplements, such as vitamin C, E or other anti-oxidants.

[0204] The agent(s) (or pharmaceutically acceptable salts, esters or amides thereof) can be administered per se or in the form of a pharmaceutical composition wherein the active agent(s) is in an admixture or mixture with one or more pharmaceutically acceptable carriers. A pharmaceutical composition, as used herein, can be any composition prepared for administration to a subject. Pharmaceutical compositions for use in accordance with the present disclosure can be formulated in conventional manner using one or more physiologically acceptable carriers, comprising excipients, diluents, and/or auxiliaries, e.g., which facilitate processing of the active agents into preparations that can be administered. Proper formulation can depend at least in part upon the route of administration chosen. The

agent(s) useful in the present disclosure, or pharmaceutically acceptable salts, esters, or amides thereof, can be delivered to a subject using a number of routes or modes of administration, including oral, buccal, topical, rectal, transdermal, transmucosal, subcutaneous, intravenous, and intramuscular applications, as well as by inhalation.

[0205] For oral administration, the agents can be formulated readily by combining the active agent(s) with pharmaceutically acceptable carriers well known in the art. Such carriers enable the agents of the disclosure to be formulated as tablets, including chewable tablets, pills, dragees, capsules, lozenges, hard candy, liquids, gels, syrups, slurries, powders, suspensions, elixirs, wafers, and the like, for oral ingestion by a subject to be treated. Such formulations can comprise pharmaceutically acceptable carriers including solid diluents or fillers, sterile aqueous media and various non-toxic organic solvents. A solid carrier can be one or more substances which can also act as diluents, flavoring agents, solubilizers, lubricants, suspending agents, binders, preservatives, tablet disintegrating agents, or an encapsulating material. In powders, the carrier generally is a finely divided solid which is a mixture with the finely divided active component. In tablets, the active component generally is mixed with the carrier having the necessary binding capacity in suitable proportions and compacted in the shape and size desired. The powders and tablets preferably contain from about one (1) to about seventy (70) percent of the active compound. Suitable carriers include but are not limited to magnesium carbonate, magnesium stearate, talc, sugar, lactose, pectin, dextrin, starch, gelatin, tragacanth, methylcellulose, sodium carboxymethylcellulose, a low melting wax, cocoa butter, and the like. Generally, the agents of the disclosure can be included at concentration levels ranging from about 0.5%, about 5%, about 10%, about 20%, or about 30% to about 50%, about 60%, about 70%, about 80% or about 90% by weight of the total composition of oral dosage forms, in an amount sufficient to provide a desired unit of dosage.

[0206] Aqueous suspensions for oral use can contain agent(s) of this disclosure with pharmaceutically acceptable excipients, such as a suspending agent (e.g., methyl cellulose), a wetting agent (e.g., lecithin, lysolecithin and/or a long-chain fatty alcohol), as well as coloring agents, preservatives, flavoring agents, and the like.

[0207] Oils or non-aqueous solvents can be used to bring the agents into solution, due to, for example, the presence of large lipophilic moieties. Alternatively, emulsions, suspensions, or other preparations, for example, liposomal preparations, can be used. With respect to liposomal preparations, any known methods for preparing liposomes for treatment of a condition can be used. See, for example, Bangham et al., J. Mol. Biol. 23: 238-252 (1965) and Szoka et al., Proc. Natl Acad. Sci. USA 75: 4194-4198 (1978). Ligands can also be attached to the liposomes to direct these compositions to particular sites of action. Agents of this disclosure can also be integrated into foodstuffs, e.g., cream cheese, butter, salad dressing, or ice cream to facilitate solubilization, administration, and/or compliance in certain subject populations.

[0208] Pharmaceutical preparations for oral use can be obtained as a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; flavoring elements, cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl cellulose, sodium carboxymethylcellulose, and/or polyvinyl pyrrolidone (PVP). If desired, disintegrating agents can be added, such as the cross linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. The agents can also be formulated as a sustained release preparation.

[0209] Dragee cores can be provided with suitable coatings. For this purpose, concentrated sugar solutions can be used, which can optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments can be added to the tablets or dragee coatings for identification or to characterize different combinations of active agents.

[0210] Pharmaceutical preparations that can be used orally include push fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active agents can be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers can be added. All formulations for oral administration should be in dosages suitable for administration.

[0211] Other forms suitable for oral administration include liquid form preparations including emulsions, syrups, elixirs, aqueous solutions, aqueous suspensions, or solid form preparations which are intended to be converted shortly before use to liquid form preparations. Emulsions can be prepared in solutions, for example, in aqueous propylene glycol solutions or can contain emulsifying agents, for example, such as lecithin, sorbitan monooleate, or acacia. Aqueous solutions can be prepared by dissolving the active component in water and adding suitable colorants, flavors, stabilizers, and thickening agents. Aqueous suspensions can be prepared by dispersing the finely divided active component in water with viscous material, such as natural or synthetic gums, resins, methylcellulose, sodium carboxymethylcellulose, and other well known suspending agents. Suitable fillers or carriers with which the compositions can be administered include agar, alcohol, fats, lactose, starch, cellulose derivatives, polysaccharides, polyvinylpyrrolidone, silica, sterile saline and the like, or mixtures thereof used in suitable amounts. Solid form preparations include solutions, suspensions, and emulsions, and can contain, in addition to the active component, colorants, flavors, stabilizers, buffers, artificial and natural sweeteners, dispersants, thickeners, solubilizing agents, and the like.

[0212] A syrup or suspension can be made by adding the active compound to a concentrated, aqueous solution of a sugar, e.g., sucrose, to which can also be added any accessory ingredients. Such accessory ingredients can include flavoring, an agent to retard crystallization

of the sugar or an agent to increase the solubility of any other ingredient, e.g., as a polyhydric alcohol, for example, glycerol or sorbitol.

[0213] When formulating compounds of the disclosure for oral administration, it can be desirable to utilize gastroretentive formulations to enhance absorption from the gastrointestinal (GI) tract. A formulation which is retained in the stomach for several hours can release compounds of the disclosure slowly and provide a sustained release that can be preferred. Disclosure of such gastro-retentive formulations are found in 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, and Chavanpatil, M.D. et al., Int. J. Pharm. (2006). Expandable, floating and bioadhesive techniques can be utilized to maximize absorption of the compounds of the disclosure.

[0214] The compounds of the disclosure can be formulated for parenteral administration (e.g., by injection, for example, bolus injection or continuous infusion) and can be presented in unit dose form in ampoules, pre-filled syringes, small volume infusion or in multi-dose containers with an added preservative. The compositions can take such forms as suspensions, solutions, or emulsions in oily or aqueous vehicles, for example, solutions in aqueous polyethylene glycol.

[0215] For injectable formulations, the vehicle can be chosen from those known in art to be suitable, including aqueous solutions or oil suspensions, or emulsions, with sesame oil, corn oil, cottonseed oil, or peanut oil, as well as elixirs, mannitol, dextrose, or a sterile aqueous solution, and similar pharmaceutical vehicles. The formulation can also comprise polymer compositions which are biocompatible, biodegradable, such as poly(lactic-co-glycolic)acid. These materials can be made into micro or nanospheres, loaded with drug and further coated or derivatized to provide superior sustained release performance. Vehicles suitable for periocular or intraocular injection include, for example, suspensions of therapeutic agent in injection grade water, liposomes and vehicles suitable for lipophilic substances. Other vehicles for periocular or intraocular injection are well known in the art.

[0216] The composition may be formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition can also include a solubilizing agent and a local anesthetic such as lidocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients can be mixed prior to administration.

[0217] When administration is by injection, the active compound can be formulated in aqueous solutions, specifically in physiologically compatible buffers such as Hanks solution, Ringer's solution, or physiological saline buffer. The solution can contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active compound can be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use. The pharmaceutical composition may not comprise an adjuvant or any other substance added to enhance the immune response stimulated by the peptide. the pharmaceutical composition may comprise a substance that inhibits an immune response to the peptide. Methods of formulation are known in the art, for example, as disclosed in Remington's Pharmaceutical Sciences, latest edition, Mack Publishing Co., Easton P.

[0218] In addition to the formulations described previously, the agents can also be formulated as a depot preparation. Such long acting formulations can be administered by implantation or transcutaneous delivery (for example, subcutaneously or intramuscularly), intramuscular injection or use of a transdermal patch. Thus, for example, the agents can be formulated with suitable polymeric or hydrophobic materials (for example, as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

[0219] pharmaceutical compositions comprising one or more agents of the present disclosure may exert local and regional effects when administered topically or injected at or near particular sites of infection. Direct topical application, e.g., of a viscous liquid, solution, suspension, dimethylsulfoxide (DMSO)-based solutions, liposomal formulations, gel, jelly, cream, lotion, ointment, suppository, foam, or aerosol spray, can be used for local administration, to produce for example, local and/or regional effects. Pharmaceutically appropriate vehicles for such formulation include, for example, lower aliphatic alcohols, polyglycols (e.g., glycerol or polyethylene glycol), esters of fatty acids, oils, fats, silicones, and the like. Such preparations can also include preservatives (e.g., p-hydroxybenzoic acid esters) and/or antioxidants (e.g., ascorbic acid and tocopherol). See also Dermatological Formulations: Percutaneous absorption, Barry (Ed.), Marcel Dekker Incl, 1983.

[0220] Pharmaceutical compositions of the present disclosure can contain a cosmetically or dermatologically acceptable carrier. Such carriers are compatible with skin, nails, mucous membranes, tissues and/or hair, and can include any conventionally used cosmetic or dermatological carrier meeting these requirements. Such carriers can be readily selected by one of ordinary skill in the art. In formulating skin ointments, an agent or combination of agents of the instant disclosure can be formulated in an oleaginous hydrocarbon base, an anhydrous absorption base, a water-in-oil absorption base, an oil-in-water water-removable base and/or a water-soluble base. Examples of such carriers and excipients include, but are not limited to, humectants (e.g., urea), glycols (e.g., propylene glycol), alcohols (e.g., ethanol), fatty acids (e.g., oleic acid), surfactants (e.g., isopropyl myristate and sodium lauryl sulfate), pyrrolidones, glycerol monolaurate, sulfoxides, terpenes (e.g., menthol), amines, amides, alkanes, alkanols, water, calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols.

[0221] Ointments and creams can, for example, be formulated with an aqueous or oily base with the addition of suitable thickening and/or gelling agents. Lotions can be formulated with an aqueous or oily base and can in general also containing one or more emulsifying agents, stabilizing agents, dispersing agents, suspending agents, thickening agents, or coloring agents. The construction and use of transdermal patches for the delivery of pharmaceutical agents is well known in the art. See, e.g., U.S. Pat. Nos. 5,023,252, 4,992,445 and 5,001,139. Such patches can be constructed for continuous, pulsatile, or on demand delivery of pharmaceutical agents.

[0222] Lubricants which can be used to form pharmaceutical compositions and dosage forms of the disclosure include, but are not limited to, calcium stearate, magnesium stearate, mineral oil, light mineral oil, glycerin, sorbitol, mannitol, polyethylene glycol, other glycols, stearic acid, sodium lauryl sulfate, talc, hydrogenated vegetable oil (e.g., peanut oil, cottonseed oil, sunflower oil, sesame oil, olive oil, corn oil, and soybean oil), zinc stearate, ethyl oleate, ethyl laureate, agar, or mixtures thereof. Additional lubricants include, for example, a syloid silica gel, a coagulated aerosol of synthetic silica, or mixtures thereof. A lubricant can optionally be added, in an amount of less than about 1 weight percent of the pharmaceutical composition.

[0223] The compositions according to the present disclosure can be in any form suitable for topical application, including aqueous, aqueous-alcoholic or oily solutions, lotion or serum dispersions, aqueous, anhydrous or oily gels, emulsions obtained by dispersion of a fatty phase in an aqueous phase (O/W or oil in water) or, conversely, (W/O or water in oil), microemulsions or alternatively microcapsules, microparticles or lipid vesicle dispersions of ionic and/or nonionic type. These compositions can be prepared according to conventional methods. Other than the agents of the disclosure, the amounts of the various constituents of the compositions according to the disclosure are those conventionally used in the art. These compositions in particular constitute protection, treatment or care creams, milks, lotions, gels or foams for the face, for the hands, for the body and/or for the mucous membranes, or for cleansing the skin. The compositions can also consist of solid preparations constituting soaps or cleansing bars.

[0224] Compositions of the present disclosure can also contain adjuvants common to the cosmetic and dermatological fields, such as hydrophilic or lipophilic gelling agents, hydrophilic or lipophilic active agents, preserving agents, antioxidants, solvents, fragrances, fillers, sunscreens, odor-absorbers and dyestuffs. The amounts of these various adjuvants are those conventionally used in the fields considered and, for example, are from about 0.01% to about 20% of the total weight of the composition. Depending on their nature, these adjuvants can be introduced into the fatty phase, into the aqueous phase and/or into the lipid vesicles.

[0225] Ocular viral infections can be effectively treated with ophthalmic solutions, suspensions, ointments or inserts comprising an agent or combination of agents of the present disclosure. Eye drops can be prepared by dissolving the active ingredient in a sterile aqueous solution such as physiological saline, buffering solution, etc., or by combining powder compositions to be dissolved before use. Other vehicles can be chosen, as is known in the art, including but not limited to: balance salt solution, saline solution, water soluble polyethers such as polyethyene glycol, polyvinyls, such as polyvinyl alcohol and povidone, cellulose derivatives such as methylcellulose and hydroxypropyl methylcellulose, petroleum derivatives such as mineral oil and white petrolatum, animal fats such as lanolin, polymers of acrylic acid such as carboxypolymethylene gel, vegetable fats such as peanut oil and polysaccharides such as dextrans, and glycosaminoglycans such as sodium hyaluronate. If desired, additives ordinarily used in the eye drops can be added. Such additives include isotonizing agents (e.g., sodium chloride, etc.), buffer agent (e.g., boric acid, sodium monohydrogen phosphate, sodium dihydrogen phosphate, etc.), preservatives (e.g., benzalkonium chloride, benzethonium chloride, chlorobutanol, etc.), thickeners (e.g., saccharide such as lactose, mannitol, maltose, etc.; e.g., hyaluronic acid or its salt such as sodium hyaluronate, potassium hyaluronate, etc.; e.g., mucopolysaccharide such as chondroitin sulfate, etc.; e.g., sodium polyacrylate, carboxyvinyl polymer, crosslinked polyacrylate, polyvinyl alcohol, polyvinyl pyrrolidone, methyl cellulose, hydroxy propyl methylcellulose, hydroxyethyl cellulose, carboxymethyl cellulose, hydroxy propyl cellulose or other agents known to those skilled in the art).

[0226] The solubility of the components of the present compositions can be enhanced by a surfactant or other appropriate co-solvent in the composition. Such cosolvents include polysorbate 20, 60, and 80, Pluronic F68, F-84 and P-103, cyclodextrin, or other agents known to those skilled in the art. Such cosolvents can be employed at a level of from about 0.01% to 2% by weight.

[0227] The compositions of the disclosure can be packaged in multidose form. Preservatives can be preferred to prevent microbial contamination during use. Suitable preservatives include: benzalkonium chloride, thimerosal, chlorobutanol, methyl paraben, propyl paraben, phenylethyl alcohol, edetate disodium, sorbic acid, Onamer M, or other agents known to those skilled in the art. In the prior art ophthalmic products, such preservatives can be employed at a level of from 0.004% to 0.02%. In the compositions of the present application the preservative, preferably benzalkonium chloride, can be employed at a level of from 0.001% to less than 0.01%, e.g., from 0.001% to 0.008%, preferably about 0.005% by weight. It has been found that a concentration of benzalkonium chloride of 0.005% can be sufficient to preserve the compositions of the present disclosure from microbial attack.

[0228] The agents of the present disclosure may be delivered in soluble rather than suspension form, which allows for more rapid and quantitative absorption to the sites of action. In general, formulations such as jellies, creams, lotions, suppositories and ointments can provide an area with more extended exposure to the agents of the present disclosure, while formulations in solution, e.g., sprays, provide more immediate, short-term exposure.

[0229] the pharmaceutical compositions that relate to topical/local application can include one or more penetration enhancers. For example, the formulations can comprise suitable solid or gel phase carriers or excipients that increase penetration or help delivery of agents or combinations of agents of the disclosure across a permeability barrier, e.g., the skin. Many of these penetration-enhancing compounds are known in the art of topical formulation, and include, e.g., water, alcohols (e.g., terpenes like methanol, ethanol, 2-

propanol), sulfoxides (e.g., dimethyl sulfoxide, decylmethyl sulfoxide, tetradecylmethyl sulfoxide), pyrrolidones (e.g., 2-pyrrolidone, N-methyl-2-pyrrolidone, N-(2-hydroxyethyl)pyrrolidone), laurocapram, acetone, dimethylacetamide, dimethylformamide, tetrahydrofurfuryl alcohol, L-α-amino acids, anionic, cationic, amphoteric or nonionic surfactants (e.g., isopropyl myristate and sodium lauryl sulfate), fatty acids, fatty alcohols (e.g., oleic acid), amines, amides, clofibric acid amides, hexamethylene lauramide, proteolytic enzymes, α-bisabolol, d-limonene, urea and N,N-diethyl-m-toluamide, and the like. Additional examples include humectants (e.g., urea), glycols (e.g., propylene glycol and polyethylene glycol), glycerol monolaurate, alkanes, alkanols, ORGELASE, calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and/or other polymers. The pharmaceutical compositions can include one or more such penetration enhancers.

[0230] The pharmaceutical compositions for local/topical application can include one or more antimicrobial preservatives such as quaternary ammonium compounds, organic mercurials, p-hydroxy benzoates, aromatic alcohols, chlorobutanol, and the like.

[0231] The pharmaceutical compositions can be orally- or rectally delivered solutions, suspensions, ointments, enemas and/or suppositories comprising an agent or combination of agents of the present disclosure.

[0232] The pharmaceutical compositions can be aerosol solutions, suspensions or dry powders comprising an agent or combination of agents of the present disclosure. The aerosol can be administered through the respiratory system or nasal passages. For example, one skilled in the art can recognize that a composition of the present disclosure can be suspended or dissolved in an appropriate carrier, e.g., a pharmaceutically acceptable propellant, and administered directly into the lungs using a nasal spray or inhalant. For example, an aerosol formulation comprising an agent can be dissolved, suspended or emulsified in a propellant or a mixture of solvent and propellant, e.g., for administration as a nasal spray or inhalant. Aerosol formulations can contain any acceptable propellant under pressure, such as a cosmetically or dermatologically or pharmaceutically acceptable propellant, as conventionally used in the art.

[0233] An aerosol formulation for nasal administration is generally an aqueous solution designed to be administered to the nasal passages in drops or sprays. Nasal solutions can be similar to nasal secretions in that they are generally isotonic and slightly buffered to maintain a pH of about 5.5 to about 6.5, although pH values outside of this range can additionally be used. Antimicrobial agents or preservatives can also be included in the formulation.

[0234] An aerosol formulation for inhalations and inhalants can be designed so that the agent or combination of agents of the present disclosure is carried into the respiratory tree of the subject when administered by the nasal or oral respiratory route. Inhalation solutions can be administered, for example, by a nebulizer. Inhalations or insufflations, comprising finely powdered or liquid drugs, can be delivered to the respiratory system as a pharmaceutical aerosol of a solution or suspension of the agent or combination of agents in a propellant, e.g., to aid in disbursement. Propellants can be liquefied gases, including halocarbons, for example, fluorocarbons such as fluorinated chlorinated hydrocarbons, hydrochlorofluorocarbons, and hydrochlorocarbons, as well as hydrocarbons and hydrocarbon ethers.

[0235] Halocarbon propellants useful in the present disclosure include fluorocarbon propellants in which all hydrogens are replaced with fluorine, chlorofluorocarbon propellants in which all hydrogens are replaced with chlorine and at least one fluorine, hydrogen-containing fluorocarbon propellants, and hydrogen-containing chlorofluorocarbon propellants. Halocarbon propellants are described in Johnson, U.S. Pat. No. 5,376,359; Byron et al., U.S. Pat. No. 5,190,029; and Purewal et al., U.S. Pat. No. 5,776,434. Hydrocarbon propellants useful in the disclosure include, for example, propane, isobutane, n-butane, pentane, isopentane and neopentane. A blend of hydrocarbons can also be used as a propellant. Ether propellants include, for example, dimethyl ether as well as the ethers. An aerosol formulation of the disclosure can also comprise more than one propellant. For example, the aerosol formulation can comprise more than one propellant from the same class, such as two or more fluorocarbons; or more than one, more than two, more than three propellants from different classes, such as a fluorohydrocarbon and a hydrocarbon. Pharmaceutical compositions of the present disclosure can also be dispensed with a compressed gas, e.g., an inert gas such as carbon dioxide, nitrous oxide or nitrogen.

[0236] Aerosol formulations can also include other components, for example, ethanol, isopropanol, propylene glycol, as well as surfactants or other components such as oils and detergents. These components can serve to stabilize the formulation and/or lubricate valve components.

[0237] The aerosol formulation can be packaged under pressure and can be formulated as an aerosol using solutions, suspensions, emulsions, powders and semisolid preparations. For example, a solution aerosol formulation can comprise a solution of an agent of the disclosure in (substantially) pure propellant or as a mixture of propellant and solvent. The solvent can be used to dissolve the agent and/or retard the evaporation of the propellant. Solvents useful in the disclosure include, for example, water, ethanol and glycols. Any combination of suitable solvents can be use, optionally combined with preservatives, antioxidants, and/or other aerosol components.

[0238] An aerosol formulation can also be a dispersion or suspension. A suspension aerosol formulation can comprise a suspension of an agent or combination of agents of the instant disclosure. Dispersing agents useful in the disclosure include, for example, sorbitan trioleate, oleyl alcohol, oleic acid, lecithin and corn oil. A suspension aerosol formulation can also include lubricants, preservatives, antioxidant, and/or other aerosol components.

[0239] An aerosol formulation can similarly be formulated as an emulsion. An emulsion aerosol formulation can include, for example, an alcohol such as ethanol, a surfactant, water and a propellant, as well as an agent or combination of agents of the disclosure. The surfactant used can be nonionic, anionic or cationic. One example of an emulsion aerosol formulation comprises, for example, ethanol,

surfactant, water and propellant. Another example of an emulsion aerosol formulation comprises, for example, vegetable oil, glyceryl monostearate and propane.

[0240] The compounds of the disclosure can be formulated for administration as suppositories. A low melting wax, such as a mixture of triglycerides, fatty acid glycerides, Witepsol S55 (trademark of Dynamite Nobel Chemical, Germany), or cocoa butter is first melted and the active component is dispersed homogeneously, for example, by stirring. The molten homogeneous mixture is then poured into convenient sized molds, allowed to cool, and to solidify.

[0241] The compounds of the disclosure can be formulated for vaginal administration. Pessaries, tampons, creams, gels, pastes, foams or sprays containing in addition to the active ingredient such carriers as are known in the art to be appropriate.

[0242] It is envisioned additionally, that the compounds of the disclosure can be attached releasably to biocompatible polymers for use in sustained release formulations on, in or attached to inserts for topical, intraocular, periocular, or systemic administration. The controlled release from a biocompatible polymer can be utilized with a water soluble polymer to form an instillable formulation, as well. The controlled release from a biocompatible polymer, such as for example, PLGA microspheres or nanospheres, can be utilized in a formulation suitable for intra ocular implantation or injection for sustained release administration, as well any suitable biodegradable and biocompatible polymer can be used.

[0243] In one aspect of the disclosure, the subject's carrier status of any of the genetic variation risk variants described herein, or genetic variants identified via other analysis methods within the genes or regulatory loci that are identified by the CNVs or SNVs described herein, can be used to help determine whether a particular treatment modality, such as any one of the above, or a combination thereof, should be administered. Whether a treatment option such as any of the above mentioned treatment options is administered can be determined based on the presence or absence of a particular genetic variation risk variant in the individual, or by monitoring expression of genes that are associated with the variants of the present disclosure. Expression levels and/or mRNA levels can thus be determined before and during treatment to monitor its effectiveness. Alternatively, or concomitantly, the status with respect to a genetic variation, and or genotype and/or haplotype status of at least one risk variant for PML presented herein can be determined before and during treatment to monitor its effectiveness. It can also be appreciated by those skilled in the art that aberrant expression levels of a gene impacted by a CNV or other mutations found as a consequence of targeted sequencing of the CNV-identified gene can be assayed or diagnostically tested for by measuring the polypeptide expression level of said aberrantly expressed gene. Aberrant expression levels of a gene may result from a CNV impacting a DNA sequence (e.g., transcription factor binding site) that regulates a gene whose aberrant expression level is involved in or causes PML, or other mutations found as a consequence of targeted sequencing of the CNV-identified gene regulatory sequence, can be assayed or diagnostically tested for by measuring the polypeptide expression level of the gene involved in or causative of PML. A specific CNV mutation within a gene, or other specific mutations found upon targeted sequencing of a CNV-identified gene found to be involved in or causative of PML, may cause an aberrant structural change in the expressed polypeptide that results from said gene mutations and the altered polypeptide structure(s) can be assayed via various methods know to those skilled in the art.

[0244] Alternatively, biological networks or metabolic pathways related to the genes within, or associated with, the genetic variations described herein can be monitored by determining mRNA and/or polypeptide levels. This can be done for example, by monitoring expression levels of polypeptides for several genes belonging to the network and/or pathway in nucleic acid samples taken before and during treatment. Alternatively, metabolites belonging to the biological network or metabolic pathway can be determined before and during treatment. Effectiveness of the treatment is determined by comparing observed changes in expression levels/metabolite levels during treatment to corresponding data from healthy subjects.

[0245] The genetic variations described herein and/or those subsequently found (e.g., via other genetic analysis methods such as sequencing) via targeted analysis of those genes initially identified by the genetic variations described herein, can be used to prevent adverse effects associated with a therapeutic agent, such as during clinical trials. For example, individuals who are carriers of at least one at-risk genetic variation can be more likely to respond negatively to a therapeutic agent, such as an immunosuppressive agent. For example, carriers of certain genetic variants may be more likely to show an adverse response to the therapeutic agent. One or more of the genetic variations employed during clinical trials for a given therapeutic agent can be used in a companion diagnostic test that is administered to the patient prior to administration of the therapeutic agent to determine if the patient is likely to have a favorable or an adverse response to the therapeutic agent.

[0246] The genetic variations described herein can be used for determining whether a subject is administered a pharmaceutical agent, such as an immunosuppressive drug. Certain combinations of variants, including those described herein, but also combinations with other risk variants for PML, can be suitable for one selection of treatment options, while other variant combinations can be suitable for selection of other treatment options. Such combinations of variants can include one variant, two variants, three variants, or four or more variants, as needed to determine with clinically reliable accuracy the selection of treatment module. Information from testing for the genetic variations described herein, or other rare genetic variations in or near the genes described herein, may be combined with information from other types of testing (e.g., a JCV antibody test, CD62L test, or CSF IgM oligoclonal bands test) for selection of treatment options.

Kits

[0247] The present invention relates to an immunosuppressive medicament for use or methods as defined in the claims, for which the

following provides useful disclosure relating to kits (in that, for instance, immunosuppressive medicament for use may be comprised in a kit, or reagents for carrying out the claimed method may be provided in a kit), but such kits do not form the subject matter being claimed herein. Kits useful in the methods of the disclosure comprise components useful in any of the methods described herein, including for example, primers for nucleic acid amplification, hybridization probes for detecting genetic variation, or other marker detection, restriction enzymes, nucleic acid probes, optionally labeled with suitable labels, allele-specific oligonucleotides, antibodies that bind to an altered polypeptide encoded by a nucleic acid of the disclosure as described herein or to a wild type polypeptide encoded by a nucleic acid of the disclosure as described herein, means for amplification of genetic variations or fragments thereof, means for analyzing the nucleic acid sequence of nucleic acids comprising genetic variations as described herein, means for analyzing the amino acid sequence of a polypeptide encoded by a genetic variation, or a nucleic acid associated with a genetic variation, etc. The kits can for example, include necessary buffers, nucleic acid primers for amplifying nucleic acids, and reagents for allele-specific detection of the fragments amplified using such primers and necessary enzymes (e.g., DNA polymerase). Additionally, kits can provide reagents for assays to be used in combination with the methods of the present disclosure, for example, reagents for use with other screening assays for PML.

[0248] The disclosure pertains to a kit for assaying a nucleic acid sample from a subject to detect the presence of a genetic variation, wherein the kit comprises reagents necessary for selectively detecting at least one particular genetic variation in the genome of the individual. The disclosure pertains to a kit for assaying a nucleic acid sample from a subject to detect the presence of at least one particular allele of at least one polymorphism associated with a genetic variation in the genome of the subject. The reagents may comprise at least one contiguous oligonucleotide that hybridizes to a fragment of the genome of the individual comprising at least genetic variation. The reagents may comprise at least one pair of oligonucleotides that hybridize to opposite strands of a genomic segment obtained from a subject, wherein each oligonucleotide primer pair is designed to selectively amplify a fragment of the genome of the individual that includes at least one genetic variation, or a fragment of a genetic variation. Such oligonucleotides or nucleic acids can be designed using the methods described herein. The kit may comprise one or more labeled nucleic acids capable of allele-specific detection of one or more specific polymorphic markers or haplotypes with a genetic variation, and reagents for detection of the label. A kit for detecting SNP markers can comprise a detection oligonucleotide probe, that hybridizes to a segment of template DNA containing a SNP polymorphism to be detected, an enhancer oligonucleotide probe, detection probe, primer and/or an endonuclease, for example, as described by Kutyavin et al., (Nucleic Acid Res. 34:e128 (2006)). The kit can contain reagents for detecting SNVs and/or CNVs.

[0249] The DNA template may be amplified by any means of the present disclosure, prior to assessment for the presence of specific genetic variations as described herein. Standard methods well known to the skilled person for performing these methods can be utilized, and are within scope of the disclosure. Reagents for performing these methods can be included in the reagent kit.

[0250] In a further aspect of the present disclosure, a pharmaceutical pack (kit) is provided, the pack comprising a therapeutic agent and a set of instructions for administration of the therapeutic agent to humans screened for one or more variants of the present disclosure, as disclosed herein. The therapeutic agent can be a small molecule drug, an antibody, a peptide, an antisense or RNAi molecule, or other therapeutic molecules as described herein. An individual identified as a non-carrier of at least one variant of the present disclosure may be instructed to take the therapeutic agent. An individual identified as a non-carrier of at least one variant of the present disclosure may be instructed to take a prescribed dose of the therapeutic agent. An individual identified as a carrier of at least one variant of the present disclosure may be instructed not to take the therapeutic agent. An individual identified as a carrier of at least one variant of the present disclosure may be instructed not to take a prescribed dose of the therapeutic agent. An individual identified as a carrier of at least one variant of the present disclosure may be instructed to take an agent that targets the JC Virus. For example, an individual identified as a carrier of at least one variant of the present disclosure can be instructed to take an agent that targets the JC Virus prior to or in conjunction with, taking an immunosuppressive agent.

[0251] Also provided herein are articles of manufacture, comprising a probe that hybridizes with a region of human chromosome as described herein and can be used to detect a polymorphism described herein. For example, any of the probes for detecting polymorphisms or genetic variations described herein can be combined with packaging material to generate articles of manufacture or kits. The kit can include one or more other elements including: instructions for use; and other reagents such as a label or an agent useful for attaching a label to the probe. Instructions for use can include instructions for screening applications of the probe for making a diagnosis, prognosis, or theranosis to PML in a method described herein. Other instructions can include instructions for attaching a label to the probe, instructions for performing in situ analysis with the probe, and/or instructions for obtaining a nucleic acid sample to be analyzed from a subject. In some cases, the kit can include a labeled probe that hybridizes to a region of human chromosome as described herein.

[0252] The kit can also include one or more additional reference or control probes that hybridize to the same chromosome or another chromosome or portion thereof that can have an abnormality associated with a particular endophenotype. A kit that includes additional probes can further include labels, e.g., one or more of the same or different labels for the probes. The additional probe or probes provided with the kit can be a labeled probe or probes. When the kit further includes one or more additional probe or probes, the kit can further provide instructions for the use of the additional probe or probes. Kits for use in self-testing can also be provided. Such test kits can include devices and instructions that a subject can use to obtain a nucleic acid sample (e.g., buccal cells, blood) without the aid of a health care provider. For example, buccal cells can be obtained using a buccal swab or brush, or using mouthwash.

[0253] Kits as provided herein can also include a mailer (e.g., a postage paid envelope or mailing pack) that can be used to return the nucleic acid sample for analysis, e.g., to a laboratory. The kit can include one or more containers for the nucleic acid sample, or the nucleic acid sample can be in a standard blood collection vial. The kit can also include one or more of an informed consent form, a test

requisition form, and instructions on how to use the kit in a method described herein. Methods for using such kits are also included herein. One or more of the forms (e.g., the test requisition form) and the container holding the nucleic acid sample can be coded, for example, with a bar code for identifying the subject who provided the nucleic acid sample.

[0254] An in vitro screening test can comprise one or more devices, tools, and equipment configured to collect a nucleic acid sample from an individual. An in vitro screening test, tools to collect a nucleic acid sample can include one or more of a swab, a scalpel, a syringe, a scraper, a container, and other devices and reagents designed to facilitate the collection, storage, and transport of a nucleic acid sample. An in vitro screening test can include reagents or solutions for collecting, stabilizing, storing, and processing a nucleic acid sample.

[0255] Such reagents and solutions for nucleotide collecting, stabilizing, storing, and processing are well known by those of skill in the art and can be indicated by specific methods used by an in vitro screening test as described herein. An in vitro screening test as disclosed herein, can comprise a microarray apparatus and reagents, a flow cell apparatus and reagents, a multiplex nucleotide sequencer and reagents, and additional hardware and software necessary to assay a nucleic acid sample for certain genetic markers and to detect and visualize certain genetic markers.

[0256] The present disclosure further relates to kits for using antibodies in the methods described herein. This includes, but is not limited to, kits for detecting the presence of a variant polypeptide in a test nucleic acid sample. A preferred kit comprises antibodies such as a labeled or labelable antibody and a compound or agent for detecting variant polypeptides in a nucleic acid sample, means for determining the amount or the presence and/or absence of variant polypeptide in the nucleic acid sample, and means for comparing the amount of variant polypeptide in the nucleic acid sample with a standard, as well as instructions for use of the kit. The kit may further comprise a set of instructions for using the reagents comprising the kit.

Computer-Implemented Aspects

[0257] The present invention relates to an immunosuppressive medicament for use or methods as defined in the claims, for which the following provides useful disclosure relating to computer-implemented aspects (which may, for instance, be used to determine the risk of PML), but such computer-implemented aspects do not form the subject matter being claimed herein. As understood by those of ordinary skill in the art, the methods and information described herein (genetic variation association with PML) can be implemented, in all or in part, as computer executable instructions on known computer readable media. For example, the methods described herein can be implemented in hardware. Alternatively, the method can be implemented in software stored in, for example, one or more memories or other computer readable medium and implemented on one or more processors. As is known, the processors can be associated with one or more controllers, calculation units and/or other units of a computer system, or implanted in firmware as desired. If implemented in software, the routines can be stored in any computer readable memory such as in RAM, ROM, flash memory, a magnetic disk, a laser disk, or other storage medium, as is also known. Likewise, this software can be delivered to a computing device via any known delivery method including, for example, over a communication channel such as a telephone line, the Internet, a wireless connection, etc., or via a transportable medium, such as a computer readable disk, flash drive, etc.

[0258] More generally, and as understood by those of ordinary skill in the art, the various steps described above can be implemented as various blocks, operations, tools, modules and techniques which, in turn, can be implemented in hardware, firmware, software, or any combination of hardware, firmware, and/or software. When implemented in hardware, some or all of the blocks, operations, techniques, etc. can be implemented in, for example, a custom integrated circuit (IC), an application specific integrated circuit (ASIC), a field programmable logic array (FPGA), a programmable logic array (PLA), etc.

[0259] Results from such genotyping can be stored in a data storage unit, such as a data carrier, including computer databases, data storage disks, or by other convenient data storage means. The computer database may be an object database, a relational database or a post- relational database. Data can be retrieved from the data storage unit using any convenient data query method.

[0260] When implemented in software, the software can be stored in any known computer readable medium such as on a magnetic disk, an optical disk, or other storage medium, in a RAM or ROM or flash memory of a computer, processor, hard disk drive, optical disk drive, tape drive, etc. Likewise, the software can be delivered to a user or a computing system via any known delivery method including, for example, on a computer readable disk or other transportable computer storage mechanism.

[0261] The steps of the claimed methods can be operational with numerous other general purpose or special purpose computing system environments or configurations. Examples of well known computing systems, environments, and/or configurations that can be suitable for use with the methods or system of the claims include, but are not limited to, personal computers, server computers, hand-held or laptop devices, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

[0262] The steps of the claimed method and system can be described in the general context of computer-executable instructions, such as program modules, being executed by a computer. Generally, program modules include routines, programs, objects, components, and/or data structures that perform particular tasks or implement particular abstract data types. The methods and apparatus can also be practiced in distributed computing environments where tasks are performed by remote processing devices that are linked through a communications network. In both integrated and distributed computing environments, program modules can be located in both local and

remote computer storage media including memory storage devices.

[0263] While the risk evaluation system and method, and other elements, have been described as preferably being implemented in software, they can be implemented in hardware, firmware, etc., and can be implemented by any other processor. Thus, the elements described herein can be implemented in a standard multi-purpose CPU or on specifically designed hardware or firmware such as an application-specific integrated circuit (ASIC) or other hard-wired device as desired. When implemented in software, the software routine can be stored in any computer readable memory such as on a magnetic disk, a laser disk, or other storage medium, in a RAM or ROM of a computer or processor, in any database, etc. Likewise, this software can be delivered to a user or a screening system via any known or desired delivery method including, for example, on a computer readable disk or other transportable computer storage mechanism or over a communication channel, for example, a telephone line, the internet, or wireless communication. Modifications and variations can be made in the techniques and structures described and illustrated herein without departing from the spirit and scope of the present disclosure.

[0264] Unless otherwise explained, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. The following references contain teaching of the methods and compositions that can be used herein: The Merck Manual of Diagnosis and Therapy, 18th Edition, published by Merck Research Laboratories, 2006 (ISBN 0-911910-18-2); Benjamin Lewin, Genes IX, published by Jones & Bartlett Publishing, 2007 (ISBN-13: 9780763740634); Kendrew et al., (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0-632-02182-9); and Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk Reference, published by VCH Publishers, Inc., 1995 (ISBN 1-56081-569-8).

[0265] Standard procedures of the present disclosure are described, e.g., in 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., New York, USA (1986); or Methods in Enzymology: Guide to Molecular Cloning Techniques Vol. 152, S. L. Berger and 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.), 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 by R. Ian Freshney, Publisher: Wiley-Liss; 5th edition (2005), and Animal Cell Culture Methods (Methods in Cell Biology, Vol. 57, Jennie P. Mather and David Barnes editors, Academic Press, 1st edition, 1998).

[0266] Unless defined otherwise, all technical and scientific terms used herein have the meanings that would be commonly understood by one of skill in the art in the context of the present specification.

[0267] It should be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, reference to "a nucleotide" includes a plurality of such nucleotides; reference to "the nucleotide" is a reference to one or more nucleotides and equivalents thereof known to those skilled in the art, and so forth.

[0268] The term "and/or" shall in the present context be understood to indicate that either or both of the items connected by it are involved.

EXAMPLES

[0269] The present invention involves a genomic variation selected from chr9: 137779251, G>A and chr1:160769595, AG>A. The following Examples, in as much as they do not relate to genomic variations chr9:137779251, G>A and chr1:160769595, AG>A, are included for reference.

Example 1 - Experimental approach

[0270] In the present study, a set of genes were identified, deleterious variants within which increase susceptibility to PML. The relevant genes were discovered on the basis of a combined CNV plus sequence analysis approach. Two sets of genes were compiled (see Table 6 and corresponding description):

1. A. A set based on a detailed literature review of genes involved in the immune system and JC virus biology, along with genes described in the context of PML via case reports.
2. B. A set based on the observation of rare CNVs within the PML cohort.

[0271] A non-redundant list of 419 genes was generated (see **Table 6**), which contains 245 curated from immune deficiency

(immunodeficiency) reviews (**Table 6**, `Public db'), 169 identified via rare CNVs using the methods described herein (**Table 6**, 'PBio'), and 6 genes that were found using both methods (**Table 6**, 'Both'). See **Table 6** and description below for further information).

[0272] Using this set of 419 genes, it was determined whether:

- Rare CNVs were present that might explain the susceptibility to PML;
- Rare sequence variants (determined via whole exome sequencing analysis - WES) were present that might explain the susceptibility to PML;
- Combinations of CNVs, SNVs and/or CNVs and SNVs might explain the susceptibility;
- Individual variants might be present at higher frequency in the PML cohort (variant burden analysis - Tables 14, 15);
- Total numbers of heterozygous, damaging variants were high for any specific genes (gene burden analysis - **Table 13**).

[0273] In all cases, due consideration was given to:

- Pathogenic/deleterious nature of the variants observed (e.g., whether gene function was highly likely to be affected);
- Rarity of the variants or variant combinations (e.g., those that would be expected to be present in 1% or less of the normal population were considered);
- Ethnicity of the PML cases to account for potential frequency differences in one population subgroup vs. another. Ethnicities (e.g., ancestry) for the PML patients are reported in Table 7. For Sample ID identifiers beginning with 'MVGS', ethnicities were not reported but all patients were from the USA and their ethnicities were assumed to be of European (EUR) ancestry. However, PML case MVGS811-13a is potentially of African (AFR) ancestry on the basis of common SNVs that are also found in PML cases known to be of AFR ancestry. In one instance, ethnic-specific frequency data from the ExAC database was used to assess relative frequencies of variants found in PML patients vs. an unselected population (ExAC subjects). ExAC ethnicities were designated as follows: African/African American (AFR), Latino (LAT, also known as AMR), East Asian (EAS), Finnish (FIN), Non-Finnish European (EUR, also known as NFE), South Asian (SAS), and Other (OTH). For some PML cases reported in Table 7, the ethnicities were alternately reported as Subsaharan, North African (MGB), Caribbean (CAR), or Hispanic (HISP). For interpretation of variants found in these patients, the assignments of ancestry using ExAC db designations were as follows: AFR = MGB or Subsaharan; LAT = CAR or HISP. Ancestry was unknown for two PML cases (PML02 and PML28) and, for frequency interpretation purposes (using ExAC db), they were assumed to be of European (EUR) ancestry.

[0274] While the primary genetic mechanism that was considered was autosomal recessive (AR) inheritance, a number of solutions were based on autosomal dominant (AD) inheritance but only in cases for which prior evidence was found that heterozygous variants in the relevant gene had previously been associated with an immune deficiency syndrome. It can be appreciated by those skilled in the art that some genes may contain both AR and AD model pathogenic variants (e.g., see Table 6 entries marked as `AD AR' in the `Disease Model' column).

[0275] For AR inheritance (-40% of genes in **Table 6** fall into this category, AR or AD AR), the following were considered:

- Homozygous or compound heterozygous gene-disruptive CNVs;
- Homozygous or compound heterozygous sequence variants; e.g. single nucleotide variants (SNVs). Compound heterozygosity was only inferred when either phasing was available or one of the pairs of SNVs was itself homozygous;
- Compound heterozygosity for a CNV and SNV. Such calls were only possible in cases for which the SNV was in trans to a deletion (e.g., DUSP16 SNV in **Table 10** and the CNV in **Table 1**).

Example 2 - Copy Number Variant (CNV) Analysis

[0276] The data presented herein was generated on the basis of a comparison of copy number variants (CNVs) identified in 2 cohorts:

1. 1) 1,005 Normal individuals (Normal Variation Engine - NVE);
2. 2) 71 Progressive Multifocal Leukoencephalopathy (PML) cases along with 6 Human Immunodeficiency Virus (HIV) cases without a diagnosis of PML (in order to aid in distinguishing germline variants vs. acquired variants that result from HIV infection). Total cohort size = 77.

Genomic DNA sample hybridization - NVE and PML, HIV cohorts

[0277] Genomic DNA samples from individuals within the Normal cohort (NVE 'test' subjects, also referred to as `NVE cases' in some

tables herein) and from the PML, HIV cohort (PML, HIV 'test' subjects) were hybridized against a single, sex-matched reference individual. Reference DNA samples were labeled with Cy5 and test subject DNA samples were labeled with Cy3. After labeling, samples were combined and co-hybridized to Agilent 1M feature oligonucleotide microarrays, design ID 021529 (Agilent Product Number G4447A) using standard conditions (array Comparative Genomic Hybridization - aCGH). Post-hybridization, arrays were scanned at 2µm resolution, using Agilent's DNA microarray scanner, generating tiff images for later analysis.

[0278] All tiff images were analyzed using Agilent Feature Extraction (FE) software, with the following settings:

• Human Genome Freeze:	hg18:NCBI36:Mar2006
• FE version:	10.7.3.1
• Grid/design file:	021529 D F 20091001
• Protocol:	CGH 107 Sep09

[0279] This procedure generates a variety of output files, one of which is a text-tab delimited file, containing ~1,000,000 rows of data, each corresponding to a specific feature on the array. This *.txt file was used to perform CNV calling using DNAcopy, an open source software package implemented in R via BioConductor (http://www.bioconductor.org/packages/release/bioc/html/DNAcopy.html). Heterozygous losses (het loss), homozygous losses (hom loss) or gains were determined according to a threshold log2ratio, which was set at:

- hom loss min = -1000;
- hom loss max = -2;
- het loss min = -2;
- het loss max = -0.5;
- gain min = 0.5;
- gain max = 1000;

[0280] With very few exceptions, all CNVs with a log2ratio value between -0.5 and +0.5 were not considered. All log2ratio values were determined according to Cy3/Cy5 (Test/Reference). A minimum probe threshold for CNV-calling was set at 2 (2 consecutive probes were sufficient to call a CNV). A CNV list was generated for each individual in the 3 cohorts (NVE, PML, and HIV).

[0281] Using custom scripts, CNVs identified in the NVE and PML cohorts (many of which appeared in multiple individuals) were (separately) 'merged' into master lists of non-redundant CNV-subregions, according to the presence or absence of the CNV-subregion in individuals within the cohort. Using this approach, the NVE-master lists have:

• 7778	het loss
• 653	hom loss
• 4862	gain

distinct CNV-subregions, respectively. The PML + HIV cohort of 77 individuals master lists contained:

• 2523	het loss
• 314	hom loss
• 1639	gain

distinct CNV-subregions, respectively.

[0282] Those skilled in the art can appreciate that CNVs can be acquired in an individual's genome that are not inherited. Such `acquired CNVs' often occur in a tissue specific manner, such as in solid tumors compared to a patient's normal tissue. In blood-derived genomic DNA samples, which are what was used for both the NVE and PML subjects in the studies described herein, acquired CNVs can be the result of blood cancers such as leukemia and lymphoma, but also due to HIV infection. Many of the PML cases in this study had HIV as their primary disease (see **Table 7**). In order to aid in the interpretation of acquired vs. germline CNVs, an HIV sub-cohort of 6 cases was included in the primary, genome-wide CNV comparison but rare CNVs in the 6 HIV (non-PML) cases were not considered as relevant to PML susceptibility. The purpose of generating data on the 6 HIV cases was to determine whether some changes seen in PML patients who developed the disorder on a background of HIV (PML/HIV) were likely related to the underlying HIV and not the PML susceptibility itself. In other words, the HIV cases served as a general control for the large number of PML/HIV cases.

[0283] For example, consider 3 individuals within the NVE cohort with the following hypothetical CNVs:
Chrl: 1-100,000; Chrl: 10,001-100,000; and Chrl: 1-89,999. In the master list, these would be merged into 3 distinct CNV subregions, as follows:

• CNV-subregion 1	Chr1:1-10,000	Subjects A, C
• CNV-subregion 2	Chr1:10,001-89,999	Subjects A, B, C
• CNV-subregion 3	Chr90,000: 1-100,000	Subjects A, B

Comparison of the corresponding NVE and PML master lists of CNV-subregions was performed (het loss versus het loss, hom loss

versus hom loss and gain versus gain), resulting in a combined file with totals for NVE and PML for each distinct CNV-subregion in the study.

[0284] The data are subsequently curated as follows (The example calculation below was based on an original PML cohort of 80 cases, of which 6 are non-PML HIV controls and 3 PML cases that were duplicate samples. In some instances, the OR and FET values reported in **Table 2** were used as `relative' guidelines when considering the relevance of a CNV. In nearly all instances, a CNV was considered as a potential cause or contributing factor to PML if it was absent from the NVE database of CNVs).

- Annotation using custom designed scripts in order to attach relevant information to each CNV region regarding overlap with known genes and exons, overlap with genes involved in the immune system and overlap with regulatory regions, including transcription factor binding sites.
- A calculation of the odds ratio (OR) and Fisher's Exact test (FET) for each CNV-subregion, according to the following formula:
  - $OR = (PML/(77-PML))/(NVE/(1005-NVE))$
  - where:
  - PML = number of PML individuals with CNV-subregion of interest
  - NVE = number of NVE individuals with CNV-subregion of interest

[0285] As an illustrative example, consider the CNV subregion gain involving chr2:55764753-55771586, which is found in 3 individuals in the PML cohort and 1 individual in the NVE cohort (see **Table 2**). The OR is: (3/74)/(1/1004) = 40.7

[0286] Note that, by one convention, if either of NVE or PML = 0, a value of 0.5 is added to all 4 entries in the main formula above, in order to avoid dealing with infinities (*see* Deeks and Higgins, Statistical algorithms in Review Manager 5, Statistical Methods Group of The Cochrane Collaboration, (2010)). This has the effect of artificially lowering OR values in cases where no individuals within the NVE have the CNV. This method is applicable to all the calculations in **Table 2.** This method is also used when calculating the Fisher's 2-tailed Exact Test (FET) in the event that any one of the variables is zero. For convenience in analysis, the sub-cohort of 6 HIV (non-PML) cases were retained in **Table 2.** Therefore, the OR values reported in **Table 2** are slightly different from the OR calculations for the actual number of PML cases (n = 71). Using the example above for a CNV-subregion gain involving chr2:55764753-55771586, the actual OR using 71 PML cases vs. 1005 NVE subjects was: (3/68)/(1/(1004) = 44.29. In some instances, a non-PML HIV control (see **Table 11**, identified as 3280, 3281, 3283, 3284, 3285, and 3286) is found to have a CNV of potential relevance in PML subjects. This can also impact the OR calculation. For example, for CNV-subregion loss chr19:55247874-55250186 the OR in **Table 2** is listed as 17.38 but one case is a non-PML HIV control (**Table 11**, PML70 control = 3280). For this example, the actual OR using 71 PML cases vs. 1005 NVE subjects, and excluding the non-PML HIV case, was: (4/67)/(4/(1001) = 14.94.

[0287] The CNV-subregions/genes that are listed herein (e.g., in one or more of Tables 1-4), fulfill one of the following criteria:

- Strong biology linking the gene that a CNV-subregion impacts or is near, with known immune deficiency pathways/mechanisms or biology in PML (e.g., JC virus related biology). That is, in some cases, statistical evidence is lacking but does not exclude the CNV-subregion as a candidate;
- Statistical analysis combined with medium to strong biology (e.g., links in the peer-reviewed literature to PML, JC virus, host defense, immune deficiency, or neuropathology) without obvious biological connection (best FET in this category was 3.25E-10);

[0288] It can be appreciated by those skilled in the art that the number of PML candidate CNV-subregions, irrespective of category, may increase or decrease as additional PML cohorts are analyzed.

Example 3 - Whole Exome Sequencing (WES) and Case Level Analysis.

[0289] WES data was obtained on a total of 70 PML cases (non-PML HIV cases were not sequenced - they were used simply to help in the interpretation of complex CNVs observed in PML patients who also had HIV).

[0290] Variant annotation reports were further interrogated against the full set of genes detailed above. Synonymous variants and variants predicted to be modifiers (outside coding regions) were not considered. For all other variants, further filtering was performed so that only those predicted by at least one *in silico* prediction algorithm (e.g., Polyphen2, SIFT, MutationTaster) to be pathogenic were considered for further evaluation. Finally, only variants or variant combinations that would be expected to be present in 1% or less of the normal population were evaluated for case level analysis (Tables 7-10). Data from the Exome Aggregation Consortium (ExAC) was used to obtain ethnic-specific frequency data for variants under consideration (*see*, Lek et al., Nature, 17;536(7616):285-91) (2016)).

Example 4 - Description of Sequence Data

[0291] The sequence file 56969-701.601 ST25.txt contains genomic sequence information for (in the following order):

1. A. All distinct CNVs listed in **Table 1**;
2. B. The full genomic extent of the transcripts listed in **Table 4**;
3. C. Sequence variants detailed in **Table 5.**
4. D. The full genomic extent of the transcripts listed in **Table 12**

[0292] Note that:

1. 1. SEQ ID 1-172 are the CNV sequences from **Table 1;**
2. 2. SEQ ID 173-455 are the transcript sequences from **Table 4;**
3. 3. SEQ ID 1000-1329 are the sequence variants from **Table 5;**
4. 4. SEQ ID 1500-2177 are the transcript sequences from **Table 12.**

[0293] Examples of sequences submitted:

Sequence entry starts:

Table 1: SEQ ID 1 = 49,653bp CNV (het loss) at chr1:1086119-1135772 involving genes MIR200A,MIR200B,MIR429,TNFRSF18,TTLL10:

<210> 1
<211> 49654
<212> DNA
<213> Homo sapiens
<400> 1
cttctggggt ctaaggccag aagtgacctt tcttctcacg gaggcacccc cacatcacag 60 gccccaagct cccaccagga gtccccaggc 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

Sequence entry ends.

Sequence entry starts:

Table 4: SEQ ID 173 = MIR200B, transcript NR 029639, which is 95bp in length:

<210> 173
<211> 95
<212> DNA
<213> Homo sapiens
<400> 173
ccagctcggg cagccgtggc catcttactg ggcagcattg gatggagtca ggtctctaat 60 actgcctggt aatgatgacg gcggagccct gcacg 95

Sequence entry ends.

Sequence entry starts:

Table 5: SEQ ID 1148 = chr 9:304628 reference allele = G; alternate allele = A

<210> 1148
<211> 40
<212> DNA
<213> Homo sapiens
<220>
<221> variant
<222> (20)..(20)

<223> G->A
<400> 1148
tttaaaaaga ctggatctcg aaaagatttt cacaagacgc 40

Sequence entry ends.

Sequence entry starts:

Table 12: SEQ ID 1500 = ACADM, transcript NM 000016, which is 39,313 bp in length:

<210> 1500
<211> 39313
<212> DNA
<213> Homo sapiens
<400> 1500
cgcaagtccc cccaccgttc agcgcaaccg ggccctccca gccccgccgc cgtccccctc 60 ccccgccctg gctctctttc cgcgctgcgg tcagcctcgg cgtcccacag agagggccag 120 .....(sequence truncated for brevity) gtaatagtgt atatttcttg tatttactat gatgaaaaaa ggtcgtttta attttgaatt 39240 gaataaagtt acctgttcat tttttattag atattttaaa gacttcagaa aatataaata 39300 tgaaataatt taa 39313

Sequence entry ends.

Example 5

[0294] Those skilled in the art can appreciate that genes can be impacted by acquired or germline genetic variants (e.g., CNVs), wherein each gene has the potential to contain genetic variants that are acquired (e.g., via a disease process such as HIV infection, or cancers such as leukemia and lymphoma) or present in the germ line (e.g., inherited from a parent or are de novo, e.g. not inherited from a parent). In Figure 1, the PRKCB gene was impacted by germ line variants in 2 PML cases and acquired variants in 6 PML cases. The invention described herein is focused on detection of germline variants that are present in PML patient genomes. Therefore, no solutions/explanations for a given patient's PML was based on an acquired CNV, although another PML patient could potentially be 'solved' by one or two germline rare variants impacting the gene.

[0295] For this PRKCB example, no CNV-based solutions were found (an AR model was assumed), but 1 SNV solution is reported in **Table 8** (het SNV, an AD model is assumed for this PML case). Further supporting evidence was assessed for the PRKCB gene by performing String analysis (high confidence = 0.7, 1st shell = up to 10 interactors; string-db.org; see Szklarczyk *et al.,* (2015), and references therein). String analysis showed that PRKCB interacts with PML-419 genes CARD11, IKBKB, and RBCK1 (see **Table 6**).

[0296] In Figure 2, both TNFRSF13C and CENPM are disrupted and/or gained by a set of acquired CNV gains. Acquired CNVs can be very complex, such as the high copy number gains often identified in tumor-derived DNA samples (as compared to the patient's normal genome). In the PML gene discovery described herein, blood-derived genomic DNA obtained from several PML-diagnosed HIV patients, or PML cases with a primary disease of leukemia and lymphoma (reported as `Other' in Table 7), showed complex genomic changes (e.g., gains exhibiting a dup-trip-dup pattern). In some PML cases, the acquired gains passed the log2 ratio cutoff (>0.5) that was selected for this study, but in other PML cases the log2 ratios for the gains were <0.5 and this data was filtered out from the main analyses that were performed to ascertain rare germline CNVs.

[0297] In one instance, a set of 6 non-PML HIV cases (3 African ancestry, 3 European ancestry) were used to aid in the interpretation of whether a CNV was an acquired or germline event. The non-PML 'PML cases' are labeled with ' control' in **Table 11** and correspond to 'PML Case ID' numbers 3280, 3281, 3283, 3284, 3285, and 3286. While some CNVs are reported in Tables 1 and 2 for this set of non-PML control HIV subjects, none of these genetic findings were used to nominate a gene discovered on the basis of rare CNVs (as compared to the NVE db) as a potential PML gene (PBio genes reported in **Table 6**). In other words, these rare CNVs were only used to aid in determining if a particular genomic region containing multiple overlapping CNVs was potentially due to an acquired genetic event. Those skilled in the art can appreciate that the set of experiments described herein do not necessarily fully rule in or out that a given genomic region contained only acquired CNVs vs. only germline CNVs (e.g. it's possible that the same region can contain an acquired CNV in one individual and a germline CNV in another).

[0298] For the CNV data shown in Figure 2, both the TNFRSF13C and CENPM genes were included in PML-419 gene list (**Table 6**) on the basis of their immune or neurological related biology reported in the literature. No CNV or SNV PML solutions were found for these two genes, but String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) shows that TNFRSF13C interacts with PML-419 genes TRAF3 (**Table 7** solution) and TNFRSF13B (**Table 8** solution), as well as BTK (a known PML gene, see **Table 6**).

[0299] Figure 3 shows another example of a gene that is impacted by both germline and acquired CNVs. While no PML cases were

solved on the basis of the acquired or germline CNVs shown to impact the PKHD1 gene, nomination of this gene to **Table 6** on the basis of its biology resulted in finding 3 potential alternate solutions (AR model) for 3 other PML cases (see **Table 8**). However, String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) did not reveal any PML-419 gene interactions with PKHD1.

Example 6

[0300] Those skilled in the art can appreciate that an AR disease model would involve ascertaining whether both alleles (for a gene or genetic locus) are impacted by a genetic variant in individuals affected by the disorder. The types of genetic variants can be SNVs, CNVs, indels, etc. In the study describe herein, if an AR disease model was invoked for a gene (see **Table 6**), we assessed the PML patient's CGH data for CNVs (heterozygous or homozygous) and their exome data for SNVs (heterozygous or homozygous). Thus, each patient may be solved for one of the PML-419 genes (**Table 6**) with one of the following scenarios: homozygous deletion, homozygous duplication (log2 ratio will appear comparable to that typically found for triplications), homozygous SNV, compound heterozygous SNVs, compound heterozygous CNVs, or compound heterozygous SNV and CNV. Those skilled in the art know that, for an AR disease mechanism, a pathogenic SNV or CNV may have appreciable frequency in the general population (e.g., up to 1% frequency) with little to no impact on the individual's health, but when present with a second pathogenic variant on the other allele, can cause disease.

[0301] Figure 4 shows an example of an intronic loss impacting the BMPR2 gene. Patient PML29 was found to have a homozygous deletion, whereas as patients PML58 and MVGS811-13a have a heterozygous deletion. Assuming an AR disease model, no SNV solutions were found for this gene; however, PML29 is potentially solved due to the homozygous deletion that was detected. While immune-related biology is reported for studies on BMPR2 (see **Table 6**), String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) did not reveal any PML-419 gene interactions with BMPR2.

[0302] Figure 5 shows an example of an exonic gain that disrupts the COMMD6 gene. Two PML patients were found to have homozygous duplications of this CNV. Interestingly, while String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) did not reveal any PML-419 gene interactions with COMMD6, recent studies (see **Table 6**, PMIDs 25355947 and 27441653) show a potential link between COMMD6 and known PML gene WAS via the WASH gene.

[0303] Figure 6 shows an example of an exonic gain that disrupts the KCTD7 gene and its right breakpoint is upstream of RABGEF1 (e.g. one or both genes may be causing/contributing to PML). A recently annotated non-coding RNA (see hg19 assembly, LOC100996437) may also be impacted by this CNV. Both genes have immune and neurological links (see **Table 6**) and since patient PML29 has a homozygous duplication, it was added as a PML solution in **Table 7.** String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) did not reveal any PML-419 gene interactions for either gene, but they are linked together in a joint String analysis.

[0304] Figure 7 shows an example of a gain that disrupts FPR2 (left breakpoint) and ZNF616 (right breakpoint, gene not labeled), and other genes are fully encompassed by this CNV. There is strong supporting biology for FPR2 (see **Table 6**) and it is listed as a PML solution in **Table 7.** String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) did not reveal any PML-419 gene interactions for FPR2, but a joint analysis of **Table 7** genes did reveal an interaction (see Figure 13).

[0305] Figure 8 shows an example of an exonic loss impacting the PIK3CD and PIK3CD-AS1 genes. Patient MVGS811-13a has a homozygous deletion and is reported as a solution in **Table 7** based on the strong immune-related biology for PIK3CD (see **Table 6**). String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) reveals PML-419 gene interactions for PTEN and PIK3R1.

Example 7

[0306] A subset of the rare CNVs found in our PML study were located in intergenic regions. While those skilled in the art can appreciate that intergenic variants (CNVs, SNVs, etc.) can have long range effects on the expression of genes (e.g., gene regulatory elements can be located several kilobases away from the genes under their influence), in our study we assumed that intergenic CNVs were potentially impacting one or both adjacent genes if they were located <~100Kb away, either upstream or downstream. The ENCODE project has revealed a wealth of information, such as transcription factor binding sites, and rare CNVs that were identified in the study herein were checked for their potential impact on these sites (hg19 assembly ENCODE annotation was checked) and were often found to impact transcription factor binding sites and/or were located in conserved DNA regions.

[0307] Figure 9 shows an intergenic gain that is upstream of CD 180. Patient MVGS995-4a has a homozygous duplication and, while not considered as a PML solution in **Table 7**, is potentially an alternate solution that may be causing or contributing to the patient's PML based on altered expression of CD180. The gene has immune-related biology (see **Table 6**) and String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) reveals a PML-419 gene interaction with PLCG2 (see **Table 7**, 2 PML cases have a solution for this gene).

[0308] Figure 10 shows an intergenic loss that is upstream of VDAC1. Patient PML30 has a homozygous deletion and, while not considered as a PML solution in **Table 7**, is potentially an alternate solution that may be causing or contributing to the patient's PML based on altered expression of VDAC1. String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) did not reveal any PML-

419 gene interactions for VDAC1.

[0309] Figure 11 shows an intergenic loss that is downstream of EGR1 and ETF1. Patient PML69 has a homozygous deletion and, based on links for EGR1 to PML-419 genes (**Table 6**) and its proximity to EGR1 (~4Kb away), it was added as a potential PML solution in **Table 7.** String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) reveals PML-419 gene interactions with JUN, PTEN, and TP53), but nothing of note was found for String analysis of ETF1.

[0310] Figure 12 shows an intergenic loss that is upstream of ITSN2. Patient PML65 has a homozygous deletion and, based on links for ITSN2 to a known PML gene (WAS) in the PML-419 gene list (**Table** 6), it was added as a potential PML solution in **Table 7.** Interestingly, another PML case was found to have a rare homozygous SNV in ITSN2, so this gene has 2 PML solutions reported in **Table 7.** String analysis (high confidence = 0.7, 1st shell = up to 10 interactors) did not reveal any PML-419 gene interactions.

Example 8

[0311] Pathway analyses, such as protein-protein interactions, are providing valuable insights into the underlying biology for complex diseases. While PML is a very rare disease that requires several concurrent factors (e.g., infection by the JC virus), multiple genes may be independently causing or increasing the risk of developing this neurodegenerative disorder based on the presence of a genetic variant in a given gene (e.g., a heterozygous variant wherein one deleterious variant is present on the maternally or paternally inherited allele, a homozygous variant wherein the same deleterious variant is present on both alleles, or compound heterozygous variants wherein a pair of deleterious variants are present but one is found on the maternally inherited allele and the other is found on the paternally inherited allele). As hypothesized, presence of an immune deficiency genetic disorder was another prerequisite. Indeed, in the PML study described herein, 43 genes were proposed as solutions for 61 of 71 PML cases (see **Table 7**) that were assessed using array CGH and whole exome sequencing. Numerous algorithms and associated databases have been developed to investigate molecular pathways, such as String (*see*, Szklarczyk *et al.,* (2015), and references therein).

[0312] Figure 13 shows an example of String analysis performed on the 43 genes considered as PML solutions on the basis of an AD or AR disease model. A series of interactions were found for 21 of 43 genes, and in several instances this included interactions for genes implicated in 2 or more PML cases that are reported in **Table 7** (9 cases for TNFRSF11A, 4 cases for PLCG2, 3 cases for ZAP70 and NOD2, and 2 PML cases for TICAM1).

Example 9

[0313] To determine the likelihood that a randomly selected individual would harbor one of the variants described herein, the following analysis was performed: For each variant or combination of variants, the ethnic-specific frequency quoted in **Table 7** was used to determine the probability that a randomly selected individual of the same ethnicity would be expected not to harbor the variant or combination of variants. The product of all such probabilities was calculated (e.g., the probability that a randomly selected individual would not harbor any of the variants) and subtracted from 1, yielding the probability that a random individual would harbor at least one of the variants. It was found that, for HIV cases, the probability of a random individual harboring at least one of the variants was ~5%, which is consistent with the pre-HAART risk of PML in the context of HIV. For non-HIV cases (mostly MS/NTZ), the risk was ~ 1%, which, again, is consistent with the risk of PML in MS/NTZ, especially after long-term therapy.

[0314] These analyses support the notion that the frequencies of the variants identified as relevant to PML risk are consistent with the actual observed risks for unselected individuals. The analyses are predicated on the reasonable assumption that there is no PML-relevant connection with the risk of developing HIV (an acquired infection) and/or MS (e.g., this implies that treatment of healthy individuals with Natalizumab, for example, would result in similar risks of PML). Any deviations (e.g., variants found in a slightly higher number of normal individuals than expected according to the numbers actually observed to be affected by PML) may be due to: penetrance (e.g., not everyone with the variants will be at maximal risk of PML); the assumption that individuals with MS, HIV and other underlying conditions represented a normal (e.g., with respect to PML risk) cross-section of the general population, prior to developing the underlying disorders HIV, MS etc; and under ascertainment of PML, even in patients with HIV, MS/NTZ.

Example 10 - Tables referenced in this study

[0315]

Table 1: CNVs of interest in this study

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
1	1086119	1135772	49653	het loss	3009	MIR200A	1
						MIR200B	
						MIR429	

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
						TNFRSF18	
						TTLL10	
1	9634094	9635206	1112	hom loss	3009	PIK3CD	2
1	12018512	12032581	14069	gain	3205		3
1	19593401	19602807	9406	het loss	3203	CAPZB	4
1	21695957	21700243	4286	het loss	3161		5
1	24364786	24391166	26380	gain	3199	IFNLR1	6
1	28666669	28737671	71002	gain	3161	PHACTR4	7
						RCC1	
						SNHG3	
1	49372054	49380088	8034	het loss	3145	AGBL4	8
1	153816159	15382769 8	11539	het loss	3168		9
1	205607255	20561034 1	3086	gain	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	gain	3145	RHOQ	19
2	55764753	55790559	25806	gain	3143	PNPT1	20
2	55764753	55790559	25806	gain	3193	PNPT1	20
2	55764753	55790559	25806	gain	3282	PNPT1	20
2	55764753	55790559	25806	gain	3143	PNPT1	20
2	55764753	55790559	25806	gain	3193	PNPT1	20
2	55764753	55790559	25806	gain	3282	PNPT1	20
2	55764753	55790559	25806	gain	3143	PNPT1	20
2	55764753	55790559	25806	gain	3193	PNPT1	20
2	55764753	55790559	25806	gain	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	gain	3118	HK2	24
2	105418748	10543527 4	16526	het loss	3193	FHL2	25
2	110182348	11021024 9	27901	gain	3174	MALL	26
						MIR4267	
						MIR4436B1	
						MIR4436B2	
2	127823042	12782841 0	5368	het loss	3273		27
2	134911636	13491425 4	2618	het loss	3273	MGAT5	28
2	203005216	20301993 3	14717	het loss	3009	BMPR2	29
2	203005216	20301993 3	14717	het loss	3192	BMPR2	29
2	203005216	20301993 3	14717	hom loss	3152	BMPR2	29

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
2	230212897	23021633 9	3442	het loss	3154	DNER	30
3	122979920	12299440 2	14482	gain	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
4	26565071	26566345	1274	hom loss	3273	STIM2	32
4	54838623	54873909	35286	gain	3153	PDGFRA	33
4	90791460	90843887	52427	gain	3168		34
4	90800863	90808258	7395	het loss	3009		35
4	90800863	90808258	7395	het loss	3284		35
5	45331278	46150784	819506	gain	3157	HCN1	36
5	49771219	49774457	3238	gain	3273	EMB	37
5	66619415	66636116	16701	gain	3010		38
5	78480194	78497296	17102	gain	3205		39
5	78497296	78531091	33795	gain	3132		40
5	78497296	78521408	24112	gain	3185		41
5	78497296	78531091	33795	gain	3132		40
5	78497296	78521408	24112	gain	3185		41
5	78500552	78526637	26085	gain	3205		42
5	78497296	78531091	33795	gain	3132		40
5	78500552	78526637	26085	gain	3205		42
5	78497296	78531091	33795	gain	3132		40
5	83490494	83495169	4675	het loss	3204	EDIL3	43
5	133372071	13337972 7	7656	hom loss	3153		44
5	137836466	13784330 9	6843	hom loss	3279		45
5	150159466	15020260 1	43135	het loss	3117		46
5	150159466	15020413 4	44668	het loss	3180		47
5	150159466	15020260 1	43135	het loss	3199		46
5	150159466	15020413 4	44668	het loss	3278		47
5	150159466	15020260 1	43135	het loss	3117		46
5	150159466	15020413 4	44668	het loss	3180		47
5	150159466	15020260 1	43135	het loss	3199		46
5	150159466	15020413 4	44668	het loss	3278		47
5	150159466	15020260 1	43135	het loss	3117		46
5	150159466	15020413 4	44668	het loss	3180		47
5	150159466	15020260 1	43135	het loss	3199		46
5	150159466	15020413 4	44668	het loss	3278		47
5	150185190	15020114 5	15955	hom loss	3009		48
5	150185190	15020114 5	15955	hom loss	3143		48
5	150185190	15020260 1	17411	hom loss	3152		49
5	150185190	15020260 1	17411	hom loss	3154		49
5	150185190	15020260 1	17411	hom loss	3193		49
5	150159466	15020114 5	41679	hom loss	3196		50
5	150185190	15020114 5	15955	hom loss	3281		48
5	150185190	15020114 5	15955	hom loss	3009		48
5	150185190	15020114 5	15955	hom loss	3143		48
5	150185190	15020260 1	17411	hom loss	3152		49
5	150185190	15020260 1	17411	hom loss	3154		49

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
5	150185190	15020260 1	17411	hom loss	3193		49
5	150159466	15020114 5	41679	hom loss	3196		50
5	150185190	15020114 5	15955	hom loss	3281		48
5	150185190	15020260 1	17411	hom loss	3152		49
5	150185190	15020260 1	17411	hom loss	3154		49
5	150185190	15020260 1	17411	hom loss	3193		49
5	150185190	15020413 4	18944	het loss	3132		51
5	150159466	15020413 4	44668	het loss	3180		47
5	150202601	15020413 4	1533	het loss	3196		52
5	150191322	15020413 4	12812	het loss	3273		53
5	150185190	15020413 4	18944	het loss	3277		51
5	150159466	15020413 4	44668	het loss	3278		47
5	150185190	15020413 4	18944	het loss	3280		51
5	150185190	15020413 4	18944	het loss	3282		51
5	179590681	17962666 0	35979	het loss	3172	MAPK9	54
6	2882577	2947403	64826	het loss	3196	DKFZP686I15217	55
						NQO2	
						SERPINB6	
6	2964646	2966011	1365	het loss	3193	HTATSF1P2	56
						NQO2	
6	51766024	51773250	7226	het loss	3167	PKHD1	57
6	51952217	51969378	17161	gain	3127	PKHD1	58
6	51952217	51969378	17161	gain	3127	PKHD1	58
6	51953476	51965723	12247	gain	3205	PKHD1	59
6	51952217	51969378	17161	gain	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	10789060 5	8238	het loss	3201	PDSS2	64
6	166418511	16642238 6	3875	het loss	3125		65
6	166418511	16642238 6	3875	het loss	3163		65
6	166418511	16642238 6	3875	het loss	3192		65
6	166418511	16642238 6	3875	het loss	3193		65
6	166418511	16642238 6	3875	het loss	3194		65
6	166418511	16642238 6	3875	het loss	3200		65
6	166418511	16642238 6	3875	het loss	3205		65
6	166418511	16642238 6	3875	het loss	3280		65
6	166418511	16642238 6	3875	het loss	3281		65
6	166418511	16642238 6	3875	het loss	3284		65
6	166418511	16642238 6	3875	hom loss	3009		65
6	166418511	16642238 6	3875	hom loss	3152		65
6	166418511	16642238 6	3875	hom loss	3175		65
7	65741238	65768682	27444	gain	3152	KCTD7	66
7	65741238	65768682	27444	gain	3202	KCTD7	66
7	157174966	15717784 3	2877	het loss	3009	PTPRN2	67
7	157425841	15749623 8	70397	gain	3189	PTPRN2	68
7	158000082	15802456 9	24487	het loss	3279	PTPRN2	69
7	158000082	15802456 9	24487	het loss	3279	PTPRN2	69

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
7	158000082	15802456 9	24487	het loss	3279	MIR595	69
						PTPRN2	
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	10206443 1	15071	het loss	3173		74
8	102049360	10206443 1	15071	het loss	3175		74
8	102049360	10206443 1	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
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	gain	3198	AUH	79
						MIR3163	
						MIR3910-1	
						MIR3910-2	
						NFIL3	
9	118564159	11857563 3	11474	gain	3193	ASTN2	80
9	118612694	11866459 3	51899	het loss	3144	ASTN2	81
9	119220847	11923307 8	12231	gain	3005		82
10	899657	1071401	171744	gain	3161	GTPBP4	83
						IDI2	
						IDI2-AS1	
						LARP4B	
10	76217585	76411591	194006	gain	3179	KAT6B	84
10	116000069	11600438 8	4319	gain	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	91
						CDKN1B	
						CREBL2	
						DUSP16	
						GPR19	
						LOH12CR1	
12	12968705	12971310	2605	gain	3127		92
12	91786998	94313682	2526684	het loss	3126	EEA1 LOC643339	93
12	91786998	94313682	2526684	het loss	3126	LOC643339	93
						MRPL42	
						NUDT4	
						NUDT4P1	
						SOCS2	

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
						SOCS2-AS1	
						UBE2N	
12	91786998	94313682	2526684	het loss	3126	CCDC41	93
						CRADD	
						PLXNC1	
12	111061085	11106448 6	3401	het loss	3004	TRAFD1	94
13	40939924	41026908	86984	gain	3140	RGCC	95
13	75006025	75016304	10279	gain	3009	COMMD6	96
13	75006025	75016304	10279	gain	3152	COMMD6	96
13	91811087	91814369	3282	het loss	3143	GPC5	97
13	91811087	91811118	31	hom loss	3173	GPC5	98
13	110754499	11077830 1	23802	gain	3006	ARHGEF7	99
						TEX29	
14	20021118	20055469	34351	gain	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	gain	3143		104
14	21120750	21125513	4763	gain	3173		104
14	60901636	60909492	7856	het loss	3193	PRKCH	105
14	60912874	60921269	8395	het loss	3174	PRKCH	106
14	63937192	63944459	7267	gain	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	gain	3169	HEXA	110
15	75096101	75128723	32622	gain	3200	PSTPIP1	111
15	75101524	75115806	14282	gain	3132	PSTPIP1	112
15	75096101	75128723	32622	gain	3200	PSTPIP1	111
15	75105789	75115806	10017	gain	3127	PSTPIP1	113
15	75101524	75115806	14282	gain	3132	PSTPIP1	112
15	75105789	75115806	10017	gain	3199	PSTPIP1	113
15	75096101	75128723	32622	gain	3200	PSTPIP1	111
15	75105789	75115806	10017	gain	3279	PSTPIP1	113
15	75105789	75115806	10017	gain	3127	PSTPIP1	113
15	75101524	75115806	14282	gain	3132	PSTPIP1	112
15	75105789	75115806	10017	gain	3199	PSTPIP1	113
15	75096101	75128723	32622	gain	3200	PSTPIP1	111
15	75105789	75115806	10017	gain	3279	PSTPIP1	113
15	75096101	75128723	32622	gain	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

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
16	6942078	6945539	3461	gain	3173	RBFOX1	116
16	6942078	6945539	3461	gain	3175	RBFOX1	116
16	6942078	6945539	3461	gain	3282	RBFOX1	116
16	23842653	23848772	6119	het loss	3198	PRKCB	117
16	23892842	23903495	10653	gain	3199	PRKCB	118
16	23892842	23903495	10653	gain	3199	PRKCB	118
16	23893969	23908248	14279	gain	3205	PRKCB	119
16	23893969	23908248	14279	gain	3205	PRKCB	119
16	69044235	69050151	5916	gain	3174	FUK	120
16	69044235	69050151	5916	gain	3185	FUK	120
16	69052450	69081640	29190	het loss	3197	COG4 FUK	121
16	70653499	70665447	11948	gain	3143	HPR	122
16	70653499	70665447	11948	gain	3152	HPR	122
16	70653499	70665447	11948	gain	3192	HPR	122
16	70653499	70665447	11948	gain	3200	HPR	122
16	70653499	70665447	11948	gain	3282	HPR	122
16	70653499	70665447	11948	gain	3284	HPR	122
17	69341925	70202523	860598	gain	3183	BTBD17	123
						C17orf77	
						CD300A	
						CD300C	
						CD300E	
						CD300LB	
						CD300LD	
						CD300LF	
						DNAI2	
						GPR 142	
						GPRC5C	
						KIF19	
						MGC16275	
						RAB37	
						RPL38	
						TTYH2	
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	gain	3205	RPTOR	127
17	76247305	76265683	18378	gain	3127	RPTOR	128
17	76241510	76267844	26334	gain	3205	RPTOR	127
17	76241510	76267844	26334	gain	3205	RPTOR	127

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
18	9985530	10125331	139801	gain	3175		129
18	12764095	12781985	17890	gain	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	gain	3125	PIAS2	132
						ST8SIA5	
18	46917195	46945018	27823	het loss	3161		133
18	59457622	59465699	8077	het loss	3145	SERPINB4	134
19	3270755	3291144	20389	gain	3205		135
19	46386511	46388364	1853	hom loss	3175		136
19	52496536	52501292	4756	gain	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	gain	3155	FPR2	140
						FPR3	
						ZNF350	
						ZNF432	
						ZNF577	
						ZNF613	
						ZNF614	
						ZNF615	
						ZNF649	
						ZNF841	
19	56964168	57308449	344281	gain	3157	FPR2	140
						FPR3	
						ZNF350	
						ZNF432	
						ZNF577	
						ZNF613	
						ZNF614	
						ZNF615	
						ZNF649	
						ZNF841	
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

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
19	59249279	59251831	2552	hom loss	3164	VSTM1	142
19	59250742	59251831	1089	hom loss	3277	VSTM1	143
20	17844577	17954650	110073	gain	3166	MGME1	144
						OVOL2	
						SNORD17	
						SNX5	
20	42706680	42711434	4754	het loss	3125	ADA	145
21	15234620	15312960	78340	gain	3009	NRIP 1	146
21	29643302	29647950	4648	het loss	3202	BACH1	147
21	44634707	44666832	32125	gain	3200	TRPM2	148
21	44634707	44641658	6951	gain	3205	TRPM2	149
21	44634707	44671482	36775	gain	3279	TRPM2	150
21	44637544	44669596	32052	gain	3127	TRPM2	151
21	44637544	44657372	19828	gain	3185	TRPM2	152
21	44634707	44666832	32125	gain	3200	TRPM2	148
21	44634707	44641658	6951	gain	3205	TRPM2	149
21	44634707	44671482	36775	gain	3279	TRPM2	150
21	44637544	44669596	32052	gain	3127	TRPM2	151
21	44637544	44657372	19828	gain	3185	TRPM2	152
21	44634707	44666832	32125	gain	3200	TRPM2	148
21	44634707	44671482	36775	gain	3279	TRPM2	150
21	44643974	44657372	13398	het loss	3161	TRPM2	153
21	44637544	44669596	32052	gain	3127	TRPM2	151
21	44637544	44657372	19828	gain	3185	TRPM2	152
21	44634707	44666832	32125	gain	3200	TRPM2	148
21	44643974	44657372	13398	gain	3205	TRPM2	153
21	44634707	44671482	36775	gain	3279	TRPM2	150
21	44637544	44669596	32052	gain	3127	TRPM2	151
21	44634707	44666832	32125	gain	3200	TRPM2	148
21	44634707	44671482	36775	gain	3279	TRPM2	150
21	44637544	44669596	32052	gain	3127	TRPM2	151
21	44634707	44666832	32125	gain	3200	TRPM2	148
21	44660199	44681194	20995	gain	3205	TRPM2	154
21	44634707	44671482	36775	gain	3279	TRPM2	150
21	44637544	44669596	32052	gain	3127	TRPM2	151
21	44660199	44681194	20995	gain	3205	TRPM2	154
21	44634707	44671482	36775	gain	3279	TRPM2	150
21	44660199	44681194	20995	gain	3205	TRPM2	154
21	44634707	44671482	36775	gain	3279	TRPM2	150
21	44660199	44681194	20995	gain	3205	TRPM2	154
21	45348895	45354820	5925	het loss	3179	ADARB1	155
						APOBEC3A	
22	37689058	37715385	26327	gain	3169	APOBEC3A B	156
						APOBEC3B	
22	39257585	39261621	4036	het loss	3005	MKL1	157
22	40642402	40655210	12808	gain	3205	TNFRSF13C	158
22	40655820	40673250	17430	gain	3185		159
22	40655820	40675788	19968	gain	3205		160
22	40659633	40671866	12233	gain	3127		161
22	40655820	40673250	17430	gain	3185		159

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
22	40655820	40675788	19968	gain	3205		160
22	40659633	40671866	12233	gain	3127	CENPM	161
22	40655820	40673250	17430	gain	3185	CENPM	159
22	40663050	40668079	5029	gain	3190	CENPM	162
22	40663050	40668079	5029	gain	3202	CENPM	162
22	40655820	40675788	19968	gain	3205	CENPM	160
22	40659633	40671866	12233	gain	3127	CENPM	161
22	40655820	40673250	17430	gain	3185	CENPM	159
22	40655820	40675788	19968	gain	3205	CENPM	160
22	40655820	40673250	17430	gain	3185	CENPM	159
22	40655820	40675788	19968	gain	3205	CENPM	160
22	40655820	40675788	19968	gain	3205		160
23	232907	244684	11777	het loss	3007	PPP2R3B	163
23	7585301	7830994	245693	gain	3172		164
23	7585301	7830994	245693	gain	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	gain	3172		164
23	7585301	7830994	245693	gain	3172		164
23	6465033	8093113	1628080	het loss	3171	MIR651 PNPLA4	166
23	7585301	7830994	245693	gain	3172	PNPLA4	164
23	48358646	48408854	50208	het loss	3009		167
23	64710574	64725828	15254	gain	3125		168
23	73083877	73086192	2315	hom loss	3193	JPX	169
23	73083877	73086192	2315	hom loss	3200	JPX	169
23	122337025	12234087 9	3854	hom loss	3125	GRIA3	170
23	148452844	14846188 9	9045	het loss	3163		171
23	148452844	14846188 9	9045	het loss	3205		171
23	148452844	14846188 9	9045	hom loss	3144		171
23	148452844	14846188 9	9045	hom loss	3193		171
23	149901706	14990426 5	2559	gain	3117	HMGB3	172
23	149901706	14990426 5	2559	gain	3118	HMGB3	172

[0316] Table 1 lists all CNVs of interest, obtained as described in the text, with the exception that, for each entry, the original CNV start and stop positions are noted, along with original CNV size, type (heterozygous loss, homozygous loss or gain), Case ID and gene annotation (for the CNV-subregion NOT original CNV). The final column contains SEQ ID numbers. Standard chromosomal numbering used by those skilled in the art is used in **Table 1** for the autosomal chromosomes (1-22) but, for convenience with analysis methods, chromosome X is designated as chromosome 23 herein. All coordinates are based on hg 18.

Table 2: CNV-subregions of interest in this study

Chr	CNV Subregion Start	CNV Subregion Stop	CNV Subregion Size	CNV Type	PML Case ID	RefSeq Gene Symbol	Exon overlap	NVE cases	PML cases	FET	OR	CNV Subregion No (SRN)
1	1086119	1135772	49653	het loss	3009	MIR200A MIR200B MIR429 TNFRSF18 TTLL10	Y	0	1	0.005115965	39.43	1
1	9634094	9635206	1112	hom loss	3009	PIK3CD	Y	0	1	0.005115965	39.43	2
1	12018512	12032581	14069	gain	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	gain	3199	IFNLR1	Y	0	1	0.005115965	39.43	6
						PHACTR4						

1	28666669	28737671	71002	gain	3161	RCC1 SNHG3	Y	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	gain	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	gain	3145	RHOQ	N	0	1	0.005115965	39.43	19
2	55764753	55771586	6833	gain	3143	PNPT1	Y	1	3	0.001318303	40.7	20
2	55764753	55771586	6833	gain	3193	PNPT1	Y	1	3	0.001318303	40.7	21
2	55764753	55771586	6833	gain	3282	PNPT1	Y	1	3	0.001318303	40.7	22
2	55771587	55772965	1378	gain	3143	PNPT1	N	2	3	0.003126725	20.33	23
2	55771587	55772965	1378	gain	3193	PNPT1	N	2	3	0.003126725	20.33	24
2	55771587	55772965	1378	gain	3282	PNPT1	N	2	3	0.003126725	20.33	25
2	55772966	55790559	17593	gain	3143	PNPT1	Y	1	3	0.001318303	40.7	26
2	55772966	55790559	17593	gain	3193	PNPT1	Y	1	3	0.001318303	40.7	27
2	55772966	55790559	17593	gain	3282	PNPT1	Y	1	3	0.001318303	40.7	28
2	71190677	71191310	633	het loss	3175	MCEE	Y	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	gain	3118	HK2	Y	0	1	0.005115965	39.43	37
2	105418748	105435274	16526	het loss	3193	FHL2	Y	0	1	0.005115965	39.43	38
2	110182348	110210249	27901	gain	3174	MALL MIR4267 MIR4436B1 MIR4436B2	Y	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	gain	3202	IQCB1	Y	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
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	gain	3153	PDGFRA	Y	0	1	0.005115965	39.43	53
4	90791460	90843887	52427	gain	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	gain	3157	HCN1	Y	0	1	0.005115965	39.43	57
5	49771219	49774457	3238	gain	3273	EMB	Y	0	1	0.005115965	39.43	58
5	66619415	66636116	16701	gain	3010		N	0	1	0.005115965	39.43	59
5	78480194	78497296	17102	gain	3205		N	0	1	0.005115965	39.43	60
5	78497296	78500551	3255	gain	3132		N	0	2	0.005115965	66.59	61
5	78497296	78500551	3255	gain	3185		N	0	2	0.005115965	66.59	62
5	78500552	78521408	20856	gain	3132		N	0	3	2.49E-05	94.48	63
5	78500552	78521408	20856	gain	3185		N	0	3	2.49E-05	94.48	64
5	78500552	78521408	20856	gain	3205		N	0	3	2.49E-05	94.48	65
5	78521409	78526637	5228	gain	3132		N	0	2	0.005115965	66.59	66
5	78521409	78526637	5228	gain	3205		N	0	2	0.005115965	66.59	67
5	78526638	78531091	4453	gain	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
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
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	Y	0	1	0.005115965	39.43	109
6	2882577	2947403	64826	het loss	3196	DKFZP686I15217 NQO2 SERPINB6	Y	0	1	0.005115965	39.43	110
6	2964646	2966011	1365	het loss	3193	HTATSF1P2 NQO2	Y	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	gain	3127	PKHD1	N	0	1	0.005115965	39.43	113
6	51953476	51965723	12247	gain	3127	PKHD1	N	0	2	0.005115965	66.59	114
6	51953476	51965723	12247	gain	3205	PKHD1	N	0	2	0.005115965	66.59	115
6	51965724	51969378	3654	gain	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	Y	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	Y	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
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	gain	3152	KCTD7	Y	0	2	0.005115965	66.59	136
7	65741238	65768682	27444	gain	3202	KCTD7	Y	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	gain	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	Y	1	1	0.13732578	13.21	142
8	23103186	23125443	22257	het loss	3140	TNFRSF10A	Y	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
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	Y	0	1	0.005115965	39.43	158
9	93140394	93447826	307432	gain	3198	AUH MIR3163 MIR3910-1 MIR3910-2	Y	0	1	0.005115965	39.43	159
						NFIL3						
9	118564159	118575633	11474	gain	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	gain	3005		N	0	1	0.005115965	39.43	162
10	899657	1071401	171744	gain	3161	GTPBP4 IDI2 IDI2-AS1	Y	0	1	0.005115965	39.43	163

						LARP4B						
10	76217585	76411591	194006	gain	3179	KAT6B	Y	0	1	0.005115965	39.43	164
10	116000069	116004388	4319	gain	3010	VWA2	Y	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	Y	0	1	0.005115965	39.43	167
11	62382087	62398462	16375	het loss	3205	SLC3A2	Y	0	1	0.005115965	39.43	168
11	76631014	76643625	12611	het loss	3193	GDPD4	Y	0	1	0.005115965	39.43	169
12	11616557	12114030	497473	het loss	3126	ETV6	Y	0	1	0.005115965	39.43	170
12	12438904	12778142	339238	het loss	3126	APOLD1 CDKN1B CREBL2 DUSP16 GPR19 LOH12CR1	Y	0	1	0.005115965	39.43	171
12	12968705	12971310	2605	gain	3127		N	0	1	0.005115965	39.43	172
12	91845527	92201342	355815	het loss	3126	EEA1 LOC643339	Y	0	1	0.005115965	39.43	173
12	92215898	92567120	351222	het loss	3126	LOC643339 MRPL42 NUDT14 NUDT4P1 SOCS2 SOCS2-AS1 UBE2N	Y	0	1	0.005115965	39.43	174
12	92568362	93307172	738810	het loss	3126	CCDC41 CRADD PLXNC1	Y	0	1	0.005115965	39.43	175
12	111061085	111064486	3401	het loss	3004	TRAFD1	Y	0	1	0.005115965	39.43	176
13	40939924	41026908	86984	gain	3140	RGCC	Y	0	1	0.005115965	39.43	177
13	75006025	75016304	10279	gain	3009	COMMD6	Y	0	2	0.005115965	66.59	178
13	75006025	75016304	10279	gain	3152	COMMD6	Y	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	gain	3006	ARHGEF7 TEX29	Y	0	1	0.005115965	39.43	182
14	20021118	20055469	34351	gain	3205	RNASE10	Y	0	1	0.005115965	39.43	183
14	20426824	20481852	55028	hom loss	3200	ECRP RNASE3	Y	0	1	0.005115965	39.43	184
14	20430810	20458350	27540	het loss	3192	ECRP	Y	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	gain	3143		N	1	2	0.014314826	26.77	193
14	21120750	21125513	4763	gain	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	gain	3205	MTHFD1	Y	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
15	66065925	66082418	16493	het loss	3010		N	0	1	0.005115965	39.43	201
15	70432627	70443017	10390	gain	3169	HEXA	Y	0	1	0.005115965	39.43	202
15	75096101	75101523	5422	gain	3200	PSTPIP1	Y	0	1	0.005115965	39.43	203
15	75101524	75105788	4264	gain	3132	PSTPIP1	Y	0	2	0.005115965	66.59	204
15	75101524	75105788	4264	gain	3200	PSTPIP1	Y	0	2	0.005115965	66.59	205
15	75105789	75109086	3297	gain	3127	PSTPIP1	Y	0	5	1.16E-07	152.56	206
15	75105789	75109086	3297	gain	3132	PSTPIP1	Y	0	5	1.16E-07	152.56	207
15	75105789	75109086	3297	gain	3199	PSTPIP1	Y	0	5	1.16E-07	152.56	208
15	75105789	75109086	3297	gain	3200	PSTPIP1	Y	0	5	1.16E-07	152.56	209
15	75105789	75109086	3297	gain	3279	PSTPIP1	Y	0	5	1.16E-07	152.56	210
15	75109087	75115806	6719	gain	3127	PSTPIP1	Y	1	5	9.14E-06	69.72	211
15	75109087	75115806	6719	gain	3132	PSTPIP1	Y	1	5	9.14E-06	69.72	212
15	75109087	75115806	6719	gain	3199	PSTPIP1	Y	1	5	9.14E-06	69.72	213
15	75109087	75115806	6719	gain	3200	PSTPIP1	Y	1	5	9.14E-06	69.72	214
15	75109087	75115806	6719	gain	3279	PSTPIP1	Y	1	5	9.14E-06	69.72	215
15	75115807	75117798	1991	gain	3200	PSTPIP1	Y	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	gain	3173	RBFOX1	N	1	3	0,001318303	40.7	220
16	6942078	6945539	3461	gain	3175	RBFOX1	N	1	3	0,001318303	40.7	221
16	6942078	6945539	3461	gain	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	gain	3199	PRKCB	N	0	1	0,005115965	39.43	224
16	23893969	23903495	9526	gain	3199	PRKCB	N	0	2	0,005115965	66.59	225
16	23893969	23903495	9526	gain	3205	PRKCB	N	0	2	0,005115965	66.59	226
16	23903496	23908248	4752	gain	3205	PRKCB	Y	0	1	0,005115965	39.43	227
16	69047888	69050151	2263	gain	3174	FUK	N	0	2	0,005115965	66.59	228
16	69047888	69050151	2263	gain	3185	FUK	N	0	2	0,005115965	66.59	229
16	69052450	69081640	29190	het loss	3197	COG4 FUK	Y	0	1	0,005115965	39.43	230
16	70653499	70665447	11948	gain	3143	HPR	Y	0	6	1.16E-07	182.82	231
16	70653499	70665447	11948	gain	3152	HPR	Y	0	6	1.16E-07	182.82	232
16	70653499	70665447	11948	gain	3192	HPR	Y	0	6	1.16E-07	182.82	233
16	70653499	70665447	11948	gain	3200	HPR	Y	0	6	1.16E-07	182.82	234
16	70653499	70665447	11948	gain	3282	HPR	Y	0	6	1.16E-07	182.82	235
16	70653499	70665447	11948	gain	3284	HPR	Y	0	6	1.16E-07	182.82	236
17	69341925	70202523	860598	gain	3183	BTBD17 C17orf77 CD300A CD300C CD300L CD300LB CD300LD CD300LF DNAI2 GPR142 GPRC5C KIF19 MGC16275 RAB37 RPL38 TTYH2	Y	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
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	gain	3205	RPTOR	N	0	1	0,005115965	39.43	251
17	76247305	76265683	18378	gain	3127	RPTOR	N	0	2	0,005115965	66.59	252
17	76247305	76265683	18378	gain	3205	RPTOR	N	0	2	0,005115965	66.59	253
17	76265684	76267844	2160	gain	3205	RPTOR	N	0	1	0,005115965	39.43	254
18	9985530	10125331	139801	gain	3175		N	0	1	0,005115965	39.43	255
18	12764095	12781985	17890	gain	3191	PTPN2	Y	0	1	0,005115965	39.43	256
18	27026205	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	gain	3125	PIAS2 ST8SIA5	Y	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	Y	0	1	0,005115965	39.43	262
19	3270755	3291144	20389	gain	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	gain	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
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	66.59	277
19	55250187	55252420	2233	hom loss	3202	FLJ26850	N	0	2	0.005115965	66.59	278
19	56964168	57308449	344281	gain	3155	FPR2 FPR3 ZNF350 ZNF432 ZNF577 ZNF613 ZNF614 ZNF615 ZNF649 ZNF841	Y	3	2	0.043434433	8.91	279
19	56964168	57308449	344281	gain	3157	FPR2 FPR3 ZNF350 ZNF432 ZNF577 ZNF613 ZNF614 ZNF615 ZNF649 ZNF841	Y	3	2	0.043434433	8.91	280
19	59016855	59023850	6995	het loss	3117	NLRP12	Y	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	0.533838399	1.39	284
19	59250742	59251831	1089	hom loss	3160	VSTM1	N	38	4	0.533838399	1.39	285
19	59250742	59251831	1089	hom loss	3164	VSTM1	N	38	4	0.533838399	1.39	286
19	59250742	59251831	1089	hom loss	3277	VSTM1	N	38	4	0.533838399	1.39	287
20	17844577	17954650	110073	gain	3166	MGME1 OVOL2 SNORD17 SNX5	Y	0	1	0.005115965	39.43	288
20	42706680	42711434	4754	het loss	3125	ADA	N	0	1	0.005115965	39.43	289
21	15237071	15312960	75889	gain	3009	NRIP1	Y	0	1	0.005115965	39.43	290
21	29643302	29647950	4648	het loss	3202	BACH1	Y	0	1	0.005115965	39.43	291
21	44634707	44637543	2836	gain	3200	TRPM2	Y	1	3	0.001318303	40.7	292
21	44634707	44637543	2836	gain	3205	TRPM2	Y	1	3	0.001318303	40.7	293
21	44634707	44637543	2836	gain	3279	TRPM2	Y	1	3	0.001318303	40.7	294
21	44637544	44641658	4114	gain	3127	TRPM2	Y	1	5	9.14E-06	69.72	295
21	44637544	44641658	4114	gain	3185	TRPM2	Y	1	5	9.14E-06	69.72	296
21	44637544	44641658	4114	gain	3200	TRPM2	Y	1	5	9.14E-06	69.72	297
21	44637544	44641658	4114	gain	3205	TRPM2	Y	1	5	9.14E-06	69.72	298
21	44637544	44641658	4114	gain	3279	TRPM2	Y	1	5	9.14E-06	69.72	299
21	44641659	44643973	2314	gain	3127	TRPM2	Y	1	4	0.000112689	55.01	300
21	44641659	44643973	2314	gain	3185	TRPM2	Y	1	4	0.000112689	55.01	301
21	44641659	44643973	2314	gain	3200	TRPM2	Y	1	4	0.000112689	55.01	302
21	44641659	44643973	2314	gain	3279	TRPM2	Y	1	4	0.000112689	55.01	303
21	44643974	44657372	13398	het loss	3161	TRPM2	Y	1	1	0.13732578	13.21	304
21	44643975	44657372	13397	gain	3127	TRPM2	Y	0	5	1.16E-07	152.56	305
21	44643975	44657372	13397	gain	3185	TRPM2	Y	0	5	1.16E-07	152.56	306
21	44643975	44657372	13397	gain	3200	TRPM2	Y	0	5	1.16E-07	152.56	307
21	44643975	44657372	13397	gain	3205	TRPM2	Y	0	5	1.16E-07	152.56	308
21	44643975	44657372	13397	gain	3279	TRPM2	Y	0	5	1.16E-07	152.56	309
21	44657373	44660198	2825	gain	3127	TRPM2	Y	0	3	2.49E-05	94.48	310
21	44657373	44660198	2825	gain	3200	TRPM2	Y	0	3	2.49E-05	94.48	311
21	44657373	44660198	2825	gain	3279	TRPM2	Y	0	3	2.49E-05	94.48	312
21	44660199	44666832	6633	gain	3127	TRPM2	Y	0	4	2.49E-05	123.12	313
21	44660199	44666832	6633	gain	3200	TRPM2	Y	0	4	2.49E-05	123.12	314
21	44660199	44666832	6633	gain	3205	TRPM2	Y	0	4	2.49E-05	123.12	315
21	44660199	44666832	6633	gain	3279	TRPM2	Y	0	4	2.49E-05	123.12	316
21	44666833	44669596	2763	gain	3127	TRPM2	Y	0	3	2.49E-05	94.48	317
21	44666833	44669596	2763	gain	3205	TRPM2	Y	0	3	2.49E-05	94.48	318

21	44666833	44669596	2763	gain	3279	TRPM2	Y	0	3	2.49E-05	94.48	319
21	44669597	44671482	1885	gain	3205	TRPM2	Y	0	2	0.005115965	66.59	320
21	44669597	44671482	1885	gain	3279	TRPM2	Y	0	2	0.005115965	66.59	321
21	44671483	44681194	9711	gain	3205	TRPM2	Y	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	gain	3169	APOBEC3A APOBEC3A B APOBEC3B	Y	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	gain	3205	TNFRSF13C	Y	0	1	0.005115965	39.43	326
22	40655820	40659632	3812	gain	3185		N	0	2	0.005115965	66.59	327
22	40655820	40659632	3812	gain	3205		N	0	2	0.005115965	66.59	328
22	40659633	40663049	3416	gain	3127		N	0	3	2.49E-05	94.48	329
22	40659633	40663049	3416	gain	3185		N	0	3	2.49E-05	94.48	330
22	40659633	40663049	3416	gain	3205		N	0	3	2.49E-05	94.48	331
22	40663050	40668079	5029	gain	3127	CENPM	Y	0	5	1.16E-07	152.56	332
22	40663050	40668079	5029	gain	3185	CENPM	Y	0	5	1.16E-07	152.56	333
22	40663050	40668079	5029	gain	3190	CENPM	Y	0	5	1.16E-07	152.56	334
22	40663050	40668079	5029	gain	3202	CENPM	Y	0	5	1.16E-07	152.56	335
22	40663050	40668079	5029	gain	3205	CENPM	Y	0	5	1.16E-07	152.56	336
22	40668080	40671866	3786	gain	3127	CENPM	Y	0	3	2.49E-05	94.48	337
22	40668080	40671866	3786	gain	3185	CENPM	Y	0	3	2.49E-05	94.48	338
22	40668080	40671866	3786	gain	3205	CENPM	Y	0	3	2.49E-05	94.48	339

22	40671867	40673250	1383	gain	3185	CENPM	Y	0	2	0.005115965	66.59	340
22	40671867	40673250	1383	gain	3205	CENPM	Y	0	2	0.005115965	66.59	341
22	40673251	40675788	2537	gain	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	gain	3172		N	5	1	0.358539546	2.63	344
23	7769323	7773949	4626	gain	3172	VCX	Y	7	1	0.447101793	1.88	345
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	gain	3172		N	5	1	0.358539546	2.63	349
23	7779354	7815400	36046	gain	3172		N	6	1	0.404443314	2.19	350
23	7779355	8093113	313758	het loss	3171	MIR651 PNPLA4	Y	0	1	0.005115965	39.43	351
23	7815401	7830994	15593	gain	3172	PNPLA4	Y	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	gain	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	gain	3117	HMGB3	Y	0	2	0.005115965	66.59	362
23	149901706	149902701	995	gain	3118	HMGB3	Y	0	2	0.005115965	66.59	363

[0317] Table 2 is identical to **Table 1**, with a number of exceptions. Firstly, the CNV coordinates listed refer to the actual CNV-subregions found to be unique or significantly different between the disease and normal cohorts, as opposed to **Table 1**, which lists the original CNVs. Secondly, an extra column details whether genic CNV-subregions of interest overlap an exon or not. Third and fourth, 2 extra columns detail the number of normal cases and the number of disease cases that harbor the relevant CNV-subregion. Finally, 2 columns report Fisher's 2-tailed Exact Test (FET) and the odds ratio (OR). Standard chromosomal numbering used by those skilled in the art is used in **Table 2** for the autosomal chromosomes (1-22) but, for convenience with analysis methods, chromosome X is designated as chromosome 23 herein. All coordinates are in hg18.

Table 3: A non-redundant list of genes listed in **Table 2**

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
ADARB1	intronic	104	double-stranded RNA-specific editase 1 isoform 1	This gene encodes the enzyme responsible for pre-mRNA editing of the glutamate receptor subunit B by site-specific deamination of adenosines. Studies in rat found that this enzyme acted on its own pre-mRNA molecules to convert an AA dinucleotide to an AI dinucleotide which resulted in a new splice site. Alternative splicing of this gene results in several transcript variants, some of which have been characterized by the presence or absence of an ALU cassette insert and a short or long C-terminal region, [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1), also known as ADAR2a-L1 or DRADA2a, lacks the ALU cassette insert and contains the long C-terminal region, as compared to variant 2. The resulting isoform (1), also known as bRED 1-Short, lacks an internal segment, compared to isoform 2. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. ##RefSeq-Attributes-START## undergoes RNA editing :: PMID: 11717408, 12045112 ##RefSeq-Attributes-END## Transcript exon combination :: AB194370.1, U76420.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025085 [ECO:0000348]	2
AGBL4	intronic	84871	cytosolic carboxypeptid ase 6	N/A	3
APOBEC3 A	exonic	200315	DNA dC->dU-editing enzyme APOBEC-3A isoform a	This gene is a member of the cytidine deaminase gene family. It is one of seven related genes or pseudogenes found in a cluster, thought to result from gene duplication, on chromosome 22. Members of the cluster encode proteins that are structurally and functionally related to the C to U RNA-editing cytidine deaminase APOBEC1. The protein encoded by this gene lacks the zinc binding activity of other family members. The protein plays a role in immunity, by restricting transmission of foreign DNA such as viruses. One mechanism of foreign DNA restriction is deamination of foreign double-stranded DNA cytidines to uridines, which leads to DNA degradation. However, other mechanisms are also thought to be involved, as anti-viral effect is not dependent on deaminase activity. Two transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Jul 2012]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (a). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: U03891.2, BC126416.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025084 [ECO:0000348]	4
APOBEC3 AB	intronic	1009131 87	probable DNA dC->dU-editing enzyme APOBEC-3A	This gene is a member of the cytidine deaminase gene family. It is one of seven related genes or pseudogenes found in a cluster, thought to result from gene duplication, on chromosome 22. Members of the cluster encode proteins that are structurally and functionally related to the C to U RNA-editing cytidine deaminase APOBEC1. The protein encoded by this gene lacks the zinc binding activity of other family members. The protein plays a role in immunity, by restricting	5

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				transmission of foreign DNA such as viruses.	
				One mechanism of foreign DNA restriction is deamination of foreign double-stranded DNA cytidines to uridines, which leads to DNA degradation. However, other mechanisms are also thought to be involved, as anti-viral effect is not dependent on deaminase activity. The protein encoded by this gene is the same as that encoded by APOBEC3A; however, this gene is a hybrid gene that results from the deletion of approximately 29.5 kb of sequence between the APOBEC3A gene and the adjacent gene APOBEC3B. The breakpoints of the deletion are within the two genes, so the deletion hybrid is predicted to have the promoter and coding region of APOBEC3A, but the 3' UTR of APOBEC3B. [provided by RefSeq, Jul 2012]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. RNAseq introns :: single sample supports all introns ERS025081, ERS025084 [ECO:0000348]	
APOBEC3 B	exonic	9582	DNA dC->dU-editing enzyme APOBEC-3B isoform a	This gene is a member of the cytidine deaminase gene family. It is one of seven related genes or pseudogenes found in a cluster, thought to result from gene duplication, on chromosome 22. Members of the cluster encode proteins that are structurally and functionally related to the C to U RNA-editing cytidine deaminase APOBEC1. It is thought that the proteins may be RNA editing enzymes and have roles in growth or cell cycle control. A hybrid gene results from the deletion of approximately 29.5 kb of sequence between this gene, APOBEC3B, and the adjacent gene APOBEC3A. The breakpoints of the deletion are within the two genes, so the deletion allele is predicted to have the promoter and coding region of APOBEC3A, but the 3' UTR of APOBEC3B. Two transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Jul 2012]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (a). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AY743217.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	6
APOLD1	exonic	81575	apolipoprotei n L domain-containing protein 1 isoform 1	APOLD1 is an endothelial cell early response protein that may play a role in regulation of endothelial cell signaling and vascular function (Regard et al., 2004 [PubMed 15102925]). [supplied by OMIM, Dec 2008]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (1). Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The extent of this transcript is supported by transcript alignments. Transcript exon combination :: BC042478.1, DR000985.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025086 [ECO:0000348]	7
ARHGEF7	exonic	8874	rho guanine nucleotide exchange factor 7 isoform a	Rho GTPases play a fundamental role in numerous cellular processes triggered by extracellular stimuli that work through G protein coupled receptors. The encoded protein belongs	8

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				to a family of cytoplasmic proteins that activate the Ras-like family of Rho proteins by exchanging bound GDP for GTP. It forms a complex with the small GTP binding protein Rac1 and recruits Rac1 to membrane ruffles and to focal adhesions. This protein can induce membrane ruffling. Multiple alternatively spliced transcript variants encoding different isoforms have been described for this gene. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) differs in the 5' UTR, 3' UTR, coding region, and uses a downstream start codon, compared to variant 3. Both variants 1 and 5 encode isoform a, which has a shorter N-terminus and a longer and distinct C-terminus, compared to isoform c. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: D63476.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	
ASTN2	intronic	23245	astrotactin-2 isoform a precursor	This gene encodes a protein that is expressed in the brain and may function in neuronal migration, based on functional studies of the related astrotactin 1 gene in human and mouse. A deletion at this locus has been associated with schizophrenia. Multiple transcript variants encoding different proteins have been found for this locus. [provided by RefSeq, May 2010]. Transcript Variant: This variant (1) represents the longest transcript and encodes the longest isoform (a). Transcript exon combination :: BC146756.1, AB014534.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025082 [ECO:0000348]	9
AUH	exonic	549	methylglutaconyl-CoA hydratase, mitochondrial precursor	The methylglutaconyl-CoA hydratase, mitochondrial protein binds to the AU-rich element (ARE), a common element found in the 3' UTR of rapidly decaying mRNA such as c-fos, c-myc and granulocyte/ macrophage colony stimulating factor. ARE elements are involved in directing RNA to rapid degradation and deadenylation. AUH is also homologous to enol-CoA hydratase, an enzyme involved in fatty acid degradation, and has been shown to have intrinsic hydratase enzymatic activity. AUH is thus a bifunctional chimera between RNA binding and metabolic enzyme activity. A possible subcellular localization in the mitochondria has been demonstrated for the mouse homolog of this protein which shares 92% identity with the human protein. It has been suggested that AUH may have a novel role as a mitochondrial located AU-binding protein. Human AUH is expressed as a single mRNA species of 1.8 kb, and translated as a 40-kDa precursor protein which is subsequently processed to a 32-kDa mature form. [provided by RefSeq, May 2010]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. ##RefSeq-Attributes-START## gene product(s) localized to mito. :: reported by MitoCarta ##RefSeq-Attributes-END## Transcript exon combination :: X79888.1, AL533438.3 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	10
BACH1	exonic	571	BTB Domain And CNC Homolog 1	This gene encodes a transcription factor that belongs to the cap'n'collar type of basic region	11

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				leucine zipper factor family (CNC-bZip). The encoded protein contains broad complex, tramtrack, bric-a-brac/poxvirus and zinc finger (BTB/POZ) domains, which is atypical of CNC-bZip family members. These BTB/POZ domains facilitate protein-protein interactions and formation of homo- and/or hetero-oligomers. When this encoded protein forms a heterodimer with MafK, it functions as a repressor of Maf recognition element (MARE) and transcription is repressed. Multiple alternatively spliced transcript variants have been identified for this gene. [provided by RefSeq, May 2009]. Transcript Variant: This variant (3), also named BACH1t, differs in the 5' UTR, 3' coding region and 3' UTR (compared to variant 1). This variant is represented as non-coding because the use of the 5'-most supported translational start codon, as used in variant 1, renders the transcript a candidate for nonsense-mediated mRNA decay (NMD). This transcript represents the splice variant reported by Kanezaki et al. (PMID: 11069897). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. RNAseq introns :: mixed/partial sample support ERS025084, ERS025088 [ECO:0000350]	
BDKRB2	intronic	624	B2 bradykinin receptor	This gene encodes a receptor for bradykinin. The 9 aa bradykinin peptide elicits many responses including vasodilation, edema, smooth muscle spasm and pain fiber stimulation. This receptor associates with G proteins that stimulate a phosphatidylinositol-calcium second messenger system. Alternate start codons result in two isoforms of the protein, [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: DC369062.1, DC417219.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025090 [ECO:0000348]	12
BMPR2	intronic	659	bone morphogeneti c protein receptor type-2 precursor	This gene encodes a member of the bone morphogenetic protein (BMP) receptor family of transmembrane serine/threonine kinases. The ligands of this receptor are BMPs, which are members of the TGF-beta superfamily. BMPs are involved in endochondral bone formation and embryogenesis. These proteins transduce their signals through the formation of heteromeric complexes of two different types of serine (threonine) kinase receptors: type I receptors of about 50-55 kD and type II receptors of about 70-80 kD. Type II receptors bind ligands in the absence of type I receptors, but they require their respective type I receptors for signaling, whereas type I receptors require their respective type II receptors for ligand binding. Mutations in this gene have been associated with primary pulmonary hypertension, both familial and fenfluramine-associated, and with pulmonary venoocclusive disease. [provided by RefSeq, Jul 2008]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The extent of this transcript is supported by transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for	13

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC052985.2, AK292430.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	
BTBD17	exonic	388419	BTB/POZ domain-containing protein 17 precursor	N/A	14
C17orf77	exonic	146723	uncharacteriz ed protein C17orf77 precursor	N/A	15
CAPZB	intronic	832	F-actin-capping protein subunit beta isoform 1	This gene encodes the beta subunit of the barbed-end actin binding protein, which belongs to the F-actin capping protein family. The capping protein is a heterodimeric actin capping protein that blocks actin filament assembly and disassembly at the fast growing (barbed) filament ends and functions in regulating actin filament dynamics as well as in stabilizing actin filament lengths in muscle and nonmuscle cells. A pseudogene of this gene is located on the long arm of chromosome 2. Multiple alternatively spliced transcript variants encoding different isoforms have been found. [provided by RefSeq, Aug 2013]. Transcript Variant: This variant (1) encodes isoform 1. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC107752.1, BM451686.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025088 [ECO:0000348]	16
CCDC41	exonic	51134	centrosomal protein of 83 kDa	N/A	17
CD300A	exonic	11314	CMRF35-like molecule 8 isoform 1 precursor	This gene encodes a member of the CD300 glycoprotein family of cell surface proteins found on leukocytes involved in immune response signaling pathways. This gene is located on chromosome 17 in a cluster with all but one of the other family members. Multiple transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Feb 2012]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer protein (isoform 1), also referred to as IRC1a. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC032352.1, AL531420.3 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025083 [ECO:0000348]	18
CD300C	exonic	10871	CMRF35-like molecule 6 precursor	The CMRF35 antigen, which was identified by reactivity with a monoclonal antibody, is present on monocytes, neutrophils, and some T and B lymphocytes (Jackson et al., 1992 [PubMed 1349532]). [supplied by OMIM, Mar 2008]. Transcript exon combination :: BC022279.1, BM922826.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025087 [ECO:0000348]	19
CD300E	exonic	342510	CMRF35-like molecule 2 precursor	This gene encodes a member of the CD300 glycoprotein family of cell surface proteins expressed on myeloid cells. The protein interacts with the TYRO protein tyrosine kinase-binding protein and is thought to act as an activating	20

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				receptor. [provided by RefSeq, Nov 2012]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. An in-frame AUG is located 41 codons upstream of the annotated translation start site but is not being annotated as a start site since it is not conserved and is in a weak Kozak sequence context. ##RefSeq-Attributes-START## CDS uses downstream in-frame AUG :: downstream AUG is associated with N-terminal localization signal ##RefSeq-Attributes-END## Transcript exon combination :: AK303545.1, BX648376.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	
CD300LB	exonic	124599	CMRF35-like molecule 7 precursor	CD300LB is a nonclassical activating receptor of the immunoglobulin (Ig) superfamily expressed on myeloid cells (Martinez-Barriocanal and Sayos, 2006 [PubMed 16920917]). [supplied by OMIM, Mar 2008]. CCDS Note: The coding region has been updated to shorten the N-terminus to one that is more supported by available conservation data and paralogous family members. The update has a predicted N-terminal signal peptide, which is consistent with functional support for the protein (e.g., PMIDs 16920917, 19359216). Transcript exon combination :: BC028091.1, AY359025.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348] ##RefSeq-Attributes-START## CDS uses downstream in-frame AUG :: downstream AUG is associated with N-terminal localization signal ##RefSeq-Attributes-END##	21
CD300LD	exonic	1001314 39	CMRF35-like molecule 4 precursor	N/A	22
CD300LF	exonic	146722	CMRF35-like molecule 1 precursor	CD300LF is an inhibitory receptor of the Ig superfamily expressed on myeloid cells. It mediates negative regulatory signals by recruiting SHP1 (PTPN6; MIM 176883) or SHIP (INPP5D; MIM 601582) (Sui et al., 2004 [PubMed 15184070]; Alvarez-Errico et al., 2004 [PubMed 15549731]). [supplied by OMIM, Mar 2008]. Sequence Note: The RefSeq transcript and protein were derived from genomic sequence to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AF251706.1, AY358545.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084 [ECO:0000348]	23
CDKN1B	exonic	1027	cyclin-dependent kinase inhibitor 1B	This gene encodes a cyclin-dependent kinase inhibitor, which shares a limited similarity with CDK inhibitor CDKN1A/p21. The encoded protein binds to and prevents the activation of cyclin E-CDK2 or cyclin D-CDK4 complexes, and thus controls the cell cycle progression at G1. The degradation of this protein, which is triggered by its CDK dependent phosphorylation and subsequent ubiquitination by SCF complexes, is required for the cellular transition from quiescence to the proliferative state. [provided by RefSeq, Jul	24

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC001971.1, AY004255.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	
CENPM	exonic	79019	centromere protein M isoform a	The centromere is a specialized chromatin domain, present throughout the cell cycle, that acts as a platform on which the transient assembly of the kinetochore occurs during mitosis. All active centromeres are characterized by the presence of long arrays of nucleosomes in which CENPA (MIM 117139) replaces histone H3 (see MIM 601128). CENPM is an additional factor required for centromere assembly (Foltz et al., 2006 [PubMed 16622419]). [supplied by OMIM, Mar 2008]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (a). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC000705.2, BC007495.2 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025085, ERS025088 [ECO:0000348]	25
COG4	exonic	25839	conserved oligomeric Golgi complex subunit 4 isoform 1	The protein encoded by this gene is a component of an oligomeric protein complex involved in the structure and function of the Golgi apparatus. Defects in this gene may be a cause of congenital disorder of glycosylation type IIj. Two transcript variants encoding different isoforms have been found for this gene.[provided by RefSeq, Aug 2010]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (1). Transcript exon combination :: BC072438.1, AK022874.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	26
CONIMD6	exonic	170622	COMM domain-containing protein 6 isoform a	COMMD6 belongs to a family of NF-kappa-B (see RELA; MIM 164014)-inhibiting proteins characterized by the presence of a COMM domain (see COMMD1; MIM 607238) (de Bie et al., 2006 [PubMed 16573520]).[supplied by OMIM, Mar 2009]. Transcript exon combination :: HY028175.1, DW440523.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025088 [ECO:0000348]	27
CRADD	exonic	8738	death domain-containing protein CRADD	The protein encoded by this gene is a death domain (CARD/DD)-containing protein and has been shown to induce cell apoptosis. Through its CARD domain, this protein interacts with, and thus recruits, caspase 2/ICH1 to the cell death signal transduction complex that includes tumor necrosis factor receptor 1 (TNFR1A), RIPK1/RIP kinase, and numbers of other CARD domain-containing proteins. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BX480215.1, BC017042.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025083 [ECO:0000348]	28
CREBL2	exonic	1389	cAMP-responsive element-binding protein-like 2	cAMP response element (CRE)-binding protein-like-2 (CREBL2) was identified in a search to find genes in a commonly deleted region on	29

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				chromosome 12p13 flanked by ETV6 and CDKN1B genes, frequently associated with hematopoietic malignancies, as well as breast, non-small-cell lung and ovarian cancers. CREBL2 shares a 41%	
				identity with CRE-binding protein (CREB) over a 48-base long region which encodes the bZip domain of CREB. The bZip domain consists of about 30 amino acids rich in basic residues involved in DNA binding, followed by a leucine zipper motif involved in protein dimerization. This suggests that CREBL2 encodes a protein with DNA binding capabilities. The occurance of CREBL2 deletion in malignancy suggests that CREBL2 may act as a tumor suppressor gene. [provided by RefSeq, Jul 2008]. Transcript exon combination :: BC106052.1, AF039081.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	
DNAI2	exonic	64446	dynein intermediate chain 2, axonemal isoform 1	The protein encoded by this gene belongs to the dynein intermediate chain family, and is part of the dynein complex of respiratory cilia and sperm flagella. Mutations in this gene are associated with primary ciliary dyskinesia type 9. Alternatively spliced transcript variants encoding different isoforms have been noted for this gene. [provided by RefSeq, Mar 2010]. Transcript Variant: This variant (1) encodes the longer isoform (1). Transcript exon combination :: AF250288.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025085 [ECO:0000348] ##RefSeq-Attributes-START## NMD candidate :: translation inferred from conservation ##RefSeq-Attributes-END##	30
DNER	intronic	92737	delta and Notch-like epidermal growth factor-related receptor precursor	N/A	31
DUSP16	exonic	80824	dual specificity protein phosphatase 16	This gene encodes a mitogen-activated protein kinase phosphatase that is a member of the dual specificity protein phosphatase subfamily. These phosphatases inactivate their target kinases by dephosphorylating both the phosphoserine/threonine and phosphotyrosine residues. The encoded protein specifically regulates the c-Jun amino-terminal kinase (JNK) and extracellular signal-regulated kinase (ERK) pathways.[provided by RefSeq, May 2010]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AF506796.1, AB052156.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	32
ECRP	exonic	643332	N/A	N/A	33
EDIL3	intronic	10085	EGF-like repeat and discoidin I-like domain-containing protein 3 isoform 1	The protein encoded by this gene is an integrin ligand. It plays an important role in mediating angiogenesis and may be important in vessel wall remodeling and development. It also influences endothelial cell behavior, [provided by RefSeq, Jul	34

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
			precursor	2008]. Transcript Variant: This variant (1) encodes the longer isoform (1). Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC030828.1, U70312.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	
EEA1	exonic	8411	early endosome antigen 1	N/A	35
EHF	both	26298	ETS homologous factor isoform 1 precursor	This gene encodes a protein that belongs to an ETS transcription factor subfamily characterized by epithelial-specific expression (ESEs). The encoded protein acts as a transcriptional repressor and may be involved in epithelial differentiation and carcinogenesis. Three transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Jun 2011]. Transcript Variant: This variant (1) encodes the longest isoform (1). Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AK310867.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025086 [ECO:0000348]	36
EMB	exonic	133418	embigin precursor	This gene encodes a transmembrane glycoprotein that is a member of the immunoglobulin superfamily. The encoded protein may be involved in cell growth and development by mediating interactions between the cell and extracellular matrix. A pseudogene of this gene is found on chromosome 1. [provided by RefSeq, Jan 2009]. Transcript exon combination :: BC059398.1, AK300860.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	37
ETV6	exonic	2120	transcription factor ETV6	This gene encodes an ETS family transcription factor. The product of this gene contains two functional domains: a N-terminal pointed (PNT) domain that is involved in protein-protein interactions with itself and other proteins, and a C-terminal DNA-binding domain. Gene knockout studies in mice suggest that it is required for hematopoiesis and maintenance of the developing vascular network. This gene is known to be involved in a large number of chromosomal rearrangements associated with leukemia and congenital fibrosarcoma. [provided by RefSeq, Sep 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC043399.1, U11732.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081,	38

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				ERS025082 [ECO:0000348]	
FHL2	exonic	2274	four and a half LIM domains protein 2	This gene encodes a member of the four-and-a-half-LIM-only protein family. Family members contain two highly conserved, tandemly arranged, zinc finger domains with four highly conserved cysteines binding a zinc atom in each zinc finger. This protein is thought to have a role in the assembly of extracellular membranes. Also, this gene is down-regulated during transformation of normal myoblasts to rhabdomyosarcoma cells and the encoded protein may function as a link between presenilin-2 and an intracellular signaling pathway. Multiple alternatively spliced variants, encoding the same protein, have been identified, [provided by RefSeq, Aug 2011]. Transcript Variant: This variant (1) differs in the 5' UTR compared to variant 2. Variants 1, 2, 4 and 5 encode the same isoform. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. ##RefSeq-Attributes-START## CDS uses downstream in-frame AUG :: lack of evidence for use of upstream AUG ##RefSeq-Attributes-END## Transcript exon combination :: BC093049.1, AL523628.3 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	39
FLJ26850	intronic	400710	N/A	N/A	40
FPR2	exonic	2358	N-formyl peptide receptor 2	N/A	41
FPR3	exonic	2359	N-formyl peptide receptor 3	N/A	42
FUK	both	197258	L-fucose kinase	The protein encoded by this gene belongs to the GHMP (galacto-, homoserine, mevalonate and phosphomevalonate) kinase family and catalyzes the phosphorylation of L-fucose to form beta-L-fucose 1-phosphate. This enzyme catalyzes the first step in the utilization of free L-fucose in glycoprotein and glycolipid synthesis. L-fucose may be important in mediating a number of cell-cell interactions such as blood group antigen recognition, inflammation, and metastatis. While several transcript variants may exist for this gene, the full-length nature of only one has been described to date. [provided by RefSeq, Jul 2008]. Transcript exon combination :: AJ441184.1, BC032542.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	43
GDA	exonic	9615	guanine deaminase isoform a	This gene encodes an enzyme responsible for the hydrolytic deamination of guanine. Studies in rat ortholog suggest this gene plays a role in microtubule assembly. Multiple transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Nov 2011]. Transcript Variant: This variant (1) encodes the longest isoform (a). Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. RNAseq introns :: mixed/partial sample support ERS025082, ERS025083 [ECO:0000350]	44
GDPD4	exonic	220032	glycerophosphodiester phosphodiesterase domain-	N/A	45

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
			containing protein 4		
GPATCH2	intronic	55105	G patch domain-containing protein 2	N/A	46
GPC5	intronic	2262	glypican-5 precursor	Cell surface heparan sulfate proteoglycans are composed of a membrane-associated protein core substituted with a variable number of heparan sulfate chains. Members of the glypican-related integral membrane proteoglycan family (GRIPS) contain a core protein anchored to the cytoplasmic membrane via a glycosyl phosphatidylinositol linkage. These proteins may play a role in the control of cell division and growth regulation. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC030584.1, BC039730.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025082, ERS025083 [ECO:0000350]	47
GPR19	exonic	2842	probable G-protein coupled receptor 19	N/A	48
GPR142	exonic	350383	probable G-protein coupled receptor 142	GPR142 is a member of the rhodopsin family of G protein-coupled receptors (GPRs) (Fredriksson et al., 2003 [PubMed 14623098]).[supplied by OMIM, Mar 2008]. Transcript exon combination :: AB196530.1, AY288421.1 [ECO:0000332]	49
GPRC5C	exonic	55890	G-protein coupled receptor family C group 5 member C isoform a	The protein encoded by this gene is a member of the type 3 G protein-coupled receptor family. Members of this superfamily are characterized by a signature 7-transmembrane domain motif. The specific function of this protein is unknown; however, this protein may mediate the cellular effects of retinoic acid on the G protein signal transduction cascade. Two transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (a). Transcript exon combination :: BC110848.1, AK131210.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025083 [ECO:0000348]	50
GRIA3	intronic	2892	glutamate receptor 3 isoform 1 precursor	Glutamate receptors are the predominant excitatory neurotransmitter receptors in the mammalian brain and are activated in a variety of normal neurophysiologic processes. These receptors are heteromeric protein complexes composed of multiple subunits, arranged to form ligand-gated ion channels. The classification of glutamate receptors is based on their activation by different pharmacologic agonists. The subunit encoded by this gene belongs to a family of AMPA (alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionate)-sensitive glutamate receptors, and is subject to RNA editing (AGA->GGA; R->G). Alternative splicing at this locus results in different isoforms, which may vary in their signal transduction properties, [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) encodes isoform 1 (also known as flip isoform). RNA editing (AGA->GGA) changes Arg775Gly. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene	51

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				record to access additional publications. ##RefSeq-Attributes-START## undergoes RNA editing :: PMID: 10688364, 7992055 ##RefSeq-Attributes-END## Transcript exon combination :: U10301.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025082, ERS025084 [ECO:0000348]	
GTPBP4	exonic	23560	nucleolar GTP-binding protein 1	GTP-binding proteins are GTPases and function as molecular switches that can flip between two states: active, when GTP is bound, and inactive, when GDP is bound. 'Active' in this context usually means that the molecule acts as a signal to trigger other events in the cell. When an extracellular ligand binds to a G-protein-linked receptor, the receptor changes its conformation and switches on the trimeric G proteins that associate with it by causing them to eject their GDP and replace it with GTP. The switch is turned off when the G protein hydrolyzes its own bound GTP, converting it back to GDP. But before that occurs, the active protein has an opportunity to diffuse away from the receptor and deliver its message for a prolonged period to its downstream target. [provided by RefSeq, Jul 2008]. Transcript exon combination :: AK001552.1, AK222861.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	52
HCN1	exonic	348980	potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 1	The membrane protein encoded by this gene is a hyperpolarization-activated cation channel that contributes to the native pacemaker currents in heart and neurons. The encoded protein can homodimerize or heterodimerize with other pore-forming subunits to form a potassium channel. This channel may act as a receptor for sour tastes. [provided by RefSeq, Oct 2011]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AF488549.1, AF064876.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	53
HEXA	exonic	3073	beta-hexosaminidase subunit alpha preproprotein	This gene encodes the alpha subunit of the lysosomal enzyme beta-hexosaminidase that, together with the cofactor GM2 activator protein, catalyzes the degradation of the ganglioside GM2, and other molecules containing terminal N-acetyl hexosamines. Beta-hexosaminidase is composed of two subunits, alpha and beta, which are encoded by separate genes. Both beta-hexosaminidase alpha and beta subunits are members of family 20 of glycosyl hydrolases. Mutations in the alpha or beta subunit genes lead to an accumulation of GM2 ganglioside in neurons and neurodegenerative disorders termed the GM2 gangliosidoses. Alpha subunit gene mutations lead to Tay-Sachs disease (GM2-gangliosidosis type I). [provided by RefSeq, Jul 2009]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data because no single transcript was available for the full length of the gene. The extent of this transcript is supported by transcript alignments. Publication Note: This RefSeq record includes a subset of the	54

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: M13520.1, CR627386.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	
HK2	exonic	3099	hexokinase-2	Hexokinases phosphorylate glucose to produce glucose-6-phosphate, the first step in most glucose metabolism pathways. This gene encodes hexokinase 2, the predominant form found in skeletal muscle. It localizes to the outer membrane of mitochondria. Expression of this gene is insulin-responsive, and studies in rat suggest that it is involved in the increased rate of glycolysis seen in rapidly growing cancer cells. [provided by RefSeq, Apr 2009]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC064369.1, AF148513.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025083, ERS025084 [ECO:0000348]	55
HMGB3	exonic	3149	high mobility group protein B3	HMGB3 belongs to the high mobility group (HMG) protein superfamily. Like HMG1 (MIM 163905) and HMG2 (MIM 163906), HMGB3 contains DNA-binding HMG box domains and is classified into the HMG box subfamily. Members of the HMG box subfamily are thought to play a fundamental role in DNA replication, nucleosome assembly and transcription (Wilke et al., 1997 [PubMed 9370291]; Nemeth et al., 2006 [PubMed 16945912]). [supplied by OMIM, Mar 2008]. Transcript exon combination :: Y10043.1, BG176733.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	56
HPR	exonic	3250	haptoglobin-related protein precursor	This gene encodes a haptoglobin-related protein that binds hemoglobin as efficiently as haptoglobin. Unlike haptoglobin, plasma concentration of this protein is unaffected in patients with sickle cell anemia and extensive intravascular hemolysis, suggesting a difference in binding between haptoglobin-hemoglobin and haptoglobin-related protein-hemoglobin complexes to CD163, the hemoglobin scavenger receptor. This protein may also be a clinically important predictor of recurrence of breast cancer. [provided by RefSeq, Oct 2011]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: CB147217.1, CB122261.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	57
HTATSFIP 2	exonic	401233	N/A	N/A	58
IDI2	exonic	91734	isopentenyl-diphosphate Delta-isomerase 2	N/A	59
IDI2-AS1	exonic	55853	N/A	N/A	60
IDO2	intronic	169355	indoleamine 2,3-dioxygenase 2	Along with the enzymes encoded by the INDO (MIM 147435) and TDO2 (MIM 191070) genes, the enzyme encoded by the INDOL 1 gene metabolizes tryptophan in the kynurenine pathway (Ball et al., 2007 [PubMed 17499941]).[supplied by OMIM, Feb 2011]. Sequence Note: The RefSeq transcript 3' UTR was derived from	61

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				genomic	
				sequence to make the sequence consistent with the reference genome assembly. The genomic coordinates used were based on transcript alignments.	
IFNLR1	exonic	163702	interferon lambda receptor 1 isoform 1 precursor	The protein encoded by this gene belongs to the class II cytokine receptor family. This protein forms a receptor complex with interleukine 10 receptor, beta (IL10RB). The receptor complex has been shown to interact with three closely related cytokines, including interleukin 28A (IL28A), interleukin 28B (IL28B), and interleukin 29 (IL29). The expression of all three cytokines can be induced by viral infection. The cells overexpressing this protein have been found to have enhanced responses to IL28A and IL29, but decreased response to IL28B. Three alternatively spliced transcript variants encoding distinct isoforms have been reported. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) represents the longest transcript and it encodes the longest protein (isoform 1). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AF439325.1, AK160364.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084 [ECO:0000348]	62
IQCB1	exonic	9657	IQ calmodulin-binding motif-containing protein 1 isoform a	This gene encodes a nephrocystin protein that interacts with calmodulin and the retinitis pigmentosa GTPase regulator protein. The encoded protein has a central coiled-coil region and two calmodulin-binding IQ domains. It is localized to the primary cilia of renal epithelial cells and connecting cilia of photoreceptor cells. The protein is thought to play a role in ciliary function. Defects in this gene result in Senior-Loken syndrome type 5. Alternative splicing results in multiple transcript variants. [provided by RefSeq, Nov 2009]. Transcript Variant: This variant (1) encodes the longer isoform (a). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: D25278.1, AY714228.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	63
JPX	intronic	554203		JPX is a nonprotein-coding RNA transcribed from a gene within the X-inactivation center (XIC; MIM 314670) that appears to participate in X chromosome inactivation (Tian et al., 2010 [PubMed 21029862]). [supplied by OMIM, Feb 2011]. Transcript exon combination :: BC071776.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	64
KANK1	intronic	23189	KN motif and ankyrin repeat domain-containing protein 1 isoform a	The protein encoded by this gene belongs to the Kank family of proteins, which contain multiple ankyrin repeat domains. This family member functions in cytoskeleton formation by regulating actin polymerization. This gene is a candidate tumor suppressor for renal cell carcinoma. Mutations in this gene cause cerebral palsy spastic quadriplegic type 2, a central nervous system development disorder. A t(5;9) translocation results in fusion of the platelet-derived growth factor receptor beta gene	65

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				(PDGFRB) on chromosome 5 with this gene in a myeloproliferative neoplasm featuring severe thrombocythemia. Alternative splicing of this gene results in multiple transcript variants. A related pseuodgene has been identified on chromosome 20. [provided by RefSeq, Mar 2012]. Transcript Variant: This variant (1) represents the shortest transcript but encodes the longer isoform (a, also known as Kank-L). Variants 1, 3 and 4 all encode isoform a. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AL833161.1, AK292989.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025085 [ECO:0000348]	
KAT6B	exonic	23522	histone acetyltransfer ase KAT6B isoform 1	The protein encoded by this gene is a histone acetyltransferase and component of the MOZ/MORF protein complex. In addition to its acetyltransferase activity, the encoded protein has transcriptional activation activity in its N-terminal end and transcriptional repression activity in its C-terminal end. This protein is necessary for RUNX2-dependent transcriptional activation and could be involved in brain development. Mutations have been found in patients with genitopatellar syndrome. A translocation of this gene and the CREBBP gene results in acute myeloid leukemias. Three transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Mar 2012]. Transcript Variant: This variant (1) represents the longest transcript and encodes the longest isoform (1). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AF217500.1, BC150618.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025083, ERS025084 [ECO:0000348]	66
KCTD7	exonic	154881	BTB/POZ domain-containing protein KCTD7 isoform 1	This gene encodes a member of the potassium channel tetramerization domain-containing protein family. Family members are identified on a structural basis and contain an amino-terminal domain similar to the T1 domain present in the voltage-gated potassium channel. Mutations in this gene have been associated with progressive myoclonic epilepsy-3. Alternative splicing results in multiple transcript variants.[provided by RefSeq, Jan 2011]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (1). Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Transcript exon combination :: AK056631.1, BU902852.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025083 [ECO:0000348]	67
KIF19	exonic	124602	kinesin-like protein KIF19	N/A	68

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
LARP4B	exonic	23185	la-related protein 4B	This gene encodes a member of an evolutionarily conserved protein family implicated in RNA metabolism and translation. Members of this family are characterized by the presence of an La motif, which is often located adjacent to one or more RNA recognition motifs (RRM). Together, the two motifs constitute the functional region of the protein and enable its interaction with the RNA substrate. This protein family is divided into five sub-families: the genuine La proteins and four La-related protein (LARP) sub-families. The protein encoded by this gene belongs to LARP subfamily 4. It is a cytoplasmic protein that may play a stimulatory role in translation. [provided by RefSeq, Oct 2012]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. CDS exon combination :: BC152443.1, D86971.2 [ECO:0000331] RNAseq introns :: mixed/partial sample support ERS025088 [ECO:0000350]	69
LOC64333 9	exonic	643339	N/A	N/A	70
LOH12CR1	exonic	118426	loss of heterozygosit y 12 chromosomal region 1 protein	N/A	71
MALL	exonic	7851	MAL-like protein	This gene encodes an element of the machinery for raft-mediated trafficking in endothelial cells. The encoded protein, a member of the MAL proteolipid family, predominantly localizes in glycolipid- and cholesterol-enriched membrane (GEM) rafts. It interacts with caveolin-1. [provided by RefSeq, Jul 2008]. Transcript exon combination :: AK125647.1, AK056616.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	72
MAPK9	exonic	5601	mitogen-activated protein kinase 9 isoform alpha1	The protein encoded by this gene is a member of the MAP kinase family. MAP kinases act as an integration point for multiple biochemical signals, and are involved in a wide variety of cellular processes such as proliferation, differentiation, transcription regulation and development. This kinase targets specific transcription factors, and thus mediates immediate-early gene expression in response to various cell stimuli. It is most closely related to MAPK8, both of which are involved in UV radiation induced apoptosis, thought to be related to the cytochrome c-mediated cell death pathway. This gene and MAPK8 are also known as c-Jun N-terminal kinases. This kinase blocks the ubiquitination of tumor suppressor p53, and thus it increases the stability of p53 in nonstressed cells. Studies of this gene's mouse counterpart suggest a key role in T-cell differentiation. Several alternatively spliced transcript variants encoding distinct isoforms have been reported. [provided by RefSeq, Sep 2008]. Transcript Variant: This variant (JNK2-a1) uses a different acceptor splice site in the last coding exon compared to transcript variant JNK2-a2, resulting in a frameshift and a shorter isoform (JNK2 alpha1) with a different C-terminus, compared to isoform JNK2 alpha2. The JNK2-a1 variant differs from the JNK2-b1 variant in the use of an alternate internal coding exon of the same length. Thus, JNK2 alpha1 isoform is the same length as JNK2 beta1 isoform, with a	73

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				few aa differences in an internal protein segment. Sequence Note: This RefSeq record was created from transcript and genomic sequence data because no single transcript was available for the full length of the gene. The extent of this transcript is supported by transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. CDS exon combination :: U34821.1 [ECO:0000331] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	
MCEE	both	84693	methylmalony 1-CoA epimerase, mitochondrial precursor	The product of this gene catalyzes the interconversion of D- and L-methylmalonyl-CoA during the degradation of branched chain amino acids. odd chain-length fatty acids, and other metabolites. Mutations in this gene result in methylmalonyl-CoA epimerase deficiency, which is presented as mild to moderate methylmalonic aciduria. [provided by RefSeq, Jul 2008]. Transcript exon combination :: BC020825.1, BG567074.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	74
MGAT5	intronic	4249	alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase A	The protein encoded by this gene belongs to the glycosyltransferase family. It catalyzes the addition of beta-1,6-N-acetylglucosamine to the alpha-linked mannose of biantennary N-linked oligosaccharides present on the newly synthesized glycoproteins. It is one of the most important enzymes involved in the regulation of the biosynthesis of glycoprotein oligosaccharides. Alterations of the oligosaccharides on cell surface glycoproteins cause significant changes in the adhesive or migratory behavior of a cell. Increase in the activity of this enzyme has been correlated with the progression of invasive malignancies. [provided by RefSeq, Oct 2011]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: D17716.1, AF113921.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	75
MGC16275	exonic	85001	N/A	N/A	76
MGME1	exonic	92667	mitochondrial genome maintenance exonuclease 1	N/A	77
MIR200A	exonic	406983		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the	78

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.	
MIR200B	exonic	406984		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.	79
MIR429	exonic	554210		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009].	80

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications.	
MIR595	exonic	693180		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage.	81
MIR651	exonic	723779		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage.	82
MIR3163	exonic	1004230 29		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or	83

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature	
				miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage.	
MIR3910-1	exonic	1005008 21		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage.	84
MIR3910-2	exonic	1005009 02		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by	85

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage.	
MIR4267	exonic	1004229 94		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage.	86
MIR4436B 1	exonic	1006161 23		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star (miRNA*) products. The mature miRNA is incorporated into a RNA-induced silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage.	87
MIR4436B 2	exonic	1008470 33		microRNAs (miRNAs) are short (20-24 nt) non-coding RNAs that are involved in post-transcriptional regulation of gene expression in multicellular organisms by affecting both the stability and translation of mRNAs. miRNAs are transcribed by RNA polymerase II as part of capped and polyadenylated primary transcripts (pri-miRNAs) that can be either protein-coding or non-coding. The primary transcript is cleaved by the Drosha ribonuclease III enzyme to produce an approximately 70-nt stem-loop precursor miRNA (pre-miRNA), which is further cleaved by the cytoplasmic Dicer ribonuclease to generate the mature miRNA and antisense miRNA star	88

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				(miRNA*) products. The mature miRNA is incorporated into a RNA-induced	
				silencing complex (RISC), which recognizes target mRNAs through imperfect base pairing with the miRNA and most commonly results in translational inhibition or destabilization of the target mRNA. The RefSeq represents the predicted microRNA stem-loop. [provided by RefSeq, Sep 2009]. Sequence Note: This record represents a predicted microRNA stem-loop as defined by miRBase. Some sequence at the 5' and 3' ends may not be included in the intermediate precursor miRNA produced by Drosha cleavage.	
MKL1	intronic	57591	MKL/myocar din-like protein 1	The protein encoded by this gene interacts with the transcription factor myocardin, a key regulator of smooth muscle cell differentiation. The encoded protein is predominantly nuclear and may help transduce signals from the cytoskeleton to the nucleus. This gene is involved in a specific translocation event that creates a fusion of this gene and the RNA-binding motif protein-15 gene. This translocation has been associated with acute megakaryocytic leukemia. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AB037859.2, AJ297258.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	89
MRPL42	exonic	28977	39S ribosomal protein L42, mitochondrial precursor	Mammalian mitochondrial ribosomal proteins are encoded by nuclear genes and help in protein synthesis within the mitochondrion. Mitochondrial ribosomes (mitoribosomes) consist of a small 28S subunit and a large 39S subunit. They have an estimated 75% protein to rRNA composition compared to prokaryotic ribosomes, where this ratio is reversed. Another difference between mammalian mitoribosomes and prokaryotic ribosomes is that the latter contain a 5S rRNA. Among different species, the proteins comprising the mitoribosome differ greatly in sequence, and sometimes in biochemical properties, which prevents easy recognition by sequence homology. This gene encodes a protein identified as belonging to both the 28S and the 39S subunits. Alternative splicing results in multiple transcript variants. Pseudogenes corresponding to this gene are found on chromosomes 4q, 6p, 6q, 7p, and 15q. [provided by RefSeq, May 2011]. Transcript Variant: This variant (1) encodes the supported protein. Both variants 1 and 2 encode the same protein. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. ##RefSeq-Attributes-START## gene product(s) localized to mito. :: reported by MitoCarta ##RefSeq-Attributes-END## Transcript exon combination :: AK000285.1, AF151038.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	90
MTHFD1	exonic	4522	C-1-tetrahydrofola te synthase, cytoplasmic	This gene encodes a protein that possesses three distinct enzymatic activities, 5,10-methylenetetrahydrofolate dehydrogenase, 5,10-methenyltetrahydrofolate cyclohydrolase and 10-	91

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				formyltetrahydrofolate synthetase. Each of these activities catalyzes one of three sequential reactions in the interconversion of 1-carbon derivatives of	
				tetrahydrofolate, which are substrates for methionine, thymidylate, and de novo purine syntheses. The trifunctional enzymatic activities are conferred by two major domains, an aminoterminal portion containing the dehydrogenase and cyclohydrolase activities and a larger synthetase domain. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. ##RefSeq-Attributes-START## CDS uses downstream in-frame AUG :: experimental evidence (PMID:3053686) ##RefSeq-Attributes-END## Transcript exon combination :: BC050420.1, J04031.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	
NFIL3	exonic	4783	nuclear factor interleukin-3-regulated protein	Expression of interleukin-3 (IL3; MIM 147740) is restricted to activated T cells, natural killer (NK) cells, and mast cell lines. Transcription initiation depends on the activating capacity of specific protein factors, such as NFIL3, that bind to regulatory regions of the gene, usually upstream of the transcription start site (Zhang et al., 1995 [PubMed 7565758]).[supplied by OMIM, Feb 2009]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: S79880.1, U26173.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	92
NLRP12	exonic	91662	NACHT, LRR and PYD domains-containing protein 12 isoform 2	This gene encodes a member of the CATERPILLER family of cytoplasmic proteins. The encoded protein, which contains an N-terminal pyrin domain, a NACHT domain, a NACHT-associated domain, and a C-terminus leucine-rich repeat region, functions as an attenuating factor of inflammation by suppressing inflammatory responses in activated monocytes. Mutations in this gene cause familial cold autoinflammatory syndrome type 2. Alternative splicing results in multiple transcript variants. [provided by RefSeq, Mar 2013]. Transcript Variant: This variant (2) uses an alternate splice site in the central coding region, compared to variant 3, resulting in an isoform (2) that is 1 aa shorter than isoform 3. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AY095146.1, BC028069.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025089 [ECO:0000348]	93
NQO2	exonic	4835	ribosyldihydronicotinamide dehydrogenase [quinone]	NQO2 (EC 1.10.99.2) is a flavoprotein that catalyzes the 2-electron reduction of various quinones, redox dyes, and the vitamin K menadione. NQO2 predominantly uses dihydronicotinamide riboside (NRH) as the electron donor (summary by Wu et al., 1997 [PubMed 9367528]).[supplied by OMIM, Jul 2010]. Publication Note: This RefSeq record includes a subset of the publications that are	94

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: J02888.1, AK311746.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	
NRIP 1	exonic	8204	nuclear receptor-interacting protein 1	Nuclear receptor interacting protein 1 (NRIP1) is a nuclear protein that specifically interacts with the hormone-dependent activation domain AF2 of nuclear receptors. Also known as RIP140, this protein modulates transcriptional activity of the estrogen receptor. [provided by RefSeq, Jul 2008]. Sequence Note: The RefSeq transcript and protein were derived from transcript and genomic sequence to make the sequence consistent with the reference genome assembly. The extent of this RefSeq transcript is supported by transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AK289786.1, DA230125.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025098 [ECO:0000348]	95
NUDT4	exonic	11163	diphosphoino sitol polyphosphat e phosphohydro lase 2 isoform alpha	The protein encoded by this gene regulates the turnover of diphosphoinositol polyphosphates. The turnover of these high-energy diphosphoinositol polyphosphates represents a molecular switching activity with important regulatory consequences. Molecular switching by diphosphoinositol polyphosphates may contribute to regulating intracellular trafficking. Several alternatively spliced transcript variants have been described, but the full-length nature of some variants has not been determined. Isoforms DIPP2alpha and DIPP2beta are distinguishable from each other solely by DIPP2beta possessing one additional amino acid due to intron boundary skidding in alternate splicing. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) encodes the predominant isoform (alpha). Transcript exon combination :: AF191651.1, AF191650.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	96
NUDT4P1	exonic	440672	N/A	N/A	97
OVOL2	exonic	58495	transcription factor Ovo-like 2	N/A	98
PDE3B	intronic	5140	cGMP-inhibited 3',5'-cyclic phosphodieste rase B	N/A	99
PDGFRA	exonic	5156	platelet-derived growth factor receptor alpha precursor	This gene encodes a cell surface tyrosine kinase receptor for members of the platelet-derived growth factor family. These growth factors are mitogens for cells of mesenchymal origin. The identity of the growth factor bound to a receptor monomer determines whether the functional receptor is a homodimer or a heterodimer, composed of both platelet-derived growth factor receptor alpha and beta polypeptides. Studies suggest that this gene plays a role in organ development, wound healing, and tumor progression. Mutations in this gene have been associated with idiopathic hypereosinophilic syndrome, somatic and familial gastrointestinal stromal tumors, and a variety of other cancers. [provided by RefSeq, Mar 2012]. Sequence Note:	100

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				This RefSeq record was created from transcript and genomic sequence data because no single transcript was available for the full length of the gene. The extent of this transcript is supported by transcript alignments and orthologous data. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: M21574.1, M22734.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025084 [ECO:0000348]	
PDSS2	exonic	57107	decaprenyl-diphosphate synthase subunit 2	The protein encoded by this gene is an enzyme that synthesizes the prenyl side-chain of coenzyme Q, or ubiquinone, one of the key elements in the respiratory chain. The gene product catalyzes the formation of all trans-polyprenyl pyrophosphates from isopentyl diphosphate in the assembly of polyisoprenoid side chains, the first step in coenzyme Q biosynthesis. Defects in this gene are a cause of coenzyme Q10 deficiency.[provided by RefSeq, Oct 2009]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC039906.1, AF254956.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	101
PHACTR4	exonic	65979	phosphatase and actin regulator 4 isoform 1	This gene encodes a member of the phosphatase and actin regulator (PHACTR) family. Other PHACTR family members have been shown to inhibit protein phosphatase 1 (PP1) activity, and the homolog of this gene in the mouse has been shown to interact with actin and PP1. Multiple transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) represents the longer transcript but encodes the shorter isoform (1). Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Transcript exon combination :: CR749449.1, BC029266.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025087 [ECO:0000348]	102
PIAS2	exonic	9063	E3 SUMO-protein ligase PIAS2 isoform alpha	This gene encodes a member of the protein inhibitor of activated STAT (PIAS) family. PIAS proteins function as SUMO E3 ligases and play important roles in many cellular processes by mediating the sumoylation of target proteins. Alternatively spliced transcript variants encoding multiple isoforms have been observed for this gene. Isoforms of the encoded protein enhance the sumoylation of specific target proteins including the p53 tumor suppressor protein, c-Jun, and the androgen receptor. A pseudogene of this gene is located on the short arm of chromosome 4. The symbol MIZ 1 has also been associated with ZBTB 17 which is a different gene located on chromosome 1. [provided by RefSeq, Aug 2011]. Transcript Variant: This variant (alpha) utilizes an alternate 3' coding exon, compared to variant beta, resulting in a shorter isoform (alpha) that has a unique C-terminus compared to isoform beta. Sequence Note: This RefSeq record was	103

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC015190.1 [ECO:0000332] RNAseq	
				introns :: single sample supports all introns ERS025088 [ECO:0000348]	
PIK3CD	exonic	5293	phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit delta isoform	Phosphoinositide 3-kinases (PI3Ks) phosphorylate inositol lipids and are involved in the immune response. The protein encoded by this gene is a class I PI3K found primarily in leukocytes. Like other class I PI3Ks (p110-alpha p110-beta, and p 110-gamma), the encoded protein binds p85 adapter proteins and GTP-bound RAS. However, unlike the other class I PI3Ks, this protein phosphorylates itself, not p85 protein, [provided by RefSeq, Jul 2010]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: U86453.1, Y10055.2 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025089 [ECO:0000348]	104
PKHD1	intronic	5314	fibrocystin isoform 1 precursor	The protein encoded by this gene is predicted to have a single transmembrane (TM)-spanning domain and multiple copies of an immunoglobulin-like plexin-transcription-factor domain. Alternative splicing results in two transcript variants encoding different isoforms. Other alternatively spliced transcripts have been described, but the full length sequences have not been determined. Several of these transcripts are predicted to encode truncated products which lack the TM and may be secreted. Mutations in this gene cause autosomal recessive polycystic kidney disease, also known as polycystic kidney and hepatic disease-1. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) encodes the longer isoform of this protein. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AY074797.1, AF480064.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025084, ERS025085 [ECO:0000350]	105
PLXNC1	exonic	10154	Plexin C1	This gene encodes a member of the plexin family. Plexins are transmembrane receptors for semaphorins, a large family of proteins that regulate axon guidance, cell motility and migration, and the immune response. The encoded protein and its ligand regulate melanocyte adhesion, and viral semaphorins may modulate the immune response by binding to this receptor. The encoded protein may be a tumor suppressor protein for melanoma. Alternatively spliced transcript variants have been observed for this gene. [provided by RefSeq, Jan 2011]. Transcript Variant: This variant (2) lacks multiple 5' exons but contains an alternate 5' exon, compared to variant 1. This variant is represented as non-coding due to the presence of an upstream ORF that is predicted to interfere with translation of the longest in-frame ORF.	106

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				Translation of the upstream ORF renders the transcript a candidate for nonsense-mediated mRNA decay (NMD). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	
PNPLA4	exonic	8228	patatin-like phospholipase domain-containing protein 4 isoform 1 precursor	This gene encodes a member of the patatin-like family of phospholipases. The encoded enzyme has both triacylglycerol lipase and transacylase activities and may be involved in adipocyte triglyceride homeostasis. Alternate splicing results in multiple transcript variants. A pseudogene of this gene is found on chromosome Y. [provided by RefSeq, Feb 2010]. Transcript Variant: This variant (1) represents the longest transcript and encodes the longer isoform (1). Variants 1 and 2 encode the same isoform (1). Sequence Note: The RefSeq transcript and protein were derived from transcript and genomic sequence to make the sequence consistent with the reference genome assembly. The extent of this transcript is supported by transcript alignments. Transcript exon combination :: U03886.1, AK289888.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025091, ERS025098 [ECO:0000348]	107
PNPT1	both	87178	polyribonucleotide nucleotidyltransferase 1, mitochondrial precursor	The protein encoded by this gene belongs to the evolutionary conserved polynucleotide phosphorylase family comprised of phosphate dependent 3'-to-5' exoribonucleases implicated in RNA processing and degradation. This enzyme is predominantly localized in the mitochondrial intermembrane space and is involved in import of RNA to mitochondria. Mutations in this gene have been associated with combined oxidative phosphorylation deficiency-13 and autosomal recessive nonsyndromic deafness-70. Related pseudogenes are found on chromosomes 3 and 7. [provided by RefSeq, Dec 2012]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC053660.1, AJ458465.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350] ##RefSeq-Attributes-START## gene product(s) localized to mito. :: PMID: 12798676; reported by MitoCarta ##RefSeq-Attributes-END##	108
PPP2R3B	intronic	28227	serine/threonine-protein phosphatase 2A regulatory subunit B" subunit beta	Protein phosphatase 2 (formerly named type 2A) is one of the four major Ser/Thr phosphatases and is implicated in the negative control of cell growth and division. Protein phosphatase 2 holoenzymes are heterotrimeric proteins composed of a structural subunit A, a catalytic subunit C, and a regulatory subunit B. The regulatory subunit is encoded by a diverse set of genes that have been grouped into the B/PR55, B'/PR61, and B"/PR72 families. These different regulatory subunits confer distinct enzymatic specificities and intracellular localizations to the holozenzyme. The product of this gene belongs to the B" family. The B" family has been further divided into subfamilies. The product of this gene belongs to the beta subfamily of regulatory subunit B". [provided by RefSeq, Apr 2010]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to	109

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Transcript exon combination :: BK000521.1, BC063429.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084 [ECO:0000348]	
PRKCB	both	5579	protein kinase C beta type isoform 1	Protein kinase C (PKC) is a family of serine- and threonine-specific protein kinases that can be activated by calcium and second messenger diacylglycerol. PKC family members phosphorylate a wide variety of protein targets and are known to be involved in diverse cellular signaling pathways. PKC family members also serve as major receptors for phorbol esters, a class of tumor promoters. Each member of the PKC family has a specific expression profile and is believed to play a distinct role in cells. The protein encoded by this gene is one of the PKC family members. This protein kinase has been reported to be involved in many different cellular functions, such as B cell activation, apoptosis induction, endothelial cell proliferation, and intestinal sugar absorption. Studies in mice also suggest that this kinase may also regulate neuronal functions and correlate fear-induced conflict behavior after stress. Alternatively spliced transcript variants encoding distinct isoforms have been reported. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) uses an alternate splice junction at the 5' end of the last exon compared to variant 2. The resulting isoform (1) has a distinct and shorter C-terminus compared to isoform 2. Sequence Note: This RefSeq record was created from transcript and genomic sequence data because no single transcript was available for the full length of the gene. The extent of this transcript is supported by transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: X06318.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025082, ERS025084 [ECO:0000348]	110
PRKCH	intronic	5583	protein kinase C eta type	Protein kinase C (PKC) is a family of serine- and threonine-specific protein kinases that can be activated by calcium and the second messenger diacylglycerol. PKC family members phosphorylate a wide variety of protein targets and are known to be involved in diverse cellular signaling pathways. PKC family members also serve as major receptors for phorbol esters, a class of tumor promoters. Each member of the PKC family has a specific expression profile and is believed to play a distinct role in cells. The protein encoded by this gene is one of the PKC family members. It is a calcium-independent and phospholipids-dependent protein kinase. It is predominantly expressed in epithelial tissues and has been shown to reside specifically in the cell nucleus. This protein kinase can regulate keratinocyte differentiation by activating the MAP kinase MAPK13 (p38delta)-activated protein kinase cascade that targets CCAAT/enhancer-binding protein alpha (CEBPA). It is also found to mediate the transcription activation of the transglutaminase 1 (TGM1) gene. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that	111

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC037268.1, AK290183.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025083 [ECO:0000348]	
PSTPIP1	exonic	9051	proline-serine-threonine phosphatase-interacting protein 1	The protein encoded by this gene binds to the cytoplasmic tail of CD2, an effector of T cell activation and adhesion, negatively affecting CD2-triggered T cell activation. The encoded protein appears to be a scaffold protein and a regulator of the actin cytoskeleton. It has also been shown to bind ABL1, PTPN18, WAS, CD2AP, and PTPN12. Mutations in this gene are a cause of PAPA syndrome. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC008602.1, U94778.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	112
PTPN2	exonic	5771	tyrosine-protein phosphatase non-receptor type 2 isoform 1	The protein encoded by this gene is a member of the protein tyrosine phosphatase (PTP) family. Members of the PTP family share a highly conserved catalytic motif, which is essential for the catalytic activity. PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. Epidermal growth factor receptor and the adaptor protein Shc were reported to be substrates of this PTP, which suggested the roles in growth factor mediated cell signaling. Multiple alternatively spliced transcript variants encoding different isoforms have been found. Two highly related but distinctly processed pseudogenes that localize to chromosomes 1 and 13, respectively, have been reported. [provided by RefSeq, May 2011]. Transcript Variant: This variant (1) encodes the longest isoform (1). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: M25393.1, AK292570.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	113
PTPRN2	intronic	5799	receptor-type tyrosine-protein phosphatase N2 isoform 1 precursor	The protein encoded by this gene is a member of the protein tyrosine phosphatase (PTP) family. PTPs are known to be signaling molecules that regulate a variety of cellular processes including cell growth, differentiation, mitotic cycle, and oncogenic transformation. This PTP possesses an extracellular region, a single transmembrane region, and a single intracellular catalytic domain, and thus represents a receptor-type PTP. The catalytic domain of this PTP is most closely related to PTPRN/IA-2beta. This PTP and PTPRN are both found to be major autoantigens associated with insulin-dependent diabetes mellitus. Three alternatively spliced transcript variants of this gene, which encode distinct proteins, have been reported. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) encodes the longest isoform (1). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination ::	114

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				U66702.1, AF007555.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	
RAB37	exonic	326624	ras-related protein Rab-37 isoform 2	Rab proteins are low molecular mass GTPases that are critical regulators of vesicle trafficking. For additional background information on Rab proteins, see MIM 179508. [supplied by OMIM, Apr 2006]. Transcript Variant: This variant (2) represents use of an alternate promoter, 5' UTR, and alternate start codon, and includes an alternate coding exon, compared to variant 3. The resulting isoform (2) has a distinct and longer N-terminus, compared to isoform 3. Transcript exon combination :: AK098068.1, BX332255.2 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	115
RBFOX1	intronic	54715	RNA binding protein fox-1 homolog 1 isoform 1	The Fox-1 family of RNA-binding proteins is evolutionarily conserved, and regulates tissue-specific alternative splicing in metazoa. Fox-1 recognizes a (U)GCAUG stretch in regulated exons or in flanking introns. The protein binds to the C-terminus of ataxin-2 and may contribute to the restricted pathology of spinocerebellar ataxia type 2 (SCA2). Ataxin-2 is the product of the SCA2 gene which causes familial neurodegenerative diseases. Fox-1 and ataxin-2 are both localized in the trans-Golgi network. Several alternatively spliced transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Nov 2011]. Transcript Variant: This variant (1), also known as gamma, encodes the longest isoform (1). Sequence Note: This RefSeq record was created from transcript and genomic sequence data because no single transcript was available for the full length of the gene. The extent of this transcript is supported by transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AF229057.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025083, ERS025099 [ECO:0000348]	116
RCC1	exonic	1104	N/A	N/A	117
RGCC	exonic	28984	regulator of cell cycle RGCC	This gene is thought to regulate cell cycle progression. It is induced by p53 in response to DNA damage, or by sublytic levels of complement system proteins that result in activation of the cell cycle. The encoded protein localizes to the cytoplasm during interphase and to centrosomes during mitosis. The protein forms a complex with polo-like kinase 1. The protein also translocates to the nucleus in response to treatment with complement system proteins, and can associate with and increase the kinase activity of cell division cycle 2 protein. In different assays and cell types, overexpression of this protein has been shown to activate or suppress cell cycle progression. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC066334.1, BG037019.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025092 [ECO:0000348]	118
RHOQ	intronic	23433	rho-related GTP-	This gene encodes a member of the Rho family of	119

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
			binding protein RhoQ precursor	small GTPases, which cycle between inactive GDP-bound and active GTP-bound states and function as molecular switches in signal transduction cascades. Rho proteins promote reorganization of the actin cytoskeleton and regulate cell shape, attachment, and motility. The encoded protein is an important signalling protein for sarcomere assembly and has been shown to play a significant role in the exocytosis of the solute carrier family 2, facilitated glucose transporter member 4 and other proteins, possibly acting as the signal that turns on the membrane fusion machinery. Three related pseudogene have been identified on chromosomes 2 and 14. [provided by RefSeq, Aug 2011]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BX428852.2, BC013135.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	
RNASE3	exonic	6037	eosinophil cationic protein precursor	N/A	120
RNASE10	exonic	338879	inactive ribonuclease-like protein 10 precursor	N/A	121
RPL38	exonic	6169	60S ribosomal protein L38	Ribosomes, the organelles that catalyze protein synthesis, consist of a small 40S subunit and a large 60S subunit. Together these subunits are composed of 4 RNA species and approximately 80 structurally distinct proteins. This gene encodes a ribosomal protein that is a component of the 60S subunit. The protein belongs to the L38E family of ribosomal proteins. It is located in the cytoplasm. Alternative splice variants have been identified, both encoding the same protein. As is typical for genes encoding ribosomal proteins, there are multiple processed pseudogenes of this gene dispersed through the genome, including one located in the promoter region of the type 1 angiotensin II receptor gene. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) is the longer and predominant transcript. Variants 1 and 2 encode the same protein. Transcript exon combination :: BQ276548.1, BU569438.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	122
RPTOR	intronic	57521	regulatory-associated protein of mTOR isoform 1	This gene encodes a component of a signaling pathway that regulates cell growth in response to nutrient and insulin levels. The encoded protein forms a stoichiometric complex with the mTOR kinase, and also associates with eukaryotic initiation factor 4E-binding protein-1 and ribosomal protein S6 kinase. The protein positively regulates the downstream effector ribosomal protein S6 kinase, and negatively regulates the mTOR kinase. Multiple transcript variants encoding different isoforms have been found for this gene. [provided by RefSeq, Sep 2009]. Transcript Variant: This variant (1) represents the longer transcript and encodes the longer isoform (1). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AY090663.1,	123

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				BC136652.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025083, ERS025085 [ECO:0000348]	
SERPINB4	exonic	6318	serpin B4	N/A	124
SERPINB6	exonic	5269	serpin B6 isoform a	The protein encoded by this gene is a member of the serpin (serine proteinase inhibitor) superfamily, and ovalbumin(ov)-serpin subfamily. It was originally discovered as a placental thrombin inhibitor. The mouse homolog was found to be expressed in the hair cells of the inner ear. Mutations in this gene are associated with nonsyndromic progressive hearing loss, suggesting that this serpin plays an important role in the inner ear in the protection against leakage of lysosomal content during stress, and that loss of this protection results in cell death and sensorineural hearing loss. Alternatively spliced transcript variants have been found for this gene. [provided by RefSeq, Sep 2010]. Transcript Variant: This variant (1) represents the predominant transcript. Variants 1, 5 and 6 encode the same isoform (a). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AK314578.1, BC098564.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	125
SLC3A2	both	6520	4F2 cell-surface antigen heavy chain isoform b	This gene is a member of the solute carrier family and encodes a cell surface, transmembrane protein. The protein exists as the heavy chain of a heterodimer, covalently bound through di-sulfide bonds to one of several possible light chains. The encoded transporter plays a role in regulation of intracellular calcium levels and transports L-type amino acids. Alternatively spliced transcript variants, encoding different isoforms, have been characterized. [provided by RefSeq, Nov 2010]. Transcript Variant: This variant (2) represents the longest transcript and encodes the longest isoform (b). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AK025584.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025082, ERS025084 [ECO:0000350]	126
SLC17A5	both	26503	sialin	This gene encodes a membrane transporter that exports free sialic acids that have been cleaved off of cell surface lipids and proteins from lysosomes. Mutations in this gene cause sialic acid storage diseases, including infantile sialic acid storage disorder and and Salla disease, an adult form. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC020961.2, AJ387747.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	127
SNHG3	exonic	8420	N/A	N/A	128
SNORD17	exonic	692086	N/A	N/A	129
SNX5	exonic	27131	sorting nexin-5 isoform a	This gene encodes a member of the sorting nexin family. Members of this family contain a phox (PX) domain, which is a phosphoinositide binding	130

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				domain, and are involved in intracellular trafficking. This protein functions in endosomal sorting, the phosphoinositide-signaling pathway, and macropinocytosis. This gene may play a role in the tumorigenesis of papillary thyroid carcinoma. Alternative splicing results in multiple transcript variants encoding different isoforms. [provided by RefSeq, Sep 2013]. Transcript Variant: This variant (1) differs in the 5' UTR, compared to variant 2. Variants 1 and 2 encode the same protein (isoform a). Transcript exon combination :: BC000100.3, AF121855.1 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	
SOCS2	exonic	8835	suppressor of cytokine signaling 2	This gene encodes a member of the suppressor of cytokine signaling (SOCS) family. SOCS family members are cytokine-inducible negative regulators of cytokine receptor signaling via the Janus kinase/signal transducer and activation of transcription pathway (the JAK/STAT pathway). SOCS family proteins interact with major molecules of signaling complexes to block further signal transduction, in part, by proteasomal depletion of receptors or signal-transducing proteins via ubiquitination. The expression of this gene can be induced by a subset of cytokines, including erythropoietin, GM-CSF, IL10, interferon (IFN)-gamma and by cytokine receptors such as growth horomone receptor. The protein encoded by this gene interacts with the cytoplasmic domain of insulin-like growth factor-1 receptor (IGF1R) and is thought to be involved in the regulation of IGFIR mediated cell signaling. This gene has pseudogenes on chromosomes 20 and 22. Alternative splicing results in multiple transcript variants. [provided by RefSeq, Jul 2012]. Transcript Variant: This variant (1) differs in the 5' UTR, compared to variant 5. Variants 1-6 encode the same protein. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AK313165.1, AL522912.3 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025082 [ECO:0000348]	131
SOCS2-AS1	exonic	144481	N/A	N/A	132
ST8SIA5	exonic	29906	alpha-2,8-sialyltransfera se 8E	The protein encoded by this gene is a type II membrane protein that may be present in the Golgi apparatus. The encoded protein, which is a member of glycosyltransferase family 29, may be involved in the synthesis of gangliosides GT1a, GQ1b, and GT3 from GD1a, GT1b, GM1b, and GD3, respectively. [provided by RefSeq, Jul 2008]. Transcript exon combination :: AK056270.1, BC108910.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025082, ERS025084 [ECO:0000348]	133
STIM2	intronic	57620	stromal interaction molecule 2 isoform 1 precursor	This gene is a member of the stromal interaction molecule (STIM) family and likely arose, along with related family member STIM1, from a common ancestral gene. The encoded protein functions to regulate calcium concentrations in the cytosol and endoplasmic reticulum, and is involved in the activation of plasma membrane Orai Ca(2+) entry channels. This gene initiates translation from a non-AUG (UUG) start site. A signal peptide is cleaved from the resulting protein. Multiple transcript variants result from	134

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				alternative splicing. [provided by RefSeq, Dec 2009]. Transcript Variant: This variant (1) encodes the longest isoform (1). Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC136449.1, AK096846.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025084 [ECO:0000348] ##RefSeq-Attributes-START## CDS uses downstream in-frame AUG :: experimental evidence (PMID:11463338) non-AUG initiation codon :: PMID: 11463338 ##RefSeq-Attributes-END##	
TBC1D16	intronic	125058	TBC1 domain family member 16 isoform a	N/A	136
TEX29	exonic	121793	testis-expressed sequence 29 protein	N/A	137
TNFRSF10 A	exonic	8797	tumor necrosis factor receptor superfamily member 10A	The protein encoded by this gene is a member of the TNF-receptor superfamily. This receptor is activated by tumor necrosis factor-related apoptosis inducing ligand (TNFSF10/TRAIL), and thus transduces cell death signal and induces cell apoptosis. Studies with FADD-deficient mice suggested that FADD, a death domain containing adaptor protein, is required for the apoptosis mediated by this protein, [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC012866.1, AK291299.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025081, ERS025084 [ECO:0000348]	138
TNFRSF13 C	exonic	115650	tumor necrosis factor receptor superfamily member 13C	B cell-activating factor (BAFF) enhances B-cell survival in vitro and is a regulator of the peripheral B-cell population. Overexpression of Baff in mice results in mature B-cell hyperplasia and symptoms of systemic lupus erythematosus (SLE). Also, some SLE patients have increased levels of BAFF in serum. Therefore, it has been proposed that abnormally high levels of BAFF may contribute to the pathogenesis of autoimmune diseases by enhancing the survival of autoreactive B cells. The protein encoded by this gene is a receptor for BAFF and is a type III transmembrane protein containing a single extracellular cysteine-rich domain. It is thought that this receptor is the principal receptor required for BAFF-mediated mature B-cell survival. [provided by RefSeq, Jul 2008]. Sequence Note: The RefSeq transcript and protein were derived from genomic sequence to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AF373846.1, BC112030.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025089 [ECO:0000348]	139
TNFRSF18	exonic	8784	tumor necrosis factor receptor superfamily	This gene encodes a member of the TNF-receptor superfamily. The encoded receptor has been shown to have increased expression upon T-cell	140

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
			member 18 isoform 1 precursor	activation, and it is thought to play a key role in dominant immunological self-tolerance maintained by CD25(+)CD4(+) regulatory T cells. Knockout studies in mice also suggest the role of this receptor is in the regulation of CD3-driven T-cell activation and programmed cell death. Three alternatively spliced transcript variants of this gene encoding distinct isoforms have been reported. [provided by RefSeq, Feb 2011]. Transcript Variant: This variant (1) represents the longest transcript. It contains an extra coding segment, which leads to a frame shift, compared to variant 2. The resulting preotein (isoform 1) contains a distinct and shorter C-terminus, as compared to isoform 2. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: AY358877.1, AF125304.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025089, ERS025093 [ECO:0000348]	
TRAFD1	exonic	10906	TRAF-type zinc finger domain-containing protein 1	The innate immune system confers host defense against viral and microbial infection, and TRAFD 1 is a negative feedback regulator that controls excessive immune responses (Sanada et al., 2008 [PubMed 18849341]).[supplied by OMIM, Dec 2009]. Transcript Variant: This variant (1) represents the longer transcript. Variants 1 and 2 both encode the same protein. Transcript exon combination :: AK122620.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025083, ERS025084 [ECO:0000348]	141
TRPM2	exonic	7226	Transient Receptor Potential Cation Channel Subfamily M Member 2	The protein encoded by this gene is a calcium-permeable cation channel that is regulated by free intracellular ADP-ribose. The encoded protein is activated by oxidative stress and confers susceptibility to cell death. Several alternatively spliced transcript variants of this gene have been described, but their full-length nature is not known. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (2) uses an alternate in-frame splice junction at the 5' end of an exon compared to variant 1. This results in the introduction of a premature stop codon and renders the transcript a nonsense-mediated mRNA decay (NMD) candidate. Therefore, this transcript is not thought to be protein-coding. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. RNAseq introns :: mixed/partial sample support ERS025081, ERS025082 [ECO:0000350]	142
TTLL10	exonic	254173	inactive polyglycylase TTLL10 isoform 1	N/A	143
TTYH2	exonic	94015	protein tweety homolog 2 isoform 1	This gene encodes a member of the tweety family of proteins. Members of this family function as chloride anion channels. The encoded protein functions as a calcium(2+)-activated large conductance chloride(-) channel, and may play a role in kidney tumorigenesis. Two transcript variants encoding distinct isoforms have been	144

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				identified for this gene. [provided by RefSeq, Jul 2008]. Transcript Variant: This variant (1) represents the longer transcript, and encodes the longer isoform (1). Transcript exon combination :: AF319952.1, BC107492.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025082, ERS025083 [ECO:0000348]	
UBE2N	exonic	7334	ubiquitin-conjugating enzyme E2 N	The modification of proteins with ubiquitin is an important cellular mechanism for targeting abnormal or short-lived proteins for degradation. Ubiquitination involves at least three classes of enzymes: ubiquitin-activating enzymes, or	145
				E1s, ubiquitin-conjugating enzymes, or E2s, and ubiquitin-protein ligases, or E3s. This gene encodes a member of the E2 ubiquitin-conjugating enzyme family. Studies in mouse suggest that this protein plays a role in DNA postreplication repair. [provided by RefSeq, Jul 2008]. Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination :: BC000396.2, D83004.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084 [ECO:0000348]	
VCX	exonic	26609	variable charge X-linked protein 1	This gene belongs to the VCX/Y gene family, which has multiple members on both X and Y chromosomes, and all are expressed exclusively in male germ cells. The X-linked members are clustered on chromosome Xp22 and Y-linked members are two identical copies of the gene within a palindromic region on Yq11. The family members share a high degree of sequence identity, with the exception that a 30-bp unit is tandemly repeated in X-linked members but occurs only once in Y-linked members. The VCX gene cluster is polymorphic in terms of copy number; different individuals may have a different number of VCX genes. VCX/Y genes encode small and highly charged proteins of unknown function. The presence of a putative bipartite nuclear localization signal suggests that VCX/Y members are nuclear proteins. This gene contains 10 repeats of the 30-bp unit. [provided by RefSeq, Jul 2008]. Transcript exon combination :: AF167081.2 [ECO:0000332]	146
VSTM1	intronic	284415	V-set and transmembran e domain-containing protein 1 precursor	N/A	147
VWA2	exonic	340706	von Willebrand factor A domain-containing protein 2 precursor	This gene encodes a member of the von Willebrand factor A-like domain protein superfamily. The encoded protein is localized to the extracellular matrix and may serve as a structural component in basement membranes or in anchoring structures on scaffolds of collagen VII or fibrillin. This gene has been linked to type 1A diabetes and is a candidate serological marker for colon cancer. [provided by RefSeq, Jan 2013]. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments. CCDS Note: The coding region has been updated to represent an alternative 3' splicing pattern that is more supported by the available transcript and protein	148

RefSeq Gene Symbol	Exon overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene # (GN)
				data. Transcript exon combination :: AY572972.1, AJ536328.2 [ECO:0000332] RNAseq introns :: mixed/partial sample support ERS025081, ERS025084 [ECO:0000350]	
ZNF350	exonic	59348	zinc finger protein 350	N/A	149
ZNF432	exonic	9668	zinc finger protein 432	N/A	150
ZNF577	exonic	84765	N/A	N/A	151
ZNF613	exonic	79898	zinc finger protein 613 isoform 1	N/A	152
ZNF614	exonic	80110	zinc finger protein 614	N/A	153
ZNF615	exonic	284370	zinc finger protein 615 isoform 1	N/A	154
ZNF649	exonic	65251	zinc finger protein 649	N/A	155
ZNF841	exonic	284371	zinc finger protein 841	N/A	156

[0318] For all genes listed in **Table 2** (namely, those relevant to CNV-subregions of interest), **Table 3** represents a non-redundant list.

Table 4: A non-redundant list of transcript variants that correspond to the genes in **Table 3**

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
MIR200B	exonic	NR 029639	Homo sapiens microRNA 200b (MIR200B), microRNA.	173
MIR200A	exonic	NR 029834	Homo sapiens microRNA 200a (MIR200A), microRNA.	174
MIR429	exonic	NR 029957	Homo sapiens microRNA 429 (MIR429), microRNA.	175
TTLL10	exonic	NM 001130045	Homo sapiens tubulin tyrosine ligase-like family, member 10 (TTLL10), transcript variant 1, mRNA.	176
TTLL10	exonic	NM 153254	Homo sapiens tubulin tyrosine ligase-like family, member 10 (TTLL10), transcript variant 2, mRNA.	177
TNFRSF18	exonic	NM 004195	Homo sapiens tumor necrosis factor receptor superfamily, member 18 (TNFRSF18), transcript variant 1, mRNA.	178
TNFRSF18	exonic	NM 148901	Homo sapiens tumor necrosis factor receptor superfamily, member 18 (TNFRSF18), transcript variant 2, mRNA.	179
TNFRSF18	exonic	NM 148902	Homo sapiens tumor necrosis factor receptor superfamily, member 18 (TNFRSF18), transcript variant 3, mRNA.	180
PIK3CD	exonic	NM 005026	Homo sapiens phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit delta (PIK3CD), mRNA.	181
CAPZB	intronic	NR 038125	Homo sapiens capping protein (actin filament) muscle Z-line, beta (CAPZB), transcript variant 4, non-coding RNA.	182
CAPZB	intronic	NM 001206540	Homo sapiens capping protein (actin filament) muscle Z-line, beta (CAPZB), transcript variant 2, mRNA.	183
CAPZB	intronic	NM 004930	Homo sapiens capping protein (actin filament) muscle Z-line, beta (CAPZB), transcript variant 1, mRNA.	184
IFNLR1	exonic	NM 170743	Homo sapiens interferon, lambda receptor 1 (IFNLR1), transcript variant 1, mRNA.	185
IFNLR1	exonic	NM 173064	Homo sapiens interferon, lambda receptor 1 (IFNLR1), transcript variant 2, mRNA.	186
IFNLR1	exonic	NM 173065	Homo sapiens interferon, lambda receptor 1 (IFNLR1), transcript variant 3, mRNA.	187
PHACTR4	exonic	NM 001048183	Homo sapiens phosphatase and actin regulator 4 (PHACTR4), transcript variant 1, mRNA.	188
PHACTR4	exonic	NM 023923	Homo sapiens phosphatase and actin regulator 4 (PHACTR4), transcript variant 2, mRNA.	189
SNHG3	exonic	NR 002909	Homo sapiens small nucleolar RNA host gene 3 (non-protein coding) (SNHG3), transcript variant 2, non-coding RNA.	190

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
SNHG3	exonic	NR 036473	Homo sapiens small nucleolar RNA host gene 3 (non-protein coding) (SNHG3), transcript variant 1, non-coding RNA.	191
RCC1	exonic	NM 001048199	Homo sapiens regulator of chromosome condensation 1 (RCC1), transcript variant 4, mRNA.	192
RCC1	exonic	NR 030725	Homo sapiens regulator of chromosome condensation 1 (RCC1), transcript variant 5, non-coding RNA.	193
RCC1	exonic	NR 030726	Homo sapiens regulator of chromosome condensation 1 (RCC1), transcript variant 6, non-coding RNA.	194
RCC1	exonic	NM 001048194	Homo sapiens regulator of chromosome condensation 1 (RCC1), transcript variant 1, mRNA.	195
RCC1	exonic	NM 001048195	Homo sapiens regulator of chromosome condensation 1 (RCC1), transcript variant 2, mRNA.	196
RCC1	exonic	NM 001269	Homo sapiens regulator of chromosome condensation 1 (RCC1), transcript variant 3, mRNA.	197
AGBL4	intronic	NM 032785	Homo sapiens ATP/GTP binding protein-like 4 (AGBL4), mRNA.	198
GPATCH2	intronic	NM 018040	Homo sapiens G patch domain containing 2 (GPATCH2), mRNA.	199
RHOQ	intronic	NM 012249	Homo sapiens ras homolog family member Q (RHOQ), mRNA.	200
PNPT1	both	NM 033109	Homo sapiens polyribonucleotide nucleotidyltransferase 1 (PNPT1), mRNA.	201
MCEE	both	NM 032601	Homo sapiens methylmalonyl CoA epimerase (MCEE), mRNA.	202
HK2	exonic	NM 000189	Homo sapiens hexokinase 2 (HK2), mRNA.	203
FHL2	exonic	NM 201557	Homo sapiens four and a half LIM domains 2 (FHL2), transcript variant 4, mRNA.	204
FHL2	intronic	NM 001039492	Homo sapiens four and a half LIM domains 2 (FHL2), transcript variant 5, mRNA.	205
FHL2	intronic	NM 001450	Homo sapiens four and a half LIM domains 2 (FHL2), transcript variant 1, mRNA.	206
FHL2	intronic	NM 201555	Homo sapiens four and a half LIM domains 2 (FHL2), transcript variant 2, mRNA.	207
MIR4267	exonic	NR 036225	Homo sapiens microRNA 4267 (MIR4267), microRNA.	208
MALL	exonic	NM 005434	Homo sapiens mal, T-cell differentiation protein-like (MALL), mRNA.	209
MIR4436B1	exonic	NR 039941	Homo sapiens microRNA 4436b-1 (MIR4436B1), microRNA.	210
MIR4436B2	exonic	NR 049830	Homo sapiens microRNA 4436b-2 (MIR4436B2), microRNA.	211
MGAT5	intronic	NM 002410	Homo sapiens mannosyl (alpha-1,6-)-glycoprotein beta-1,6-N-acetyl-glucosaminyltransferase (MGAT5), mRNA.	212
BMPR2	intronic	NM 001204	Homo sapiens bone morphogenetic protein receptor, type II (serine/threonine kinase) (BMPR2), mRNA.	213
DNER	intronic	NM 139072	Homo sapiens delta/notch-like EGF repeat containing (DNER), mRNA.	214
IQCB1	exonic	NM 001023570	Homo sapiens IQ motif containing B 1 (IQCB 1), transcript variant 1, mRNA.	215
IQCB1	exonic	NM 001023571	Homo sapiens IQ motif containing B 1 (IQCB 1), transcript variant 3, mRNA.	216
STIM2	intronic	NM 001169117	Homo sapiens stromal interaction molecule 2 (STIM2), transcript variant 3, mRNA.	217
STIM2	intronic	NM 00116911 8	Homo sapiens stromal interaction molecule 2 (STIM2), transcript variant 1, mRNA.	218
STIM2	intronic	NM 020860	Homo sapiens stromal interaction molecule 2 (STIM2), transcript variant 2, mRNA.	219
PDGFRA	exonic	NM 006206	Homo sapiens platelet-derived growth factor receptor, alpha polypeptide (PDGFRA), mRNA.	220
HCN1	exonic	NM 021072	Homo sapiens hyperpolarization activated cyclic nucleotide-gated potassium channel 1 (HCN1), mRNA.	221
EMB	exonic	NM 198449	Homo sapiens embigin (EMB), mRNA.	222

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
EDIL3	intronic	NM 005711	Homo sapiens EGF-like repeats and discoidin I-like domains 3 (EDIL3), transcript variant 1, mRNA.	223
MAPK9	exonic	NM 002752	Homo sapiens mitogen-activated protein kinase 9 (MAPK9), transcript variant JNK2-a2, mRNA.	224
MAPK9	exonic	NM 139068	Homo sapiens mitogen-activated protein kinase 9 (MAPK9), transcript variant JNK2-a1, mRNA.	225
MAPK9	exonic	NM 139069	Homo sapiens mitogen-activated protein kinase 9 (MAPK9), transcript variant JNK2-b1, mRNA.	226
MAPK9	exonic	NM 139070	Homo sapiens mitogen-activated protein kinase 9 (MAPK9), transcript variant JNK2-b2, mRNA.	227
MAPK9	exonic	NM 001135044	Homo sapiens mitogen-activated protein kinase 9 (MAPK9), transcript variant JNK2-g, mRNA.	228
SERPINB6	exonic	NM 001271 825	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 6 (SERPINB6), transcript variant 6, mRNA.	229
SERPINB6	exonic	NM 001271823	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 6 (SERPINB6), transcript variant 4, mRNA.	230
SERPINB6	exonic	NM 001271 822	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 6 (SERPINB6), transcript variant 3, mRNA.	231
SERPINB6	exonic	NM 001195291	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 6 (SERPINB6), transcript variant 2, mRNA.	232
SERPINB6	exonic	NM 001271824	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 6 (SERPINB6), transcript variant 5, mRNA.	233
SERPINB6	exonic	NM 004568	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 6 (SERPINB6), transcript variant 1, mRNA.	234
DKFZP686I 15217	exonic	NR 026855	Homo sapiens long intergenic non-protein coding RNA 1011 (LINC01011), transcript variant 1, non-coding RNA.	235
DKFZP686I 15217	exonic	NR 026856	Homo sapiens long intergenic non-protein coding RNA 1011 (LINC01011), transcript variant 2, non-coding RNA.	236
NQO2	exonic	NM 000904	Homo sapiens NAD(P)H dehydrogenase, quinone 2 (NQO2), mRNA.	237
HTATSFIP 2	exonic	NR 033884	Homo sapiens HIV-1 Tat specific factor 1 pseudogene 2 (HTATSF1P2), non-coding RNA.	238
PKHD1	intronic	NM 138694	Homo sapiens polycystic kidney and hepatic disease 1 (autosomal recessive) (PKHD1), transcript variant 1, mRNA.	239
PKHD1	intronic	NM 170724	Homo sapiens polycystic kidney and hepatic disease 1 (autosomal recessive) (PKHD1), transcript variant 2, mRNA.	240
SLC17A5	both	NM 012434	Homo sapiens solute carrier family 17 (acidic sugar transporter), member 5 (SLC17A5), mRNA.	241
PDSS2	exonic	NM 020381	Homo sapiens prenyl (decaprenyl) diphosphate synthase, subunit 2 (PDSS2), mRNA.	242
KCTD7	exonic	NM 001167961	Homo sapiens potassium channel tetramerization domain containing 7 (KCTD7), transcript variant 2, mRNA.	243
KCTD7	exonic	NM 153033	Homo sapiens potassium channel tetramerization domain containing 7 (KCTD7), transcript variant 1, mRNA.	244
PTPRN2	intronic	NM 002847	Homo sapiens protein tyrosine phosphatase, receptor type, N polypeptide 2 (PTPRN2), transcript variant 1, mRNA.	245
PTPRN2	intronic	NM 130842	Homo sapiens protein tyrosine phosphatase, receptor type, N polypeptide 2 (PTPRN2), transcript variant 2, mRNA.	246
PTPRN2	intronic	NM 130843	Homo sapiens protein tyrosine phosphatase, receptor type, N polypeptide 2 (PTPRN2), transcript variant 3, mRNA.	247
MIR595	exonic	NR 030325	Homo sapiens microRNA 595 (MIR595), microRNA.	248
TNFRSF10 A	exonic	NM 003844	Homo sapiens tumor necrosis factor receptor superfamily, member 10a (TNFRSF10A), mRNA.	249
IDO2	intronic	NM 194294	Homo sapiens indoleamine 2,3-dioxygenase 2 (IDO2), mRNA.	250
STK3	intronic	NM 001256313	Homo sapiens serine/threonine kinase 3 (STK3), transcript variant 3, mRNA.	251
STK3	intronic	NM 006281	Homo sapiens serine/threonine kinase 3 (STK3), transcript variant 1, mRNA.	252

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
STK3	intronic	NM 001256312	Homo sapiens serine/threonine kinase 3 (STK3), transcript variant 2, mRNA.	253
KANK1	intronic	NM 001256876	Homo sapiens KN motif and ankyrin repeat domains 1 (KANK1), transcript variant 3, mRNA.	254
KANK1	intronic	NM 001256877	Homo sapiens KN motif and ankyrin repeat domains 1 (KANK1), transcript variant 4, mRNA.	255
KANK1	intronic	NM 015158	Homo sapiens KN motif and ankyrin repeat domains 1 (KANK1), transcript variant 1, mRNA.	256
KANK1	intronic	NM 153186	Homo sapiens KN motif and ankyrin repeat domains 1 (KANK1), transcript variant 2, mRNA.	257
GDA	exonic	NM 001242507	Homo sapiens guanine deaminase (GDA), transcript variant 4, mRNA.	258
GDA	exonic	NM 001242505	Homo sapiens guanine deaminase (GDA), transcript variant 1, mRNA.	259
GDA	exonic	NM 001242506	Homo sapiens guanine deaminase (GDA), transcript variant 3, mRNA.	260
GDA	exonic	NM 004293	Homo sapiens guanine deaminase (GDA), transcript variant 2, mRNA.	261
AUH	exonic	NM 001698	Homo sapiens AU RNA binding protein/enoyl-CoA hydratase (AUH), mRNA.	262
MIR3163	exonic	NR 036121	Homo sapiens microRNA 3163 (MIR3163), microRNA.	263
NFIL3	exonic	NM 005384	Homo sapiens nuclear factor, interleukin 3 regulated (NFIL3), mRNA.	264
MIR3910-1	exonic	NR 037472	Homo sapiens microRNA 3910-1 (MIR3910-1), microRNA.	265
MIR3910-2	exonic	NR 037489	Homo sapiens microRNA 3910-2 (MIR3910-2), microRNA.	266
ASTN2	intronic	NM 014010	Homo sapiens astrotactin 2 (ASTN2), transcript variant 1, mRNA.	267
ASTN2	intronic	NM 198186	Homo sapiens astrotactin 2 (ASTN2), transcript variant 2, mRNA.	268
ASTN2	intronic	NM 001184734	Homo sapiens astrotactin 2 (ASTN2), transcript variant 5, mRNA.	269
ASTN2	intronic	NM 198187	Homo sapiens astrotactin 2 (ASTN2), transcript variant 3, mRNA.	270
ASTN2	intronic	NM 198188	Homo sapiens astrotactin 2 (ASTN2), transcript variant 4, mRNA.	271
ASTN2	intronic	NM 001184735	Homo sapiens astrotactin 2 (ASTN2), transcript variant 6, mRNA.	272
LARP4B	exonic	NM 015155	Homo sapiens La ribonucleoprotein domain family, member 4B (LARP4B), mRNA.	273
GTPBP4	exonic	NM 012341	Homo sapiens GTP binding protein 4 (GTPBP4), mRNA.	274
IDI2	exonic	NM 033261	Homo sapiens isopentenyl-diphosphate delta isomerase 2 (IDI2), mRNA.	275
IDI2-AS1	exonic	NR 024628	Homo sapiens IDI2 antisense RNA 1 (IDI2-AS1), transcript variant 1, non-coding RNA.	276
IDI2-AS1	exonic	NR 024629	Homo sapiens IDI2 antisense RNA 1 (IDI2-AS1), transcript variant 2, non-coding RNA.	277
IDI2-AS1	exonic	NR 027708	Homo sapiens IDI2 antisense RNA 1 (IDI2-AS1), transcript variant 3, non-coding RNA.	278
IDI2-AS1	exonic	NR 027709	Homo sapiens IDI2 antisense RNA 1 (IDI2-AS1), transcript variant 4, non-coding RNA.	279
KAT6B	exonic	NM 001256468	Homo sapiens K(lysine) acetyltransferase 6B (KAT6B), transcript variant 2, mRNA.	280
KAT6B	exonic	NM 001256469	Homo sapiens K(lysine) acetyltransferase 6B (KAT6B), transcript variant 3, mRNA.	281
KAT6B	exonic	NM 012330	Homo sapiens K(lysine) acetyltransferase 6B (KAT6B), transcript variant 1, mRNA.	282
VWA2	exonic	NM 001272046	Homo sapiens von Willebrand factor A domain containing 2 (VWA2), mRNA.	283

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
PDE3B	intronic	NM 000922	Homo sapiens phosphodiesterase 3B, cGMP-inhibited (PDE3B), mRNA.	284
EHF	intronic	NM 001206615	Homo sapiens ets homologous factor (EHF), transcript variant 3, mRNA.	285
EHF	intronic	NM 012153	Homo sapiens ets homologous factor (EHF), transcript variant 2, mRNA.	286
EHF	exonic	NM 001206616	Homo sapiens ets homologous factor (EHF), transcript variant 1, mRNA.	287
SLC3A2	exonic	NM 001012662	Homo sapiens solute carrier family 3 (amino acid transporter heavy chain), member 2 (SLC3A2), transcript variant 2, mRNA.	288
SLC3A2	intronic	NM 001012664	Homo sapiens solute carrier family 3 (amino acid transporter heavy chain), member 2 (SLC3A2), transcript variant 5, mRNA.	289
SLC3A2	exonic	NM 002394	Homo sapiens solute carrier family 3 (amino acid transporter heavy chain), member 2 (SLC3A2), transcript variant 3, mRNA.	290
SLC3A2	intronic	NM 001013251	Homo sapiens solute carrier family 3 (amino acid transporter heavy chain), member 2 (SLC3A2), transcript variant 6, mRNA.	291
SLC3A2	intronic	NR 037193	Homo sapiens solute carrier family 3 (amino acid transporter heavy chain), member 2 (SLC3A2), transcript variant 7, non-coding RNA.	292
GDPD4	exonic	NM 182833	Homo sapiens glycerophosphodiester phosphodiesterase domain containing 4 (GDPD4), mRNA.	293
ETV6	exonic	NM 001987	Homo sapiens ets variant 6 (ETV6), mRNA.	294
LOH12CR1	exonic	NM 058169	Homo sapiens loss of heterozygosity, 12, chromosomal region 1 (LOH12CR1), mRNA.	295
DUSP16	exonic	NM 030640	Homo sapiens dual specificity phosphatase 16 (DUSP16), mRNA.	296
CREBL2	exonic	NM 001310	Homo sapiens cAMP responsive element binding protein-like 2 (CREBL2), mRNA.	297
GPR19	exonic	NM 006143	Homo sapiens G protein-coupled receptor 19 (GPR19), mRNA.	298
CDKN1B	exonic	NM 004064	Homo sapiens cyclin-dependent kinase inhibitor 1B (p27, Kip1) (CDKN1B), mRNA.	299
APOLD1	exonic	NM 001130415	Homo sapiens apolipoprotein L domain containing 1 (APOLD 1), transcript variant 1, mRNA.	300
APOLD1	intronic	NM 030817	Homo sapiens apolipoprotein L domain containing 1 (APOLD 1), transcript variant 2, mRNA.	301
EEA1	exonic	NM 003566	Homo sapiens early endosome antigen 1 (EEA1), mRNA.	302
LOC643339	exonic	NR 040096	Homo sapiens uncharacterized LOC643339 (LOC643339), non-coding RNA.	303
NUDT4	exonic	NM 019094	Homo sapiens nudix (nucleoside diphosphate linked moiety X)-type motif 4 (NUDT4), transcript variant 1, mRNA.	304
NUDT4	exonic	NM 199040	Homo sapiens nudix (nucleoside diphosphate linked moiety X)-type motif 4 (NUDT4), transcript variant 2, mRNA.	305
NUDT4P1	exonic	NR 002212	Homo sapiens nudix (nucleoside diphosphate linked moiety X)-type motif 4 pseudogene 1 (NUDT4P1), non-coding RNA.	306
UBE2N	exonic	NM 003348	Homo sapiens ubiquitin-conjugating enzyme E2N (UBE2N), mRNA.	307
MRPL42	exonic	NM 014050	Homo sapiens mitochondrial ribosomal protein L42 (MRPL42), transcript variant 1, mRNA.	308
MRPL42	exonic	NM 172177	Homo sapiens mitochondrial ribosomal protein L42 (MRPL42), transcript variant 2, mRNA.	309
MRPL42	exonic	NR 038159	Homo sapiens mitochondrial ribosomal protein L42 (MRPL42), transcript variant 3, non-coding RNA.	310
MRPL42	exonic	NR 038160	Homo sapiens mitochondrial ribosomal protein L42 (MRPL42), transcript variant 4, non-coding RNA.	311
MRPL42	exonic	NR 038161	Homo sapiens mitochondrial ribosomal protein L42 (MRPL42), transcript variant 5, non-coding RNA.	312
SOCS2-AS1	exonic	NR 038263	Homo sapiens SOCS2 antisense RNA 1 (SOCS2-AS1), non-coding RNA.	313

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
SOCS2	exonic	NM 003877	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), transcript variant 1, mRNA.	314
SOCS2	exonic	NM 001270467	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), transcript variant 2, mRNA.	315
SOCS2	exonic	NM 001270468	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), transcript variant 3, mRNA.	316
SOCS2	exonic	NM 001270469	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), transcript variant 4, mRNA.	317
SOCS2	exonic	NM 001270470	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), transcript variant 5, mRNA.	318
SOCS2	exonic	NM 001270471	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), transcript variant 6, mRNA.	319
CRADD	exonic	NM 003805	Homo sapiens CASP2 and RIPK1 domain containing adaptor with death domain (CRADD), mRNA.	320
PLXNC1	exonic	NM 005761	Homo sapiens plexin C1 (PLXNC1), transcript variant 1, mRNA.	321
PLXNC1	exonic	NR 037687	Homo sapiens plexin C1 (PLXNC1), transcript variant 2, non-coding RNA.	322
CCDC41	exonic	NM 001042399	Homo sapiens coiled-coil domain containing 41 (CCDC41), transcript variant 2, mRNA.	323
CCDC41	exonic	NM 016122	Homo sapiens coiled-coil domain containing 41 (CCDC41), transcript variant 1, mRNA.	324
TRAFD1	exonic	NM 001143906	Homo sapiens TRAF-type zinc finger domain containing 1 (TRAFD1), transcript variant 1, mRNA.	325
TRAFD1	exonic	NM 006700	Homo sapiens TRAF-type zinc finger domain containing 1 (TRAFD1), transcript variant 2, mRNA.	326
RGCC	exonic	NM 014059	Homo sapiens regulator of cell cycle (RGCC), mRNA.	327
CONIMD6	exonic	NM 203495	Homo sapiens COMM domain containing 6 (COMMD6), transcript variant 2, mRNA.	328
CONIMD6	exonic	NM 203497	Homo sapiens COMM domain containing 6 (COMMD6), transcript variant 1, mRNA.	329
GPC5	intronic	NM 004466	Homo sapiens glypican 5 (GPC5), mRNA.	330
ARHGEF7	exonic	NM 003899	Homo sapiens Rho guanine nucleotide exchange factor (GEF) 7 (ARHGEF7), transcript variant 1, mRNA.	331
ARHGEF7	exonic	NM 001113513	Homo sapiens Rho guanine nucleotide exchange factor (GEF) 7 (ARHGEF7), transcript variant 5, mRNA.	332
TEX29	exonic	NM 152324	Homo sapiens testis expressed 29 (TEX29), mRNA.	333
ARHGEF7	intronic	NM 001113511	Homo sapiens Rho guanine nucleotide exchange factor (GEF) 7 (ARHGEF7), transcript variant 3, mRNA.	334
ARHGEF7	intronic	NM 001113512	Homo sapiens Rho guanine nucleotide exchange factor (GEF) 7 (ARHGEF7), transcript variant 4, mRNA.	335
ARHGEF7	intronic	NM 145735	Homo sapiens Rho guanine nucleotide exchange factor (GEF) 7 (ARHGEF7), transcript variant 2, mRNA.	336
RNASE10	exonic	NM 001012975	Homo sapiens ribonuclease, RNase A family, 10 (non-active) (RNASE10), mRNA.	337
RNASE3	exonic	NM 002935	Homo sapiens ribonuclease, RNase A family, 3 (RNASE3), mRNA.	338
ECRP	exonic	NR 033909	Homo sapiens ribonuclease, RNase A family, 2 (liver, eosinophil-derived neurotoxin) pseudogene (ECRP), non-coding RNA.	339
PRKCH	intronic	NM 006255	Homo sapiens protein kinase C, eta (PRKCH), mRNA.	340
MTHFD1	exonic	NM 005956	Homo sapiens methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase (MTHFD1), mRNA.	341
BDKRB2	intronic	NM 000623	Homo sapiens bradykinin receptor B2 (BDKRB2), mRNA.	342
HEXA	exonic	NM 000520	Homo sapiens hexosaminidase A (alpha polypeptide) (HEXA), mRNA.	343

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
PSTPIP1	exonic	NM 003978	Homo sapiens proline-serine-threonine phosphatase interacting protein 1 (PSTPIP1), mRNA.	344
RBFOX 1	intronic	NM 001142333	Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 1 (RBFOX1), transcript variant 5, mRNA.	345
RBFOX 1	intronic	NM 018723	Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 1 (RBFOX1), transcript variant 4, mRNA.	346
RBFOX 1	intronic	NM 001142334	Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 1 (RBFOX1), transcript variant 6, mRNA.	347
RBFOX 1	intronic	NM 145891	Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 1 (RBFOX1), transcript variant 1, mRNA.	348
RBFOX 1	intronic	NM 145892	Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 1 (RBFOX1), transcript variant 2, mRNA.	349
RBFOX 1	intronic	NM 145893	Homo sapiens RNA binding protein, fox-1 homolog (C. elegans) 1 (RBFOX1), transcript variant 3, mRNA.	350
PRKCB	both	NM 002738	Homo sapiens protein kinase C, beta (PRKCB), transcript variant 2, mRNA.	351
PRKCB	both	NM 212535	Homo sapiens protein kinase C, beta (PRKCB), transcript variant 1, mRNA.	352
FUK	both	NM 145059	Homo sapiens fucokinase (FUK), mRNA.	353
COG4	exonic	NM 001195139	Homo sapiens component of oligomeric golgi complex 4 (COG4), transcript variant 2, mRNA.	354
COG4	exonic	NM 015386	Homo sapiens component of oligomeric golgi complex 4 (COG4), transcript variant 1, mRNA.	355
HPR	exonic	NM 020995	Homo sapiens haptoglobin-related protein (HPR), mRNA.	356
RPL38	exonic	NM 000999	Homo sapiens ribosomal protein L38 (RPL38), transcript variant 1, mRNA.	357
RPL38	exonic	NM 001035258	Homo sapiens ribosomal protein L38 (RPL38), transcript variant 2, mRNA.	358
MGC16275	exonic	NR 026914	Homo sapiens uncharacterized protein MGC16275 (MGC16275), non-coding RNA.	359
TTYH2	exonic	NM 032646	Homo sapiens tweety family member 2 (TTYH2), transcript variant 1, mRNA.	360
TTYH2	exonic	NM 052869	Homo sapiens tweety family member 2 (TTYH2), transcript variant 2, mRNA.	361
DNAI2	exonic	NM 001172810	Homo sapiens dynein, axonemal, intermediate chain 2 (DNAI2), transcript variant 2, mRNA.	362
DNAI2	exonic	NM 023036	Homo sapiens dynein, axonemal, intermediate chain 2 (DNAI2), transcript variant 1, mRNA.	363
KIF19	exonic	NM 153209	Homo sapiens kinesin family member 19 (KIF19), mRNA.	364
BTBD17	exonic	NM 001080466	Homo sapiens BTB (POZ) domain containing 17 (BTBD17), mRNA.	365
GPR142	exonic	NM 181790	Homo sapiens G protein-coupled receptor 142 (GPR142), mRNA.	366
GPRC5C	exonic	NM 022036	Homo sapiens G protein-coupled receptor, family C, group 5, member C (GPRC5C), transcript variant 1, mRNA.	367
GPRC5C	exonic	NM 018653	Homo sapiens G protein-coupled receptor, family C, group 5, member C (GPRC5C), transcript variant 2, mRNA.	368
CD300A	exonic	NM 001256841	Homo sapiens CD300a molecule (CD300A), transcript variant 2, mRNA.	369
CD300A	exonic	NM 007261	Homo sapiens CD300a molecule (CD300A), transcript variant 1, mRNA.	370
CD300LB	exonic	NM 174892	Homo sapiens CD300 molecule-like family member b (CD300LB), mRNA.	371
CD300C	exonic	NM 006678	Homo sapiens CD300c molecule (CD300C), mRNA.	372
CD300LD	exonic	NM 001115152	Homo sapiens CD300 molecule-like family member d (CD300LD), mRNA.	373
C17orf77	exonic	NM 152460	Homo sapiens chromosome 17 open reading frame 77	374

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
			(C17orf77), mRNA.	
CD300E	exonic	NM 181449	Homo sapiens CD300e molecule (CD300E), mRNA.	375
RAB37	exonic	NM 175738	Homo sapiens RAB37, member RAS oncogene family (RAB37), transcript variant 3, mRNA.	376
CD300LF	exonic	NM 139018	Homo sapiens CD300 molecule-like family member f (CD300LF), mRNA.	377
RAB37	intronic	NM 001163989	Homo sapiens RAB37, member RAS oncogene family (RAB37), transcript variant 4, mRNA.	378
RAB37	intronic	NM 001006638	Homo sapiens RAB37, member RAS oncogene family (RAB37), transcript variant 2, mRNA.	379
RAB37	intronic	NM 001163990	Homo sapiens RAB37, member RAS oncogene family (RAB37), transcript variant 5, mRNA.	380
TBC1D16	intronic	NM 019020	Homo sapiens TBC1 domain family, member 16 (TBC1D16), transcript variant 1, mRNA.	381
TBC1D16	intronic	NM 001271844	Homo sapiens TBC1 domain family, member 16 (TBC1D16), transcript variant 2, mRNA.	382
TBC1D16	intronic	NM 001271845	Homo sapiens TBC1 domain family, member 16 (TBC1D16), transcript variant 3, mRNA.	383
TBC1D16	intronic	NM 001271846	Homo sapiens TBC1 domain family, member 16 (TBC1D16), transcript variant 4, mRNA.	384
RPTOR	intronic	NM 001163034	Homo sapiens regulatory associated protein of MTOR, complex 1 (RPTOR), transcript variant 2, mRNA.	385
RPTOR	intronic	NM 020761	Homo sapiens regulatory associated protein of MTOR, complex 1 (RPTOR), transcript variant 1, mRNA.	386
PTPN2	exonic	NM 001207013	Homo sapiens protein tyrosine phosphatase, non-receptor type 2 (PTPN2), transcript variant 4, mRNA.	387
PTPN2	exonic	NM 080422	Homo sapiens protein tyrosine phosphatase, non-receptor type 2 (PTPN2), transcript variant 2, mRNA.	388
PTPN2	exonic	NM 080423	Homo sapiens protein tyrosine phosphatase, non-receptor type 2 (PTPN2), transcript variant 3, mRNA.	389
PTPN2	intronic	NM 002828	Homo sapiens protein tyrosine phosphatase, non-receptor type 2 (PTPN2), transcript variant 1, mRNA.	390
ST8SIA5	exonic	NM 013305	Homo sapiens ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 5 (ST8SIA5), mRNA.	391
PIAS2	exonic	NM 004671	Homo sapiens protein inhibitor of activated STAT, 2 (PIAS2), transcript variant beta, mRNA.	392
PIAS2	exonic	NM 173206	Homo sapiens protein inhibitor of activated STAT, 2 (PIAS2), transcript variant alpha, mRNA.	393
SERPINB4	exonic	NM 002974	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 4 (SERPINB4), mRNA.	394
FLJ26850	intronic	NR 027257	Homo sapiens FLJ26850 protein (FLJ26850), non-coding RNA.	395
FPR2	exonic	NM 001005738	Homo sapiens formyl peptide receptor 2 (FPR2), transcript variant 2, mRNA.	396
FPR2	exonic	NM 001462	Homo sapiens formyl peptide receptor 2 (FPR2), transcript variant 1, mRNA.	397
FPR3	exonic	NM 002030	Homo sapiens formyl peptide receptor 3 (FPR3), mRNA.	398
ZNF577	exonic	NR 024181	Homo sapiens zinc finger protein 577 (ZNF577), transcript variant 3, non-coding RNA.	399
ZNF577	exonic	NM 001135590	Homo sapiens zinc finger protein 577 (ZNF577), transcript variant 2, mRNA.	400
ZNF577	exonic	NM 032679	Homo sapiens zinc finger protein 577 (ZNF577), transcript variant 1, mRNA.	401
ZNF649	exonic	NM 023074	Homo sapiens zinc finger protein 649 (ZNF649), mRNA.	402
ZNF613	exonic	NM 001031721	Homo sapiens zinc finger protein 613 (ZNF613), transcript variant 1, mRNA.	403
ZNF613	exonic	NM 024840	Homo sapiens zinc finger protein 613 (ZNF613), transcript variant 2, mRNA.	404

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
ZNF350	exonic	NM 021632	Homo sapiens zinc finger protein 350 (ZNF350), mRNA.	405
ZNF615	exonic	NM 001199324	Homo sapiens zinc finger protein 615 (ZNF615), transcript variant 1, mRNA.	406
ZNF615	exonic	NM 198480	Homo sapiens zinc finger protein 615 (ZNF615), transcript variant 2, mRNA.	407
ZNF614	exonic	NM 025040	Homo sapiens zinc finger protein 614 (ZNF614), mRNA.	408
ZNF432	exonic	NM 014650	Homo sapiens zinc finger protein 432 (ZNF432), mRNA.	409
ZNF841	exonic	NM 001136499	Homo sapiens zinc finger protein 841 (ZNF841), mRNA.	410
NLRP12	exonic	NM 001277126	Homo sapiens NLR family, pyrin domain containing 12 (NLRP12), transcript variant 3, mRNA.	411
NLRP12	exonic	NM 001277129	Homo sapiens NLR family, pyrin domain containing 12 (NLRP12), transcript variant 4, mRNA.	412
NLRP12	exonic	NM 144687	Homo sapiens NLR family, pyrin domain containing 12 (NLRP12), transcript variant 2, mRNA.	413
VSTM1	intronic	NM 198481	Homo sapiens V-set and transmembrane domain containing 1 (VSTM1), mRNA.	414
SNX5	exonic	NM 014426	Homo sapiens sorting nexin 5 (SNX5), transcript variant 2, mRNA.	415
SNX5	exonic	NM 152227	Homo sapiens sorting nexin 5 (SNX5), transcript variant 1, mRNA.	416
SNORD 17	exonic	NR 003045	Homo sapiens small nucleolar RNA, C/D box 17 (SNORD17), small nucleolar RNA.	417
MGME1	exonic	NM 052865	Homo sapiens mitochondrial genome maintenance exonuclease 1 (MGME1), mRNA.	418
OVOL2	exonic	NM 021220	Homo sapiens ovo-like 2 (Drosophila) (OVOL2), mRNA.	419
ADA	intronic	NM 000022	Homo sapiens adenosine deaminase (ADA), mRNA.	420
NRIP 1	exonic	NM 003489	Homo sapiens nuclear receptor interacting protein 1 (NRIP1), mRNA.	421
BACH1	exonic	NR 027655	Homo sapiens BTB and CNC homology 1, basic leucine zipper transcription factor 1 (BACH1), transcript variant 3, non-coding RNA.	422
BACH1	intronic	NM 001186	Homo sapiens BTB and CNC homology 1, basic leucine zipper transcription factor 1 (BACH1), transcript variant 2, mRNA.	423
BACH1	intronic	NM 206866	Homo sapiens BTB and CNC homology 1, basic leucine zipper transcription factor 1 (BACH1), transcript variant 1, mRNA.	424
TRPM2	exonic	NM 003307	Homo sapiens transient receptor potential cation channel, subfamily M, member 2 (TRPM2), transcript variant 1, mRNA.	425
TRPM2	exonic	NR 038257	Homo sapiens transient receptor potential cation channel, subfamily M, member 2 (TRPM2), transcript variant 2, non-coding RNA.	426
ADARB1	intronic	NM 001112	Homo sapiens adenosine deaminase, RNA-specific, B 1 (ADARB1), transcript variant 1, mRNA.	427
ADARB1	intronic	NM 001160230	Homo sapiens adenosine deaminase, RNA-specific, B 1 (ADARB1), transcript variant 7, mRNA.	428
ADARB1	intronic	NM 015833	Homo sapiens adenosine deaminase, RNA-specific, B 1 (ADARB1), transcript variant 2, mRNA.	429
ADARB1	intronic	NM 015834	Homo sapiens adenosine deaminase, RNA-specific, B 1 (ADARB1), transcript variant 3, mRNA.	430
ADARB1	intronic	NR 027672	Homo sapiens adenosine deaminase, RNA-specific, B1 (ADARB1), transcript variant 5, non-coding RNA.	431
ADARB1	intronic	NR 027673	Homo sapiens adenosine deaminase, RNA-specific, B1 (ADARB1), transcript variant 4, non-coding RNA.	432
ADARB1	intronic	NR 027674	Homo sapiens adenosine deaminase, RNA-specific, B1 (ADARB1), transcript variant 6, non-coding RNA.	433
ADARB1	intronic	NR 073200	Homo sapiens adenosine deaminase, RNA-specific, B1 (ADARB1), transcript variant 8, non-coding RNA.	434
APOBEC3A	exonic	NM 001270406	Homo sapiens apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A (APOBEC3A), transcript variant 3,	435

RefSeq Gene Symbol	Exon overlap	RefSeq Accession Number	mRNA Description	SEQ ID
			mRNA.	
APOBEC3A	exonic	NM 145699	Homo sapiens apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3A (APOBEC3A), transcript variant 1, mRNA.	436
APOBEC3A B	intronic	NM 001193289	Homo sapiens APOBEC3A and APOBEC3B deletion hybrid (APOBEC3A B), mRNA.	437
APOBEC3B	exonic	NM 001270411	Homo sapiens apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B (APOBEC3B), transcript variant 2, mRNA.	438
APOBEC3B	exonic	NM 004900	Homo sapiens apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B (APOBEC3B), transcript variant 1, mRNA.	439
MKL1	intronic	NM 020831	Homo sapiens megakaryoblastic leukemia (translocation) 1 (MKL1), mRNA.	440
TNFRSF13 C	exonic	NM 052945	Homo sapiens tumor necrosis factor receptor superfamily, member 13C (TNFRSF13C), mRNA.	441
CENPM	exonic	NM 001110215	Homo sapiens centromere protein M (CENPM), transcript variant 3, mRNA.	442
CENPM	exonic	NM 001002876	Homo sapiens centromere protein M (CENPM), transcript variant 2, mRNA.	443
CENPM	exonic	NM 024053	Homo sapiens centromere protein M (CENPM), transcript variant 1, mRNA.	444
PPP2R3B	intronic	NM 013239	Homo sapiens protein phosphatase 2, regulatory subunit B", beta (PPP2R3B), mRNA.	445
VCX	exonic	NM 013452	Homo sapiens variable charge, X-linked (VCX), mRNA.	446
PNPLA4	exonic	NM 004650	Homo sapiens patatin-like phospholipase domain containing 4 (PNPLA4), transcript variant 1, mRNA.	447
PNPLA4	exonic	NM 001142389	Homo sapiens patatin-like phospholipase domain containing 4 (PNPLA4), transcript variant 2, mRNA.	448
PNPLA4	exonic	NM 001172672	Homo sapiens patatin-like phospholipase domain containing 4 (PNPLA4), transcript variant 3, mRNA.	449
MIR651	exonic	NR 030380	Homo sapiens microRNA 651 (MIR651), microRNA.	450
JPX	intronic	NR 024582	Homo sapiens JPX transcript, XIST activator (non-protein coding) (JPX), non-coding RNA.	451
GRIA3	intronic	NM 000828	Homo sapiens glutamate receptor, ionotropic, AMPA 3 (GRIA3), transcript variant 2, mRNA.	452
GRIA3	intronic	NM 007325	Homo sapiens glutamate receptor, ionotropic, AMPA 3 (GRIA3), transcript variant 1, mRNA.	453
GRIA3	intronic	NM 001256743	Homo sapiens glutamate receptor, ionotropic, AMPA 3 (GRIA3), transcript variant 3, mRNA.	454
HMGB3	exonic	NM 005342	Homo sapiens high mobility group box 3 (HMGB3), mRNA.	455

[0319] For all genes listed in **Table 2** (namely, those relevant to CNV-subregions of interest), **Table 4** represents a non-redundant list.

Table 5: The set of SNVs reported in Tables 7-10, 14, or 15 that were found in the 70 PML cases in this study for which WES data were generated.

Chromosome	Position 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

Chromosome	Position hg19	REF	ALT	SEQ ID
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
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

Chromosome	Position hg19	REF	ALT	SEQ ID
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
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
6	51524339	C	G	1108
6	51524409	G	T	1109

Chromosome	Position hg19	REF	ALT	SEQ ID
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
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

Chromosome	Position hg19	REF	ALT	SEQ ID
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	1179
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
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

Chromosome	Position hg19	REF	ALT	SEQ ID
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
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

Chromosome	Position hg19	REF	ALT	SEQ ID
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
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
21	45811411	G	T	1308
21	45811438	C	T	1309
21	45815307	T	C	1310
21	45815331	G	A	1311

Chromosome	Position hg19	REF	ALT	SEQ ID
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

[0320] Table 5 lists, in order of genomic coordinates, all single nucleotide variants (SNVs) that are relevant to the present study, whether as case-level solutions (Tables 7, 8) or potential solutions (Tables 9, 10), or at the level of variant burden analysis (Tables 14, 15). All genome coordinates are based on hg19.

Table 6: Non-redundant list of 419 genes involved in the immune system and/or linked to PML via a CNV

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
ACADM	AR	Public db	MySql	157
ACKR1	AD	Public db	MySql	158
ACP5	AR	Public db	PMID: 26052098, 27260006, 27821552	159
ADAR	ADAR	Public db	PMID: 26052098, 27260006, 27821552	160
ADARB1	unknown	PBio	PMID: 16227093, 17376196, 19482597, 20220309, 21682836, 21809195, 22001568, 22085847, 22113393, 24586166, 24725957, 24760760, 25826567	2
ADK	AR	PBio	PMID: 17205396, 23592612, 25654762, 25720338, 25979489, 26341819, 26642971	161
AGBL4	unknown	PBio	PMID: 17244818, 21074048, 23085998, 25416787, 25332286,26502776	3
AICDA	ADAR	Public db	MySql; PMID: 23765059	162
AK2	AR	Public db	PMID: 19043417, 19782549, 20008220, 23765059, 24135998, 24753205, 26454313	163
ALG12	AR	Public db	MySql	164
ALPL	ADAR	PBio	PMID: 18821074, 20049532, 20977932, 21191615, 21289095, 23091474, 23454488, 23860646, 26219705, 26219711,26219717	165
AP3B1	AR	Public db	MySql; PMID: 11590544, 19782549, 24302998, 24753205, 24916509, 25980904, 27889060	166
AP3B2	AR	Public db	PMID: 26377319, 27889060	167
AP3D1	AR	Public db	PMID: 26744459, 27889060	168
APOBEC3A	unknown	PBio	PMID: 16720547, 17303427, 20062055, 20615867, 22896697, 23344558, 23640892, 25262471, 25576866, 26416327,26489798,26678087	4

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
APOBEC3B	unknown	PBio	PMID: 16720547, 17303427, 20062055, 20615867, 22896697, 23344558, 23640892, 25262471, 25576866, 26416327,26489798,26678087	6
APOL1	association	Public db	PMID: 27042682	169
ARHGEF7	unknown	PBio	PMID: 11160719, 16983070, 18378701, 19861492, 25284783,25500533	8
ASH1L	unknown	PBio	PMID: 17923682, 17981149, 22541069, 24012418, 24244179, 25866973, 26002201, 27154821, 27229316, 27434206	170
ASTN2	unknown	PBio	PMID: 2223091,8602532,20573900,24357807,24381304,25146927, 25410587,26514622,26586575	9
ATL2	unknown	PBio	PMID: 18270207, 19665976, 25773277	171
ATM	AR	Public db	MySql; PMID: 19903823, 20301790, 23765059, 24799566, 25692705, 27042682, 27484032, 27884168, 27895165	172
ATR	ADAR	Public db	PMID: 17564965,17151099, 19903823, 20506465, 21615334, 24799566, 25910481	173
AUH	AR	PBio	PMID: 12434311, 12655555, 17130438, 20855850, 25280001,25597510	10
BACH1	unknown	PBio	PMID: 15068237, 18555605, 22024395, 22791292, 23456643, 23562577, 24752012, 25344725, 25391381, 24752012, 26045540, 26894991	11
BACH2	unknown	PBio	PMID: 17262715, 17991429, 18769450, 22791292, 23728300, 24367030, 24608439, 24681888, 24694524, 25123280, 25344725, 25665584, 25686607, 26444573, 26620562, 26731475, 26894991, 26981933	174
BCL10	AR	Public db	MySql	176
BDKRB2	unknown	PBio	PMID: 7787759, 18930543, 22047990, 22095814, 24925394	12
BLM	AR	PBio	PMID: 15137905, 15493327, 17210642, 17321898, 19109166, 19709744, 2032252, 23572515, 24606147	177
BLNK	AR	Public db	PMID: 23765059	178
BLOC1S6	AR	Public db	MySql	179
BMPR2	AD	PBio	PMID: 15877825, 19191909, 23733693, 24334027	13
C11orf65	unknown	Public db	MySql	181
C1QA	AR	Public db	PMID: 27821552	182
C1QB	AR	Public db	PMID: 27821552	183
C1QC	AR	Public db	PMID: 27821552	184
C5AR1	unknown	PBio	PMID: 1847994, 22964232, 25041699, 25174320, 25455139, 25539817,25769922,26059553,26283482,26537334	185
CAPZB	unknown	PBio	PMID: 99354614, 19806181, 22493691, 22706086, 22710966, 22918941, 23178720, 26758871	186
CARD 11	ADAR	Public db	MySql; PMID: 23765059, 25645939, 26525107	187
CARD9	AR	Public db	PMID: 27222657	188
CASP8	AR	Public db	PMID: 22365665, 26454313, 27873163, 27999438	189
CCL11	AD	Public db	MySql	190
CCL2	association	Public db	MySql	191
CCL5	association	Public db	MySql	192
CCR2	association	Public db	MySql	193
CCR5	association	Public db	MySql	194

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
CD 180	unknown	PBio	PMID: 9763566, 10880523, 21918197, 21959264, 22484241, 23103284, 23483427, 24019553, 25749095, 26371254, 26384474, 26482097, 26555723, 26371254	195
CD19	AR	Public db	MySql; PMID: 23765059, 26453379	196
CD209	association	Public db	MySql	197
CD247	AR	Public db	PMID: 26454313	198
CD27	AR	Public db	MySql; PMID: 23765059	199
CD27-AS1	unknown	Public db	MySql	200
CD300LF	unknown	PBio	PMID: 15184070, 15549731, 17202342, 18688020, 19592130,22288587,23072861,23293083,24035150	23
CD34	unknown	Public db	PMID: 27042682	201
CD3D	AR	Public db	PMID: 23765059, 26454313	202
CD3E	AR	Public db	PMID: 23765059, 26454313	203
CD3G	AR	Public db	PMID: 23765059, 26454313	204
CD40	AR	Public db	MySql; PMID: 23765059, 26453379	205
CD55	unknown	PBio	PMID: 12417446, 1385527, 16406700, 16503113, 17678954, 18424707, 19660813, 21143144, 22795896, 24588829, 24639397,25156074,25954012,26423932	207
CD59	AR	Public db	MySql	208
CD79A	AR	Public db	PMID: 23765059	209
CD79B	AR	Public db	PMID: 23765059	210
CD81	AR	Public db	MySql	211
CD8A	AR	Public db	PMID: 26454313	212
CDCA7	AR	Public db	PMID: 26216346	213
CDKN1B	AD	PBio	PMID: 10799578, 10825149, 10916090, 11123298, 11123306, 15557280, 16410832, 17273559, 20854895, 21078910, 22454463, 24317118, 25213837	24
CEBPB	unknown	Public db	PMID: 27042682	214
CENPM	unknown	PBio	PMID: 15183305,16391015,19711193,25006165	25
CHD7	AD	Public db	PMID: 18505430, 18976358, 26454313, 27484032	215
CHEK1	unknown	Public db	PMID: 19903823, 27042682	216
CIITA	AR	Public db	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	Public db	MySql	219
COMMD6	unknown	PBio	PMID: 14685242, 15799966, 16573520, 20126548, 25355947,27441653	27

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
CORO1A	AR	Public db	PMID: 23887241, 26454313	220
CR2	AR	Public db	MySql	221
CRADD	AR	PBio	PMID: 11573962, 21242994, 22323537, 24958727, 26190521,27135977	28
CRTC3	unknown	PBio	PMID: 15466468, 2032252, 21536665, 23033494, 23241891, 25114223,25316186,25351958,26937622	222
CSF3R	AR	Public db	PMID: 24753537, 26324699, 27789332	223
CTLA4	AD	Public db	PMID: 18219311, 25213377, 25329329	224
CTPS1	AR	Public db	MySql	225
CTSC	AR	Public db	PMID: 27222657	226
CX3CR1	association	Public db	MySql	227
CXCL12	association	Public db	MySql	228
CXCL9	unknown	Public db	PMID: 27042682	229
CXCR1	association	Public db	MySql	230
CXCR4	AD	Public db	PMID: 19782549, 19950235, 23765059, 24753205, 25645939,28009639	231
CXorf40A	unknown	PBio	PMID: 15541360, 24916366, 26881174	232
CYBB	XLR	Public db	PMID: 27222657	233
CYP2S1	unknown	PBio	PMID: 15681441,23933117	234
DCLRE1C	AR	Public db	PMID: 26454313, 26476407, 27484032	235
DDX1	unknown	Public db	PMID: 27042682	236
DDX58	AD	Public db	PMID: 23592984, 25692705, 25794939, 26052098, 26748340, 26848516, 26987611, 27260006, 27821552	237
DHX58	association	Public db	PMID: 25794939, 26748340, 26848516	238
DKC1	XLR	Public db	MySql; PMID: 23765059	239
DNER	unknown	PBio	PMID: 15965470, 16298139, 16997755, 17765022, 18474614, 20058045, 20367751, 22447725, 23041955, 23328254,24248099,24935874,26869529	31
DOCK2	AR	Public db	MySql	241
DSC1	unknown	PBio	PMID: 16048752, 20222919, 22692770, 24460202, 24680560, 25078507, 25244249, 26043694, 26758100	243
DUSP16	unknown	PBio	PMID: 15284860, 21613215, 24311790, 25716993, 26381291,27162525	32
ECRP	unknown	PBio	PMID: 9826755, 12855582, 25271100, 26184157	33
EDIL3	unknown	PBio	PMID: 22601780, 23518061, 24060278, 24504014, 25385367,26038125	34
EEA1	unknown	PBio	PMID: 16670179, 22591512, 24491918, 24561067, 26478006,26909655,27077111	35
EGR1	unknown	PBio	PMID: 15308739, 19050264, 19812322, 20414733, 21368226, 21622185, 22554935, 25613134, 26052046, 26980486, 11910893, 14647476, 18203138, 24627779, 25368162,27192563	244
EHF	unknown	PBio	PMID: 16380452, 17027647, 19801549, 20879862, 21617703,24219556,25217163	36
ELANE	AD	Public db	PMID: 20008220, 24145314, 27222657	245

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
EMB	unknown	PBio	PMID: 8432389, 15917240, 18209069, 19164284, 25773908	37
EPG5	AR	Public db	MySql; PMID: 21965116, 23222957, 23838600, 26917586, 26927810,27588602	246
ETF1	unknown	PBio	PMID: 20418372, 25606970, 26384426, 26833392, 27387891	247
ETV6	AD	PBio	PMID: 19264918, 20350664, 21714648, 22438058, 25581430, 25807284, 26718572, 27365488	38
F9	XLR	Public db	MySql	248
FAS	AD	Public db	PMID: 27222657	249
FASLG	AD	Public db	PMID: 27222657	250
FCGR2A	ADAR	Public db	MySql	251
FCGR3A	AR	Public db	MySql	252
FCN3	AR	Public db	MySql	253
FEZ1	unknown	Public db	PMID: 27042682	254
FHL2	unknown	PBio	PMID: 16389449, 20592280, 22417706, 22633286, 23212909	39
FOS	unknown	Public db	PMID: 27042682	255
FOXH1	unknown	Public db	PMID: 27042682	256
FOXN1	AR	Public db	MySql	257
FOXP3	XLR	Public db	PMID: 26454313	258
FPR1	unknown	PBio	PMID: 8994115, 10229829, 10611407, 17084101, 22934745, 23230437, 25605714, 25826286, 26101324, 26701131, 27034344, 27100350, 27131862, 27154726	259
FPR2	unknown	PBio	PMID: 8994115, 10229829, 10611407, 17084101, 22934745, 23230437, 25605714, 25826286, 26101324, 26701131, 27034344, 27100350, 27131862, 27154726	41
FPR3	unknown	PBio	PMID: 8994115, 10229829, 10611407, 17084101, 22934745, 23230437, 25605714, 25826286, 26101324, 26701131, 27034344, 27100350, 27131862, 27154726	42
FUK	unknown	PBio	PMID: 11753075, 12651883, 15774760, 19394435, 19647987, 20363321, 22134107, 22203233, 22276660, 22461019,24239607	43
G6PC3	AR	Public db	PMID: 20008220, 24145314, 25879134, 26479985	260
GATA2	AD	Public db	PMID: 23765059, 23887241	261
GDA	unknown	PBio	PMID: 10595517, 18600524, 20826431, 23838888, 24834013	44
GDPD4	unknown	PBio	PMID: 24373430, 24977479, 24977485, 24977489, 25528375,25596343	45
GFI1	AD	Public db	PMID: 20008220, 24145314	262
GOLGB1	unknown	PBio	PMID: 17475246, 21217069, 22034594, 23555793, 24046448	263
GPATCH2	unknown	PBio	PMID: 19432882, 25353171, 25376275	46
GPC5	unknown	PBio	PMID: 24130709, 24943672, 25354479, 26224662, 26349477	47
GPRC5A	unknown	PBio	PMID: 19593893, 20959490, 22239913, 25621293, 25714996,26165721	264
GRAP2	unknown	Public db	PMID: 25452106, 25636200, 26246585	265
GRIA3	XLR	PBio	PMID: 10441169, 12682273, 17202328, 18590483, 25904555,26648591	51
GTPBP4	unknown	PBio	PMID: 17785438, 26015807	52

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
HAX1	AR	Public db	PMID: 20008220, 24145314	266
HCN1	AD	PBio	PMID: 9405696, 9630217, 9634236, 9921901, 11133998, 23042740, 23077068, 23319474, 24403084, 24747641, 24756635, 25580535, 26578877	53
HELLS	AR	Public db	PMID: 26216346	267
HEXA	AR	PBio	PMID: 20301397, 21997228, 23727835, 24445368	54
HIVEP1	association	Public db	MySql; PMID 20226436, 26117544	268
HIVEP2	AD	Public db	MySql; PMID: 21475200, 21936769, 23389689, 24366360, 26153216, 26483320, 27003583	269
HIVEP3	unknown	Public db	MySql	270
HK2	unknown	PBio	PMID: 2749583, 4265132, 19519254, 2496891, 23874603, 25525876,25602755	55
HMGB3	unknown	PBio	PMID: 12714519, 15259015, 15358624, 16945912, 22014684,23994280,26553261	56
HNRNPLL	unknown	PBio	PMID: 18669861, 18719244, 19100700, 20505149, 22073166, 23934048, 24476532, 25825742	271
HP	unknown	PBio	PMID: 16046400, 19380867, 19795414, 26445729	272
HPCAL1	unknown	PBio	PMID: 12445467, 24699524, 25519916, 26659654, 26729710	273
HPR	unknown	PBio	PMID: 16046400, 19380867, 19795414, 26445729	57
HTR2A	association	Public db	PMID: 19204164, 24089568, 25078361, 26056932, 27003757,27042682	274
ICOS	AR	Public db	MySql; PMID: 19380800, 23765059, 27250108	275
IDI1	unknown	PBio	PMID: 14629038, 17202134, 19454010, 20955688, 22579571,23585482,25950736	276
IDI2	unknown	PBio	PMID: 14629038, 17202134, 19454010, 20955688, 22579571,23585482,25950736	59
IDI2-AS1	unknown	PBio	See IDI2	60
IDO2	unknown	PBio	PMID: 18219311, 18364004, 19487973, 19799997, 20197554, 20484729, 20693847, 21084489, 21406395, 21990421, 22754762, 24391212, 24402311, 24844751, 25477879, 25478733, 25541686, 25949913, 26308414, 27183624	61
IFIH 1	AD	Public db	PMID: 21156324, 24686847, 24995871, 25794939, 26052098, 26748340, 27260006, 27821552	277
IFNAR 1	association	Public db	PMID: 27821552	278
IFNAR2	AR	Public db	PMID: 26424569, 27821552	279
IFNG	association	Public db	MySql	280
IFNGR 1	ADAR	Public db	MySql	281
IFNGR2	ADAR	Public db	MySql; PMID: 15356149, 23161749	282
IFNLR1	unknown	PBio	PMID: 12469119, 12483210, 15166220, 22386267, 22891284, 25634147, 25904743, 25941255	62
IGLL1	AR	Public db	PMID: 25502423	283
IKBKB	ADAR	Public db	MySql; PMID: 17047224, 17072332, 25764117, 25930993, 26117626,26525107	284
IKBKG	XLDXLR	Public db	MySql; PMID: 17047224, 21455173, 21455181, 23765059, 25764117,25886387,25930993,26117626,26525107	285
IKZF 1	AD	Public db	PMID: 26454313, 26981933	286

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
IL10	AR	Public db	PMID: 23887241	287
IL10RA	AR	Public db	PMID: 23887241	288
IL10RB	AR	Public db	PMID: 23887241	289
IL12B	AR	Public db	MySql	290
IL12RB1	AR	Public db	MySql	291
IL 17F	AD	Public db	PMID: 22284928, 23887241, 24240291, 24690400, 25890879,27144517	292
IL17RA	AR	Public db	PMID: 23887241	293
IL 1B	AD	Public db	PMID: 15327898, 20543597, 24248593, 26525107, 27730320,27873163,27999438	294
IL21	AR	Public db	MySql	295
IL21R	ADAR	Public db	PMID: 23765059, 23887241	296
IL2RA	AR	Public db	MySql	297
IL2RG	XLR	Public db	PMID: 23765059, 26454313, 27484032	298
IL4R	association	Public db	MySql	299
IL7	unknown	PBio	PMID: 21508983, 22288682, 24507157, 24979548, 25130296,25214510,25411246,25734144, 26537673, 26675348,26908786	300
IL7R	AR	Public db	PMID: 26454313, 27484032	301
IRAK4	ADAR	Public db	PMID: 23766853, 25232776, 25344726, 25764117, 25886387,25930993,26785681,27845762	302
IRF3	AD	Public db	PMID: 23388631, 26513235, 26748340	303
IRF7	AR	Public db	MySql; PMID: 26748340	304
IRF8	AD AR	Public db	PMID: 23887241	305
IRGM	association	PBio	PMID: 14707092, 17911638, 22174682, 22722598, 23084913,23335927	306
ISG15	AR	Public db	MySql; PMID: 26052098, 27260006, 27821552	307
ITSN2	unknown	PBio	PMID: 11748279, 15020715, 17696400, 17696407, 22558309, 22975684, 23986746, 24097067, 24284073, 25797047,26479042	309
JAGN 1	AR	Public db	PMID: 25129144	310
JAK3	AR	Public db	PMID: 23765059, 26454313	311
JMY	unknown	PBio	PMID: 19287377, 20573979, 20574148, 20888769, 21965285, 23291261, 25015719, 26223951, 26305109	312
JPX	association	PBio	PMID: 21029862, 23791181, 23943155	64
JUN	unknown	Public db	PMID: 16928756, 27042682	313
KANK1	association	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	Public db	PMID: 27042682	314

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
LAMTOR2	AR	Public db	MySql; PMID 19782549, 20008220, 24753205	315
LARP4B	unknown	PBio	PMID: 20573744, 23815932, 25534202, 26001795, 26501340,26644407	69
LCP2	unknown	Public db	PMID: 12874226, 18219311, 19056831,23494777, 26246585	317
LIG1	AR	Public db	MySql	318
LIG4	AR	Public db	MySql	319
LOC102724297	unknown	Public db	MySql	320
LOC400710	unknown	PBio	ncRNA, limited gene information; see SNAR gene family (adjacent locus)	321
LRBA	AR	Public db	MySql; PMID: 23765059, 27873163, 27192563	322
LYST	AR	Public db	PMID: 19302049, 19782549, 20008220, 24753205, 24916509,26454313,27881733	323
MAGEA9	unknown	PBio	PMID: 15222021, 15900605, 21093980, 21791470, 25315972, 25445503, 25755744	324
MAGEA9B	unknown	PBio	PMID: 15222021, 15900605, 21093980, 21791470, 25315972, 25445503, 25755744	325
MAGT1	XLR	Public db	PMID: 23887241, 27873163, 25504528	326
MALL	unknown	PBio	PMID: 11294831, 19064697, 24101378, 24746959, 26109641, 26622604, 26641089, 26772392, 27583248, 27846891	72
MALT 1	AR	Public db	MySql; PMID: 26525107	327
MAP3K2	unknown	PBio	PMID: 11032806, 11278622, 12138187, 14734742, 16430878, 21333552, 2437584, 24847879, 25012295, 26056008	328
MAPK1	unknown	Public db	PMID: 14671106, 27042682	329
MAPK3	unknown	Public db	PMID: 14671106, 27042682	330
MAPK9	unknown	PBio	PMID: 15023353, 23685277, 24673683, 25762148, 26141991	73
MAVS	association	Public db	PMID: 23582325, 26513235, 26987611	331
MCEE	AR	PBio	PMID: 17846917, 20301409, 21365456, 23726524, 24532006,25763508,26725562	74
MECP2	XLDXLR	Public db	PMID: 27042682	332
MEX3C	unknown	PBio	PMID: 18779327, 22357625, 22658931, 22863774, 22927639, 23140835, 23446422, 23999169, 24706898, 24741071	333
MGAT5	unknown	PBio	PMID: 12417426, 15585841, 18292539, 20089585, 20117844, 25768892, 26972830	75
MKL1	AR	PBio	PMID: 12944485, 22626970, 26098208, 26098211, 26221020,26241940,26405212,26224645,26554816	89
MRE11A	AR	Public db	PMID: 23388631, 23765059	334
MS4A1	AR	Public db	MySql; PMID: 23765059	335
MSN	unknown	PBio	PMID: 9070665, 10444190, 11777944, 12445265, 14758359, 16368573, 17110458, 18025306, 18725395, 21486194, 23526587, 23613524, 24250818, 24358210, 24760896, 25746045	336
MTHFD1	AR	Both	PMID: 26454313	337
MYD88	ADAR	Public db	PMID: 23766853, 25344726, 25764117, 25886387, 25930993,26371186,27435819	338
NBN	ADAR	Public db	MySql; PMID: 23765059	339

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
NFIC	unknown	PBio	PMID: 11559801, 15327898, 16928756, 18474555, 19058033,22205750	340
NFIL3	unknown	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	Public db	PMID: 22081022, 26279205	341
NFKB2	AD	Public db	MySql; PMID: 25764117	342
NFKBIA	AD	Public db	MySql; PMID: 23765059, 25645939, 25764117	343
NHEJ1	AR	Public db	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	Public db	PMID: 16724804, 19302049, 23592984, 26848516, 27999438	345
NOD2	AD	Public db	PMID: 16724804, 19302049, 23584365, 26509073, 26848516,26953272	346
NQO2	unknown	PBio	PMID: 16253210, 16905546, 17720881, 18552348, 26046590	94
NRIP1	unknown	PBio	PMID: 18267075, 23241901, 24969109, 25066731, 25697398,25879677,26937622	95
ORAI1	ADAR	Public db	PMID: 19075015, 20004786, 21790973, 22144678, 23765059,26454313,26469693	347
OSTM1	AR	Public db	PMID: 16813530, 19507210, 21107136, 23685543	348
OVOL2	AD	PBio	PMID: 16423343, 25267199, 26619963, 26749309	98
PDE3B	unknown	PBio	PMID: 17220874, 23276671, 25816736, 26203135, 26297880,26374610	99
PDGFRA	association	PBio	PMID: 12660384, 18634583, 18701889, 19246520, 19839938, 20032375, 20569695, 21123584, 21975205, 22449623, 22523564, 23771592, 25319708, 25940087	100
PDSS2	AR	PBio	PMID: 17186472, 18437205, 18784258, 21567994, 21871565,21983691,23150520	101
PGM3	AR	Public db	MySql; PMID: 25502423	349
PHACTR4	unknown	PBio	PMID: 15107502, 17609112, 22215804, 22215812, 22766235, 23076051, 23203801, 23319639, 24748504, 26850007	102
PIAS1	unknown	PBio	PMID: 10805787, 10858346, 14644436, 15297606, 15311277, 17065208, 17540171, 18056374, 19857525, 20966256, 22969086, 22982248, 23299081, 24036127	103
PIAS2	unknown	PBio	PMID: 9724754, 11117529, 12077349, 12764129, 14514699, 15582666, 16460827, 19549844, 21156324, 21779164, 22210188, 22982248, 24344134, 25484205, 25434787, 26223632	350
PIK3CD	AD	Both	MySql; PMID: 24165795, 25133419, 25645939, 26437962, 26453379, 27379089, 27426521, 27873163, 14647476, 27192563	104
PIK3R1	ADAR	Public db	PMID: 23765059, 23887241, 25645939, 26246585, 26453379, 27076228, 14647476, 27192563	351
PKHD1	AR	PBio	PMID: 8178487, 15052665, 17450421, 23423256, 24964219, 24984783,25186187,26502924	105
PLCG2	AD	Public db	PMID: 19056831, 23000145, 23765059, 23887241, 25452106, 25636200, 25645939, 26246585, 27192563	352
PMS2	AR	Public db	MySql; PMID: 23765059	353
PNPLA4	unknown	PBio	PMID: 22289388, 23741432, 26017929, 26164793, 26713677,26741492,26968210	107
PNPT1	AR	PBio	PMID: 14563561, 15492272, 16410805, 16687933, 17983748, 19580345, 23084291, 23221631, 24143183, 24729470, 25457163	108
POLA1	XLR	Public	PMID: 27019227, 27821552	355

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
		db		
POLE	AR	Public db	MySql; PMID: 23230001, 23765059, 25948378	356
PPP2R3B	unknown	PBio	PMID: 9847399, 10629059, 11593413, 18353419, 20485545, 26683421	109
PRF1	ADAR	Public db	PMID: 17311987, 19302049, 21881043, 24916509, 25776844,26454313,26864340,27391055	357
PRKCB	unknown	PBio	PMID: 10872892, 15488737, 16935002, 17060474, 17395590, 19907441, 21997316, 22994860, 23959874, 24550541, 25548371, 24550541, 25808972, 26509731, 26510741	110
PRKCD	AR	Public db	MySql; PMID: 23319571, 27250108, 27873163	358
PRKCH	unknown	PBio	PMID: 15327898, 16571806, 18353419, 22114277, 22155788, 22892130, 23868949, 24705298, 25617472, 25889880	111
PRKDC	AD AR	Public db	PMID: 12847277, 23722905, 26454313, 26838362, 27980111	359
PROC	ADAR	PBio	PMID: 2437584, 18751723, 21114396, 22447930, 24162617	360
PSMB8	AR	Public db	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	Public db	PMID: 26246517, 27426521	362
PTPN2	unknown	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	Public db	PMID: 26454313	363
PTPRN2	unknown	PBio	PMID: 9714834, 10426369, 11086001, 11086294, 11793386, 15114673, 19361477, 23595248, 24988487, 26141787, 26609326	114
PURA	AD	Public db	PMID: 27042682	364
RAB27A	AR	Public db	PMID: 19302049, 20008220, 21881043, 23810987	365
RAB37	unknown	PBio	PMID: 21805469, 22899725, 26931073, 27798165	115
RAB7A	AD		PMID: 25992615, 27588602	366
RABGEF1	unknown	PBio	PMID: 12505986, 15143060, 15235600, 16499958, 16533754, 16605131, 17341663, 20829437, 22846990, 23552075, 24569883, 24957337, 25427001, 26567216, 26588713,27791468	367
RAC2	AD	Public db	MySql	368
RAD51	AD	Public db	PMID: 25310191, 27042682	369
RAG2	AR	Public db	PMID: 23765059, 23887241, 26454313, 27808398	371
RBCK1	AR	Public db	MySql; PMID: 21455173, 21455181, 23765059, 23969028, 24958845, 25764117, 25930993, 26008899, 26525107, 27810922	372
RBFOX 1	unknown	PBio	PMID: 23350840, 24039908, 25043849, 26500751, 26687839	116
RCC1	unknown	PBio	PMID: 1961752, 18442486, 19060893, 20347844, 23536659, 25452301,26864624	117
RFX5	AR	Public db	PMID: 23765059, 26454313	373
RFXANK	AR	Public db	PMID: 23765059, 26454313	374
RFXAP	AR	Public db	PMID: 23765059, 26454313	375
RGCC	unknown	PBio	PMID: 19158077, 19652095, 23000427, 24973210, 25770350,26134570	118

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
RHOQ	unknown	PBio	PMID: 10490598, 12456725, 14734537, 16246732, 17016434, 19258391, 22916134, 24223996, 24297911, 24663214,24667291	119
RIPK1	association	Public db	PMID: 21455173, 27999438	376
RIPK3	association	Public db	PMID: 22365665, 27999438	377
RMRP	AR	Public db	MySql; PMID 19782549, 20008220, 24753205	378
RNASE3	unknown	PBio	PMID: 19515815, 26184157	120
RNASEH2A	AR	Public db	PMID: 26052098, 27260006, 27821552	379
RNASEH2B	AR	Public db	PMID: 26052098, 27260006, 27821552	380
RNASEH2C	AR	Public db	PMID: 26052098, 27260006, 27821552	381
RNASEL	association	Public db	PMID: 24995003, 27525044	382
RNF168	AR	Public db	MySql; PMID: 23765059	383
RNF31	AR	Public db	PMID: 21455173, 21455181, 23969028, 24958845, 26008899, 26525107, 27810922	384
RNU4ATAC	AR	Public db	PMID: 27222657	385
RPTOR	unknown	PBio	PMID: 16959881, 22810227, 23349361, 23812589, 24287405, 24303063, 24671993, 24948799, 26678875	123
RTEL1	AR	Public db	MySql; PMID: 23329068, 23765059, 24009516, 25607374, 26810774	386
RTEL1-TNFRSF6B	unknown	Public db	MySql; PMID: 23329068, 23765059, 25607374	387
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	Public db	PMID: 26052098, 27260006, 27821552	389
SBDS	AR	Public db	PMID: 20008220, 21062271, 27418648, 27658964	390
SERPINB4	unknown	PBio	PMID: 15203215, 19070595, 21857942, 22451727, 22808225, 24560885, 24635038, 25111616, 25133778, 25213322	124
SERPINB6	AR	PBio	PMID: 14670919, 20451170, 24172014, 24359430	125
SH2D1A	XLR	Public db	MySql; PMID: 19302049, 23765059, 25744037	391
SHARPIN	unknown	Public db	PMID: 21455181, 22901541, 23969028, 24958845, 26525107, 26848516, 27810922, 27892465	392
SKIV2L	AR	Public db	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	Public db	PMID: 20008220, 20301489	394
SLC3A2	unknown	PBio	PMID: 22588539, 22624878, 23297381, 24491544, 25002078, 26172215, 26439699, 26444422	126
SLC46A1	AR	Public db	PMID: 26454313	395
SLC8A1	unknown	PBio	PMID: 23224883, 23224887, 23224890, 23224891, 26045217, 26418956, 26775040, 26859825, 26924806	396
SMAD2	unknown	Public db	PMID: 27042682	397

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
SMAD3	AD	Public db	PMID: 27042682	398
SMAD4	AD	Both	PMID: 12202226, 14987161, 16800882, 19420158, 25637015,25705527,26454313,27042682	399
SNAP29	AR	Public db	PMID: 15968592, 21073448, 27588602	400
SNAR-A1	unknown	PBio	PMID: 25327818, 25447144	401
SNAR-A10	unknown	PBio	PMID: 25327818,25447144	402
SNAR-A 11	unknown	PBio	PMID: 25327818, 25447144	403
SNAR-A12	unknown	PBio	PMID: 25327818, 25447144	404
SNAR-A13	unknown	PBio	PMID: 25327818, 25447144	405
SNAR-A14	unknown	PBio	PMID: 25327818, 25447144	406
SNAR-A2	unknown	PBio	PMID: 25327818, 25447144	407
SNAR-A3	unknown	PBio	PMID: 25327818, 25447144	408
SNAR-A4	unknown	PBio	PMID: 25327818, 25447144	409
SNAR-A5	unknown	PBio	PMID: 25327818, 25447144	410
SNAR-A6	unknown	PBio	PMID: 25327818, 25447144	411
SNAR-A7	unknown	PBio	PMID: 25327818, 25447144	412
SNAR-A8	unknown	PBio	PMID: 25327818, 25447144	413
SNAR-A9	unknown	PBio	PMID: 25327818, 25447144	414
SNAR-B1	unknown	PBio	PMID: 25327818, 25447144	415
SNAR-B2	unknown	PBio	PMID: 25327818, 25447144	416
SNAR-C1	unknown	PBio	PMID: 25327818, 25447144	417
SNAR-C2	unknown	PBio	PMID: 25327818, 25447144	418
SNAR-C3	unknown	PBio	PMID: 25327818, 25447144	419
SNAR-C4	unknown	PBio	PMID: 25327818, 25447144	420
SNAR-C5	unknown	PBio	PMID: 25327818, 25447144	421
SNAR-D	unknown	PBio	PMID: 25327818, 25447144	422
SNAR-E	unknown	PBio	PMID: 25327818, 25447144	423
SNAR-F	unknown	PBio	PMID: 25327818, 25447144	424
SNAR-G1	unknown	PBio	PMID: 25327818, 25447144	425
SNAR-G2	unknown	PBio	PMID: 25327818, 25447144	426
SNAR-H	unknown	PBio	PMID: 25327818, 25447144	427
SNAR-I	unknown	PBio	PMID: 25327818, 25447144	428
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	unknown	PBio	PMID: 22308462, 22843687, 26373735	128
SNX 10	AR	Public db	PMID: 22499339, 23123320	430
SNX5	unknown	PBio	PMID: 10600472, 11128621, 14499622, 15133132, 15561769, 16857196, 18596235, 21725319, 21903422, 21943487, 23213485, 24820351, 26220253	130
SOCS2	unknown	PBio	PMID: 19279332, 21403007, 22693634, 22795647, 23455506, 24400794, 26216515, 26709655, 26765997, 27071013,27158906, 27330188, 27338192	131
SP110	AR	Public db	MySql	431
SP140	unknown	Public db	MySql	432
SPINK5	AR	Public db	PMID: 19683336, 26865388, 27222657, 27905021	433

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
SQSTM1	ADAR	Public db	PMID: 19229298, 27715390	434
SRSF1	unknown	Public db	PMID: 27042682	435
ST8SIA5	unknown	PBio	PMID: 11089916, 15829700	133
STAT2	AR	Public db	PMID: 23391734, 26122121, 27821552	437
STAT5B	AR	Public db	MySql	439
STIM1	ADAR	Public db	PMID: 20004786, 21790973, 23765059, 26454313, 26469693	440
STIM2	unknown	PBio	PMID: 20004786, 21790973, 21880262, 22129055, 22477146,22914293,25157823,26109647,26469693	134
STK4	AR	Public db	PMID: 19782549, 23765059, 23887241, 24753205, 26029204	441
STX11	AR	Public db	PMID: 19302049, 21881043, 24916509, 26454313	442
STXBP2	ADAR	Public db	PMID: 21881043, 24916509, 25564401, 26454313	443
SYNCRIP	unknown	PBio	PMID: 10734137, 18045242, 19331829, 19232660, 22493061, 22935615, 23679954, 23700384, 24844655, 25100733,26641092	444
T	ADAR	PBio	PMID: 11897834, 17438107, 23064415, 23662285, 24253444, 24556085, 25186612, 26210634, 26919728	445
TAP1	AR	Public db	PMID: 26454313	446
TAP2	AR	Public db	PMID: 26454313	447
TAPBP	unknown	Public db	PMID: 26454313	448
TAZ	XLR	Public db	PMID: 20008220	449
TBC1D16	unknown	PBio	PMID: 16923123, 19077034, 21250943, 23019362, 23485563,23812537,24513270,26030178	136
TBK1	AD	Public db	PMID: 23887241, 25930993, 26513235, 28049150	450
TBX1	AD	Public db	PMID: 26454313	451
TCIRG1	ADAR	Public db	MySql; PMID 19507210, 19782549, 24753205, 27233968	452
TICAM1	AD AR	Public db	PMID: 22105173, 23887241, 25764117, 25930993, 26513235,28049150	453
TLR3	AD	Public db	PMID: 23592984, 23887241, 25930993, 26513235, 27810922,27873163,27881733	454
TLR4	association	Both	PMID: 12124407, 17893200, 18946062, 19843948, 20521908, 21677132, 22474023, 22962435, 23055527, 23890253, 25365308, 25454804, 25930993, 26189680, 26453379,27881733	455
TMEM173	AD	Public db	PMID: 23388631, 25645939, 25692705, 26052098, 27260006, 27801882, 27821552	456
TNF	association	Public db	MySql; PMID: 27042682	457
TNFAIP3	AD	Public db	PMID: 23969028, 26642243, 27845235	458
TNFRSF10A	unknown	PBio	PMID: 10889508, 11602752, 11704827, 11777550, 11844843, 12390973, 12694389, 14975593, 15007095, 16394652, 16554480, 17671142, 19690337, 20921531	138
TNFRSF 11A	ADAR	Public db	PMID: 17088646, 17360404, 18281276, 18606301, 19380800, 19507210, 25102334, 25393853, 27003757, 27016605	459
TNFRSF 11B	AR	Public db	PMID: 19507210, 25102334, 25393853, 27003757	460

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
TNFRSF13B	ADAR	Public db	MySql; PMID: 17467261, 17492055, 18978466, 18981294, 19629655, 20889194, 21458042, 22697072, 23765059, 25454804,25930993,26727773,27123465	461
TNFRSF13C	AR	Both	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	unknown	PBio	PMID: 16439533, 19162554, 19363449, 22017440, 23432692,24484736,25738498	140
TNFRSF4	AR	Public db	MySql	462
TNFRSF8	unknown	PBio	PMID: 10921351, 15990453, 16472805, 18852356, 20141444, 20378007, 21933041, 23115213, 23307550, 23654079,24809535,25999451	463
TNFSF11	AR	Public db	PMID: 17088646, 17360404, 18281276, 18606301, 19507210,25992615,27003757	464
TNFSF12	association	Public db	PMID: 23765059	465
TP53	ADAR	Public db	MySql; PMID: 11048806, 11079782, 12009037, 19282432, 26870672	466
TRAF3	AD	Public db	PMID: 20832341, 23887241, 25764117, 25930993, 28049150	467
TRAF6	unknown	Public db	PMID: 10215628, 10421844, 25200954, 27808398, 27999438	468
TRAFD1	unknown	PBio	PMID: 16221674, 18849341, 23913580, 25909814, 25992615,26283173	141
TREX1	ADAR	Public db	PMID: 26052098, 27260006, 27821552	469
TRNT1	AR	Public db	MySql; PMID: 25193871	470
TRPM2	unknown	PBio	PMID: 9806837, 16585058, 18569867, 19411837, 20107186, 25012489, 25049394, 25088676, 26300888, 26558786, 26679996, 26942016, 26969190, 27405665, 27872485	142
TTC7A	AR	Public db	MySql; PMID: 27873163	471
UBE2N	unknown	PBio	PMID: 21512573, 23159053, 24906799, 25343992, 25503582, 25548215, 26085214, 26150489, 26212332, 26518362	145
UNC119	AD	Public db	MySql	472
UNC13D	AR	Public db	PMID: 19302049, 21881043, 24916509, 25564401, 25980904,26454313	473
UNC93B 1	association	Public db	PMID: 23810987, 23887241, 25930993, 27873163	474
UNG	AR	Public db	MySql; PMID: 23765059	475
USP18	AR	Public db	PMID: 27016605, 27325888, 27801882, 27821552	476
USP20	unknown	Public db	PMID: 27801882	477
VAPA	unknown	PBio	PMID: 9657962, 10523508, 10655491, 11511104, 12931207, 18713837, 23536298, 24076421, 24569996, 25015719	478
VCP	AD	Public db	PMID: 24248593, 27730320	479
VDAC1	unknown	PBio	PMID: 10620603, 25874870, 26322231, 26542804, 26616244, 26758954, 26878172	480
VPS13B	AR	Public db	PMID: 20008220, 20301655	481
VPS45	AR	Public db	MySql; PMID: 23738510, 24145314, 24164830, 26358756	482
VSTM1	unknown	PBio	PMID: 22960280, 23436183, 24205237, 25351446, 25887911,26760041	147

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
VWA2	unknown	PBio	PMID: 14506275, 18434322, 21385852, 23443151, 23960233,26121272	148
WEE1	unknown	Public db	PMID: 19903823, 25088202, 26598692, 26881506, 27042682	484
WIPF1	AR	Public db	PMID: 23765059, 26029204, 26453379	485
XIAP	XLD XLR	Public db	MySql; PMID: 22365665, 25744037, 26953272	486
YBX1	unknown	Public db	PMID: 27042682	487
YWHAZ	unknown	Both	PMID: 25894827, 27042682	488
ZAP70	AD AR	Public db	PMID: 18219311, 23494777, 23765059, 24164480, 26454313	489
ZBTB24	AR	Public db	MySql; PMID: 23486536, 23765059, 26851945, 27098601	490

[0321] Table 6 is a comprehensive list of 419 exemplary genes (referred to herein as 'PML-419 genes' or 'PML-419 gene list') interrogated in the present study, along with information related to the inheritance pattern assumed for analysis and the reason for inclusion of the gene. Gene sources for **Table 6** (column heading 'Gene Source'): 1) nominated on the basis of being linked to immune deficiency, as curated from public databases (indicated by 'Public db') such as PubMed and ClinVar, 2) PBio CNV-identified genes ('PBio', see **Table 6** column heading `Gene Source') from a genome-wide array CGH gene discovery study of 71 PML cases, or 3) curated from public databases and identified in PBio's PML gene discovery study (indicated by 'Both'). A genetic predisposition to PML on the basis of the host's genome was proposed; that is, germline genetic variant(s) in the PML patient's genome, rather than genetic variants that are present in the JC virus, are the cause of the patient's PML (Hatchwell, Front Immunol., 6:216 (2015). Details on the source of the genes in the PML-419 gene list can be found in the following immunodeficiency and immune-related gene sources: 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); and Tsujita et al., J Allergy Clin Immunol. (2016). MySql' genes are derived from the ClinVar database. ClinVar was searched using the terms "immune deficiency" and "immunodeficiency." Entries that described large genomic rearrangements, containing multiple genes, were excluded. A non-redundant list of 125 genes was compiled by combining the output of the two searches and deposited into a MySQL database. NOTE: A subset of these genes are not flagged as 'MySql' if they appeared in one or more of the immune gene review papers noted above. van der Kolk et al., Ann Clin Transl Neurol.;3(3):226-32 (2016) was the source of known BAG3 PML gene (see below) and 28 candidate PML genes on the basis of connection to JCV. Van der Kolk *et al.,* cite a method as follows: "the latter was performed by searching for JCV in NCBI, and selecting for genes in humans." This yielded 30 human genes, 5 of which overlapped with the PML gene list and 2 genes (HLA-DQB1, HLA-DRB1) were excluded because HLA loci are difficult to interpret. The genes ADA, BAG3, BTK, CD40LG, DOCK8, STAT1, WAS, and WIPF1 were derived from Hatchwell, Front Immunol., 6:216 (2015) (see **Table 1** for primary references); van der Kolk et al., Ann Clin Transl Neurol., 3(3):226-32 (2016); and Zerbe et al., Clin Infect Dis., 62(8):986-94 (2016). PBio genes are based on CNV studies and a subset overlap the immune review gene lists (annotated as 'Both' in column heading `Gene Source'). Tier 1 genes were used as potential solutions for PML cases. Determination of Autosomal Dominant (AD), Autosomal Recessive (AR), X-linked dominant (XLD), or X-linked (XLR) disease model for each gene was derived from the immunodeficiency review papers and/or OMIM annotations. Entries marked 'association' denotes variants were found to be associated with an immune-related condition; 'unknown' denotes no evidence reported in the literature for an AD or AR model.

Table 7: Potential cause of PML in each patient in the study

SampleID	Ethnicity	Gender	Primary Disease	RefSeq Gene Symbol	Variant Type	Varian Frequency Details (Ethnic specific)	Frequency (Reciprocol) (Ethnic specific)	SEQ ID
MVGS1116-8a	EUR	F	MS (NZ Rx)	DOCK8	SNV hom SNV het	0.499 0.00447	1 in 1,792	1147 1148
MVGS1359	EUR	F	MS (NZ Rx)	IL17F	SNV het	0.00024	1 in 4,170	1114
MVGS1368	EUR	F	MS (NZ Rx)	IDO2	SNV hom SNV het	0.508 0.065	1 in 121	1125 1126
MVGS540-374b	EUR	M	MS (NZ Rx)	SHARPIN	SNV hom	0.00217	1 in 461	1142
MVGS540-393b	EUR	F	MS (NZ Rx)	DOCK8	SNV hom SNV het SNV het	0.499 0.00153 0.194	1 in 5,246	1147 1154 1152
MVGS694-6a	EUR	F	Other	CHD7	SNV het	0.00028	1 in 3,528	1135

MVGS811-13a	EUR	M	HIV	PIK3CD PIK3CD-AS1	CNV hom	novel	0	2
MVGS995-4a	EUR	M	MS (NZ Rx)	EPG5	SNV hom SNV het	0.495 0.000251	1 in 32,224	1279 1273
PML-01	EUR	F	HIV	ITSN2	SNV hom	0.00183	1 in 547	1028

PML01	EUR	F	HIV	ITSN2	SNV hom	0.00183	1 in 547	1028
PML02	EUR	M	Other	IKBKB	SNV het	novel	0	1127
PML03	EUR	F	MS (NZ Rx)	FPR2	CNV hom	2.23E-06	1 in 448,833	140
PML04	EUR	M	HIV	*unsolved*	n/a	n/a	n/a	n/a
PML05	LAT	M	HIV	TBK1	SNV het	novel	0	1203
PML06	AFR	M	HIV	TICAM1	SNV het	0.000777	1 in 1,287	1289
PML09	EUR	M	HIV	LIG4	SNV SNV comp het	0.00399 0.287	1 in 3497	1221 1222
PML10	EUR	F	HIV	TNFRSF11A	SNV het	novel	0	1287
PML12	LAT	F	HIV	BLM	SNV hom	0.000874	1 in 1,144	1235
PML13	AFR	M	HIV	PLCG2	SNV SNV comp het	0.00167 0.0187	1 in 128,105	1261 1263
PML14	EUR	M	HIV	PLCG2	SNV SNV comp het	0.00998 0.0159	1 in 25,259	1261 1263
PML15	LAT	M	HIV	NOD2	SNV het	novel	0	1255
PML16	AFR	F	HIV	TNFRSF11A	SNV het	novel	0	1287
PML17	EUR	M	HIV	ZAP70	SNV het	0.00009	1 in 11,110	1035
PML18	EUR	M	HIV	*unsolved*	n/a	n/a	n/a	n/a
PML19	AFR	M	HIV	ATM	SNV SNV comp het	0.0479 novel	0	1193 1194
PML20	AFR	M	HIV	NFKB1	SNV het	0.00173	1 in 577	1069
PML21	EUR	M	HIV	ZAP70	SNV het	0.0000602	1 in 16,623	1034
PML22	EUR	M	HIV	*unsolved*	n/a	n/a	n/a	n/a
PML23	EUR	F	HIV	DCLRE1C	SNV hom	novel	0	1167
PML25	EUR	F	HIV	PLCG2	SNV het	0.000150	1 in 6,672	1259
PML26	EUR	M	HIV	TRAFD1	SNV hom	0.000689	1 in 1,451	1208
PML27	EUR	M	HIV	TAP2	SNV hom	0.00837	1 in 120	1101
PML28	EUR	F	MS (NZ Rx)	TRPM2	SNV hom	novel	0	1311
PML29	AFR	M	HIV	KCTD7 RABGEF1	CNV hom	0.000387	1 in 2,584	65
PML30	EUR	M	HIV	TNFRSF11A	SNV het	novel	0	1287
PML31	AFR	F	HIV	DDX58	SNV het	0.000779	1 in 1,283	1157
PML32	EUR	M	HIV	*unsolved*	n/a	n/a	n/a	n/a
PML33	EUR	M	HIV	TNFRSF11A	SNV het	novel	0	1287
PML35	EUR	F	HIV	TNFRSF11A	SNV het	novel	0	1287
PML36	AFR	F	HIV	TCIRG1	SNV het	0.002134	1 in 469	1184
PML37	AFR	M	HIV	GATA2	SNV het	novel	0	1056
PML38	EUR	M	HIV	MALL	CNV hom	3.95E-06	1 in 253,036	26
PML39	AFR	M	HIV	*unsolved*	n/a	n/a	n/a	n/a

PML40	LAT	F	HIV	PNPT1	SNV hom	novel	0	1032
PML41	AFR	M	HIV	ZAP70	SNV het	novel	0	1036
PML43	EUR	M	HIV	PTPRC	SNV hom	novel	0	1020
PML44	EUR	M	HIV	TNFRSF11A	SNV het	novel	0	1287
PML45	EUR	F	Other	CARD11	SNV het	0.0024	1 in 417	1123
PML46	LAT	M	HIV	EPG5	SNV SNV comp het	0.0123 0.436	1 in 745	1278 1279
PML48	EUR	M	HIV	SMAD4	SNV het	0.000901	1 in 11,100	1284
PML49	EUR	M	HIV	STIM1	SNV het	novel	0	1174
PML50	AFR	M	HIV	NOD2	SNV het	novel	0	1256
PML51	EUR	M	HIV	TICAM1	SNV het	0.00265	1 in 377	1289
PML52	EUR	F	Other	*unsolved*	n/a	n/a	n/a	n/a
PML53	EUR	M	Other	GFI1	SNV het	0.00003	1 in 32,635	1011
PML54	EUR	F	HIV	TNFRSF11A	SNV het	novel	0	1287
PML55	EUR	F	HIV	RTEL1	SNV het	0.00326	1 in 307	1299
PML56	EUR	M	HIV	TNFRSF11A	SNV het	novel	0	1287
PML57	EUR	F	Other	TRAF3	SNV het	0.00093	1 in 1,075	1229
PML58	AFR	M	HIV	DOCK8	SNV SNV comp het	0.0575 0.478	1 in 146	1146 1147
PML59	AFR	M	HIV	IFIH1	SNV het	0.00281	1 in 356	1040
PML60	EUR	M	HIV	*unsolved*	n/a	n/a	n/a	n/a
PML61	AFR	F	HIV	TNFRSF11A	SNV het	novel	0	1287
PML62	AFR	F	HIV	*unsolved*	n/a	n/a	n/a	n/a
PML63	AFR	M	HIV	PLCG2	SNV het	0.00195	1 in 514	1260
PML64	AFR	M	HIV	PIK3R1	SNV het	novel	0	1077
PML65	AFR	M	HIV	ITSN2	CNV hom	0.00313	1 in 319	14
PML66	AFR	M	HIV	*unsolved*	n/a	n/a	n/a	n/a
PML67	EUR	F	MS (NZ Rx)	*unsolved* (CNV data, no WES data)	n/a	n/a	n/a	n/a
PML68	EUR	F	MS (NZ Rx)	LRBA	SNV hom	0.00162	1 in 618	1073

PML69	EUR	M	Other	EGR1 ETF1	CNV hom	0.001	1 in 1,005	45
PML72	AFR	F	HIV	NOD2	SNV het	0.004036	1 in 248	1252

[0322] Table 7 contains a single genetic solution/explanation that is the potential cause of PML in each patient in the study (71 cases were assessed with genome-wide array CGH and 71 were also assessed by whole exome sequencing), with the exception of 19 'unsolved' cases. Solutions are based on a combination of CNV and SNV variants, connected by SEQ IDs to tables 1, 4 and 5. For homozygous or compound heterozygous variant solutions, expected population frequencies were calculated as follows:

Expected population frequency for variant a (freq p) and variant b (freq q) = pq/4.

[0323] For example, PML09 has 2 variants, SEQID 1221 and 1222, with individual frequencies in the normal population of 0.00399, 0.287. The expected frequency in an ethnically-matched normal population for this combination is (0.00399*0.287)*0.25 = 0.000286283 = 1/3,497.

[0324] The Primary Disease identifiers in **Table 7** are: HIV, infection with human immunodeficiency virus; MS (NZ Rx), multiple sclerosis treated with natalizumab; Other, which includes a variety of disorders/conditions (MVGS694-6a had aplastic anemia, PML02 and PML52 had lymphoma, PML45 and PML 57 had chronic lymphocytic leukemia, PML53 had sarcoidosis, and PML69 is a kidney transplant patient who was on belatacept).

[0325] Solutions were considered on the basis of presence of rare variants (CNVs and/or SNVs) in or near genes that are listed in **Table 6.** Both autosomal recessive (AR) and autosomal dominant (AD) disease models comprise this set of solutions, based on finding homozygous SNVs, homozygous CNVs, compound heterozygous SNVs, or heterozygous SNVs. Nine PML cases in **Table 7** were considered 'unsolved' on the basis of analyzing both CNV and SNV data, and one case (PML67) was assessed for CNVs only since WES data were unavailable. In some instances, a case was considered unsolved for a best solution (**Table 7**) but alternate solutions were reported in **Table 8** (see below).

[0326] For PML cases that had more than one potential solution. In these instances, the 'best' solution (**Table 7**) was determined on the basis of rarity of the genetic variant(s) and the relative strength of the biology for the PML-419 genes (**Table 6**). Alternate solutions are reported in **Table 8.** For example, for PML case MVGS1116-8a, three solutions were found, which impacted genes DOCK8, HIVEP2, and RNF168. In this example, DOCK8 compound heterozygous SNVs (**Table 7**, SNV hom and SNV het) were selected as the best solution because DOCK8 is a known PML gene. In another example, PML case MVGS1359 has IL17F (het SNV) listed as the best solution in **Table 7** because it is rarer than alternate solutions for the ATR and STXBP2 genes.

[0327] While some PML patients may have multiple genes/variants causing and/or contributing to their PML, in many PML patients only a single gene will be the primary cause analogous to patients diagnosed with primary immunodeficiency disorders. In addition to the alternate solutions reported in **Table 8,** which are based on SNV genetic findings only, additional alternate solutions based on CNV genetic findings are reported in **Table 1.**

Table 8: Alternate genetic solutions/explanations as the potential cause of PML in the study

Sample ID	RefSeq Gene Symbol	Variant Type	Variant Frequency Details (Ethnic specific)	Frequency (Reciprocol) (Ethnic specific)	SEQ ID
MVGS1116-8a	HIVEP2	SNV het	novel	0	1118
MVGS1116-8a	RNF168	SNV hom	0.469	1 in 1,041	1063
		SNV het	0.00818		1066
MVGS1359	ATR	SNV het	0.00393	1 in 254	1058
MVGS1359	STXBP2	SNV het	0.00501	1 in 199	1291
MVGS540-374b	MKL1	CNV hom	3.99E-08	1 in 25,081,515	157
MVGS540-393b	PRKDC	SNV het	0.00097	1 in 1,031	1130
MVGS811-13a	CLCN7	SNV het	0.00028	1 in 3,571	1239
MVGS995-4a	KAT6B	SNV het	0.00003	1 in 33,357	1169
MVGS995-4a	PRF1	SNV het	0.00243	1 in 412	1168
PML03	CDKN1B	SNV het	0.00003	1 in 32,209	1200
PML05	ATR	SNV het	novel	0	1061
PML05	NFKB1	SNV het	0.00501	1 in 200	1070
PML06	CHD7	SNV het	0.00797	1 in 125	1136
PML06	DOCK8	SNV hom	0.478	1 in 267	1147
		SNV het	0.0313		1152
PML09	RIPK3	SNV het	0.00398	1 in 251	1227
PML10	JUN	SNV het	0.00103	1 in 968	1009

Sample ID	RefSeq Gene Symbol	Variant Type	Variant Frequency Details (Ethnic specific)	Frequency (Reciprocol) (Ethnic specific)	SEQ ID
PML10	RAG1	SNV het	0.00039	1 in 2,566	1179
PML12	CARD11	SNV het	novel	0	1122
PML12	PRKDC	SNV het	novel	0	1128
PML13	DOCK8	SNV hom	0.478	1 in 267	1147
		SNV het	0.0313		1152
PML13	IRAK4	SNV het	novel	0	1202
PML13	PIK3CD	SNV het	0.00679	1 in 147	1000
PML14	NBN	SNV het	0.0039	1 in 256	1138
PML14	NFKB1	SNV het	novel	0	1071
PML15	ASH1L	SNV	Novel	0	1016
		SNV comp het	0.0019		1017
PML15	CHD7	SNV het	0.00176	1 in 568	1133
PML15	HIVEP2	SNV het	novel	0	1116
PML15	STIM1	SNV het	0.00587	1 in 170	1175
PML16	TBK1	SNV het	novel	0	1204
PML16	TLR3	SNV het	0.00136	1 in 738	1076
PML17	APOL1	SNV het	0.0021	1 in 475	1327
PML18	PKHD1	SNV hom	0.498	1 in 171	1104
		SNV het	0.0471		1107
PML19	DOCK8	SNV	0.0575	1 in 146	1146
		SNV comp het	0.478		1147
PML19	IFIH1	SNV het	0.00444	1 in 225	1041
PML20	JUN	SNV het	0.00535	1 in 187	1010
PML21	PRKCH	SNV het	novel	0	1228
PML21	PSTPIP1	SNV het	0.00093	1 in 1,074	1232
PML21	RAG2	SNV het	novel	0	1182
PML22	RIPK3	SNV hom	0.00309	1 in 324	1226
PML22	VPS45	SNV het	0.00114	1 in 878	1014
PML23	NOD2	SNV het	novel	0	1251
PML23	RAG1	SNV het	0.00003	1 in 33,317	1180
PML28	PKHD1	SNV hom	0.498	1 in 171	1104
		SNV het	0.0471		1107
PML28	TNFRSF13B	SNV het	0.00929	1 in 108	1267
PML30	RTEL1	SNV het	0.000124	1 in 8,068	1300
PML31	AP3B1	SNV het	novel	0	1084
PML31	PRKDC	SNV het	novel	0	1129
PML33	STIM2	SNV het	0.00003	1 in 32,688	1068
PML33	TLR3	SNV hom	0.413	1 in 2,227	1075
		SNV het	0.00435		1074
PML33	TLR4	SNV hom	0.00283	1 in 354	1161
		SNV hom	0.00285		1160
PML35	PRKCB	SNV het	0.00276	1 in 362	1247
PML36	NOD2	SNV het	0.00871	1 in 115	1254
PML36	PIK3CD	SNV het	0.00679	1 in 147	1000
PML37	AP3B1	SNV het	novel	0	1080
PML37	ATR	SNV het	0.00038	1 in 2,601	1059
PML37	WEE1	SNV het	0.00825	1 in 121	1177
PML38	MYD88	SNV het	novel	0	1051

Sample ID	RefSeq Gene Symbol	Variant Type	Variant Frequency Details (Ethnic specific)	Frequency (Reciprocol) (Ethnic specific)	SEQ ID
PML40	MCEE	SNV hom	0.01	1 in 100	1033
PML41	AP3B1	SNV het	0.00173	1 in 577	1082
PML41	CHD7	SNV het	novel	0	1137
PML41	DOCK8	SNV	0.0575	1 in 146	1146
		SNV comp het	0.478		1147
PML41	POLE	SNV hom	0.00019	1 in 5,203	1219
PML41	RNF168	SNV	0.412	0	1063
		SNV comp het	novel		1062
PML43	DOCK8	SNV hom	0.499	0	1147
		SNV het	novel		1150
PML44	DCLRE1C	SNV hom	0.0287	1 in 174	1166
		SNV hom	0.00575		1165
PML44	GFI1	SNV het	0.00708	1 in 141	1012
PML45	POLA1	SNV het	novel	0	1328
PML46	AP3B1	SNV het	0.00587	1 in 170	1082
PML46	IL21R	SNV het	0.00573	1 in 175	1248
PML46	PRKDC	SNV het	0.00017	1 in 5,781	1131
PML48	TNFRSF11A	SNV het	0.00233	1 in 429	1286
PML49	DCLRE1C	SNV hom	0.00575	1 in 174	1166
		SNV hom	0.0287		1165
PML49	PTEN	SNV het	novel	0	1171
PML49	RIPK1	SNV het	0.00090	1 in 1,112	1090
PML50	AP3B1	SNV het	0.00387	1 in 259	1078
PML50	PIAS2	SNV het	0.00357	1 in 280	1283
PML50	STXBP2	SNV het	0.00038	1 in 2,598	1290
PML52	GFI1	SNV het	0.00708	1 in 141	1012
PML53	IL1B	SNV het	novel	0	1037
PML53	STXBP2	SNV het	0.00501	1 in 199	1291
PML54	EPG5	SNV	0.0638	1 in 127	1278
		SNV comp het	0.495		1279
PML54	IFNGR2	SNV het	0.00009	1 in 11,096	1304
PML54	RAG1	SNV het	0.00003	1 in 33,352	1178
PML54	RAG2	SNV het	novel	0	1183
PML57	PIAS1	SNV het	novel	0	1231
PML57	PKHD1	SNV hom	0.498	1 in 171	1104
		SNV het	0.0471		1107
PML57	SKIV2L	SNV hom	0.157	1 in 538	1098
		SNV hom	0.214		1100
		SNV het	0.0471		1099
PML58	GFI1	SNV het	0.00144	1 in 693	1012
PML59	IFNLR1	SNV het	novel	0	1002
PML59	NOD2	SNV het	0.00404	1 in 248	1252
PML59	NRIP1	SNV hom	0.00711	1 in 141	1301
PML59	RAD51	SNV het	0.00865	1 in 116	1230
PML60	MAPK3	SNV het	novel	0	1250
PML60	TP53	SNV het	0.00048	1 in 2,085	1266
PML61	GATA2	SNV het	0.00024	1 in 4,139	1057
PML61	PTPRC	SNV hom	novel	0	1019

Sample ID	RefSeq Gene Symbol	Variant Type	Variant Frequency Details (Ethnic specific)	Frequency (Reciprocol) (Ethnic specific)	SEQ ID
PML61	TNFRSF8	SNV het	novel	0	1001
PML62	PRKCD	SNV het	novel	0	1054
PML63	HTR2A	SNV hom	0.00519	1 in 193	1220
PML63	MAPK3	SNV het	0.00193	1 in 518	1249
PML64	PLCG2	SNV het	0.00044	1 in 2,276	1264
PML64	WEE1	SNV het	novel	0	1176
PML65	IRAK4	SNV het	0.00118	1 in 850	1201
PML66	PIK3CD	SNV het	0.00679	1 in 147	1000
PML68	RAG1	SNV het	0.00586	1 in 171	1181
PML72	CARD11	SNV het	0.00242	1 in 413	1121
PML72	HIVEP1	SNV hom	0.00164	1 in 610	1092
PML72	IFIH 1	SNV het	0.00843	1 in 119	1043

[0328] **Table 8** contains analogous information to **Table 7,** with the exception that Ethnicity, Gender and Primary Disease are not repeated. **Table 8** contains alternate genetic solutions/explanations as the potential cause of PML for the patients in the study (71 cases were assessed with genome-wide array CGH and 70 were also assessed by whole exome sequencing). Solutions in **Table 8** are also case-level and represent secondary, alternative solutions for the cases listed (using the same criteria used to identify potential solutions reported in **Table 7**). In other words, for some individuals, more than one reasonable solution was identified and, while those in **Table 7** are considered the most likely, those in **Table 8** are also potential solutions. It can be appreciated by those skilled in the art that further data on new PML cases, patients with genetic-based immunodeficiency disorders, or functional studies on a given gene may support selection of a **Table 8** solution as the 'best' single solution (e.g., a current **Table 7** solution could be considered instead as a **Table 8** solution, and vice versa).

Table 9: Pairs of SNVs impacting the same gene

Sample ID	RefSeq Gene Symbol	Variant Type	Chromosome	Position	Ref Allele	Alt Allele	Amino Acid Change	Variant Frequency Details (Ethnic specific)	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	RNF168	SNV het	3	196199204	G	T	P401Q	0.46947	1063
MVGS1368	RNF168	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	EEA1	SNV het	12	93196332	C	T	E840K	0.01949	1206
MVGS995-4a	EEA1	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	novel	1246
PML04	POLE	SNV het	12	133220526	T	C	N1369S	0.22363	1213
PML04	POLE	SNV het	12	133237658	T	G	Q766P	novel	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	novel	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	R235W	0.50282	1125
PML12	IDO2	SNV het	8	39862893	T	A	S239T	0.02384	1126
PML13	STX11	SNV het	6	144508353	G	A	V197M	novel	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
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	novel	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	novel	1212
PML16	POLE	SNV het	12	133220526	T	C	N1369S	0.24889	1213
PML17	RNF168	SNV het	3	196199204	G	T	P401Q	0.46947	1063
PML17	RNF168	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
PML17	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
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	novel	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
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	novel	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	novel	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	novel	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
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*	novel	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	novel	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	RNF168	SNV het	3	196199204	G	T	P401Q	0.46947	1063
PML35	RNF168	SNV het	3	196214320	C	T	E170K	0.00818	1066
PML36	IDO2	SNV het	8	39862881	C	T	R235W	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
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	R235W	0.46108	1125
PML45	VPS13B	SNV het	8	100791158	G	A	E2560K	0.00964	1140
PML45	VPS13B	SNV het	8	100865941	G	A	A3442T	novel	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	novel	1315
PML51	TRPM2	SNV het	21	45855099	C	T	R1300W	0.00021	1322
PML53	EPG5	SNV het	18	43445580	C	T	D181N	novel	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
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	novel	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
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	51712759	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	novel	1243

[0329] Table 9 lists, for each case (in multiple rows), variants for which it was not possible, using the whole exome sequencing (WES) data available, to determine phase (e.g., whether two variants are in cis - on the same chromosome - or trans - on opposite chromosomes). Determining phase is an important consideration when dealing with disorders that are being evaluated on an autosomal recessive (AR) basis. If two variants are known to be present but it is impossible to determine whether they are in cis or trans, then it is

impossible to conclude that both gene copies are affected, as opposed to only one (albeit with 2 variants). This problem does not arise in the case of homozygous variants, for which it is obvious that the variants must be in trans (e.g., it is only an issue for non-identical variants). All genome coordinates are based on hg19 build.

[0330] In summary, **Table 9** lists all unphased case-level compound heterozygous SNV solutions, which might represent further case-level solutions, were phasing to have been possible. Furthermore, it can be appreciated by those skilled in the art that unphased solutions reported in **Table 9** (2 het SNVs per gene) or **Table 10** (see below, which reports het SNVs in patients that also have a CNV reported in **Table 1**) can potentially cause or contribute to the patient's PML if follow up genetic analysis reveals the pair of variants are on different alleles (e.g., each gene copy impacted by a variant). Variants reported in **Tables 1, 9,** or **10** may also be found to be significantly deleterious on their own (e.g., in functional studies on patient-derived cells, animal models, etc.) and thus constitute an AD model solution (e.g., genes presently listed as 'AR' model in **Table 6**) may be causal or contributing to disease via an AD or AR model, like several genes already known to be AD or AR (**Table 6**, 'AD AR' disease model).

Table 10: SNVs found in genes suspected of being impacted by acquired CNVs

Sample ID	RefSeq Gene Symbol	Variant Type	Chr	Position	Ref Allele	Alt Allele	Amino Acid Change	Variant Frequency Details (Ethnic specific)	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	Q4048R	0.49837	1104
PML02	DUSP16 [1]	SNV het	12	12673965	G	A	T23M	0.00015	1199
PML39	SALL2	SNV het	14	22004996	G	T	S13R	0.00231	1225
PML51	JMY	SNV het	5	78596018	G	C	D524H	novel	1086
PML65	SALL2	SNV het	14	21992397	T	C	S347G	0.07709	1223

[1] The DUSP16 SNV (chr12:12673965) was in trans with a chr12 deletion of DUSP16 in this patient (PML02), whose primary diagnosis was lymphoma.

[0331] Table 10 is a list of all heterozygous SNVs that are potentially compound heterozygotes with a CNV on the allele. See text for a fuller explanation. All genome coordinates are based on hg19 build.

Table 11: Key that maps Sample ID for the PML cases to the PML Case ID numbers

Sample ID	PML Case ID
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

Sample ID	PML Case ID
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
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

Sample ID	PML Case ID
PML73 control	3283
PML74 control	3284
PML75 control	3285
PML76 control	3286

[0332] Table 11 provides the Sample ID and PML Case ID (experimental ID for CGH data) for 77 'PML cases' (includes 6 non-PML HIV cases listed as controls).

Table 12: Non-redundant list of transcript variants that correspond to the set of genes that no CNV 'solutions' have been reported in the 71 PML cases

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
ACADM	NM 000016	Homo sapiens acyl-CoA dehydrogenase, C-4 to C-12 straight chain (ACADM), transcript variant 1, mRNA.	1500
ACADM	NM 001127328	Homo sapiens atypical chemokine receptor 1 (Duffy blood group) (ACKR1), transcript variant 1, mRNA.	1501
ACKR1	NM 002036	Homo sapiens atypical chemokine receptor 1 (Duffy blood group) (ACKR1), transcript variant 2, mRNA.	1502
ACKR1	NM 001122951	Homo sapiens atypical chemokine receptor 1 isoform a	1503
ACP5	NM 001611	Homo sapiens acid phosphatase 5, tartrate resistant (ACP5), transcript variant 4, mRNA.	1504
ACP5	NM 001111034	Homo sapiens acid phosphatase 5, tartrate resistant (ACP5), transcript variant 2, mRNA.	1505
ACP5	NM 001111035	Homo sapiens acid phosphatase 5, tartrate resistant (ACP5), transcript variant 1, mRNA.	1506
ACP5	NM 001111036	Homo sapiens acid phosphatase 5, tartrate resistant (ACP5), transcript variant 3, mRNA.	1507
ADAR	NM 001111	Homo sapiens adenosine deaminase, RNA-specific (ADAR), transcript variant 1, mRNA.	1508
ADAR	NM 015840	Homo sapiens adenosine deaminase, RNA-specific (ADAR), transcript variant 2, mRNA.	1509
ADAR	NM 015841	Homo sapiens adenosine deaminase, RNA-specific (ADAR), transcript variant 3, mRNA.	1510
ADAR	NM 001025107	Homo sapiens adenosine deaminase, RNA-specific (ADAR), transcript variant 4, mRNA.	1511
ADAR	NM 001193495	Homo sapiens adenosine deaminase, RNA-specific (ADAR), transcript variant 5, mRNA.	1512
ADK	NM 001202450	Homo sapiens adenosine kinase (ADK), transcript variant 4, mRNA.	1513
ADK	NM 006721	Homo sapiens adenosine kinase (ADK), transcript variant 2, mRNA.	1514
ADK	NM 001123	Homo sapiens adenosine kinase (ADK), transcript variant 1, mRNA.	1515
ADK	NM 001202449	Homo sapiens adenosine kinase (ADK), transcript variant 3, mRNA.	1516
AICDA	NM 020661	Homo sapiens activation-induced cytidine deaminase (AICDA), mRNA.	1517
AK2	NM 001199199	Homo sapiens adenylate kinase 2 (AK2), transcript variant 3, mRNA.	1518
AK2	NM 013411	Homo sapiens adenylate kinase 2 (AK2), transcript variant 2, mRNA.	1519
AK2	NM 001625	Homo sapiens adenylate kinase 2 (AK2), transcript variant 1, mRNA.	1520
ALG12	NM 024105	Homo sapiens ALG12, alpha-1,6-mannosyltransferase (ALG12), mRNA.	1521
ALPL	NM 000478	Homo sapiens alkaline phosphatase, liver/bone/kidney (ALPL), transcript variant 1, mRNA.	1522
ALPL	NM 001127501	Homo sapiens alkaline phosphatase, liver/bone/kidney (ALPL), transcript variant 2, mRNA.	1523

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
ALPL	NM 001177520	Homo sapiens alkaline phosphatase, liver/bone/kidney (ALPL), transcript variant 3, mRNA.	1524
AP3B1	NM 001271769	Homo sapiens adaptor related protein complex 3 beta 1 subunit (AP3B 1), transcript variant 2, mRNA.	1525
AP3B1	NM 003664	Homo sapiens adaptor related protein complex 3 beta 1 subunit (AP3B 1), transcript variant 1, mRNA.	1526
AP3B2	NM 004644	Homo sapiens adaptor-related protein complex 3, beta 2 subunit (AP3B2), transcript variant 2, mRNA.	1527
AP3D1	NM 003938	Homo sapiens adaptor-related protein complex 3, delta 1 subunit (AP3D1), transcript variant 2, mRNA.	1528
AP3D1	NM 001261826	Homo sapiens adaptor-related protein complex 3, delta 1 subunit (AP3D1), transcript variant 3, mRNA.	1529
APOL1	NM 001136540	Homo sapiens apolipoprotein L1 (APOL1), transcript variant 3, mRNA.	1530
APOL1	NM 001136541	Homo sapiens apolipoprotein L1 (APOL 1), transcript variant 4, mRNA.	1531
APOL1	NM 003661	Homo sapiens apolipoprotein L1 (APOL1), transcript variant 1, mRNA.	1532
APOL1	NM 145343	Homo sapiens apolipoprotein L1 (APOL 1), transcript variant 2, mRNA.	1533
ASH1L	NM 018489	Homo sapiens ASH1 like histone lysine methyltransferase (ASH1L), mRNA.	1534
ATL2	NM 001135673	Homo sapiens atlastin GTPase 2 (ATL2), transcript variant 2, mRNA.	1535
ATL2	NM 022374	Homo sapiens atlastin GTPase 2 (ATL2), transcript variant 1, mRNA.	1536
ATL2	NR 024191	Homo sapiens atlastin GTPase 2 (ATL2), transcript variant 3, non-coding RNA.	1537
ATM	NM 000051	Homo sapiens ATM serine/threonine kinase (ATM), mRNA.	1538
ATR	NM 001184	Homo sapiens ATR serine/threonine kinase (ATR), mRNA.	1539
BACH2	NM 001170794	Homo sapiens BTB domain and CNC homolog 2 (BACH2), transcript variant 2, mRNA.	1540
BACH2	NM 021813	Homo sapiens BTB domain and CNC homolog 2 (BACH2), transcript variant 1, mRNA.	1541
BAG3	NM 004281	Homo sapiens BCL2 associated athanogene 3 (BAG3), mRNA.	1542
BCL10	NM 003921	Homo sapiens B-cell CLL/lymphoma 10 (BCL10), transcript variant 1, mRNA.	1543
BLM	NM 000057	Homo sapiens Bloom syndrome RecQ like helicase (BLM), transcript variant 1, mRNA.	1544
BLNK	NM 001114094	Homo sapiens B-cell linker (BLNK), transcript variant 2, mRNA.	1545
BLNK	NM 001258440	Homo sapiens B-cell linker (BLNK), transcript variant 3, mRNA.	1546
BLNK	NM 001258441	Homo sapiens B-cell linker (BLNK), transcript variant 4, mRNA.	1547
BLNK	NM 001258442	Homo sapiens B-cell linker (BLNK), transcript variant 5, mRNA.	1548
BLNK	NM 013314	Homo sapiens B-cell linker (BLNK), transcript variant 1, mRNA.	1549
BLNK	NR 047680	Homo sapiens B-cell linker (BLNK), transcript variant 6, non-coding RNA.	1550
BLNK	NR 047681	Homo sapiens B-cell linker (BLNK), transcript variant 7, non-coding RNA.	1551
BLNK	NR 047682	Homo sapiens B-cell linker (BLNK), transcript variant 8, non-coding RNA.	1552
BLNK	NR 047683	Homo sapiens B-cell linker (BLNK), transcript variant 9, non-coding RNA.	1553
BLOCIS 6	NM 012388	Homo sapiens biogenesis of lysosomal organelles complex 1	1554

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		subunit 6 (BLOC1S6), transcript variant 2, mRNA.	
BTK	NM 000061	Homo sapiens Bruton tyrosine kinase (BTK), transcript variant 1, mRNA.	1555
C11orf65	NM 152587	Homo sapiens chromosome 11 open reading frame 65 (C11orf65), mRNA.	1556
C1QA	NM 015991	Homo sapiens complement component 1, q subcomponent, A chain (C1QA), mRNA.	1557
C1QB	NM 000491	Homo sapiens complement component 1, q subcomponent, B chain (C1QB), mRNA.	1558
ClQC	NM 001114101	Homo sapiens complement component 1, q subcomponent, C chain (C1QC), transcript variant 1, mRNA.	1559
ClQC	NM 172369	Homo sapiens complement component 1, q subcomponent, C chain (C1QC), transcript variant 2, mRNA.	1560
C5AR1	NM 001736	Homo sapiens complement component 5a receptor 1 (C5AR1), mRNA.	1561
CARD11	NM 032415	Homo sapiens caspase recruitment domain family member 11 (CARD11), transcript variant 2, mRNA.	1562
CARD9	NM 052813	Homo sapiens caspase recruitment domain family, member 9 (CARD9), transcript variant 1, mRNA.	1563
CARD9	NM 052814	Homo sapiens caspase recruitment domain family, member 9 (CARD9), transcript variant 2, mRNA.	1564
CASP8	NM 001080124	Homo sapiens caspase 8 (CASP8), transcript variant F, mRNA.	1565
CASP8	NM 001228	Homo sapiens caspase 8 (CASP8), transcript variant A, mRNA.	1566
CASP8	NM 033355	Homo sapiens caspase 8 (CASP8), transcript variant B, mRNA.	1567
CASP8	NM 033358	Homo sapiens caspase 8 (CASP8), transcript variant E, mRNA.	1568
CASP8	NM 001080125	Homo sapiens caspase 8 (CASP8), transcript variant G, mRNA.	1569
CASP8	NM 033356	Homo sapiens caspase 8 (CASP8), transcript variant C, mRNA.	1570
CCL11	NM 002986	Homo sapiens C-C motif chemokine ligand 11 (CCL11), mRNA.	1571
CCL2	NM 002982	Homo sapiens C-C motif chemokine ligand 2 (CCL2), mRNA.	1572
CCL5	NM 002985	Homo sapiens C-C motif chemokine ligand 5 (CCL5), transcript variant 1, mRNA.	1573
CCR2	NM 001123041	Homo sapiens C-C motif chemokine receptor 2 (CCR2), transcript variant A, mRNA.	1574
CCR2	NM 001123396	Homo sapiens C-C motif chemokine receptor 2 (CCR2), transcript variant B, mRNA.	1575
CCR5	NM 000579	Homo sapiens C-C motif chemokine receptor 5 (gene/pseudogene) (CCR5), transcript variant A, mRNA.	1576
CCR5	NM 001100168	Homo sapiens C-C motif chemokine receptor 5 (gene/pseudogene) (CCR5), transcript variant B, mRNA.	1577
CD 180	NM 005582	Homo sapiens CD180 molecule (CD180), mRNA.	1578
CD19	NM 001178098	Homo sapiens CD19 molecule (CD19), transcript variant 1, mRNA.	1579
CD19	NM 001770	Homo sapiens CD19 molecule (CD19), transcript variant 2, mRNA.	1580
CD209	NM 001144893	Homo sapiens CD209 molecule (CD209), transcript variant 5, mRNA.	1581
CD209	NM 001144894	Homo sapiens CD209 molecule (CD209), transcript variant 6, mRNA.	1582
CD209	NM 001144895	Homo sapiens CD209 molecule (CD209), transcript variant 7, mRNA.	1583
CD209	NM 001144896	Homo sapiens CD209 molecule (CD209), transcript variant 3, mRNA.	1584

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
CD209	NM 001144897	Homo sapiens CD209 molecule (CD209), transcript variant 4, mRNA.	1585
CD209	NM 001144899	Homo sapiens CD209 molecule (CD209), transcript variant 8, mRNA.	1586
CD209	NM 021155	Homo sapiens CD209 molecule (CD209), transcript variant 1, mRNA.	1587
CD209	NR 026692	Homo sapiens CD209 molecule (CD209), transcript variant 2, non-coding RNA.	1588
CD247	NM 000734	Homo sapiens CD247 molecule (CD247), transcript variant 2, mRNA.	1589
CD247	NM 198053	Homo sapiens CD247 molecule (CD247), transcript variant 1, mRNA.	1590
CD27	NM 001242	Homo sapiens CD27 molecule (CD27), mRNA.	1591
CD27-AS1	NR 015382	Homo sapiens CD27 antisense RNA 1 (CD27-AS1), long non-coding RNA.	1592
CD34	NM 001025109	Homo sapiens CD34 molecule (CD34), transcript variant 1, mRNA.	1593
CD34	NM 001773	Homo sapiens CD34 molecule (CD34), transcript variant 2, mRNA.	1594
CD3D	NM 000732	Homo sapiens CD3d molecule (CD3D), transcript variant 1, mRNA.	1595
CD3D	NM 001040651	Homo sapiens CD3d molecule (CD3D), transcript variant 2, mRNA.	1596
CD3E	NM 000733	Homo sapiens CD3e molecule (CD3E), mRNA.	1597
CD3G	NM 000073	Homo sapiens CD3g molecule (CD3G), mRNA.	1598
CD40	NM 001250	Homo sapiens CD40 molecule (CD40), transcript variant 1, mRNA.	1599
CD40	NM 152854	Homo sapiens CD40 molecule (CD40), transcript variant 2, mRNA.	1600
CD40LG	NM 000074	Homo sapiens CD40 ligand (CD40LG), mRNA.	1601
CD55	NM 000574	Homo sapiens CD55 molecule (Cromer blood group) (CD55), transcript variant 1, mRNA.	1602
CD55	NM 001114752	Homo sapiens CD55 molecule (Cromer blood group) (CD55), transcript variant 2, mRNA.	1603
CD59	NM 000611	Homo sapiens CD59 molecule (CD59), transcript variant 2, mRNA.	1604
CD59	NM 001127223	Homo sapiens CD59 molecule (CD59), transcript variant 5, mRNA.	1605
CD59	NM 001127225	Homo sapiens CD59 molecule (CD59), transcript variant 6, mRNA.	1606
CD59	NM 001127226	Homo sapiens CD59 molecule (CD59), transcript variant 7, mRNA.	1607
CD59	NM 001127227	Homo sapiens CD59 molecule (CD59), transcript variant 8, mRNA.	1608
CD59	NM 203329	Homo sapiens CD59 molecule (CD59), transcript variant 3, mRNA.	1609
CD59	NM 203330	Homo sapiens CD59 molecule (CD59), transcript variant 1, mRNA.	1610
CD59	NM 203331	Homo sapiens CD59 molecule (CD59), transcript variant 4, mRNA.	1611
CD79A	NM 001783	Homo sapiens CD79a molecule (CD79A), transcript variant 1, mRNA.	1612
CD79A	NM 021601	Homo sapiens CD79a molecule (CD79A), transcript variant 2, mRNA.	1613
CD79B	NM 000626	Homo sapiens CD79b molecule (CD79B), transcript variant 1, mRNA.	1614
CD79B	NM 001039933	Homo sapiens CD79b molecule (CD79B), transcript variant 3, mRNA.	1615

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
CD79B	NM 021602	Homo sapiens CD79b molecule (CD79B), transcript variant 2, mRNA.	1616
CD81	NM 004356	Homo sapiens CD81 molecule (CD81), transcript variant 1, mRNA.	1617
CD8A	NM 001145873	Homo sapiens CD8a molecule (CD8A), transcript variant 3, mRNA.	1618
CD8A	NM 001768	Homo sapiens CD8a molecule (CD8A), transcript variant 1, mRNA.	1619
CD8A	NM 171827	Homo sapiens CD8a molecule (CD8A), transcript variant 2, mRNA.	1620
CD8A	NR 027353	Homo sapiens CD8a molecule (CD8A), transcript variant 4, non-coding RNA.	1621
CDCA7	NM 031942	Homo sapiens cell division cycle associated 7 (CDCA7), transcript variant 1, mRNA.	1622
CDCA7	NM 145810	Homo sapiens cell division cycle associated 7 (CDCA7), transcript variant 2, mRNA.	1623
CEBPB	NM 005194	Homo sapiens CCAAT/enhancer binding protein beta (CEBPB), transcript variant 1, mRNA.	1624
CHD7	NM 017780	Homo sapiens chromodomain helicase DNA binding protein 7 (CHD7), transcript variant 1, mRNA.	1625
CHEK1	NM 001114121	Homo sapiens checkpoint kinase 1 (CHEK1), transcript variant 2, mRNA.	1626
CHEK1	NM 001114122	Homo sapiens checkpoint kinase 1 (CHEK1), transcript variant 1, mRNA.	1627
CHEK1	NM 001244846	Homo sapiens checkpoint kinase 1 (CHEK1), transcript variant 4, mRNA.	1628
CHEK1	NR 045204	Homo sapiens checkpoint kinase 1 (CHEK1), transcript variant 5, non-coding RNA.	1629
CHEK1	NR 045205	Homo sapiens checkpoint kinase 1 (CHEK1), transcript variant 6, non-coding RNA.	1630
CHEK1	NM 001274	Homo sapiens checkpoint kinase 1 (CHEK1), transcript variant 3, mRNA.	1631
CIITA	NM 000246	Homo sapiens class II major histocompatibility complex transactivator (CIITA), transcript variant 2, mRNA.	1632
CLCN7	NM 001114331	Homo sapiens chloride channel, voltage-sensitive 7 (CLCN7), transcript variant 2, mRNA.	1633
CLCN7	NM 001287	Homo sapiens chloride channel, voltage-sensitive 7 (CLCN7), transcript variant 1, mRNA.	1634
COG6	NM 001145079	Homo sapiens component of oligomeric golgi complex 6 (COG6), transcript variant 2, mRNA.	1635
COG6	NM 020751	Homo sapiens component of oligomeric golgi complex 6 (COG6), transcript variant 1, mRNA.	1636
COG6	NR 026745	Homo sapiens component of oligomeric golgi complex 6 (COG6), transcript variant 3, non-coding RNA.	1637
CORO1 A	NM 001193333	Homo sapiens coronin 1A (CORO1A), transcript variant 1, mRNA.	1638
CORO1 A	NM 007074	Homo sapiens coronin 1A (CORO1A), transcript variant 2, mRNA.	1639
CR2	NM 001006658	Homo sapiens complement component 3d receptor 2 (CR2), transcript variant 1, mRNA.	1640
CR2	NM 001877	Homo sapiens complement component 3d receptor 2 (CR2), transcript variant 2, mRNA.	1641
CRTC3	NM 001042574	Homo sapiens CREB regulated transcription coactivator 3 (CRTC3), transcript variant 2, mRNA.	1642
CRTC3	NM 022769	Homo sapiens CREB regulated transcription coactivator 3 (CRTC3), transcript variant 1, mRNA.	1643
CSF3R	NM 000760	Homo sapiens colony stimulating factor 3 receptor (granulocyte) (CSF3R), transcript variant 1, mRNA.	1644
CSF3R	NM 156039	Homo sapiens colony stimulating factor 3 receptor	1645

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		(granulocyte) (CSF3R), transcript variant 3, mRNA.	
CSF3R	NM 172313	Homo sapiens colony stimulating factor 3 receptor (granulocyte) (CSF3R), transcript variant 4, mRNA.	1646
CTLA4	NM 005214	Homo sapiens cytotoxic T-lymphocyte-associated protein 4 (CTLA4), transcript variant 1, mRNA.	1647
CTLA4	NM 001037631	Homo sapiens cytotoxic T-lymphocyte-associated protein 4 (CTLA4), transcript variant 2, mRNA.	1648
CTPS1	NM 001905	Homo sapiens CTP synthase 1 (CTPS1), transcript variant 1, mRNA.	1649
CTSC	NM 148170	Homo sapiens cathepsin C (CTSC), transcript variant 2, mRNA.	1650
CTSC	NM 001114173	Homo sapiens cathepsin C (CTSC), transcript variant 3, mRNA.	1651
CTSC	NM 001814	Homo sapiens cathepsin C (CTSC), transcript variant 1, mRNA.	1652
CX3CR1	NN 001171171	Homo sapiens C-X3-C motif chemokine receptor 1 (CX3CR1), transcript variant 2, mRNA.	1653
CX3CR1	NM 001171172	Homo sapiens C-X3-C motif chemokine receptor 1 (CX3CR1), transcript variant 3, mRNA.	1654
CX3CR1	NM 001171174	Homo sapiens C-X3-C motif chemokine receptor 1 (CX3CR1), transcript variant 1, mRNA.	1655
CX3CR1	NM 001337	Homo sapiens C-X3-C motif chemokine receptor 1 (CX3CR1), transcript variant 4, mRNA.	1656
CXCL12	NM 000609	Homo sapiens C-X-C motif chemokine ligand 12 (CXCL12), transcript variant 2, mRNA.	1657
CXCL12	NM 001033886	Homo sapiens C-X-C motif chemokine ligand 12 (CXCL12), transcript variant 3, mRNA.	1658
CXCL12	NM 001178134	Homo sapiens C-X-C motif chemokine ligand 12 (CXCL12), transcript variant 4, mRNA.	1659
CXCL12	NM 199168	Homo sapiens C-X-C motif chemokine ligand 12 (CXCL12), transcript variant 1, mRNA.	1660
CXCL9	NM 002416	Homo sapiens C-X-C motif chemokine ligand 9 (CXCL9), mRNA.	1661
CXCR1	NM 000634	Homo sapiens C-X-C motif chemokine receptor 1 (CXCR1), mRNA.	1662
CXCR4	NM 001008540	Homo sapiens C-X-C motif chemokine receptor 4 (CXCR4), transcript variant 1, mRNA.	1663
CXCR4	NM 003467	Homo sapiens C-X-C motif chemokine receptor 4 (CXCR4), transcript variant 2, mRNA.	1664
CXorf40 A	NM 001171907	Homo sapiens chromosome X open reading frame 40A (CXorf40A), transcript variant 2, mRNA.	1665
CXorf40 A	NM 001171908	Homo sapiens chromosome X open reading frame 40A (CXorf40A), transcript variant 3, mRNA.	1666
CXorf40 A	NM 178124	Homo sapiens chromosome X open reading frame 40A (CXorf40A), transcript variant 1, mRNA.	1667
CXorf40 A	NM 001171909	Homo sapiens chromosome X open reading frame 40A (CXorf40A), transcript variant 4, mRNA.	1668
CYBB	NM 000397	Homo sapiens cytochrome b-245, beta polypeptide (CYBB), mRNA.	1669
CYP2S1	NM 030622	Homo sapiens cytochrome P450 family 2 subfamily S member 1 (CYP2S1), mRNA.	1670
DCLRE1 C	NM 001033855	Homo sapiens DNA cross-link repair 1C (DCLRE1C), transcript variant a, mRNA.	1671
DCLRE1 C	NM 001033857	Homo sapiens DNA cross-link repair 1C (DCLRE1C), transcript variant d, mRNA.	1672
DCLRE1 C	NM 001033858	Homo sapiens DNA cross-link repair 1C (DCLRE1C), transcript variant c, mRNA.	1673
DCLRE1 C	NM 022487	Homo sapiens DNA cross-link repair 1C (DCLRE1C), transcript variant b, mRNA.	1674

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
DDX1	NM 004939	Homo sapiens DEAD/H-box helicase 1 (DDX1), mRNA.	1675
DDX58	NM 014314	Homo sapiens DEAD (Asp-Glu-Ala-Asp) box polypeptide 58 (DDX58), mRNA.	1676
DHX58	NM 024119	Homo sapiens DEXH (Asp-Glu-X-His) box polypeptide 58 (DHX58), mRNA.	1677
DKC1	NM 001142463	Homo sapiens dyskerin pseudouridine synthase 1 (DKC1), transcript variant 2, mRNA.	1678
DKC1	NM 001363	Homo sapiens dyskerin pseudouridine synthase 1 (DKC1), transcript variant 1, mRNA.	1679
DNMT3 B	NM 001207055	Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 7, mRNA.	1680
DNMT3 B	NM 001207056	Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 8, mRNA.	1681
DNMT3 B	NM 006892	Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 1, mRNA.	1682
DNMT3 B	NM 175848	Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 2, mRNA.	1683
DNMT3 B	NM 175849	Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 3, mRNA.	1684
DNMT3 B	NM 175850	Homo sapiens DNA methyltransferase 3 beta (DNMT3B), transcript variant 6, mRNA.	1685
DOCK2	NM 004946	Homo sapiens dedicator of cytokinesis 2 (DOCK2), mRNA.	1686
DOCK8	NM 203447	Homo sapiens dedicator of cytokinesis 8 (DOCK8), transcript variant 1, mRNA.	1687
DOCK8	NM 001190458	Homo sapiens dedicator of cytokinesis 8 (DOCK8), transcript variant 2, mRNA.	1688
DOCK8	NM 001193536	Homo sapiens dedicator of cytokinesis 8 (DOCK8), transcript variant 3, mRNA.	1689
DSC1	NM 004948	Homo sapiens desmocollin 1 (DSC1), transcript variant Dsc1b, mRNA.	1690
DSC1	NM 024421	Homo sapiens desmocollin 1 (DSC1), transcript variant Dscla, mRNA.	1691
EGR1	NM 001964	Homo sapiens early growth response 1 (EGR1), mRNA.	1692
ELANE	NM 001972	Homo sapiens elastase, neutrophil expressed (ELANE), mRNA.	1693
EPG5	NM 020964	Homo sapiens ectopic P-granules autophagy protein 5 homolog (EPG5), mRNA.	1694
ETF1	NM 004730	Homo sapiens eukaryotic translation termination factor 1 (ETF1), transcript variant 1, mRNA.	1695
F9	NM 000133	Homo sapiens coagulation factor IX (F9), transcript variant 1, mRNA.	1696
FAS	NM 000043	Homo sapiens Fas cell surface death receptor (FAS), transcript variant 1, mRNA.	1697
FAS	NM 152871	Homo sapiens Fas cell surface death receptor (FAS), transcript variant 2, mRNA.	1698
FAS	NM 152872	Homo sapiens Fas cell surface death receptor (FAS), transcript variant 3, mRNA.	1699
FAS	NR 028033	Homo sapiens Fas cell surface death receptor (FAS), transcript variant 4, non-coding RNA.	1700
FAS	NR 028034	Homo sapiens Fas cell surface death receptor (FAS), transcript variant 5, non-coding RNA.	1701
FAS	NR 028035	Homo sapiens Fas cell surface death receptor (FAS), transcript variant 6, non-coding RNA.	1702
FAS	NR 028036	Homo sapiens Fas cell surface death receptor (FAS), transcript variant 7, non-coding RNA.	1703
FASLG	NM 000639	Homo sapiens Fas ligand (TNF superfamily, member 6) (FASLG), mRNA.	1704
FCGR2A	NM 001136219	Homo sapiens Fc fragment of IgG receptor IIa (FCGR2A), transcript variant 1, mRNA.	1705

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
FCGR2A	NM 021642	Homo sapiens Fc fragment of IgG receptor IIa (FCGR2A), transcript variant 2, mRNA.	1706
FCGR3A	NM 000569	Homo sapiens Fc fragment of IgG receptor IIIa (FCGR3A), transcript variant 1, mRNA.	1707
FCGR3A	NM 001127592	Homo sapiens Fc fragment of IgG receptor IIIa (FCGR3A), transcript variant 2, mRNA.	1708
FCGR3A	NM 001127593	Homo sapiens Fc fragment of IgG receptor IIIa (FCGR3A), transcript variant 3, mRNA.	1709
FCGR3A	NM 001127595	Homo sapiens Fc fragment of IgG receptor IIIa (FCGR3A), transcript variant 4, mRNA.	1710
FCGR3A	NM 001127596	Homo sapiens Fc fragment of IgG receptor IIIa (FCGR3A), transcript variant 5, mRNA.	1711
FCN3	NM 003665	Homo sapiens ficolin 3 (FCN3), transcript variant 1, mRNA.	1712
FCN3	NM 173452	Homo sapiens ficolin 3 (FCN3), transcript variant 2, mRNA.	1713
FEZ1	NM 005103	Homo sapiens fasciculation and elongation protein zeta 1 (FEZ1), transcript variant 1, mRNA.	1714
FEZ1	NM 022549	Homo sapiens fasciculation and elongation protein zeta 1 (FEZ1), transcript variant 2, mRNA.	1715
FOS	NM 005252	Homo sapiens Fos proto-oncogene, AP-1 transcription factor subunit (FOS), mRNA.	1716
FOXH1	NM 003923	Homo sapiens forkhead box H1 (FOXH1), mRNA.	1717
FOXN1	NM 003593	Homo sapiens forkhead box N1 (FOXN1), mRNA.	1718
FOXP3	NM 001114377	Homo sapiens forkhead box P3 (FOXP3), transcript variant 2, mRNA.	1719
FOXP3	NM 014009	Homo sapiens forkhead box P3 (FOXP3), transcript variant 1, mRNA.	1720
FPR1	NM 001193306	Homo sapiens formyl peptide receptor 1 (FPR1), transcript variant 1, mRNA.	1721
FPR1	NM 002029	Homo sapiens formyl peptide receptor 1 (FPR1), transcript variant 2, mRNA.	1722
G6PC3	NM 138387	Homo sapiens glucose 6 phosphatase, catalytic, 3 (G6PC3), transcript variant 1, mRNA.	1723
G6PC3	NR 028582	Homo sapiens glucose 6 phosphatase, catalytic, 3 (G6PC3), transcript variant 2, non-coding RNA.	1724
G6PC3	NR 028581	Homo sapiens glucose 6 phosphatase, catalytic, 3 (G6PC3), transcript variant 3, non-coding RNA.	1725
GATA2	NM 001145661	Homo sapiens GATA binding protein 2 (GATA2), transcript variant 1, mRNA.	1726
GATA2	NM 001145662	Homo sapiens GATA binding protein 2 (GATA2), transcript variant 3, mRNA.	1727
GATA2	NM 032638	Homo sapiens GATA binding protein 2 (GATA2), transcript variant 2, mRNA.	1728
GFI1	NM 005263	Homo sapiens growth factor independent 1 transcription repressor (GFI1), transcript variant 1, mRNA.	1729
GFI1	NM 001127215	Homo sapiens growth factor independent 1 transcription repressor (GFI1), transcript variant 2, mRNA.	1730
GFI1	NM 001127216	Homo sapiens growth factor independent 1 transcription repressor (GFI1), transcript variant 3, mRNA.	1731
GOLGB1	NM 001256486	Homo sapiens golgin B1 (GOLGB 1), transcript variant 1, mRNA.	1732
GOLGB1	NM 001256487	Homo sapiens golgin B 1 (GOLGB 1), transcript variant 3, mRNA.	1733
GOLGB1	NM 001256488	Homo sapiens golgin B 1 (GOLGB 1), transcript variant 4, mRNA.	1734
GOLGB1	NM 004487	Homo sapiens golgin B 1 (GOLGB 1), transcript variant 2, mRNA.	1735
GPRC5A	NM 003979	Homo sapiens G protein-coupled receptor class C group 5 member A (GPRC5A), mRNA.	1736

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
GRAP2	NM 004810	Homo sapiens GRB2-related adaptor protein 2 (GRAP2), mRNA.	1737
HAX1	NM 006118	Homo sapiens HCLS1 associated protein X-1 (HAX1), transcript variant 1, mRNA.	1738
HAX1	NM 001018837	Homo sapiens HCLS1 associated protein X-1 (HAX1), transcript variant 2, mRNA.	1739
HELLS	NM 018063	Homo sapiens helicase, lymphoid-specific (HELLS), transcript variant 1, mRNA.	1740
HIVEP1	NM 002114	Homo sapiens human immunodeficiency virus type I enhancer binding protein 1 (HIVEP1), mRNA.	1741
HIVEP2	NM 006734	Homo sapiens human immunodeficiency virus type I enhancer binding protein 2 (HIVEP2), mRNA.	1742
HIVEP3	NM 001127714	Homo sapiens human immunodeficiency virus type I enhancer binding protein 3 (HIVEP3), transcript variant 2, mRNA.	1743
HIVEP3	NM 024503	Homo sapiens human immunodeficiency virus type I enhancer binding protein 3 (HIVEP3), transcript variant 1, mRNA.	1744
HIVEP3	NR 038260	Homo sapiens human immunodeficiency virus type I enhancer binding protein 3 (HIVEP3), transcript variant 3, non-coding RNA.	1745
HIVEP3	NR 038261	Homo sapiens human immunodeficiency virus type I enhancer binding protein 3 (HIVEP3), transcript variant 4, non-coding RNA.	1746
HP	NM 001126102	Homo sapiens haptoglobin (HP), transcript variant 2, mRNA.	1747
HP	NM 005143	Homo sapiens haptoglobin (HP), transcript variant 1, mRNA.	1748
HPCAL1	NM 002149	Homo sapiens hippo calc in like 1 (HPCAL1), transcript variant 1, mRNA.	1749
HPCAL1	NM 134421	Homo sapiens hippo calc in like 1 (HPCAL1), transcript variant 2, mRNA.	1750
HPCAL1	NM 001258357	Homo sapiens hippo calc in like 1 (HPCAL1), transcript variant 3, mRNA.	1751
HPCAL1	NM 001258358	Homo sapiens hippo calc in like 1 (HPCAL 1), transcript variant 4, mRNA.	1752
HPCAL1	NM 001258359	Homo sapiens hippo calc in like 1 (HPCAL1), transcript variant 5, mRNA.	1753
HTR2A	NM 000621	Homo sapiens 5-hydroxytryptamine receptor 2A (HTR2A), transcript variant 1, mRNA.	1754
HTR2A	NM 001165947	Homo sapiens 5-hydroxytryptamine (serotonin) receptor 2A, G protein-coupled (HTR2A), transcript variant 2, mRNA.	1755
ICOS	NM 012092	Homo sapiens inducible T-cell costimulator (ICOS), mRNA.	1756
IDI1	NM 004508	Homo sapiens isopentenyl-diphosphate delta isomerase 1 (IDI1), transcript variant 1, mRNA.	1757
IFIH 1	NM 022168	Homo sapiens interferon induced with helicase C domain 1 (IFIH1), mRNA.	1758
IFNAR 1	NM 000629	Homo sapiens interferon (alpha, beta and omega) receptor 1 (IFNAR1), mRNA.	1759
IFNAR2	NM 207584	Homo sapiens interferon (alpha, beta and omega) receptor 2 (IFNAR2), transcript variant 3, mRNA.	1760
IFNAR2	NM 207585	Homo sapiens interferon (alpha, beta and omega) receptor 2 (IFNAR2), transcript variant 1, mRNA.	1761
IFNAR2	NM 000874	Homo sapiens interferon (alpha, beta and omega) receptor 2 (IFNAR2), transcript variant 2, mRNA.	1762
IFNG	NM 000619	Homo sapiens interferon gamma (IFNG), mRNA.	1763
IFNGR1	NM 000416	Homo sapiens interferon gamma receptor 1 (IFNGR1), mRNA.	1764
IFNGR2	NM 005534	Homo sapiens interferon gamma receptor 2 (interferon gamma transducer 1) (IFNGR2), transcript variant 2, mRNA.	1765
IGLL1	NM 020070	Homo sapiens immunoglobulin lambda like polypeptide 1	1766

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		(IGLL1), transcript variant 1, mRNA.	
IGLL1	NM 152855	Homo sapiens immunoglobulin lambda like polypeptide 1 (IGLL1), transcript variant 2, mRNA.	1767
IKBKB	NM 001190720	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta (IKBKB), transcript variant 2, mRNA.	1768
IKBKB	NM 001242778	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta (IKBKB), transcript variant 7, mRNA.	1769
IKBKB	NM 001556	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta (IKBKB), transcript variant 1, mRNA.	1770
IKBKB	NR 033818	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta (IKBKB), transcript variant 5, non-coding RNA.	1771
IKBKB	NR 033819	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta (IKBKB), transcript variant 6, non-coding RNA.	1772
IKBKB	NR 040009	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta (IKBKB), transcript variant 8, non-coding RNA.	1773
IKBKG	NM 001099856	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma (IKBKG), transcript variant 2, mRNA.	1774
IKBKG	NM 001099857	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma (IKBKG), transcript variant 1, mRNA.	1775
IKBKG	NM 001145255	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma (IKBKG), transcript variant 4, mRNA.	1776
IKBKG	NM 003639	Homo sapiens inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase gamma (IKBKG), transcript variant 3, mRNA.	1777
IKZF1	NM 001220765	Homo sapiens IKAROS family zinc finger 1 (IKZF1), transcript variant 2, mRNA.	1778
IKZF1	NM 001220767	Homo sapiens IKAROS family zinc finger 1 (IKZF1), transcript variant 4, mRNA.	1779
IKZF1	NM 001220768	Homo sapiens IKAROS family zinc finger 1 (IKZF1), transcript variant 5, mRNA.	1780
IKZF1	NM 001220770	Homo sapiens IKAROS family zinc finger 1 (IKZF1), transcript variant 7, mRNA.	1781
IKZF1	NM 001220771	Homo sapiens IKAROS family zinc finger 1 (IKZF1), transcript variant 8, mRNA.	1782
IKZF1	NM 006060	Homo sapiens IKAROS family zinc finger 1 (IKZF1), transcript variant 1, mRNA.	1783
IL10	NM 000572	Homo sapiens interleukin 10 (IL10), mRNA.	1784
IL10RA	NM 001558	Homo sapiens interleukin 10 receptor subunit alpha (IL10RA), transcript variant 1, mRNA.	1785
IL10RA	NR 026691	Homo sapiens interleukin 10 receptor subunit alpha (IL10RA), transcript variant 2, non-coding RNA.	1786
IL10RB	NM 000628	Homo sapiens interleukin 10 receptor subunit beta (IL10RB), mRNA.	1787
IL12B	NM 002187	Homo sapiens interleukin 12B (IL12B), mRNA.	1788
IL12RB1	NM 005535	Homo sapiens interleukin 12 receptor subunit beta 1 (IL12RB1), transcript variant 1, mRNA.	1789
IL12RB1	NM 153701	Homo sapiens interleukin 12 receptor subunit beta 1 (IL12RB1), transcript variant 2, mRNA.	1790
IL17F	NM 052872	Homo sapiens interleukin 17F (IL17F), mRNA.	1791
IL17RA	NM 014339	Homo sapiens interleukin 17 receptor A (IL17RA), transcript variant 1, mRNA.	1792

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
IL1B	NM 000576	Homo sapiens interleukin 1, beta (IL1B), mRNA.	1793
IL21	NM 001207006	Homo sapiens interleukin 21 (IL21), transcript variant 2, mRNA.	1794
IL21	NM 021803	Homo sapiens interleukin 21 (IL21), transcript variant 1, mRNA.	1795
IL21R	NM 181078	Homo sapiens interleukin 21 receptor (IL21R), transcript variant 2, mRNA.	1796
IL21R	NM 181079	Homo sapiens interleukin 21 receptor (IL21R), transcript variant 3, mRNA.	1797
IL21R	NM 021798	Homo sapiens interleukin 21 receptor (IL21R), transcript variant 1, mRNA.	1798
IL2RA	NM 000417	Homo sapiens interleukin 2 receptor, alpha (IL2RA), transcript variant 1, mRNA.	1799
IL2RG	NM 000206	Homo sapiens interleukin 2 receptor subunit gamma (IL2RG), mRNA.	1800
IL4R	NM 000418	Homo sapiens interleukin 4 receptor (IL4R), transcript variant 1, mRNA.	1801
IL4R	NM 001257406	Homo sapiens interleukin 4 receptor (IL4R), transcript variant 3, mRNA.	1802
IL4R	NM 001257407	Homo sapiens interleukin 4 receptor (IL4R), transcript variant 4, mRNA.	1803
IL4R	NM 001257997	Homo sapiens interleukin 4 receptor (IL4R), transcript variant 5, mRNA.	1804
IL7	NM 000880	Homo sapiens interleukin 7 (IL7), transcript variant 1, mRNA.	1805
IL7	NM 001199886	Homo sapiens interleukin 7 (IL7), transcript variant 2, mRNA.	1806
IL7	NM 001199887	Homo sapiens interleukin 7 (IL7), transcript variant 3, mRNA.	1807
IL7	NM 001199888	Homo sapiens interleukin 7 (IL7), transcript variant 4, mRNA.	1808
IL7R	NM 002185	Homo sapiens interleukin 7 receptor (IL7R), transcript variant 1, mRNA.	1809
IRAK4	NM 001114182	Homo sapiens interleukin 1 receptor associated kinase 4 (IRAK4), transcript variant 1, mRNA.	1810
IRAK4	NM 001145256	Homo sapiens interleukin 1 receptor associated kinase 4 (IRAK4), transcript variant 3, mRNA.	1811
IRAK4	NM 001145257	Homo sapiens interleukin 1 receptor associated kinase 4 (IRAK4), transcript variant 4, mRNA.	1812
IRAK4	NM 001145258	Homo sapiens interleukin 1 receptor associated kinase 4 (IRAK4), transcript variant 5, mRNA.	1813
IRAK4	NM 016123	Homo sapiens interleukin 1 receptor associated kinase 4 (IRAK4), transcript variant 2, mRNA.	1814
IRF3	NM 001197122	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 2, mRNA.	1815
IRF3	NM 001197123	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 3, mRNA.	1816
IRF3	NM 001197124	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 4, mRNA.	1817
IRF3	NM 001197125	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 5, mRNA.	1818
IRF3	NM 001197126	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 6, mRNA.	1819
IRF3	NM 001197127	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 7, mRNA.	1820
IRF3	NM 001197128	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 8, mRNA.	1821
IRF3	NM 001571	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 1, mRNA.	1822
IRF3	NR 045568	Homo sapiens interferon regulatory factor 3 (IRF3), transcript variant 9, non-coding RNA.	1823
IRF7	NM 001572	Homo sapiens interferon regulatory factor 7 (IRF7), transcript	1824

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		variant a, mRNA.	
IRF7	NM 004029	Homo sapiens interferon regulatory factor 7 (IRF7), transcript variant b, mRNA.	1825
IRF7	NM 004031	Homo sapiens interferon regulatory factor 7 (IRF7), transcript variant d, mRNA.	1826
IRF8	NM 002163	Homo sapiens interferon regulatory factor 8 (IRF8), mRNA.	1827
IRGM	NM 001145805	Homo sapiens immunity related GTPase M (IRGM), mRNA.	1828
ISG15	NM 005101	Homo sapiens ISG15 ubiquitin-like modifier (ISG15), mRNA.	1829
ITK	NM 005546	Homo sapiens IL2 inducible T-cell kinase (ITK), mRNA.	1830
ITSN2	NM 006277	Homo sapiens intersectin 2 (ITSN2), transcript variant 1, mRNA.	1831
ITSN2	NM 019595	Homo sapiens intersectin 2 (ITSN2), transcript variant 3, mRNA.	1832
ITSN2	NM 147152	Homo sapiens intersectin 2 (ITSN2), transcript variant 2, mRNA.	1833
JAGN1	NM 032492	Homo sapiens jagunal homolog 1 (Drosophila) (JAGN1), mRNA.	1834
JAK3	NM 000215	Homo sapiens Janus kinase 3 (JAK3), mRNA.	1835
JMY	NM 152405	Homo sapiens junction mediating and regulatory protein, p53 cofactor (JMY), mRNA.	1836
JUN	NM 002228	Homo sapiens Jun proto-oncogene, AP-1 transcription factor subunit (JUN), mRNA.	1837
KITLG	NM 000899	Homo sapiens KIT ligand (KITLG), transcript variant b, mRNA.	1838
KITLG	NM 003994	Homo sapiens KIT ligand (KITLG), transcript variant a, mRNA.	1839
LAMTO R2	NM 001145264	Homo sapiens late endosomal/lysosomal adaptor, MAPK and MTOR activator 2 (LAMTOR2), transcript variant 2, mRNA.	1840
LAMTO R2	NM 014017	Homo sapiens late endosomal/lysosomal adaptor, MAPK and MTOR activator 2 (LAMTOR2), transcript variant 1, mRNA.	1841
LCK	NM 005356	Homo sapiens LCK proto-oncogene, Src family tyrosine kinase (LCK), transcript variant 2, mRNA.	1842
LCK	NM 001042771	Homo sapiens LCK proto-oncogene, Src family tyrosine kinase (LCK), transcript variant 1, mRNA.	1843
LCP2	NM 005565	Homo sapiens lymphocyte cytosolic protein 2 (SH2 domain containing leukocyte protein of 76kDa) (LCP2), mRNA.	1844
LIG1	NM 000234	Homo sapiens DNA ligase 1 (LIG1), transcript variant 1, mRNA.	1845
LIG4	NM 001098268	Homo sapiens DNA ligase 4 (LIG4), transcript variant 3, mRNA.	1846
LIG4	NM 002312	Homo sapiens DNA ligase 4 (LIG4), transcript variant 1, mRNA.	1847
LIG4	NM 206937	Homo sapiens DNA ligase 4 (LIG4), transcript variant 2, mRNA.	1848
LRBA	NM 001199282	Homo sapiens LPS responsive beige-like anchor protein (LRBA), transcript variant 1, mRNA.	1849
LRBA	NM 006726	Homo sapiens LPS responsive beige-like anchor protein (LRBA), transcript variant 2, mRNA.	1850
LYST	NM 000081	Homo sapiens lysosomal trafficking regulator (LYST), transcript variant 1, mRNA.	1851
MAGEA 9	NM 005365	Homo sapiens MAGE family member A9 (MAGEA9), mRNA.	1852
MAGEA 9B	NM 001080790	Homo sapiens MAGE family member A9B (MAGEA9B), mRNA.	1853
MAGT1	NM 032121	Homo sapiens magnesium transporter 1 (MAGT1), mRNA.	1854
MALT1	NM 006785	Homo sapiens MALT1 paracaspase (MALT1), transcript variant 1, mRNA.	1855
MALT1	NM 173844	Homo sapiens MALT1 paracaspase (MALT1), transcript variant 2, mRNA.	1856

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
MAP3K2	NM 006609	Homo sapiens mitogen-activated protein kinase kinase kinase 2 (MAP3K2), mRNA.	1857
MAPK1	NM 002745	Homo sapiens mitogen-activated protein kinase 1 (MAPK1), transcript variant 1, mRNA.	1858
MAPK1	NM 138957	Homo sapiens mitogen-activated protein kinase 1 (MAPK1), transcript variant 2, mRNA.	1859
MAPK3	NM 001040056	Homo sapiens mitogen-activated protein kinase 3 (MAPK3), transcript variant 2, mRNA.	1860
MAPK3	NM 001109891	Homo sapiens mitogen-activated protein kinase 3 (MAPK3), transcript variant 3, mRNA.	1861
MAPK3	NM 002746	Homo sapiens mitogen-activated protein kinase 3 (MAPK3), transcript variant 1, mRNA.	1862
MAVS	NM 020746	Homo sapiens mitochondrial antiviral signaling protein (MAVS), transcript variant 1, mRNA.	1863
MAVS	NM 001206491	Homo sapiens mitochondrial antiviral signaling protein (MAVS), transcript variant 3, mRNA.	1864
MAVS	NR 037921	Homo sapiens mitochondrial antiviral signaling protein (MAVS), transcript variant 2, non-coding RNA.	1865
MECP2	NM 004992	Homo sapiens methyl-CpG binding protein 2 (MECP2), transcript variant 1, mRNA.	1866
MECP2	NM 001110792	Homo sapiens methyl-CpG binding protein 2 (MECP2), transcript variant 2, mRNA.	1867
MEX3C	NM 016626	Homo sapiens mex-3 RNA binding family member C (MEX3C), mRNA.	1868
MRE11A	NM 005590	Homo sapiens MRE11 homolog A, double strand break repair nuclease (MRE11A), transcript variant 2, mRNA.	1869
MRE11A	NM 005591	Homo sapiens MRE11 homolog A, double strand break repair nuclease (MRE11A), transcript variant 1, mRNA.	1870
MS4A1	NM 021950	Homo sapiens membrane spanning 4-domains A1 (MS4A1), transcript variant 3, mRNA.	1871
MS4A1	NM 152866	Homo sapiens membrane spanning 4-domains A1 (MS4A1), transcript variant 1, mRNA.	1872
MSN	NM 002444	Homo sapiens moesin (MSN), mRNA.	1873
MYD88	NM 001172566	Homo sapiens myeloid differentiation primary response 88 (MYD88), transcript variant 5, mRNA.	1874
MYD88	NM 001172567	Homo sapiens myeloid differentiation primary response 88 (MYD88), transcript variant 1, mRNA.	1875
MYD88	NM 001172568	Homo sapiens myeloid differentiation primary response 88 (MYD88), transcript variant 3, mRNA.	1876
MYD88	NM 001172569	Homo sapiens myeloid differentiation primary response 88 (MYD88), transcript variant 4, mRNA.	1877
MYD88	NM 002468	Homo sapiens myeloid differentiation primary response 88 (MYD88), transcript variant 2, mRNA.	1878
NBN	NM 002485	Homo sapiens nibrin (NBN), mRNA.	1879
NFIC	NM 001245005	Homo sapiens nuclear factor I C (NFIC), transcript variant 4, mRNA.	1880
NFIC	NM 205843	Homo sapiens nuclear factor I C (NFIC), transcript variant 2, mRNA.	1881
NFIC	NM 001245002	Homo sapiens nuclear factor I C (NFIC), transcript variant 1, mRNA.	1882
NFIC	NM 001245004	Homo sapiens nuclear factor I C (NFIC), transcript variant 3, mRNA.	1883
NFIC	NM 005597	Homo sapiens nuclear factor I C (NFIC), transcript variant 5, mRNA.	1884
NFKB1	NM 003998	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NFKB 1), transcript variant 1, mRNA.	1885
NFKB1	NM 001165412	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (NFKB 1), transcript variant 2, mRNA.	1886

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
NFKB2	NM 001077494	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (NFKB2), transcript variant 1, mRNA.	1887
NFKB2	NM 002502	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (NFKB2), transcript variant 2, mRNA.	1888
NFKB2	NM 001261403	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 2 (NFKB2), transcript variant 4, mRNA.	1889
NFKBIA	NM 020529	Homo sapiens NFKB inhibitor alpha (NFKBIA), mRNA.	1890
NHEJ1	NM 024782	Homo sapiens non-homologous end joining factor 1 (NHEJ1), mRNA.	1891
NLRP3	NM 183395	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 2, mRNA.	1892
NLRP3	NM 004895	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 1, mRNA.	1893
NLRP3	NM 001127462	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 5, mRNA.	1894
NLRP3	NM 001127461	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 4, mRNA.	1895
NLRP3	NM 001079821	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 3, mRNA.	1896
NLRP3	NM 001243133	Homo sapiens NLR family, pyrin domain containing 3 (NLRP3), transcript variant 6, mRNA.	1897
NOD2	NM 022162	Homo sapiens nucleotide-binding oligomerization domain containing 2 (NOD2), mRNA.	1898
ORAI1	NM 032790	Homo sapiens ORAI calcium release-activated calcium modulator 1 (ORAI1), mRNA.	1899
OSTM1	NM 014028	Homo sapiens osteopetrosis associated transmembrane protein 1 (OSTM1), mRNA.	1900
PGM3	NM 001199917	Homo sapiens phosphoglucomutase 3 (PGM3), transcript variant 1, mRNA.	1901
PGM3	NM 001199918	Homo sapiens phosphoglucomutase 3 (PGM3), transcript variant 3, mRNA.	1902
PGM3	NM 015599	Homo sapiens phosphoglucomutase 3 (PGM3), transcript variant 2, mRNA.	1903
PGM3	NM 001199919	Homo sapiens phosphoglucomutase 3 (PGM3), transcript variant 4, mRNA.	1904
PIAS1	NM 016166	Homo sapiens protein inhibitor of activated STAT 1 (PIAS1), transcript variant 2, mRNA.	1905
PIK3R1	NM 181523	Homo sapiens phosphoinositide-3-kinase regulatory subunit 1 (PIK3R1), transcript variant 1, mRNA.	1906
PIK3R1	NM 181524	Homo sapiens phosphoinositide-3-kinase regulatory subunit 1 (PIK3R1), transcript variant 3, mRNA.	1907
PIK3R1	NM 181504	Homo sapiens phosphoinositide-3-kinase regulatory subunit 1 (PIK3R1), transcript variant 2, mRNA.	1908
PIK3R1	NM 001242466	Homo sapiens phosphoinositide-3-kinase regulatory subunit 1 (PIK3R1), transcript variant 4, mRNA.	1909
PLCG2	NM 002661	Homo sapiens phospholipase C gamma 2 (PLCG2), mRNA.	1910
PMS2	NM 000535	Homo sapiens PMS1 homolog 2, mismatch repair system component (PMS2), transcript variant 1, mRNA.	1911
PNP	NM 000270	Homo sapiens purine nucleoside phosphorylase (PNP), mRNA.	1912
POLA1	NM 016937	Homo sapiens polymerase (DNA directed), alpha 1, catalytic subunit (POLA1), mRNA.	1913
POLE	NM 006231	Homo sapiens DNA polymerase epsilon, catalytic subunit (POLE), mRNA.	1914
PRF1	NM 001083116	Homo sapiens perforin 1 (PRF1), transcript variant 2, mRNA.	1915
PRF1	NM 005041	Homo sapiens perforin 1 (PRF1), transcript variant 1, mRNA.	1916
PRKCD	NM 006254	Homo sapiens protein kinase C delta (PRKCD), transcript variant 1, mRNA.	1917

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
PRKCD	NM 212539	Homo sapiens protein kinase C delta (PRKCD), transcript variant 2, mRNA.	1918
PRKDC	NM 001081640	Homo sapiens protein kinase, DNA-activated, catalytic polypeptide (PRKDC), transcript variant 2, mRNA.	1919
PRKDC	NM 006904	Homo sapiens protein kinase, DNA-activated, catalytic polypeptide (PRKDC), transcript variant 1, mRNA.	1920
PROC	NM 000312	Homo sapiens protein C, inactivator of coagulation factors Va and Villa (PROC), mRNA.	1921
PSMB8	NM 004159	Homo sapiens proteasome (prosome, macropain) subunit, beta type, 8 (PSMB8), transcript variant 1, mRNA.	1922
PSMB8	NM 148919	Homo sapiens proteasome (prosome, macropain) subunit, beta type, 8 (PSMB8), transcript variant 2, mRNA.	1923
PTEN	NM 000314	Homo sapiens phosphatase and tensin homolog (PTEN), transcript variant 1, mRNA.	1924
PTPRC	NM 001267798	Homo sapiens protein tyrosine phosphatase, receptor type C (PTPRC), transcript variant 5, mRNA.	1925
PTPRC	NM 002838	Homo sapiens protein tyrosine phosphatase, receptor type C (PTPRC), transcript variant 1, mRNA.	1926
PTPRC	NM 080921	Homo sapiens protein tyrosine phosphatase, receptor type C (PTPRC), transcript variant 2, mRNA.	1927
PTPRC	NR 052021	Homo sapiens protein tyrosine phosphatase, receptor type C (PTPRC), transcript variant 4, non-coding RNA.	1928
PURA	NM 005859	Homo sapiens purine rich element binding protein A (PURA), mRNA.	1929
RAB27A	NM 183235	Homo sapiens RAB27A, member RAS oncogene family (RAB27A), transcript variant 3, mRNA.	1930
RAB27A	NM 183236	Homo sapiens RAB27A, member RAS oncogene family (RAB27A), transcript variant 4, mRNA.	1931
RAB27A	NM 004580	Homo sapiens RAB27A, member RAS oncogene family (RAB27A), transcript variant 1, mRNA.	1932
RAB27A	NM 183234	Homo sapiens RAB27A, member RAS oncogene family (RAB27A), transcript variant 2, mRNA.	1933
RAB7A	NM 004637	Homo sapiens RAB7A, member RAS oncogene family (RAB7A), mRNA.	1934
RABGEF 1	NM 014504	Homo sapiens RAB guanine nucleotide exchange factor (GEF) 1 (RABGEF1), transcript variant 4, mRNA.	1935
RAC2	NM 002872	Homo sapiens ras-related C3 botulinum toxin substrate 2 (rho family, small GTP binding protein Rac2) (RAC2), mRNA.	1936
RAD51	NM 001164270	Homo sapiens RAD51 recombinase (RAD51), transcript variant 3, mRNA.	1937
RAD51	NM 002875	Homo sapiens RAD51 recombinase (RAD51), transcript variant 1, mRNA.	1938
RAD51	NM 133487	Homo sapiens RAD51 recombinase (RAD51), transcript variant 2, mRNA.	1939
RAD51	NM 001164269	Homo sapiens RAD51 recombinase (RAD51), transcript variant 4, mRNA.	1940
RAG1	NM 000448	Homo sapiens recombination activating gene 1 (RAG1), mRNA.	1941
RAG2	NM 000536	Homo sapiens recombination activating gene 2 (RAG2), transcript variant 1, mRNA.	1942
RAG2	NM 001243785	Homo sapiens recombination activating gene 2 (RAG2), transcript variant 3, mRNA.	1943
RAG2	NM 001243786	Homo sapiens recombination activating gene 2 (RAG2), transcript variant 4, mRNA.	1944
RBCK1	NM 006462	Homo sapiens RANBP2-type and C3HC4-type zinc finger containing 1 (RBCK1), transcript variant 1, mRNA.	1945
RBCK1	NM 031229	Homo sapiens RANBP2-type and C3HC4-type zinc finger containing 1 (RBCK1), transcript variant 2, mRNA.	1946
RFX5	NM 000449	Homo sapiens regulatory factor X5 (RFX5), transcript variant	1947

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		1, mRNA.	
RFX5	NM 001025603	Homo sapiens regulatory factor X5 (RFX5), transcript variant 2, mRNA.	1948
RFXAN K	NM 003721	Homo sapiens regulatory factor X associated ankyrin containing protein (RFXANK), transcript variant 1, mRNA.	1949
RFXAN K	NM 134440	Homo sapiens regulatory factor X associated ankyrin containing protein (RFXANK), transcript variant 2, mRNA.	1950
RFXAP	NM 000538	Homo sapiens regulatory factor X associated protein (RFXAP), mRNA.	1951
RIPK1	NM 003804	Homo sapiens receptor (TNFRSF)-interacting serine-threonine kinase 1 (RIPK1), mRNA.	1952
RIPK3	NM 006871	Homo sapiens receptor-interacting serine-threonine kinase 3 (RIPK3), mRNA.	1953
RMRP	NR 003051	Homo sapiens RNA component of mitochondrial RNA processing endoribonuclease (RMRP), RNase MRP RNA.	1954
RNASE H2A	NM 006397	Homo sapiens ribonuclease H2, subunit A (RNASEH2A), mRNA.	1955
RNASE H2B	NM 001142279	Homo sapiens ribonuclease H2, subunit B (RNASEH2B), transcript variant 2, mRNA.	1956
RNASE H2B	NM 024570	Homo sapiens ribonuclease H2, subunit B (RNASEH2B), transcript variant 1, mRNA.	1957
RNASE H2C	NM 032193	Homo sapiens ribonuclease H2, subunit C (RNASEH2C), mRNA.	1958
RNASEL	NM 021133	Homo sapiens ribonuclease L (2',5'-oligoisoadenylate synthetase-dependent) (RNASEL), mRNA.	1959
RNF168	NM 152617	Homo sapiens ring finger protein 168 (RNF168), mRNA.	1960
RNF31	NM 017999	Homo sapiens ring finger protein 31 (RNF31), mRNA.	1961
RNU4AT AC	NR 023343	Homo sapiens RNA, U4atac small nuclear (U12-dependent splicing) (RNU4ATAC), small nuclear RNA.	1962
RTEL1	NM 016434	Homo sapiens regulator of telomere elongation helicase 1 (RTEL1), transcript variant 1, mRNA.	1963
RTEL1	NM 032957	Homo sapiens regulator of telomere elongation helicase 1 (RTEL1), transcript variant 2, mRNA.	1964
RTEL1-TNFRSF 6B	NR 037882	Homo sapiens RTEL1-TNFRSF6B readthrough (NMD candidate) (RTEL1-TNFRSF6B), long non-coding RNA.	1965
SALL2	NM 005407	Homo sapiens spalt like transcription factor 2 (SALL2), transcript variant 1, mRNA.	1966
SAMHD 1	NM 015474	Homo sapiens SAM domain and HD domain 1 (SAMHD1), mRNA.	1967
SBDS	NM 016038	Homo sapiens Shwachman-Bodian-Diamond syndrome (SBDS), mRNA.	1968
SH2D1A	NM 001114937	Homo sapiens SH2 domain containing 1A (SH2D1A), transcript variant 2, mRNA.	1969
SH2D1A	NM 002351	Homo sapiens SH2 domain containing 1A (SH2D 1A), transcript variant 1, mRNA.	1970
SHARPI N	NM 030974	Homo sapiens SHANK-associated RH domain interactor (SHARPIN), transcript variant 1, mRNA.	1971
SHARPI N	NR 038270	Homo sapiens SHANK-associated RH domain interactor (SHARPIN), transcript variant 2, non-coding RNA.	1972
SKIV2L	NM 006929	Homo sapiens superkiller viralicidic activity 2-like (S. cerevisiae) (SKIV2L), mRNA.	1973
SLC37A 4	NM 001164277	Homo sapiens solute carrier family 37 (glucose-6-phosphate transporter), member 4 (SLC37A4), transcript variant 1, mRNA.	1974
SLC37A 4	NM 001164278	Homo sapiens solute carrier family 37 (glucose-6-phosphate transporter), member 4 (SLC37A4), transcript variant 2, mRNA.	1975
SLC37A 4	NM 001164279	Homo sapiens solute carrier family 37 (glucose-6-phosphate transporter), member 4 (SLC37A4), transcript variant 3,	1976

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		mRNA.	
SLC37A 4	NM 001467	Homo sapiens solute carrier family 37 (glucose-6-phosphate transporter), member 4 (SLC37A4), transcript variant 4, mRNA.	1977
SLC37A 4	NM 001164280	Homo sapiens solute carrier family 37 (glucose-6-phosphate transporter), member 4 (SLC37A4), transcript variant 5, mRNA.	1978
SLC46A 1	NM 001242366	Homo sapiens solute carrier family 46 member 1 (SLC46A1), transcript variant 2, mRNA.	1979
SLC46A 1	NM 080669	Homo sapiens solute carrier family 46 member 1 (SLC46A1), transcript variant 1, mRNA.	1980
SLC8A1	NM 001112800	Homo sapiens solute carrier family 8 member A1 (SLC8A1), transcript variant B, mRNA.	1981
SLC8A1	NM 001112801	Homo sapiens solute carrier family 8 member A1 (SLC8A1), transcript variant C, mRNA.	1982
SLC8A1	NM 001112802	Homo sapiens solute carrier family 8 member A1 (SLC8A1), transcript variant D, mRNA.	1983
SLC8A1	NM 001252624	Homo sapiens solute carrier family 8 member A1 (SLC8A1), transcript variant E, mRNA.	1984
SLC8A1	NM 021097	Homo sapiens solute carrier family 8 member A1 (SLC8A1), transcript variant A, mRNA.	1985
SMAD2	NM 001003652	Homo sapiens SMAD family member 2 (SMAD2), transcript variant 2, mRNA.	1986
SMAD2	NM 001135937	Homo sapiens SMAD family member 2 (SMAD2), transcript variant 3, mRNA.	1987
SMAD2	NM 005901	Homo sapiens SMAD family member 2 (SMAD2), transcript variant 1, mRNA.	1988
SMAD3	NM 005902	Homo sapiens SMAD family member 3 (SMAD3), transcript variant 1, mRNA.	1989
SMAD3	NM 001145102	Homo sapiens SMAD family member 3 (SMAD3), transcript variant 2, mRNA.	1990
SMAD3	NM 001145103	Homo sapiens SMAD family member 3 (SMAD3), transcript variant 3, mRNA.	1991
SMAD3	NM 001145104	Homo sapiens SMAD family member 3 (SMAD3), transcript variant 4, mRNA.	1992
SMAD4	NM 005359	Homo sapiens SMAD family member 4 (SMAD4), mRNA.	1993
SNAP29	NM 004782	Homo sapiens synaptosomal-associated protein, 29kDa (SNAP29), mRNA.	1994
SNAR-A1	NR 004435	Homo sapiens small ILF3/NF90-associated RNA A1 (SNAR-A1), small nuclear RNA.	1995
SNAR-A10	NR 024229	Homo sapiens small ILF3/NF90-associated RNA A10 (SNAR-A10), small nuclear RNA.	1996
SNAR-A11	NR 024225	Homo sapiens small ILF3/NF90-associated RNA A11 (SNAR-A11), small nuclear RNA.	1997
SNAR-A12	NR 004437	Homo sapiens small ILF3/NF90-associated RNA A12 (SNAR-A12), small nuclear RNA.	1998
SNAR-A13	NR 024216	Homo sapiens small ILF3/NF90-associated RNA A13 (SNAR-A13), small nuclear RNA.	1999
SNAR-A14	NR 024242	Homo sapiens small ILF3/NF90-associated RNA A14 (SNAR-A14), small nuclear RNA.	2000
SNAR-A2	NR 004436	Homo sapiens small ILF3/NF90-associated RNA A2 (SNAR-A2), small nuclear RNA.	2001
SNAR-A3	NR 024214	Homo sapiens small ILF3/NF90-associated RNA A3 (SNAR-A3), small nuclear RNA.	2002
SNAR-A4	NR 024215	Homo sapiens small ILF3/NF90-associated RNA A4 (SNAR-A4), small nuclear RNA.	2003
SNAR-A5	NR 024223	Homo sapiens small ILF3/NF90-associated RNA A5 (SNAR-A5), small nuclear RNA.	2004
SNAR-A6	NR 024227	Homo sapiens small ILF3/NF90-associated RNA A6 (SNAR-	2005

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		A6), small nuclear RNA.	
SNAR-A7	NR 024224	Homo sapiens small ILF3/NF90-associated RNA A7 (SNAR-A7), small nuclear RNA.	2006
SNAR-A8	NR 024228	Homo sapiens small ILF3/NF90-associated RNA A8 (SNAR-A8), small nuclear RNA.	2007
SNAR-A9	NR 024226	Homo sapiens small ILF3/NF90-associated RNA A9 (SNAR-A9), small nuclear RNA.	2008
SNAR-B1	NR 024231	Homo sapiens small ILF3/NF90-associated RNA B1 (SNAR-B1), small nuclear RNA.	2009
SNAR-B2	NR 024230	Homo sapiens small ILF3/NF90-associated RNA B2 (SNAR-B2), small nuclear RNA.	2010
SNAR-C1	NR 024220	Homo sapiens small ILF3/NF90-associated RNA C1 (SNAR-C1), small nuclear RNA.	2011
SNAR-C2	NR 024217	Homo sapiens small ILF3/NF90-associated RNA C2 (SNAR-C2), small nuclear RNA.	2012
SNAR-C3	NR 024221	Homo sapiens small ILF3/NF90-associated RNA C3 (SNAR-C3), small nuclear RNA.	2013
SNAR-C4	NR 024218	Homo sapiens small ILF3/NF90-associated RNA C4 (SNAR-C4), small nuclear RNA.	2014
SNAR-C5	NR 024219	Homo sapiens small ILF3/NF90-associated RNA C5 (SNAR-C5), small nuclear RNA.	2015
SNAR-D	NR 024243	Homo sapiens small ILF3/NF90-associated RNA D (SNAR-D), small nuclear RNA.	2016
SNAR-E	NR 024258	Homo sapiens small ILF3/NF90-associated RNA E (SNAR-E), small nuclear RNA.	2017
SNAR-F	NR 004384	Homo sapiens small ILF3/NF90-associated RNA F (SNAR-F), small nuclear RNA.	2018
SNAR-G1	NR 004383	Homo sapiens small ILF3/NF90-associated RNA G1 (SNAR-G1), small nuclear RNA.	2019
SNAR-G2	NR 024244	Homo sapiens small ILF3/NF90-associated RNA G2 (SNAR-G2), small nuclear RNA.	2020
SNAR-H	NR 024342	Homo sapiens small ILF3/NF90-associated RNA H (SNAR-H), small nuclear RNA.	2021
SNAR-I	NR 024343	Homo sapiens small ILF3/NF90-associated RNA I (SNAR-I), small nuclear RNA.	2022
SNCA	NM 000345	Homo sapiens synuclein, alpha (non A4 component of amyloid precursor) (SNCA), transcript variant 1, mRNA.	2023
SNCA	NM 001146054	Homo sapiens synuclein, alpha (non A4 component of amyloid precursor) (SNCA), transcript variant 2, mRNA.	2024
SNCA	NM 001146055	Homo sapiens synuclein, alpha (non A4 component of amyloid precursor) (SNCA), transcript variant 3, mRNA.	2025
SNCA	NM 007308	Homo sapiens synuclein, alpha (non A4 component of amyloid precursor) (SNCA), transcript variant 4, mRNA.	2026
SNX 10	NM 013322	Homo sapiens sorting nexin 10 (SNX10), transcript variant 2, mRNA.	2027
SNX 10	NM 001199835	Homo sapiens sorting nexin 10 (SNX10), transcript variant 1, mRNA.	2028
SNX10	NM 001199837	Homo sapiens sorting nexin 10 (SNX10), transcript variant 3, mRNA.	2029
SNX 10	NM 001199838	Homo sapiens sorting nexin 10 (SNX10), transcript variant 4, mRNA.	2030
SNX10	NR 037670	Homo sapiens sorting nexin 10 (SNX10), transcript variant 5, non-coding RNA.	2031
SP110	NM 004509	Homo sapiens SP110 nuclear body protein (SP110), transcript variant a, mRNA.	2032
SP110	NM 080424	Homo sapiens SP110 nuclear body protein (SP110), transcript variant c, mRNA.	2033
SP110	NM 001185015	Homo sapiens SP110 nuclear body protein (SP110), transcript variant d, mRNA.	2034

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
SP110	NM 004510	Homo sapiens SP110 nuclear body protein (SP110), transcript variant b, mRNA.	2035
SP140	NM 001005176	Homo sapiens SP140 nuclear body protein (SP140), transcript variant 2, mRNA.	2036
SP140	NM 007237	Homo sapiens SP140 nuclear body protein (SP140), transcript variant 1, mRNA.	2037
SPINK5	NM 001127698	Homo sapiens serine peptidase inhibitor, Kazal type 5 (SPINK5), transcript variant 1, mRNA.	2038
SPINK5	NM 006846	Homo sapiens serine peptidase inhibitor, Kazal type 5 (SPINK5), transcript variant 2, mRNA.	2039
SPINK5	NM 001127699	Homo sapiens serine peptidase inhibitor, Kazal type 5 (SPINK5), transcript variant 3, mRNA.	2040
SQSTM1	NM 003900	Homo sapiens sequestosome 1 (SQSTM1), transcript variant 1, mRNA.	2041
SQSTM1	NM 001142298	Homo sapiens sequestosome 1 (SQSTM1), transcript variant 2, mRNA.	2042
SQSTM1	NM 001142299	Homo sapiens sequestosome 1 (SQSTM1), transcript variant 3, mRNA.	2043
SRSF1	NM 001078166	Homo sapiens serine and arginine rich splicing factor 1 (SRSF1), transcript variant 2, mRNA.	2044
SRSF1	NM 006924	Homo sapiens serine and arginine rich splicing factor 1 (SRSF1), transcript variant 1, mRNA.	2045
SRSF1	NR 034041	Homo sapiens serine and arginine rich splicing factor 1 (SRSF1), transcript variant 3, non-coding RNA.	2046
STAT1	NM 007315	Homo sapiens signal transducer and activator of transcription 1 (STAT1), transcript variant alpha, mRNA.	2047
STAT1	NM 139266	Homo sapiens signal transducer and activator of transcription 1 (STAT1), transcript variant beta, mRNA.	2048
STAT2	NM 005419	Homo sapiens signal transducer and activator of transcription 2, 113kDa (STAT2), transcript variant 1, mRNA.	2049
STAT2	NM 198332	Homo sapiens signal transducer and activator of transcription 2, 113kDa (STAT2), transcript variant 2, mRNA.	2050
STAT3	NM 003150	Homo sapiens signal transducer and activator of transcription 3 (STAT3), transcript variant 2, mRNA.	2051
STAT3	NM 139276	Homo sapiens signal transducer and activator of transcription 3 (STAT3), transcript variant 1, mRNA.	2052
STAT3	NM 213662	Homo sapiens signal transducer and activator of transcription 3 (STAT3), transcript variant 3, mRNA.	2053
STAT5B	NM 012448	Homo sapiens signal transducer and activator of transcription 5B (STAT5B), mRNA.	2054
STIM1	NM 003156	Homo sapiens stromal interaction molecule 1 (STIM1), transcript variant 2, mRNA.	2055
STK4	NM 006282	Homo sapiens serine/threonine kinase 4 (STK4), mRNA.	2056
STX11	NM 003764	Homo sapiens syntaxin 11 (STX11), mRNA.	2057
STXBP2	NM 001127396	Homo sapiens syntaxin binding protein 2 (STXBP2), transcript variant 2, mRNA.	2058
STXBP2	NM 001272034	Homo sapiens syntaxin binding protein 2 (STXBP2), transcript variant 3, mRNA.	2059
STXBP2	NM 006949	Homo sapiens syntaxin binding protein 2 (STXBP2), transcript variant 1, mRNA.	2060
STXBP2	NR 073560	Homo sapiens syntaxin binding protein 2 (STXBP2), transcript variant 4, non-coding RNA.	2061
SYNCRI P	NM 001159673	Homo sapiens synaptotagmin binding cytoplasmic RNA interacting protein (SYNCRIP), transcript variant 2, mRNA.	2062
SYNCRI P	NM 001159674	Homo sapiens synaptotagmin binding cytoplasmic RNA interacting protein (SYNCRIP), transcript variant 3, mRNA.	2063
SYNCRI P	NM 001159676	Homo sapiens synaptotagmin binding cytoplasmic RNA interacting protein (SYNCRIP), transcript variant 5, mRNA.	2064
SYNCRI P	NM 001159677	Homo sapiens synaptotagmin binding cytoplasmic RNA	2065

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		interacting protein (SYNCRIP), transcript variant 6, mRNA.	
SYNCRI P	NM 001253771	Homo sapiens synaptotagmin binding cytoplasmic RNA interacting protein (SYNCRIP), transcript variant 7, mRNA.	2066
SYNCRI P	NM 001159675	Homo sapiens synaptotagmin binding cytoplasmic RNA interacting protein (SYNCRIP), transcript variant 4, mRNA.	2067
SYNCRI P	NM 006372	Homo sapiens synaptotagmin binding cytoplasmic RNA interacting protein (SYNCRIP), transcript variant 1, mRNA.	2068
T	NM 001270484	Homo sapiens T brachyury transcription factor (T), transcript variant 2, mRNA.	2069
T	NM 003181	Homo sapiens T brachyury transcription factor (T), transcript variant 1, mRNA.	2070
TAP1	NM 000593	Homo sapiens transporter 1, ATP binding cassette subfamily B member (TAP1), transcript variant 1, mRNA.	2071
TAP2	NM 018833	Homo sapiens transporter 2, ATP binding cassette subfamily B member (TAP2), transcript variant 2, mRNA.	2072
TAP2	NM 000544	Homo sapiens transporter 2, ATP binding cassette subfamily B member (TAP2), transcript variant 1, B allele, mRNA.	2073
TAPBP	NM 003190	Homo sapiens TAP binding protein (tapasin) (TAPBP), transcript variant 1, mRNA.	2074
TAPBP	NM 172209	Homo sapiens TAP binding protein (tapasin) (TAPBP), transcript variant 3, mRNA.	2075
TAPBP	NM 172208	Homo sapiens TAP binding protein (tapasin) (TAPBP), transcript variant 2, mRNA.	2076
TAZ	NM 000116	Homo sapiens tafazzin (TAZ), transcript variant 1, mRNA.	2077
TAZ	NM 181312	Homo sapiens tafazzin (TAZ), transcript variant 3, mRNA.	2078
TAZ	NM 181311	Homo sapiens tafazzin (TAZ), transcript variant 2, mRNA.	2079
TAZ	NM 181313	Homo sapiens tafazzin (TAZ), transcript variant 4, mRNA.	2080
TAZ	NR 024048	Homo sapiens tafazzin (TAZ), transcript variant 5, non-coding RNA.	2081
TBK1	NM 013254	Homo sapiens TANK binding kinase 1 (TBK1), mRNA.	2082
TBX1	NM 005992	Homo sapiens T-box 1 (TBX1), transcript variant B, mRNA.	2083
TBX1	NM 080646	Homo sapiens T-box 1 (TBX1), transcript variant A, mRNA.	2084
TBX1	NM 080647	Homo sapiens T-box 1 (TBX1), transcript variant C, mRNA.	2085
TCIRG1	NM 006019	Homo sapiens T-cell immune regulator 1, ATPase H+ transporting V0 subunit a3 (TCIRG1), transcript variant 1, mRNA.	2086
TCIRG1	NM 006053	Homo sapiens T-cell immune regulator 1, ATPase H+ transporting V0 subunit a3 (TCIRG1), transcript variant 2, mRNA.	2087
TICAM1	NM 182919	Homo sapiens toll like receptor adaptor molecule 1 (TICAM1), mRNA.	2088
TLR3	NM 003265	Homo sapiens toll like receptor 3 (TLR3), mRNA.	2089
TLR4	NM 003266	Homo sapiens toll like receptor 4 (TLR4), transcript variant 3, mRNA.	2090
TLR4	NM 138554	Homo sapiens toll like receptor 4 (TLR4), transcript variant 1, mRNA.	2091
TLR4	NM 138557	Homo sapiens toll like receptor 4 (TLR4), transcript variant 4, mRNA.	2092
TMEM 1 73	NM 198282	Homo sapiens transmembrane protein 173 (TMEM173), mRNA.	2093
TNF	NM 000594	Homo sapiens tumor necrosis factor (TNF), mRNA.	2094
TNFAIP 3	NM 001270507	Homo sapiens TNF alpha induced protein 3 (TNFAIP3), transcript variant 2, mRNA.	2095
TNFAIP 3	NM 001270508	Homo sapiens TNF alpha induced protein 3 (TNFAIP3), transcript variant 1, mRNA.	2096
TNFAIP 3	NM 006290	Homo sapiens TNF alpha induced protein 3 (TNFAIP3), transcript variant 3, mRNA.	2097

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
TNFRSF 11A	NM 003839	Homo sapiens tumor necrosis factor receptor superfamily, member 11a, NFKB activator (TNFRSF11A), transcript variant 1, mRNA.	2098
TNFRSF 11A	NM 001270949	Homo sapiens tumor necrosis factor receptor superfamily, member 11a, NFKB activator (TNFRSF11A), transcript variant 2, mRNA.	2099
TNFRSF 11A	NM 001270950	Homo sapiens tumor necrosis factor receptor superfamily, member 11a, NFKB activator (TNFRSF11A), transcript variant 3, mRNA.	2100
TNFRSF 11A	NM 001270951	Homo sapiens tumor necrosis factor receptor superfamily, member 11a, NFKB activator (TNFRSF11A), transcript variant 4, mRNA.	2101
TNFRSF 11B	NM 002546	Homo sapiens tumor necrosis factor receptor superfamily, member 11b (TNFRSF11B), mRNA.	2102
TNFRSF 13B	NM 012452	Homo sapiens TNF receptor superfamily member 13B (TNFRSF13B), mRNA.	2103
TNFRSF 4	NM 003327	Homo sapiens TNF receptor superfamily member 4 (TNFRSF4), mRNA.	2104
TNFRSF 8	NM 001243	Homo sapiens TNF receptor superfamily member 8 (TNFRSF8), transcript variant 1, mRNA.	2105
TNFSF1 1	NM 003701	Homo sapiens tumor necrosis factor (ligand) superfamily, member 11 (TNFSF11), transcript variant 1, mRNA.	2106
TNFSF1 1	NM 033012	Homo sapiens tumor necrosis factor (ligand) superfamily, member 11 (TNFSF11), transcript variant 2, mRNA.	2107
TNFSF1 2	NM 003809	Homo sapiens tumor necrosis factor superfamily member 12 (TNFSF12), transcript variant 1, mRNA.	2108
TNFSF1 2	NR 037146	Homo sapiens tumor necrosis factor superfamily member 12 (TNFSF12), transcript variant 2, non-coding RNA.	2109
TP53	NM 000546	Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA.	2110
TP53	NM 001126112	Homo sapiens tumor protein p53 (TP53), transcript variant 2, mRNA.	2111
TP53	NM 001126113	Homo sapiens tumor protein p53 (TP53), transcript variant 4, mRNA.	2112
TP53	NM 001126114	Homo sapiens tumor protein p53 (TP53), transcript variant 3, mRNA.	2113
TP53	NM 001126115	Homo sapiens tumor protein p53 (TP53), transcript variant 5, mRNA.	2114
TP53	NM 001126116	Homo sapiens tumor protein p53 (TP53), transcript variant 6, mRNA.	2115
TP53	NM 001126117	Homo sapiens tumor protein p53 (TP53), transcript variant 7, mRNA.	2116
TP53	NM 001126118	Homo sapiens tumor protein p53 (TP53), transcript variant 8, mRNA.	2117
TP53	NM 001276695	Homo sapiens tumor protein p53 (TP53), transcript variant 4, mRNA.	2118
TP53	NM 001276696	Homo sapiens tumor protein p53 (TP53), transcript variant 3, mRNA.	2119
TP53	NM 001276697	Homo sapiens tumor protein p53 (TP53), transcript variant 5, mRNA.	2120
TP53	NM 001276698	Homo sapiens tumor protein p53 (TP53), transcript variant 6, mRNA.	2121
TP53	NM 001276699	Homo sapiens tumor protein p53 (TP53), transcript variant 7, mRNA.	2122
TP53	NM 001276760	Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA.	2123
TP53	NM 001276761	Homo sapiens tumor protein p53 (TP53), transcript variant 2, mRNA.	2124
TRAF3	NM 001199427	Homo sapiens TNF receptor associated factor 3 (TRAF3), transcript variant 4, mRNA.	2125

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
TRAF3	NM 003300	Homo sapiens TNF receptor associated factor 3 (TRAF3), transcript variant 3, mRNA.	2126
TRAF3	NM 145725	Homo sapiens TNF receptor associated factor 3 (TRAF3), transcript variant 1, mRNA.	2127
TRAF3	NM 145726	Homo sapiens TNF receptor associated factor 3 (TRAF3), transcript variant 2, mRNA.	2128
TRAF6	NM 004620	Homo sapiens TNF receptor-associated factor 6, E3 ubiquitin protein ligase (TRAF6), transcript variant 2, mRNA.	2129
TRAF6	NM 145803	Homo sapiens TNF receptor-associated factor 6, E3 ubiquitin protein ligase (TRAF6), transcript variant 1, mRNA.	2130
TREX1	NM 007248	Homo sapiens three prime repair exonuclease 1 (TREX1), transcript variant 5, mRNA.	2131
TREX1	NM 033629	Homo sapiens three prime repair exonuclease 1 (TREX1), transcript variant 4, mRNA.	2132
TREX1	NM 016381	Homo sapiens three prime repair exonuclease 1 (TREX1), transcript variant 1, mRNA.	2133
TRNT1	NM 182916	Homo sapiens tRNA nucleotidyl transferase 1 (TRNT1), transcript variant 1, mRNA.	2134
TTC7A	NM 020458	Homo sapiens tetratricopeptide repeat domain 7A (TTC7A), transcript variant 2, mRNA.	2135
TYK2	NM 003331	Homo sapiens tyrosine kinase 2 (TYK2), mRNA.	2136
UNC119	NM 005148	Homo sapiens unc-119 lipid binding chaperone (UNC119), transcript variant 1, mRNA.	2137
UNC119	NM 054035	Homo sapiens unc-119 lipid binding chaperone (UNC119), transcript variant 2, mRNA.	2138
UNC13D	NM 199242	Homo sapiens unc-13 homolog D (UNC13D), mRNA.	2139
UNC93B 1	NM 030930	Homo sapiens unc-93 homolog B1 (C. elegans) (UNC93B1), mRNA.	2140
UNG	NM 080911	Homo sapiens uracil DNA glycosylase (UNG), transcript variant 2, mRNA.	2141
UNG	NM 003362	Homo sapiens uracil DNA glycosylase (UNG), transcript variant 1, mRNA.	2142
USP18	NM 017414	Homo sapiens ubiquitin specific peptidase 18 (USP18), mRNA.	2143
USP20	NM 006676	Homo sapiens ubiquitin specific peptidase 20 (USP20), transcript variant 1, mRNA.	2144
USP20	NM 001008563	Homo sapiens ubiquitin specific peptidase 20 (USP20), transcript variant 2, mRNA.	2145
USP20	NM 001110303	Homo sapiens ubiquitin specific peptidase 20 (USP20), transcript variant 3, mRNA.	2146
VAPA	NM 003574	Homo sapiens VAMP associated protein A (VAPA), transcript variant 1, mRNA.	2147
VAPA	NM 194434	Homo sapiens VAMP associated protein A (VAPA), transcript variant 2, mRNA.	2148
VCP	NM 007126	Homo sapiens valosin containing protein (VCP), mRNA.	2149
VDAC1	NM 003374	Homo sapiens voltage dependent anion channel 1 (VDAC1), transcript variant 1, mRNA.	2150
VDAC1	NR 036624	Homo sapiens voltage dependent anion channel 1 (VDAC1), transcript variant 3, non-coding RNA.	2151
VDAC1	NR 036625	Homo sapiens voltage dependent anion channel 1 (VDAC1), transcript variant 2, non-coding RNA.	2152
VPS13B	NM 017890	Homo sapiens vacuolar protein sorting 13 homolog B (yeast) (VPS13B), transcript variant 5, mRNA.	2153
VPS13B	NM 181661	Homo sapiens vacuolar protein sorting 13 homolog B (yeast) (VPS13B), transcript variant 4, mRNA.	2154
VPS13B	NM 015243	Homo sapiens vacuolar protein sorting 13 homolog B (yeast) (VPS13B), transcript variant 3, mRNA.	2155
VPS13B	NR 047582	Homo sapiens vacuolar protein sorting 13 homolog B (yeast) (VPS 13B), transcript variant 6, non-coding RNA.	2156

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
VPS13B	NM 152564	Homo sapiens vacuolar protein sorting 13 homolog B (yeast) (VPS13B), transcript variant 1, mRNA.	2157
VPS45	NM 007259	Homo sapiens vacuolar protein sorting 45 homolog (VPS45), transcript variant 1, mRNA.	2158
WAS	NM 000377	Homo sapiens Wiskott-Aldrich syndrome (WAS), mRNA.	2159
WEE1	NM 003390	Homo sapiens WEE1 G2 checkpoint kinase (WEE1), transcript variant 1, mRNA.	2160
WEE1	NM 001143976	Homo sapiens WEE1 G2 checkpoint kinase (WEE1), transcript variant 2, mRNA.	2161
WIPF1	NM 001077269	Homo sapiens WAS/WASL interacting protein family member 1 (WIPF1), transcript variant 2, mRNA.	2162
WIPF1	NM 003387	Homo sapiens WAS/WASL interacting protein family member 1 (WIPF1), transcript variant 1, mRNA.	2163
XIAP	NM 001204401	Homo sapiens X-linked inhibitor of apoptosis, E3 ubiquitin protein ligase (XIAP), transcript variant 2, mRNA.	2164
XIAP	NM 001167	Homo sapiens X-linked inhibitor of apoptosis, E3 ubiquitin protein ligase (XIAP), transcript variant 1, mRNA.	2165
XIAP	NR 037916	Homo sapiens X-linked inhibitor of apoptosis, E3 ubiquitin protein ligase (XIAP), transcript variant 3, non-coding RNA.	2166
YBX1	NM 004559	Homo sapiens Y-box binding protein 1 (YBX1), transcript variant 1, mRNA.	2167
YWHAZ	NM 001135699	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta (YWHAZ), transcript variant 3, mRNA.	2168
YWHAZ	NM 001135700	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta (YWHAZ), transcript variant 4, mRNA.	2169
YWHAZ	NM 001135701	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta (YWHAZ), transcript variant 5, mRNA.	2170
YWHAZ	NM 001135702	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta (YWHAZ), transcript variant 6, mRNA.	2171
YWHAZ	NM 003406	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta (YWHAZ), transcript variant 1, mRNA.	2172
YWHAZ	NM 145690	Homo sapiens tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein zeta (YWHAZ), transcript variant 2, mRNA.	2173
ZAP70	NM 001079	Homo sapiens zeta chain of T cell receptor associated protein kinase 70 (ZAP70), transcript variant 1, mRNA.	2174
ZAP70	NM 207519	Homo sapiens zeta chain of T cell receptor associated protein kinase 70 (ZAP70), transcript variant 2, mRNA.	2175
ZBTB24	NM 014797	Homo sapiens zinc finger and BTB domain containing 24 (ZBTB24), transcript variant 1, mRNA.	2176
ZBTB24	NM 001164313	Homo sapiens zinc finger and BTB domain containing 24 (ZBTB24), transcript variant 2, mRNA.	2177

[0333] Table 12 lists all transcript variants for genes in **Table 6** that were not 'discovered' by PBio on the basis of aCGH (CNV identified genes). The SEQ ID NOs correspond to transcript variants (oftentimes more than one per gene).

Table 13: Genes for which the total burden of heterozygous, damaging variants was found to be statistically greater in PML cases versus ExAC controls

GENE	Ave CASES	Ave EXAC CASES	Ave EXAC SAMPLES	Ave FET	Ave FET corr (419)	Ave OR	Ethnicity	Overlap
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	

GENE	Ave CASES	Ave EXAC CASES	Ave EXAC SAMPLES	Ave FET	Ave FET corr (419)	Ave OR	Ethnicity	Overlap
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	
JAGN 1	5	163	27,768	6.80E-06	2.96E-03	21.71	EUR	
TAP 1	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	
IL 12B	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	

[0334] Table 13 lists genes for which the total burden of heterozygous, damaging variants was found to be statistically greater in PML cases versus ExAC controls. Gene burden analysis was performed as described below at minor allele frequency (MAF) cutoffs of 0.01, 0.02, 0.03, 0.04 and 0.05. Not all genes survived statistical analysis at all MAF cutoffs. For each gene that survived at multiple MAF cutoffs, the averages of the Fisher's Exact Test (FET), nominal and corrected, were calculated, as were the other relevant metrics. Two genes overlapped between AFR and EUR analyses. FETs were corrected for multiple testing with the number of genes used in this study (419). Only genes for which FET corr was < 0.05 and in which variants affected > 10% of cases within the given ethnicity (>2 for AFR, >4 for EUR) were considered for inclusion.

Table 14: Top tier of variants found to be significant on the basis of variant burden analysis

Gene	Variant (hg19)	Geno type	PML EUR 44	PML AFR 21	PML LAT 5	ExA C EUR	ExA C AFR	ExA C LAT	PML EUR OR	PML EUR FET	PML AFR OR	PML AFR FET	PML ALL OR	PML ALL FET
PLC G2	chr16:81 942175, A>G	het	2	5	0	512/3 2281	88/47 07	116/5 548	2.95	0.1547 55	16.40	0.0000	6.49	0.0002
IFIH 1	chr2:163 136505, C>G	het	6	1	0	611/3 3155	23/51 82	119/5 671	8.41	0.0001 56	11.22	0.0927	6.38	0.0002
TCI RG1	chr11:67 818269, G>A	het	0	4	0	103/3 3193	200/5 170	60/57 70	NA	NA	5.85	0.0082	7.31	0.0028
IGL L1	chr22:23 917192, G>T	het	4	3	1	751/3 3348	603/5 183	236/5 782	4.34	0.0172 18	1.27	0.7286	3.47	0.0036
MA VS	chr20:38 46397, C>T	hom	4	4	0	800/3 2122	684/4 982	52/56 91	3.92	0.0238 68	1.48	0.5191	3.47	0.0036
SHA RPI N	chr8:145 154222, G>A	het	8	4	0	2916/ 33177	59/48 65	171/5 780	2.31	0.0535 26	19.17	0.0001	2.68	0.0040
CHD 7	chr8:616 54298, T>A	het	5	0	0	1103/ 33106	39/48 40	93/57 25	3.72	0.0152 68	NA	NA	2.64	0.0485
CX3 CR1	chr3:393 23163,	hom	11	4	0	4723/ 31219	193/4 376	1357/ 5491	1.87	0.0880 87	5.10	0.0128	1.51	0.1806

Gene	Variant (hg19)	Geno type	PML EUR 44	PML AFR 21	PML LAT 5	ExA C EUR	ExA C AFR	ExA C LAT	PML EUR OR	PML EUR FET	PML AFR OR	PML AFR FET	PML ALL OR	PML ALL FET
	A>C													
LRBA	chr4:151199080, 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	chr1:42047208, 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	chr2:163124051, C>T	hom	20	3	1	12107/33356	184/5199	1076/5776	1.46	0.212471	4.54	0.0374	1.21	0.4372
RNASEL	chr1:182554557, C>T	hom	7	2	0	4543/33356	78/5202	167/5785	1.20	0.658473	6.91	0.0403	1.22	0.5616

[0335] Table 14 lists the top tier of variants that were found to be significant on the basis of variant burden analysis, as described below. For each variant (genome coordinates are based on UCSC hg19), detailed information is presented of the numbers of EUR and AFR cases that carry the variant, along with the ethnic-specific and aggregate statistical metrics.

Table 15: Second tier of variants found on the basis of variant burden analysis

Gene	Variant (hgl9)	Geno type	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	chr8:145154824, A>C	het	3	0	0	2/30,670	0/4,471	0/5,302	1122.00	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	1268.41	0.0000
IGLL1	chr22:23915745, 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	chr6:838841 61, C>G	het	0	2	0	0/33,069	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	chr5:138856923, C>T	het	2	2	0	108/32,327	204/4,842	58/5,770	14.21	0.009863	2.39	0.2226	6.97	0.0033
CLCN7	chr16:1510535, C>T	het	0	2	0	1/32,898	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	803/33,206	167/5,171	46/5,779	4.04	0.021706	3.15	0.1480	3.98	0.0056
ORAI1	chr12:122064788, G>GT	het	4	0	0	371/28,708	5/3,555	16/5,354	7.64	0.002562	NA	NA	5.76	0.0064
RBFOX1	chr16:7714909, C>T	het	0	2	0	1/33,367	69/4,902	4/5,782	NA	NA	7.37	0.0361	17.48	0.0066
MALT1	chr18:56401523, C>T	het	4	0	0	466/33,239	9/5,179	40/5,760	7.03	0.003411	NA	NA	5.14	0.0093
GFI1	chr1:92946625, G>C	het	2	1	0	206/29,111	6/4,156	39/5,114	6.68	0.039391	34.58	0.0347	6.80	0.0113
DOCK2	chr5:169081453, G>C	het	0	2	0	48/33.350	43/5,201	27/5,786	NA	NA	12.63	0.0137	11.02	0.0155

Gene	Variant (hgl9)	Geno type	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
ATM	chr11:108 117787, C>T	het	2	0	0	93/33, 256	2/5,15 1	28/5,7 56	16.98	0.00704 7	NA	NA	10.53	0.0169
SNA P29	chr22:212 35389, A>G	het	3	0	0	283/3 2,917	21/5,1 49	32/5,7 40	8.44	0.00658 4	NA	NA	5.79	0.0171
TICA M1	chr19:481 7657, C>T	het	0	2	0	32/31, 437	71/4,8 14	19/5,6 87	NA	NA	7.03	0.0392	10.08	0.0183
GTP BP4	chr10:106 0218, G>A	hom	3	0	0	334/3 3,367	20/5,2 02	21/5,7 86	7.24	0.00992 5	NA	NA	5.25	0.0220
BAC H1	chr21:306 98953, T>G	het	2	0	0	134/3 3,122	4/5,10 3	6/5,77 8	11.72	0.01411 0	NA	NA	8.96	0.0227
DOC K8	chr9:3046 28, G>A	het	2	0	0	149/3 3,298	2/5,16 1	5/5,76 2	10.59	0.01702 0	NA	NA	8.31	0.0261
STX BP2	chr19:771 2287, G>C	het	2	0	0	161/3 2,104	4/4,62 6	11/5,6 86	9.45	0.02102 8	NA	NA	7.06	0.0350
FAS	chr10:907 71767, G>A	het	2	0	0	175/3 3,304	3/5,18 2	10/5,7 31	9.01	0.02290 2	NA	NA	6.89	0.0365
GOL GB1	chr3:1214 15370, T>C	het	3	2	0	1,111/ 33,349	26/5,1 79	84/5,7 79	2.12	0.18074 3	20.86	0.0055	2.71	0.0443
FUK	chr16:705 03095, A>G	het	4	0	0	741/3 3,341	23/4,8 99	73/5,7 87	4.40	0.01648 8	NA	NA	3.13	0.0449
IL 10	chr1:2069 45738, C>T	het	2	0	0	206/3 3,343	2/5,19 8	6/5,78 7	7.66	0.03078 7	NA	NA	6.06	0.0458
ITK	chr5:1565 93120, C>T	het	2	0	0	206/3 3,353	5/5,20 3	5/5,78 9	7.66	0.03077 0	NA	NA	6.01	0.0466
STIM 2	chr4:2701 9452, C>T	het	2	0	0	219/3 3,369	5/5,20 2	9/5,78 9	7.21	0.03434 6	NA	NA	5.57	0.0532
ASH1 L	chr1:1553 17682, C>T	het	2	0	0	218/3 3,367	2/5,20 3	17/5,7 89	7.24	0.03406 7	NA	NA	5.48	0.0548
TBC1 D16	chr17:779 26526, C>T	het	3	0	0	496/3 1,905	15/4,8 45	21/5,7 07	4.63	0.03126 9	NA	NA	3.53	0.0584
LYS1	chr1:2358 40495, G>T	het	3	0	0	517/3 3,239	10/5,1 56	35/5,7 56	4.63	0.03129 9	NA	NA	3.47	0.0606
SALL 2	chr14:219 93359, G>A	het	3	0	0	519/3 1,729	14/4,5 20	17/5,7 18	4.40	0.03552 8	NA	NA	3.37	0.0650
CHD 7	chr8:6175 7805, C>T	het	3	0	0	517/3 2,880	14/4,8 72	41/5,7 65	4.58	0.03216 9	NA	NA	3.36	0.0654
BLM	chr15:913 06241, G>A	het	2	0	0	266/3 3,277	10/5,0 61	17/5,7 56	5.91	0.04887 5	NA	NA	4.40	0.0799
NOD 2	chr16:507 41791, C>T	het	0	2	0	285/3 3,369	21/5,2 03	2/5,78 9	NA	NA	25.97	0.0037	4.21	0.0860
IGLL 1	chr22:239 15583, T>C	het	2	0	0	265/3 3,334	21/5,1 83	26/5,7 87	5.94	0.04840 3	NA	NA	4.15	0.0881

Gene	Variant (hgl9)	Geno type	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	2.38	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	40/5,425	NA	NA	7.80	0.0327	3.64	0.1093
CR2	chr1:207641950, C>T	het	0	2	0	391/33,363	19/5,203	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	209/5,788	1.35	0.494339	10.21	0.0202	1.64	0.2458
ITSN2	chr2:24431184, C>T	hom	3	2	0	2,019/33,339	17/5,186	55/5,784	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	0.753875	30.59	0.0028	1.50	0.3937
DOCK8	chr9:312134, G>A	het	3	2	0	2,114/33,251	79/5,180	161/5,768	1.08	0.757661	6.80	0.0415	1.37	0.4238
VPS13B	chr8:100205255, G>A	het	0	2	0	811/33,345	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	0.6698

[0336] Table 15 lists the second tier of variants that were found on the basis of variant burden analysis, as described below. For each variant (genome coordinates are UCSC hg19), detailed information is presented of the numbers of EUR and AFR cases that carry the variant, along with the ethnic-specific and aggregate statistical metrics.

Table 16: Potential testing scenario, based on top variant burden hits

Gene/Variant	Cases solved	Proportion of Cohort (n = 70)	Test Method	Patient information		
				Primary disease	Ethnicity	Gender
All 4 SNVs	28	40%	genotyping	M, H, O	A, E	both
SHARPIN, IFIH1, PLCG2 SNVs	24	34%	genotyping	M, H, O	A, E	both
IFIH1, PLCG2 SNVs	13	19%	genotyping	M, H, O	A, E	both
SHARPIN SNV	13	19%	genotyping	M, H	A, E	both
IFIH1 SNV	7	10%	genotyping	M, H, O	A, E	both
PLCG2 SNV	7	10%	genotyping	M, H	A, E	both
CHD7 SNV	5	7%	genotyping	M, H, O	E	both

[0337] Table 16 lists a potential testing scenario, based on top variant burden hits (reported in **Table 14**). The analysis is for illustrative purposes only, it being acknowledged that greater diagnostic yields can be obtained by assaying for a larger number of variants, including those listed in **Table 15.** Examples are given for diagnostic yield using singleton variants, as well as a variety of combinations, including the use of the top 4 variants. For this set of variants, the test method is described as genotyping, as opposed to whole gene sequencing (e.g., determination of the status at each of the bases, which yields a binary output, as opposed to identification of variants elsewhere in the relevant genes).

Table 17: Potential testing scenario using genes identified as having a greater burden of damaging, heterozygous variants in the PML cohort

Gene	Ave Cases	Ethnicity	Ethnic-Specific Yield (%)	Overall Yield (Eur + Afr) (%)	Test Method
PLCG2	17/44	EUR	38	38	Gene sequencing
PLCG2	8/21	AFR	38		Gene sequencing
POLE	8/21	AFR	38		Gene sequencing
POLE	11/44	EUR	25	29	Gene sequencing
LRBA	11/44	EUR	25		Gene sequencing
EPG5	9/44	EUR	20		Gene sequencing
SHARPIN	8/44	EUR	18		Gene sequencing

[0338] Table 17 lists a potential testing scenario using genes identified as having a greater burden of damaging, heterozygous variants in the PML cohort (see **Table 13**). The nature of the testing method is 'gene sequencing' since the variants are not known in advance - any and all potentially damaging variants need to be considered in such an assay.

Table 18: Summary of genes that survive case-level, gene burden and/or variant burden analyses

Gene	Case Level	Variant Burden	Gene Burden
PLCG2	Yes	Yes	Yes
CHD7	Yes	Yes	
IFIH1	Yes	Yes	
AP3B1	Yes		Yes
EPG5	Yes		Yes
PIK3CD	Yes		Yes
LRBA		Yes	Yes
SHARPIN		Yes	Yes

[0339] Table 18 represents a summary of genes that survive case-level (2 or more examples in Tables 7, 8), gene burden and/or variant burden analyses (based on **Tables 13 and 14**). Of note is that PLCG2 satisfies all 3 criteria (2 or more examples, in **Table 8,** presence in Tables 13, 14). This summary demonstrates that many genes have been identified as significant on the basis of independent analysis methods.

Example 11 - Figures referenced in this study

[0340] Figures 1-12 represent example CNV data from the PML gene discovery study (71 PML cases, see **Table 7** for patient information) using array CGH (methods described herein). In each figure/drawing: 1) genome coordinates are listed at the top (hg18 assembly, chromosome number and position depicted); 2) data track 1 (labeled 'Genes') depicts the location of the RefSeq genes (exons are dark gray portions of the bars, introns are light gray portions of the bars); 3) data track 2 (labeled `Normal Cohort') depicts the size and location of CNVs found in the NVE cohort (PBio's proprietary control database consisting of CNV findings in apparently healthy - e.g. normal - subjects, see methods herein) with the y-axis corresponding to the number of NVE subjects that have the CNV; and 4) remaining data tracks are CNV data found in individual PML patients wherein the y-axis corresponds to the log2 ratio (see methods herein), points represent individual probes on the microarray, and line segments are shifted positive (copy number gain) or negative (copy number loss) based on the output of DNAcopy, the CNV calling algorithm. Typical log2 ratios for gains and losses on the Agilent 1M microarray (see methods herein) and our experimental protocols are: 0.6 for duplications, 1.0 for triplications (or homozygous duplications), -1.0 for heterozygous deletions, and <-2 (often -4.0 to -6.0) for homozygous deletions. Relevant genes are labeled in the 'Genes' data track.

[0341] Figure 1 represents an example of a gene impacted by germline and acquired CNVs. Germline CNVs that impact the PRKCB gene include patient PML50 with a 4.8Kb intronic heterozygous loss (also found in 7 Normal subjects) and patient PML11 with a 7.3Kb intronic gain (also found in 1 Normal subject). Acquired CNVs were found in 6 PML patients, a series of gains at ~23.9Mb with varying log2 ratios, suggestive of a mixed cell population (array CGH experiments were performed on blood-derived genomic DNA).

[0342] Figure 2 represents an example of potentially PML-relevant genes (TNFRSF13C and CENPM) impacted by acquired CNVs. Acquired CNVs were found in 9 PML patients, a series of gains at ~40.6Mb with varying log2 ratios, suggestive of a mixed cell population (array CGH experiments were performed on blood-derived genomic DNA). All 9 PML patients (see **Table 7** for patient information) had a primary diagnosis of HIV and were mixed gender (3 females and 6 males) and ethnicity (4 African ancestry and 5 European ancestry).

[0343] Figure 3 represents an example of a gene impacted by CNVs. A 7.2Kb intronic heterozygous loss (not found in Normal subjects, but an adjacent loss is found in 8 Normal subjects) that impacts the PKHD1 gene, was detected in patient PML26. CNVs were found in 3 PML patients, a series of gains at ~51.9Mb with varying log2 ratios, suggestive of a mixed cell population (array CGH experiments were

performed on blood-derived genomic DNA).

[0344] Figure 4 represents an example of a gene impacted by a CNV loss. The 14.7Kb intronic deletion impacts the BMPR2 gene. Heterozygous deletions were detected in patients PML58 and MVGS811-13a (also found in 2 Normal subjects), and a homozygous deletion was detected in patient PML29 (none found in Normal subjects). All three PML patients are males and their primary disease is HIV (see **Table 7**).

[0345] Figure 5 represents an example of a gene impacted by a CNV gain. The 10.2Kb exonic gain disrupts the COMMD6 gene. Two PML patients, PML29 and MVGS811-13a, have a homozygous duplication (log2 ratio comparable to triplications) based on the observation that 1000genomes subjects are reported to have this gain (see hg19 assembly DGV variant esv3632749, which reports 148 of 2504 subjects as having this gain; no Normals were found in PBio's NVE db). Both PML patients are males and their primary disease is HIV (see **Table 7**).

[0346] Figure 6 represents an example of a gene impacted by a CNV gain. The 27.4Kb exonic gain disrupts the KCTD7 gene and the right breakpoint is 16-90Kb upstream of RABGEF1 transcript variants (RefSeq: NM 001287060, NR 104676, NM 014504, NM 001287062, NM 001287061). Patient PML29 has a homozygous duplication (log2 ratio comparable to triplications) based on the observation that 1000genomes subjects are reported to have this gain (see hg19 assembly DGV variant esv3613515, which reports 28 of 2504 subjects as having this gain; no Normals were found in PBio's NVE db). Patient PML63 has a duplication. Both PML patients are males of African ancestry and their primary disease is HIV (see **Table 7**).

[0347] Figure 7 represents an example of a gene impacted by a CNV gain. The 344Kb exonic gain disrupts the FPR2 and ZNF616 genes (via left and right breakpoints) and additional genes fully encompassed by this CNV are: FPR3, ZNF350, ZNF350-AS1, ZNF432, ZNF577, ZNF613, ZNF614, ZNF615, ZNF649, ZNF649-AS1, ZNF841. Patient PML03 has a homozygous duplication (log2 ratio comparable to triplications) based on the observation that 3 Normal subjects (PBio's NVE db) are found to have a duplication of this region, along with patient PML10. Both PML patients are females of European ancestry and their primary diseases are HIV and MS (see **Table 7**).

[0348] Figure 8 represents an example of a gene impacted by a CNV loss. The 1.1Kb exonic deletion impacts the PIK3CD and PIK3CD-AS1 (previous gene symbol was C1orf200) genes. A homozygous deletion was detected in patient MVGS811-13a and this loss (heterozygous or homozygous) was not found in Normal subjects or the DGV public CNV database. The PML patient is a male and his primary disease is HIV (see **Table 7**). He is presumed to be of EUR ancestry (ethnicities were not available for MVGS samples).

[0349] Figure 9 represents an example of a gene impacted by an intergenic CNV gain. The 16.7Kb intergenic gain has a left breakpoint that is 105Kb upstream of the CD 180 gene (RefSeq transcript variant NM 005582). Patient MVGS995-4a has a homozygous duplication (log2 ratio comparable to triplications) based on the observation that 1000genomes subjects are reported to have this gain (see hg19 assembly DGV variant esv3605336, which reports 2 of 2504 subjects as having this gain; no Normals were found in PBio's NVE db). The PML patient is a male of European ancestry and his primary disease is MS (see **Table 7**).

[0350] Figure 10 represents an example of a gene impacted by an intergenic CNV loss. The 7.7Kb intergenic homozygous deletion has a left breakpoint that is 3-4Kb upstream of VDAC1 transcript variants (RefSeq: NM 003374, NR 036625, NR 036624). This loss (heterozygous or homozygous) was not found in Normal subjects or the DGV public CNV database. Patient PML30 is a male of European ancestry and his primary disease is HIV (see **Table 7**).

[0351] Figure 11 represents an example of a gene impacted by an intergenicCNV loss. The 6.8Kb intergenic homozygous deletion has a left breakpoint that is 4Kb downstream of EGR1 transcript variant (RefSeq: NM 001964) and 26Kb downstream of ETF1 transcript variants (RefSeq: NM 001256302, NM 004730, NM 001282185, NM 001291975, NM 001291974). This loss was found to be homozygous in 1 Normal subject and the loss was also reported in the DGV public CNV database (see hg19 assembly DGV variant esv3606925, which reports 33 of 2504 subjects as having this loss, homozygous vs. heterozygous subjects are unknown). Patient PML69 is a male of European ancestry and his primary disease (condition) is kidney transplant (see **Table 7**, reported as 'Other'). Patient PML69 was treated with CTLA4-Ig (belatacept, a CD28-B7 costimulation blocker and T-cell anergy inducer). The CD28 pathway includes links to the patient's genetic finding (e.g., homozygous deletion adjacent to the EGR1 gene) and several other genes that may be related to immunodeficiency (e.g., CD40LG, ITK, LCK, LRBA, PIK3CD, PIK3R1, PLCG2, WAS, and ZAP70) (Dekeyser M et al. Open Forum Infect Diseases, 2016, Refractory T-Cell Anergy and Rapidly Fatal Progressive Multifocal

Leukoencephalopathy following Prolonged CTLA4 Therapy).

[0352] Figure 12 represents an example of a gene impacted by an intergenic CNV loss. The 5.6Kb intergenic homozygous deletion has a left breakpoint that is 20Kb upstream of ITSN2 transcript variants (RefSeq: NM 019595, NM 006277, NM 147152). Heterozygous losses were found in 50 Normal subjects and the loss was also reported in the DGV public CNV database (see hg19 assembly DGV variant esv3590068, which reports 222 of 2504 subjects as having this loss, homozygous vs. heterozygous subjects are unknown). Patient PML65 is a male of African ancestry and his primary disease is HIV (see **Table 7**).

[0353] Figure 13 represents an example of known and/or predicted protein interactions using the String database (string-db.org; see Szklarczyk *et al.*, (2015) and references therein). A non-redundant list of all genes reported in **Table 7** (43 genes, which included those

whose expression was inferred to be impacted by a nearby intergenic CNV) as best solutions/explanations for 61 of 71 PML cases (11 PML cases are reported as 'unsolved', including 1 case for which only CGH data was obtained) was assessed using the String db. The `minimum required interaction score' was set to `high confidence (0.7)' and no additional 'interactors' were added. Of the 43 input genes, 21 were found to have high confidence interactions, as shown in the figure, along with annotation of the number of PML cases that had each of these genes as a solution/explanation (e.g., 3 PML cases in **Table 7** were found to have a PLCG2 solution).

Example 12 - Gene Burden Analysis

[0354] Gene burden analysis was performed as follows. Using a variety of in-house scripts, and data downloaded from ExAC (exac.broadinstitute.org), a count was performed for all variants occurring in each of the 419 genes listed in **Table 6.** Each variant was classified according to whether it was deemed damaging (on the basis of at least one of the prediction algorithms SIFT, PolyPhen2 or MutationTaster) or non-damaging, heterozygous or homozygous. This was performed in parallel for PML variants and those found in ExAC. ExAC data for which quality/coverage was < 80% of expected was not used and gene burden analysis could not therefore be performed.

[0355] An ethnic-specific (EUR or AFR only, there were too few LAT cases for this type of analysis) comparison was then performed for each of 4 categories:

- Homozygous damaging
- Homozygous non-damaging
- Heterozygous damaging
- Heterozygous non-damaging

[0356] For all 4 categories, variants with minor allele frequency (MAF) cutoffs of 0.01, 0.02. 0.03. 0.04, 0.05 and 0.1 were considered.

[0357] For each comparison, odds ratios (OR) and Fisher's exact test (FET) were calculated for the comparison of numbers of PML cases with at least one variant of the type under consideration and those in ExAC. Correction for multiple testing was performed by multiplying the FET by the number of genes being considered (419). Only genes for whom the FET corrected was < 0.05 were included in **Table 13,** which contains data on the average values for a given gene at all MAFs that passed FET correction. In practice, only the category of heterozygous damaging yielded significant genes.

Example 13 - Variant Burden Analysis

[0358] For each variant identified in at least one PML case, a count was performed in order to obtain the frequency of the same variant in the cohort as a whole. This aggregate data was compared to counts for the same variant as reported in ExAC. ExAC data was filtered for quality/coverage and variant burden analysis was not performed if ExAC coverage was < 80% expected.

[0359] Variant burden analysis was performed separately for EUR (n = 44 cases) and AFR (n = 21 cases) cohorts (LAT cohort was too small) and the OR and FET values calculated. From this analysis, only variants with OR>1 (e.g., potentially indicative of increased risk for PML) for both ethnicities (AFR and EUR) and for which the ExAC frequency of the variant was <5% were considered. Furthermore, only those variants for which the frequency in the ethnic-specific cohort was >10% (5 or more EUR cases, 3 or more AFR cases) were considered top-tier (Table 14), although other variants have been tabulated in **Table 15.**

Example 14 - Exemplary PML risk prediction tests

[0360] Table 16 provides exemplary markers for creating a low-cost, simple (genotype specific SNVs) PML risk prediction test. Other markers could be similarly devised from other SNVs reported in **Tables 14 and 15.** Different combinations of SNVs from Tables 14, 15 could be utilized in tests of varying complexity, to develop a test that would yield higher diagnostic yields than the top example listed in **Table 16** (e.g., 40%).

[0361] Table 17 provides exemplary genes that could be included in a gene panel sequencing test for PML risk prediction. Other gene panels could be similarly devised from genes reported in **Table 13,** or from other tables disclosed herein.

[0362] Table 9 contains `example' variants that may be considered as 'AD' causes of immunodeficiency (e.g., presence of just 1 of the 2 reported het SNVs in a given patient may be causing immunodeficiency), which may increase the risk for PML. For example, this may be a more likely scenario for het SNVs that are `novel' in the ExAC db (e.g., not found in the general population), and even more likely if such novel SNVs are found in >=2 PML cases (irrespective of the invoked disease model). Examples of this include the following 3 genes:

- AK2, 2 cases (**Table 9**), chr1:33476435, C>A, novel in ExAC PML20 and PML33, AFR and EUR, both HIV

- EPG5, 2 cases (**Table 9**), chr18: 43445601, T>G, novel in ExAC PML25 and PML27, both EUR, both HIV
- TNFRSF11A, 9 cases (**Table 7**), chr18: 60052034, A>C, novel in ExAC, see **Table 7** for case IDs, 2 AFR and 7 EUR, all HIV

[0363] It can be appreciated by those skilled in the art that immunodeficiency genes presently known to cause AR disease may potentially cause AD disease. Numerous examples have been reported in the literature, including several of the genes listed in **Table 6** (e.g., Disease model is indicated as AD AR for 32 genes, such as ADAR and TICAM1).

Example 15 - Exemplary 96-gene panel PML risk prediction tests

[0364] Table 19 contains an exemplary 96-gene panel based on genes that were found to have at least one PML case count from Tables 7 and 8. The "Genes" and "Case level solutions" columns showed genes and total number of PML cases (with at least one 'case level' solution) reported in Tables 7 and 8. In addition, the top 7 genes (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) from **Table 14** with SNVs based on `PML ALL FET' values <0.05 (column O) were also included in **Table 19.** Among these 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) with SNVs were based on 'PML ALL FET' values <0.05 (column O) from **Table 15.**

Table 19: exemplary 96-gene panel

Genes	Case level solutions
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
EPG5	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

Genes	Case level solutions
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
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
STIM 1	2
STIM2	1
STXBP2	3
TAP2	1
TBK1	2
TCIRG1	1
TICAM1	2
TLR3	2
TLR4	1
TNFRSF11A	10

Genes	Case level solutions
TNFRSF13B	1
TNFRSF8	1
TP53	1
TRAF3	1
TRAFD1	1
TRPM2	1
VPS45	1
WEE 1	2
ZAP70	3
TOTAL (96 genes)	**172**
Non-redundant cases	**67**
Dx yield for PML cohort (n = 70)	**95.7%**

[0365] The non-redundant number of PML cases and diagnostic yield are listed in the last 2 rows of **Table 19.** Specifically, a test including the 96 genes had a diagnostic yield of 95.7% based on the genetic findings from the 70 PML cases used in the present study.

Example 16 - Exemplary 39-gene panel PML risk prediction tests

[0366] Table 20 contains an exemplary 39-gene panel based on genes that were found to have multiple PML case count from Tables 7 and 8. The "Genes" and "Case level solutions" columns showed genes and total number of PML cases (with at least two 'case level' solutions) reported in Tables 7 and 8. In addition, the top 7 genes (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) from **Table 14** with SNVs based on `PML ALL FET' values <0.05 (column O) were also included in **Table 20.** Among these 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) with SNVs were based on 'PML ALL FET' values <0.05 (column O) from **Table 15.**

Table 20: exemplary 39-gene panel

Genes	Case level solutions
AP3B1	5
ATR	3
CARD11	3
CHD7	4
DCLRE1C	3
DOCK8	8
EPG5	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
RAG1	4
RAG2	2

Genes	Case level solutions
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**
Non-redundant cases	**57**
Dx yield for PML cohort (n = 70)	**81.4%**

[0367] The non-redundant number of PML cases and diagnostic yield are listed in the last 2 rows of **Table 20.** Specifically, a test including the 39 genes had a diagnostic yield of 81.4% based on the genetic findings from the 70 PML cases used in the present study.

Example 17 - Exemplary 23-gene panel PML risk prediction tests

[0368] Table 21 contains an exemplary 23-gene panel based on genes that were found to have multiple PML case count from Tables 7 and 8. The "Genes" and "Case level solutions" columns showed genes and total number of PML cases (with at least three 'case level' solutions) reported in Tables 7 and 8. In addition, the top 7 genes (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) from **Table 14** with SNVs based on `PML ALL FET' values <0.05 (column O) were also included in **Table 21.** Among these 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) with SNVs were based on 'PML ALL FET' values <0.05 (column O) from **Table 15.**

Table 21: exemplary 23-gene panel

Genes	Case level solutions
AP3B1	5
ATR	3
CARD11	3
CHD7	4
DCLRE1C	3
DOCK8	8
EPGS	3
GFI1	4
IFIH1	3
IGLL1	0
MAVS	0
NFKB1	3
NOD2	6
PIK3CD	4
PKHD1	3
PLCG2	5
PRKDC	4
RAG1	4
SHARPIN	1
STXBP2	3
TCIRG1	1
TNFRSF11A	10

Genes	Case level solutions
ZAP70	3
TOTAL (23 genes)	**83**
Non-redundant cases	**50**
Dx yield for PML cohort (n = 70)	**71.4%**

[0369] The non-redundant number of PML cases and diagnostic yield are listed in the last 2 rows of **Table 21.** Specifically, a test including the 23 genes had a diagnostic yield of 71.4% based on the genetic findings from the 70 PML cases used in the present study.

Example 18 - Exemplary 15-gene panel PML risk prediction tests

[0370] Table 22 contains an exemplary 15-gene panel based on genes that were found to have multiple PML case count from Tables 7 and 8. The "Genes" and "Case level solutions" columns showed genes and total number of PML cases (with at least four 'case level' solutions) reported in Tables 7 and 8. In addition, the top 7 genes (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) from **Table 14** with SNVs based on `PML ALL FET' values <0.05 (column O) were also included in **Table 22.** Among these 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) with SNVs were based on 'PML ALL FET' values <0.05 (column O) from **Table 15.**

Table 22: exemplary 15-gene panel

Genes	Case level solutions
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
TOTAL (15 genes)	**59**
Non-redundant cases	**39**
Dx yield for PML cohort (n = 70)	**55.7%**

[0371] The non-redundant number of PML cases and diagnostic yield are listed in the last 2 rows of **Table 22.** Specifically, a test including the 15 genes had a diagnostic yield of 55.7% based on the genetic findings from the 70 PML cases used in the present study.

Example 19 - Exemplary 11-gene panel PML risk prediction tests

[0372] Table 23 contains an exemplary 11-gene panel based on genes that were found to have multiple PML case count from Tables 7 and 8. The "Genes" and "Case level solutions" columns showed genes and total number of PML cases (with at least five 'case level' solutions) reported in Tables 7 and 8. In addition, the top 7 genes (CHD7, IFIH1, IGLL1, MAVS, PLCG2, SHARPIN, TCIRG1) from **Table 14** with SNVs based on `PML ALL FET' values <0.05 (column O) were also included in **Table 23.** Among these 7 genes, 3 genes (IGLL1, MAVS, SHARPIN) with SNVs were based on `PML ALL FET' values <0.05 (column O) from **Table 15.**

Table 23: exemplary 11-gene panel

Genes	Case level solutions
AP3B1	5
CHD7	4
DOCK8	8

Genes	Case level solutions
IFIH1	3
IGLL1	0
MAVS	0
NOD2	6
PLCG2	5
SHARPIN	1
TCIRG1	1
TNFRSF11A	10
TOTAL (11 genes)	**43**
Non-redundant cases	**33**
Dx yield for PML cohort (n = 70)	**47.1%**

[0373] The non-redundant number of PML cases and diagnostic yield are listed in the last 2 rows of **Table 23.** Specifically, a test including the 11 genes had a diagnostic yield of 47.1% based on the genetic findings from the 70 PML cases used in the present study.

Example 20 - Exemplary 10-gene panel PML risk prediction tests

[0374] **Table 24** contains an exemplary 10-SNV panel based on top 7 SNVs in **Table 14** and 3 SNVs from **Table 15** (based on overlapping genes between 14 and 15: IGLL1, MAVS, SHARPIN). Specifically, Using the top 10 SNVs (7 from **Table 14,** along with 3 from **Table 15,** residing in genes already selected from **Table 14),** an additive count (column "Case total additive (non-redundant)") was performed to determine how many PML cases had at least one of the variants when these were considered in order (e.g., column "Order (FET)": '1', first, followed by '1' + '2', followed by ` 1' + '2' + '3', etc). Since some individuals harbor more than one variant, the additive count is not equal to the simple sum of PML case numbers for each variant (column "Case total per SNV"). All genome coordinates are based on hg19 build.

[0375] An additive count was performed for ExAC subjects (column "ExAC subjects total additive (redundant)"), as follows: i) The average cohort size for ExAC for all variants was calculated; ii) Each total subject count (all ethnicities) was normalized to this average cohort size. The ExAC additive count represents a simple addition: labeled as "redundant" in column "ExAC subjects total additive (redundant)", because information regarding the possible presence of multiple variants in the same individual is not available; iii) Odds Ratios (ORs) and Fisher's Exact test (FET) values were calculated (columns "PML ALL OR additive" and "PML ALL FET additive").

Table 24: exemplary 10-gene panel

Order (FET)[1]	Table source	Gene	Variant (hg19)	Genotype	Case total per SNV	Case total additive (non-redundant)[2]	Dx yield (non-redundant)	ExAC subjects total additive (redundant)[3]	PML ALL OR additive	PML ALL FET additive
1	14	PLCG2	chr16:81942175, A>G	het	7	7	10%	730	6.50	2.00E-04
2	14	IFIH1	chr2:163136505, 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%	3,388	5.42	9.41E-09
5	14	MAVS	chr20:3846397, 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	chr20:3843027, C>A	hom	6	38	54%	10,393	3.77	5.26E-08

[1] SNV order based on lowest FET value reported in **Tables 14 and 15** for combined ethnicities

[2] PML case total = 70

[3] ExAC subject total = 43,419 (average for the 10 SNVs)

[0376] It can be appreciated by those skilled in the art that the above gene panels were selected based on the present genetic findings in 70 PML cases. Furthermore, a gene not presently selected for any of these exemplary gene panels may be added to the gene panel. For example, genes in which only 1 PML case was found to have variants fulfilling the criteria may be added to the gene panel if genetic validation in additional PML cases shows a 'n=1 case' gene is impacted by more than 1 PML case when the data are examined for a new set of PML cases. In some cases, additional genes (e.g., PML-linked genes such as DOCK8, BAG3, STAT1) may be added to the gene panel.

Example 21 - Identify additional genetic variations

[0377] The methods and protocols described in the previous examples can be used to identify any possible genetic variations. Data can be generated by comparing genetic variations identified in 2 cohorts: 1) non-diseased cohort including 1000 non-diseased individuals (e.g., individuals without PML); and 2) diseased cohort including 100 diseased individuals (e.g., individuals with PML). The individuals in the cohorts can be gender and/or ethnically matched. The genetic variations present in the non-diseased cohort and diseased cohort can be identified using CNV analysis (e.g., described in Example 2) or whole exome sequencing (e.g., described in Example 3). Two new genetic variations, CNV 1 (located on gene #1) and CNV 2 (located on gene #2), are identified, for example, by comparing the sequence data with a reference sequence (e.g., UCSC hg19). Data from a CNV database created using genome-wide CNV data on healthy subjects (or individuals without PML) such as the Normal Variation Engine (NVE) described herein can be used to determine if a CNV found in a PML cohort occurs at higher frequency or not compared to the NVE. Similarly, SNVs identified in a PML cohort can be interpreted for the potential relevance to PML using the Exome Aggregation Consortium (ExAC) or Genome Aggregation Database (gnomAD) publicly available resources [Lek M et al. Nature. 2016 Aug 18;536(7616):285-91]; that is, to obtain frequency data (e.g., ethnic-specific frequency) for variants under consideration. NVE databases for CNV assessment can be created for a variety of ethnicities (e.g., African and Latino ancestry subjects) to determine relevance of a CNV in a PML cohort compared to an ethnically matched CNV database on individuals without PML and/or using ethnic-specific data from publicly available CNV databases such as the Database of Genomic Variants [MacDonald J et al. Nucleic Acids Res. 2014 Jan;42(Database issue):D986-92]. In one example, 20 out of the 100 diseased individuals have CNV 1, and only 10 out of the 1000 non-diseased individuals have CNV 1. In another example, 10 out of the 100 diseased individuals have CNV 2, and only 5 out of the 1000 non-diseased individuals have CNV 2. The p-value can be calculated using standard tests, such as the Fisher's Exact Test (FET) and data can be selected using certain significance value. For example, genetic variations with a p-value of less than 0.05 are included. Further, the frequency and odds ratio of the two genetic variations can be calculated and summarized in an exemplary table:

Genetic variation	Genes	Frequency in diseased cohort	Frequency in non-diseased cohort	OR
CNV 1	Gene # 1	20/100 = 20%	10/1000 = 1%	(20/80) / (10/990) = 24.75
CNV 2	Gene #2	10/100 = 10%	5/1000 = 0.5%	(10/90) / (5/995) = 22.11

[0378] A subject without CNV 1, CNV 2, or both, may have a decreased risk of PML due to an infection of the brain by John Cunningham virus (JCV), and thus may be administered an immunosuppressive medicament, such as natalizumab.

[0379] Other genetic variations such as SNV can be identify using the same method as described above.

Example 22 - Analysis of the PML Cohort Array CGH Data for Additional CNVs

[0380] The array CGH data on the 71 PML cases was assessed (70 of 71 cases have array CGH and WES data, but PML67 only has array CGH data) to identify recurrent CNVs meeting the criteria of an OR >= 3 and a FET <= 0.05. The benefit of looking at higher frequency, recurrent CNVs is that PML risk can be easily assessed with PCR based, high-throughput screening methods that are also cost-effective. More frequent CNVs also have the potential to identify increased risk for PML in a larger proportion of patients, similar to the variant burden SNVs reported in multiple PML cases (e.g., Tables 14, 15, 38, and 39). For example, a genetic variant panel test for identifying a patient's risk of developing PML might contain one or more CNVs from Tables 1 or 28A and/or one or more SNVs from Tables 14, 15, 34, or 35.

[0381] Four recurrent losses of potential relevance in assessing PML risk were found and are reported in **Table 28A** (GRCh36/hg18 genome coordinates), three of which are also reported in **Table 28B.** One loss impacts exonic regions of CFHR1 and CFHR3 (SEQ ID 2200) and two other losses occur in intronic regions of the genes FUT8 (SEQ ID 2203) and ZBTB20 (SEQ ID 2202). The fourth loss is intergenic (SEQ ID 2201) and is located between two Ensembl genes (GRCh37/hg19 genome assembly): ENSG00000229703 (Ensembl transcript variant ENST00000437830, which overlaps RefSeq gene LOC101928226 and its corresponding RefSeq transcript NR 125950.1) and ENSG00000232694 (Ensembl transcript variant ENST00000436484). While the intergenic deletion may be impacting the expression of one or both adjacent genes (e.g., via loss of a transcription factor binding site), those skilled in the art will also consider whether long-range interactions are involved such that expression levels of genes located further away are modulated up or down (e.g., see Mifsud B et al. 2015, PMID 25938943). Similarly, intronic variants may have a similar impact and there is also the possibility that genetic variants (SNPs/SNVs, CNVs, etc.) are not directly involved in modulating gene expression but are in linkage disequilibrium (LD) with other genetic variants that are directly causing an effect (e.g., such variants serve as tags of causal or protective variants).

[0382] It can be appreciated by those skilled in the art that genetic variants (CNVs, SNVs, etc.) can be found in multiple, independent studies and many are reported in public databases, such as the Database of Genomic Variants (MacDonald J et al. 2014, PMID 24174537), the Exome Aggregation Consortium (ExAC, see Lek M et al. 2016, PMID 27535533), and ClinVar (Landrum M et al. 2018, PMID 29165669). Recurrent CNVs (gains or losses), such as those reported in **Table 28A** (and a subset that are reported in **Table 1),** are often found to approximately match the chromosomal breakpoints for CNVs reported from the 1000 Genomes Consortium Phase 3

study (Mills R et al. 2011, PMID 21293372; 1000 Genomes Project Consortium 2015, PMID 26432245). Since the CGH array (Agilent, Santa Clara, CA; array design ID AMADID 021529) used for genome-wide detection of CNVs in the PML cases contains ~1 million probes each spaced ~3Kb apart, the breakpoints can each map anywhere between the reported genomic locations of probes **(Tables 1 and 28A)** and the next nearest probe. Use of publically available CNV data is sometimes useful for mapping breakpoints via sequencing and designing assays when they are similar in size and location of the CNVs detected in an independent study. For example, the three recurrent CNVs reported in **Table 28A** are approximately the same and location as deletion reported in the 1000 Genomes Consortium Phase 3: SEQ ID 2200 is similar to esv3588469, SEQ ID 2201 is similar to esv3589567, SEQ ID 2202 is similar to esv3597466, and SEQ ID 2203 is similar to esv3634776.

[0383] **Table 28B** lists the CNV Subregion Numbers (SRN, numbers 364-366) for the three variants that are present as heterozygous and/or homozygous losses and pass statistical filters, as described below. For CNVs that are not identical, the CNV subregion corresponds to the overlapping portions with other CNVs. In the case of recurrent CNVs, the CNV subregion may be identical to the original CNV (e.g., as in Tables 28A and 28B). **Table 28B** also contains the OR and FET values in the comparison of the frequency of the CNV in NVE cases (e.g., control subjects that do not have PML) vs. PML cases. The NVE cohort size is 1,000 subjects (unaffecteds) and the PML cohort size is 71 patients (affecteds). All reported CNVs/CNV subregions in **Table 28B** had FET < 0.004, which is well below the 0.05 cutoff filter that was applied for this analysis. The corresponding OR values for the heterozygous losses (het loss) were 3.61 to 32.62 and for the homozygous losses (hom loss) were 42.57 to 71.98, which all are greater than the OR cutoff of 3. **Tables 29 and 30** contain the gene and transcript variant information for the **Table 28A** genes that are directly impacted by a CNV (FUT8 and ZBTB20) except for the CFHR1 and CFHR3 genes, which are reported on in **Tables 31 and 32.**

Immune Dysregulation Genes

[0384] Immune dysregulation disorders, which are one of the underlying factors of PML (Hatchwell E 2015, PMID 26029204), are caused by mutations in one or more genes. Numerous genes have already been reported in the literature (e.g., NCBI PubMed) or other public databases (e.g., OMIM) and new ones continue to be discovered. Consequently, a state-of-the-art comprehensive genetic risk prediction test for PML will periodically involve assessment of information in the field for newly reported immune dysregulation genes, as well as performing new discovery studies (e.g., using a CNV-based gene discovery approach, as described herein, on a new cohort of PML cases). Furthermore, newly discovered genes linked to a disease or condition are oftentimes independently validated by other studies. For example, in our first PML study, our discovery of a homozygous deletion (SEQ ID 1028) upstream of ITSN2 (GN 309) as a potential cause of a PML patient's underlying immunodeficiency condition was subsequently supported by evidence in a mouse model that this gene is involved in immune dysregulation (Burbage M et al. 2018, PMID 29337666).

[0385] Our original PML risk gene/variant discovery study involved assessment of a total of 435 genes (see Tables 3, 6, 25A, 25B, and 26). The published literature was reassessed for new genes that cause or contribute to immune dysregulation disorders and a curated list of 270 genes was assembled (see **Table 31).** For example, one source for new genes to consider in our PML cohort was derived from a December, 2017 update by the International Union of Immunological Societies (IUIS) on primary immunodeficiencies (Bousfiha A et al. 2017, PMID 29226301; Picard C et al. 2017, PMID 29226302), which described 334 genes linked to inborn errors of immunity and 137 of these were not included in our original gene lists (Tables 3, 6, 25A, 25B, and 26). Another source was a study of common variable immunodeficiency genes (CVID) and review of the literature (de Valles-Ibanez G et al. 2017, PMID 29867916), which described 92 genes and 40 were not included in our original list of 435 genes. Interestingly, 37 of these 92 genes were not included in the updated IUIS list (PMIDs 29226301 and 29226302), which underscores the importance of consulting multiple, independent sources when curating an updated list of new immune dysregulation genes to consider. One other major source of new genes that comprise our set of 270 genes was an exome sequencing study on patients with primary antibody deficiency (PAD), a frequent form of primary immunodeficiency (PID). This study (Abolhassani H et al. 2018, PMID 29921932) assessed 202 genes of which 76 genes were not included in our original set of 435 genes and 45 of the 202 genes were not included in the updated IUIS list (PMIDs 29226301 and 29226302). The role of complement genetics in immune dysregulation disorders has been discussed (Mayilyan K 2012, PMID 22773339). Some complement system genes were already assessed in our original set of PML genes (e.g., C1QA, C1QB, C1QC, C5AR1, CD55, CD59, CR2, and FCN3 listed in **Table 6)** and 30 new ones were included in the new set of immune dysregulation genes **(Table 31,** see genes annotated with PMID 22773339).

[0386] Since another key underlying factor of PML is infection with the JC virus (Hatchwell E 2015, PMID 26029204), we also reassessed the literature and other sources (e.g., interactions via the String db, Szklarczyk D et al. 2017, PMID 27924014) for JCV-related biology and corresponding genes. For example, this was previously done to identify candidate PML risk genes (e.g., as described by van der Kolk N et al. 2016, PMID 27042682) in which a subset was subsequently found to harbor rare variants in one or more PML cases in our study (see **Table 27,** JC virus biology column). This new assessment resulted in inclusion of 18 new genes that are linked to JCV biology: B2M, BRD4, CXCR3, CXCL10, HERC5, IFI35, IFIT2, IFIT3, IGHMBP2, MX1, MX2, OAS1, OAS2, OAS3, OASL, RELA, RSAD2, XAF1. Published source annotation for this set of genes are reported in **Table 31.** For example, the Type I interferon pathway, which was implicated as a factor in PML on the basis of genetic variants found in genes involved in the pathway (e.g., IFIH1 and MAVS) in our prior study of the 70 PML cases (e.g., see **Table 27),** was linked to JCV biology via String db analysis of other genes in the pathway (Assetta B et al. 2016, PMID 27381292). In another example, supporting biology for the OAS gene family was found in a study of HIV-associated neurocognitive disorder (HAND) patients (Puccini J et al. 2015, PMID 25834052; Sanfilippo C et al. 2018, PMID 28236279), which is another type of neurological disorder that can manifest in HIV patients (Kolson D 2018, PMID 28820724).

[0387] Potential links between Parkinson's Disease (PD) genes and PML were further explored since SNCA was in our first gene list (see

Table 6) and our separate investigation of SNCA triplication breakpoints identified the HERC5 and HERC6 genes as potentially relevant to PD (see Zafar F. et al. 2018, PMID 29928688; Im E et al. 2016, PMID 27534820). The potential overlap in pathological mechanisms between PML and PD is further underscored by the link of HERC5 to JCV (Assetta B et al. 2016, PMID 27381292) and neuroinflammation in general (Gelders G et al. 2018, PMID 29850629). Our curation of independent PD studies (including our own) or review papers have now linked a total of 37 PD-related genes to our candidate PML genes: 16 genes (CXCL12, CXCR4, DDX58, IFIH1, IL1B, MAVS, MYD88, RAB7A, RABGEF1, RAG1, SNCA, SQSTM1, TBK1, TLR3, TLR4, TMEM173) in our original list of 435 genes **(Table 6)** and 21 genes (ATG9A, CCZ1, FIS1, HERC5, HERC6, IFIT1, IFIT2, IFIT3, LRRK2, MFN1, MFN2, MON1A, MON1B, PINK1, PARK2/PRKN, RAB5A, RAB5B, RAB5C RSAD2, TBC1D15, TBC1D17) in our new list of 270 genes **(Table 31).** In addition to the previously mentioned PD citations (PMIDs 29850629 and 29928688), supporting literature for linking known and candidate PD genes to candidate PML genes include: Torres-Odio S et al. 2017, PMID 28768533; Yamano K et al. 2018, PMID 29360040; Paparisto E et al. 2018, PMID 29669830. In **Table 31,** the disease model for LRRK2 is listed as unknown since mutations in this gene are not presently known to cause an immunodeficiency disorder (e.g., it does not appear in the updated IUIS gene lists, see PMIDs 29226301 and 29226302). It is known to cause PD via an AD model but, without further studies, this cannot be assumed with respect to immune dysregulation in the context of other disorders. Interestingly, recent studies (e.g., see Yan R and Liu Z 2017, PMID 27830463; Lee H et al. 2017, PMID 28202666; Witoelar A et al. 2017, PMID 28586827; Hui K et al. 2018, PMID 29321258; Sheng D et al. 2018, PMID 29499195; Toledo Pinto T et al. 2018, PMID 29755459; Kim K et al. 2018, PMID 29760073) increasingly support a role for LRRK2 in immune function (e.g., Type I interferon pathway) and immune disorders (e.g., Crohn's disease and leprosy).

[0388] We also revisited the literature for any new findings related to genes in our original list of 419 candidate PML genes **(Table 6)** that were found to harbor genetics variants of interest in our original analysis of the array CGH CNV data and/or WES data on the 70 PML cases in our cohort. Supporting evidence (Magna M et al. 2014, PMID 24531836; Jiang R et al. 2014, PMID 25339669; Minguet S et al. 2017, PMID 28805811; Ratajczak M et al. 2018, PMID 29541038; Lee G et al. 2018, PMID 29674451) was found for TLR4-linked genes (CXCL12, CXCR4, MYD88) in our previous list of 435 genes **(Table 6),** plus two new genes (CAV1 and HMGB1) in our updated list of 270 genes **(Table 31).** Finally, MB21D1 (gene alias cGAS) was added to our updated gene list **(Table 31)** on the basis of its direct link to TMEM173 (gene alias STING), which was included in our original gene list **(Table 6);** for example, as highlighted in a recent review (Chen Q et al. 2016, PMID 27648547). Finally, ITSN1 was included in our list of 270 genes **(Table 31)** based on a supporting study (Burbage M et al. 2018, PMID 2933766), cited herein as supporting functional data for one of our CNV-discovered PML candidate genes (ITSN2), plus three other more recent studies (Alvisi G et al. 2018, PMID 29599122; Dergai O et al. 2018, PMID 29851086; Gryaznova T et al. 2018, PMID 29958948).

[0389] It is well appreciated by those skilled in the art that many common disorders have a genetic basis and can be caused by one (autosomal, X-linked dominant, or X-linked recessive) or two mutations (autosomal recessive) in a gene, and that the mutation(s) may occur in any one of numerous genes. Furthermore, an increasing number of genetic studies are revealing that a given disorder can be caused or modified (e.g., age of onset or severity) by multiple variants present within a patient's genome. For example, autism has been linked to the presence of two or more genetic variants (CNVs or SNVs) within an individual (Marshall C et al. 2008, PMID 18252227; Yuen R et al. 2015, PMID 25621899). Accumulating evidence suggests that a monogenic disease model for immune dysregulation also provides an incomplete explanation of the pathology. For example, de Valles-Ibanez et al. 2018 (PMID 29867916) discuss the importance of considering other disease models (e.g., oligogenic or polygenic), particularly since the majority of CVID patients remain undiagnosed despite genetic assessment with exome or whole genome sequencing approaches.

[0390] Based on advances in genetic sequencing technologies and rapidly growing disease gene lists, many genetic testing providers now use gene panel tests, which are used to assess and deliver clinical information only for the subset of genes that have been definitively reported to cause the disease. Gene panel tests can be performed by genome capture methods for just the genes of interest or by performing WES or WGS but only interpreting the relevant disease genes (e.g., an *in silico* capture of a subset of the genes in the human genome). It can be appreciated by those skilled in the art that a PML genetic risk prediction test will likely involve testing for deleterious genetic variants in several immune dysregulation genes and that new genes/variants may be added to the testing panel as new studies are reported on immune disorder patients and/or PML patients.

[0391] There are several genetic testing providers that now offer gene panel testing. An example of a commercial provider of several different disease gene panels is Invitae (San Francisco, CA). Given their experience with testing and interpretation of a wide variety of gene panels (e.g., inherited cancers, metabolic disorders, and cardiology disorders), we compared their set of immunology gene panels to our curated PML candidate gene lists **(Tables 6 and 31)** to determine if their existing gene panel could be one potential option for assessing the genetic risk of developing PML. As of June 2018, Invitae had a total of 210 genes in their immunology panels and 157 of these were already included in our first list of 435 PML candidate genes (Tables 3, 6, 25A, 25B, and 26) and 206 were present in IUIS's December, 2017 update on primary immunodeficiencies (Bousfiha A et al. 2017, PMID 29226301; Picard C et al. 2017, PMID 29226302). However, 3 genes (ACD, PMM2, and SLC7A7) in Invitae's immunology gene panels were not found in our CNV-based gene discovery studies or in the various curation resources described herein. Therefore, for completeness, we included these 3 genes in our updated panel of 270 genes **(Table 31).**

[0392] There are 9 genes in **Table 31** that have had gene symbol changes since the PML WES data were generated, these are (current RefSeq gene symbol list first and WES data gene symbol second):

ADA2	CECR1
ADGRL2	LPHN2

CXCL8	IL8
FAAP24	C19orf40
NSMCE3	NDNL2
OTULIN	FAM105B
PRKN	PARK2
STN1	OBFC1
WASHC5	KIAA0196

[0393] Transcript variants and SEQ ID numbers for the additional 270 immune dysregulation genes **(Table 31)** are reported in **Table 32.** No RefSeq transcript variants were found for the IGHM and IGKC genes, so these are omitted from **Table 32.**

Analysis of the PML Cohort WES Data for 270 Additional Immune Dysregulation Genes

[0394] Using our previous methods for mining genetic variants from WES data on 70 PML cases (described herein), we have identified several new variants in a subset of the 270 genes (see Tables 32-37, 38, and 39) that may be causing or contributing to a PML patient's immune-compromised condition. As was done for the original 435 immune dysregulation genes (Tables 3, 6, 25A, 25B, and 26), three types of genetic analysis methods were used for the new set of 270 genes **(Table 31):** 1) case-solving approach, 2) gene burden analysis, and 3) variant burden analysis. A total of 275 SNVs (see **Table 33)** in 113 of the 270 genes that were assessed were identified in the cohort of 70 PML cases.

[0395] The case-solving method, which involves assessing WES data for each PML patient for the presence of rare, deleterious SNVs in comparison to an unselected control set of WES data (publicly available data from the Exome Aggregation Consortium, exac.broadinstitute.org), was applied to all 70 PML cases for which WES data were collected (see **Table 7** for Sample ID, Ethnicity, Gender, and Primary Disease for each case). In the first study of this cohort that involved assessment of 435 genes (Tables 3, 6, 25A, 25B, and 26), the case-solving findings were reported in Tables 7-10 wherein **Table 7** reported the top genetic variation(s) and gene for each case on the basis of variant frequency (< 1 in 100 frequency cutoff in comparison to the ExAC subjects) and the known biology of the gene at the time of the study. Additional potential solutions were reported in Tables 8-10. Given the growing evidence for multiple variants/genes causing or contributing to immune dysregulation, case-solving results for the 270 additional genes were partitioned into 4 tables as follows:

Table 34, SNVs (het, hom, or phased comp het) with a frequency of <= 1/1,000 ExAC subjects or are novel;

Table 35, SNVs (het, hom, or phased comp het) with a frequency of <= 1/100 but > 1/1,000 ExAC subjects;

Table 36, Un-phased het SNVs with a potential compound heterozygous frequency of <= 1/100 ExAC subjects (a subset may prove to be in trans with further validation work);

wherein SNVs are referred to as heterozygous (het), homozygous (hom), or compound heterozygous (comp het).

[0396] As before, the frequency information for each variant was assessed in an ethnic-specific manner (AFR, EUR, or LAT). Variants were excluded from the analysis if the allele number in the ExAC subjects for a given ethnicity was < 75% of the maximal number (-10,300 for AFR subjects, -11,300 for LAT subjects, or -66,500 for EUR subjects). An exception was made for variant CFHR3 chr1: 196759282, C>T, which has lower AN but only in AFR ExAC subjects. Other studies have reported that a deletion impacting CFHR3 is present at higher frequency in African ancestry subjects (e.g., see Cantsilieris S et al. 2018, PMID 29686068).

[0397] We observed a recurrent CFHR3 deletion (~56Kb in size) in our NVE control subjects and PML cases that maps to: chr1:196742735-196799244 (GRCh37/hgl9). Both heterozygous and homozygous losses were found but only the homozygous deletions found in PML cases are reported in **Table 28A** (SEQ ID 2200). The loss overlaps the 1000 Genomes Phase 3 deletion esv3588469 (~80Kb in size): chr1:196728877-196808865 (GRCh37/hg19). It is possible that these two deletions are the same, which can be verified by sequencing the breakpoints of the deletion found in our NVE controls and PML cases. We observed the CFHR3 deletion in the homozygous state in 52 NVE controls (all EUR ancestry) and 5 PML cases (2 AFR and 1 LAT ancestry). Furthermore, we found a CFHR3 hom SNV in our variant burden analysis (described herein, see **Table 38).** Give the high frequency of the deletion in the general population, it is possible that some individuals called as homozygotes for the SNV are compound heterozygotes for the SNV and the deletion. The nature of short read sequencing means that homozygosity for a variant is inferred if a) the data is of sufficient quality (sufficient number of reads present) and b) only one of the two possible sequences at a given base position is observed. However, if an individual has only one allele (the other having been deleted), there is only one possible 'state' at each given base. This is known as loss of heterozygosity and is distinct from homozygosity. For this reason, the analysis of the homozygous state for the SNV has to be treated cautiously. However, the compound heterozygous state (SNV plus deletion) may well result in the same functional consequences as the homozygous SNV. Further laboratory work would be required to deconvolute the genotypes in this cohort.

[0398] Interestingly, the CFHR3 deletion has been linked to atypical hemolytic uremic syndrome (aHUS), as described in a recent case report (Bitzan M. et al. 2018, PMID 29728803). Furthermore, a 2015 case report (Gomez-Cibeira E et al. 2015, PMID 26718572) links aHUS to PML. No genetic information was reported for this aHUS patient, but she developed PML after receiving eculizumab.

[0399] A recessive or dominant disease model was applied according to the information reported in Table 31. For genes that have an 'unknown' or 'association' entry, the disease model was assumed to be AD and/or AR. Furthermore, if an AD model is currently the standard for a given gene (e.g., in OMIM), we did not rule out the potential for an AR model. For example, MBL2 is reported in **Table 31** as an AD model gene but **Table 34** reports on a rare homozygous (om) SNV (amino acid change G57E) in the PML15 case. The frequency cutoff of < 1 in 100 was relaxed for FAT4 SNV A807V (SEQ ID 3086) in **Tables 35 and 36** since this variant is present as a solution for multiple PML cases (PML09, PML28, PML31, PML32, PML37, PML65) in **Table 34.** In instances wherein a PML case had 3 or more SNVs, the frequency calculations were as follows: a) for a hom SNV paired with multiple het SNVs (e.g., obligate comp hets), the pairwise frequency was calculated for the hom SNV and each het SNV, and b) for 3 or more het SNVs, pairwise frequency was calculated using the rarest SNV paired with each higher frequency SNV.

[0400] Gene and variant burden analyses were performed on the updated set of 270 genes **(Table 31)** according to the methods described herein for the original set of 435 genes (Tables 3, 6, 25A, 25B, and 26). The results are reported in the following tables:

Table 37, gene burden results (same criteria used to generate the **Table 13** results);

Table 38, variant burden results (same criteria used to generate the **Table 14** results);

Table 39, variant burden results (same criteria used to generate the **Table 15** results);

[0401] The top SNV in **Table 38** is CFHR3 chr1:196759282, C>T (rs138675433 in dbSNP build 150). The genotype is homozygous in 3 AFR and 3 EUR ancestry PML cases. Therefore, we have found a homozygous deletion (SEQ ID 2200) and a homozygous SNV (SEQ ID 3025) impacting the CFHR3 gene. Both the deletion and SNV are present at higher frequency in AFR ancestry subjects. Therefore, genotyping both variants in control subjects and PML patients of different ancestries will help define their role in increasing or decreasing PML risk. This is potentially highly relevant in light of a PML case report on a patient that had aHUS (Gomez-Cibeira E et al. 2015, PMID 26718572) and the link between aHUS and the CFHR3 gene (Pouw R et al. 2018, PMID 29740447).

[0402] Genetic findings in the first set of 435 genes for the 70 PML cases were summarized in a variety of formats (see Tables 7-10 and 13-27). Genetic findings in the updated set of 270 genes are summarized in **Table 40** for all three analysis methods. For some PML cases, an SNV can appear in more than one case solving table (Tables 34-36), such as when an 'unknown' disease model gene has 2 or more SNVs and a given SNV is ascribed an AD model solution and an AR model solution. However, the total PML case counts in **Table 40** correspond to a non-redundant number of PML cases to accurately reflect the total number of potentially solved PML cases per gene.

[0403] One of the top genes summarized in **Table 40** is FAT4, which is the only gene that has genetic findings reported for all three types of analysis (case-solving, gene burden, variant burden) as reported in Tables 34-39. The FAT4 gene also has the highest number of potentially solved PML cases, 23 solutions based on Tables 34-36. Other genes with a higher number of case-solving solutions are: PRRC2A (22 solutions), MSH5 (9 solutions), LRRK2 (8 solutions), and MX1 (8 solutions). Autosomal recessive mutations in the FAT4 gene (OMIM 612411) are involved in two disorders: Hennekam lymphangiectasia-lymphedema syndrome 2 (OMIM 616006) and Van Maldergem syndrome 2 (OMIM 615546). Hennekam lymphangiectasia-lymphedema syndrome 2 is of particular interest since a clinical feature of these patients is intestinal lymphangiectasia, which has been reported in a case of PML (Gomez-Cibeira E et al. 2015, PMID 26718572), which is also noted herein for the link to the CFHR3 gene via the PML patient's aHUS condition.

[0404] Another top gene in **Table 40** is LRRK2, which has 8 solutions based on Tables 34-36 and was significant in gene burden analysis **(Table 37).** LRRK2 variants (e.g., the G2019S gain-of-function mutation) were initially found to cause or contribute to Parkinson's Disease (PD), but additional studies have linked the LRRK2 gene to immune function and disorders, as reported herein and in **Table 31** (see PMID citations). We also found a subset of PML patients harboring a pair of LRRK2 SNVs (N299K, also known as N551K, SEQ ID 3192; R1398H, SEQ ID 3197), with heterozygous or homozygous genotype (see **Tables 34 and 35),** that are reported to be protective against PD (e.g., see Ross O et al. 2011, PMID 21885347; Heckman M et al. 2014, PMID 23962496; Heckman M et al. 2016, PMID 27521182). Whether one or both variants modify PML risk will require further studies.

Table 28A: SEQ ID 2200-2203, four recurrent CNVs

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
1	195009358	195065867	56509	hom loss	3143	CFHR1, CFHR3	2200
1	195009358	195065867	56509	hom loss	3159	CFHR1, CFHR3	2200
1	195009358	195065867	56509	hom loss	3193	CFHR1, CFHR3	2200
1	195009358	195065867	56509	hom loss	3202	CFHR1, CFHR3	2200
1	195009358	195065867	56509	hom loss	3273	CFHR1, CFHR3	2200
1	246933929	246940630	6701	het loss	3009		2201

Chr	Original CNV Start	Original CNV Stop	Original CNV Size	CNV Type	PML Case ID	RefSeq Gene Symbol	SEQ ID
1	246933929	246940630	6701	het loss	3125		2201
1	246933929	246940630	6701	het loss	3152		2201
1	246933929	246940630	6701	het loss	3173		2201
1	246933929	246940630	6701	het loss	3175		2201
1	246933929	246940630	6701	het loss	3188		2201
1	246933929	246940630	6701	het loss	3193		2201
1	246933929	246940630	6701	het loss	3202		2201
1	246933929	246940630	6701	het loss	3203		2201
1	246933929	246940630	6701	het loss	3205		2201
1	246933929	246940630	6701	hom loss	3163		2201
1	246933929	246940630	6701	hom loss	3273		2201
3	116140997	116146946	5949	het loss	3009	ZBTB20	2202
3	116140997	116146946	5949	het loss	3144	ZBTB20	2202
3	116140997	116146946	5949	het loss	3152	ZBTB20	2202
3	116140997	116146946	5949	het loss	3154	ZBTB20	2202
3	116140997	116146946	5949	het loss	3159	ZBTB20	2202
3	116140997	116146946	5949	het loss	3192	ZBTB20	2202
3	116140997	116146946	5949	het loss	3194	ZBTB20	2202
3	116140997	116146946	5949	het loss	3200	ZBTB20	2202
3	116140997	116146946	5949	het loss	3273	ZBTB20	2202
3	116140997	116146946	5949	hom loss	3173	ZBTB20	2202
14	65143023	65145120	2097	het loss	3006	FUT8	2203
14	65143023	65145120	2097	het loss	3010	FUT8	2203
14	65143023	65145120	2097	het loss	3160	FUT8	2203
14	65143023	65145120	2097	het loss	3178	FUT8	2203
14	65143023	65145120	2097	het loss	3183	FUT8	2203
14	65143023	65145120	2097	het loss	3189	FUT8	2203
14	65143023	65145120	2097	het loss	3194	FUT8	2203
14	65143023	65145120	2097	het loss	3201	FUT8	2203

Table 28A lists recurrent CNVs of interest in this study.

Table 28B: SRN 364-366, three recurrent CNVs with values of OR >= 3 and FET <= 0.05 (SRN = original CNV since these are recurrent CNVs)

Chr	CNV Sub-region Start	CNV Sub-region Stop	CNV Sub-region Size	CNV Type	PML Case ID	Ref-Seq Gene Symbol	Exon Over-lap	NVE Cases	PML Cases	FET	OR	CNV Sub-region No. cases
1	246933929	246940630	6701	het loss	3009		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3125		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3152		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3173		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3175		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3188		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3193		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3202		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3203		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	het loss	3205		N	5	10	2.02E-09	32.62	364
1	246933929	246940630	6701	hom loss	3163		N	0	2	0,004453	71.98	364
1	246933929	246940630	6701	hom loss	3273		N	0	2	0,004453	71.98	364
3	116140997	116146946	5949	het loss	3009	ZBTB20	N	23	9	0,000122	6.17	365
3	116140997	116146946	5949	het loss	3144	ZBTB20	N	23	9	0,000122	6.17	365
3	116140997	116146946	5949	het loss	3152	ZBTB20	N	23	9	0,000122	6.17	365
3	116140997	116146946	5949	het loss	3154	ZBTB20	N	23	9	0,000122	6.17	365
3	116140997	116146946	5949	het loss	3159	ZBTB20	N	23	9	0,000122	6.17	365
3	116140997	116146946	5949	het loss	3192	ZBTB20	N	23	9	0,000122	6.17	365
3	116140997	116146946	5949	het loss	3194	ZBTB20	N	23	9	0,000122	6.17	365

3	116140997	116146946	5949	het loss	3200	ZBTB20	N	23	9	0,000122	6.17	365
3	116140997	116146946	5949	het loss	3273	ZBTB20	N	23	9	0,000122	6.17	365
3	116140997	116146946	5949	hom loss	3173	ZBTB20	N	0	1	0,004453	42.57	365
14	65143023	65145120	2097	het loss	3006	FUT8	N	34	8	0,004761	3.61	366
14	65143023	65145120	2097	het loss	3010	FUT8	N	34	8	0,004761	3.61	366
14	65143023	65145120	2097	het loss	3160	FUT8	N	34	8	0,004761	3.61	366
14	65143023	65145120	2097	het loss	3178	FUT8	N	34	8	0,004761	3.61	366
14	65143023	65145120	2097	het loss	3183	FUT8	N	34	8	0,004761	3.61	366
14	65143023	65145120	2097	het loss	3189	FUT8	N	34	8	0,004761	3.61	366
14	65143023	65145120	2097	het loss	3194	FUT8	N	34	8	0,004761	3.61	366
14	65143023	65145120	2097	het loss	3201	FUT8	N	34	8	0,004761	3.61	366

Table 28B lists CNVs/CNV subregions of interest with OR and FET values.

Table 29: GN 491 and 492, NCBI Gene ID, descriptions, RefSeq summary for 2 of 3 **Table 28A** genes

RefSeq Gene Symbol	Exon Overlap	NCBI Gene ID	Gene Description	RefSeq Summary	Gene No. (GN)
FUT8	intronic	2530	alpha-(1,6)-fucosyltransferase isoform a	This gene encodes an enzyme belonging to the family of fucosyltransferases. The product of this gene catalyzes the transfer of fucose from GDP-fucose to N-linked type complex glycopeptides. This enzyme is distinct from other fucosyltransferases which catalyze alpha1-2, alpha1-3, and alpha1-4 fucose addition. The expression of this gene may contribute to the malignancy of cancer cells and to their invasive and metastatic capabilities.	491
				Alternative splicing results in multiple transcript variants, [provided by RefSeq, May 2011]. Transcript Variant: This variant (1, also known as B6) represents the longest transcript and encodes the longer isoform (a). Variants 1 and 3 encode the same isoform. Sequence Note: This RefSeq record was created from transcript and genomic sequence data to make the sequence consistent with the reference genome assembly. The genomic coordinates used for the transcript record were based on transcript alignments.	
				Publication Note: This RefSeq record includes a subset of the publications that are available for this gene. Please see the Gene record to access additional publications. Transcript exon combination:: AJ536056.1, AJ536054.1 [ECO:0000332] RNAseq introns :: single sample supports all introns ERS025084, ERS025088 [ECO:0000348]	
ZBTB20	intronic	26137	zinc finger and BTB domain-containing protein 20 isoform 1	N/A	492
Table 29 lists gene information for genes in **Table 28B**					

Table 30: SEQ ID 2204-2215, transcript variants for **Table 29** genes (FUT8 and ZBTB20, third CNV is intergenic)

RefSeq Gene Symbol	Exon Overlap	RefSeq Accession Number	mRNA Description	SEQ ID
FUT8	intronic	NM 004480	Homo sapiens fucosyltransferase 8 (alpha (1,6) fucosyltransferase) (FUT8), transcript variant 4, mRNA.	2204
FUT8	intronic	NM 178155	Homo sapiens fucosyltransferase 8 (alpha (1,6) fucosyltransferase) (FUT8), transcript variant 1, mRNA.	2205
FUT8	intronic	NM 178156	Homo sapiens fucosyltransferase 8 (alpha (1,6) fucosyltransferase) (FUT8), transcript variant 3, mRNA.	2206
FUT8	intronic	NR 038167	Homo sapiens fucosyltransferase 8 (alpha (1,6) fucosyltransferase) (FUT8), transcript variant 5, non-coding RNA.	2207

RefSeq Gene Symbol	Exon Overlap	RefSeq Accession Number	mRNA Description	SEQ ID
FUT8	intronic	NR 038170	Homo sapiens fucosyltransferase 8 (alpha (1,6) fucosyltransferase) (FUT8), transcript variant 2, non-coding RNA.	2208
ZBTB20	intronic	NM 001164343	Homo sapiens zinc finger and BTB domain containing 20 (ZBTB20), transcript variant 3, mRNA.	2209
ZBTB20	intronic	NM 015642	Homo sapiens zinc finger and BTB domain containing 20 (ZBTB20), transcript variant 2, mRNA.	2210
ZBTB20	intronic	NM 001164342	Homo sapiens zinc finger and BTB domain containing 20 (ZBTB20), transcript variant 1, mRNA.	2211
ZBTB20	intronic	NM 001164347	Homo sapiens zinc finger and BTB domain containing 20 (ZBTB20), transcript variant 7, mRNA.	2212
ZBTB20	intronic	NM 001164346	Homo sapiens zinc finger and BTB domain containing 20 (ZBTB20), transcript variant 6, mRNA.	2213
ZBTB20	intronic	NM 001164344	Homo sapiens zinc finger and BTB domain containing 20 (ZBTB20), transcript variant 4, mRNA.	2214
ZBTB20	intronic	NM 001164345	Homo sapiens zinc finger and BTB domain containing 20 (ZBTB20), transcript variant 5, mRNA.	2215

Table 30 represents a non-redundant list of transcript variants that correspond to the **Table 29** genes

Table 31: GN 493-762, 270 candidate PML genes, updated list from various public databases (e.g., PubMed)

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
ACD	ADAR	Public db	PMID: 25205116, 25233904, 26810774	493
ACTB	AD	Public db	PMID: 29226301, 29226302	494
ACTN4	AD	Public db	PMID: 29921932	495
ADA2	AR	Public db	PMID: 29226301, 29226302, 29867916	496
ADAM17	AR	Public db	PMID: 29226301, 29226302	497
ADGRL2	unknown	Public db	PMID: 29921932	498
AIRE	ADAR	Public db	PMID: 29226301, 29226302, 29867916	499
ANP32B	unknown	Public db	PMID: 29867916	500
AP1S3	AR	Public db	PMID: 29226301, 29226302	501
ARPC1B	AR	Public db	PMID: 29226301, 29226302	502
ATG12	unknown	Public db	PMID: 28141795, 28295214	503
ATG16L1	association	Public db	PMID: 28141795	504
ATG5	AR	Public db	PMID: 28141795,28295214	505
ATG7	unknown	Public db	PMID: 28141795	506
ATG9A	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	507
ATP6AP1	XLR	Public db	PMID: 29226301, 29226302	508
B2M	AR	Public db	PMID: 27381292, 29226301, 29226302	509
BCL11B	AD	Public db	PMID: 29226301, 29226302	510
BCL2	association	Public db	PMID: 29867916	511
BLK	AD	Public db	PMID: 29867916, 29921932	512
BRD4	unknown	Public db	PMID: 27007123, 29475942	513
BTLA	unknown	Public db	PMID: 29867916	514
C1R	AR	Public db	PMID: 29226301, 29226302, 22773339	515
C1S	AR	Public db	PMID: 29226301, 29226302, 22773339	516
C2	AR	Public db	PMID: 29226301, 29226302, 22773339	517
C3	ADAR	Public db	PMID: 29226301, 29226302, 22773339	518
C4A	AR	Public db	PMID: 29226301,29226302,22773339	519
C4B	AR	Public db	PMID: 29226301, 29226302, 22773339	520
C5	AR	Public db	PMID: 29226301, 29226302, 22773339	521
C6	AR	Public db	PMID: 29226301, 29226302, 22773339	522
C7	AR	Public db	PMID: 29226301, 29226302, 22773339	523

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
C8A	AR	Public db	PMID: 29226301, 29226302, 22773339	524
C8B	AR	Public db	PMID: 29226301, 29226302, 22773339	525
C8G	AR	Public db	PMID: 29226301, 29226302, 22773339	526
C9	AR	Public db	PMID: 29226301, 29226302, 22773339	527
CAD	AR	Public db	PMID: 29921932	528
CAMLG	unknown	Public db	PMID: 29921932	529
CARD14	AD	Public db	PMID: 29226301, 29226302	530
CASP10	AD	Public db	PMID: 29226301, 29226302	531
CAV1	AD	Public db	PMID: 24531836, 25339669, 28805811, 29541038,29674451	532
CCBE1	AR	Public db	PMID: 29226301, 29226302	533
CCDC22	XLR	Public db	PMID: 29921932	534
CCZ1	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	535
CD22	unknown	Public db	PMID: 29867916	536
CD274	unknown	Public db	PMID: 29867916	537
CD276	unknown	Public db	PMID: 29867916	538
CD36	AR	Public db	PMID: 29921932	539
CD37	unknown	Public db	PMID: 29867916	540
CD38	unknown	Public db	PMID: 29867916	541
CD46	AD	Public db	PMID: 29226301, 29226302, 22773339	542
CD5	unknown	Public db	PMID: 29867916	543
CD70	AR	Public db	PMID: 29226301, 29226302	544
CD72	unknown	Public db	PMID: 29867916	545
CD74	unknown	Public db	PMID: 29867916	546
CD84	unknown	Public db	PMID: 29867916	547
CD93	unknown	Public db	PMID: 29867916, 22773339	548
CEBPE	AR	Public db	PMID: 27042682, 29226301, 29226302	549
CFB	ADAR	Public db	PMID: 29226301, 29226302, 22773339	550
CFD	AR	Public db	PMID: 29226301, 29226302, 22773339	551
CFH	ADAR	Public db	PMID: 29226301, 29226302, 22773339	552
CFHR1	ADAR	Public db	PMID: 29226301, 29226302, 22773339	553
CFHR2	ADAR	Public db	PMID: 29226301, 29226302, 22773339	554
CFHR3	ADAR	Public db	PMID: 29226301, 29226302, 22773339	555
CFHR4	ADAR	Public db	PMID: 29226301, 29226302, 22773339	556
CFHR5	ADAR	Public db	PMID: 29226301, 29226302, 22773339	557
CFI	ADAR	Public db	PMID: 29226301, 29226302, 22773339	558
CFP	XLR	Public db	PMID: 29226301, 29226302, 22773339	559
CFTR	AR	Public db	PMID: 29226301, 29226302	560
CHD2	AD	Public db	PMID: 29921932	561
CLEC16A	unknown	Public db	PMID: 29867916, 29921932	562
CLPB	AR	Public db	PMID: 29226301, 29226302	563
COPA	AD	Public db	PMID: 29226301, 29226302	564
CSF2RA	XLR	Public db	PMID: 29226301, 29226302	565
CSF2RB	AR	Public db	PMID: 29226301, 29226302	566
CTC1	AR	Public db	PMID: 29226301, 29226302	567
CXCL1	unknown	Public db	PMID: 28681388	568
CXCL10	unknown	Public db	PMID: 23086711, 28237728	569
CXCL5	unknown	Public db	PMID: 28681388	570
CXCL8	unknown	Public db	PMID: 28681388	571
CXCR3	unknown	Public db	PMID: 28237728	572

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
CYBA	AR	Public db	PMID: 29226301, 29226302	573
DCLRE1B	AR	Public db	PMID: 29226301, 29226302	574
DNAJC21	AR	Public db	PMID: 29226301, 29226302	575
DNASE 1L3	AR	Public db	PMID: 29226301, 29226302	576
DNASE2	unknown	Public db	PMID: 29226301, 29226302	577
EBF1	unknown	Public db	PMID: 29921932	578
EGF	association	Public db	PMID: 29921932	579
ERCC6L2	AR	Public db	PMID: 29226301, 29226302	580
EXTL3	AR	Public db	PMID: 29226301, 29226302	581
FAAP24	AR	Public db	PMID: 29226301, 29226302	582
FADD	AR	Public db	PMID: 29226301, 29226302	583
FAT4	AR	Public db	PMID: 29226301, 29226302	584
FCER2	unknown	Public db	PMID: 29867916	585
FERMT3	AR	Public db	PMID: 29226301, 29226302	586
FIS1	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	587
G6PD	XLD	Public db	PMID: 29226301, 29226302	588
GINS 1	AR	Public db	PMID: 29226301, 29226302	589
HERC5	unknown	Public db	PMID: 27381292, 28768533, 29360040, 29669830, 29928688	590
HERC6	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	591
HMGB1	unknown	Public db	PMID: 24531836, 25339669, 28805811, 29541038,29674451	592
HMOX1	AR	Public db	PMID: 29226301, 29226302	593
HYOU1	AR	Public db	PMID: 29226301, 29226302	594
ICAM1	association	Public db	PMID: 29921932	595
ICOSLG	unknown	Public db	PMID: 29867916	596
IFI3 5	unknown	Public db	PMID: 27381292, 28236279	597
IFIT1	unknown	Public db	PMID: 28236279, 28768533, 29360040, 29669830, 29928688	598
IFIT2	unknown	Public db	PMID: 27381292, 28236279, 28768533, 29360040, 29669830, 29921932, 29928688	599
IFIT3	unknown	Public db	PMID: 27381292, 28236279, 28768533, 29360040,29669830,29928688	600
IGHM	AR	Public db	PMID: 29226301, 29226302, 29867916	601
IGHMBP2	AR	Public db	PMID: 1714899, 9034313, 11889755, 28202949,29272405	602
IGKC	AR	Public db	PMID: 29226301, 29226302, 29867916	603
IL17RC	AR	Public db	PMID: 29226301, 29226302	604
IL1RN	AR	Public db	PMID: 29226301, 29226302	605
IL3	unknown	Public db	PMID: 29867916	606
IL36RN	AR	Public db	PMID: 29226301, 29226302	607
IL4	unknown	Public db	PMID: 29867916	608
INO80	AR	Public db	PMID: 29226301, 29226302	609
INPP5D	unknown	Public db	PMID: 29921932	610
IRAK 1	XLR	Public db	PMID: 29226301, 29226302	611
IRF2BP2	AD	Public db	PMID: 29226301, 29226302, 29867916	612
ITCH	AR	Public db	PMID: 28295214, 29226301, 29226302, 29867916	613
ITGAM	association	Public db	PMID: 29921932, 22773339	614
ITGB2	AR	Public db	PMID: 29226301, 29226302, 22773339	615
ITPKB	unknown	Public db	PMID: 29921932	616
ITSN1	unknown	Public db	PMID: 29337666, 29599122, 29851086, 29958948	617
JAK 1	ADAR	Public db	PMID: 27648547, 29226301, 29226302	618
KDM6A	XLD XLR	Public db	PMID: 29226301, 29226302	619
KMT2D	AD	Public db	PMID: 29226301, 29226302	620

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
KRAS	AD	Public db	PMID: 29921932	621
LAT	AR	Public db	PMID: 29226301, 29226302	622
LPIN2	AR	Public db	PMID: 29226301, 29226302	623
LRRK2	unknown	Public db	PMID: 26844546, 27830463, 28202666, 28768533, 29321258, 29360040, 29499195, 29669830, 29755459, 29760073,29850629	624
MAP3K14	AR	Public db	PMID: 29226301, 29226302	625
MASP2	AR	Public db	PMID: 29226301, 29226302, 22773339	626
MB21D1	unknown	Public db	PMID: 27648547	627
MBL2	AD	Public db	PMID: 29867916, 22773339	628
MCM4	AR	Public db	PMID: 29226301, 29226302	629
MCM5	AR	Public db	PMID: 29921932	630
MDC1	unknown	Public db	PMID: 29921932	631
MEF2C	AD	Public db	PMID: 29921932	632
MEFV	ADAR	Public db	PMID: 29226301, 29226302	633
MFN1	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	634
MFN2	ADAR	Public db	PMID: 28768533, 29360040, 29669830, 29928688	635
MLH1	ADAR	Public db	PMID: 29867916, 29921932	636
MMP9	association	Public db	PMID: 28681388	637
MOGS	AR	Public db	PMID: 29226301, 29226302	638
MON1A	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	639
MON1B	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	640
MSH2	ADAR	Public db	PMID: 29867916, 29921932	641
MSH5	AR	Public db	PMID: 29867916, 29921932	642
MSH6	AR	Public db	PMID: 29226301, 29226302	643
MVK	ADAR	Public db	PMID: 29226301, 29226302	644
MX1	unknown	Public db	PMID: 27381292, 28236279	645
MX2	unknown	Public db	PMID: 27381292, 28236279	646
MYSM1	AR	Public db	PMID: 29226301, 29226302	647
NBAS	AR	Public db	PMID: 29226301, 29226302	648
NCF1	AR	Public db	PMID: 29226301, 29226302	649
NCF2	AR	Public db	PMID: 29226301, 29226302	650
NCF4	AR	Public db	PMID: 29226301, 29226302	651
NCSTN	AD	Public db	PMID: 29226301, 29226302	652
NFAT5	AD	Public db	PMID: 29226301, 29226302	653
NHP2	AR	Public db	PMID: 29226301, 29226302	654
NLRC4	AD	Public db	PMID: 29226301, 29226302	655
NLRP1	AR	Public db	PMID: 29226301, 29226302	656
NLRP2	unknown	Public db	PMID: 29921932	657
NLRX1	unknown	Public db	PMID: 28295214	658
NOD1	unknown	Public db	PMID: 29921932	659
NOP10	AR	Public db	PMID: 29226301, 29226302	660
NSMCE3	AR	Public db	PMID: 29226301, 29226302	661
OAS1	AR	Public db	PMID: 27381292, 28236279	662
OAS2	unknown	Public db	PMID: 27381292, 28236279	663
OAS3	unknown	Public db	PMID: 27381292, 28236279	664
OASL	unknown	Public db	PMID: 27381292, 28236279	665
ORC4	AR	Public db	PMID: 29867916, 29921932	666
OTULIN	AR	Public db	PMID: 29226301, 29226302	667
PARN	ADAR	Public db	PMID: 26810774, 29226301, 29226302	668

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
PCCA	AR	Public db	PMID: 29921932	669
PCCB	AR	Public db	PMID: 29921932	670
PDCD1	association	Public db	PMID: 29867916	671
PDCD 1LG2	unknown	Public db	PMID: 29867916	672
PEPD	AR	Public db	PMID: 29226301, 29226302	673
PINK 1	AR	Public db	PMID: 28768533, 29360040, 29669830, 29928688	674
PLAU	AD	Public db	PMID: 28681388	675
PLAUR	unknown	Public db	PMID: 28681388	676
PLCG1	unknown	Public db	PMID: 29921932	677
PLD1	AR	Public db	PMID: 29921932	678
PLEKHM1	AR	Public db	PMID: 29226301, 29226302	679
PLK1	unknown	Public db	PMID: 28295214	680
PLXNB1	unknown	Public db	PMID: 29921932	681
PMM2	AR	Public db	PMID: 19862844, 24474243, 27415628	682
POLE2	AR	Public db	PMID: 29226301, 29226302	683
PPM1A	unknown	Public db	PMID: 28295214	684
PRKN	AR	Public db	PMID: 28768533, 29360040, 29669830, 29928688	685
PRRC2A	unknown	Public db	PMID: 29921932	686
PSEN1	AD	Public db	PMID: 29226301, 29226302	687
PSENEN	AD	Public db	PMID: 29226301, 29226302	688
PSMA7	unknown	Public db	PMID: 28295214	689
RAB5A	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	690
RAB5B	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	691
RAB5C	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	692
RAD50	association	Public db	PMID: 29867916, 29921932	693
RANBP2	AD	Public db	PMID: 29226301, 29226302	694
RASGRP1	AR	Public db	PMID: 29226301, 29226302	695
RELA	unknown	Public db	PMID: 27042682, 27648547	696
RELB	AR	Public db	PMID: 29226301, 29226302	697
RHOH	AR	Public db	PMID: 29226301, 29226302	698
RLTPR	AR	Public db	PMID: 29226301, 29226302	699
RNF125	AD	Public db	PMID: 28295214	700
RORC	AR	Public db	PMID: 29226301, 29226302	701
RPSA	AD	Public db	PMID: 29226301, 29226302	702
RSAD2	unknown	Public db	PMID: 27381292, 28236279, 28768533, 29360040,29669830,29928688	703
SAMD9	AD	Public db	PMID: 29226301, 29226302	704
SAMD9L	AD	Public db	PMID: 29226301, 29226302	705
SEMA3E	AD	Public db	PMID: 29226301, 29226302	706
SERPINA1	AR	Public db	PMID: 29867916	707
SERPINB2	unknown	Public db	PMID: 28681388	708
SERPING1	AD	Public db	PMID: 29226301, 29226302, 22773339	709
SH3BP2	AD	Public db	PMID: 29226301, 29226302	710
SLC29A3	AR	Public db	PMID: 29226301, 29226302	711
SLC35C1	AR	Public db	PMID: 29226301, 29226302	712
SLC7A7	AR	Public db	PMID: 10655553, 20301535	713
SLC9A1	AR	Public db	PMID: 29921932	714
SMARCAL1	AR	Public db	PMID: 29226301, 29226302	715
SMARCD2	AR	Public db	PMID: 29226301, 29226302	716
SMC3	AD	Public db	PMID: 29921932	717

RefSeq Gene Symbol	Disease Model	Gene Source	Source Annotation	Gene Number (GN)
SMURF2	unknown	Public db	PMID: 28295214	718
SRP54	unknown	Public db	PMID: 29226301, 29226302	719
STN1	AR	Public db	PMID: 29226301, 29226302	720
TBC1D15	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	721
TBC1D17	unknown	Public db	PMID: 28768533, 29360040, 29669830, 29928688	722
TCF3	AD	Public db	PMID: 29226301, 29226302	723
TCN2	AR	Public db	PMID: 29226301, 29226302	724
TEK	AD	Public db	PMID: 29921932	725
TERC	AD	Public db	PMID: 29226301, 29226302	726
TERT	ADAR	Public db	PMID: 26810774, 29226301, 29226302	727
TFPI	unknown	Public db	PMID: 28681388	728
TFRC	AR	Public db	PMID: 29226301, 29226302	729
THBD	AD	Public db	PMID: 28681388, 29226301, 29226302	730
THBS1	unknown	Public db	PMID: 29921932	731
TINF2	AD	Public db	PMID: 26810774, 29226301, 29226302	732
TIRAP	AR	Public db	PMID: 29226301, 29226302	733
TMC6	AR	Public db	PMID: 29226301, 29226302	734
TMC8	AR	Public db	PMID: 29226301, 29226302	735
TNFRSF17	unknown	Public db	PMID: 29867916	736
TNFRSF1A	AD	Public db	PMID: 29226301, 29226302	737
TNFSF10	unknown	Public db	PMID: 29867916	738
TNFSF13	unknown	Public db	PMID: 29867916	739
TNFSF13B	unknown	Public db	PMID: 29867916	740
TNIP1	unknown	Public db	PMID: 29921932	741
TP53AIP1	unknown	Public db	PMID: 29921932	742
TPP1	ADAR	Public db	PMID: 29226301, 29226302	743
TPP2	AR	Public db	PMID: 29226301, 29226302	744
TRAC	AR	Public db	PMID: 29226301, 29226302	745
TRAF3IP2	AR	Public db	PMID: 29226301, 29226302	746
TRIM25	unknown	Public db	PMID: 28295214	747
TRIM37	AR	Public db	PMID: 29921932	748
TTC37	AR	Public db	PMID: 29226301, 29226302	749
UBD	unknown	Public db	PMID: 28295214	750
USB1	AR	Public db	PMID: 29226301, 29226302	751
USP15	unknown	Public db	PMID: 28295214	752
USP21	unknown	Public db	PMID: 28295214	753
USP25	unknown	Public db	PMID: 28295214	754
USP3	unknown	Public db	PMID: 28295214	755
VAV1	unknown	Public db	PMID: 29867916, 29921932	756
VDR	ADAR	Public db	PMID: 29867916	757
VEGFA	association	Public db	PMID: 28681388	758
WASHC5	ADAR	Public db	PMID: 29921932	759
WDR1	AR	Public db	PMID: 29226301, 29226302	760
WRAP53	AR	Public db	PMID: 29226301, 29226302	761
XAF1	unknown	Public db	PMID: 27381292, 28236279	762

Table 31 represents a non-redundant list of 270 genes involved in the immune system that are not listed in **Table 6.**

Table 32: SEQ ID 2300-2893, transcript variants for **Table 31** genes

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
ACD	NM 001082486	Homo sapiens adrenocortical dysplasia homolog (mouse) (ACD), transcript variant 1, mRNA.	2300
ACD	NM 001082487	Homo sapiens adrenocortical dysplasia homolog (mouse) (ACD), transcript variant 3, mRNA.	2301
ACD	NM 022914	Homo sapiens adrenocortical dysplasia homolog (mouse) (ACD), transcript variant 2, mRNA.	2302
ACTB	NM 001101	Homo sapiens actin, beta (ACTB), mRNA.	2303
ACTN4	NM 004924	Homo sapiens actinin, alpha 4 (ACTN4), mRNA.	2304
ADA2	NM 001282225.1	Homo sapiens adenosine deaminase 2, transcript variant 3	2305
ADA2	NM 001282226.1	Homo sapiens adenosine deaminase 2, transcript variant 4	2306
ADA2	NM 001282227.1	Homo sapiens adenosine deaminase 2, transcript variant 5	2307
ADA2	NM 001282228.1	Homo sapiens adenosine deaminase 2, transcript variant 6	2308
ADA2	NM 001282229.1	Homo sapiens adenosine deaminase 2, transcript variant 7	2309
ADA2	NM 177405.2	Homo sapiens adenosine deaminase 2, transcript variant 2	2310
ADAM17	NM 003183	Homo sapiens ADAM metallopeptidase domain 17 (ADAM17), mRNA.	2311
ADGRL2	NM 001297704.2	Homo sapiens adhesion G protein-coupled receptor L2, transcript variant 2	2312
ADGRL2	NM 001297705.2	Homo sapiens adhesion G protein-coupled receptor L2, transcript variant 3	2313
ADGRL2	NM 001297706.2	Homo sapiens adhesion G protein-coupled receptor L2, transcript variant 4	2314
ADGRL2	NM 001330645.2	Homo sapiens adhesion G protein-coupled receptor L2, transcript variant 5	2315
ADGRL2	NM 001350698.1	Homo sapiens adhesion G protein-coupled receptor L2, transcript variant 6	2316
ADGRL2	NM 001350699.1	Homo sapiens adhesion G protein-coupled receptor L2, transcript variant 7	2317
ADGRL2	NM 012302.4	Homo sapiens adhesion G protein-coupled receptor L2, transcript variant 1	2318
AIRE	NM 000383	Homo sapiens autoimmune regulator (AIRE), mRNA.	2319
ANP32B	NM 006401	Homo sapiens acidic (leucine-rich) nuclear phosphoprotein 32 family, member B (ANP32B), mRNA.	2320
AP1S3	NM 001039569	Homo sapiens adaptor-related protein complex 1, sigma 3 subunit (AP1S3), mRNA.	2321
ARPC1B	NM 005720	Homo sapiens actin related protein 2/3 complex, subunit 1B, 41kDa (ARPC1B), mRNA.	2322
ATG12	NM 004707	Homo sapiens autophagy related 12 (ATG12), transcript variant 1, mRNA.	2323
ATG12	NR 033362	Homo sapiens autophagy related 12 (ATG12), transcript variant 2, non-coding RNA.	2324
ATG12	NR 033363	Homo sapiens autophagy related 12 (ATG12), transcript variant 3, non-coding RNA.	2325
ATG12	NR 073603	Homo sapiens autophagy related 12 (ATG12), transcript variant 4, non-coding RNA.	2326
ATG12	NR 073604	Homo sapiens autophagy related 12 (ATG12), transcript variant 5, non-coding RNA.	2327
ATG12	NR 073605	Homo sapiens autophagy related 12 (ATG12), transcript variant 5, non-coding RNA.	2328
ATG16L1	NM 001190266	Homo sapiens autophagy related 16-like 1 (S. cerevisiae) (ATG16L1), transcript variant 4, mRNA.	2329
ATG16L1	NM 001190267	Homo sapiens autophagy related 16-like 1 (S. cerevisiae) (ATG16L1), transcript variant 5, mRNA.	2330
ATG16L1	NM 017974	Homo sapiens autophagy related 16-like 1 (S. cerevisiae) (ATG16L1), transcript variant 2, mRNA.	2331

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
ATG16L1	NM 030803	Homo sapiens autophagy related 16-like 1 (S. cerevisiae) (ATG16L1), transcript variant 1, mRNA.	2332
ATG16L1	NM 198890	Homo sapiens autophagy related 16-like 1 (S. cerevisiae) (ATG16L1), transcript variant 3, mRNA.	2333
ATG5	NM 004849	Homo sapiens autophagy related 5 (ATG5), mRNA.	2334
ATG7	NM 001136031	Homo sapiens autophagy related 7 (ATG7), transcript variant 2, mRNA.	2335
ATG7	NM 001144912	Homo sapiens autophagy related 7 (ATG7), transcript variant 3, mRNA.	2336
ATG7	NM 006395	Homo sapiens autophagy related 7 (ATG7), transcript variant 1, mRNA.	2337
ATG9A	NM 0010771 98	Homo sapiens autophagy related 9A (ATG9A), transcript variant 1, mRNA.	2338
ATG9A	NM 024085	Homo sapiens autophagy related 9A (ATG9A), transcript variant 2, mRNA.	2339
ATP6AP1	NM 001183	Homo sapiens ATPase, H+ transporting, lysosomal accessory protein 1 (ATP6AP1), mRNA.	2340
B2M	NM 004048	Homo sapiens beta-2-microglobulin (B2M), mRNA.	2341
BCL11B	NM 022898	Homo sapiens B-cell CLL/lymphoma 11B (zinc finger protein) (BCL11B), transcript variant 2, mRNA.	2342
BCL11B	NM 138576	Homo sapiens B-cell CLL/lymphoma 11B (zinc finger protein) (BCL11B), transcript variant 1, mRNA.	2343
BCL2	NM 000633	Homo sapiens B-cell CLL/lymphoma 2 (BCL2), transcript variant alpha, mRNA.	2344
BCL2	NM 000657	Homo sapiens B-cell CLL/lymphoma 2 (BCL2), transcript variant beta, mRNA.	2345
BLK	NM 001715	Homo sapiens B lymphoid tyrosine kinase (BLK), mRNA.	2346
BRD4	NM 014299	Homo sapiens bromodomain containing 4 (BRD4), transcript variant short, mRNA.	2347
BRD4	NM 058243	Homo sapiens bromodomain containing 4 (BRD4), transcript variant long, mRNA.	2348
BTLA	NM 001085357	Homo sapiens B and T lymphocyte associated (BTLA), transcript variant 2, mRNA.	2349
BTLA	NM 181780	Homo sapiens B and T lymphocyte associated (BTLA), transcript variant 1, mRNA.	2350
C1R	NM 001733	Homo sapiens complement component 1, r subcomponent (C1R), mRNA.	2351
C1S	NM 001734	Homo sapiens complement component 1, s subcomponent (C1S), transcript variant 2, mRNA.	2352
C1S	NM 201442	Homo sapiens complement component 1, s subcomponent (C1S), transcript variant 1, mRNA.	2353
C2	NM 000063	Homo sapiens complement component 2 (C2), transcript variant 1, mRNA.	2354
C2	NM 001145903	Homo sapiens complement component 2 (C2), transcript variant 2, mRNA.	2355
C2	NM 001178063	Homo sapiens complement component 2 (C2), transcript variant 3, mRNA.	2356
C2	NR 073063	Homo sapiens complement component 2 (C2), transcript variant 4, non-coding RNA.	2357
C3	NM 000064	Homo sapiens complement component 3 (C3), mRNA.	2358
C4A	NM 001252204	Homo sapiens complement component 4A (Rodgers blood group) (C4A), transcript variant 2, mRNA.	2359
C4A	NM 007293	Homo sapiens complement component 4A (Rodgers blood group) (C4A), transcript variant 1, mRNA.	2360
C4B	NM 001002029	Homo sapiens complement component 4B (Chido blood group) (C4B), mRNA.	2361
C5	NM 001735	Homo sapiens complement component 5 (C5), mRNA.	2362
C6	NM 000065	Homo sapiens complement component 6 (C6), transcript variant 1, mRNA.	2363
C6	NM 001115131	Homo sapiens complement component 6 (C6), transcript variant 2, mRNA.	2364
C7	NM 000587	Homo sapiens complement component 7 (C7), mRNA.	2365
C8A	NM 000562	Homo sapiens complement component 8, alpha polypeptide (C8A), mRNA.	2366
C8B	NM 000066	Homo sapiens complement component 8, beta polypeptide (C8B), transcript variant 1, mRNA.	2367
C8G	NM 000606	Homo sapiens complement component 8, gamma polypeptide (C8G), mRNA.	2368

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
C9	NM 001737	Homo sapiens complement component 9 (C9), mRNA.	2369
CAD	NM 004341	Homo sapiens carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase (CAD), mRNA.	2370
CAMI,G	NM 001745	Homo sapiens calcium modulating ligand (CAMLG), mRNA.	2371
CARD14	NM 001257970	Homo sapiens caspase recruitment domain family, member 14 (CARD14), transcript variant 3, mRNA.	2372
CARD14	NM 024110	Homo sapiens caspase recruitment domain family, member 14 (CARD14), transcript variant 1, mRNA.	2373
CARD14	NM 052819	Homo sapiens caspase recruitment domain family, member 14 (CARD 14), transcript variant 2, mRNA.	2374
CARD14	NR 047566	Homo sapiens caspase recruitment domain family, member 14 (CARD 14), transcript variant 4, non-coding RNA.	2375
CASP10	NM 001206524	Homo sapiens caspase 10, apoptosis-related cysteine peptidase (CASP10), transcript variant 6, mRNA.	2376
CASP10	NM 001206542	Homo sapiens caspase 10, apoptosis-related cysteine peptidase (CASP10), transcript variant 5, mRNA.	2377
CASP10	NM 001230	Homo sapiens caspase 10, apoptosis-related cysteine peptidase (CASP10), transcript variant 3, mRNA.	2378
CASP10	NM 032974	Homo sapiens caspase 10, apoptosis-related cysteine peptidase (CASP10), transcript variant 2, mRNA.	2379
CASP10	NM 032976	Homo sapiens caspase 10, apoptosis-related cysteine peptidase (CASP10), transcript variant 4, mRNA.	2380
CASP10	NM 032977	Homo sapiens caspase 10, apoptosis-related cysteine peptidase (CASP10), transcript variant 1, mRNA.	2381
CAV1	NM 001172895	Homo sapiens caveolin 1, caveolae protein, 22kDa (CAV1), transcript variant 2, mRNA.	2382
CAV1	NM 001172896	Homo sapiens caveolin 1, caveolae protein, 22kDa (CAV1), transcript variant 3, mRNA.	2383
CAV1	NM 001172897	Homo sapiens caveolin 1, caveolae protein, 22kDa (CAV1), transcript variant 4, mRNA.	2384
CAV1	NM 001753	Homo sapiens caveolin 1, caveolae protein, 22kDa (CAV1), transcript variant 1, mRNA.	2385
CCBE1	NM 133459	Homo sapiens collagen and calcium binding EGF domains 1 (CCBE1), mRNA.	2386
CCDC22	NM 014008	Homo sapiens coiled-coil domain containing 22 (CCDC22), mRNA.	2387
CCZ1	NM 015622	Homo sapiens CCZ1 vacuolar protein trafficking and biogenesis associated homolog (S. cerevisiae) (CCZ1), mRNA.	2388
CD22	NM 001185099	Homo sapiens CD22 molecule (CD22), transcript variant 2, mRNA.	2389
CD22	NM 001185100	Homo sapiens CD22 molecule (CD22), transcript variant 3, mRNA.	2390
CD22	NM 001185101	Homo sapiens CD22 molecule (CD22), transcript variant 4, mRNA.	2391
CD22	NM 00 1771	Homo sapiens CD22 molecule (CD22), transcript variant 1, mRNA.	2392
CD274	NM 001267706	Homo sapiens CD274 molecule (CD274), transcript variant 2, mRNA.	2393
CD274	NM 014143	Homo sapiens CD274 molecule (CD274), transcript variant 1, mRNA.	2394
CD274	NR 052005	Homo sapiens CD274 molecule (CD274), transcript variant 3, non-coding RNA.	2395
CD276	NM 001024736	Homo sapiens CD276 molecule (CD276), transcript variant 1, mRNA.	2396
CD276	NM 025240	Homo sapiens CD276 molecule (CD276), transcript variant 2, mRNA.	2397
CD36	NM 000072	Homo sapiens CD36 molecule (thrombospondin receptor) (CD36), transcript variant 3, mRNA.	2398
CD36	NM 001001547	Homo sapiens CD36 molecule (thrombospondin receptor) (CD36), transcript variant 2, mRNA.	2399
CD36	NM 001001548	Homo sapiens CD36 molecule (thrombospondin receptor) (CD36), transcript variant 1, mRNA.	2400
CD36	NM 001127443	Homo sapiens CD36 molecule (thrombospondin receptor) (CD36), transcript variant 4, mRNA.	2401
CD36	NM 001127444	Homo sapiens CD36 molecule (thrombospondin receptor) (CD36), transcript	2402

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		variant 5, mRNA.	
CD37	NM 001040031	Homo sapiens CD37 molecule (CD37), transcript variant 2, mRNA.	2403
CD37	NM 001774	Homo sapiens CD37 molecule (CD37), transcript variant 1, mRNA.	2404
CD38	NM 001775	Homo sapiens CD38 molecule (CD38), mRNA.	2405
CD46	NM 002389	Homo sapiens CD46 molecule, complement regulatory protein (CD46), transcript variant a, mRNA.	2406
CD46	NM 153826	Homo sapiens CD46 molecule, complement regulatory protein (CD46), transcript variant d, mRNA.	2407
CD46	NM 172350	Homo sapiens CD46 molecule, complement regulatory protein (CD46), transcript variant n, mRNA.	2408
CD46	NM 172351	Homo sapiens CD46 molecule, complement regulatory protein (CD46), transcript variant c, mRNA.	2409
CD46	NM 172352	Homo sapiens CD46 molecule, complement regulatory protein (CD46), transcript variant e, mRNA.	2410
CD46	NM 172353	Homo sapiens CD46 molecule, complement regulatory protein (CD46), transcript variant f, mRNA.	2411
CD46	NM 172359	Homo sapiens CD46 molecule, complement regulatory protein (CD46), transcript variant b, mRNA.	2412
CD46	NM 172361	Homo sapiens CD46 molecule, complement regulatory protein (CD46), transcript variant 1, mRNA.	2413
CD5	NM 014207	Homo sapiens CD5 molecule (CD5), mRNA.	2414
CD70	NM 001252	Homo sapiens CD70 molecule (CD70), mRNA.	2415
CD72	NM 001782	Homo sapiens CD72 molecule (CD72), mRNA.	2416
CD74	NM 001025158	Homo sapiens CD74 molecule, major histocompatibility complex, class II invariant chain (CD74), transcript variant 3, mRNA.	2417
CD74	NM 001025159	Homo sapiens CD74 molecule, major histocompatibility complex, class II invariant chain (CD74), transcript variant 1, mRNA.	2418
CD74	NM 004355	Homo sapiens CD74 molecule, major histocompatibility complex, class II invariant chain (CD74), transcript variant 2, mRNA.	2419
CD84	NM 001184879	Homo sapiens CD84 molecule (CD84), transcript variant 1, mRNA.	2420
CD84	NM 001184881	Homo sapiens CD84 molecule (CD84), transcript variant 3, mRNA.	2421
CD84	NM 001184882	Homo sapiens CD84 molecule (CD84), transcript variant 4, mRNA.	2422
CD84	NM 003874	Homo sapiens CD84 molecule (CD84), transcript variant 2, mRNA.	2423
CD93	NM 012072	Homo sapiens CD93 molecule (CD93), mRNA.	2424
CEBPE	NM 001805	Homo sapiens CCAAT/enhancer binding protein (C/EBP), epsilon (CEBPE), mRNA.	2425
CFB	NM 001710	Homo sapiens complement factor B (CFB), mRNA.	2426
CFD	NM 001928	Homo sapiens complement factor D (adipsin) (CFD), mRNA.	2427
CFH	NM 000186	Homo sapiens complement factor H (CFH), transcript variant 1, mRNA.	2428
CFH	NM 001014975	Homo sapiens complement factor H (CFH), transcript variant 2, mRNA.	2429
CFHR1	NM 002113	Homo sapiens complement factor H-related 1 (CFHR1), mRNA.	2430
CFHR2	NM 005666	Homo sapiens complement factor H-related 2 (CFHR2), mRNA.	2431
CFHR3	NM 001166624	Homo sapiens complement factor H-related 3 (CFHR3), transcript variant 2, mRNA.	2432
CFHR3	NM 021023	Homo sapiens complement factor H-related 3 (CFHR3), transcript variant 1, mRNA.	2433
CFHR4	NM 001201550	Homo sapiens complement factor H-related 4 (CFHR4), transcript variant 1, mRNA.	2434
CFHR4	NM 001201551	Homo sapiens complement factor H-related 4 (CFHR4), transcript variant 2, mRNA.	2435
CFHR4	NM 006684	Homo sapiens complement factor H-related 4 (CFHR4), transcript variant 3, mRNA.	2436
CFHR5	NM 030787	Homo sapiens complement factor H-related 5 (CFHR5), mRNA.	2437
CFI	NM 000204	Homo sapiens complement factor I (CFI), mRNA.	2438
CFP	NM 001145252	Homo sapiens complement factor properdin (CFP), transcript variant 2,	2439

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		mRNA.	
CFP	NM 002621	Homo sapiens complement factor properdin (CFP), transcript variant 1, mRNA.	2440
CFTR	NM 000492	Homo sapiens cystic fibrosis transmembrane conductance regulator (ATP-binding cassette sub-family C, member 7) (CFTR), mRNA.	2441
CHD2	NM 001042572	Homo sapiens chromodomain helicase DNA binding protein 2 (CHD2), transcript variant 2, mRNA.	2442
CHD2	NM 001271	Homo sapiens chromodomain helicase DNA binding protein 2 (CHD2), transcript variant 1, mRNA.	2443
CLEC16A	NM 001243403	Homo sapiens C-type lectin domain family 16, member A (CLEC16A), transcript variant 2, mRNA.	2444
CLEC16A	NM 015226	Homo sapiens C-type lectin domain family 16, member A (CLEC16A), transcript variant 1, mRNA.	2445
CLPB	NM 001258392	Homo sapiens ClpB caseinolytic peptidase B homolog (E. coli) (CLPB), transcript variant 2, mRNA.	2446
CLPB	NM 001258393	Homo sapiens ClpB caseinolytic peptidase B homolog (E. coli) (CLPB), transcript variant 3, mRNA.	2447
CLPB	NM 001258394	Homo sapiens ClpB caseinolytic peptidase B homolog (E. coli) (CLPB), transcript variant 4, mRNA.	2448
CLPB	NM 030813	Homo sapiens ClpB caseinolytic peptidase B homolog (E. coli) (CLPB), transcript variant 1, mRNA.	2449
COPA	NM 001098398	Homo sapiens coatomer protein complex, subunit alpha (COPA), transcript variant 1, mRNA.	2450
COPA	NM 004371	Homo sapiens coatomer protein complex, subunit alpha (COPA), transcript variant 2, mRNA.	2451
CSF2RA	NM 001161529	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 7, mRNA.	2452
CSF2RA	NM 001161530	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 8, mRNA.	2453
CSF2RA	NM 001161531	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 9, mRNA.	2454
CSF2RA	NM 001161532	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 10, mRNA.	2455
CSF2RA	NM 006140	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 1, mRNA.	2456
CSF2RA	NM 172245	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 2, mRNA.	2457
CSF2RA	NM 172246	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 3, mRNA.	2458
CSF2RA	NM 172247	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 4, mRNA.	2459
CSF2RA	NM 172249	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 6, mRNA.	2460
CSF2RA	NR 027760	Homo sapiens colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage) (CSF2RA), transcript variant 11, non-coding RNA.	2461
CSF2RB	NM 000395	Homo sapiens colony stimulating factor 2 receptor, beta, low-affinity (granulocyte-macrophage) (CSF2RB), mRNA.	2462
CTC1	NM 025099	Homo sapiens CTS telomere maintenance complex component 1 (CTC1), transcript variant 1, mRNA.	2463
CTC1	NR 046431	Homo sapiens CTS telomere maintenance complex component 1 (CTC1), transcript variant 2, non-coding RNA.	2464
CXCL1	NM 001511	Homo sapiens chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha) (CXCL1), transcript variant 1, mRNA.	2465
CXCL1	NR 046035	Homo sapiens chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha) (CXCL1), transcript variant 2, non-coding RNA.	2466
CXCL10	NM 001565	Homo sapiens chemokine (C-X-C motif) ligand 10 (CXCL10), mRNA.	2467
CXCL5	NM 002994	Homo sapiens chemokine (C-X-C motif) ligand 5 (CXCL5), mRNA.	2468
CXCL8	NM 000584.3	Homo sapiens C-X-C motif chemokine ligand 8, transcript variant 1	2469

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
CXCL8	NM 001354840.1	Homo sapiens C-X-C motif chemokine ligand 8, transcript variant 2	2470
CXCR3	NM 001142797	Homo sapiens chemokine (C-X-C motif) receptor 3 (CXCR3), transcript variant 2, mRNA.	2471
CXCR3	NM 001504	Homo sapiens chemokine (C-X-C motif) receptor 3 (CXCR3), transcript variant 1, mRNA.	2472
CYBA	NM 000101	Homo sapiens cytochrome b-245, alpha polypeptide (CYBA), mRNA.	2473
DCLRE1B	NM 022836	Homo sapiens DNA cross-link repair 1B (DCLRE1B), mRNA.	2474
DNAJC21	NM 001012339	Homo sapiens DnaJ (Hsp40) homolog, subfamily C, member 21 (DNAJC21), transcript variant 2, mRNA.	2475
DNAJC21	NM 194283	Homo sapiens DnaJ (Hsp40) homolog, subfamily C, member 21 (DNAJC21), transcript variant 1, mRNA.	2476
DNASE1L3	NM 001256560	Homo sapiens deoxyribonuclease I-like 3 (DNASE1L3), transcript variant 2, mRNA.	2477
DNASE1L3	NM 004944	Homo sapiens deoxyribonuclease I-like 3 (DNASE1L3), transcript variant 1, mRNA.	2478
DNASE2	NM 001375	Homo sapiens deoxyribonuclease II, lysosomal (DNASE2), mRNA.	2479
EBF1	NM 024007	Homo sapiens early B-cell factor 1 (EBF1), mRNA.	2480
EGF	NM 001178130	Homo sapiens epidermal growth factor (EGF), transcript variant 2, mRNA.	2481
EGF	NM 001178131	Homo sapiens epidermal growth factor (EGF), transcript variant 3, mRNA.	2482
EGF	NM 001963	Homo sapiens epidermal growth factor (EGF), transcript variant 1, mRNA.	2483
ERCC6L2	NM 001010895	Homo sapiens excision repair cross-complementing rodent repair deficiency, complementation group 6-like 2 (ERCC6L2), mRNA.	2484
EXTL3	NM 001440	Homo sapiens exostosin-like glycosyltransferase 3 (EXTL3), transcript variant 1, mRNA.	2485
EXTL3	NR 073468	Homo sapiens exostosin-like glycosyltransferase 3 (EXTL3), transcript variant 2, non-coding RNA.	2486
EXTL3	NR 073469	Homo sapiens exostosin-like glycosyltransferase 3 (EXTL3), transcript variant 3, non-coding RNA.	2487
FAAP24	NM 001300978.1	Homo sapiens Fanconi anemia core complex associated protein 24, transcript variant 2	2488
FAAP24	NM 152266.4	Homo sapiens Fanconi anemia core complex associated protein 24, transcript variant 1	2489
FADD	NM 003824	Homo sapiens Fas (TNFRSF6)-associated via death domain (FADD), mRNA.	2490
FAT4	NM 024582	Homo sapiens FAT atypical cadherin 4 (FAT4), mRNA.	2491
FCER2	NM 001207019	Homo sapiens Fc fragment of IgE, low affinity II, receptor for (CD23) (FCER2), transcript variant 2, mRNA.	2492
FCER2	NM 001220500	Homo sapiens Fc fragment of IgE, low affinity II, receptor for (CD23) (FCER2), transcript variant 3, mRNA.	2493
FCER2	NM 002002	Homo sapiens Fc fragment of IgE, low affinity II, receptor for (CD23) (FCER2), transcript variant 1, mRNA.	2494
FERMT3	NM 031471	Homo sapiens fermitin family member 3 (FERMT3), transcript variant URP2SF, mRNA.	2495
FERMT3	NM 178443	Homo sapiens fermitin family member 3 (FERMT3), transcript variant URP2LF, mRNA.	2496
FIS1	NM 016068	Homo sapiens fission 1 (mitochondrial outer membrane) homolog (S. cerevisiae) (FIS1), mRNA.	2497
G6PD	NM 000402	Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), transcript variant 1, mRNA.	2498
G6PD	NM 001042351	Homo sapiens glucose-6-phosphate dehydrogenase (G6PD), transcript variant 2, mRNA.	2499
GINS1	NM 021067	Homo sapiens GINS complex subunit 1 (Psf1 homolog) (GINS1), mRNA.	2500
HERC5	NM 016323	Homo sapiens HECT and RLD domain containing E3 ubiquitin protein ligase 5 (HERC5), mRNA.	2501
HERC6	NM 001165136	Homo sapiens HECT and RLD domain containing E3 ubiquitin protein ligase family member 6 (HERC6), transcript variant 2, mRNA.	2502

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
HERC6	NM 017912	Homo sapiens HECT and RLD domain containing E3 ubiquitin protein ligase family member 6 (HERC6), transcript variant 1, mRNA.	2503
HMGB1	NM 002128	Homo sapiens high mobility group box 1 (HMGB 1), mRNA.	2504
HMOX1	NM 002133	Homo sapiens heme oxygenase (decycling) 1 (HMOX1), mRNA.	2505
HYOU1	NM 001130991	Homo sapiens hypoxia up-regulated 1 (HYOU1), transcript variant 2, mRNA.	2506
HYOU1	NM 006389	Homo sapiens hypoxia up-regulated 1 (HYOU1), transcript variant 1, mRNA.	2507
ICAM1	NM 000201	Homo sapiens intercellular adhesion molecule 1 (ICAM1), mRNA.	2508
ICOSLG	NM 015259	Homo sapiens inducible T-cell co-stimulator ligand (ICOSLG), mRNA.	2509
IFI35	NM 005533	Homo sapiens interferon-induced protein 35 (IFI35), mRNA.	2510
IFIT1	NM 001270927	Homo sapiens interferon-induced protein with tetratricopeptide repeats 1 (IFIT1), transcript variant 2, mRNA.	2511
IFIT1	NM 001270928	Homo sapiens interferon-induced protein with tetratricopeptide repeats 1 (IFIT1), transcript variant 3, mRNA.	2512
IFIT1	NM 001270929	Homo sapiens interferon-induced protein with tetratricopeptide repeats 1 (IFIT1), transcript variant 4, mRNA.	2513
IFIT1	NM 001270930	Homo sapiens interferon-induced protein with tetratricopeptide repeats 1 (IFIT1), transcript variant 5, mRNA.	2514
IFIT1	NM 001548	Homo sapiens interferon-induced protein with tetratricopeptide repeats 1 (IFIT1), transcript variant 1, mRNA.	2515
IFIT2	NM 001547	Homo sapiens interferon-induced protein with tetratricopeptide repeats 2 (IFIT2), mRNA.	2516
IFIT3	NM 001031683	Homo sapiens interferon-induced protein with tetratricopeptide repeats 3 (IFIT3), transcript variant 2, mRNA.	2517
IFIT3	NM 001549	Homo sapiens interferon-induced protein with tetratricopeptide repeats 3 (IFIT3), transcript variant 1, mRNA.	2518
IGHMBP2	NM 002180	Homo sapiens immunoglobulin mu binding protein 2 (IGHMBP2), mRNA.	2519
IL17RC	NM 001203263	Homo sapiens interleukin 17 receptor C (IL17RC), transcript variant 4, mRNA.	2520
IL17RC	NM 001203264	Homo sapiens interleukin 17 receptor C (IL17RC), transcript variant 5, mRNA.	2521
IL17RC	NM 001203265	Homo sapiens interleukin 17 receptor C (IL17RC), transcript variant 6, mRNA.	2522
IL17RC	NM 032732	Homo sapiens interleukin 17 receptor C (IL17RC), transcript variant 3, mRNA.	2523
IL17RC	NM 153460	Homo sapiens interleukin 17 receptor C (IL17RC), transcript variant 1, mRNA.	2524
IL17RC	NM 153461	Homo sapiens interleukin 17 receptor C (IL17RC), transcript variant 2, mRNA.	2525
IL17RC	NR 037807	Homo sapiens interleukin 17 receptor C (IL17RC), transcript variant 7, non-coding RNA.	2526
IL 1RN	NM 000577	Homo sapiens interleukin 1 receptor antagonist (IL1RN), transcript variant 3, mRNA.	2527
IL1RN	NM 173841	Homo sapiens interleukin 1 receptor antagonist (IL1RN), transcript variant 2, mRNA.	2528
IL1RN	NM 173842	Homo sapiens interleukin 1 receptor antagonist (IL1RN), transcript variant 1, mRNA.	2529
IL1RN	NM 173843	Homo sapiens interleukin 1 receptor antagonist (IL1RN), transcript variant 4, mRNA.	2530
IL3	NM 000588	Homo sapiens interleukin 3 (colony-stimulating factor, multiple) (IL3), mRNA.	2531
IL36RN	NM 012275	Homo sapiens interleukin 36 receptor antagonist (IL36RN), transcript variant 1, mRNA.	2532
IL36RN	NM 173170	Homo sapiens interleukin 36 receptor antagonist (IL36RN), transcript variant 2, mRNA.	2533
IL4	NM 000589	Homo sapiens interleukin 4 (IL4), transcript variant 1, mRNA.	2534
IL4	NM 172348	Homo sapiens interleukin 4 (IL4), transcript variant 2, mRNA.	2535
INO80	NM 017553	Homo sapiens INO80 complex subunit (INO80), transcript variant 1, mRNA.	2536

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
INPP5D	NM 001017915	Homo sapiens inositol polyphosphate-5-phosphatase, 145kDa (INPP5D), transcript variant 1, mRNA.	2537
INPP5D	NM 005541	Homo sapiens inositol polyphosphate-5-phosphatase, 145kDa (INPP5D), transcript variant 2, mRNA.	2538
IRAK1	NM 001025242	Homo sapiens interleukin-1 receptor-associated kinase 1 (IRAK1), transcript variant 2, mRNA.	2539
IRAK1	NM 001025243	Homo sapiens interleukin-1 receptor-associated kinase 1 (IRAK1), transcript variant 3, mRNA.	2540
IRAK1	NM 001569	Homo sapiens interleukin-1 receptor-associated kinase 1 (IRAK1), transcript variant 1, mRNA.	2541
IRF2BP2	NM 001077397	Homo sapiens interferon regulatory factor 2 binding protein 2 (IRF2BP2), transcript variant 2, mRNA.	2542
IRF2BP2	NM 182972	Homo sapiens interferon regulatory factor 2 binding protein 2 (IRF2BP2), transcript variant 1, mRNA.	2543
ITCH	NM 001257137	Homo sapiens itchy E3 ubiquitin protein ligase (ITCH), transcript variant 1, mRNA.	2544
ITCH	NM 001257138	Homo sapiens itchy E3 ubiquitin protein ligase (ITCH), transcript variant 3, mRNA.	2545
ITCH	NM 031483	Homo sapiens itchy E3 ubiquitin protein ligase (ITCH), transcript variant 2, mRNA.	2546
ITGAM	NM 000632	Homo sapiens integrin, alpha M (complement component 3 receptor 3 subunit) (ITGAM), transcript variant 2, mRNA.	2547
ITGAM	NM 001145808	Homo sapiens integrin, alpha M (complement component 3 receptor 3 subunit) (ITGAM), transcript variant 1, mRNA.	2548
ITGB2	NM 000211	Homo sapiens integrin, beta 2 (complement component 3 receptor 3 and 4 subunit) (ITGB2), transcript variant 1, mRNA.	2549
ITGB2	NM 001127491	Homo sapiens integrin, beta 2 (complement component 3 receptor 3 and 4 subunit) (ITGB2), transcript variant 2, mRNA.	2550
ITPKB	NM 002221	Homo sapiens inositol-trisphosphate 3-kinase B (ITPKB), mRNA.	2551
ITSN1	NM 001001132	Homo sapiens intersectin 1 (SH3 domain protein) (ITSN1), transcript variant 2, mRNA.	2552
ITSN1	NM 003024	Homo sapiens intersectin 1 (SH3 domain protein) (ITSN1), transcript variant 1, mRNA.	2553
JAK1	NM 002227	Homo sapiens Janus kinase 1 (JAK1), mRNA.	2554
KDM6A	NM 021140	Homo sapiens lysine (K)-specific demethylase 6A (KDM6A), mRNA.	2555
KMT2D	NM 003482	Homo sapiens lysine (K)-specific methyltransferase 2D (KMT2D), mRNA.	2556
KRAS	NM 004985	Homo sapiens Kirsten rat sarcoma viral oncogene homolog (KRAS), transcript variant b, mRNA.	2557
KRAS	NM 033360	Homo sapiens Kirsten rat sarcoma viral oncogene homolog (KRAS), transcript variant a, mRNA.	2558
LAT	NM 001014987	Homo sapiens linker for activation of T cells (LAT), transcript variant 2, mRNA.	2559
LAT	NM 001014988	Homo sapiens linker for activation of T cells (LAT), transcript variant 3, mRNA.	2560
LAT	NM 001014989	Homo sapiens linker for activation of T cells (LAT), transcript variant 4, mRNA.	2561
LAT	NM 014387	Homo sapiens linker for activation of T cells (LAT), transcript variant 1, mRNA.	2562
LPIN2	NM 014646	Homo sapiens lipin 2 (LPIN2), mRNA.	2563
LRRK2	NM 198578	Homo sapiens leucine-rich repeat kinase 2 (LRRK2), mRNA.	2564
MAP3K14	NM 003954	Homo sapiens mitogen-activated protein kinase kinase kinase 14 (MAP3K14), mRNA.	2565
MASP2	NM 006610	Homo sapiens mannan-binding lectin serine peptidase 2 (MASP2), transcript variant 1, mRNA.	2566
MASP2	NM 139208	Homo sapiens mannan-binding lectin serine peptidase 2 (MASP2), transcript variant 2, mRNA.	2567
MB21D1	NM 138441	Homo sapiens Mab-21 domain containing 1 (MB21D1), mRNA.	2568

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
MBL2	NM 000242	Homo sapiens mannose-binding lectin (protein C) 2, soluble (MBL2), mRNA.	2569
MCM4	NM 005914	Homo sapiens minichromosome maintenance complex component 4 (MCM4), transcript variant 1, mRNA.	2570
MCM4	NM 182746	Homo sapiens minichromosome maintenance complex component 4 (MCM4), transcript variant 2, mRNA.	2571
MCM5	NM 006739	Homo sapiens minichromosome maintenance complex component 5 (MCM5), mRNA.	2572
MDC1	NM 014641	Homo sapiens mediator of DNA-damage checkpoint 1 (MDC1), mRNA.	2573
MEF2C	NM 001131005	Homo sapiens myocyte enhancer factor 2C (MEF2C), transcript variant 2, mRNA.	2574
MEF2C	NM 001193347	Homo sapiens myocyte enhancer factor 2C (MEF2C), transcript variant 3, mRNA.	2575
MEF2C	NM 001193348	Homo sapiens myocyte enhancer factor 2C (MEF2C), transcript variant 4, mRNA.	2576
MEF2C	NM 001193349	Homo sapiens myocyte enhancer factor 2C (MEF2C), transcript variant 5, mRNA.	2577
MEF2C	NM 001193350	Homo sapiens myocyte enhancer factor 2C (MEF2C), transcript variant 6, mRNA.	2578
MEF2C	NM 002397	Homo sapiens myocyte enhancer factor 2C (MEF2C), transcript variant 1, mRNA.	2579
MEFV	NM 000243	Homo sapiens Mediterranean fever (MEFV), transcript variant 1, mRNA.	2580
MEFV	NM 001198536	Homo sapiens Mediterranean fever (MEFV), transcript variant 2, mRNA.	2581
MFN1	NM 033540	Homo sapiens mitofusin 1 (MFN1), mRNA.	2582
MFN2	NM 001127660	Homo sapiens mitofusin 2 (MFN2), transcript variant 2, mRNA.	2583
MFN2	NM 014874	Homo sapiens mitofusin 2 (MFN2), transcript variant 1, mRNA.	2584
MLH1	NM 000249	Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), transcript variant 1, mRNA.	2585
MLH1	NM 001167617	Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), transcript variant 2, mRNA.	2586
MLH1	NM 001167618	Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), transcript variant 3, mRNA.	2587
MLH1	NM 001167619	Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), transcript variant 4, mRNA.	2588
MLH1	NM 001258271	Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), transcript variant 5, mRNA.	2589
MLH1	NM 001258273	Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), transcript variant 6, mRNA.	2590
MLH1	NM 001258274	Homo sapiens mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli) (MLH1), transcript variant 7, mRNA.	2591
MMP9	NM 004994	Homo sapiens matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase) (MMP9), mRNA.	2592
MOGS	NM 001146158	Homo sapiens mannosyl-oligosaccharide glucosidase (MOGS), transcript variant 2, mRNA.	2593
MOGS	NM 006302	Homo sapiens mannosyl-oligosaccharide glucosidase (MOGS), transcript variant 1, mRNA.	2594
MON1A	NM 001142501	Homo sapiens MON1 secretory trafficking family member A (MON1A), transcript variant 2, mRNA.	2595
MON1A	NM 032355	Homo sapiens MON1 secretory trafficking family member A (MON1A), transcript variant 1, mRNA.	2596
MON1B	NM 014940	Homo sapiens MON1 secretory trafficking family member B (MON1B), mRNA.	2597
MSH2	NM 000251	Homo sapiens mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli) (MSH2), transcript variant 1, mRNA.	2598
MSH2	NM 001258281	Homo sapiens mutS homolog 2, colon cancer, nonpolyposis type 1 (E. coli) (MSH2), transcript variant 2, mRNA.	2599
MSH5	NM 002441	Homo sapiens mutS homolog 5 (E. coli) (MSH5), transcript variant 3, mRNA.	2600
MSH5	NM 025259	Homo sapiens mutS homolog 5 (E. coli) (MSH5), transcript variant 1, mRNA.	2601

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
MSH5	NM 172165	Homo sapiens mutS homolog 5 (E. coli) (MSH5), transcript variant 2, mRNA.	2602
MSH5	NM 172166	Homo sapiens mutS homolog 5 (E. coli) (MSH5), transcript variant 4, mRNA.	2603
MSH6	NM 000179	Homo sapiens mutS homolog 6 (E. coli) (MSH6), transcript variant 1, mRNA.	2604
MVK	NM 000431	Homo sapiens mevalonate kinase (MVK), transcript variant 1, mRNA.	2605
MVK	NM 001114185	Homo sapiens mevalonate kinase (MVK), transcript variant 2, mRNA.	2606
MX1	NM 001144925	Homo sapiens myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) (MX1), transcript variant 1, mRNA.	2607
MX1	NM 001178046	Homo sapiens myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) (MX1), transcript variant 3, mRNA.	2608
MX1	NM 002462	Homo sapiens myxovirus (influenza virus) resistance 1, interferon-inducible protein p78 (mouse) (MX1), transcript variant 2, mRNA.	2609
MX2	NM 002463	Homo sapiens myxovirus (influenza virus) resistance 2 (mouse) (MX2), mRNA.	2610
MYSM1	NM 001085487	Homo sapiens Myb-like, SWIRM and MPN domains 1 (MYSM1), mRNA.	2611
NBAS	NM 015909	Homo sapiens neuroblastoma amplified sequence (NBAS), transcript variant 1, mRNA.	2612
NBAS	NR 052013	Homo sapiens neuroblastoma amplified sequence (NBAS), transcript variant 2, non-coding RNA.	2613
NCF1	NM 000265	Homo sapiens neutrophil cytosolic factor 1 (NCF1), mRNA.	2614
NCF2	NM 000433	Homo sapiens neutrophil cytosolic factor 2 (NCF2), transcript variant 1, mRNA.	2615
NCF2	NM 001127651	Homo sapiens neutrophil cytosolic factor 2 (NCF2), transcript variant 2, mRNA.	2616
NCF2	NM 001190789	Homo sapiens neutrophil cytosolic factor 2 (NCF2), transcript variant 4, mRNA.	2617
NCF2	NM 001190794	Homo sapiens neutrophil cytosolic factor 2 (NCF2), transcript variant 3, mRNA.	2618
NCF4	NM 000631	Homo sapiens neutrophil cytosolic factor 4, 40kDa (NCF4), transcript variant 1, mRNA.	2619
NCF4	NM 013416	Homo sapiens neutrophil cytosolic factor 4, 40kDa (NCF4), transcript variant 2, mRNA.	2620
NCSTN	NM 015331	Homo sapiens nicastrin (NCSTN), mRNA.	2621
NFAT5	NM 001113178	Homo sapiens nuclear factor of activated T-cells 5, tonicity-responsive (NFAT5), transcript variant 6, mRNA.	2622
NFAT5	NM 006599	Homo sapiens nuclear factor of activated T-cells 5, tonicity-responsive (NFAT5), transcript variant 3, mRNA.	2623
NFAT5	NM 138713	Homo sapiens nuclear factor of activated T-cells 5, tonicity-responsive (NFAT5), transcript variant 2, mRNA.	2624
NFAT5	NM 138714	Homo sapiens nuclear factor of activated T-cells 5, tonicity-responsive (NFAT5), transcript variant 1, mRNA.	2625
NFAT5	NM 173214	Homo sapiens nuclear factor of activated T-cells 5, tonicity-responsive (NFAT5), transcript variant 4, mRNA.	2626
NFAT5	NM 173215	Homo sapiens nuclear factor of activated T-cells 5, tonicity-responsive (NFAT5), transcript variant 5, mRNA.	2627
NHP2	NM 001034833	Homo sapiens NHP2 ribonucleoprotein (NHP2), transcript variant 2, mRNA.	2628
NHP2	NM 017838	Homo sapiens NHP2 ribonucleoprotein (NHP2), transcript variant 1, mRNA.	2629
NLRC4	NM 001199138	Homo sapiens NLR family, CARD domain containing 4 (NLRC4), transcript variant 2, mRNA.	2630
NLRC4	NM 001199139	Homo sapiens NLR family, CARD domain containing 4 (NLRC4), transcript variant 3, mRNA.	2631
NLRC4	NM 021209	Homo sapiens NLR family, CARD domain containing 4 (NLRC4), transcript variant 1, mRNA.	2632
NLRP1	NM 001033053	Homo sapiens NLR family, pyrin domain containing 1 (NLRP1), transcript variant 5, mRNA.	2633
NLRP1	NM 014922	Homo sapiens NLR family, pyrin domain containing 1 (NLRP1), transcript variant 2, mRNA.	2634

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
NLRP1	NM 033004	Homo sapiens NLR family, pyrin domain containing 1 (NLRP1), transcript variant 1, mRNA.	2635
NLRP1	NM 033006	Homo sapiens NLR family, pyrin domain containing 1 (NLRP1), transcript variant 3, mRNA.	2636
NLRP1	NM 033007	Homo sapiens NLR family, pyrin domain containing 1 (NLRP1), transcript variant 4, mRNA.	2637
NLRP2	NM 001174081	Homo sapiens NLR family, pyrin domain containing 2 (NLRP2), transcript variant 2, mRNA.	2638
NLRP2	NM 001174082	Homo sapiens NLR family, pyrin domain containing 2 (NLRP2), transcript variant 3, mRNA.	2639
NLRP2	NM 001174083	Homo sapiens NLR family, pyrin domain containing 2 (NLRP2), transcript variant 4, mRNA.	2640
NLRP2	NM 017852	Homo sapiens NLR family, pyrin domain containing 2 (NLRP2), transcript variant 1, mRNA.	2641
NLRX1	NM 024618	Homo sapiens NLR family member X1 (NLRX1), transcript variant 1, mRNA.	2642
NLRX1	NM 170722	Homo sapiens NLR family member X1 (NLRX1), transcript variant 2, mRNA.	2643
NOD1	NM 006092	Homo sapiens nucleotide-binding oligomerization domain containing 1 (NOD1), mRNA.	2644
NOP10	NM 018648	Homo sapiens NOP10 ribonucleoprotein (NOP10), mRNA.	2645
NSMCE3	NM 138704.3	Homo sapiens NSE3 homolog, SMC5-SMC6 complex component	2646
OAS1	NM 001032409	Homo sapiens 2'-5'-oligoadenylate synthetase 1, 40/46kDa (OAS1), transcript variant 3, mRNA.	2647
OAS1	NM 002534	Homo sapiens 2'-5'-oligoadenylate synthetase 1, 40/46kDa (OAS1), transcript variant 2, mRNA.	2648
OAS1	NM 016816	Homo sapiens 2'-5'-oligoadenylate synthetase 1, 40/46kDa (OAS1), transcript variant 1, mRNA.	2649
OAS2	NM 001032731	Homo sapiens 2'-5'-oligoadenylate synthetase 2, 69/71kDa (OAS2), transcript variant 3, mRNA.	2650
OAS2	NM 002535	Homo sapiens 2'-5'-oligoadenylate synthetase 2, 69/71kDa (OAS2), transcript variant 2, mRNA.	2651
OAS2	NM 016817	Homo sapiens 2'-5'-oligoadenylate synthetase 2, 69/71kDa (OAS2), transcript variant 1, mRNA.	2652
OAS3	NM 006187	Homo sapiens 2'-5'-oligoadenylate synthetase 3, 100kDa (OAS3), mRNA.	2653
OASL	NM 001261825	Homo sapiens 2'-5'-oligoadenylate synthetase-like (OASL), transcript variant 3, mRNA.	2654
OASL	NM 003733	Homo sapiens 2'-5'-oligoadenylate synthetase-like (OASL), transcript variant 1, mRNA.	2655
OASL	NM 198213	Homo sapiens 2'-5'-oligoadenylate synthetase-like (OASL), transcript variant 2, mRNA.	2656
ORC4	NM 001190879	Homo sapiens origin recognition complex, subunit 4 (ORC4), transcript variant 4, mRNA.	2657
ORC4	NM 001190881	Homo sapiens origin recognition complex, subunit 4 (ORC4), transcript variant 5, mRNA.	2658
ORC4	NM 001190882	Homo sapiens origin recognition complex, subunit 4 (ORC4), transcript variant 6, mRNA.	2659
ORC4	NM 002552	Homo sapiens origin recognition complex, subunit 4 (ORC4), transcript variant 2, mRNA.	2660
ORC4	NM 181741	Homo sapiens origin recognition complex, subunit 4 (ORC4), transcript variant 1, mRNA.	2661
ORC4	NM 181742	Homo sapiens origin recognition complex, subunit 4 (ORC4), transcript variant 3, mRNA.	2662
OTULIN	NM 138348.5	Homo sapiens OTU deubiquitinase with linear linkage specificity	2663
PARN	NM 001134477	Homo sapiens poly(A)-specific ribonuclease (PARN), transcript variant 2, mRNA.	2664
PARN	NM 001242992	Homo sapiens poly(A)-specific ribonuclease (PARN), transcript variant 3, mRNA.	2665
PARN	NM 002582	Homo sapiens poly(A)-specific ribonuclease (PARN), transcript variant 1,	2666

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		mRNA.	
PCCA	NM 000282	Homo sapiens propionyl CoA carboxylase, alpha polypeptide (PCCA), transcript variant 1, mRNA.	2667
PCCA	NM 001127692	Homo sapiens propionyl CoA carboxylase, alpha polypeptide (PCCA), transcript variant 2, mRNA.	2668
PCCA	NM 001178004	Homo sapiens propionyl CoA carboxylase, alpha polypeptide (PCCA), transcript variant 3, mRNA.	2669
PCCB	NM 000532	Homo sapiens propionyl CoA carboxylase, beta polypeptide (PCCB), transcript variant 1, mRNA.	2670
PCCB	NM 001178014	Homo sapiens propionyl CoA carboxylase, beta polypeptide (PCCB), transcript variant 2, mRNA.	2671
PDCD1	NM 005018	Homo sapiens programmed cell death 1 (PDCD1), mRNA.	2672
PDCD1LG2	NM 025239	Homo sapiens programmed cell death 1 ligand 2 (PDCD1LG2), mRNA.	2673
PEPD	NM 000285	Homo sapiens peptidase D (PEPD), transcript variant 1, mRNA.	2674
PEPD	NM 001166056	Homo sapiens peptidase D (PEPD), transcript variant 2, mRNA.	2675
PEPD	NM 00 1166057	Homo sapiens peptidase D (PEPD), transcript variant 3, mRNA.	2676
PINK1	NM 032409	Homo sapiens PTEN induced putative kinase 1 (PINK1), mRNA.	2677
PLAU	NM 001145031	Homo sapiens plasminogen activator, urokinase (PLAU), transcript variant 2, mRNA.	2678
PLAU	NM 002658	Homo sapiens plasminogen activator, urokinase (PLAU), transcript variant 1, mRNA.	2679
PLAUR	NM 001005376	Homo sapiens plasminogen activator, urokinase receptor (PLAUR), transcript variant 2, mRNA.	2680
PLAUR	NM 001005377	Homo sapiens plasminogen activator, urokinase receptor (PLAUR), transcript variant 3, mRNA.	2681
PLAUR	NM 002659	Homo sapiens plasminogen activator, urokinase receptor (PLAUR), transcript variant 1, mRNA.	2682
PLCG1	NM 002660	Homo sapiens phospholipase C, gamma 1 (PLCG1), transcript variant 1, mRNA.	2683
PLCG1	NM 182811	Homo sapiens phospholipase C, gamma 1 (PLCG1), transcript variant 2, mRNA.	2684
PLD1	NM 001130081	Homo sapiens phospholipase D1, phosphatidylcholine-specific (PLD 1), transcript variant 2, mRNA.	2685
PLD1	NM 002662	Homo sapiens phospholipase D1, phosphatidylcholine-specific (PLD1), transcript variant 1, mRNA.	2686
PLEKHM1	NM 014798	Homo sapiens pleckstrin homology domain containing, family M (with RUN domain) member 1 (PLEKHM1), transcript variant 1, mRNA.	2687
PLEKHM1	NR 027774	Homo sapiens pleckstrin homology domain containing, family M (with RUN domain) member 1 (PLEKHM1), transcript variant 2, non-coding RNA.	2688
PLEKHM1	NR 027782	Homo sapiens pleckstrin homology domain containing, family M (with RUN domain) member 1 (PLEKHM1), transcript variant 3, non-coding RNA.	2689
PLK1	NM 005030	Homo sapiens polo-like kinase 1 (PLK1), mRNA.	2690
PLXNB1	NM 001130082	Homo sapiens plexin B1 (PLXNB 1), transcript variant 2, mRNA.	2691
PLXNB1	NM 002673	Homo sapiens plexin B1 (PLXNB 1), transcript variant 1, mRNA.	2692
PMM2	NM 000303	Homo sapiens phosphomannomutase 2 (PMM2), mRNA.	2693
POLE2	NM 001197330	Homo sapiens polymerase (DNA directed), epsilon 2, accessory subunit (POLE2), transcript variant 2, mRNA.	2694
POLE2	NM 001197331	Homo sapiens polymerase (DNA directed), epsilon 2, accessory subunit (POLE2), transcript variant 3, mRNA.	2695
POLE2	NM 002692	Homo sapiens polymerase (DNA directed), epsilon 2, accessory subunit (POLE2), transcript variant 1, mRNA.	2696
PPM1A	NM 021003	Homo sapiens protein phosphatase, Mg2+/Mn2+ dependent, 1A (PPM1A), transcript variant 1, mRNA.	2697
PPM1A	NM 177951	Homo sapiens protein phosphatase, Mg2+/Mn2+ dependent, 1A (PPM1A), transcript variant 2, mRNA.	2698
PPM1A	NM 177952	Homo sapiens protein phosphatase, Mg2+/Mn2+ dependent, 1A (PPM1A),	2699

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		transcript variant 3, mRNA.	
PRKN	NM 004562.2	Homo sapiens parkin RBR E3 ubiquitin protein ligase, transcript variant 1	2700
PRKN	NM 013987.2	Homo sapiens parkin RBR E3 ubiquitin protein ligase, transcript variant 2	2701
PRKN	NM 013988.2	Homo sapiens parkin RBR E3 ubiquitin protein ligase, transcript variant 3	2702
PRRC2A	NM 004638	Homo sapiens proline-rich coiled-coil 2A (PRRC2A), transcript variant 2, mRNA.	2703
PRRC2A	NM 080686	Homo sapiens proline-rich coiled-coil 2A (PRRC2A), transcript variant 1, mRNA.	2704
PSEN1	NM 000021	Homo sapiens presenilin 1 (PSEN1), transcript variant 1, mRNA.	2705
PSEN1	NM 007318	Homo sapiens presenilin 1 (PSEN1), transcript variant 2, mRNA.	2706
PSENEN	NM 172341	Homo sapiens presenilin enhancer 2 homolog (C. elegans) (PSENEN), transcript variant 1, mRNA.	2707
PSMA7	NM 002792	Homo sapiens proteasome (prosome, macropain) subunit, alpha type, 7 (PSMA7), mRNA.	2708
RAB5A	NM 004162	Homo sapiens RAB5A, member RAS oncogene family (RAB5A), mRNA.	2709
RAB5B	NM 001252036	Homo sapiens RAB5B, member RAS oncogene family (RAB5B), transcript variant 2, mRNA.	2710
RAB5B	NM 001252037	Homo sapiens RAB5B, member RAS oncogene family (RAB5B), transcript variant 3, mRNA.	2711
RAB5B	NM 002868	Homo sapiens RAB5B, member RAS oncogene family (RAB5B), transcript variant 1, mRNA.	2712
RAB5C	NM 001252039	Homo sapiens RAB5C, member RAS oncogene family (RAB5C), transcript variant 3, mRNA.	2713
RAB5C	NM 004583	Homo sapiens RAB5C, member RAS oncogene family (RAB5C), transcript variant 2, mRNA.	2714
RAB5C	NM 201434	Homo sapiens RAB5C, member RAS oncogene family (RAB5C), transcript variant 1, mRNA.	2715
RAD50	NM 005732	Homo sapiens RAD50 homolog (S. cerevisiae) (RAD50), mRNA.	2716
RANBP2	NM 006267	Homo sapiens RAN binding protein 2 (RANBP2), mRNA.	2717
RASGRP1	NM 001128602	Homo sapiens RAS guanyl releasing protein 1 (calcium and DAG-regulated) (RASGRP1), transcript variant 2, mRNA.	2718
RASGRP1	NM 005739	Homo sapiens RAS guanyl releasing protein 1 (calcium and DAG-regulated) (RASGRP1), transcript variant 1, mRNA.	2719
RELA	NM 001145138	Homo sapiens v-rel avian reticuloendotheliosis viral oncogene homolog A (RELA), transcript variant 2, mRNA.	2720
RELA	NM 001243984	Homo sapiens v-rel avian reticuloendotheliosis viral oncogene homolog A (RELA), transcript variant 3, mRNA.	2721
RELA	NM 001243985	Homo sapiens v-rel avian reticuloendotheliosis viral oncogene homolog A (RELA), transcript variant 4, mRNA.	2722
RELA	NM 021975	Homo sapiens v-rel avian reticuloendotheliosis viral oncogene homolog A (RELA), transcript variant 1, mRNA.	2723
RELB	NM 006509	Homo sapiens v-rel avian reticuloendotheliosis viral oncogene homolog B (RELB), mRNA.	2724
RHOH	NM 004310	Homo sapiens ras homolog family member H (RHOH), transcript variant 6, mRNA.	2725
RLTPR	NM 001013838	Homo sapiens RGD motif, leucine rich repeats, tropomodulin domain and proline-rich containing (RLTPR), mRNA.	2726
RNF125	NM 017831	Homo sapiens ring finger protein 125, E3 ubiquitin protein ligase (RNF125), mRNA.	2727
RORC	NM 001001523	Homo sapiens RAR-related orphan receptor C (RORC), transcript variant 2, mRNA.	2728
RORC	NM 005060	Homo sapiens RAR-related orphan receptor C (RORC), transcript variant 1, mRNA.	2729
RPSA	NM 001012321	Homo sapiens ribosomal protein SA (RPSA), transcript variant 2, mRNA.	2730
RPSA	NM 002295	Homo sapiens ribosomal protein SA (RPSA), transcript variant 1, mRNA.	2731
RSAD2	NM 080657	Homo sapiens radical S-adenosyl methionine domain containing 2 (RSAD2),	2732

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		mRNA.	
SAMD9	NM 001193307	Homo sapiens sterile alpha motif domain containing 9 (SAMD9), transcript variant 2, mRNA.	2733
SAMD9	NM 017654	Homo sapiens sterile alpha motif domain containing 9 (SAMD9), transcript variant 1, mRNA.	2734
SAMD9L	NM 152703	Homo sapiens sterile alpha motif domain containing 9-like (SAMD9L), mRNA.	2735
SEMA3E	NM 001178129	Homo sapiens sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E (SEMA3E), transcript variant 2, mRNA.	2736
SEMA3E	NM 012431	Homo sapiens sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E (SEMA3E), transcript variant 1, mRNA.	2737
SERPINA1	NM 000295	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 1, mRNA.	2738
SERPINA1	NM 001002235	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 3, mRNA.	2739
SERPINA1	NM 001002236	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 2, mRNA.	2740
SERPINA1	NM 001127700	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 4, mRNA.	2741
SERPINA1	NM 001127701	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 5, mRNA.	2742
SERPINA1	NM 001127702	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 6, mRNA.	2743
SERPINA1	NM 001127703	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 7, mRNA.	2744
SERPINA1	NM 001127704	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 8, mRNA.	2745
SERPINA1	NM 001127705	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 9, mRNA.	2746
SERPINA1	NM 001127706	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 10, mRNA.	2747
SERPINA1	NM 001127707	Homo sapiens serpin peptidase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1 (SERPINA1), transcript variant 11, mRNA.	2748
SERPINB2	NM 001143818	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 2 (SERPINB2), transcript variant 1, mRNA.	2749
SERPINB2	NM 002575	Homo sapiens serpin peptidase inhibitor, clade B (ovalbumin), member 2 (SERPINB2), transcript variant 2, mRNA.	2750
SERPING1	NM 000062	Homo sapiens serpin peptidase inhibitor, clade G (C1 inhibitor), member 1 (SERPING1), transcript variant 1, mRNA.	2751
SERPING1	NM 001032295	Homo sapiens serpin peptidase inhibitor, clade G (C1 inhibitor), member 1 (SERPING1), transcript variant 2, mRNA.	2752
SH3BP2	NM 001122681	Homo sapiens SH3-domain binding protein 2 (SH3BP2), transcript variant 2, mRNA.	2753
SH3BP2	NM 001145855	Homo sapiens SH3-domain binding protein 2 (SH3BP2), transcript variant 4, mRNA.	2754
SH3BP2	NM 001145856	Homo sapiens SH3-domain binding protein 2 (SH3BP2), transcript variant 3, mRNA.	2755
SH3BP2	NM 003023	Homo sapiens SH3-domain binding protein 2 (SH3BP2), transcript variant 1, mRNA.	2756
SLC29A3	NM 001174098	Homo sapiens solute carrier family 29 (equilibrative nucleoside transporter), member 3 (SLC29A3), transcript variant 2, mRNA.	2757
SLC29A3	NM 018344	Homo sapiens solute carrier family 29 (equilibrative nucleoside transporter), member 3 (SLC29A3), transcript variant 1, mRNA.	2758
SLC29A3	NR 033413	Homo sapiens solute carrier family 29 (equilibrative nucleoside transporter), member 3 (SLC29A3), transcript variant 3, non-coding RNA.	2759
SLC29A3	NR 033414	Homo sapiens solute carrier family 29 (equilibrative nucleoside transporter), member 3 (SLC29A3), transcript variant 4, non-coding RNA.	2760
SLC35C1	NM 001145265	Homo sapiens solute carrier family 35 (GDP-fucose transporter), member C1 (SLC35C1), transcript variant 2, mRNA.	2761

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
SLC35C1	NM 001145266	Homo sapiens solute carrier family 35 (GDP-fucose transporter), member C1 (SLC35C1), transcript variant 3, mRNA.	2762
SLC35C1	NM 018389	Homo sapiens solute carrier family 35 (GDP-fucose transporter), member C1 (SLC35C1), transcript variant 1, mRNA.	2763
SLC7A7	NM 001126105	Homo sapiens solute carrier family 7 (amino acid transporter light chain, y+L system), member 7 (SLC7A7), transcript variant 2, mRNA.	2764
SLC7A7	NM 001126106	Homo sapiens solute carrier family 7 (amino acid transporter light chain, y+L system), member 7 (SLC7A7), transcript variant 3, mRNA.	2765
SLC7A7	NR 040448	Homo sapiens solute carrier family 7 (amino acid transporter light chain, y+L system), member 7 (SLC7A7), transcript variant 1, non-coding RNA.	2766
SLC9A1	NM 003047	Homo sapiens solute carrier family 9, subfamily A (NHE1, cation proton antiporter 1), member 1 (SLC9A1), transcript variant 1, mRNA.	2767
SLC9A1	NR 046474	Homo sapiens solute carrier family 9, subfamily A (NHE1, cation proton antiporter 1), member 1 (SLC9A1), transcript variant 2, non-coding RNA.	2768
SMARCAL1	NM 001127207	Homo sapiens SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a-like 1 (SMARCAL1), transcript variant 2, mRNA.	2769
SMARCAL1	NM 014140	Homo sapiens SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a-like 1 (SMARCAL1), transcript variant 1, mRNA.	2770
SMARCD2	NM 001098426	Homo sapiens SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 2 (SMARCD2), mRNA.	2771
SMC3	NM 005445	Homo sapiens structural maintenance of chromosomes 3 (SMC3), mRNA.	2772
SMURF2	NM 022739	Homo sapiens SMAD specific E3 ubiquitin protein ligase 2 (SMURF2), mRNA.	2773
SRP54	NM 001146282	Homo sapiens signal recognition particle 54kDa (SRP54), transcript variant 2, mRNA.	2774
SRP54	NM 003136	Homo sapiens signal recognition particle 54kDa (SRP54), transcript variant 1, mRNA.	2775
STN1	NM 024928.4	Homo sapiens STN1, CST complex subunit	2776
TBC1D15	NM 001146213	Homo sapiens TBC1 domain family, member 15 (TBC1D15), transcript variant 3, mRNA.	2777
TBC1D15	NM 001146214	Homo sapiens TBC1 domain family, member 15 (TBC1D15), transcript variant 2, mRNA.	2778
TBC1D15	NM 022771	Homo sapiens TBC1 domain family, member 15 (TBC1D15), transcript variant 1, mRNA.	2779
TBC1D15	NR 027449	Homo sapiens TBC1 domain family, member 15 (TBC1D15), transcript variant 4, non-coding RNA.	2780
TBC1D17	NM 001168222	Homo sapiens TBC1 domain family, member 17 (TBC1D17), transcript variant 2, mRNA.	2781
TBC1D17	NM 024682	Homo sapiens TBC1 domain family, member 17 (TBC1D17), transcript variant 1, mRNA.	2782
TCF3	NM 001136139	Homo sapiens transcription factor 3 (TCF3), transcript variant 2, mRNA.	2783
TCF3	NM 003200	Homo sapiens transcription factor 3 (TCF3), transcript variant 1, mRNA.	2784
TCN2	NM 000355	Homo sapiens transcobalamin II (TCN2), transcript variant 1, mRNA.	2785
TCN2	NM 001184726	Homo sapiens transcobalamin II (TCN2), transcript variant 2, mRNA.	2786
TEK	NM 000459	Homo sapiens TEK tyrosine kinase, endothelial (TEK), mRNA.	2787
TERC	NR 001566	Homo sapiens telomerase RNA component (TERC), telomerase RNA.	2788
TERT	NM 001193376	Homo sapiens telomerase reverse transcriptase (TERT), transcript variant 2, mRNA.	2789
TERT	NM 198253	Homo sapiens telomerase reverse transcriptase (TERT), transcript variant 1, mRNA.	2790
TFPI	NM 001032281	Homo sapiens tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor) (TFPI), transcript variant 2, mRNA.	2791
TFPI	NM 006287	Homo sapiens tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor) (TFPI), transcript variant 1, mRNA.	2792
TFRC	NM 001128148	Homo sapiens transferrin receptor (p90, CD71) (TFRC), transcript variant 2,	2793

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		mRNA.	
TFRC	NM 003234	Homo sapiens transferrin receptor (p90, CD71) (TFRC), transcript variant 1, mRNA.	2794
THBD	NM 000361	Homo sapiens thrombomodulin (THBD), mRNA.	2795
THBS1	NM 003246	Homo sapiens thrombospondin 1 (THBS1), mRNA.	2796
TINF2	NM 001099274	Homo sapiens TERF1 (TRF1)-interacting nuclear factor 2 (TINF2), transcript variant 1, mRNA.	2797
TINF2	NM 012461	Homo sapiens TERF1 (TRF1)-interacting nuclear factor 2 (TINF2), transcript variant 2, mRNA.	2798
TIRAP	NM 001039661	Homo sapiens toll-interleukin 1 receptor (TIR) domain containing adaptor protein (TIRAP), transcript variant 3, mRNA.	2799
TIRAP	NM 148910	Homo sapiens toll-interleukin 1 receptor (TIR) domain containing adaptor protein (TIRAP), transcript variant 2, mRNA.	2800
TMC6	NM 0011271 98	Homo sapiens transmembrane channel-like 6 (TMC6), transcript variant 1, mRNA.	2801
TMC6	NM 007267	Homo sapiens transmembrane channel-like 6 (TMC6), transcript variant 2, mRNA.	2802
TMC8	NM 152468	Homo sapiens transmembrane channel-like 8 (TMC8), mRNA.	2803
TNFRSF17	NM 001192	Homo sapiens tumor necrosis factor receptor superfamily, member 17 (TNFRSF17), mRNA.	2804
TNFRSF1A	NM 001065	Homo sapiens tumor necrosis factor receptor superfamily, member 1A (TNFRSF1A), mRNA.	2805
TNFSF10	NM 001190942	Homo sapiens tumor necrosis factor (ligand) superfamily, member 10 (TNFSF10), transcript variant 2, mRNA.	2806
TNFSF10	NM 001190943	Homo sapiens tumor necrosis factor (ligand) superfamily, member 10 (TNFSF10), transcript variant 3, mRNA.	2807
TNFSF10	NM 003810	Homo sapiens tumor necrosis factor (ligand) superfamily, member 10 (TNFSF10), transcript variant 1, mRNA.	2808
TNFSF10	NR 033994	Homo sapiens tumor necrosis factor (ligand) superfamily, member 10 (TNFSF10), transcript variant 4, non-coding RNA.	2809
TNFSF13	NM 001198622	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13 (TNFSF13), transcript variant delta, mRNA.	2810
TNFSF13	NM 001198623	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13 (TNFSF13), transcript variant zeta, mRNA.	2811
TNFSF13	NM 001198624	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13 (TNFSF13), transcript variant eta, mRNA.	2812
TNFSF13	NM 003808	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13 (TNFSF13), transcript variant alpha, mRNA.	2813
TNFSF13	NM 172087	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13 (TNFSF13), transcript variant beta, mRNA.	2814
TNFSF13	NM 172088	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13 (TNFSF13), transcript variant gamma, mRNA.	2815
TNFSF13	NR 073490	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13 (TNFSF13), transcript variant episilon, non-coding RNA.	2816
TNFSF13B	NM 001145645	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13b (TNFSF13B), transcript variant 2, mRNA.	2817
TNFSF13B	NM 006573	Homo sapiens tumor necrosis factor (ligand) superfamily, member 13b (TNFSF13B), transcript variant 1, mRNA.	2818
TNIP 1	NM 001252385	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 1, mRNA.	2819
TNIP 1	NM 001252386	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 2, mRNA.	2820
TNIP 1	NM 001252390	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 3, mRNA.	2821
TNIP 1	NM 001252391	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 4, mRNA.	2822
TNIP 1	NM 001252392	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 6, mRNA.	2823

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
TNIP 1	NM 001252393	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 7, mRNA.	2824
TNIP 1	NM 001258454	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 8, mRNA.	2825
TNIP 1	NM 001258455	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 9, mRNA.	2826
TNIP 1	NM 001258456	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 10, mRNA.	2827
TNIP 1	NM 006058	Homo sapiens TNFAIP3 interacting protein 1 (TNIP1), transcript variant 5, mRNA.	2828
TP53AIP1	NM 001195194	Homo sapiens tumor protein p53 regulated apoptosis inducing protein 1 (TP53AIP1), transcript variant 3, mRNA.	2829
TP53AIP1	NM 001195195	Homo sapiens tumor protein p53 regulated apoptosis inducing protein 1 (TP53AIP1), transcript variant 2, mRNA.	2830
TP53AIP1	NM 001251964	Homo sapiens tumor protein p53 regulated apoptosis inducing protein 1 (TP53AIP1), transcript variant 4, mRNA.	2831
TP53AIP1	NM 022112	Homo sapiens tumor protein p53 regulated apoptosis inducing protein 1 (TP53AIP1), transcript variant 1, mRNA.	2832
TPP1	NM 000391	Homo sapiens tripeptidyl peptidase I (TPP1), mRNA.	2833
TPP2	NM 003291	Homo sapiens tripeptidyl peptidase II (TPP2), mRNA.	2834
TRAF3IP2	NM 001164281	Homo sapiens TRAF3 interacting protein 2 (TRAF3IP2), transcript variant 3, mRNA.	2835
TRAF3IP2	NM 001164283	Homo sapiens TRAF3 interacting protein 2 (TRAF3IP2), transcript variant 5, mRNA.	2836
TRAF3IP2	NM 147686	Homo sapiens TRAF3 interacting protein 2 (TRAF3IP2), transcript variant 2, mRNA.	2837
TRAF3IP2	NR 028338	Homo sapiens TRAF3 interacting protein 2 (TRAF3IP2), transcript variant 1, non-coding RNA.	2838
TRIM25	NM 005082	Homo sapiens tripartite motif containing 25 (TRIM25), mRNA.	2839
TRIM37	NM 001005207	Homo sapiens tripartite motif containing 37 (TRIM37), transcript variant 2, mRNA.	2840
TRIM37	NM 015294	Homo sapiens tripartite motif containing 37 (TRIM37), transcript variant 1, mRNA.	2841
TTC37	NM 014639	Homo sapiens tetratricopeptide repeat domain 37 (TTC37), mRNA.	2842
UBD	NM 006398	Homo sapiens ubiquitin D (UBD), mRNA.	2843
USB1	NM 001195302	Homo sapiens U6 snRNA biogenesis 1 (USB1), transcript variant 2, mRNA.	2844
USB1	NM 001204911	Homo sapiens U6 snRNA biogenesis 1 (USB1), transcript variant 3, mRNA.	2845
USB1	NM 024598	Homo sapiens U6 snRNA biogenesis 1 (USB1), transcript variant 1, mRNA.	2846
USP15	NM 001252078	Homo sapiens ubiquitin specific peptidase 15 (USP15), transcript variant 1, mRNA.	2847
USP15	NM 001252079	Homo sapiens ubiquitin specific peptidase 15 (USP15), transcript variant 3, mRNA.	2848
USP15	NM 006313	Homo sapiens ubiquitin specific peptidase 15 (USP15), transcript variant 2, mRNA.	2849
USP21	NM 001014443	Homo sapiens ubiquitin specific peptidase 21 (USP21), transcript variant 3, mRNA.	2850
USP21	NM 012475	Homo sapiens ubiquitin specific peptidase 21 (USP21), transcript variant 1, mRNA.	2851
USP25	NM 013396	Homo sapiens ubiquitin specific peptidase 25 (USP25), mRNA.	2852
USP3	NM 001256702	Homo sapiens ubiquitin specific peptidase 3 (USP3), transcript variant 2, mRNA.	2853
USP3	NM 006537	Homo sapiens ubiquitin specific peptidase 3 (USP3), transcript variant 1, mRNA.	2854
USP3	NR 046341	Homo sapiens ubiquitin specific peptidase 3 (USP3), transcript variant 3, non-coding RNA.	2855
USP3	NR 046342	Homo sapiens ubiquitin specific peptidase 3 (USP3), transcript variant 4,	2856

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
		non-coding RNA.	
VAV1	NM 001258206	Homo sapiens vav 1 guanine nucleotide exchange factor (VAV1), transcript variant 2, mRNA.	2857
VAV1	NM 001258207	Homo sapiens vav 1 guanine nucleotide exchange factor (VAV1), transcript variant 3, mRNA.	2858
VAV1	NM 005428	Homo sapiens vav 1 guanine nucleotide exchange factor (VAV1), transcript variant 1, mRNA.	2859
VDR	NM 000376	Homo sapiens vitamin D (1,25- dihydroxyvitamin D3) receptor (VDR), transcript variant 1, mRNA.	2860
VDR	NM 001017535	Homo sapiens vitamin D (1,25- dihydroxyvitamin D3) receptor (VDR), transcript variant 2, mRNA.	2861
VDR	NM 001017536	Homo sapiens vitamin D (1,25- dihydroxyvitamin D3) receptor (VDR), transcript variant 3, mRNA.	2862
VEGFA	NM 001025366	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 1, mRNA.	2863
VEGFA	NM 001025367	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 3, mRNA.	2864
VEGFA	NM 001025368	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 4, mRNA.	2865
VEGFA	NM 001025369	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 5, mRNA.	2866
VEGFA	NM 001025370	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 6, mRNA.	2867
VEGFA	NM 001033756	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 7, mRNA.	2868
VEGFA	NM 001171622	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 8, mRNA.	2869
VEGFA	NM 001171623	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 1, mRNA.	2870
VEGFA	NM 001171624	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 2, mRNA.	2871
VEGFA	NM 001171625	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 3, mRNA.	2872
VEGFA	NM 001171626	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 4, mRNA.	2873
VEGFA	NM 001171627	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 5, mRNA.	2874
VEGFA	NM 001171628	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 6, mRNA.	2875
VEGFA	NM 001171629	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 7, mRNA.	2876
VEGFA	NM 001171630	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 8, mRNA.	2877
VEGFA	NM 001204384	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 9, mRNA.	2878
VEGFA	NM 001204385	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 9, mRNA.	2879
VEGFA	NM 003376	Homo sapiens vascular endothelial growth factor A (VEGFA), transcript variant 2, mRNA.	2880
WASHC5	NM 001330609.1	Homo sapiens WASH complex subunit 5, transcript variant 2	2881
WASHC5	NM 014846.3	Homo sapiens WASH complex subunit 5, transcript variant 1	2882
WDR1	NM 005112	Homo sapiens WD repeat domain 1 (WDR1), transcript variant 2, mRNA.	2883
WDR1	NM 017491	Homo sapiens WD repeat domain 1 (WDR1), transcript variant 1, mRNA.	2884
WRAP53	NM 001143990	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), transcript variant 2, mRNA.	2885
WRAP53	NM 001143991	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), transcript variant 3, mRNA.	2886

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
WRAP53	NM 001143992	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), transcript variant 4, mRNA.	2887
WRAP53	NM 018081	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), transcript variant 1, mRNA.	2888
XAF1	NM 017523	Homo sapiens XIAP associated factor 1 (XAF1), transcript variant 1, mRNA.	2889
XAF1	NM 199139	Homo sapiens XIAP associated factor 1 (XAF1), transcript variant 2, mRNA.	2890
XAF1	NR 046396	Homo sapiens XIAP associated factor 1 (XAF1), transcript variant 3, non-coding RNA.	2891
XAF1	NR 046397	Homo sapiens XIAP associated factor 1 (XAF1), transcript variant 4, non-coding RNA.	2892
XAF1	NR 046398	Homo sapiens XIAP associated factor 1 (XAF1), transcript variant 5, non-coding RNA.	2893

Table 32 represents a non-redundant list of transcript variants that correspond to the **Table 31** genes.

Table 33: SEQ ID 3000-3274, SNV list (Tables 34-36, 38, 39) with SEQ ID numbers (similar to **Table 5)**

Chromosome	Position (hg!9)	Ref Allele	Alt Allele	SEQ ID
1	11087369	T	C	3000
1	11090287	C	T	3001
1	11090916	C	A	3002
1	11094908	T	A	3003
1	11106648	G	A	3004
1	11106666	T	C	3005
1	11106673	G	A	3006
1	12049283	C	T	3007
1	12064892	G	A	3008
1	12064931	G	A	3009
1	57333311	C	A	3010
1	57372463	C	T	3011
1	57373778	G	A	3012
1	57378149	G	T	3013
1	57383295	G	A	3014
1	57409459	C	A	3015
1	57422511	C	T	3016
1	59131311	G	T	3017
1	82409337	A	G	3018
1	82416040	C	T	3019
1	82450271	G	A	3020
1	82456165	G	T	3021
1	160281740	T	C	3022
1	183532364	T	A	3023
1	196709774	G	T	3024
1	196759282	C	T	3025
1	196794681	G	T	3026
1	196799813	G	A	3027
1	196871610	A	T	3028
1	196918732	G	T	3029
1	196920123	G	A	3030
1	196973890	G	A	3031
2	7027110	G	T	3032
2	15432775	C	T	3033
2	15519924	C	T	3034
2	15542352	C	T	3035

Chromosome	Position (hg!9)	Ref Allele	Alt Allele	SEQ ID
2	15607842	T	C	3036
2	15674686	T	C	3037
2	47637246	A	G	3038
2	47702191	A	G	3039
2	74688563	C	T	3040
2	74688884	G	A	3041
2	74689335	G	T	3042
2	74690039	G	A	3043
2	74690371	C	T	3044
2	74690378	C	T	3045
2	109381493	G	C	3046
2	109382448	C	T	3047
2	109384800	C	T	3048
2	188349523	A	G	3049
2	217285060	C	G	3050
2	217285104	C	G	3051
2	217288388	G	C	3052
3	11382205	A	C	3053
3	11399970	C	T	3054
3	11402163	G	A	3055
3	11468330	A	G	3056
3	20017123	A	C	3057
3	37061893	T	C	3058
3	37061929	A	G	3059
3	37092025	C	T	3060
3	39452455	C	T	3061
3	48457498	G	C	3062
3	48461135	C	T	3063
3	48463544	G	A	3064
3	49949430	C	T	3065
3	171321023	C	T	3066
3	171379953	C	T	3067
3	171431726	C	G	3068
3	171455697	G	C	3069
3	171455739	A	G	3070
4	89352440	A	C	3071
4	89414196	G	A	3072
4	110864533	C	T	3073
4	110865044	G	C	3074
4	110866260	G	C	3075
4	110897252	C	G	3076
4	110929301	T	C	3077
4	110932508	C	A	3078
4	126237697	A	C	3079
4	126238090	G	T	3080
4	126238305	C	A	3081
4	126239241	G	A	3082
4	126239253	G	C	3083
4	126239421	C	T	3084
4	126239623	G	A	3085

Chromosome	Position (hg!9)	Ref Allele	Alt Allele	SEQ ID
4	126239986	C	T	3086
4	126240255	A	G	3087
4	126240377	G	T	3088
4	126240390	A	G	3089
4	126240968	A	T	3090
4	126241248	C	G	3091
4	126241720	T	C	3092
4	126241785	G	A	3093
4	126328170	C	T	3094
4	126336758	G	A	3095
4	126336851	G	A	3096
4	126372003	A	G	3097
4	126372975	A	C	3098
4	126373570	C	T	3099
4	126389832	G	A	3100
4	126408663	A	G	3101
4	126411179	C	T	3102
4	126411493	C	T	3103
4	126412106	C	G	3104
4	126412154	G	A	3105
4	126412226	G	A	3106
4	126412634	C	G	3107
5	34929974	A	T	3108
5	34937524	C	T	3109
5	39311336	A	T	3110
5	40955561	G	C	3111
5	40959622	C	T	3112
5	40964852	A	C	3113
5	131925413	A	G	3114
5	132015535	C	T	3115
5	134076987	G	C	3116
6	30673403	A	G	3117
6	30675830	T	A	3118
6	30680721	G	A	3119
6	31593603	A	G	3120
6	31595795	C	T	3121
6	31595926	C	T	3122
6	31597451	G	A	3123
6	31597469	A	C	3124
6	31598524	G	A	3125
6	31599370	G	C	3126
6	31600118	G	A	3127
6	31600558	C	T	3128
6	31600696	G	A	3129
6	31600708	G	A	3130
6	31601735	G	A	3131
6	31602967	G	A	3132
6	31603045	A	G	3133
6	31604591	C	T	3134
6	31604610	T	C	3135
6	31604894	C	T	3136

Chromosome	Position (hg!9)	Ref Allele	Alt Allele	SEQ ID
6	31605016	T	C	3137
6	31605278	C	T	3138
6	31709045	C	T	3139
6	31725978	C	G	3140
6	31729925	C	T	3141
6	43748510	G	A	3142
6	106740989	T	C	3143
6	111913058	G	A	3144
6	111913262	C	T	3145
7	5959528	T	C	3146
7	30491421	G	T	3147
7	30491693	C	T	3148
7	74193620	G	A	3149
7	80300449	T	G	3150
7	92732769	T	C	3151
7	92733766	C	A	3152
7	92761257	A	T	3153
7	92763720	G	A	3154
7	117230454	G	C	3155
7	117232086	G	A	3156
7	117232223	C	T	3157
7	117246776	T	C	3158
8	11400805	C	A	3159
8	11400849	C	T	3160
8	11407690	C	T	3161
8	11412934	G	A	3162
8	11415492	A	C	3163
8	11418856	C	T	3164
8	126071726	G	A	3165
8	126095371	T	C	3166
9	27109600	A	C	3167
9	27158131	C	T	3168
9	27183598	C	T	3169
9	27197588	C	T	3170
9	35612978	G	A	3171
9	98678698	G	C	3172
9	98691137	T	C	3173
10	54531226	C	T	3174
10	54531235	C	T	3175
10	54531242	G	A	3176
10	75672059	G	A	3177
10	75673748	A	C	3178
10	75675086	T	C	3179
10	91098614	G	A	3180
10	91162073	T	C	3181
11	60891358	C	T	3182
11	60893235	C	T	3183
11	65423327	G	A	3184
11	72145307	C	G	3185
11	119045378	C	T	3186

Chromosome	Position (hg!9)	Ref Allele	Alt Allele	SEQ ID
11	119045951	C	T	3187
11	119052975	G	T	3188
11	128807550	C	G	3189
12	25362788	T	A	3190
12	25368449	A	T	3191
12	40657700	C	G	3192
12	40671773	A	G	3193
12	40677813	G	T	3194
12	40702420	A	G	3195
12	40702910	C	T	3196
12	40702911	G	A	3197
12	40740686	A	G	3198
12	49421042	C	T	3199
12	49421811	C	A	3200
12	49428694	T	C	3201
12	49433083	G	A	3202
12	49433533	G	C	3203
12	49440564	C	T	3204
12	56383856	A	G	3205
12	56385915	GGGA	G	3206
12	113357209	G	A	3207
12	113357237	G	C	3208
12	113357442	G	A	3209
12	113403675	C	T	3210
12	113405825	G	A	3211
12	113448288	A	G	3212
14	35497285	T	C	3213
14	94847262	T	A	3214
15	63866584	C	A	3215
15	93545488	T	A	3216
16	11145457	G	A	3217
16	11272435	G	A	3218
16	14704607	G	A	3219
16	23693443	C	T	3220
16	23700676	T	A	3221
16	58054099	C	G	3222
16	67680806	G	A	3223
16	67685730	A	T	3224
16	67694044	C	T	3225
16	67694078	G	T	3226
16	69725697	T	G	3227
16	77225440	G	T	3228
17	6659427	C	A	3229
17	7592168	C	G	3230
17	8138569	C	G	3231
17	41165632	G	A	3232
17	43555253	A	G	3233
17	62582265	T	C	3234
17	76129619	C	T	3235
17	76130947	G	T	3236
18	29645930	A	T	3237

Chromosome	Position (hg!9)	Ref Allele	Alt Allele	SEQ ID
18	57103126	G	A	3238
18	61570402	G	A	3239
19	6707129	G	A	3240
19	7754222	G	A	3241
19	10394724	C	T	3242
19	10395141	G	A	3243
19	15383894	C	G	3244
19	33892731	A	G	3245
19	33968991	T	A	3246
19	33980963	G	A	3247
19	35823528	C	T	3248
19	44153248	T	C	3249
19	49843566	G	A	3250
19	50385633	G	C	3251
19	55494157	G	A	3252
20	39794388	T	C	3253
20	44640275	G	A	3254
20	44640959	G	A	3255
21	35122475	C	T	3256
21	35239562	A	G	3257
21	42752030	G	A	3258
21	42807815	A	G	3259
21	42807881	G	A	3260
21	42812952	G	A	3261
21	42813652	C	A	3262
21	42815731	G	A	3263
21	42817460	C	T	3264
21	42830468	G	A	3265
21	42830657	G	A	3266
21	45708270	T	C	3267
21	45708278	G	A	3268
21	45713696	G	A	3269
22	31006882	T	G	3270
22	31008867	T	C	3271
22	31018975	T	C	3272
22	35806756	G	A	3273
22	37271882	T	C	3274

Table 33 lists a set of SNVs reported in Tables 34-36, 38, or 39 that were found in the 70 PML cases for the genes listed in **Table 31.**

Table 34: Case-solving Tier 1 SNVs (het, hom, or phased comp het) with frequency <=1/1,000 or novel

Sample ID	RefSeq Gene Symbol	Variant Type	Chromosome	Position (hg19)	Ref Allele	Alt Allele	Amino Acid Change	Variant Frequency Details (Ethnic specific)	Frequency (Reciprocal, Ethnic specific)	Compound Variant Frequency Details (Ethnic specific)	Compound Frequency (Reciprocal, Ethnic specific)	SEQ ID
MVGS1116-8a	PRRC2A	SNV het	6	31601735	G	A	R1563Q	0.087118	1 in 11	7.40E-04	1 in 1,350	3131
MVGS1116-8a	PRRC2A	SNV het	6	31604894	C	T	R2075W	0.034000	1 in 29	7.40E-04	1 in 1,350	3136
MVGS1368	KMT2D	SNV het	12	49421811	C	A	K4832N	0.000240	1 in 4,167	n/a	n/a	3200
MVGS540-374b	ATG7	SNV het	3	11382205	A	C	K287Q	6.14E-05	1 in 16,292	n/a	n/a	3053
MVGS540-374b	KMT2D	SNV het	12	49433533	G	C	Q2674E	0	infinite	n/a	n/a	3203
MVGS694-6a	AIRE	SNV het	21	45708270	T	C	M194T	0	infinite	n/a	n/a	3267
MVGS811-13a	LRRK2	SNV het	12	40702910	C	T	R1398C	6.01E-05	1 in 16,646	2.09E-06	1 in 477,620	3196
MVGS811-		SNV						0.1394			1 in	

MVGS811-13a	LRRK2	SNV het	12	40702911	G	A	R1398H	0.1954 08	1 in 7	2.09E-06	1 in 477,620	3197
MVGS811-13a	USP3	SNV het	15	63866584	C	A	P111T	0.0005 13	1 in 1,950	n/a	n/a	3215
MVGS995-4a	VEGFA	SNV het	6	43748510	G	A	R155H	3.00E-05	1 in 33,364	n/a	n/a	3142
PML04	RSAD2	SNV het	2	7027110	G	T	E117*	0	infinite	n/a	n/a	3032
PML04	IFIT1	SNV het	10	91162073	T	C	L14P	0	infinite	n/a	n/a	3181
PML04	CLEC16A	SNV het	16	11145457	G	A	V634M	0.0001 20	1 in 8,339	n/a	n/a	3217
PML06	LRRK2	SNV hom	12	40657700	C	G	N299K	0.0185 54	1 in 54	1.19E-05	1 in 84,337	3192
PML06	LRRK2	SNV het	12	40671773	A	G	I423M	0.0001 94	1 in 5,159	see het SNVs	see het SNVs	3193
PML06	LRRK2	SNV hom	12	40702911	G	A	R1398H	0.0196 46	1 in 51	1.17E-05	1 in 85,236	3197
PML06	CD37	SNV het	19	49843566	G	A	R208Q	0.0001 94	1 in 5,165	n/a	n/a	3250
PML09	FAT4	SNV hom	4	126239986	C	T	A807V	0.1960 43	1 in 5	0.000273	1 in 3,665	3086
PML09	FAT4	SNV het	4	126241248	C	G	Q1228E	0.0022 27	1 in 449	0.000273	1 in 3,665	3091
PML09	PRRC2A	SNV het	6	31600118	G	A	R1223H	0	infinite	n/a	n/a	3127
PML10	ATG7	SNV het	3	11468330	A	G	N643S	0	infinite	n/a	n/a	3056
PML10	PRRC2A	SNV het	6	31597469	A	C	K701Q	0.0001 86	1 in 5,385	n/a	n/a	3124
PML14	KRAS	SNV het	12	25362788	T	A	M170L	3.03E-05	1 in 33,044	n/a	n/a	3190
PML15	CFHR1	SNV hom	1	196794681	G	T	V45F	0.0001 82	1 in 5,509	n/a	n/a	3026
PML15	CFHR1	SNV hom	1	196799813	G	A	n/a	0.0001 75	1 in 5,711	n/a	n/a	3027
PML15	PRRC2A	SNV het	6	31598524	G	A	R804H	0	infinite	n/a	n/a	3125
PML15	MBL2	SNV hom	10	54531226	C	T	G57E	0.0005 18	1 in 1,929	n/a	n/a	3174

PML16	IFIT3	SNV het	10	91098614	G	A	E68K	0	infinite	n/a	n/a	3180
PML17	TEK	SNV het	9	27109600	A	C	L4F	8.99E-05	1 in 11,123	n/a	n/a	3167
PML17	TEK	SNV het	9	27158131	C	T	R119C	0	infinite	n/a	n/a	3168
PML18	OAS3	SNV hom	12	113403675	C	T	R844*	0.0002 77	1 in 3,611	n/a	n/a	3210
PML19	MASP2	SNV het	1	11087369	T	C	N545S	0	infinite	0	infinite	3000
PML19	MASP2	SNV hom	1	11090916	C	A	D371Y	0.0707 13	1 in 14	see het SNV	see het SNV	3002
PML19	ADGRL2	SNV het	1	82450271	G	A	R1017Q	0	infinite	n/a	n/a	3020
PML23	RANBP2	SNV het	2	109382448	C	T	P1818L	3.00E-05	1 in 33,362	n/a	n/a	3047
PML25	ATG7	SNV het	3	11399970	C	T	R416C	6.00E-05	1 in 16,676	n/a	n/a	3054
PML25	IL4	SNV het	5	132015535	C	T	R105*	0	infinite	n/a	n/a	3115
PML27	MLL1	SNV het	3	37061929	A	G	N240S	0.0001 80	1 in 5,543	n/a	n/a	3059
PML28	MSH2	SNV het	2	47702191	A	G	N630S	0.0006 31	1 in 1,586	n/a	n/a	3039
PML28	FAT4	SNV het	4	126239623	G	A	S686N	5.99E-05	1 in 16,681	7.35E-06	1 in 136,129	3085
PML28	FAT4	SNV het	4	126239986	C	T	A807V	0.4902 41	1 in 2	7.35E-06	1 in 136,129	3086
PML30	MON1A	SNV het	3	49949410	C	T	G145S	0.0007 20	1 in 1,388	n/a	n/a	3065
PML31	CFHR5	SNV het	1	196973890	G	A	R477H	0.0005 77	1 in 1,734	n/a	n/a	3031
PML31	FAT4	SNV het	4	126239253	G	C	V563L	0	infinite	0	infinite	3083
PML31	FAT4	SNV hom	4	126239986	C	T	A807V	0.3255 10	1 in 3	see het SNVs	see het SNVs	3086
PML31	FAT4	SNV het	4	126328170	C	T	R113C	0	infinite	0	infinite	3094
PML31	FAT4	SNV het	4	126372003	A	G	R1576G	0	infinite	0	infinite	3097
PML31	PRRC2A	SNV het	6	31593601	A	G	Y265C	0	infinite	n/a	n/a	3120
PML31	PRRC2A	SNV hom	6	31602967	G	A	R1740H	0.2583 62	1 in 4	see het SNV	see het SNV	3132
PML31	RNF125	SNV hom	18	29645930	A	T	R190S	0.0001 92	1 in 5,203	n/a	n/a	3257
PML31	PLCG1	SNV het	20	39794388	T	C	I574T	0	infinite	n/a	n/a	3253
PML32	FAT4	SNV hom	4	126239986	C	T	A807V	0.1960 43	1 in 5	see het SNV	see het SNV	3086
PML32	FAT4	SNV het	4	126412226	G	A	G2991E	0	infinite	0	infinite	3106
PML33	PRRC2A	SNV het	6	31600558	C	T	R1370C	3.16E-05	1 in 21,613	n/a	n/a	3128
PML33	KMT2D	SNV het	12	49440564	C	T	V1416M	0	infinite	n/a	n/a	3204
PML35	TP53AIP1	SNV het	11	128807550	C	G	R55P	0.0002 10	1 in 4,763	n/a	n/a	3189
PML35	MX1	SNV het	21	42813652	C	A	T224K	0	infinite	n/a	n/a	3262
PML37	FAT4	SNV hom	4	126239986	C	T	A807V	0.3255 10	1 in 3	2.36E-05	1 in 42,385	3086
PML37	FAT4	SNV het	4	126336758	G	A	A2214T	0.0001 92	1 in 5,203	2.36E-05	1 in 42,385	3095
PML38	PRRC2A	SNV het	6	31597451	G	A	A695T	0	infinite	n/a	n/a	3123
PML38	MX1	SNV het	21	42830657	G	A	R631Q	3.07E-05	1 in 32,566	n/a	n/a	3266
PML39	PRRC2A	SNV het	6	31600708	G	A	G1420R	0	infinite	n/a	n/a	3130
PML40	COPA	SNV het	1	160281710	T	C	M332V	0.0006 92	1 in 1,444	n/a	n/a	3022
PML40	MSH2	SNV hom	2	47637246	A	G	N127S	0	infinite	n/a	n/a	3038
PML40	TBC1D17	SNV het	19	50383633	G	C	E223D	0	infinite	n/a	n/a	3251

PML41	MLH1	SNV het	3	37061893	T	C	V228A	0.0005 77	1 in 1,734	n/a	n/a	3058
PML41	SAMD9L	SNV het	7	92761257	A	T	L1343H	0	infinite	n/a	n/a	3153
PML41	SMURF2	SNV het	17	62582265	T	C	I129V	0	infinite	n/a	n/a	3234
PML43	KRAS	SNV het	12	25368449	A	T	Y166N	0.0003 00	1 in 3,335	n/a	n/a	3191
PML44	RAB5A	SNV het	3	20017123	A	C	D65A	3.01E-05	1 in 33,252	n/a	n/a	3057
PML44	PLXNB1	SNV het	3	48461135	C	T	E854K	0.0007 05	1 in 1,419	n/a	n/a	3063

PML44	NLRX1	SNV het	11	119045378	C	T	R256W	0.000992	1 in 1,008	n/a	n/a	3186
PML45	C7	SNV het	5	40955561	G	C	S389T	0.477838	1 in 2	3.97E-06	1 in 251,629	3111
PML45	C7	SNV het	5	40959622	C	T	R521C	3.36E-05	1 in 29,715	3.97E-06	1 in 251,629	3112
PML45	WASHC5	SNV het	8	126071726	G	A	T379I	3.03E-05	1 in 32,985	n/a	n/a	3165
PML45	BRD4	SNV het	19	15383894	C	G	G6A	9.24E-05	1 in 10,823	n/a	n/a	3244
PML48	NCF1	SNV hom	7	74193620	G	A	G83R	6.32E-05	1 in 15,829	n/a	n/a	3149
PML48	MX1	SNV het	21	42807815	A	G	I52V	0.000120	1 in 8,339	n/a	n/a	3259
PML49	RPSA	SNV het	3	39452455	C	T	R155C	0	infinite	n/a	n/a	3061
PML49	BLK	SNV het	8	11400805	C	A	S24R	3.02E-05	1 in 33,067	n/a	n/a	3159
PML51	NBAS	SNV hom	2	15519974	C	T	R1011H	0.000839	1 in 1,191	n/a	n/a	3034
PML52	KMT2D	SNV het	12	49433083	G	A	P2763L	0	infinite	n/a	n/a	3202
PML52	MX1	SNV het	21	42830468	G	A	R568H	0.000391	1 in 2,559	n/a	n/a	3265
PML54	PARN	SNV het	16	14704607	G	A	R104C	0.000120	1 in 8,342	n/a	n/a	3219
PML55	PLXNB1	SNV het	3	48457498	G	C	L1004V	0	infinite	n/a	n/a	3062
PML55	FAT4	SNV het	4	126412106	C	G	S2951C	0.037914	1 in 26	0.000400	1 in 2,501	3104
PML55	FAT4	SNV het	4	126412154	G	A	R2967K	0.042189	1 in 24	0.000400	1 in 2,501	3105
PML56	PRRC2A	SNV hom	6	31603045	A	G	D1766G	0.001000	1 in 1,000	n/a	n/a	3133
PML56	OAS3	SNV het	12	113405825	G	A	G155R	6.00E-05	1 in 16,672	n/a	n/a	3211
PML57	CD5	SNV het	11	60891358	C	T	R410C	0.000183	1 in 5,475	n/a	n/a	3182
PML57	ITSN1	SNV het	21	35122475	C	T	P125L	3.04E-05	1 in 32,871	n/a	n/a	3256
PML58	MASP2	SNV het	1	11106648	G	A	P126L	0.216611	1 in 5	see het SNV	see het SNV	3004
PML58	MASP2	SNV het	1	11106673	G	A	R118C	0	infinite	0	infinite	3006
PML59	CAMLG	SNV het	5	134076987	G	C	R136T	0	infinite	n/a	n/a	3116
PML60	RELA	SNV het	11	65423327	G	A	n/a	0.000300	1 in 3,337	n/a	n/a	3184
PML62	KMT2D	SNV het	12	49421042	C	T	V4903M	0	infinite	n/a	n/a	3199
PML63	PLD1	SNV het	3	171455697	G	C	P49A	0.033699	1 in 30	3.05E-04	1 in 3,276	3069
PML63	PLD1	SNV het	3	171455739	A	G	F35L	0.011729	1 in 85	3.05E-04	1 in 3,276	3070
PML64	MFN2	SNV het	1	12049283	C	T	H20Y	0	infinite	n/a	n/a	3007
PML65	FAT4	SNV het	4	126239421	C	T	R619C	0.006117	1 in 163	0.000751	1 in 1,332	3084
PML65	FAT4	SNV hom	4	126239986	C	T	A807V	0.325510	1 in 3	0.000751	1 in 1,332	3086
PML65	XAF1	SNV het	17	6659427	C	A	N10K	0	infinite	n/a	n/a	3229
PML66	MFN2	SNV het	1	12064931	G	A	R481H	0	infinite	n/a	n/a	3009
PML66	ADGRL2	SNV het	1	82409337	A	G	Q228R	0.000192	1 in 5,291	n/a	n/a	3018
PML66	PLXNB1	SNV het	3	48463544	G	A	P497L	0	infinite	n/a	n/a	3064
PML66	MX1	SNV het	21	42817460	C	T	P342L	0	infinite	n/a	n/a	3264
PML72	MX1	SNV het	21	42812952	G	A	G221R	0	infinite	n/a	n/a	3261

Table 34 lists potential cause(s) of PML in the study, SNVs (het, hom, or phased comp het) for genes in **Table 31** with frequency <=1/1,000 or novel.

Table 35: Case-solving Tier 2 SNVs (het, hom, or phased comp het) with frequency <=1/1,00 but >1/1,000

Sample ID	RefSeq Gene Symbol	Variant Type	Chromosome	Position (hg19)	Ref Allele	Alt Allele	Amino Acid Change	Variant Frequency Details (Ethnic specific)	Frequency (Reciprocal, Ethnic specific)	Compound Variant Frequency Details (Ethnic specific)	Compound Frequency (Reciprocal, Ethnic specific)	SEQ ID
MVGS111-6-8a	RNF125	SNV het	18	29645930	A	T	R190S	0.004555	1 in 220	n/a	n/a	3237
MVGS540-374b	BLK	SNV het	8	11415492	A	C	K254T	0.004497	1 in 222	n/a	n/a	3163
MVGS540-393b	PRRC2A	SNV het	6	31604591	C	T	P2006S	0.282199	1 in 4	0.002399	1 in 417	3134
MVGS540-393b	PRRC2A	SNV het	6	31604894	C	T	R2075W	0.034000	1 in 29	0.002399	1 in 417	3136
MVGS811-13a	PRRC2A	SNV hom	6	31602967	G	A	R1740H	0.115236	1 in 9	see het SNV	see het SNV	3132
MVGS811-13a	PRRC2A	SNV het	6	31605278	C	T	P2130L	0.045577	1 in 22	0.005004	1 in 200	3138
PML01	PLK1	SNV het	16	23693443	C	T	L21F	0.006119	1 in 163	n/a	n/a	3220
PML02	PRRC2A	SNV het	6	31595926	C	T	P559S	0.007481	1 in 134	n/a	n/a	3122
PML04	ADGRL2	SNV het	1	82456165	G	T	G1130V	0.002578	1 in 388	n/a	n/a	3021
PML05	PRRC2A	SNV het	6	31595795	C	T	P515L	0.009777	1 in 102	0.001131	1 in 884	3121
PML05	PRRC2A	SNV hom	6	31602967	G	A	R1740H	0.149550	1 in 7	see het SNV	see het SNV	3132
PML05	NOD1	SNV het	7	30491693	C	T	R447H	0.006746	1 in 148	n/a	n/a	3148
PML		SNV							1 in			

05	CD72	hom	9	35612978	G	A	P254L	0.001558	642	n/a	n/a	3171
PML 05	MON1B	SNV het	16	77225440	G	T	D20Y	0.009826	1 in 102	n/a	n/a	3228
PML 05	MX2	SNV het	21	42752030	G	A	E177K	0.002074	1 in 482	n/a	n/a	3258
PML 09	MASP2	SNV hom	1	11090916	C	A	D371Y	0.655818	1 in 2	see het SNV	see het SNV	3002
PML 09	MASP2	SNV het	1	11106666	T	C	D120G	0.062380	1 in 16	0.0047 83	1 in 209	3005
PML 10	MASP2	SNV hom	1	11090916	C	A	D371Y	0.655818	1 in 2	see het SNV	see het SNV	3002
PML 10	MASP2	SNV het	1	11106666	T	C	D120G	0.062380	1 in 16	0.0047 83	1 in 209	3005

PML 10	MFN2	SNV het	1	12064892	G	A	R468H	0.006444	1 in 155	n/a	n/a	3008
PML 10	BLK	SNV het	8	11400849	C	T	P39L	0.006656	1 in 150	n/a	n/a	3160
PML 10	RNF125	SNV het	18	29645930	A	T	R190S	0.004555	1 in 220	n/a	n/a	3237
PML 13	ADGRL 2	SNV het	1	82416040	C	T	P323S	0.005425	1 in 184	n/a	n/a	3019
PML 13	LRRK2	SNV het	12	40740686	A	G	N2081D	0.005082	1 in 197	n/a	n/a	3198
PML 13	MMP9	SNV het	20	44640275	G	A	G296S	0.001154	1 in 867	n/a	n/a	3254
PML 14	MBL2	SNV het	10	54531235	C	T	G54D	0.248456	1 in 4	0.0086 25	1 in 116	3175
PML 14	MBL2	SNV het	10	54531242	G	A	R52C	0.138866	1 in 7	0.0086 25	1 in 116	3176
PML 15	CFHR2	SNV het	1	19692012 3	G	A	R116Q	0.001567	1 in 638	n/a	n/a	3030
PML 16	RANBP 2	SNV het	2	10938149 3	G	C	A1500P	0.008681	1 in 115	n/a	n/a	3046
PML 16	PRRC2A	SNV het	6	31604610	T	C	V2012A	0.001126	1 in 888	n/a	n/a	3135
PML 16	AIRE	SNV het	21	45708278	G	A	G197R	0.002143	1 in 467	n/a	n/a	3268
PML 18	WASHC 5	SNV het	8	12609537 1	T	C	I106V	0.001990	1 in 502	n/a	n/a	3166
PML 19	SAMD9 L	SNV het	7	92763720	G	A	A522V	0.006869	1 in 146	n/a	n/a	3154
PML 20	C8B	SNV hom	1	57422511	C	T	E108K	0.002691	1 in 372	n/a	n/a	3016
PML 20	HERC5	SNV het	4	89414196	G	A	G361R	0.003652	1 in 274	n/a	n/a	3072
PML 20	FAT4	SNV het	4	12623998 6	C	T	A807V	0.491020	1 in 2	0.0044 87	1 in 223	3086
PML 20	FAT4	SNV het	4	12624178 5	G	A	V1407I	0.036553	1 in 27	0.0044 87	1 in 223	3093
PML 20	PRRC2A	SNV het	6	31595795	C	T	P515L	0.022579	1 in 44	0.0028 72	1 in 348	3121
PML 20	PRRC2A	SNV hom	6	31602967	G	A	R1740H	0.258362	1 in 4	see het SNVs	see het SNVs	3132
PML 20	CCZ1	SNV het	7	5959528	T	C	V203A	0.007045	1 in 142	n/a	n/a	3146
PML 22	C8A	SNV hom	1	57378149	G	T	R485L	0.001712	1 in 584	n/a	n/a	3013
PML 23	LRRK2	SNV hom	12	40657700	C	G	N299K	0.006768	1 in 148	n/a	n/a	3192
PML 23	LRRK2	SNV hom	12	40702911	G	A	R1398H	0.006697	1 in 149	n/a	n/a	3197
PML 26	FAT4	SNV hom	4	12623998 6	C	T	A807V	0.196043	1 in 5	see het SNV	see het SNV	3086
PML 26	FAT4	SNV het	4	12637357 0	C	T	S2098F	0.088450	1 in 11	0.0108 40	1 in 92	3099
PML 31	FAT4	SNV hom	4	12623998 6	C	T	A807V	0.325510	1 in 3	see het SNV	see het SNV	3086
PML 31	FAT4	SNV het	4	12638983 2	G	A	R2285Q	0.056972	1 in 18	0.0069 94	1 in 143	3100
PML 31	RAB5B	SNV het	12	56383856	A	G	I97V	0.003416	1 in 293	n/a	n/a	3205
PML 31	FCER2	SNV het	19	7754222	G	A	R274C	0.004261	1 in 235	n/a	n/a	3241
PML 32	C3	SNV het	19	6707129	G	A	R735W	0.006398	1 in 156	n/a	n/a	3240
PML 33	BLK	SNV het	8	11412934	G	A	R167Q	0.008846	1 in 113	n/a	n/a	3162
PML 36	CHD2	SNV het	15	93545488	T	A	S1407T	0.004813	1 in 208	n/a	n/a	3216
PML 36	PLK1	SNV het	16	23700676	T	A	L463H	0.004422	1 in 226	n/a	n/a	3221

PML 37	MASP2	SNV het	1	11090287	C	T	D415N	0.036717	1 in 27	0.0033 04	1 in 303	3001
PML 37	MASP2	SNV het	1	11090916	C	A	D371Y	0.359923	1 in 3	0.0033 04	1 in 303	3002
PML 37	MOGS	SNV hom	2	74688884	G	A	R572W	0.008575	1 in 117	n/a	n/a	3041
PML 37	FAT4	SNV hom	4	12623998 6	C	T	A807V	0.325510	1 in 3	see het SNVs	see het SNVs	3086
PML 37	FAT4	SNV het	4	12624037 7	G	T	K937N	0.021429	1 in 47	0.0026 30	1 in 380	3088
PML 37	FAT4	SNV het	4	12640866 3	A	G	D2568G	0.031611	1 in 32	0.0038 80	1 in 258	3101
PML 38	LRRK2	SNV het	12	40702420	A	G	I1371V	0.001055	1 in 948	n/a	n/a	3195
PML 38	MMP9	SNV het	20	44640275	G	A	G296S	0.005222	1 in 191	n/a	n/a	3254
PML 39	CFHR4	SNV het	1	19687161 0	A	T	R40*	0.005310	1 in 188	n/a	n/a	3028
PML 39	CD36	SNV hom	7	80300449	T	G	Y249*	0.009686	1 in 103	n/a	n/a	3150
PML 39	NLRX1	SNV hom	11	11905297 5	G	T	A843S	0.001182	1 in 846	n/a	n/a	3188
PML 39	ICAM1	SNV het	19	10394724	C	T	P218L	0.003194	1 in 313	n/a	n/a	3242
PML 40	MLH1	SNV het	3	37092025	C	T	H477Y	0.003985	1 in 251	n/a	n/a	3060
PML 40	NOD1	SNV het	7	30491693	C	T	R447H	0.006746	1 in 148	n/a	n/a	3148
PML 40	TEK	SNV het	9	27183598	C	T	T201I	0.002950	1 in 339	n/a	n/a	3169
PML 40	NLRX1	SNV het	11	11905297 5	G	T	A843S	0.001734	1 in 577	n/a	n/a	3188
PML 40	MX1	SNV het	21	42815731	G	A	G293R	0.002418	1 in 414	n/a	n/a	3263
PML		SNV							1 in			

PML 41	PLAU	SNV het	10	75675086	T	C	Y314H	0.001730	1 in 578	n/a	n/a	3179
PML 41	IFI35	SNV het	17	41165632	G	A	R172Q	0.002886	1 in 347	n/a	n/a	3232
PML 43	SERPINB2	SNV het	18	61570402	G	A	A371T	0.002162	1 in 463	n/a	n/a	3239
PML 45	RAD50	SNV het	5	13192541 3	A	G	K307E	0.001430	1 in 700	n/a	n/a	3114
PML 45	MX1	SNV het	21	42807881	G	A	V75I	0.001409	1 in 710	n/a	n/a	3260
PML 46	NLRX1	SNV het	11	11904595 1	C	T	R547W	0.001904	1 in 525	n/a	n/a	3187
PML 48	NFAT5	SNV het	16	69725697	T	G	S563A	0.005437	1 in 184	n/a	n/a	3227
PML 49	LRRK2	SNV het	12	40677813	G	T	R793M	0.001409	1 in 710	n/a	n/a	3194
PML 51	TEK	SNV het	9	27197588	C	T	L444F	0.002245	1 in 446	n/a	n/a	3170
PML 52	USB1	SNV hom	16	58054099	C	G	Q199E	0.007344	1 in 136	n/a	n/a	3222
PML 53	PRRC2A	SNV het	6	31599370	G	C	E974Q	0.002791	1 in 358	n/a	n/a	3126
PML 54	RANBP2	SNV het	2	10938480 0	C	T	T2602M	0.003726	1 in 268	n/a	n/a	3048
PML 54	FAT4	SNV hom	4	12623998 6	C	T	A807V	0.196043	1 in 5	see het SNVs	see het SNVs	3086
PML 54	FAT4	SNV het	4	12637297 5	A	C	I1900L	0.022291	1 in 45	0.0027 32	1 in 366	3098
PML 54	FAT4	SNV het	4	12637357 0	C	T	S2098F	0.088450	1 in 11	0.0108 40	1 in 92	3099
PML 56	BLK	SNV het	8	11418856	C	T	R288C	0.002025	1 in 494	n/a	n/a	3164
PML 56	MSH5	SNV hom	6	31709045	C	T	L85F	0.001200	1 in 833	n/a	n/a	3139
PML 56	MSH5	SNV hom	6	31729925	C	T	P108S	0.001409	1 in 710	n/a	n/a	3141
PML 56	KMT2D	SNV het	12	49428694	T	C	D3419G	0.004830	1 in 207	n/a	n/a	3201
PML 61	CD22	SNV het	19	35823528	C	T	A36V	0.002712	1 in 369	n/a	n/a	3248
PML 62	CFH	SNV hom	1	19670977 4	G	T	E936D	0.004036	1 in 248	n/a	n/a	3024
PML 63	FAT4	SNV hom	4	12623998 6	C	T	A807V	0.325510	1 in 3	see het SNV	see het SNV	3086
PML 63	FAT4	SNV het	4	12638983 2	G	A	R2285Q	0.056972	1 in 18	0.0069 94	1 in 143	3100
PML 63	ACD	SNV het	16	67694078	G	T	L102M	0.002049	1 in 488	n/a	n/a	3226
PML 63	AIRE	SNV het	21	45713696	G	A	G228R	0.001432	1 in 698	n/a	n/a	3269
PML 64	PLAU	SNV het	10	75672059	G	A	G41R	0.001953	1 in 512	n/a	n/a	3177
PML 66	HERC6	SNV het	4	89352440	A	C	M709L	0.002213	1 in 452	n/a	n/a	3071
PML 68	RANBP2	SNV het	2	10938480 0	C	T	T2602M	0.003726	1 in 268	n/a	n/a	3048
PML 68	BLK	SNV het	8	11407690	C	T	R131W	0.001589	1 in 629	n/a	n/a	3161
PML 72	CFHR2	SNV het	1	19691873 2	G	T	R69L	0.007311	1 in 137	n/a	n/a	3029
PML 72	CLEC16A	SNV het	16	11272435	G	A	S1017N	0.008537	1 in 117	n/a	n/a	3218

Table 35 lists potential cause(s) of PML in the study, SNVs (het, hom, or phased comp het) for genes in **Table 31** with frequency <=1/1,00 but >1/1,000.

Table 36: Case-solving un-phased potential comp het SNVs with comp het frequency <=1/1,00

Sample ID	RefSeq Gene Symbol	Variant Type	Chromosome	Position (hg19)	Ref Allele	Alt Allele	Amino Acid Change	Variant Frequency Details (Ethnic specific)	Frequency (Reciprocal, Ethnic specific)	Predicted Compound Variant Frequency Details (Ethnic specific)	Predicted Compound Frequency (Reciprocal, Ethnic specific)	SEQ ID
MVGS1359	DNAJC21	SNV het	5	34929974	A	T	E17V	0	infinite	0	infinite	3108
MVGS1359	DNAJC21	SNV het	5	34937524	C	T	R178C	0	infinite	0	infinite	3109
MVGS1368	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.439204	1 in 2	0.005004	1 in 200	3132
MVGS1368	PRRC2A	SNV het	6	31605278	C	T	P2130L	0.045577	1 in 22	0.005004	1 in 200	3138
MVGS540-393b	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.439204	1 in 2	0.003733	1 in 268	3132
MVGS540-393b	PRRC2A	SNV het	6	31604894	C	T	R2075W	0.034000	1 in 29	0.003733	1 in 268	3136
MVGS540-393b	CFTR	SNV het	7	117230454	G	C	G515A	0.015182	1 in 66	6.99E-05	1 in 14,314	3155
MVGS540-393b	CFTR	SNV het	7	117232223	C	T	R607C	0.018407	1 in 54	6.99E-05	1 in 14,314	3157
MVGS811-13a	LRRK2	SNV het	12	40657700	C	G	N299K	0.139067	1 in 7	2.09E-06	1 in 479,386	3192
MVGS811-13a	LRRK2	SNV het	12	40702910	C	T	R1398C	6.01E-05	1 in 16,646	2.09E-06	1 in 479,386	3196
MVGS995-4a	MSH5	SNV het	6	31709045	C	T	L85F	0.041127	1 in 24	0.000446	1 in 2,242	3139

MVGS995-4a	MSH5	SNV het	6	31729925	C	T	P108S	0.043373	1 in 23	0.000446	1 in 2,242	3141
PML01	FAT4	SNV het	4	126239986	C	T	A807V	0.490241	1 in 2	0.000273	1 in 3,665	3086
PML01	FAT4	SNV het	4	126241248	C	G	Q1228E	0.002227	1 in 449	0.000273	1 in 3,665	3091
PML01	MSH5	SNV het	6	31709045	C	T	L85F	0.041127	1 in 24	0.000310	1 in 3,224	3139

PML01	MSH5	SNV het	6	31725978	C	G	R188G	0.030171	1 in 33	see other het SNVs	see other het SNVs	3140
PML01	MSH5	SNV het	6	31729925	C	T	P108S	0.043373	1 in 23	0.0003 27	1 in 3,057	3141
PML02	PRRC2A	SNV het	6	31595926	C	T	P559S	0.007481	1 in 134	6.54E-05	1 in 15,280	3122
PML02	PRRC2A	SNV het	6	31603045	A	G	D1766G	0.034995	1 in 29	6.54E-05	1 in 15,280	3133
PML02	MSH5	SNV het	6	31709045	C	T	L85F	0.041127	1 in 24	0.0004 46	1 in 2,242	3139
PML02	MSH5	SNV het	6	31729925	C	T	P108S	0.043373	1 in 23	0.0004 46	1 in 2,242	3141
PML04	FAT4	SNV het	4	126237697	A	C	E44A	0.000939	1 in 1,065	see other het SNV	see other het SNV	3079
PML04	FAT4	SNV het	4	126240968	A	T	E1134D	0.004166	1 in 240	see other het SNV	see other het SNV	3090
PML04	FAT4	SNV het	4	126411179	C	T	P2642L	0	infinite	0	infinite	3102
PML05	C8A	SNV het	1	57372463	C	T	T407I	0.003976	1 in 252	1.15E-05	1 in 86,659	3011
PML05	C8A	SNV het	1	57373778	G	A	D458N	0.011610	1 in 86	1.15E-05	1 in 86,659	3012
PML05	MOGS	SNV het	2	74688884	G	A	R572W	0.006223	1 in 161	0.000581	1 in 1,720	3041
PML05	MOGS	SNV het	2	74690378	C	T	D120N	0.373681	1 in 3	0.000581	1 in 1,720	3045
PML05	CFTR	SNV het	7	117232086	G	A	G561D	0	infinite	0	infinite	3156
PML05	CFTR	SNV het	7	117246776	T	C	L925P	0	infinite	0	infinite	3158
PML05	ERCC6L2	SNV het	9	98678698	G	C	L202F	0.001389	1 in 720	0.000115	1 in 8,675	3172
PML05	ERCC6L2	SNV het	9	98691137	T	C	V403A	0.331883	1 in 3	0.000115	1 in 8,675	3173
PML09	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.439204	1 in 2	0.002116	1 in 473	3132
PML09	PRRC2A	SNV het	6	31605016	T	C	F2083S	0.019275	1 in 52	see other het SNVs	see other het SNVs	3137
PML09	PRRC2A	SNV het	6	31605278	C	T	P2130L	0.045577	1 in 22	0.000220	1 in 4,553	3138
PML09	MSH5	SNV het	6	31709045	C	T	L85F	0.041127	1 in 24	0.0004 46	1 in 2,242	3139
PML09	MSH5	SNV het	6	31729925	C	T	P108S	0.043373	1 in 23	0.0004 46	1 in 2,242	3141
PML10	FAT4	SNV het	4	126239986	C	T	A807V	0.490241	1 in 2	0.004647	1 in 215	3086
PML10	FAT4	SNV het	4	126241106	C	G	S2951C	0.037914	1 in 26	0.004647	1 in 215	3104
PML10	PRRC2A	SNV het	6	31597469	A	C	K701Q	0.000186	1 in 5,385	1.62E-06	1 in 615,476	3124
PML10	PRRC2A	SNV het	6	31603045	A	G	D1766G	0.034995	1 in 29	1.62E-06	1 in 615,476	3133

PML10	MSH5	SNV het	6	31709045	C	T	L85F	0.041127	1 in 24	0.0004 46	1 in 2,242	3139
PML10	MSH5	SNV het	6	31729925	C	T	P108S	0.043373	1 in 23	0.0004 46	1 in 2,242	3141
PML15	SMARCAL1	SNV het	2	217285060	C	G	P165A	0.002972	1 in 337	1.58E-05	1 in 63,333	3050
PML15	SMARCAL1	SNV het	2	217288388	G	C	E241Q	0.021255	1 in 47	1.58E-05	1 in 63,333	3052
PML15	TCN2	SNV het	22	31006882	T	G	L30R	0.000691	1 in 1,447	6.93E-06	1 in 144,401	3270
PML15	TCN2	SNV het	22	31018975	T	C	L349S	0.040090	1 in 25	6.93E-06	1 in 144,401	3272
PML16	FAT4	SNV het	4	126239986	C	T	A807V	0.491020	1 in 2	0.005173	1 in 193	3086
PML16	FAT4	SNV het	4	126241720	T	C	L1385S	0.042138	1 in 24	0.005173	1 in 193	3092
PML16	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.508843	1 in 2	0.000143	1 in 6,979	3132
PML16	PRRC2A	SNV het	6	31604610	T	C	V2012A	0.001126	1 in 888	see other het SNVs	see other het SNVs	3135
PML16	PRRC2A	SNV het	6	31605016	T	C	F2083S	0.013115	1 in 76	3.69E-06	1 in 270,779	3137
PML16	PRRC2A	SNV het	6	31605278	C	T	P2130L	0.220930	1 in 5	6.22E-05	1 in 16,074	3138
PML19	MOGS	SNV het	2	74689335	G	T	D421E	0.000817	1 in 1,223	7.83E-05	1 in 12,779	3042
PML19	MOGS	SNV het	2	74690378	C	T	D120N	0.382905	1 in 3	7.83E-05	1 in 12,779	3045
PML20	PLD1	SNV het	3	171321023	C	T	V1024I	0.036230	1 in 28	0.000305	1 in 3,276	3066
PML20	PLD1	SNV het	3	171455697	G	C	P49A	0.033699	1 in 30	0.000305	1 in 3,276	3069
PML20	PRRC2A	SNV het	6	31595795	C	T	P515L	0.022579	1 in 44	0.001247	1 in 802	3121
PML20	PRRC2A	SNV het	6	31605278	C	T	P2130L	0.220930	1 in 5	0.001247	1 in 802	3138
PML21	PRRC2A	SNV het	6	31601735	G	A	R1563Q	0.087118	1 in 11	0.009566	1 in 105	3131
PML21	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.439204	1 in 2	0.009566	1 in 105	3132
PML25	NBAS	SNV het	2	15432775	C	T	R1518H	0.009143	1 in 109	0.000247	1 in 4,053	3033
PML25	NBAS	SNV het	2	15542352	C	T	R1004Q	0.107957	1 in 9	0.000247	1 in 4,053	3035
PML27	LRRK2	SNV het	12	40657700	C	G	N299K	0.139067	1 in 7	0.004847	1 in 206	3192
PML27	LRRK2	SNV het	12	40702911	G	A	R1398H	0.139408	1 in 7	0.004847	1 in 206	3197
PML28	FAT4	SNV het	4	126239623	G	A	S686N	5.99E-05	1 in 16,684	see other het SNVs	see other het SNVs	3085
PML28	FAT4	SNV het	4	126240255	A	G	I897V	3.00E-05	1 in 33,302	4.50E-10	1 in 2,222,442,272	3087
PML28	FAT4	SNV het	4	126240390	A	G	I942V	3.00E-05	1 in 33,354	4.50E-10	1 in 2,222,442,272	3089

PML30	MOGS	SNV het	2	74690371	C	T	R122H	0	infinite	0	infinite	3044
PML30	MOGS	SNV het	2	74690378	C	T	D120N	0.236 303	1 in 4	see other het SNV	see other het SNV	3045
PML30	MSH5	SNV het	6	31709045	C	T	L85F	0.041 127	1 in 24	0.0004 46	1 in 2,242	3139

PML30	MSH5	SNV het	6	31729925	C	T	P108S	0.043 373	1 in 23	0.0004 46	1 in 2,242	3141
PML31	PRRC2A	SNV het	6	31593603	A	G	Y265C	0	infinite	0	infinite	3120
PML31	PRRC2A	SNV het	6	31605278	C	T	P2130L	0.220 930	1 in 5	see other het SNV	see other het SNV	3138
PML33	PRRC2A	SNV het	6	31600558	C	T	R1370C	3.16E -05	1 in 31,613	see other het SNVs	see other het SNVs	3128
PML33	PRRC2A	SNV het	6	31600696	G	A	G1416S	0.024 081	1 in 42	1.90E-07	1 in 5,251,204	3129
PML33	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.439 204	1 in 2	3.47E-06	1 in 287,912	3132
PML33	PRRC2A	SNV het	6	31604591	C	T	P2006S	0.282 199	1 in 4	2.23E-06	1 in 448,095	3134
PML36	C8A	SNV het	1	57333311	C	A	A36E	0.002 695	1 in 371	0.0002 33	1 in 4,298	3010
PML36	C8A	SNV het	1	57378149	G	T	R485L	0.345 286	1 in 3	0.0002 33	1 in 4,298	3013
PML36	FAT4	SNV het	4	126238305	C	A	P247T	0.003 905	1 in 256	0.0004 79	1 in 2,086	3081
PML36	FAT4	SNV het	4	126239986	C	T	A807V	0.491 020	1 in 2	0.0004 79	1 in 2,086	3086
PML36	MSH5	SNV het	6	31709045	C	T	L85F	0.136 473	1 in 7	0.0072 28	1 in 138	3139
PML36	MSH5	SNV het	6	31729925	C	T	P108S	0.211 842	1 in 5	0.0072 28	1 in 138	3141
PML38	MASP2	SNV het	1	11090916	C	A	D371Y	0.306 713	1 in 3	0.0047 83	1 in 209	3002
PML38	MASP2	SNV het	1	11106666	T	C	D120G	0.062 380	1 in 16	0.0047 83	1 in 209	3005
PML38	PRRC2A	SNV het	6	31597451	G	A	A695T	0	infinite	0	infinite	3123
PML38	PRRC2A	SNV het	6	31600696	G	A	G1416S	0.024 081	1 in 42	see other het SNV	see other het SNV	3129
PML38	RLTPR	SNV het	16	67680806	G	A	V181M	0.090 170	1 in 11	0.0003 35	1 in 2,989	3223
PML38	RLTPR	SNV het	16	67685730	A	T	D821V	0.014 840	1 in 67	0.0003 35	1 in 2,989	3224
PML39	MASP2	SNV het	1	11094908	T	A	D355V	0.017 983	1 in 56	0.0009 74	1 in 1,027	3003
PML39	MASP2	SNV het	1	11106648	G	A	P126L	0.216 611	1 in 5	0.0009 74	1 in 1,027	3004
PML39	FAT4	SNV het	4	126239986	C	T	A807V	0.491 020	1 in 2	0.0037 66	1 in 266	3086
PML39	FAT4	SNV het	4	126336851	G	A	V2245I	0.030 677	1 in 33	0.0037 66	1 in 266	3096
PML40	C8A	SNV het	1	57378149	G	T	R485L	0.161 223	1 in 6	6.97E-06	1 in 143,528	3013
PML40	C8A	SNV het	1	57383295	G	A	G554D	0.000 173	1 in 5,785	6.97E-06	1 in 143,528	3014
PML40	MSH5	SNV het	6	31709045	C	T	L85F	0.090 941	1 in 11	0.0022 40	1 in 446	3139
PML40	MSH5	SNV het	6	31729925	C	T	P108S	0.098 531	1 in 10	0.0022 40	1 in 446	3141
PML41	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.508 843	1 in 2	0.0016 68	1 in 599	3132
PML41	PRRC2A	SNV het	6	31605016	T	C	P2083S	0.013 115	1 in 76	see other het SNVs	see other het SNVs	3137

PML41	PRRC2A	SNV het	6	31605278	C	T	P2130L	0.220 930	1 in 5	0.0007 24	1 in 1,381	3138
PML41	TRAF3IP2	SNV het	6	111913058	G	A	R78W	0.014 991	1 in 67	0.0009 15	1 in 1,093	3144
PML41	TRAF3IP2	SNV het	6	111913262	C	T	D10N	0.244 114	1 in 4	0.0009 15	1 in 1,093	3145
PML43	PRRC2A	SNV het	6	31601735	G	A	R1563Q	0.087 118	1 in 11	see other het SNVs	see other het SNVs	3131
PML43	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.439 204	1 in 2	0.0095 66	1 in 105	3132
PML43	PRRC2A	SNV het	6	31604591	C	T	P2006S	0.282 199	1 in 4	0.0061 46	1 in 163	3134
PML43	LRRK2	SNV het	12	40657700	C	G	N299K	0.139 067	1 in 7	0.0048 47	1 in 206	3192
PML43	LRRK2	SNV het	12	40702911	G	A	R1398H	0.139 408	1 in 7	0.0048 47	1 in 206	3197
PML44	C8A	SNV het	1	57333311	C	A	A36E	0.015 385	1 in 65	0.0003 27	1 in 3,060	3010
PML44	C8A	SNV het	1	57378149	G	T	R485L	0.084 963	1 in 12	0.0003 27	1 in 3,060	3013
PML44	PLD1	SNV het	3	171379953	C	T	R708H	3.00E -05	1 in 33,309	see other het SNV	see other het SNV	3067
PML44	PLD1	SNV het	3	171431726	C	G	E290Q	0	infinite	0	infinite	3068
PML44	FAT4	SNV het	4	126239986	C	T	A807V	0.490 241	1 in 2	0.0051 71	1 in 193	3086
PML44	FAT4	SNV het	4	126412154	G	A	R2967K	0.042 189	1 in 24	0.0051 71	1 in 193	3105
PML46	SAMD9	SNV het	7	92732769	T	C	D881G	0.037 350	1 in 27	0.0011 01	1 in 908	3151
PML46	SAMD9	SNV het	7	92733766	C	A	V549L	0.117 932	1 in 8	0.0011 01	1 in 908	3152
PML46	CFTR	SNV het	7	117230454	G	C	G515A	0.008 531	1 in 117	2.00E-05	1 in 49,970	3155
PML46	CFTR	SNV het	7	117232223	C	T	R607C	0.009 383	1 in 107	2.00E-05	1 in 49,970	3157
		SNV						0.490		0.0108		

PML49	FAT4	SNV het	4	126239986	C	T	A807V	241	1 in 2	40	1 in 92	3086	
PML49	FAT4	SNV het	4	126373570	C	T	S2098F	0.088 450	1 in 11	0.0108 40	1 in 92	3099	
PML50	PRRC2A	SNV het	6	31601735	G	A	R1563Q	0.010 692	1 in 94	see other het SNVs	see other het SNVs	3131	
PML50	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.508 843	1 in 2	0.0013 60	1 in 735	3132	
PML50	PRRC2A	SNV het	6	31604591	C	T	P2006S	0.324 640	1 in 3	0.0008 68	1 in 1,152	3134	
PML51	PRRC2A	SNV het	6	31600696	G	A	G1416S	0.024 081	1 in 42	see other het SNVs	see other het SNVs	3129	
PML51	PRRC2A	SNV het	6	31602967	G	A	R1740H	0.439 204	1 in 2	0.0026 44	1 in 378	3132	
PML51	PRRC2A	SNV het	6	31604591	C	T	P2006S	0.282 199	1 in 4	0.0016 99	1 in 589	3134	
PML55	FAT4	SNV het	4	126239986	C	T	A807V	0.490 241	1 in 2	see other het SNVs	see other het SNVs	3086	
PML55	FAT4	SNV het	4	126412106	C	G	S2951C	0.037 914	1 in 26	0.0046 47	1 in 215	3104	
PML55	FAT4	SNV het	4	126412154	G	A	R2967K	0.042 189	1 in 24	0.0051 71	1 in 193	3105	
PML55	SAMD9	SNV het	7	92732769	T	C	D881G	0.046 152	1 in 22	0.0021 09	1 in 474	3151	
PML55	SAMD9	SNV het	7	92733766	C	A	V549L	0.182 832	1 in 5	0.0021 09	1 in 474	3152	
PML56	FAT4	SNV het	4	126239986	C	T	A807V	0.490 241	1 in 2	0.0051 71	1 in 193	3086	
PML56	FAT4	SNV het	4	126412154	G	A	R2967K	0.042 189	1 in 24	0.0051 71	1 in 193	3105	
PML56	SAMD9	SNV het	7	92732769	T	C	D881G	0.046 152	1 in 22	0.0021 09	1 in 474	3151	
PML56	SAMD9	SNV het	7	92733766	C	A	V549L	0.182 832	1 in 5	0.0021 09	1 in 474	3152	
PML57	ITSN1	SNV het	21	35122475	C	T	P125L	3.04E -05	1 in 32,871	1.01E-06	1 in 987,921	3256	
PML57	ITSN1	SNV het	21	35239562	A	G	K102R	0.133 092	1 in 8	1.01E-06	1 in 987,921	3257	
PML58	MASP2	SNV het	1	11090916	C	A	D371Y	0.359 923	1 in 3	see other het SNVs	see other het SNVs	3002	
PML58	MASP2	SNV het	1	11106648	G	A	P126L	0.216 611	1 in 5	see other het SNVs	see other het SNVs	3004	
PML58	MASP2	SNV het	1	11106673	G	A	R118C	0	infinite	0	infinite	3006	
PML58	SMARCA L1	SNV het	2	217285104	C	G	S179R	0.148 576	1 in 7	0.0048 48	1 in 206	3051	
PML58	SMARCA L1	SNV het	2	217288388	G	C	E241Q	0.130 527	1 in 8	0.0048 48	1 in 206	3052	
PML58	FAT4	SNV het	4	126238090	G	T	R175L	0.001 241	1 in 806	0.0001 52	1 in 6,565	3080	
PML58	FAT4	SNV het	4	126239986	C	T	A807V	0.491 020	1 in 2	0.0001 52	1 in 6,565	3086	
PML59	MOGS	SNV het	2	74688563	C	T	G679S	0.091 076	1 in 11	0.0087 18	1 in 115	3040	
PML59	MOGS	SNV het	2	74690378	C	T	D120N	0.382 905	1 in 3	0.0087 18	1 in 115	3045	
PML59	PEPD	SNV het	19	33968991	T	A	L106V	0.096 099	1 in 10	0.0017 47	1 in 572	3246	
PML59	PEPD	SNV het	19	33980963	G	A	R148C	0.072 732	1 in 14	0.0017 47	1 in 572	3247	
PML60	FAT4	SNV het	4	126239241	G	A	V559I	0	infinite	0	infinite	3082	
PML60	FAT4	SNV het	4	126411493	C	T	P2747S	0	infinite	0	infinite	3103	
PML61	C8A	SNV het	1	57372463	C	T	T407I	0.087 199	1 in 11	0.0075 27	1 in 133	3011	
PML61	C8A	SNV het	1	57378149	G	T	R485L	0.345 286	1 in 3	0.0075 27	1 in 133	3013	
PML63	C8A	SNV het	1	57372463	C	T	T407I	0.087 199	1 in 11	0.0075 27	1 in 133	3011	
PML63	C8A	SNV het	1	57378149	G	T	R485L	0.345 286	1 in 3	0.0075 27	1 in 133	3013	
PML63	PLD1	SNV het	3	171321023	C	T	V1024I	0.036 230	1 in 28	see other het SNVs	see other het SNVs	3066	
PML63	PLD1	SNV het	3	171455697	G	C	P49A	0.033 699	1 in 30	0.0003 05	1 in 3,276	3069	
PML63	PLD1	SNV het	3	171455739	A	G	F35L	0.011 729	1 in 85	0.0001 06	1 in 9,413	3070	
PML63	PEPD	SNV het	19	33892731	A	G	I118T	0.042 743	1 in 23	see other het SNVs	see other het SNVs	3245	
PML63	PEPD	SNV het	19	33968991	T	A	L106V	0.096 099	1 in 10	0.0010 27	1 in 974	3246	
PML63	PEPD	SNV het	19	33980963	G	A	R148C	0.072 732	1 in 14	0.0007 77	1 in 1,287	3247	
PML64	FAT4	SNV het	4	126239986	C	T	A807V	0.491 020	1 in 2	0.0017 00	1 in 588	3086	
PML64	FAT4	SNV het	4	126373570	C	T	S2098F	0.013 846	1 in 72	0.0017 00	1 in 588	3099	
PML72	EGF	SNV het	4	110866260	G	C	D257H	0.042 672	1 in 23	0.0004 80	1 in 2,084	3075	
PML72	EGF	SNV het	4	110897252	C	G	S596R	0.044 974	1 in 22	0.0004 80	1 in 2,084	3076	

Table 36 lists potential cause(s) of PML in the study, un-phased potential comp het SNVs for genes in **Table 31** with frequency <=1/100.

Table 37: Gene burden genes (same criteria as **Table 13**)

RefSeq Gene Symbol	Ave PML Cases	Ave ExAC Cases	Ave ExAC Samples	Ave FET	Ave FET corr (270)	Ave OR	Ethnicity	Ethnicity Overlap
ORC4	8	247	33171	1.51E-09	4.08E-07	29.62	EUR	
PRRC2A	17	729	30571	1.54E-08	4.17E-06	27.58	EUR	EUR + AFR
PINK1	7	315	24054	1.79E-06	4.82E-04	14.26	EUR	
FAT4	23	6454	32974	2.16E-05	5.84E-03	4.64	EUR	EUR + AFR
LRRK2	9	1431	32539	7.96E-05	2.15E-02	6.41	EUR	
PLD1	12	257	5158	5.87E-11	1.58E-08	25.43	AFR	
PRRC2A	8	52	4536	3.27E-08	8.83E-06	57.80	AFR	EUR + AFR
FAT4	16	1259	4867	5.64E-06	1.52E-03	9.14	AFR	EUR + AFR
CTC1	6	159	4730	5.64E-05	1.52E-02	11.50	AFR	

Table 37 lists genes in **Table 31** for which the total burden of damaging variants (heterogzygous) was found to be statistically greater in PML cases versus ExAC controls. As was observed in gene burden analysis performed on the original set of 435 genes, only the category of heterozygous damaging variants yielded significant genes for the new set of 270 immune dysregulation genes.

Table 38: Variant burden Tier 1 SNVs (same criteria as **Table 14**)

RefSeq Gene Symbol	Variant (hg19)	Genotype	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
CFHR3	chr1:196759282, C>T	hom	3	3	0	57/ 28,802	41/ 2,986	16/ 4,514	36.90	0.000107	11.97	0.003235	29.76	1.250E-07
FAT4	chr4:126412634, C>G	het	0	4	0	76/ 33,339	176/ 5,194	19/ 5,784	NA	NA	6.71	0.005182	9.85	0.000959
PLAU	chr10:75673748, A>C	hom	0	3	0	0/ 33,321	141/ 5,198	0/ 5,780	NA	NA	5.98	0.019009	14.02	0.001563
MDC1	chr6:30675830, T>A	het	1	4	1	711/ 31,685	118/ 4,585	148/ 5,730	1.01	1.000000	8.91	0.001977	3.94	0.005909
MDC1	chr6:30680721, G>A	het	1	4	1	712/ 31,652	125/ 4,581	148/ 5,730	1.01	1.000000	8.39	0.002435	3.90	0.006166
RAB5B	chr12:56385915, GGGA>G	het	0	3	0	189/ 33,331	26/ 5,198	35/ 5,779	NA	NA	33.15	0.000192	7.89	0.007567
CCBE1	chr18:57103126, G>A	het	2	3	0	768/ 33,368	21/ 5,201	83/ 5,789	2.02	0.269436	41.11	0.000107	3.84	0.012489
PLD1	chr3:171321023, C>T	het	0	3	0	101/ 33,301	188/ 5,189	20/ 5,742	NA	NA	4.43	0.039674	6.36	0.013380
MYSM1	chr1:59151311, G>T	het	16	7	0	7,209/ 33,225	1,248/ 4,889	645/ 5,765	2.06	0.026498	1.46	0.451989	1.87	0.017528
TCN2	chr22:31008867, T>C	het	6	0	0	1,216/ 33,109	41/ 5,144	69/ 5,776	4.14	0.005279	NA	NA	3.02	0.019108
MOGS	chr2:74690378, C>T	het	17	10	1	7,884/ 33,364	1,877/ 4,902	2,161/ 5,783	2.03	0.031024	1.47	0.378471	1.80	0.021224
WRAP53	chr17:7592168, C>G	het	15	8	2	6,828/ 33,192	1,712/ 5,136	1,880/ 5,783	2.00	0.037960	1.23	0.647415	1.80	0.023425
MOGS	chr2:74690039, G>A	het	17	9	1	8,286/ 33,133	1,610/ 4,889	2,178/ 5,776	1.89	0.053205	1.53	0.356038	1.65	0.044663
OAS2	chr12:113448288, A>G	het	26	8	1	14,781/ 33,329	560/ 5,190	1,544/ 5,784	1.81	0.067220	5.09	0.001060	1.62	0.048285
TMC8	chr17:76130947, G>T	het	7	9	0	5,032/ 33,022	848/ 5,148	492/ 5,758	1.05	0.834739	3.80	0.004015	1.75	0.059738
NBAS	chr2:15674686, T>C	hom	25	8	2	15,155/ 33,366	406/ 5,202	1,556/ 5,786	1.58	0.132924	7.27	0.000118	1.59	0.064503
NBAS	chr2:15607842, T>C	hom	25	7	2	15,009/ 32,894	221/ 5,133	1,491/ 5,637	1.57	0.172287	11.11	2.070E-05	1.52	0.085087
SERPINA1	chr14:94847262, T>A	het	6	1	0	1,942/ 33,369	81/ 5,203	253/ 5,789	2.56	0.041248	3.16	0.283151	2.05	0.092465
FAT4	chr4:126412154, G>A	het	5	0	0	1,407/ 33,350	99/ 5,192	112/ 5,785	2.91	0.037262	NA	NA	2.03	0.112929
NCF4	chr22:37271882, T>C	het	17	5	0	9,101/ 33,366	522/ 5,202	1,402/ 5,788	1.68	0.092908	2.80	0.053500	1.39	0.213191
MMP9	chr20:44640959, G>A	hom	18	5	0	10,856/ 32,887	329/ 4,800	410/ 5,748	1.40	0.265273	4.25	0.012513	1.34	0.278668
CTC1	chr17:8138569, C>G	het	0	4	0	1,235/ 32,722	42/ 4,721	92/ 5,747	NA	NA	26.21	4.110E-05	1.85	0.286669
OAS1	chr12:113357237, G>C	het	22	6	3	14,796/ 33,345	648/ 5,201	1,548/ 5,784	1.25	0.452928	2.81	0.039148	1.28	0.325931
OAS1	chr12:113357209, G>A	het	22	6	3	14,788/ 33,270	648/ 5,201	1,548/ 5,780	1.25	0.544210	2.81	0.039148	1.28	0.326337
CD5	chr11:60893235, C>T	hom	15	10	0	9,028/ 33,233	1,386/ 5,189	3,098/ 5,770	1.39	0.310432	2.49	0.044954	1.26	0.364341
OAS1	chr12:113357442, G>A	het	22	6	3	13,136/ 28,065	594/ 4,740	1,355/ 4,892	1.14	0.762822	2.79	0.040180	1.19	0.466504
C7	chr5:40964852, A>C	het	16	9	0	12,068/ 33,346	552/ 4,899	1,966/ 5,755	1.01	1.000000	5.91	0.000253	1.12	0.703241
SRP54	chr14:35497285, T>C	het	14	5	0	9,945/ 33,265	391/ 5,189	887/ 5,776	1.09	0.744651	3.83	0.018196	1.10	0.783330
NLRP2	chr19:55494157, G>A	het	9	0	0	2,739/ 33,004	1,241/ 5,094	1,774/ 5,706	2.84	0.009214	NA	NA	0.98	1.000000

Table 38 lists top tier of variants found to be significant on the basis of variant burden analysis for genes in **Table 31**.

Table 39: Variant burden Tier 2 SNVs (same criteria as **Table 15**)

R	Va		P	P	P	E	E	E	P	PM	PM	PM	PM	PM

efSeq Gene Symbol	riant (hg19)	Genotype	ML EUR 44	dL AFR 21	ML LAT 5	ExAC EUR	ExAC AFR	ExAC LAT	ML EUR OR	IL EUR FET	IL AFR OR	IL AFR FET	IL ALL OR	L ALL FET
CLPB	chr11: 72145307, C>G	het	0	2	1	5/ 32,307	27/ 4,744	1/ 5,748	NA	NA	18.39	0.006992	58.03	2.860E-05
NOD1	chr7: 30491421, G>T	het	0	2	0	3/ 33,152	1/ 5,131	1/ 5,782	NA	NA	540.00	4.740E-05	259.18	5.180E-05

MDC1	chr6: 30673403, A>G	het	2	1	0	32/ 32,318	7/ 4,973	7/ 5,692	48.04	0.000986	35.47	0.033172	41.79	7.190E-05
PLAUR	chr19: 44153248, T>C	het	4	0	0	295/ 33,361	8/ 5,200	41/ 5,787	11.21	0.000644	NA	NA	7.75	0.002261
PLEKHM1	chr17: 43555253, A>G	hom	4	1	1	735/ 30,909	8/ 4,484	52/ 5,406	4.11	0.020571	27.98	0.041216	4.72	0.002518
TFPI	chr2: 188349523, A>G	het	0	2	0	2/ 32,880	63/ 5,103	2/ 5,620	NA	NA	8.42	0.028496	19.11	0.005543
C8B	chr1: 57409459, C>A	het	2	2	0	408/ 33,367	18/ 5,203	31/ 5,788	3.85	0.101548	30.32	0.002800	5.82	0.006119
FAT4	chr4: 126241248, C>G	het	2	0	0	74/ 33,236	3/ 4,867	7/ 5,779	21.34	0.004576	NA	NA	15.34	0.008383
C9	chr5: 39311336, A>T	het	2	0	0	85/ 33,354	4/ 5,201	5/ 5,770	18.63	0.005917	NA	NA	13.83	0.010166
TMC8	chr17: 76129619, C>T	het	2	0	0	115/ 33,116	2/ 5,174	3/ 5,778	13.67	0.010601	NA	NA	10.77	0.016190
EGF	chr4: 110929301, T>C	het	2	0	0	109/ 33,340	2/ 5,201	10/ 5,739	14.52	0.009462	NA	NA	10.73	0.016294
ATG7	chr3: 11402163, G>A	het	0	2	0	48/ 33,370	47/ 5,203	27/ 5,789	NA	NA	11.55	0.016157	10.67	0.016487
ACD	chr16: 67694044, C>T	het	2	0	0	111/ 31,250	10/ 4,297	7/ 5,689	13.36	0.011075	NA	NA	9.45	0.020647
ICAM1	chr19: 10395141, G>A	het	2	0	0	146/ 32,810	6/ 4,915	14/ 5,732	10.65	0.016849	NA	NA	7.96	0.028184
ATG5	chr6: 106740989, T>C	het	3	0	0	404/ 33,200	6/ 5,111	22/ 5,751	5.94	0.016651	NA	NA	4.52	0.032035
NCF2	chr1: 183532364, T>A	het	3	0	0	390/ 33,366	13/ 5,203	48/ 5,787	6.19	0.014980	NA	NA	4.36	0.035082
MCM5	chr22: 35806756, G>A	het	3	0	0	405/ 33,370	14/ 5,203	64/ 5,789	5.96	0.016535	NA	NA	4.07	0.041554
EGF	chr4: 110865044, G>C	het	2	0	1	20/ 33,364	506/ 5,202	25/ 5,787	79.39	0.000385	NA	NA	3.56	0.057209
EGF	chr4: 110864533, C>T	het	2	0	1	22/ 33,276	549/ 5,149	29/ 5,777	71.98	0.000462	NA	NA	3.25	0.070527
FAT4	chr4: 126238090, G>T	het	0	2	0	266/ 33,051	6/ 4,835	19/ 5,744	NA	NA	84.72	0.000491	4.38	0.080399
EGF	chr4: 110932508, C>A	het	2	0	0	13/ 33,325	261/ 5,185	22/ 5,767	122.02	0.000176	NA	NA	4.37	0.080707
PRRC2A	chr6: 31605016, T>C	het	1	2	0	610/ 31,648	60/ 4,575	102/ 5,730	1.18	0.580000	7.92	0.031896	2.39	0.139143
SAMD9	chr7: 92733766, C>A	hom	0	2	0	371/ 33,364	87/ 5,111	30/ 5,783	NA	NA	6.08	0.050453	2.64	0.181292
RNF125	chr18: 29645930, A>T	het	2	0	0	152/ 33,368	324/ 5,203	44/ 5,789	10.41	0.017592	NA	NA	2.48	0.198848

Table 39 lists second tier of variants found to be significant on the basis of variant burden analysis for genes in **Table 31**.

Table 40: Summary of the subset of genes found to harbor variants of interest in the 70 PML cases (from Tables 34-39)

RefSeq Gene Symbol	Tables 34-36 Total Case Level Solutions	Table 34 SNV Frequency (Het, Hom, Comp Het) <= 1/1,000	Table 35 SNV Frequency (Het, Hom, Comp Het) <= 1/100 to > 1/1,000	Table 36 Unphased SNV Frequency (Comp Het) <= 1/100	Table 37 Gene Burden	Table 38 Variant Burden, PML Cases >= 5 EUR and/or 3 AFR	Table 39 Variant Burden, PML Cases 2-4 EUR and/or 2 AFR
ACD	1		1 AFR				EUR
ADGRL2	4	2 AFR	1 AFR, 1 EUR				
AIRE	3	1 EUR	2 AFR				
ATG5	0						EUR
ATG7	3	3 EUR					AFR
BLK	6	1 EUR	5 EUR				
BRD4	1	1 EUR					
C3	1		1 EUR				

RefSeq Gene Symbol	Tables 34-36 Total Case Level Solutions	Table 34 SNV Frequency (Het, Hom, Comp Het) <= 1/1,000	Table 35 SNV Frequency (Het, Hom, Comp Het) <= 1/100 to > 1/1,000	Table 36 Un-phased SNV Frequency (Comp Het) <= 1/100	Table 37 Gene Burden	Table 38 Variant Burden, PML Cases >= 5 EUR and/or 3 AFR	Table 39 Variant Burden, PML Cases 2-4 EUR and/or 2 AFR
C7	1	1 EUR				AFR, EUR	
C8A	7		1 EUR	3 AFR, 1EUR, 2 LAT			
C8B	1		1 AFR				AFR, EUR
C9	0						EUR
CAMLG	1	1 AFR					
CCBE1	0					AFR, EUR	
CCZ1	1		1 AFR				
CD22	1		1 AFR				
CD36	1		1 AFR				
CD37	1	1 AFR					
CD5	1	1 EUR					
CD72	1		1LAT				
CFH	1		1 AFR				
CFHR1	1	1LAT					
CFHR2	2		1 AFR, 1 LAT				
CFHR3	0					AFR, EUR	
CFHR4	1		1 AFR				
CFHR5	1	1 AFR					
CFTR	3			1 EUR, 2 LAT			
CHD2	1		1 AFR				
CLEC16A	2	1 EUR	1 AFR				
CLPB	0						AFR, EUR
COPA	1	1LAT					
CTC1	0				AFR	AFR	
DNAJC21	1			1 EUR			
EGF	1			1 AFR			EUR, LAT
ERCC6L2	1			1LAT			
FAT4	23	3 AFR, 4 EUR	4 AFR, 2 EUR	5 AFR, 9 EUR	AFR, EUR	AFR, EUR	AFR, EUR
FCER2	1		1 AFR				
HERC5	1		1 AFR				
HERC6	1		1 AFR				
ICAM1	1		1 AFR				EUR
IFI3 5	1		1 AFR				
IFIT1	1	1 EUR					
IFIT3	1	1 AFR					
IL4	1	1 EUR					
ITSN1	1	1 EUR		1 EUR			
KMT2D	6	1 AFR, 4 EUR	1 EUR				
KRAS	2	2 EUR					
LRRK2	8	1 AFR, 1 EUR	1 AFR, 3 EUR	3 EUR	EUR		
MASP2	7	2 AFR	1 AFR, 2 EUR	2 AFR, 1 EUR			
MBL2	2	1LAT	1 EUR				
MCM5	0						EUR
MDC1	0					AFR, EUR, LAT	AFR, EUR
MFN2	3	2 AFR	1 EUR				
MLH1	3	1 AFR, 1 EUR	1LAT				

RefSeq Gene Symbol	Tables 34-36 Total Case Level Solutions	Table 34 SNV Frequency (Het, Hom, Comp Het) <= 1/1,000	Table 35 SNV Frequency (Het, Hom, Comp Het) <= 1/100 to > 1/1,000	Table 36 Un-phased SNV Frequency (Comp Het) <= 1/100	Table 37 Gene Burden	Table 38 Variant Burden, PML Cases >= 5 EUR and/or 3 AFR	Table 39 Variant Burden, PML Cases 2-4 EUR and/or 2 AFR
MMP9	2		1 AFR, 1 EUR			AFR, EUR	
MOGS	5		1 AFR	2 AFR, 1 EUR, 1 LAT		AFR, EUR, LAT	
MON1A	1	1 EUR					
MON1B	1		1LAT				
MSH2	2	1 EUR, 1 LAT					
MSH5	9		1 EUR	1 AFR, 6 EUR, 1 LAT			
MX1	8	2 AFR, 4 EUR	1 EUR, 1 LAT				
MX2	1		1LAT				
MYSM1	0					AFR, EUR	
NBAS	2	1 EUR		1 EUR		AFR, EUR, LAT	
NCF1	1	1 EUR					
NCF2	0						EUR
NCF4	0					AFR, EUR	
NFAT5	1		1 EUR				
NLRP2	0					EUR	
NLRX1	4	1 EUR	1 AFR, 2 LAT				
NOD1	2		2 LAT				AFR
OAS1	0					AFR, EUR, LAT	
OAS2	0					AFR, EUR, LAT	
OAS3	2	2 EUR					
ORC4	0				EUR		
PARN	1	1 EUR					
PEPD	2			2 AFR			
PINK 1	0				EUR		
PLAU	2		2 AFR			AFR	
PLAUR	0						EUR
PLCG1	1	1 AFR					
PLD1	3	1 AFR		2 AFR, 1 EUR	AFR	AFR	
PLEKHM1	0						AFR, EUR, LAT
PLK1	2		1 AFR, 1 EUR				
PLXNB1	3	1 AFR, 2 EUR					
PRRC2A	22	2 AFR, 6 EUR, 1 LAT	2 AFR, 4 EUR, 1 LAT	5 AFR, 10 EUR	AFR, EUR		AFR, EUR
RAB5A	1	1 EUR					
RAB5B	1		1 AFR			AFR	
RAD50	1		1 EUR				
RANBP2	4	1 EUR	1 AFR, 2 EUR				
RELA	1	1 EUR					
RLTPR	1			1 EUR			
RNF125	3	1 AFR	2 EUR				EUR
RPSA	1	1 EUR					
RSAD2	1	1 EUR					
SAMD9	3			2 EUR, 1 LAT			AFR

RefSeq Gene Symbol	Tables 34-36 Total Case Level Solutions	Table 34 SNV Frequency (Het, Hom, Comp Het) <= 1/1,000	Table 35 SNV Frequency (Het, Hom, Comp Het) <= 1/100 to > 1/1,000	Table 36 Un-phased SNV Frequency (Comp Het) <= 1/100	Table 37 Gene Burden	Table 38 Variant Burden, PML Cases >= 5 EUR and/or 3 AFR	Table 39 Variant Burden, PML Cases 2-4 EUR and/or 2 AFR
SAMD9L	2	1 AFR	1 AFR				
SERPINA1	0					AFR, EUR	
SERPINB2	1		1 EUR				
SMARCAL1	2			1 AFR, 1 LAT			
SMURF2	1	1 AFR					
SRP54	0					AFR, EUR	
TBC1D17	1	1 LAT					
TCN2	1			1LAT		EUR	
TEK	3	1 EUR	1 EUR, 1 LAT				
TFPI	0						AFR
TMC8	0					AFR, EUR	EUR
TP53AIP1	1	1 EUR					
TRAF3IP2	1			1 AFR			
USB1	1		1 EUR				
USP3	1	1 EUR					
VEGFA	1	1 EUR					
WASHC5	2	1 EUR	1 EUR				
WRAP53	0					AFR, EUR, LAT	
XAF1	1	1 AFR					

Table 40 lists a summary of the subset of 270 genes found to harbor variants of interest in the 70 PML cases.

Example 23 - Protein-Protein Interaction Analysis

[0405] This example contains analysis of protein-protein interactions (e.g., pathway analysis) for the two sets of PML candidate genes: PML-435 (Tables 6, 25A, and 25B) and PML-270 (**Table 31**).

[0406] An integrated list of top candidate PML genes was generated from the three analysis methods used herein: 1) case-solving approach (e.g., Tables 7-9 and 34-36); 2) gene burden analysis (e.g., **Tables 13 and 37**); and 3) variant burden analysis for top findings (e.g., tier 1, **Tables 14 and 38**) and second tier findings (e.g., tier 2, **Tables 15 and 39**). A gene-scoring metric was applied (see **Table 42** for results) as follows: 1) case-solving, the total number of unique PML cases that are potentially solved were summed up for each gene, 2) gene burden, the number of ethnicities that were found to be significant for a given gene are reported (e.g., if AFR and EUR cases were significant in their respective analyses, then the reported number is 2), 3) variant burden (tier 1 or tier 2), the number of ethnicities that were found to be significant for a given variant are reported, and 4) gene score total, the sum of steps 1-3 (case-solving, gene burden, and variant burden) were summed up to yield the total score for a given gene wherein a higher number is a stronger candidate gene (e.g., FAT4 was the top-scoring gene with a gene score total of 29). A total of 255 genes (derived from Tables 6, 25A, 25B, and 31) were found to have variants of interest in the set of 70 PML cases for which WES and array CGH data were obtained. A gene score of >3 was set as the cutoff for pathway analysis, which yielded the 74 genes that are reported in

Table 42.

[0407] Pathway analysis was performed on the set of 74 genes reported in **Table 42** using the String database resource (Szklarczyk D et al. 2017, PMID 27924014). Default settings were used except for the "minimum required interaction score", which was set to "high confidence" (the default setting is "medium confidence").

[0408] The String database analysis yielded three main interaction networks:

1. 1) 24-gene network (the output is depicted in Figure 14 and annotated in **Table 42**): BLK, CARD11, CFTR, EGF, IFIH1, ITSN2, MAVS, MMP9, MX1, NFKB1, NLRX1, NOD1, NOD2, OAS1, OAS2, PIK3CD, PLAU, PLCG2, RNF125, SAMD9, TEK, TICAM1, TLR4, ZAP70;

2. 2) 13-gene network: ATM, ATR, BLM, DCLRE1C, LRRK2, MDC1, MLH1, MSH5, POLE, PRKDC, RANBP2, RNF168, RTEL1; and
3. 3) 3-gene network: C7, C8A, C8B.

Table 42: Analysis of Protein-Protein Interactions

RefSeq Gene Symbol	Gene Source (Table)	Case-Solving	Gene Burden	Variant Burden Tier 1	Variant Burden Tier 2	Gene Score Total	24-Gene Network	GO:0002250	GO:0045087	GO:0006955	GO:0002252	GO:0002253	GO:0042113	GO:0032479	GO:0030217	GO:0006958
ADGRL2	31	4				4										
AIRE	31	3				3										
AP3B1	6	5	1			6										
ATG7	31	3			1	4					X					
ATM	6	8			2	10							X			
ATR	6	3			1	4										
BLK	31	6				6	X		X	X		X			X	
BLM	6	2			1	3									X	
C7	31	1		2		3		X	X	X	X	X				X
C8A	31	7				7		X	X	X	X	X				X
C8B	31	1			2	3		X	X	X	X	X				X
CARD11	6	3				3	X		X	X		X			X	
CFTR	31	3				3	X									
CHD7	6	4		1	1	6									X	
DCLRE1C	6	10				10							X			
DOCK8	25A, 25B	11			2	13										
EGF	31	1			2	3	X		X	X						
EPG5	6	12	1			13										
FAT4	31	23	2	2	2	29										
GFI1	6	4			2	6										
HIVEP1	6	6	1			7										
HIVEP2	6	2			2	4										
HIVEP3	6			3		3										
IDO2	6	5				5										
IFIH1	6	3		3		6	X		X	X	X		X			
IGLL1	6			3	2	5				X						
ITSN2	6	2			2	4	X									
KMT2D	31	6				6										
LRBA	6	1	1	2		4										
LRRK2	31	8	1			9										
LYST	6	4	1		1	6		X	X							
MASP2	31	7				7		X	X	X	X	X				X
MAVS	6			2	2	4	X		X	X		X		X		
MDC1	31			3	2	5										
MFN2	31	3				3										
MLH1	31	3				3		X		X	X		X			
MMP9	31	2		2		4	X									
MOGS	31	5		3		8										
MSH5	31	9				9										
MX1	31	8				8	X		X	X	X					
NBAS	31	2		3		5										
NFKB1	6	3				3	X		X	X		X		X		
NLRX1	31	4				4	X		X	X						
NOD1	31	2			1	3	X		X	X		X				
NOD2	6	6			1	7	X	X		X	X	X				
OAS1	31			3		3	X		X	X	X					
OAS2	31			3		3	X		X	X	X					
PIK3CD	6	4	1			5	X	X	X	X	X	X	X		X	
PKHD1	6	11				11										
PLAU	31	2		1		3	X									
PLCG2	6	5	2	2		9	X		X			X		X		
PLD1	31	3	1	1		5										
PLEKHM1	31				3	3										
PLXNB1	31	3				3										
POLE	6	7	2			9										
PRKDC	6	4				4			X	X			X	X	X	
PRRC2A	31	22	2		2	26										
RAG1	25A, 25B	4				4		X					X		X	
RANBP2	31	4				4										
RBFOX1	6	2			1	3										
RNF125	31	3			1	4	X		X	X				X		
RNF168	6	5				5		X		X	X		X			
RTEL1	6	2			1	3										
SAMD9	31	3			1	4	X									
SHARPIN	6	2	1	2	1	6										
SKIV2L	6	3				3										
STXBP2	6	3			1	4				X	X					
TEK	31	3				3	X									
TICAM1	6	2			1	3	X		X		X	X		X		
TLR4	6	9				9	X	X	X	X	X	X	X	X		
TNFRSF11A	6	10				10		X		X						
TRPM2	6	5				5										
WRAP53	31			3		3										
ZAP70	6	3				3	X	X	X	X		X	X			

[0409] Table 42 contains the gene scoring results, genes found in the 24-gene network, plus the top 5 "Biological Process" gene ontology (GO) pathways (based on the "false discovery rate" values reported in the String db output) and representative additional GO pathways of interest. Column headings and GO results are reported below:

RefSeq Gene Symbol, reports genes from Tables 6, 25A, 25B, and 31 that had a "Gene Score Total" of >= 3.

Gene Source, reports the original source table(s) of immune dysregulation genes.

Case-Solving, reports the total number unique PML cases that are potentially solved using the case-solving approach (described herein for Tables 7-9 and 34-36).

Gene Burden, reports the number of ethnicities (AFR, EUR, and/or LAT) that were found with the gene burden method (described herein for **Tables 13 and 37**).

Variant Burden Tier 1, reports the number of ethnicities (AFR, EUR, and/or LAT) that were found with the variant burden method (described herein for **Tables 14 and 38**).

Variant Burden Tier 2, reports the number of ethnicities (AFR, EUR, and/or LAT) that were found with the variant burden method (described herein for **Tables 15 and 39**).

Gene Score Total, reports the sum of Case-Solving, Gene Burden, Variant Burden Tier 1, and Variant Burden Tier 2 entries (described herein), wherein only genes with a score of >=3 were included.

24-Gene Network, identifies the genes (marked with an X) found in the largest network from the String db analysis (described herein), see Figure 14 for a graphical depiction.

[0410] The next 9 column headings list the "pathway ID" GO identifier numbers and identify the genes (marked with an X) that were found in each pathway from the total set of 74 genes in **Table 42.** Also listed below is the GO "pathway description", the "count in gene set" (number of genes from **Table 42**), and the "false discovery rate" (association value):

GO:0002250, adaptive immune response, 13 genes, 1.12e-10

GO:0045087, innate immune response, 23 genes, 1.35e-10

GO:0006955, immune response, 26 genes, 1.54e-10

GO:0002252, immune effector process, 16 genes, 2.70e-09

GO:0002253, activation of immune response, 16 genes, 2.70e-09

GO:0042113, B cell activation, 9 genes, 5.24e-07

GO:0032479, regulation of type I interferon production, 8 genes, 3.70e-06

GO:0030217, T cell differentiation, 7 genes, 6.65e-05

GO:0006958, complement activation, classical pathway, 4 genes, 6.10e-04

Example 24 - Analysis of Deleterious/Protective Variants

[0411] The WES data on the 70 PML cases were also analyzed for variants that occur at a statistically significant rate among the 705 total (435 from the original analysis and 270 from the second analysis) genes that have herein been identified as playing a role in immune dysregulation. In this analysis, "statistically significant" can be having an FET P-Value, after Bonferroni correction, of $P < 4.95E^{-6}$. Inclusion criteria included any SNV showing statistical significance at in one or more ethnic group (European, African or Latino) or across all groups. Some of these variants may occur with high frequency among the 70 PML cases. Some of these variants may occur in patients with HIV, MS, or other diseases as described herein, or any combination thereof. Observation of these variants in the 70 PML cases suggests that there may be a link between these variants and PML. **Table 43** lists statistically significant deleterious variants (such as variant 18-60052034-A-C, SEQ ID No. 1287), observed in the 70 PML cases with WES data. **Table 44** lists statistically significant protective variants observed in the 70 PML cases with WES data. Column headers: Variant; Gene; Total PML; Ethnicities; Diseases; FET P-Values: EUR; FET P-Values: AFR; FET P-Values: LAT; FET P-Values: ALL; Odds Ratios: EUR; Odds Ratios: AFR; Odds Ratios: LAT; Odds Ratios: ALL.

Table 43: Summary of statistically significant deleterious variants that have been observed in the 70 PML cases with WES data

Variant	Gene	PML Total	Ethnicities	Diseases	FET P-Values				Odds Ratios			
					EUR	AFR	LAT	All	EUR	AFR	LAT	All
1-154600405-T-C	ADAR	69	3	3	1.2E-28	2.3E-34	7.0E-09	1.4E-53	165.6	Inf	Inf	359.5
1-154600394-T-C	ADAR	4	2	1	2.4E-02	1.2E-07	1.0E+00	1.4E-09	45.1	Inf	0.0	352.0
1-154600411-A-C	ADAR	65	3	3	1.7E-29	1.0E-35	9.7E-08	4.7E-53	72.3	1510.3	344.0	93.3
1-33476387-C-T	AK2	70	3	3	1.6E-16	3.5E-07	8.4E-03	7.4E-26	Inf	Inf	Inf	Inf
1-33476385-C-G	AK2	70	3	3	6.4E-16	9.5E-07	1.1E-02	7.3E-25	Inf	Inf	Inf	Inf
1-33476353-C-T	AK2	70	3	3	6.7E-07	2. 1E-03	1.7E-01	8.0E-11	Inf	Inf	Inf	Inf
1-33476435-C-A	AK2	40	3	2	2.7E-82	2.0E-25	3.6E-10	6.9E-118	Inf	Inf	Inf	Inf
1-33475721-G-C	AK2	67	3	3	2.1E-11	1.5E-06	5.2E-02	7.0E-18	22.9	29.8	Inf	25.0
1-33476404-T-G	AK2	64	3	3	2.9E-14	1.6E-07	4.0E-02	2.4E-22	18.1	29.9	8.5	19.2
1-33476396-G-A	AK2	64	3	3	7.9E-14	2.5E-07	4.3E-02	1.2E-21	17.4	28.8	8.2	18.5
1-33475687-C-G	AK2	49	3	3	9.8E-15	2.2E-04	2.0E-02	6.9E-18	12.7	6.5	12.3	9.0
1-33476143-G-A	AK2	24	3	3	9.9E-08	6.9E-02	6.0E-02	2.3E-08	6.6	3.5	10.7	5.2
19-2129067-A-G	AP3D1	21	3	3	2.0E-25	8.2E-13	3.5E-02	2.7E-39	Inf	Inf	Inf	Inf
22-39357634-A-G	APOBEC3A	16	3	2	9.2E-09	4.5E-06	2.7E-02	5.9E-18	31.9	12.0	46.4	31.9
22-39358241-C-T	APOBEC3A	17	3	2	1.0E-07	1.2E-02	8.2E-06	7.1E-12	10.2	7.0	162.2	11.3
11-108114883-T-C	ATM	9	2	1	1.0E+00	2.4E-01	4.0E-06	1.8E-05	0.0	1.9	208.8	6.8
6-31997129-G-C	C4B	26	2	3	3.1E-08	4.1E-04	6.0E-01	4.4E-08	6.2	7.8	0.0	4.4
11-60891305-A-C	CDS	31	3	3	3.7E-51	3.0E-20	1.1E-95	2.1E-76	4176.5	Inf	1278.4	4457.8
1-160519815-C-CT	CD84	31	3	3	4.8E-06	1.6E-02	1.3E-03	1.2E-10	4.6	3.0	27.4	5.3
2-174229756-A-G	CDCA7	3	1	2	3.6E-06	1.6E-01	1.0E+00	8.5E-02	124.7	0.0	0.0	3.0
14-23588326-	CEBPE	17	3	2	1.4E-25	3.8E-09	5.2E-03	1.8E-37	Inf	Inf	Inf	Inf

Variant	Gene	PML Total	Ethnicities	Diseases	FET P-Values				Odds Ratios			
					EUR	AFR	LAT	All	EUR	AFR	LAT	All
G-C												
14-23588316-T-C	CEBPE	9	2	2	2.9E-15	1.5E-05	1.0E+00	7.1E-20	Inf	199.6	0.0	1010.1
1-196642969-CT-C	CFH	2	1	1	1.0E+00	1.0E+00	3.8E-06	3.3E-02	0.0	0.0	1726.0	7.2
1-196797357-A-G	CFHR1	37	3	3	3.4E-10	3.0E-02	1.0E+00	7.2E-12	7.2	2.6	1.2	5.5
16-88710039-T-C	CYBA	5	2	1	3.2E-02	4.0E-04	1.0E+00	1.5E-07	Inf	Inf	0.0	Inf
1-161518336-C-T	FCGR3A	30	3	3	1.6E-06	2.8E-03	6.5E-01	2.2E-06	4.7	4.1	0.4	3.4
4-89400457-C-A	HERC5	6	2	1	3.4E-08	4.1E-05	1.0E+00	6.0E-12	167.2	Inf	0.0	195.2
4-89400460-T-A	HERC5	6	2	1	1.8E-06	3.3E-04	1.0E+00	6.8E-09	55.9	197.4	0.0	52.3
1-42049542-G-A	HIVEP3	6	2	1	1.0E+00	1.0E+00	3.0E-06	2.3E-03	0.0	0.9	228.1	4.8
1-42050364-G-T	HIVEP3	6	2	1	1.0E+00	1.0E+00	2.9E-06	2.5E-03	0.0	0.9	234.9	4.7
21-45650009-T-TG	ICOSLG	19	3	3	2.8E-17	3.8E-14	1.4E-03	1.2E-34	Inf	Inf	Inf	Inf
21-45649595-T-C	ICOSLG	9	2	2	2.6E-06	2.5E-03	1.0E+00	1.1E-08	23.1	Inf	0.0	24.2
21-45649580-A-G	ICOSLG	9	2	2	2.2E-06	1.9E-02	1.0E+00	2.7E-07	16.7	Inf	0.0	14.6
10-1061646-T-C	Π7Ι2	39	3	3	2.5E-67	1.1E-21	3.1E-03	4.0E-93	Inf	Inf	Inf	Inf
10-1061650-C-G	IDE	40	3	3	6.5E-67	2.6E-20	8.1E-06	1.0E-94	Inf	Inf	Inf	Inf
8-42128955-A-G	IKBKB	5	2	3	2.5E-06	3.6E-01	1.0E+00	3.1E-06	Inf	4.2	0.0	87.4
5-78610478-A-C	JMY	10	3	1	8.4E-14	2.2E-07	2.4E-03	1.5E-22	468.4	Inf	910.6	581.7
5-78610472-T-C	JMY	12	3	3	5.4E-11	2.2E-09	1.6E-06	6.5E-23	127.6	Inf	Inf	225.3
19-48613788-T-G	LIG1	35	3	3	1.5E-31	1.9E-21	6.6E-05	8.3E-59	Inf	261.6	169.8	590.8
21-42775180-C-CT	MX2	9	2	2	3.7E-18	5.9E-08	1.0E+00	3.7E-26	Inf	Inf	0.0	Inf
11-119050352-T-C	NLRX1	3	2	2	1.8E-06	4.2E-03	1.0E+00	4.2E-09	Inf	Inf	0	Inf
11-	NLRX1	4	2	1	2.0E-06	1.5E-01	1.0E+00	6.0E-	154.9	0.2	0.0	2.8

Variant	Gene	PML Total	Ethnicities	Diseases	FET P-Values				Odds Ratios			
					EUR	AFR	LAT	All	EUR	AFR	LAT	All
119044844-C-T								02				
11-119044158-C-T	NLRX1	4	2	1	1.6E-06	2.4E-01	1.0E+00	6.5E-02	166.3	0.2	0.0	2.8
11-119045300-T-G	NLRX1	5	2	1	4.2E-06	2.0E-01	1.0E+00	4.8E-02	115.8	0.3	0.0	2.6
11-119045431-T-C	NLRX1	4	2	1	1.4E-06	1.5E-01	1.0E+00	7.8E-02	173.4	0.2	0.0	2.6
6-51503623-T-A	PKHD1	21	3	3	2.9E-03	5.7E-08	1.6E-01	5.2E-06	2.8	28.4	7.2	3.7
17-43552812-A-G	PLEKHM1	21	3	3	2.5E-09	2.5E-06	9.1E-02	5.3E-14	10.4	20.5	13.0	11.3
22-37622880-G-GT	RAC2	22	3	3	5.0E-24	6.6E-01	4.4E-05	4.3E-12	283.4	0.7	90.6	8.4
8-145154824-A-C	SHARPIN	3	1	1	2.7E-08	1.0E+00	1.0E+00	4.9E-08	1082.8	0.0	0.0	892.3
11-62655878-C-T	SLC3A2	7	2	1	1.0E+00	5.4E-01	1.0E-06	3.4E-04	0.0	1.3	336.4	5.8
19-50394219-G-C	TBC1D17	20	3	3	1.7E-14	1.3E-05	4.1E-02	1.2E-20	36.4	19.2	41.4	33.0
18-60052034-A-C	TNFRSF11A	9	2	1	1.4E-20	2.4E-05	1.0E+00	1.7E-25	Inf	Inf	0.0	Inf

Table 44: Summary of statistically significant protective variants that have been observed in the 70 PML cases with WES data

Variant	Gene	PML Total	Ethnicities	Diseases	FET P-Values				Odds Ratios			
					EUR	AFR	LAT	All	EUR	AFR	LAT	All
1-154562624-CG-C	ADAR	7	2	3	1.2E-37	3.8E-13	5.9E-08	2.6E-60	0.0	0.0	0.0	0.0
1-154562625-G-C	ADAR	4	1	1	2.5E-120	1.3E-34	3.7E-14	3.1E-168	0.0	0.0	0.0	0.0
1-33476223-TAC-T	AK2	58	3	3	1.0E-04	3.7E-04	2.1E-01	1.8E-08	0.1	0.1	0.2	0.1
1-33475967-G-A	AK2	41	3	3	9.6E-20	3.1E-05	2.3E-02	2.2E-28	0.0	0.1	0.0	0.0
1-33478931-G-C	AK2	1	1	1	8.9E-24	1.0E-08	1.5E-03	5.9E-35	0.0	0.0	0.0	0.0
1-33475982-C-A	AK2	4	2	1	8.6E-46	3.8E-10	8.9E-06	1.1E-63	0.0	0.0	0.0	0.0
1-33478959-C-G	AK2	1	1	1	8.6E-37	3.8E-13	1.2E-05	1.9E-55	0.0	0.0	0.0	0.0
9-100756891-C-CT	ANP32B	12	3	3	6.0E-18	2.0E-07	1.7E-01	7.7E-23	0.1	0.1	0.2	0.1
19-2129473-T-C	AP3D1	3	2	1	3.1E-18	4.0E-03	3.8E-01	3.9E-23	0.0	0.0	0.3	0.0
19-2129474-C-G	AP3D1	3	1	2	8.2E-17	4.0E-03	6.3E-02	3.6E-23	0.0	0.0	0.0	0.0
22-39387655-G-T	APOBEC3B	20	3	3	1.2E-10	6.5E-03	2.1E-02	5.4E-16	0.1	0.3	0.1	0.1
9-119491277-C-T	ASTN2	2	2	1	1.0E-09	1.0E+00	1.0E+00	1.3E-09	0.0	0.7	0.0	0.1
5-115167595-CT-C	ATG12	17	2	3	9.2E-15	3.8E-11	1.7E-06	5.6E-22	0.1	0.0	0.0	0.1
3-142231081-G-A	ATR	8	3	2	9.9E-20	5.5E-15	5.9E-07	7.6E-34	0.0	0.0	0.0	0.0
6-31994974-G-A	C4B	8	2	2	8.3E-02	1.1E-06	3.4E-01	3.5E-04	0.5	0.1	0.0	0.3
6-31994782-C-T	C4B	7	3	3	4.5E-06	5.0E-01	5.3E-01	1.9E-06	0.2	0.4	1.5	0.2
6-31997401-G-A	C4B	8	2	3	1.2E-06	7.1E-01	4.3E-03	2.0E-11	0.2	0.5	0.0	0.1
6-31997321-C-T	C4B	19	3	3	8.4E-21	1.2E-03	5.9E-02	1.9E-25	0.0	0.2	0.1	0.1
6-31994750-T-C	C4B	3	2	1	2.3E-05	6.0E-07	5.2E-03	3.4E-11	0.1	0.1	0.0	0.1
6-31994742-A-G	C4B	3	2	1	2.3E-05	6.8E-07	4.7E-03	3.1E-11	0.1	0.1	0.0	0.1
6-31997600-C-G	C4B	27	3	3	4.8E-16	2.6E-06	2.7E-05	9.4E-27	0.1	0.1	0.0	0.1
6-31994723-G-A	C4B	3	1	2	5.3E-09	1.6E-01	1.4E-03	7.1E-15	0.1	0.0	0.0	0.1
9-123762321-G-A	C5	12	2	3	5.7E-11	7.1E-02	7.9E-03	6.2E-17	0.1	0.2	0.0	0.1
2-87012399-TA-T	CD8A	1	1	1	1.6E-11	1.4E-04	1.4E-01	4.9E-18	0.0	0.0	0.0	0.0
2-174230974-G-T	CDCA7	15	3	2	8.5E-37	3.7E-01	2.9E-05	9.0E-34	0.0	0.6	0.0	0.0
20-48808011-C-T	CEBPB	41	3	3	2.4E-07	2.4E-01	5.1E-02	2.6E-10	0.2	0.6	0.1	0.2
22-42343091-G-A	CENPM	36	3	3	7.6E-01	1.9E-06	5.6E-01	1.0E+00	1.1	0.1	0.7	1.0
16-1493491-T-G	CLCN7	33	3	3	1.3E-07	1.0E+00	3.8E-03	4.2E-06	0.2	1.1	0.0	0.3
13-40252190-C-T	COG6	25	3	3	4.7E-06	3.5E-01	1.6E-01	1.8E-07	0.2	0.6	0.3	0.3
13-40326284-CA-C	COG6	22	3	3	7.9E-11	7.7E-05	3.0E-02	1.8E-18	0.1	0.1	0.1	0.1

9-399275-A-AC	DOCK8	20	3	3	1.0E+00	2.3E-06	3.8E-01	2.1E-02	1.0	0.1	0.3	0.5
12-93247775-G-GAT	EEA1	47	3	3	5.8E-02	4.5E-01	4.2E-06	1.6E-04	0.5	0.7	0.0	0.4
5-137847084-CACACACAT-C	ETF1	3	3	1	2.8E-09	8.0E-05	6.5E-01	1.7E-12	0.0	0.1	0.4	0.1
5-137847090-CAT-C	ETF1	2	2	1	5.3E-12	3.8E-03	1.8E-01	4.3E-16	0.0	0.1	0.0	0.0
12-11903753-C-T	ETV6	2	2	1	1.0E+00	4.2E-12	4.3E-02	2.0E-03	0.0	0.0	0.1	0.2
12-11903752-G-T	ETV6	2	2	1	1.0E+00	2.6E-15	3.0E-02	5.7E-04	0.0	0.0	0.1	0.1
12-11903751-TGC-T	ETV6	1	1	1	8.9E-14	7.6E-15	4.9E-04	1.0E-26	0.0	0.0	0.0	0.0
8-39842443-C-CTTAT	IDO2	12	2	2	3.2E-10	3.3E-02	9.6E-03	1.4E-14	0.1	0.3	0.0	0.1
22-17590180-G-A	IL17RA	19	3	2	1.3E-11	6.9E-04	4.0E-01	2.2E-11	0.1	0.2	3.0	0.2
1-82402364-TAATC-	LPHN2	10	3	3	5.3E-10	2.4E-02	1.0E+00	1.7E-11	0.1	0.3	0.6	0.1
6-74161762-G-T	MB21D1	3	2	2	7.3E-07	1.4E-01	5.9E-01	8.8E-07	0.0	3.6	0.0	0.1
21-42775180-C-T	MX2	30	3	3	4.2E-38	7.8E-18	1.1E-10	1.5E-62	0.0	0.0	0.0	0.0
1-59150941-G-A	MYSM1	54	3	3	9.6E-13	1.5E-05	1.0E+00	2.5E-16	0.0	0.0	Inf	0.0
7-74197598-T-A	NCF1	5	2	2	3.0E-26	2.1E-01	4.8E-04	2.3E-27	0.0	0.4	0.0	0.0
22-37268257-C-T	NCF4	52	3	3	5.7E-07	1.0E+00	1.5E-01	5.2E-07	0.1	0.9	0.1	0.2
11-119050352-T-TGAACAGGCA CATGGAAGGCC	NLRX1	2	2	1	2.5E-02	6.6E-07	1.0E+00	2.4E-04	0.2	0.0	0.0	0.1
12-113376543-C-CCAAAGGG	OAS3	29	3	3	2.2E-09	1.7E-18	5.3E-03	2.7E-17	0.2	0.0	0.1	0.1
1-9714543-A-ACCCC	PIK3CD	18	3	3	2.3E-09	1.2E-09	5.7E-03	5.2E-12	0.1	0.0	0.1	0.2
1-9714544-A-AG	PIK3CD	18	3	3	1.7E-09	1.8E-12	4.8E-03	2.7E-12	0.1	0.0	0.1	0.2
1-9714541-C-CCA	PIK3CD	18	3	3	1.7E-09	3.6E-12	4.5E-03	2.3E-12	0.1	0.0	0.1	0.2
12-133195625-G-A	POLE	56	3	3	6.2E-12	2.7E-03	1.7E-02	1.2E-15	0.0	0.1	0.0	0.0
3-53223286-GCTGGT-G	PRKCD	12	2	2	4.6E-21	1.2E-08	1.7E-01	1.1E-27	0.0	0.1	0.0	0.1
3-53223291-T-G	PRKCD	57	3	3	1.0E-20	1.7E-03	1.0E+00	6.3E-18	0.0	0.1	Inf	0.0
2-128181234-GC-G	PROC	10	3	2	1.1E-09	1.7E-02	1.0E+00	1.7E-11	0.0	0.0	0.0	0.0
12-56380689-T-C	RAB5B	6	3	2	8.7E-18	1.5E-06	8.7E-03	2.5E-24	0.0	0.1	0.1	0.0
11-62638333-T-A	SLC3A2	12	3	3	3.1E-06	1.5E-01	1.0E+00	3.9E-09	0.2	0.2	1.0	0.2
11-62638330-T-A	SLC3A2	15	3	3	1.1E-09	3.3E-02	1.0E+00	7.2E-14	0.1	0.1	0.7	0.1
11-62638200-T-A	SLC3A2	3	2	1	6.9E-08	4.9E-03	1.7E-01	3.0E-11	0.0	0.0	0.0	0.0
5-147504315-A-T	SPINK5	41	3	3	1.9E-18	5.1E-07	4.3E-02	1.0E-24	0.0	0.0	0.0	0.0
4-26862782-C-G	STIM2	19	3	3	9.1E-10	6.1E-03	3.4E-03	4.5E-17	0.1	0.2	0.0	0.1
4-26862754-C-T	STIM2	18	3	3	1.2E-18	2.8E-07	2.8E-04	1.9E-32	0.1	0.1	0.0	0.0
19-7705236-C-CTG	STXBP2	34	3	3	5.4E-01	7.7E-08	3.6E-01	8.1E-02	0.8	0.1	3.6	0.6
19-50394205-T-C	TBC1D17	1	1	1	2.1E-08	4.4E-05	1.0E+00	3.9E-13	0.0	0.0	0.0	0.0
22-19754091-A-C	TBX1	16	2	2	1.7E-07	7.6E-01	6.5E-05	7.9E-13	0.2	1.2	0.0	0.2
2-47277208-A-G	TTC7A	2	2	1	1.4E-17	1.3E-23	2.7E-04	8.6E-32	0.0	0.0	0.0	0.0
9-132640726-T-G	USP20	24	3	3	5.0E-06	1.0E+00	6.8E-01	2.2E-07	0.2	1.0	0.6	0.3
9-132632162-G-GC	USP20	53	3	3	3.3E-12	3.0E-02	1.0E-04	1.3E-15	0.0	0.3	0.0	0.1
11-9595768-C-T	WEE1	13	3	3	1.7E-07	9.0E-05	5.8E-01	5.9E-11	0.1	0.1	0.5	0.1

Example 25 - Description of Further Sequence Data

[0412] The sequence file 56969-701.601 ST25.txt contains genomic information for:

1. 1. The genetic sequence information referenced in Example 4 (SEQ IDs 1-2177);
2. 2. SEQ ID 2200-2203 are the distinct CNV sequences for the CNVs in **Table 28A**;
3. 3. SEQ ID 2204-2215 are the full genomic extent of the transcript sequences for the transcripts in **Table 30**;
4. 4. SEQ ID 2300-2893 are the full genomic extent of the transcript sequences for the transcripts in **Table 32**;
5. 5. SEQ ID 3000-3274 are the sequence variants listed in **Table 33**;
6. 6. SEQ ID 3275-3281 are the full genomic extent of the transcript sequences for the transcripts in **Table 49**;
7. 7. SEQ ID 3300-3351 are the sequence variants listed in **Table 45A**;
8. 8. SEQ ID 3400-3467 are the sequence variants listed in **Table 45B**;
9. 9. SEQ ID 3500-3526 are the sequence variants listed in **Table 45C.**

Table 45A: SEQ ID 3300-3351, SNV list (**Table 43**) with SEQ ID numbers

Chromosome	Position (hg19)	Ref Allele	Alt Allele	SEQ ID
1	33475687	C	G	3310
1	33475721	G	C	3307
1	33476143	G	A	3311
1	33476353	C	T	3305
1	33476385	C	G	3304
1	33476387	C	T	3303
1	33476396	G	A	3309
1	33476404	T	G	3308
1	33476435	C	A	3306
1	42049542	G	A	3328
1	42050364	G	T	3329
1	154600394	T	C	3301
1	154600405	T	C	3300

Chromosome	Position (hg19)	Ref Allele	Alt Allele	SEQ ID
1	154600411	A	C	3302
1	160519815	C	CT	3318
1	161518336	C	T	3325
1	196642969	CT	C	3322
1	196797357	A	G	3323
2	174229756	A	G	3319
4	89400457	C	A	3326
4	89400460	T	A	3327
5	78610472	T	C	3337
5	78610478	A	C	3336
6	31997129	G	C	3316
6	51503623	T	A	3345
8	42128955	A	G	3335
8	145154824	A	C	3348
10	1061646	T	C	3333
10	1061650	C	G	3334
11	60891305	A	C	3317
11	62655878	C	T	3349
11	108114883	T	C	3315
11	119044158	C	T	3342
11	119044844	C	T	3341
11	119045300	T	G	3343
11	119045431	T	C	3344
11	119050352	T	C	3340
14	23588316	T	C	3321
14	23588326	G	C	3320
16	88710039	T	C	3324
17	43552812	A	G	3346
18	60052034	A	C	3351
19	2129067	A	G	3312
19	48613788	T	G	3338
19	50394219	G	C	3350
21	42775180	C	CT	3339
21	45649580	A	G	3332
21	45649595	T	C	3331
21	45650009	T	TG	3330
22	37622880	G	GT	3347
22	39357634	A	G	3313
22	39358241	C	T	3314

Table 45B: SEQ ID 3400-3467, SNV list (**Table 44**) with SEQ ID numbers

Chromosome	Position (hg19)	Ref Allele	Alt Allele	SEQ ID
1	9714541	C	CCA	3449
1	9714543	A	ACCCC	3447
1	9714544	A	AG	3448
1	33475967	G	A	3403
1	33475982	C	A	3405
1	33476223	TAC	T	3402
1	33478931	G	C	3404
1	33478959	C	G	3406
1	59150941	G	A	3442
1	82402364	TAATC	T	3439

Chromosome	Position (hg19)	Ref Allele	Alt Allele	SEQ ID
1	154562624	CG	C	3400
1	154562625	G	C	3401
2	47277208	A	G	3464
2	87012399	TA	T	3423
2	128181234	GC	G	3453
2	174230974	G	T	3424
3	53223286	GCTGGT	G	3451
3	53223291	T	G	3452
3	142231081	G	A	3413
4	26862754	C	T	3460
4	26862782	C	G	3459
5	115167595	CT	C	3412
5	137847084	CACACACAT	C	3432
5	137847090	CAT	C	3433
5	147504315	A	T	3458
6	31994723	G	A	3421
6	31994742	A	G	3419
6	31994750	T	C	3418
6	31994782	C	T	3415
6	31994974	G	A	3414
6	31997321	C	T	3417
6	31997401	G	A	3416
6	31997600	C	G	3420
6	74161762	G	T	3440
7	74197598	T	A	3443
8	39842443	C	CTTAT	3437
9	399275	A	AC	3430
9	100756891	C	CT	3407
9	119491277	C	T	3411
9	123762321	G	A	3422
9	132632162	G	GC	3466
9	132640726	T	G	3465
11	9595768	C	T	3467
11	62638200	T	A	3457
11	62638330	T	A	3456
11	62638333	T	A	3455
11	119050352	T	TGAACAGGCACATG GAAGGCC	3445
12	11903751	TGC	T	3436
12	11903752	G	T	3435
12	11903753	C	T	3434
12	56380689	T	C	3454
12	93247775	G	GATA	3431
12	113376543	C	CCAAAGGG	3446
12	133195625	G	A	3450
13	40252190	C	T	3428
13	40326284	CA	C	3429
16	1493491	T	G	3427
19	2129473	T	C	3408
19	2129474	C	G	3409
19	7705236	C	CTG	3461
19	50394205	T	C	3462

Chromosome	Position (hg19)	Ref Allele	Alt Allele	SEQ ID
20	48808011	C	T	3425
21	42775180	C	T	3441
22	17590180	G	A	3438
22	19754091	A	C	3463
22	37268257	C	T	3444
22	39387655	G	T	3410
22	42343091	G	A	3426

Table 45C: SEQ ID 3500-3526, SNV list (**Table 50A**) with SEQ ID numbers

Chromosome	Position (hg19)	Ref Allele	Alt Allele	SEQ ID
1	27699670	AG	A	3500
1	42047208	C	G	3501
1	57409459	C	A	3502
1	92946625	G	C	3503
1	160769595	AG	A	3504
1	196918605	A	G	3505
2	163136505	C	G	3506
2	230579019	G	A	3507
3	58191230	G	T	3508
4	151793903	T	C	3509
6	3015818	G	A	3510
6	30673359	T	G	3511
6	32814942	C	T	3512
6	32816772	C	A	3513
6	51798908	C	T	3514
9	137779251	G	A	3515
11	67818269	G	A	3516
11	108106443	T	A	3517
14	94847262	T	A	3518
16	81942175	A	G	3519
19	7712287	G	C	3520
19	8564523	T	G	3521
19	48643270	C	T	3522
21	45708278	G	A	3523
22	23915583	T	C	3524
22	23915745	G	A	3525
22	35806756	G	A	3526

Table 48: GN 763-765, NCBI Gene ID, descriptions, RefSeq summary for 3 genes from all-genes variant burden analyses (**Table 50A**)

RefSeq Gene Symbol	NCBI Gene ID	Gene Description	RefSeq Summmary	Gene No. (GN)
FCN2	2220	ficolin-2 isoform a precursor	The product of this gene belongs to the ficolin family of proteins. This family is characterized by the presence of a leader peptide, a short N-terminal segment, followed by a collagen-like region, and a C-terminal fibrinogen-like domain. This gene is predominantly expressed in the liver, and has been shown to have carbohydrate binding and opsonic activities. Alternatively spliced transcript variants encoding different isoforms have been identified. [provided by RefSeq, Jul 2008].	763
LY9	4063	T-lymphocyte surface antigen Ly-9 isoform a precursor	LY9 belongs to the SLAM family of immunomodulatory receptors (see SLAMF1; MIM 603492) and interacts with the adaptor molecule SAP (SH2D1A; MIM 300490) (Graham et al., 2006 [PubMed 16365421]). [supplied by OMIM, Mar 2008].	764
PRAM 1	84106	PML-RARA-regulated adapter molecule 1	The protein encoded by this gene is similar to FYN binding protein (FYB/SLAP-130), an adaptor protein involved in T cell receptor mediated signaling. This gene is expressed and regulated during normal myelopoiesis. The expression of this gene is induced by	765

RefSeq Gene Symbol	NCBI Gene ID	Gene Description	RefSeq Summmary	Gene No. (GN)
			retinoic acid and is inhibited by the expression of PML-RARalpha, a fusion protein of promyelocytic leukemia (PML) and the retinoic acid receptor-alpha (RARalpha). [provided by RefSeq, Jul 2008].	

Table 49: SEQ ID 3275-3281, non-redundant list of transcript variants that correspond to the genes in **Table 48**

RefSeq Gene Symbol	RefSeq Accession Number	mRNA Description	SEQ ID
LY9	NM 001033667	Homo sapiens lymphocyte antigen 9 (LY9), transcript variant 2, mRNA.	3275
LY9	NM 001261456	Homo sapiens lymphocyte antigen 9 (LY9), transcript variant 3, mRNA.	3276
LY9	NM 001261457	Homo sapiens lymphocyte antigen 9 (LY9), transcript variant 4, mRNA.	3277
LY9	NM 002348	Homo sapiens lymphocyte antigen 9 (LY9), transcript variant 1, mRNA.	3278
FCN2	NM 004108	Homo sapiens ficolin (collagen/fibrinogen domain containing lectin) 2 (hucolin) (FCN2), transcript variant SV0, mRNA.	3279
FCN2	NM 015837	Homo sapiens ficolin (collagen/fibrinogen domain containing lectin) 2 (hucolin) (FCN2), transcript variant SV1, mRNA.	3280
PRAM 1	NM 032152	Homo sapiens PML-RARA regulated adaptor molecule 1 (PRAM1), mRNA.	3281

Example 26 - PML Discovery and Replication Cohorts

[0413] The cohort of 70 PML cases whose analysis was described previously (e.g., the genetic results reported in **Table 7**) can be considered a 'Discovery' (Dis) cohort. An additional study was performed on a 'Replication' (Rep) cohort comprising 115 additional PML cases. These 115 cases were obtained from 4 different sources, only one of which was the same as one of the sources for the Dis cohort. **Table 46** lists information for all 185 PML cases (Dis cohort = 70 cases, Rep cohort = 115 cases). As was done for the Dis PML cohort (Example 3), WES data was obtained on the 115 Rep PML cases. Unlike the Dis cohort, aCGH was not performed on the Rep cohort. However, those skilled in the art know that CNVs (e.g., those reported in **Tables 1 and 28**A) can be detected by other methods besides aCGH. For example, the recurrent CNVS reported in **Table 1** and/or **Table 28A** can, at some stage, be easily screened for in the Rep cohort using PCR-based methods (e.g., junction fragment PCR assays).

[0414] To confirm the ancestry of each PML case, which was reportd by the clinician, WGS data (Gencove, NY, USA) was generated on genomic DNA for all 185 cases. The Gencove platform uses low pass sequencing at a read depth of 0.1x that yields sufficient data (on the basis of SNPs) to ascertain ethnicity and "uniqueness" of a sample. This data enabled determination of:

1. A. Actual ethnicity, as opposed to relying on the clinician-reported ethnicity;
2. B. Whether any of the 185 PML samples were duplicates. Because PML is a rare disorder, it is possible that affected individuals may have been seen multiple times by different medical centers. None of the 185 samples were found to be duplicates.

Gencove reports the following categories of ancestries/ethnicities (full name is following by abbreviation used in parentheses): Americas (AMERICAS), Ashkenazi Jewish (ASHKENAZI), Bengal (BENGALI), Central Africa (CAFRICA), Central Asia (CASIA), Eastern Africa (EAFRICA), East Asia (EASIA), Eastern Mediterranean (EMED), Finland (FINLAND), Central Indian subcontinent (INDPAK), Northern Africa (NAFRICA), North-central Asia (NCASIA), Middle East (NEAREAST), Northeast Asia (NEASIA), Northeast Europe (NEEUROPE), Northern and Central Europe (NEUROPE), Northern Italy (NITALY), Northern British Isles (NNEUROPE), Oceania (OCEANIA), Southern Africa (SAFRICA), Scandinavia (SCANDINAVIA), Southeast Asia (SEASIA), Southern Indian subcontinent (SSASIA), Southwestern Europe (SWEUROPE), Anatolia, Caucasus, Iranian Plateau (TURK-IRAN-CAUCASUS), and Western Africa (WAFRICA). To simplify interpretation of the PML patient variants using public databases, such as gnomAD and the 1000genomes project (TGP), Gencove ethnicities were combined to determine the highest percentage ethnicity for either European ancestry (corresponding to NFE in gnomAD and EUR in TGP) or African ancestry (corresponding to AFR in gnomAD and in TGP). The following Gencove ethnicities were combined: NFE/EUR = (EMED + NEEUROPE + NEUROPE + NITALY + NNEUROPE + SCANDINAVIA + SWEUROPE) and AFR = (CAFRICA + EAFRICA + NAFRICA + SAFRICA + WAFRICA). The majority of PML patients could be readily assigned as NFE/EUR or AFR ethnicity, but a few individuals were binned according to their closest ethnicity (e.g., patients with predominantly ASHKENAZI and TURK-IRAN-CAUCASUS ancestry were assigned as NFE/EUR).

Table 46: Complete list of 185 PML cases in the Discovery (n = 70) and Replication (n = 115) cohorts

Cohort	Sample ID	Gender	aCGH Expt ID	Primary Disease	Primary Ethnicity
Dis	MVGS1116-8a	F	3006	MS	NFE
Dis	MVGS1359	F	3117	MS	NFE
Dis	MVGS1368	F	3118	MS	NFE
Dis	MVGS540-374b	M	3005	MS	NFE
Dis	MVGS540-393b	F	3004	MS	NFE
Dis	MVGS694-6a	F	3007	Other	NFE
Dis	MVGS811-13a	M	3009	HIV	AFR
Dis	MVGS995-4a	M	3010	MS	NFE
Dis	PML1	F	3127	HIV	NFE
Dis	PML10	F	3157	HIV	NFE
Dis	PML12	F	3159	HIV	NFE
Dis	PML13	M	3160	HIV	NFE
Dis	PML14	M	3161	HIV	NFE
Dis	PML15	M	3194	HIV	AFR
Dis	PML16	F	3163	HIV	AFR
Dis	PML17	M	3140	HIV	NFE
Dis	PML18	M	3141	HIV	NFE
Dis	PML19	M	3164	HIV	AFR
Dis	PML2	M	3126	Blood Cancers	NFE
Dis	PML20	M	3143	HIV	AFR
Dis	PML21	M	3144	HIV	NFE
Dis	PML22	M	3145	HIV	NFE
Dis	PML23	F	3165	HIV	NFE
Dis	PML25	F	3166	HIV	NFE
Dis	PML26	M	3167	HIV	NFE
Dis	PML27	M	3168	HIV	NFE
Dis	PML28	F	3151	MS	NFE
Dis	PML29	M	3152	HIV	AFR
Dis	PML3	F	3155	MS	NFE
Dis	PML30	M	3153	HIV	NFE
Dis	PML31	F	3154	HIV	AFR
Dis	PML32	M	3169	HIV	NFE
Dis	PML33	M	3170	HIV	NFE
Dis	PML35	F	3171	HIV	NFE
Dis	PML36	F	3172	HIV	NFE
Dis	PML37	M	3173	HIV	AFR
Dis	PML38	M	3174	HIV	NFE
Dis	PML39	M	3175	HIV	AFR
Dis	PML4	M	3156	HIV	NFE
Dis	PML40	F	3273	HIV	AFR
Dis	PML41	M	3177	HIV	NFE
Dis	PML43	M	3178	HIV	NFE
Dis	PML44	M	3179	HIV	NFE
Dis	PML45	F	3180	Blood Cancers	NFE
Dis	PML46	M	3196	HIV	NFE
Dis	PML48	M	3197	Other	NFE
Dis	PML49	M	3183	HIV	NFE
Dis	PML5	M	3125	HIV	AFR
Dis	PML50	M	3198	HIV	AFR
Dis	PML51	M	3185	HIV	NFE
Dis	PML52	F	3186	Blood Cancers	NFE

Cohort	Sample ID	Gender	aCGH Expt ID	Primary Disease	Primary Ethnicity
Dis	PML53	M	3187	Other	NFE
Dis	PML54	F	3188	HIV	NFE
Dis	PML55	F	3189	HIV	NFE
Dis	PML56	M	3190	HIV	NFE
Dis	PML57	F	3191	Blood Cancers	NFE
Dis	PML58	M	3192	HIV	AFR
Dis	PML59	M	3193	HIV	AFR
Dis	PML6	M	3124	HIV	NFE
Dis	PML60	M	3199	HIV	NFE
Dis	PML61	F	3200	HIV	AFR
Dis	PML62	F	3201	HIV	AFR
Dis	PML63	M	3202	HIV	AFR
Dis	PML64	M	3203	HIV	AFR
Dis	PML65	M	3204	HIV	AFR
Dis	PML66	M	3205	HIV	AFR
Dis	PML68	F	3278	MS	NFE
Dis	PML69	M	3279	Other	NFE
Dis	PML72	F	3282	HIV	AFR
Dis	PML9	M	3132	HIV	NFE
Rep	PBPML100	F	n/a	Blood Cancers	NFE
Rep	PBPML101	M	n/a	Blood Cancers	NFE
Rep	PBPML 102	F	n/a	Blood Cancers	NFE
Rep	PBPML 103	M	n/a	Other	NFE
Rep	PBPML 104	F	n/a	Blood Cancers	NFE
Rep	PBPML 105	M	n/a	Other	NFE
Rep	PBPML 106	M	n/a	Other	NFE
Rep	PBPML 107	M	n/a	Blood Cancers	NFE
Rep	PBPML108	F	n/a	Blood Cancers	NFE
Rep	PBPML109	M	n/a	Blood Cancers	NFE
Rep	PBPML110	M	n/a	HIV	AFR
Rep	PBPML111	M	n/a	HIV	AFR
Rep	PBPML112	M	n/a	Blood Cancers	NFE
Rep	PBPML113	M	n/a	HIV	NFE
Rep	PBPML114	M	n/a	Blood Cancers	NFE
Rep	PBPML115	F	n/a	Other	NFE
Rep	PBPML116	F	n/a	HIV	NFE
Rep	PBPML117	M	n/a	HIV	NFE
Rep	PBPML118	M	n/a	HIV	NFE
Rep	PBPML119	M	n/a	Other	NFE
Rep	PBPML120	M	n/a	HIV	NFE
Rep	PBPML121	M	n/a	HIV	NFE
Rep	PBPML122	M	n/a	HIV	NFE
Rep	PBPML123	F	n/a	MS	NFE
Rep	PBPML124	M	n/a	Other	NFE
Rep	PBPML125	M	n/a	MS	NFE
Rep	PBPML126	M	n/a	Other	NFE
Rep	PBPML127	F	n/a	Other	NFE
Rep	PBPML128	M	n/a	HIV	AFR
Rep	PBPML129	F	n/a	HIV	AFR
Rep	PBPML130	M	n/a	Blood Cancers	NFE

Cohort	Sample ID	Gender	aCGH Expt ID	Primary Disease	Primary Ethnicity
Rep	PBPML131	M	n/a	HIV	NFE
Rep	PBPML132	F	n/a	HIV	AFR
Rep	PBPML133	M	n/a	HIV	NFE
Rep	PBPML134	F	n/a	HIV	AFR
Rep	PBPML135	F	n/a	HIV	NFE
Rep	PBPML136	F	n/a	Other	NFE
Rep	PBPML137	M	n/a	HIV	NFE
Rep	PBPML138	M	n/a	HIV	NFE
Rep	PBPML139	F	n/a	MS	NFE
Rep	PBPML 140	M	n/a	HIV	NFE
Rep	PBPML141	M	n/a	HIV	AFR
Rep	PBPML142	M	n/a	HIV	NFE
Rep	PBPML143	M	n/a	HIV	NFE
Rep	PBPML144	F	n/a	HIV	AFR
Rep	PBPML145	M	n/a	HIV	NFE
Rep	PBPML146	M	n/a	Other	NFE
Rep	PBPML147	M	n/a	HIV	NFE
Rep	PBPML148	M	n/a	HIV	AFR
Rep	PBPML149	M	n/a	Blood Cancers	NFE
Rep	PBPML150	M	n/a	HIV	NFE
Rep	PBPML151	M	n/a	HIV	AFR
Rep	PBPML152	M	n/a	Other	NFE
Rep	PBPML153	M	n/a	HIV	NFE
Rep	PBPML154	F	n/a	MS	NFE
Rep	PBPML155	M	n/a	HIV	NFE
Rep	PBPML156	F	n/a	Blood Cancers	NFE
Rep	PBPML157	M	n/a	HIV	NFE
Rep	PBPML158	F	n/a	Other	NFE
Rep	PBPML159	M	n/a	HIV	NFE
Rep	PBPML160	F	n/a	Blood Cancers	NFE
Rep	PBPML161	F	n/a	HIV	AFR
Rep	PBPML162	M	n/a	HIV	NFE
Rep	PBPML163	F	n/a	Blood Cancers	NFE
Rep	PBPML164	M	n/a	HIV	NFE
Rep	PBPML165	F	n/a	HIV	AFR
Rep	PBPML166	M	n/a	HIV	NFE
Rep	PBPML167	M	n/a	HIV	NFE
Rep	PBPML168	F	n/a	MS	NFE
Rep	PBPML169	M	n/a	HIV	NFE
Rep	PBPML170	M	n/a	HIV	AFR
Rep	PBPML171	M	n/a	HIV	NFE
Rep	PBPML172	F	n/a	Blood Cancers	NFE
Rep	PBPML173	M	n/a	Other	NFE
Rep	PBPML174	M	n/a	Other	NFE
Rep	PBPML175	F	n/a	HIV	AFR
Rep	PBPML176	F	n/a	Blood Cancers	NFE
Rep	PBPML177	M	n/a	HIV	NFE
Rep	PBPML178	M	n/a	HIV	NFE
Rep	PBPML179	M	n/a	HIV	NFE
Rep	PBPML180	F	n/a	HIV	AFR
Rep	PBPML181	M	n/a	HIV	NFE

Cohort	Sample ID	Gender	aCGH Expt ID	Primary Disease	Primary Ethnicity
Rep	PBPML182	M	n/a	HIV	NFE
Rep	PBPML183	M	n/a	HIV	AFR
Rep	PBPML184	F	n/a	HIV	AFR
Rep	PBPML185	M	n/a	HIV	NFE
Rep	PBPML186	F	n/a	HIV	AFR
Rep	PBPML187	M	n/a	HIV	NFE
Rep	PBPML188	M	n/a	HIV	AFR
Rep	PBPML189	F	n/a	HIV	NFE
Rep	PBPML190	M	n/a	HIV	AFR
Rep	PBPML191	F	n/a	HIV	AFR
Rep	PBPML77	M	n/a	HIV	NFE
Rep	PBPML78	M	n/a	HIV	NFE
Rep	PBPML79	F	n/a	HIV	NFE
Rep	PBPML80	F	n/a	HIV	AFR
Rep	PBPML81	M	n/a	HIV	NFE
Rep	PBPML82	F	n/a	HIV	NFE
Rep	PBPML83	M	n/a	HIV	AFR
Rep	PBPML84	M	n/a	HIV	AFR
Rep	PBPML85	F	n/a	HIV	AFR
Rep	PBPML86	M	n/a	Other	AFR
Rep	PBPML87	F	n/a	HIV	NFE
Rep	PBPML88	M	n/a	Other	NFE
Rep	PBPML89	M	n/a	HIV	AFR
Rep	PBPML90	M	n/a	HIV	NFE
Rep	PBPML91	F	n/a	Blood Cancers	NFE
Rep	PBPML92	F	n/a	HIV	NFE
Rep	PBPML93	M	n/a	HIV	AFR
Rep	PBPML94	M	n/a	Blood Cancers	NFE
Rep	PBPML95	F	n/a	MS	NFE
Rep	PBPML96	F	n/a	HIV	NFE
Rep	PBPML97	M	n/a	Other	NFE
Rep	PBPML98	M	n/a	Other	NFE
Rep	PBPML99	F	n/a	Other	NFE

Table 46 represents the complete set of PML patients in the Discovery (Dis) or Replication (Rep) cohorts. The Dis cohort is equivalent to the patients listed in **Table 7** except for PML67, for which insufficient DNA was available to generate WES data. **Table 46** also lists the Sample ID, Gender, aCGH Expt ID (also reported in **Table 11**), Primary Disease, and Primary Ethnicity (based on Gencove ancestry analysis). Besides MS and HIV, the Primary Disease categories included Blood Cancers (mainly leukemias and lymphomas) and Other (miscellaneous conditions such as anemia, lymphopenia, sarcoidosis, and transplant patients).

Example 27 - Updated Sequence Analysis Pipeline

[0415] High-throughput sequencing and sequence analysis methods are continually being developed to improve the quality of the data and results. A new approach, complimenting the previous approach, was used for the analysis of the Dis and Rep cohort WES data. Google's Deep Variant (DV) caller (github.com/google/deepvariant) was used for variant calling (see Poplin et al. (2018) Nat Biotechnol. Nov;36(10):983-987). Using GRCh37 as the reference genome, BWA-aligned BAM files for each PML sample (Dis or Rep cohort) was run through the DV caller. Variant inclusion criteria were: DP ≥ 10 (site depths were calculated again directly from the BAM files, with constraints for Base Quality > 10 and Mapping Quality > 20, and DP < 10 were treated as "NA"), VAF > 0.2 for heterozygous calls (VAF < 0.2 were treated as "NA"), and VAF > 0.8 for homozygous calls (VAF < 0.8 were treated as heterozygous calls). Annotation of DV-called variants was performed using dbNSFP (see Liu et al. (2016) Hum Mutat. 2016 Mar;37(3):235-41).

[0416] Variant burden analyses were performed using similar criteria as was used for variants reported in Tables 14, 15, 38, and 39, except that DV-called variants for the PML cases was used as input. The Dis and Rep cohorts were analyzed separately, with separate

output files being generated for heterozygous (het) and homozygous (horn) variants. For simplicity, the genes listed in Tables 6, 25A, 25B, 26, and 31 were combined into a non-redundant list of 705 genes. The 705-gene list was compared to all genes that were found in the annotation reports for the two cohorts (Dis and Rep), which yielded a non-redundant gene list of 663 genes that were used for the DV pipeline variant burden analyses. In addition to the 663-gene analyses, a parallel set of all-genes (all ~20,000 human genome genes) analyses were generated. To summarize, for each variant observed, a total count was made of the number of individuals in the PML cohorts that possessed the variant, and this was similarly done for comparison in three different unselected population cohorts: gnomAD exome (WES data), gnomAD genome (WGS data), and TGP exome (WES data). Fishers Exact Test (FET) and Odds Ratio (OR) were calculated to identify statistically significant variants with the highest OR values.

Table 47: 663 genes analyzed for the 185 PML cases (DV pipeline)

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
ACADM	no	unknown	Table 6	157
ACD	no	ADAR	Table 31	493
ACKR1	no	ADAR	Table 6	158
ACP5	yes	AR	Table 6	159
ACTB	yes	AD	Table 31	494
ACTN4	no	unknown	Table 31	495
ADA	yes	AR	Table 25B; Table 26	1
ADA2	yes	AR	Table 31	496
ADAM17	yes	AR	Table 31	497
ADAR	yes	ADAR	Table 6	160
ADARB1	no	unknown	Table 6	2
ADGRL2	no	unknown	Table 31	498
ADK	no	unknown	Table 6	161
AGBL4	no	unknown	Table 6	3
AICDA	yes	ADAR	Table 6	162
AIRE	yes	ADAR	Table 31	499
AK2	yes	AR	Table 6	163
ALG12	no	AR	Table 6	164
ALPL	no	unknown	Table 6	165
ANP32B	no	unknown	Table 31	500
AP1S3	yes	AR	Table 31	501
AP3B1	yes	AR	Table 6	166
AP3B2	no	unknown	Table 6	167
AP3D1	yes	AR	Table 6	168
APOBEC3A	no	unknown	Table 6	4
APOBEC3B	no	unknown	Table 6	6
APOL1	yes	unknown	Table 6	169
ARHGEF7	no	unknown	Table 6	8
ARPC1B	yes	AR	Table 31	502
ASH1L	no	unknown	Table 6	170
ASTN2	no	unknown	Table 6	9
ATG12	no	unknown	Table 31	503
ATG16L1	no	unknown	Table 31	504
ATG5	no	unknown	Table 31	505
ATG7	no	unknown	Table 31	506
ATG9A	no	unknown	Table 31	507
ATL2	no	unknown	Table 6	171
ATM	yes	AR	Table 6	172
ATP6AP1	yes	XLR	Table 31	508
ATR	no	unknown	Table 6	173
AUH	no	unknown	Table 6	10
B2M	yes	AR	Table 31	509
BACH1	no	unknown	Table 6	11

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
BACH2	yes	unknown	Table 6	174
BAG3	no	AR	Table25A; Table 25B	175
BCL10	yes	AR	Table 6	176
BCL11B	yes	AD	Table 31	510
BCL2	no	unknown	Table 31	511
BDKRB2	no	unknown	Table 6	12
BLK	no	AD	Table 31	512
BLM	yes	AR	Table 6	177
BLNK	yes	AR	Table 6	178
BLOC1S6	no	AR	Table 6	179
BMPR2	no	unknown	Table 6	13
BRD4	no	unknown	Table 31	513
BTK	yes	XLR	Table25A; Table 25B	180
BTLA	no	unknown	Table 31	514
C11orf65	no	unknown	Table 6	181
C1QA	yes	AR	Table 6	182
CIQB	yes	AR	Table 6	183
C1QC	yes	AR	Table 6	184
C1R	yes	AR	Table 31	515
C1S	yes	AR	Table 31	516
C2	yes	AR	Table 31	517
C3	yes	AD AR	Table 31	518
C4A	yes	AR	Table 31	519
C4B	yes	AR	Table 31	520
C5	yes	AR	Table 31	521
C5AR1	no	unknown	Table 6	185
C6	yes	AR	Table 31	522
C7	yes	AR	Table 31	523
C8A	yes	AR	Table 31	524
C8B	yes	AR	Table 31	525
C8G	yes	AR	Table 31	526
C9	yes	AR	Table 31	527
CAD	no	AR	Table 31	528
CAMLG	no	unknown	Table 31	529
CAPZB	no	unknown	Table 6	186
CARD11	yes	AD AR	Table 6	187
CARD14	yes	AD	Table 31	530
CARD9	yes	AR	Table 6	188
CASP10	yes	AD	Table 31	531
CASP8	yes	AR	Table 6	189
CAV1	no	unknown	Table 31	532
CCBE1	yes	AR	Table 31	533
CCDC22	no	XLR	Table 31	534
CCL11	no	unknown	Table 6	190
CCL2	no	unknown	Table 6	191
CCL5	no	unknown	Table 6	192
CCR2	no	unknown	Table 6	193
CCR5	no	unknown	Table 6	194
CCZ1	no	unknown	Table 31	535
CD180	no	unknown	Table 6	195
CD19	yes	AR	Table 6	196

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
CD209	no	unknown	Table 6	197
CD22	no	unknown	Table 31	536
CD247	yes	AR	Table 6	198
CD27	yes	AR	Table 6	199
CD274	no	unknown	Table 31	537
CD276	no	unknown	Table 31	538
CD300LF	no	unknown	Table 6	23
CD34	no	unknown	Table 6	201
CD36	no	AR	Table 31	539
CD37	no	unknown	Table 31	540
CD38	no	unknown	Table 31	541
CD3D	yes	AR	Table 6	202
CD3E	yes	AR	Table 6	203
CD3G	yes	AR	Table 6	204
CD40	yes	AR	Table 6	205
CD40LG	yes	XLR	Table25A; Table 25B	206
CD46	yes	AD	Table 31	542
CD5	no	unknown	Table 31	543
CD55	yes	unknown	Table 6	207
CD59	yes	AR	Table 6	208
CD70	yes	AR	Table 31	544
CD72	no	unknown	Table 31	545
CD74	no	unknown	Table 31	546
CD79A	yes	AR	Table 6	209
CD79B	yes	AR	Table 6	210
CD81	yes	AR	Table 6	211
CD84	no	unknown	Table 31	547
CD8A	yes	AR	Table 6	212
CD93	no	unknown	Table 31	548
CDCA7	yes	AR	Table 6	213
CDKN1B	no	unknown	Table 6	24
CEBPB	no	unknown	Table 6	214
CEBPE	yes	AR	Table 31	549
CENPM	no	unknown	Table 6	25
CFB	yes	ADAR	Table 31	550
CFD	yes	AR	Table 31	551
CFH	yes	ADAR	Table 31	552
CFHR1	yes	ADAR	Table 31	553
CFHR2	yes	ADAR	Table 31	554
CFHR3	yes	AD AR	Table 31	555
CFHR4	yes	ADAR	Table 31	556
CFHR5	yes	ADAR	Table 31	557
CFI	yes	ADAR	Table 31	558
CFP	yes	XLR	Table 31	559
CFTR	yes	AR	Table 31	560
CHD2	no	unknown	Table 31	561
CHD7	yes	AD	Table 6	215
CHEK1	no	unknown	Table 6	216
CIITA	yes	AR	Table 6	217
CLCN7	yes	AD	Table 6	218

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
CLEC16A	no	unknown	Table 31	562
CLPB	yes	AR	Table 31	563
COG4	no	unknown	Table 6	26
COG6	no	AR	Table 6	219
COMMD6	no	unknown	Table 6	27
COPA	yes	AD	Table 31	564
CORO1A	yes	AR	Table 6	220
CR2	yes	AR	Table 6	221
CRADD	no	unknown	Table 6	28
CRTC3	no	unknown	Table 6	222
CSF2RA	yes	XLR	Table 31	565
CSF2RB	yes	AR	Table 31	566
CSF3R	yes	AR	Table 6	223
CTC1	yes	AR	Table 31	567
CTLA4	yes	AD	Table 6	224
CTPS1	yes	AR	Table 6	225
CTSC	yes	AR	Table 6	226
CX3CR1	no	unknown	Table 6	227
CXCL1	no	unknown	Table 31	568
CXCL10	no	unknown	Table 31	569
CXCL12	no	unknown	Table 6	228
CXCL5	no	unknown	Table 31	570
CXCL8	no	unknown	Table 31	571
CXCL9	no	unknown	Table 6	229
CXCR1	no	unknown	Table 6	230
CXCR3	no	unknown	Table 31	572
CXCR4	yes	AD	Table 6	231
CXorf40A	no	unknown	Table 6	232
CYBA	yes	AR	Table 31	573
CYBB	yes	XLR	Table 6	233
CYP2S1	no	unknown	Table 6	234
DCLRE1B	yes	AR	Table 31	574
DCLRE1C	yes	AR	Table 6	235
DDX1	no	unknown	Table 6	236
DDX58	yes	AD	Table 6	237
DHX58	no	unknown	Table 6	238
DKC1	yes	XLR	Table 6	239
DNAJC21	yes	AR	Table 31	575
DNASE1L3	yes	AR	Table 31	576
DNASE2	yes	AR	Table 31	577
DNER	no	unknown	Table 6	31
DNMT3B	yes	AR	Table 25B	240
DOCK2	yes	AR	Table 6	241
DOCK8	yes	AR	Table25A; Table 25B	242
DSC1	no	unknown	Table 6	243
DUSP16	no	unknown	Table 6	32
EBF1	no	unknown	Table 31	578
EDIL3	no	unknown	Table 6	34
EEA1	no	unknown	Table 6	35
EGF	no	unknown	Table 31	579
EGR1	no	unknown	Table 6	244

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
EHF	no	unknown	Table 6	36
ELANE	yes	AD	Table 6	245
EMB	no	unknown	Table 6	37
EPG5	yes	AR	Table 6	246
ERCC6L2	yes	AR	Table 31	580
ETF1	no	unknown	Table 6	247
ETV6	no	AD	Table 6	38
EXTL3	yes	AR	Table 31	581
F9	no	unknown	Table 6	248
FAAP24	yes	AR	Table 31	582
FADD	yes	AR	Table 31	583
FAS	yes	AD	Table 6	249
FASLG	yes	AD	Table 6	250
FAT4	yes	AR	Table 31	584
FCER2	no	unknown	Table 31	585
FCGR2A	no	AD AR	Table 6	251
FCGR3A	yes	AR	Table 6	252
FCN3	yes	AR	Table 6	253
FERMT3	yes	AR	Table 31	586
FEZ1	no	unknown	Table 6	254
FHL2	no	unknown	Table 6	39
FIS1	no	unknown	Table 31	587
FOS	no	unknown	Table 6	255
FOXH1	no	unknown	Table 6	256
FOXN1	yes	AR	Table 6	257
FOXP3	yes	XLR	Table 6	258
FPR1	yes	unknown	Table 6	259
FPR2	no	unknown	Table 6	41
FPR3	no	unknown	Table 6	42
FUK	no	unknown	Table 6	43
G6PC3	yes	AR	Table 6	260
G6PD	yes	XLD	Table 31	588
GATA2	yes	AD	Table 6	261
GDA	no	unknown	Table 6	44
GDPD4	no	unknown	Table 6	45
GFI1	yes	AD	Table 6	262
GINS 1	yes	AR	Table 31	589
GOLGB1	no	unknown	Table 6	263
GPATCH2	no	unknown	Table 6	46
GPC5	no	unknown	Table 6	47
GPRC5A	no	unknown	Table 6	264
GRAP2	no	unknown	Table 6	265
GRIA3	no	unknown	Table 6	51
GTPBP4	no	unknown	Table 6	52
HAX1	yes	AR	Table 6	266
HCN1	no	unknown	Table 6	53
HELLS	yes	AR	Table 6	267
HERC5	no	unknown	Table 31	590
HERC6	no	unknown	Table 31	591
HEXA	no	unknown	Table 6	54

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
HIVEP1	no	unknown	Table 6	268
HIVEP2	no	AD	Table 6	269
HIVEP3	no	unknown	Table 6	270
HK2	no	unknown	Table 6	55
HMGB1	no	unknown	Table 31	592
HMOX1	yes	AR	Table 31	593
HNRNPLL	no	unknown	Table 6	271
HP	no	unknown	Table 6	272
HPCAL1	no	unknown	Table 6	273
HPR	no	unknown	Table 6	57
HTR2A	no	unknown	Table 6	274
HYOU1	yes	AR	Table 31	594
ICAM1	no	unknown	Table 31	595
ICOS	yes	AR	Table 6	275
ICOSLG	no	unknown	Table 31	596
IDI1	no	unknown	Table 6	276
IDI2	no	unknown	Table 6	59
IDO2	no	unknown	Table 6	61
IFI35	no	unknown	Table 31	597
IFIH 1	yes	AD	Table 6	277
IFIT1	no	unknown	Table 31	598
IFIT2	no	unknown	Table 31	599
IFIT3	no	unknown	Table 31	600
IFNAR1	no	unknown	Table 6	278
IFNAR2	yes	AR	Table 6	279
IFNG	no	unknown	Table 6	280
IFNGR1	yes	ADAR	Table 6	281
IFNGR2	yes	ADAR	Table 6	282
IFNLR1	no	unknown	Table 6	62
IGHM	yes	AR	Table 31	601
IGHMBP2	no	unknown	Table 31	602
IGKC	yes	AR	Table 31	603
IGLL1	yes	AR	Table 6	283
IKBKB	yes	ADAR	Table 6	284
IKBKG	yes	XLD XLR	Table 6	285
IKZF1	yes	AD	Table 6	286
IL10	yes	AR	Table 6	287
IL10RA	yes	AR	Table 6	288
IL10RB	yes	AR	Table 6	289
IL12B	yes	AR	Table 6	290
IL12RB1	yes	AR	Table 6	291
IL17F	yes	AD	Table 6	292
IL17RA	yes	AR	Table 6	293
IL17RC	yes	AR	Table 31	604
IL1B	no	unknown	Table 6	294
IL1RN	yes	AR	Table 31	605
IL21	yes	AR	Table 6	295
IL21R	yes	ADAR	Table 6	296
IL2RA	yes	AR	Table 6	297
IL2RG	yes	XLR	Table 6	298
IL3	no	unknown	Table 31	606

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
IL36RN	yes	AR	Table 31	607
IL4	no	unknown	Table 31	608
IL4R	no	unknown	Table 6	299
IL7	no	unknown	Table 6	300
IL7R	yes	AR	Table 6	301
INO80	yes	AR	Table 31	609
INPP5D	no	unknown	Table 31	610
IRAK1	yes	XLR	Table 31	611
IRAK4	yes	ADAR	Table 6	302
IRF2BP2	yes	AD	Table 31	612
IRF3	yes	AD	Table 6	303
IRF7	yes	AR	Table 6	304
IRF8	yes	ADAR	Table 6	305
IRGM	no	unknown	Table 6	306
ISG15	yes	AR	Table 6	307
ITCH	yes	AR	Table 31	613
ITGAM	no	unknown	Table 31	614
ITGB2	yes	AR	Table 31	615
ITK	yes	AR	Table 25B	308
ITPKB	no	unknown	Table 31	616
ITSN1	no	unknown	Table 31	617
ITSN2	no	unknown	Table 6	309
JAGN1	yes	AR	Table 6	310
JAK1	yes	ADAR	Table 31	618
JAK3	yes	AR	Table 6	311
JMY	no	unknown	Table 6	312
JUN	no	unknown	Table 6	313
KANK1	no	unknown	Table 6	65
KAT6B	no	unknown	Table 6	66
KCTD7	no	unknown	Table 6	67
KDM6A	yes	XLD XLR	Table 31	619
KITLG	no	unknown	Table 6	314
KMT2D	yes	AD	Table 31	620
KRAS	no	AD	Table 31	621
LAMTOR2	yes	AR	Table 6	315
LARP4B	no	unknown	Table 6	69
LAT	yes	AR	Table 31	622
LCK	yes	AR	Table 25B	316
LCP2	no	unknown	Table 6	317
LIG1	yes	AR	Table 6	318
LIG4	yes	AR	Table 6	319
LPIN2	yes	AR	Table 31	623
LRBA	yes	AR	Table 6	322
LRRK2	no	unknown	Table 31	624
LYST	yes	AR	Table 6	323
MAGEA9B	no	unknown	Table 6	325
MAGT1	yes	XLR	Table 6; Table 25A	326
MALL	no	unknown	Table 6	72
MALT1	yes	AR	Table 6	327
MAP3K14	yes	AR	Table 31	625

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
MAP3K2	no	unknown	Table 6	328
MAPK1	no	unknown	Table 6	329
MAPK3	no	unknown	Table 6	330
MAPK9	no	unknown	Table 6	73
MASP2	yes	AR	Table 31	626
MAVS	no	unknown	Table 6	331
MB21D1	no	unknown	Table 31	627
MBL2	no	AD	Table 31	628
MCEE	no	unknown	Table 6	74
MCM4	yes	AR	Table 31	629
MCM5	no	AR	Table 31	630
MDC1	no	unknown	Table 31	631
MECP2	no	XLD XLR	Table 6	332
MEF2C	no	AD	Table 31	632
MEFV	yes	ADAR	Table 31	633
MEX3C	no	unknown	Table 6	333
MFN1	no	unknown	Table 31	634
MFN2	no	unknown	Table 31	635
MGATS	no	unknown	Table 6	75
MKL1	yes	AR	Table 6	89
MLH1	no	ADAR	Table 31	636
MMP9	no	unknown	Table 31	637
MOGS	yes	AR	Table 31	638
MON1A	no	unknown	Table 31	639
MON1B	no	unknown	Table 31	640
MRE11A	no	AR	Table 6	334
MS4A1	yes	AR	Table 6	335
MSH2	no	ADAR	Table 31	641
MSH5	no	AR	Table 31	642
MSH6	yes	AR	Table 31	643
MSN	yes	unknown	Table 6	336
MTHFD1	yes	AR	Table 6	337
MVK	yes	AD AR	Table 31	644
MX1	no	unknown	Table 31	645
MX2	no	unknown	Table 31	646
MYD88	yes	ADAR	Table 6	338
MYSM1	yes	AR	Table 31	647
NBAS	yes	AR	Table 31	648
NBN	yes	ADAR	Table 6	339
NCF1	yes	AR	Table 31	649
NCF2	yes	AR	Table 31	650
NCF4	yes	AR	Table 31	651
NCSTN	yes	AD	Table 31	652
NFAT5	yes	AD	Table 31	653
NFIC	no	unknown	Table 6	340
NFIL3	no	unknown	Table 6	92
NFKB1	yes	AD	Table 6	341
NFKB2	yes	AD	Table 6	342
NFKBIA	yes	AD	Table 6	343
NHEJ1	yes	AR	Table 6	344
NHP2	yes	AR	Table 31	654

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
NLRC4	yes	AD	Table 31	655
NLRP1	yes	AR	Table 31	656
NLRP12	yes	AD	Table 6	93
NLRP2	no	unknown	Table 31	657
NLRP3	yes	AD	Table 6	345
NLRX1	no	unknown	Table 31	658
NOD1	no	unknown	Table 31	659
NOD2	yes	AD	Table 6	346
NOP10	yes	AR	Table 31	660
NQO2	no	unknown	Table 6	94
NRIP1	no	unknown	Table 6	95
NSMCE3	yes	AR	Table 31	661
OAS1	no	AD	Table 31	662
OAS2	no	unknown	Table 31	663
OAS3	no	unknown	Table 31	664
OASL	no	unknown	Table 31	665
ORAI1	yes	ADAR	Table 6	347
ORC4	no	AR	Table 31	666
OSTM1	yes	AR	Table 6	348
OTULIN	yes	AR	Table 31	667
OVOL2	no	unknown	Table 6	98
PARN	yes	ADAR	Table 31	668
PCCA	no	AR	Table 31	669
PCCB	no	AR	Table 31	670
PDCD1	no	unknown	Table 31	671
PDCD1LG2	no	unknown	Table 31	672
PDE3B	no	unknown	Table 6	99
PDGFRA	no	unknown	Table 6	100
PDSS2	no	unknown	Table 6	101
PEPD	yes	AR	Table 31	673
PGM3	yes	AR	Table 6	349
PHACTR4	no	unknown	Table 6	102
PIAS1	no	unknown	Table 6	103
PIAS2	no	unknown	Table 6	350
PIK3CD	yes	AD	Table 6	104
PIK3R1	yes	AD AR	Table 6	351
PINK1	no	unknown	Table 31	674
PKHD1	no	unknown	Table 6	105
PLAU	no	unknown	Table 31	675
PLAUR	no	unknown	Table 31	676
PLCG1	no	unknown	Table 31	677
PLCG2	yes	AD	Table 6	352
PLD1	no	AR	Table 31	678
PLEKHM1	yes	AR	Table 31	679
PLK1	no	unknown	Table 31	680
PLXNB1	no	unknown	Table 31	681
PMM2	no	AR	Table 31	682
PMS2	yes	AR	Table 6	353
PNP	yes	unknown	Table 25B	354
PNPLA4	no	unknown	Table 6	107

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
PNPT1	no	unknown	Table 6	108
POLA1	yes	XLR	Table 6	355
POLE	yes	AR	Table 6	356
POLE2	yes	AR	Table 31	683
PPM1A	no	unknown	Table 31	684
PPP2R3B	no	unknown	Table 6	109
PRF1	yes	AD AR	Table 6	357
PRKCB	no	unknown	Table 6	110
PRKCD	yes	AR	Table 6	358
PRKCH	no	unknown	Table 6	111
PRKDC	yes	ADAR	Table 6	359
PRKN	no	unknown	Table 31	685
PROC	no	unknown	Table 6	360
PRRC2A	no	unknown	Table 31	686
PSEN1	yes	AD	Table 31	687
PSENEN	yes	AD	Table 31	688
PSMA7	no	unknown	Table 31	689
PSMB8	yes	AR	Table 6	361
PSTPIP1	yes	AD	Table 6	112
PTEN	yes	AD	Table 6	362
PTPN2	no	unknown	Table 6	113
PTPRC	yes	AR	Table 6	363
PTPRN2	no	unknown	Table 6	114
PURA	no	unknown	Table 6	364
RAB27A	yes	AR	Table 6	365
RAB37	no	unknown	Table 6	115
RAB5A	no	unknown	Table 31	690
RAB5B	no	unknown	Table 31	691
RAB5C	no	unknown	Table 31	692
RAB7A	no	unknown	Table 6	366
RABGEF1	no	unknown	Table 6	367
RAC2	yes	AD	Table 6	368
RAD50	no	unknown	Table 31	693
RAD51	no	unknown	Table 6	369
RAG1	yes	AD AR	Table25A; Table 25B	370
RAG2	yes	AR	Table 6	371
RANBP2	yes	AD	Table 31	694
RASGRP1	yes	AR	Table 31	695
RBCK1	yes	AR	Table 6	372
RBFOX1	no	unknown	Table 6	116
RCC1	no	unknown	Table 6	117
RELA	no	unknown	Table 31	696
RELB	yes	AR	Table 31	697
RFX5	yes	AR	Table 6	373
RFXANK	yes	AR	Table 6	374
RFXAP	yes	AR	Table 6	375
RGCC	no	unknown	Table 6	118
RHOH	yes	AR	Table 31	698
RHOQ	no	unknown	Table 6	119
RIPK1	no	unknown	Table 6	376
RIPK3	no	unknown	Table 6	377

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
RLTPR	yes	AR	Table 31	699
RNASE3	no	unknown	Table 6	120
RNASEH2A	yes	AR	Table 6	379
RNASEH2B	yes	AR	Table 6	380
RNASEH2C	yes	AR	Table 6	381
RNASEL	no	unknown	Table 6	382
RNF125	no	unknown	Table 31	700
RNF168	yes	AR	Table 6	383
RNF31	yes	AR	Table 6	384
RORC	yes	AR	Table 31	701
RPSA	yes	AD	Table 31	702
RPTOR	no	unknown	Table 6	123
RSAD2	no	unknown	Table 31	703
RTEL1	yes	AR	Table 6	386
SALL2	no	unknown	Table 6	388
SAMD9	yes	AD	Table 31	704
SAMD9L	yes	AD	Table 31	705
SAMHD1	yes	AR	Table 6	389
SBDS	yes	AR	Table 6	390
SEMA3E	yes	AD	Table 31	706
SERPINA1	no	AR	Table 31	707
SERPINB2	no	unknown	Table 31	708
SERPINB4	no	unknown	Table 6	124
SERPINB6	no	unknown	Table 6	125
SERPING1	yes	AD	Table 31	709
SH2D1A	yes	XLR	Table 6	391
SH3BP2	yes	AD	Table 31	710
SHARPIN	no	unknown	Table 6	392
SKIV2L	no	AR	Table 6	393
SLC17A5	no	unknown	Table 6	127
SLC29A3	yes	AR	Table 31	711
SLC35C1	yes	AR	Table 31	712
SLC37A4	yes	AR	Table 6	394
SLC3A2	no	unknown	Table 6	126
SLC46A1	yes	AR	Table 6	395
SLC7A7	no	AR	Table 31	713
SLC8A1	no	unknown	Table 6	396
SLC9A1	no	unknown	Table 31	714
SMAD2	no	unknown	Table 6	397
SMAD3	no	unknown	Table 6	398
SMAD4	no	unknown	Table 6	399
SMARCAL1	yes	AR	Table 31	715
SMARCD2	yes	AR	Table 31	716
SMC3	no	AD	Table 31	717
SMURF2	no	unknown	Table 31	718
SNAP29	no	unknown	Table 6	400
SNCA	no	unknown	Table 6	429
SNHG3	no	unknown	Table 6	128
SNX10	yes	AR	Table 6	430
SNX5	no	unknown	Table 6	130

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
SOCS2	no	unknown	Table 6	131
SP110	yes	AR	Table 6	431
SP140	no	unknown	Table 6	432
SPINK5	yes	AR	Table 6	433
SQSTM1	no	unknown	Table 6	434
SRP54	yes	AD	Table 31	719
ST8SIA5	no	unknown	Table 6	133
STAT1	yes	AD AR	Table25A; Table 25B	436
STAT2	yes	AR	Table 6	437
STAT3	yes	AD	Table 25B	438
STAT5B	yes	AR	Table 6	439
STIM1	yes	ADAR	Table 6	440
STIM2	no	unknown	Table 6	134
STK3	no	unknown	Table 25B; Table 26	135
STK4	yes	AR	Table 6	441
STN1	yes	AR	Table 31	720
STX11	yes	AR	Table 6	442
STXBP2	yes	ADAR	Table 6	443
SYNCRIP	no	unknown	Table 6	444
TAP1	yes	AR	Table 6	446
TAP2	yes	AR	Table 6	447
TAPBP	yes	unknown	Table 6	448
TAZ	yes	XLR	Table 6	449
TBC1D15	no	unknown	Table 31	721
TBC1D16	no	unknown	Table 6	136
TBC1D17	no	unknown	Table 31	722
TBK1	yes	AD	Table 6	450
TBX1	yes	AD	Table 6	451
TBXT	no	unknown	Table 6	445
TCF3	yes	AD	Table 31	723
TCIRG1	yes	ADAR	Table 6	452
TCN2	yes	AR	Table 31	724
TEK	no	AD	Table 31	725
TERT	yes	ADAR	Table 31	727
TFPI	no	unknown	Table 31	728
TFRC	yes	AR	Table 31	729
THBD	yes	AD	Table 31	730
THBS1	no	unknown	Table 31	731
TICAM1	yes	AD AR	Table 6	453
TINF2	yes	AD	Table 31	732
TIRAP	yes	AR	Table 31	733
TLR3	yes	AD	Table 6	454
TLR4	no	unknown	Table 6	455
TMC6	yes	AR	Table 31	734
TMC8	yes	AR	Table 31	735
TMEM173	yes	AD	Table 6	456
TNF	no	unknown	Table 6	457
TNFAIP3	yes	AD	Table 6	458
TNFRSF10A	no	unknown	Table 6	138
TNFRSF11A	yes	AD AR	Table 6	459
TNFRSF11B	no	AR	Table 6	460

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
TNFRSF13B	yes	AD AR	Table 6	461
TNFRSF13C	yes	AR	Table 6	139
TNFRSF17	no	unknown	Table 31	736
TNFRSF18	no	unknown	Table 6	140
TNFRSF1A	yes	AD	Table 31	737
TNFRSF4	yes	AR	Table 6	462
TNFRSF8	no	unknown	Table 6	463
TNFSF10	no	unknown	Table 31	738
TNFSF11	yes	AR	Table 6	464
TNFSF12-TNFSF13	yes	AD	Table 6	465
TNFSF13B	no	unknown	Table 31	740
TNIP1	no	unknown	Table 31	741
TP53	no	AD AR	Table 6	466
TP53AIP1	no	unknown	Table 31	742
TPP1	yes	AD AR	Table 31	743
TPP2	yes	AR	Table 31	744
TRAC	yes	AR	Table 31	745
TRAF3	yes	AD	Table 6	467
TRAF3IP2	yes	AR	Table 31	746
TRAF6	no	unknown	Table 6	468
TRAFD1	no	unknown	Table 6	141
TREX1	yes	ADAR	Table 6	469
TRIM25	no	unknown	Table 31	747
TRIM37	no	AR	Table 31	748
TRNT1	yes	AR	Table 6	470
TRPM2	no	unknown	Table 6	142
TTC37	yes	AR	Table 31	749
TTC7A	yes	AR	Table 6	471
TYK2	yes	AR	Table 25B	144
UBD	no	unknown	Table 31	750
UBE2N	no	unknown	Table 6	145
UNC119	no	AD	Table 6	472
UNC13D	yes	AR	Table 6	473
UNC93B1	yes	unknown	Table 6	474
UNG	yes	AR	Table 6	475
USB1	yes	AR	Table 31	751
USP15	no	unknown	Table 31	752
USP18	yes	AR	Table 6	476
USP20	no	unknown	Table 6	477
USP21	no	unknown	Table 31	753
USP25	no	unknown	Table 31	754
USP3	no	unknown	Table 31	755
VAPA	no	unknown	Table 6	478
VAV1	no	unknown	Table 31	756
VCP	no	AD	Table 6	479
VDAC1	no	unknown	Table 6	480
VDR	no	AD AR	Table 31	757
VEGFA	no	unknown	Table 31	758
VPS13B	yes	AR	Table 6	481
VPS45	yes	AR	Table 6	482

RefSeq Gene Symbol	IUIS-334 Gene	Disease Model	Cross-referenced tables	Gene Number (GN)
VSTM1	no	unknown	Table 6	147
VWA2	no	unknown	Table 6	148
WAS	yes	XLR	Table25A; Table 25B	483
WASHC5	no	AD AR	Table 31	759
WDR1	yes	AR	Table 31	760
WEE1	no	unknown	Table 6	484
WIPF1	yes	AR	Table 6; Table 25B	485
WRAP53	yes	AR	Table 31	761
XAF1	no	unknown	Table 31	762
XIAP	yes	XLD XLR	Table 6	486
YBX1	no	unknown	Table 6	487
YWHAZ	no	unknown	Table 6	488
ZAP70	yes	AD AR	Table 6	489
ZBTB24	yes	AR	Table 6	490

For the 663-genes list in **Table 47,** the following are provided: RefSeq Gene Symbol, whether the gene is included in the IUIS-334 list (Picard et al. (2018) J Clin Immunol. 38(1):96-128; Bousfiha et al. (2018) J Clin Immunol. 38(1): 129-143), the Disease Model, Cross-referenced tables that also list a subset of the 663 genes, and Gene Number.

[0417] Stricter criteria can be used for the Disease Model designation. For example, **Table 6** uses AR, AD, XLR, and XLD disease models based on the OMIM entries irrespective of whether the genetic disease is a classical immunodeficiency disorder. Whereas, **Table 47** lists AR, AD, XLR, and XLD entries only for those genes that are on the IUIS-334 gene list. For example, the gene PKHD 1 causes AR polycystic kidney disease and is listed as an AR disease in **Table 6** but with an "unknown" Disease Model in **Table 47.** However, those skilled in the art understand that additional disease phenotypes are often added to the clinical picture once more patients are identified. In Burgmaier et al. 2019 (Sci Rep. 9(1):7919), patients are reported to have the immune phenotypes of thrombocytopenia, anemia, and leukopenia.

Example 28 - DV Pipeline Variant Burden Summary Data

[0418] Top variants from the variant burden analyses were identified on the basis of filtering criteria that were applied to both the Dis and Rep PML cohorts (including a subset of functional annotation), followed by determining the subset of filtered variants that were found in both cohorts. The filtering criteria used were:

1. A. 663-genes (het and horn files), IMPACT = HIGH or MODERATE, ≥ 2 EUR cases or ≥ 1 AFR case, FET ≤ 0.1, OR > 1
2. B. All-genes (het and horn files) used 3 sets of filters:
    1. 1. IMPACT = HIGH, ≥ 3 EUR cases or ≥ 1 AFR case, FET ≤ 0.05, OR > 1
    2. 2. IMPACT = MODERATE, VARIANT CLASS = deletion or insertion, ≥ 3 EUR cases or ≥ 1 AFR case, FET ≤ 0.05, OR > 1
    3. 3. IMPACT = MODERATE, VARIANT CLASS = SNV, PolyPhen = damaging and SIFT = D (deleterious), ≥ 3 EUR cases or ≥ 2 AFR case, FET ≤ 0.05, OR > 1

After finding the set of overlapping variants for the Dis and Rep cohorts according to the above filters, they were further assessed for biology of the genes using PubMed and ranked as follows: 1 = strong biology, 2 = medium biology, 3=unknown biology. Variants were also excluded for: 1) very high PML case counts (more likely to be false positives), 2) high frequency in one ethnicity but not the other (greater likelihood that the variant is benign as compared to a variant that is rare in both the EUR and AFR ethnicities), and 3) presence in a chr X gene (such variants are challenging to interpret as gender needs to be taken into account). The top 27 variants from this set of analyses and filtering criteria are listed in Tables 50A and 50B.

Table 50A: Top 27 variants from DV pipeline variant burden analyses

Gene	Variant (hg19)	Source	IUI S-334	Conseque nce	Impac t	Poly Phen	SIFT	SEQ ID
AIRE	21-45708278-G-A	663 het AFR	yes	missense	moder ate	prob. dam.	tol.	3268, 3523
ATM	11-108106443-T-A	663 horn AFR	yes	missense	moder ate	benign	tol.	3517
C8B	1-57409459-C-A	663 het AFR	yes	missense	moder ate	poss. dam.	del.	3015, 3502
CFHR2	1-196918605-A-G	663 het AFR	yes	missense	moder ate	benign	tol.	3505
DNASE1 L3	3-58191230-G-T	663 het AFR	yes	missense	moder ate	benign	tol.	3508

Gene	Variant (hg19)	Source	IUI S-334	Consequence	Impact	Poly Phen	SIFT	SEQ ID
DNER	2-230579019-G-A	663 het NFE	no	missense	moderate	benign	del.	3507
FCN2	9-137779251-G-A	all het NFE	no	missense	moderate	prob. dam.	del.	3515
FCN3	1-27699670-AG-A	663 het NFE	yes	frame shift	high	n/a; LOF	n/a; LOF	3500
GFI1	1-92946625-G-C	663 het AFR	yes	missense	moderate	benign	tol.	3503
HIVEP3	1-42047208-C-G	663 het AFR	no	missense	moderate	benign	tol.	3501
IFIH1	2-163136505-C-G	US 102402 05; 663 het NFE; all het NFE	yes	Splice donor	high	n/a; LOF	n/a; LOF	1041, 3506
IGLL1	22-23915583-T-C	663 het NFE	yes	missense	moderate	benign	tol.	3524
IGLL1	22-23915745-G-A	US 102402 05	yes	missense	moderate	benign	del.	1325, 3525
LIG1	19-48643270-C-T	663 het AFR	yes	missense	moderate	poss. dam.	del.	3522
LRBA	4-151793903-T-C	663 het NFE	yes	missense	moderate	benign	tol.	3509
LY9	1-160769595-AG-A	all het AFR	no	frame shift	high	n/a; LOF	n/a; LOF	3504
MCM5	22-35806756-G-A	663 het NFE	no	missense	moderate	poss. dam.	tol.	3273, 3526
MDC1	6-30673359-T-G	663 het NFE	no	missense	moderate	benign	tol.	3511
NQO2	6-3015818-G-A	663 het AFR	no	missense	moderate	prob. dam.	del.	3510
PKHD1	6-51798908-C-T	663 het AFR	no	missense	moderate	prob. dam.	del.	3514
PLCG2	16-81942175-A-G	US 102402 05	yes	missense	moderate	benign	del.	1263, 3519
PRAM1	19-8564523-T-G	all het NFE; all het AFR	no	missense	moderate	poss. dam.	del.	3521
SERPIN A1	14-94847262-T-A	663 het NFE	no	missense	moderate	prob. dam.	del.	3214, 3518
STXBP2	19-7712287-G-C	US 102402 05; 663 het NFE	yes	missense	moderate	prob. dam.	del.	1291, 3520
TAP1	6-32814942-C-T	663 het NFE; 663 horn AFR	yes	missense	moderate	prob. dam.	del.	3512
TAP1	6-32816772-C-A	663 het NFE; 663 horn AFR	yes	missense	moderate	benign	tol.	3513
TCIRG1	11-67818269-G-A	US 102402 05	yes	missense	mod	benign	tol.	1185, 3516

Table 50A lists the Gene, Variant position (hg19), Source of the variant (663-genes or all-genes het or horn variant burden files, and/or from US 10, 240,205), functional prediction (Consequence, Impact, PolyPhen, SIFT), and SEQ IDs. Loss of Function variants (LOF) are indicated in the Polyphen and SIFT columns as n/aLOF (e.g., no amino acid prediction information was available from these algorithms but the gnomAD annotation indicates "LOF: High-confidence").

Table 50B: Top 27 variants from DV pipeline variant burden analyses, distribution of PML cases by cohort (Dis and Rep) and ethnicity (EUR and AFR)

Gene	Variant (hg19)	Dis EUR (n = 49)		Rep EUR (n = 87)		Dis AFR (n = 21)		Rep AFR (n = 28)	
		Het	Hom	Het	Hom	Het	Hom	Het	Horn
AIRE	21-45708278-G-A	0	0	0	0	1	0	1	0
ATM	11-108106443-T-A	2	0	0	0	7	3	8	3
C8B	1-57409459-C-A	3	0	0	0	1	0	1	0
CFHR2	1-196918605-A-G	0	0	0	0	1	0	2	0
DNASE1L 3	3-58191230-G-T	0	0	0	0	1	0	1	0

Gene	Variant (hg19)	Dis EUR (n = 49)		Rep EUR (n = 87)		Dis AFR (n = 21)		Rep AFR (n = 28)	
		Het	Hom	Het	Hom	Het	Hom	Het	Horn
DNER	2-230579019-G-A	4	0	4	0	0	0	0	0
FCN2	9-137779251-G-A	3	0	3	0	0	0	1	0
FCN3	1-27699670-AG-A	5	0	7	0	0	0	2	0
GFI1	1-92946625-G-C	2	0	0	0	1	0	1	0
HIVEP3	1-42047208-C-G	8	0	8	0	2	0	3	0
IFIH1	2-163136505-C-G	6	0	5	0	1	0	0	0
IGLL1	22-23915583-T-C	2	0	3	0	0	0	0	0
IGLL1	22-23915745-G-A	2	0	1	0	1	0	0	0
LIG1	19-48643270-C-T	1	0	0	0	1	0	1	0
LRBA	4-151793903-T-C	6	0	9	0	0	0	1	0
LY9	1-160769595-AG-A	0	0	0	0	1	0	1	0
MCM5	22-35806756-G-A	3	0	4	0	0	0	0	0
MDC1	6-30673359-T-G	2	0	5	0	0	0	0	0
NQO2	6-3015818-G-A	0	0	0	0	1	0	2	0
PKHD1	6-51798908-C-T	0	0	0	0	2	0	1	0
PLCG2	16-81942175-A-G	3	0	3	0	4	0	1	0
PRAM1	19-8564523-T-G	27	1	14	4	6	0	4	1
SERPINA 1	14-94847262-T-A	7	0	12	0	0	0	1	0
STXBP2	19-7712287-G-C	2	0	2	0	0	0	0	0
TAP1	6-32814942-C-T	6	0	9	1	6	2	5	2
TAP1	6-32816772-C-A	6	0	9	1	6	2	5	2
TCIRG1	11-67818269-G-A	0	0	0	0	4	0	3	0

Table 50B shows the same top 27 variants listed in **Table 50A,** but with the distribution of number of PML cases that had a het or horn genotype and a breakdown by cohort (Dis or Rep) and ethnicity (EUR or AFR).

Example 29 - DV Pipeline Variant Burden Statisical Analyses

[0419] The set of 27 variants from the DV pipeline variant burden analyses (see Tables 50A and 50B) were ranked based on FET and OR values for four types of analyses: EUR only, AFR only, EUR+AFR normalized (norm.) wherein the gnomAD cohort sizes were adjusted to match the relative proportion of EUR (NFE in gnomAD) and AFR PML cases in the total cohort (n = 185), and EUR+AFR summed (sum.) wherein the subject data for the gnomAD NFE and AFR cohorts were simply added together. Three sets of statistical analyses were performed using three publically available resources (gnomAD exome, gnomAD genome, and TGP exome), but only the gnomAD exome analyses are presented in Tables 51-62 as the results were comparable for nearly all variants. For each set of 3 tables (tier 1, tier 2, and tiers 1 and 2 combined), FET and OR were calculated for all 27 variants, variants were sorted on OR (descending), and variants with FET > 0.05 were excluded. For Tier 1 variants (Tables 51, 54, 57, and 60), preference was given to variants with OR > 3 and only moderate frequency in PML cases (e.g., PRAM1 19-8564523-T-G was considered too frequent). The "Reads" listed in the tables corresponds to the read counts in the sequencing data (e.g., the number of individuals sampled for the variant). **Tables 51 and 54** also have a Panel rank wherein the top 7 EUR variants and top 10 AFR variants are assigned a rank that is used in Tables 63A-64B.

Table 51: 7 EUR variants (tier 1) with FET and OR values

Gene	Variant (hg19)	PML All EUR			gnomAD NFE			Dominant Model		Panel rank
		He t	Ho m	Read s	Het	Ho m	Read s	FET	OR	
IGLL1	22-23915745-G-A	3	0	136	42	0	5670 3	1.77E-04	30.42	1
MDC1	6-30673359-T-G	7	0	136	302	0	5460 8	1.34E-05	9.76	2
STXBP 2	19-7712287-G-C	4	0	136	267	0	5579 8	4.46E-03	6.30	3
FCN2	9-137779251-G-A	6	0	136	461	1	5645 7	9.85E-04	5.59	4
IGLL1	22-23915583-T-C	5	0	136	446	0	5681 1	4.70E-03	4.82	5
MCM5	22-35806756-G-A	7	0	136	665	7	5687 5	1.28E-03	4.54	6
IFIH1	2-163136505-C-G	11	0	136	122 5	6	5643 8	2.19E-04	3.95	7

Table 52: 9 EUR variants (tier 2) with FET and OR values

Gene	Variant (hg19)	PML All EUR			gnomAD NFE			Dominant Model	
		Het	Hom	Reads	Het	Hom	Reads	FET	OR
PRAM1	19-8564523-T-G	41	5	136	2076	250	30728	1.65E-18	6.24
ATM	11-108106443-T-A	2	0	136	143	0	56802	4.74E-02	5.91
TAP1	6-32816772-C-A	15	1	135	2250	26	55054	1.61E-04	3.12
TAP1	6-32814942-C-T	15	1	136	2258	28	55157	1.80E-04	3.08
PLCG2	16-81942175-A-G	6	0	136	831	2	56290	1.64E-02	3.07
FCN3	1-27699670-AG-A	12	0	136	1826	24	56778	1.77E-03	2.87
DNER	2-230579019-G-A	8	0	136	55	3	2670	1.32E-02	2.81
SERPINA1	14-94847262-T-A	19	0	136	3946	105	56878	4.25E-03	2.12
LRBA	4-151793903-T-C	15	0	136	3046	62	55340	1.32E-02	2.08

Table 53: 16 EUR variants (tiers 1 and 2) with FET and OR values

Gene	Variant (hg19)	PML All EUR			gnomAD NFE			Dominant Model	
		Het	Hom	Reads	Het	Hom	Reads	FET	OR
IGLL1	22-23915745-G-A	3	0	136	42	0	56703	1.77E-04	30.42
MDC1	6-30673359-T-G	7	0	136	302	0	54608	1.34E-05	9.76
STXBP2	19-7712287-G-C	4	0	136	267	0	55798	4.46E-03	6.30
PRAM1	19-8564523-T-G	41	5	136	2076	250	30728	1.65E-18	6.24
ATM	11-108106443-T-A	2	0	136	143	0	56802	4.74E-02	5.91
FCN2	9-137779251-G-A	6	0	136	461	1	56457	9.85E-04	5.59
IGLL1	22-23915583-T-C	5	0	136	446	0	56811	4.70E-03	4.82
MCM5	22-35806756-G-A	7	0	136	665	7	56875	1.28E-03	4.54
IFIH1	2-163136505-C-G	11	0	136	1225	6	56438	2.19E-04	3.95
TAP1	6-32816772-C-A	15	1	135	2250	26	55054	1.61E-04	3.12
TAP1	6-32814942-C-T	15	1	136	2258	28	55157	1.80E-04	3.08
PLCG2	16-81942175-A-G	6	0	136	831	2	56290	1.64E-02	3.07
FCN3	1-27699670-AG-A	12	0	136	1826	24	56778	1.77E-03	2.87
DNER	2-230579019-G-A	8	0	136	55	3	2670	1.32E-02	2.81
SERPINA1	14-94847262-T-A	19	0	136	3946	105	56878	4.25E-03	2.12
LRBA	4-151793903-T-C	15	0	136	3046	62	55340	1.32E-02	2.08

Table 54: 10 AFR variants (tier 1) with FET and OR values

Gene	Variant (hg19)	PML All AFR			gnomAD AFR			Dominant Model		Panel rank
		He t	Ho m	Read s	He t	Ho m	Read s	FET	OR	
LY9	1-160769595-AG-A	2	0	49	0	0	8123	3.52E-05	855.1 1	1
LIG1	19-48643270-C-T	2	0	49	1	0	8126	1.05E-04	339.7 2	2
PKHD1	6-51798908-C-T	3	0	49	15	0	8127	1.55E-04	35.15	3
AIRE	21-45708278-G-A	2	0	49	17	0	8079	5.70E-03	20.14	4
GFI1	1-92946625-G-C	2	0	49	19	0	7756	7.51E-03	17.29	5
CFHR2	1-196918605-A-G	3	0	49	37	0	8027	1.77E-03	14.07	6
NQO2	6-3015818-G-A	3	0	49	39	0	8128	1.97E-03	13.51	7
C8B	1-57409459-C-A	2	0	49	28	0	8128	1.38E-02	12.29	8
DNASE1 L3	3-58191230-G-T	2	0	49	32	0	8126	1.75E-02	10.75	9
PLCG2	16-81942175-A-G	5	0	49	12 8	0	7719	1.42E-03	6.73	10

Table 55: 3 AFR variants (tier 2) with FET and OR values

Gene	Variant (hg19)	PML All AFR			gnomAD AFR			Dominant Model	
		Het	Hom	Reads	Het	Hom	Reads	FET	OR
PRAM 1	19-8564523-T-G	10	1	49	136	4	6351	1.18E-08	12.82
HIVEP3	1-42047208-C-G	5	0	49	175	2	8041	4.67E-03	5.05
TCIRG1	11-67818269-G-A	7	0	49	318	1	8113	3.08E-03	4.07

Table 56: 13 AFR variants (tiers 1 and 2) with FET and OR values

Gene	Variant (hg19)	PML All AFR			gnomAD AFR			Dominant Model	
		Het	Hom	Reads	Het	Hom	Reads	FET	OR
LY9	1-160769595-AG-A	2	0	49	0	0	8123	3.52E-05	855.11
LIG1	19-48643270-C-T	2	0	49	1	0	8126	1.05E-04	339.72
PKHD1	6-51798908-C-T	3	0	49	15	0	8127	1.55E-04	35.15
AIRE	21-45708278-G-A	2	0	49	17	0	8079	5.70E-03	20.14
GFI1	1-92946625-G-C	2	0	49	19	0	7756	7.51E-03	17.29
CFHR2	1-196918605-A-G	3	0	49	37	0	8027	1.77E-03	14.07
NQO2	6-3015818-G-A	3	0	49	39	0	8128	1.97E-03	13.51
PRAM1	19-8564523-T-G	10	1	49	136	4	6351	1.18E-08	12.82
C8B	1-57409459-C-A	2	0	49	28	0	8128	1.38E-02	12.29
DNASE1L3	3-58191230-G-T	2	0	49	32	0	8126	1.75E-02	10.75
PLCG2	16-81942175-A-G	5	0	49	128	0	7719	1.42E-03	6.73
HIVEP3	1-42047208-C-G	5	0	49	175	2	8041	4.67E-03	5.05
TCIRG1	11-67818269-G-A	7	0	49	318	1	8113	3.08E-03	4.07

Table 57: 13 EUR+AFR variants (tier 1) with FET and OR values (normalized)

Gene	Variant (hg19)	PML All EUR+AFR			gnomAD NFE+AFR (norm.)			Dominant Model	
		Het	Ho m	Read s	Het	Hom	Reads	FET	OR
PKHD1	6-51798908-C-T	3	0	185	6	0	43916	5.99E-06	120.0 0
LY9	1-160769595-AG-A	2	0	185	6	0	43924	4.82E-04	79.75
CFHR2	1-196918605-A-G	3	0	185	10	0	43752	2.04E-05	72.02
NQO2	6-3015818-G-A	3	0	185	12	0	43941	3.18E-05	60.30
AIRE	21-45708278-G-A	2	0	185	8	0	43685	7.78E-04	59.62
IGLL1	22-23915745-G-A	4	0	185	58	0	43836	1.39E-04	16.67
TCIRG 1	11-67818269-G-A	7	0	185	227	0	43910	6.39E-05	7.57
ATM	11-108106443-TA	17	6	185	797	88	43909	6.97E-12	6.90
MDC1	6-30673359-T-G	7	0	185	251	0	42164	1.50E-04	6.57
PRAM 1	19-8564523-T-G	51	6	185	1562	185	24271	2.94E-21	5.74
FCN2	9-137779251-G-A	7	0	185	342	1	43653	7.60E-04	4.96
STXBP 2	19-7712287-G-C	4	0	185	198	0	43049	1.13E-02	4.78
PLCG2	16-81942175-A-G	11	0	185	645	1	43425	1.27E-04	4.19

Table 58: 8 EUR+AFR variants (tier 2) with FET and OR values (normalized)

Gene	Variant (hg19)	PML All EUR+AFR			gnomAD NFE+AFR (norm.)			Dominant Model	
		He t	Ho m	Read s	Het	Hom	Reads	FET	OR
TAP1	6-32816772-C-A	26	5	184	2144	52	42515	7.80E-09	3.72
TAP1	6-32814942-C-T	26	5	185	2140	53	42544	8.50E-09	3.70
GFI1	1-92946625-G-C	4	0	185	256	1	40597	3.16E-02	3.47
IGLL1	22-23915583-T-C	5	0	185	352	0	43910	1.78E-02	3.44
MCM5	22-35806756-G-A	7	0	185	494	5	43964	5.71E-03	3.43
IFIH1	2-163136505-C-G	12	0	185	910	4	43635	6.16E-04	3.24
FCN3	1-27699670-AG-A	14	0	185	1448	18	43889	5.84E-03	2.37
SERPINA 1	14-94847262-TA	20	0	185	2934	77	43966	4.05E-02	1.65

Table 59: 21 EUR+AFR variants (tiers 1 and 2) with FET and OR values (normalized)

Gene	Variant (hg19)	PML All EUR+AFR			gnomAD NFE+AFR (norm.)			Dominant Model	
		Het	Hom	Read s	Het	Hom	Reads	FET	OR
PKHD1	6-51798908-C-T	3	0	185	6	0	43916	5.99E-06	120.0 0
LY9	1-160769595-AG-A	2	0	185	6	0	43924	4.82E-04	79.75
CFHR2	1-196918605-A-G	3	0	185	10	0	43752	2.04E-05	72.02
NQO2	6-3015818-G-A	3	0	185	12	0	43941	3.18E-05	60.30
AIRE	21-45708278-G-A	2	0	185	8	0	43685	7.78E-04	59.62
IGLL1	22-23915745-G-A	4	0	185	58	0	43836	1.39E-04	16.67
TCIRG1	11-67818269-G-A	7	0	185	227	0	43910	6.39E-05	7.57

Gene	Variant (hg19)	PML All EUR+AFR			gnomAD NFE+AFR (norm.)			Dominant Model	
		Het	Hom	Read s	Het	Hom	Reads	FET	OR
ATM	11-108106443-TA	17	6	185	797	88	43909	6.97E-12	6.90
MDC1	6-30673359-T-G	7	0	185	251	0	42164	1.50E-04	6.57
PRAM1	19-8564523-T-G	51	6	185	1562	185	24271	2.94E-21	5.74
FCN2	9-137779251-G-A	7	0	185	342	1	43653	7.60E-04	4.96
STXBP2	19-7712287-G-C	4	0	185	198	0	43049	1.13E-02	4.78
PLCG2	16-81942175-A-G	11	0	185	645	1	43425	1.27E-04	4.19
TAP1	6-32816772-C-A	26	5	184	2144	52	42515	7.80E-09	3.72
TAP1	6-32814942-C-T	26	5	185	2140	53	42544	8.50E-09	3.70
GFI1	1-92946625-G-C	4	0	185	256	1	40597	3.16E-02	3.47
IGLL1	22-23915583-T-C	5	0	185	352	0	43910	1.78E-02	3.44
MCM5	22-35806756-G-A	7	0	185	494	5	43964	5.71E-03	3.43
IFIH1	2-163136505-C-G	12	0	185	910	4	43635	6.16E-04	3.24
FCN3	1-27699670-AG-A	14	0	185	1448	18	43889	5.84E-03	2.37
SERPIN A1	14-94847262-TA	20	0	185	2934	77	43966	4.05E-02	1.65

Table 60: 12 EUR+AFR variants (tier 1) with FET and OR values (summed)

Gene	Variant (hg19)	PML All EUR+AFR			gnomAD NFE+AFR (sum.)			Dominant Model	
		Het	Ho m	Reads	Het	Ho m	Read s	FET	OR
LY9	1-160769595-AG-A	2	0	185	8	0	64946	3.56E-04	88.57
PKHD1	6-51798908-C-T	3	0	185	18	0	64937	2.89E-05	59.42
AIRE	21-45708278-G-A	2	0	185	22	0	64592	2.15E-03	32.06
CFHR2	1-196918605-A-G	3	0	185	37	0	64650	2.09E-04	28.78
NQO2	6-3015818-G-A	3	0	185	41	0	64972	2.74E-04	26.11
IGLL1	22-23915745-G-A	4	0	185	146	0	64826	9.30E-04	9.79
PRAM1	19-8564523-T-G	51	6	185	2212	254	37079	5.81E-23	6.25
MDC1	6-30673359-T-G	7	0	185	411	0	62235	2.71E-04	5.91
FCN2	9-137779251-G-A	7	0	185	471	1	64574	4.92E-04	5.34
STXBP2	19-7712287-G-C	4	0	185	274	0	63462	9.08E-03	5.10
TCIRG1	11-67818269-G-A	7	0	185	512	1	64920	7.71E-04	4.94
PLCG2	16-81942175-A-G	11	0	185	959	2	64009	1.34E-04	4.15

Table 61: 11 EUR+AFR variants (tier 2) with FET and OR values (summed)

Gene	Variant (hg19)	PML All EUR+AFR			gnomAD NFE+AFR (sum.)			Dominant Model	
		Het	Hom	Reads	Het	Hom	Reads	FET	OR
LIG1	19-48643270-C-T	3	0	185	281	0	64971	4.77E-02	3.79
MCM5	22-35806756-G-A	7	0	185	683	7	65003	3.97E-03	3.67
GFI1	1-92946625-G-C	4	0	185	361	1	60185	2.67E-02	3.65
IFIH1	2-163136505-C-G	12	0	185	1262	6	64538	3.48E-04	3.46
IGLL1	22-23915583-T-C	5	0	185	537	0	64915	1.98E-02	3.33
ATM	11-108106443-T-A	17	6	185	2755	333	64928	2.83E-05	2.84
TAP1	6-32814942-C-T	26	5	185	4070	149	62693	2.63E-06	2.79
TAP1	6-32816772-C-A	26	5	184	4101	149	62767	2.65E-06	2.79
FCN3	1-27699670-AG-A	14	0	185	2224	26	64892	6.97E-03	2.28
LRBA	4-151793903-T-C	16	0	185	3116	62	63322	4.01E-02	1.79
SERPI NA1	14-94847262-T-A	20	0	185	4073	106	65006	2.30E-02	1.76

Table 62: 23 EUR+AFR variants (tiers 1 and 2) with FET and OR values (summed)

Gene	Variant (hg19)	PML All EUR+AFR			gnomAD NFE+AFR (sum.)			Dominant Model	
		Het	Horn	Reads	Het	Horn	Reads	FET	OR
LY9	1-160769595-AG-A	2	0	185	8	0	64946	3.56E-04	88.57
PKHD1	6-51798908-C-T	3	0	185	18	0	64937	2.89E-05	59.42
AIRE	21-45708278-G-A	2	0	185	22	0	64592	2.15E-03	32.06
CFHR2	1-196918605-A-G	3	0	185	37	0	64650	2.09E-04	28.78

Gene	Variant (hg19)	PML All EUR+AFR			gnomAD NFE+AFR (sum.)			Dominant Model	
		Het	Horn	Reads	Het	Horn	Reads	FET	OR
NQO2	6-3015818-G-A	3	0	185	41	0	64972	2.74E-04	26.11
IGLL1	22-23915745-G-A	4	0	185	146	0	64826	9.30E-04	9.79
PRAM1	19-8564523-T-G	51	6	185	2212	254	37079	5.81E-23	6.25
MDC1	6-30673359-T-G	7	0	185	411	0	62235	2.71E-04	5.91
FCN2	9-137779251-G-A	7	0	185	471	1	64574	4.92E-04	5.34
STXBP2	19-7712287-G-C	4	0	185	274	0	63462	9.08E-03	5.10
TCIRG1	11-67818269-G-A	7	0	185	512	1	64920	7.71E-04	4.94
PLCG2	16-81942175-A-G	11	0	185	959	2	64009	1.34E-04	4.15
LIG1	19-48643270-C-T	3	0	185	281	0	64971	4.77E-02	3.79
MCM5	22-35806756-G-A	7	0	185	683	7	65003	3.97E-03	3.67
GFI1	1-92946625-G-C	4	0	185	361	1	60185	2.67E-02	3.65
IFIH1	2-163136505-C-G	12	0	185	1262	6	64538	3.48E-04	3.46
IGLL1	22-23915583-T-C	5	0	185	537	0	64915	1.98E-02	3.33
ATM	11-108106443-T-A	17	6	185	2755	333	64928	2.83E-05	2.84
TAP1	6-32814942-C-T	26	5	185	4070	149	62693	2.63E-06	2.79
TAP1	6-32816772-C-A	26	5	184	4101	149	62767	2.65E-06	2.79
FCN3	1-27699670-AG-A	14	0	185	2224	26	64892	6.97E-03	2.28
LRBA	4-151793903-T-C	16	0	185	3116	62	63322	4.01E-02	1.79
SERPINA1	14-94847262-T-A	20	0	185	4073	106	65006	2.30E-02	1.76

Example 30 - Variant Panels with Diagnostic Yield and Population Impact Assessment

[0420] Variant panels consisting of the top 7 EUR variants (**Table 51**) and top 10 AFR variants (**Table 54**) were assessed for their non-redundant diagnostic yield in the PML cases (e.g., a PML case with ≥ 2 panel variants was counted only once). These panels were similarly assessed in TGP exome data from 440 AFR and 436 EUR subjects, whose GRCh37-referenced BAM files were downloaded and re-analized using the DV pipeline. **Tables 63A** and **64A** show the results for EUR and AFR PML cases vs. TGP controls, respectively, wherein the PML Panel Frequency corresponds to the diagnostic yield and the Frequncy Ratio reflects the level of enrichment of the panel variants in the PML cases. The FET and OR values correspond to the number of PML cases and TGP controls that have any one or more of 3 variants, 4 variants, etc. for a given panel.

[0421] Tables 63B and **64B** contain the population impact assessment results as described in Tonk et al. (2016) Pharmacogenomics J.17(4):386-392, for the following: Sensitivity, Specificity, Positive predictive value (PPV), Negative predictive value (NPV), Population attributable frequency (PAF), Number needed to treat (NNT), and Number needed to genotype (NNG).

Table 63A: EUR 7-variant panel, diagnostic yield and FET and OR values

Variants Tested	EUR PML Cases (n = 136)	PML Panel Frequency	EUR TGP Subjects (n = 436)	TGP Panel Frequency	Frequency Ratio (PML/TGP)	FET	OR
EUR Top 3	14	0.103	5	0.011	8.98	4.02E-06	9.89
EUR Top 4	19	0.140	7	0.016	8.70	6.90E-08	9.95
EUR Top 5	24	0.176	11	0.025	6.99	7.45E-09	8.28
EUR Top 6	31	0.228	16	0.037	6.21	1.32E-10	7.75
EUR Top 7	40	0.294	30	0.069	4.27	1.06E-10	5.64

Table 63B: EUR 7-variant panel, population impact assessment

Variants Tested	Sensitivity	Specificity	PPV	NPV	Relative Risk (RR)	Risk Difference (RD)	PAF	NNT	NNG
EUR Top 3	9.1%	99.0%	10.7%	98.8%	8.94	0.09	0.08	11	953
EUR Top 4	15.1%	98.2%	10.2%	98.9%	9.04	0.09	0.13	11	574
EUR Top 5	17.6%	97.5%	8.4%	98.9%	7.66	0.07	0.15	14	503
EUR Top 6	22.7%	96.4%	7.6%	99.0%	7.24	0.07	0.20	15	394
EUR Top 7	25.8%	94.2%	5.5%	99.0%	5.38	0.04	0.21	22	366

Table 64A: AFR 10-variant panel, diagnostic yield and FET and OR values

Variants Tested	AFR PML Cases (n = 49)	PML Panel Frequency	AFR TGP Subjects (n = 440)	TGP Panel Frequency	Frequency Ratio (PML/TGP)	FET	OR
AFR Top 6	14	0.286	5	0.011	25.14	1.29E-11	34.80
AFR Top 7	17	0.347	6	0.014	25.44	3.97E-14	38.43
AFR Top 8	19	0.388	6	0.014	28.44	3.18E-16	45.81
AFR Top 9	20	0.408	6	0.014	29.93	2.66E-17	49.89
AFR Top 10	23	0.469	10	0.023	20.65	1.90E-18	38.04

Table 64B: AFR 10-variant panel, population impact assessment

Variants Tested	Sensitivity	Specificity	PPV	NPV	Relative Risk (RR)	Risk Difference (RD)	PAF	NNT	NNG
AFR Top 6	22.6%	99.2%	26.4%	99.0%	25.89	0.25	0.22	4	355
AFR Top 7	31.7%	98.8%	25.9%	99.1%	28.73	0.25	0.31	4	251
AFR Top 8	40.0%	98.6%	26.9%	99.2%	33.78	0.26	0.39	4	198
AFR Top 9	47.0%	98.3%	26.2%	99.3%	37.09	0.25	0.46	4	168
AFR Top 10	56.5%	96.7%	18.4%	99.4%	31.22	0.18	0.55	6	141

Example 31 - Variant Panels with Summary Information

[0422] Table 65 shows the breakdown of PML cases with a variant from the set of 17 variants (7 EUR and 10 AFR, **Tables 51 and 54**). In general, they are distributed across genders, ethnicities, and primary diseases. For example, all 7 EUR variants are found in two or more Primary Diseases (MS, HIV, Blood Cancers, Other). In contrast, about half of the 10 AFR variants are found only in HIV PML patients, but this is not surprising given the large proporation of AFR HIV PML cases (48 of 49) in the PML cohort (see **Table 46**).

[0423] Table 66 lists the dbSNP ID numbers for the set of 17 variants (7 EUR and 10 AFR, **Tables 51 and 54**), along with PubMed PMID numbers for supporting biology in the literature (see also **Tables 6 and 31** for additional PMID citations for a subset of the genes).

Table 65: EUR 7-variant and AFR 10-variant panels, distribution of PML cases by cohort, gender, ethnicity, and primary disease

Gene	Variant (hg19)	Panel Source Rank	Dis Cohort (n = 70)	Rep Cohort (n = 115)	PML TOTAL (n = 185	Males (n = 118)	Females (n = 67)	AFR (n = 49)	EUR (n = 136)	MS (n = 15)	HIV (n = 125)	Blood Cancers (n = 22)	Other (n = 23)
IGLL1	22-23915745-G-A	EUR 1	3	1	4	2	2	1	3	1	3	0	0
MDC1	6-30673359-T-G	EUR 2	2	7	9	9	0	1	8	0	8	0	1
STXBP2	19-7712287-G-C	EUR 3	2	2	4	1	3	0	4	2	1	0	1
FCN2	9-137779251-G-A	EUR 4	3	4	7	4	3	1	6	2	4	1	0
IGLL1	22-23915583-T-C	EUR 5	2	3	5	1	4	0	5	0	1	4	0
MCM5	22-35806756-G-A	EUR 6	3	4	7	4	3	0	7	1	5	0	1
IFIH1	2-163136505-C-G	EUR7	7	5	12	9	3	1	11	1	8	1	2
LY9	1-160769595-AG-A	AFR 1	1	1	2	1	1	2	0	0	1	0	1
LIG1	19-48643270-C-T	AFR 2	2	1	3	3	0	2	1	0	3	0	0
PKHD1	6-51798908-C-T	AFR 3	2	1	3	1	2	3	0	0	3	0	0
AIRE	21-45708278-G-A	AFR 4	1	1	2	0	2	2	0	0	2	0	0
GFI1	1-92946625-G-C	AFR 5	3	1	4	2	2	2	2	0	3	1	0
CFHR2	1-196918605-A-G	AFR 6	1	2	3	2	1	3	0	0	3	0	0
NQO2	6-3015818-G-A	AFR 7	1	2	3	2	1	3	0	0	3	0	0
C8B	1-57409459-C-A	AFR 8	4	1	5	2	3	2	3	1	3	1	0
DNASE1L3	3-58191230-G-T	AFR 9	1	1	2	2	0	2	0	0	2	0	0
PLCG2	16-81942175-	AFR 10	7	4	11	9	2	5	6	1	9	0	2

PLCG2	A-G	AFR 10	7	4	11	9	2	5	6	1	8	0	2

Table 66: EUR 7-variant and AFR 10-variant panels, dbSNP IDs and supporting biology

Gene	Variant (hg19)	dbSNP ID	Panel Source Rank	Supporting Biology - PubMed (PMID)
IGLL1	22-23915745-G-A	rs143780139	EUR 1	31291582
MDC1	6-30673359-T-G	rs143258964	EUR 2	30453211,30541848
STXBP2	19-7712287-G-C	rs35490401	EUR 3	28983403, 29599780
FCN2	9-137779251-G-A	rs76267164	EUR 4	20375618, 25251245, 28894916, 30747617,30868077
IGLL1	22-23915583-T-C	rs1064421	EUR 5	31291582
MCM5	22-35806756-G-A	rs2230933	EUR 6	27414250
IFIH1	2-163136505-C-G	rs35337543	EUR 7	22648297, 28716935, 28475461, 28553952,30201512
LY9	1-160769595-AG-A	rs763811636	AFR 1	15905546, 26667173, 27054584, 27482100,30791129
LIG1	19-48643270-C-T	rs3730947	AFR 2	30395541
PKHD1	6-51798908-C-T	rs199589074	AFR 3	31138820
AIRE	21-45708278-G-A	rs148012328	AFR 4	31167928
GFI1	1-92946625-G-C	rs149914857	AFR 5	26447191,31004086
CFHR2	1-196918605-A-G	rs148175483	AFR 6	29686068
NQO2	6-3015818-G-A	rs148024596	AFR 7	27692612
C8B	1-57409459-C-A	rs139498867	AFR 8	31270218
DNASE1 L3	3-58191230-G-T	rs 12491947	AFR 9	28533778, 30593563, 30026744
PLCG2	16-81942175-A-G	rs75472618	AFR 10	28714976

Example 32 - Updated Drug Information

[0424] Immunosuppressive medicaments are a broad and expanding class of therapeutics, which includes generic versions of biologics (termed biosimilars). Drugs, such as those on the market that may be linked to PML risk, either now or in the future, include but are not limited to: diroximel fumarate (e.g. VUMERITY), siponimod (e.g. MAYZENT), golimumab (e.g. SIMPONI), elotuzumab (e.g. EMPLICITI), idebenone, ravulizimab-cwvz (e.g. (ULTOMIRIS), letrozole (e.g. FEMARA), voriconazole (e.g. VFEND), dalfampridine (e.g. AMPYRA), ambrisentan (e.g. LETAIRIS), pegfilgrastim (e.g. NEULASTA), simvastatin, adrenocorticotropic hormone, lipoic acid, and oxcarbazepine.

[0425] Immunosuppressive medicaments undergoing clinical trials that may linked to PML risk before or after drug approval include but are not limited to: elezanumab, nerispirdine, daprolizumab pegol, etrolizumab, abrilumab, evobrutinib, amiselimod, ceralifimod, clemastine, sunphenon epigallocatechin-gallate, and andrographolides (a botanical medicinal herb also known as IB-MS).

[0426] In addition, other molecules of interest include but are not limited to BNZ-1, IMU-838, SAR442168/PRN2246, BIIB033, BIIB059, BIIB061, AJM300, EK-12 (a neuropeptide combination of metenkefalin and tridecactide), and CHS-131.

[0427] Some drugs of interest as immunosuppressive medicaments have alternate names reported in FDA databases, such as fingolimod vs. fingolimod hydrochloride, mycophenolate mofetil vs. mycophenolate mofetil hydrochloride, tacrolimus vs. tacrolimus anhydrous, bendamustine vs. bendamustine hydrochloride, doxorubicin vs. doxorubicin hydrochloride, and certolizumab vs. certolizumab pegol.

[0428] Some HIV treatments have also been linked to cases of PML in the FDA database, such as ritonavir (e.g. NORVIR), enfuvirtida (e.g. FUZEON), and maraviroc (e.g. SELZENTRY).

[0429] Approved medicaments and/or molecules of interest that have the same mechanism of action as natalizumab, and therefore may linked to PML, include but are not limited to levocabastine, TR-14035, and R-411.

Table 67: Examples of biosimilars

Orignal Brand Name	Generic Name	Biosimilar Name	Biosimilar Brand/Code Name
Rituxan	rituximab	rituximab-abbs	Truxima
Rituxan	rituximab	rituximab-pvvr	Ruxience
Rituxan	rituximab		Rimimyo
Rituxan	rituximab		Rixathon

Orignal Brand Name	Generic Name	Biosimilar Name	Biosimilar Brand/Code Name
Rituxan	rituximab		ABP 798
Rituxan	rituximab		RTXM83
Remicade	infliximab	infliximab-dyyb	Inflectra
Remicade	infliximab	infliximab-abda	Renflexis
Remicade	infliximab	infliximab-qbtx	Ixifi / Zessly
Remicade	infliximab		ABP 710
Remicade	infliximab		Remsima
Humira	adalimumab	adalimumab-atto	Amjevita
Humira	adalimumab	adalimumab-adaz	Hyrimoz
Humira	adalimumab	adalimumab-bwwd	Hadalmia
Humira	adalimumab	adalimumab-adbm	Cyltezo
Humira	adalimumab		Hulio
Humira	adalimumab		Imraldi
Humira	adalimumab		PF-06410293
Humira	adalimumab		CT-P17
Humira	adalimumab		CHS-1420
Humira	adalimumab		M923/BAX923
Humira	adalimumab		MYL-1401A
Humira	adalimumab		LBAL
Humira	adalimumab		MSB 11022 (Idacio / Kromeya)
Humira	adalimumab		GP2017
Humira	adalimumab		BCD-057
Humira	adalimumab		ONS-3010
Humira	adalimumab		AVT02
Enbrel	etanercept	etanercept-szzs	Erelzi
Enbrel	etanercept	etanercept-ykro	Eticovo
Enbrel	etanercept		CT-P05
Enbrel	etanercept		Benepali
Enbrel	etanercept		Altebrel
Avastin	bevacizumab	bevacizumab-awwb	Mvasi
Avastin	bevacizumab	bevacizumab-bvzr	Zirabev
Avastin	bevacizumab		AB8
Avastin	bevacizumab		CT-P16
Solaris	eculizumab		SB 12
Solaris	eculizumab		ABP 959
Neulasta	pegfilgrastim	pegfilgrastim-jmdb	Fulphila
Neulasta	pegfilgrastim	pegfilgrastim-cbqv	Udenyca
Neulasta	pegfilgrastim		Pelmeg
Neulasta	pegfilgrastim		PF-06881894

[0430] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. The following claims define the scope of the invention.

REFERENCES CITED IN THE DESCRIPTION

Cited references

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document.

Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US20170169919A [0001]
- US6207392B [0122]
- US4683195A [0153]
- US20060024711A [0167]
- US20060024678A [0167]
- US20060012793A [0167]
- US20060012784A [0167]
- US20050100932A [0167]
- WO9210092A [0172]
- WO9511995A [0172]
- US5424186A [0172]
- US5143854A [0174]
- US5510270A [0174]
- US5527681A [0174]
- US5384261A [0174]
- US5288514A [0174]
- US9304145W [0174]
- US6858394B [0174]
- US6429027B [0174]
- US5445934A [0174]
- US5700637A [0174]
- US5744305A [0174]
- US5945334A [0174]
- US6054270A [0174]
- US6300063B [0174]
- US6733977B [0174]
- US7364856B [0174]
- EP619321A [0174]
- EP373203A [0174]
- US5451663A [0175]
- WO9820019A [0175]
- US5858659A [0177]
- US5837832A [0177]
- US4897268A [0193]
- US5075109A [0193]
- US5928647A [0193]
- US5811128A [0193]
- US5820883A [0193]
- US5853763A [0193]
- US5814344A [0193]
- US5942252A [0193]
- US4789734A [0194]
- US4906474A [0195]
- US4925673A [0195]
- US3625214A [0195]
- US5023252A [0221]
- US4992445A [0221]
- US5001139A [0221]
- US5376359A [0235]
- US5190029A [0235]
- US5776434A [0235]
- US10240205B [0418] [0418] [0418] [0418] [0418] [0418]

Non-patent literature cited in the description

- **MILLS et al.**Understanding Progressive Multifocal Leukoencephalopathy Risk in Multiple Sclerosis Patients Treated with immunomodulatory Therapies: A Bird's eye view", A Bird's Eye ViewFront. Immunol., vol. 9, 138- [0001]
- CHEMICAL ABSTRACTS, 189261-10-7 [0032]
- **MELIS et al.**CNS Drugs., 2015, vol. 29, 10879-91 [0065]
- **MAAS et al.**J Neurol., 2016, vol. 263, 102004-21 [0065]
- **COLIN et al.**Fundam Clin Pharmacol., 2016, [0065]
- **CALABRESE L. H. et al.**Nat Rev Rheumatol., 2015, [0066]
- **PAVLOVIC et al.**Ther Adv Neurol Disord., 2015, vol. 8, 6255-73 [0073]
- **BOWEN et al.**Nat Rev Neurol., 2016, vol. 12, 11662-674 [0073]
- **PAVLOVIC**Ther Adv Neurol Disord., 2015, vol. 8, 6255-73 [0073]
- **ZHANG F. et al.**Nat. Genet., 2009, [0085]
- **HASTINGS P. J. et al.**PLoS Genetics, 2009, [0085]
- **ESTIVILLARMENGOL**PLoS Genetics, 2007, [0086]
- **MACDONALD JR et al.**Nucleic Acids Res., 2014, vol. 42, D986-92 [0086]
- **SMITH M. W. et al.**Am. J. Hum. Genet., 2004, vol. 74, 1001- [0093]
- **AUSUBEL, F. et al.**Current Protocols in Molecular BiologyJohn Wiley & Sons19980000 [0111]
- **KRAUS, M.AARONSON, S.**Methods Enzymol., 1991, vol. 200, 546-556 [0111]
- **KARLIN, S.ALTSCHUL, S.**Proc. Natl. Acad. Sci. USA, 1993, vol. 90, 5873-5877 [0113]
- **ALTSCHUL, S. et al.**Nucleic Acids Res., 1997, vol. 25, 3389-3402 [0113]
- **MYERSMILLER**CABIOS, 1989, [0113]
- **NIELSEN, P. et al.**Science, 1991, vol. 254, 1497-1500 [0114]
- Principles of Fluorescence SpectroscopyPlenum Pub Corp19990700 [0123]
- **RICHARD P. HOAGLAND**Molecular Probes Handbook [0123]
- **BENJAMINI Y.HOCHBERG Y.**J. Royal Statistical Society, 1995, vol. 57, 289- [0140]
- **KUTYAVIN et al.**Nucleic Acid Res., 2006, vol. 34, e128- [0158] [0248]
- **BIER, F.F. et al.**Adv Biochem Eng Biotechnol, 2008, vol. 109, 433-53 [0174]
- **HOHEISEL, J. D.**Nat Rev Genet, 2006, vol. 7, 200-10 [0174]
- **FAN, J. B. et al.**Methods Enzymol, 2006, vol. 410, 57-73 [0174]
- **RAQOUSSIS, J.ELVIDGE, G.**Expert Rev Mol Design, 2006, vol. 6, 145-52 [0174]
- **MOCKLER, T.C. et al.**Genomics, 2005, vol. 85, 1-15 [0174]
- **SNIJDERS et al.**Nat. Genetics, 2001, vol. 29, 263-264 [0174]
- **KLEIN et al.**Proc. Natl. Acad. Sci. USA, 1999, vol. 96, 4494-4499 [0174]
- **ALBERTSON et al.**Breast Cancer Research and Treatment, 2003, vol. 78, 289-298 [0174]
- BAC microarray based comparative genomic hybridization**SNIJDERS et al.**Bacterial Artificial Chromosomes: Methods and Protocols, Methods in Molecular BiologyHumana Press20020000 [0174]
- **JELCIC et al.**Science Translational Medicine, 2015, vol. 7, 306306-150 [0187]
- **RAY et al.**Science Translational Medicine, 2015, vol. 7, 306306-151 [0187]
- **ANSEL et al.**Pharmaceutical Dosage Forms and Drug Delivery SystemsLippencott Williams & Wilkins19990000 [0191]
- **ANSEL et al.**Pharmaceutical Dosage Forms and Drug Delivery SystemsLippencott, Williams, & Wilkins19990000 [0192]
- Liposomes**G. GREGORIADIS**Drug Carriers in Biology and MedicineAcademic Press197900002- [0194]
- **JEIN**TIPS, 1998, vol. 19, 155-157 [0195]
- **REMINGTON**The Science and Practice of PharmacyLippincott Williams & Wilkins [0197]
- **BANGHAM et al.**J. Mol. Biol., 1965, vol. 23, 238-252 [0207]
- **SZOKA et al.**Proc. Natl Acad. Sci. USA, 1978, vol. 75, 4194-4198 [0207]
- **KLAUSNER E.A. et al.**Pharm. Res., 2003, vol. 20, 1466-73 [0213]
- **HOFFMAN, A. et al.**Int. J. Pharm., 2004, vol. 11, 141-53 [0213]
- **STREUBEL, A. et al.**Expert Opin. Drug Deliver., vol. 3, 217-3 [0213]
- **CHAVANPATIL, M.D. et al.**Int. J. Pharm., 2006, [0213]
- Remington's Pharmaceutical SciencesMack Publishing Co. [0217]
- Dermatological Formulations: Percutaneous absorptionMarcel Dekker19830000 [0219]
- The Merck Manual of Diagnosis and TherapyMerck Research Laboratories20060000 [0264]
- **BENJAMIN LEWIN**Genes IXJones & Bartlett Publishing20070000 [0264]
- The Encyclopedia of Molecular BiologyBlackwell Science Ltd.19940000 [0264]
- Molecular Biology and Biotechnology: a Comprehensive Desk ReferenceVCH Publishers, Inc.19950000 [0264]
- **MANIATIS et al.**Molecular Cloning: A Laboratory ManualCold Spring Harbor Laboratory Press19820000 [0265]
- **SAMBROOK et al.**Molecular Cloning: A Laboratory ManualCold Spring Harbor Laboratory Press19890000 [0265]
- **DAVIS et al.**Basic Methods in Molecular BiologyElsevier Science Publishing, Inc.19860000 [0265]
- Methods in Enzymology: Guide to Molecular Cloning TechniquesAcademic Press Inc.19870000vol. 152, [0265]
- Current Protocols in Molecular Biology (CPMB)John Wiley and Sons, Inc. [0265]
- Current Protocols in Protein Science (CPPS)John Wiley and Sons, Inc [0265]
- Current Protocols in Immunology (CPI)John Wiley and Sons, Inc. [0265]
- Current Protocols in Cell Biology (CPCB)John Wiley and Sons, Inc. [0265]

- **R. IAN FRESHNEY**ulture of Animal Cells: A Manual of Basic TechniqueWiley-Liss20050000 [0265]
- Methods in Cell BiologyAcademic Press19980000vol. 57, [0265]
- **DEEKSHIGGINS**Statistical algorithms in Review Manager 5, Statistical Methods Group of The Cochrane Collaboration, 2010, [0286]
- **LEK et al.**Nature, 2016, vol. 536, 7616285-91 [0290]
- **HATCHWELL**Front Immunol., 2015, vol. 6, 216- [0321] [0321] [0321]
- **DURANDY et al.**Nat Rev Immunol., 2013, vol. 13, 7519-33 [0321]
- **MILNER et al.**Nat Rev Immunol., 2013, vol. 13, 9635-48 [0321]
- **PACIOLLA et al.**Genes Immun., 2015, vol. 16, 4239-46 [0321]
- **THIJSSEN et al.**Nat Commun., 2015, vol. 6, 7870- [0321]
- **CHINN et al.**Immunol Allergy Clin North Am., 2015, vol. 35, 4671-94 [0321]
- **ZHOU et al.**Nat Genet., 2015, vol. 48, 167-73 [0321]
- **NAVABI et al.**Allergy Asthma Clin Immunol., 2016, vol. 12, 27- [0321]
- **TSUJITA et al.**J Allergy Clin Immunol., 2016, [0321]
- **VAN DER KOLK et al.**Ann Clin Transl Neurol., 2016, vol. 3, 3226-32 [0321] [0321]
- **ZERBE et al.**Clin Infect Dis., 2016, vol. 62, 8986-94 [0321]
- **DEKEYSER M et al.**Open Forum Infect Diseases, 2016, [0351]
- **LEK M et al.**Nature, 2016, vol. 536, 7616285-91 [0377]
- **MACDONALD J et al.**Nucleic Acids Res., 2014, vol. 42, D986-92 [0377]
- **POPLIN et al.**Nat Biotechnol., 2018, vol. 36, 10983-987 [0415]
- **LIU et al.**Hum Mutat., 2016, vol. 37, 3235-41 [0415]
- **PICARD et al.**J Clin Immunol., 2018, vol. 38, 196-128 [0416]
- **BOUSFIHA et al.**J Clin Immunol, 2018, vol. 38, 1129-143 [0416]
- **BURGMAIER et al.**Sci Rep, 2019, vol. 9, 17919- [0417]
- **TONK et al.**Pharmacogenomics J, 2016, vol. 17, 4386-392 [0421]

PATENTKRAV

1. Immunsuppressivt lægemiddel til anvendelse i behandling af en tilstand hos et individ med behov for immunsuppressiv lægemiddelterapi,

hvor individet har en reduceret risiko for progressiv multifokal leukoencephalopati (PML) på grund af en infektion i hjernen med John Cunningham-virus (JCV),

hvor individets reducerede risiko skyldes fravær af en genomisk variation valgt fra gruppen bestående af: chr9:137779251, G>A og chr1:160769595, AG>A,

hvor den genomiske variations kromosomposition defineres i forhold til UCSC hg19.

2. Immunsuppressivt lægemiddel til anvendelse ifølge krav 1, hvor tilstanden er en cancer, et organtransplantation eller en autoimmun sygdom, eventuelt hvor:

(i) den autoimmune sygdom vælges fra gruppen bestående af Addisons sygdom, anti-NMDA-receptor-encephalitis, antisyntetasesyndrom, aplastisk anæmi, autoimmune anæmier, autoimmun hæmolytisk anæmi, autoimmun pancreatitis, Behcets sygdom, bulløse hudlidelser, cøliaki - sprue (glutenfølsom enteropati), kronisk træthedssyndrom, kronisk inflammatorisk demyeliniserende polyneuropati, kronisk lymfocytisk leukæmi, Crohns sygdom, dermatomyositis, Devics sygdom, erythroblastopeni, Evans syndrom, fokal segmentær glomerulosklerose, granulomatose med polyangiitis, Graves sygdom, Graves' oftalmopati, Guillain-Barres syndrom, Hashimoto-skjoldbruskkirtelbetændelse, idiopatisk thrombocytopenisk purpura (ITP), IgA-nefropati, IgA-medierede autoimmune sygdomme, IgG4-relateret sygdom, inflammatorisk tarmsygdom, juvenil idiopatisk arthritis, multipel sklerose, myasthenia gravis, myelom, ikke-Hodgkins lymfom, opsoclonus myoclonus-syndrom (OMS), pemphigoid, pemphigus, pemphigus vulgaris, perniciøs anæmi, polymyositis, psoriasis, erytocytaplasi, reaktiv arthritis, rheumatoid arthritis, sarcoidose, sclerodermi, Sjögrens syndrom, systemisk lupus erythematosus, trombocytopenisk purpura, trombotisk trombocytopenisk purpura, type I diabetes, ulcerøs colitis, vasculitis, vitiligo, og kombinationer deraf;

(ii) den autoimmune sygdom er multipel sklerose eller Crohns sygdom; eller

(iii) den autoimmune sygdom er en tilbagevendende form af multipel sklerose.

3. Immunsuppressivt lægemiddel til anvendelse ifølge et hvilket som helst foregående krav, hvor

(i) det immunsuppressive lægemiddel omfatter abatacept, abrilumab, acalabrutinib, adalimumab, adrenocorticotropt hormon, agatolimodnatriums, aldesleukin, alefacept, alemtuzumab, alisertib, alvespimycinhydrochlorid, alvocidib, ambrisentan, aminocamptothecin,

amiselimod, anakinra, andecaliximab, andrographolider, anifrolumab, antithymocyt Ig, apatinib, apelisib, asparaginase, atacicept, atezolizumab, avelumab, azacitidine, azathioprine, bafetinib, baminercept, baricitinib, basiliximab, becatecarin, begelomab, belatacept, belimumab, bemcentinib, bendamustin, betalutin med lilotomab, bevacizumab, bimekizumab, binimetinib, bleomycin, blinatumomab, BNZ-1, bortezomib, brentuximab vedotin, bryostatin 1, bucillamin, buparlisib, busulfan, canakinumab, capecitabin, carboplatin, carfilzomib, carmustin, cediranibmaleat, cemiplimab, ceralifimod, cerdulatinib, certolizumab, cetuximab, chidamid, chlorambucil, cilengitid, cirmtuzumab, cisplatin, cladribin, clazakizumab, clemastin, clioquinol, corticosteroider, cyclophosphamid, cyclosporin, cytarabin, cytotoksisk kemoterapi, daclizumab, dalfampridin, daprolizumab-pegol, daratumumab, dasatinib, defactinib, defibrotid, denosumab, dexamethason, diacerein, dimethylfumarat, dinaciclib, diroximelfumarat, doxorubicin, durvalumab, duvelisib, duvortuxizumab, eculizumab, efalizumab, eftilagimod alpha, en neuropeptidkombination af metenkefalin og tridecactid, elezanumab, elotuzumab, encorafenib, entinostat, entospletinib, enzastaurin, epacadostat, epirubicin, epratuzumab, eritoran-tetranatrium, etanercept, etoposid, etrolizumab, everolimus, evobrutinib, filgotinib, fingolimod, firategrast, fludarabin, fluoruracil, fontolizumab, forodesinhydrochlorid, fostamatinib, galunisertib, ganetespib, ganitumab, gemcitabin, gemtuzumab-ozogamicin, gerilimzumab, glasdegib, glassia, glatirameracetat, glembatumumab-vedotin, glesatinib, golimumab, guadecitabin, hydrocortison, hydroxychloroquinsulfat, hydroxyurea, ibritumomab-tiuxetan, ibrutinib, ibudilast, idarubicin, idebenon, idelalisib, ifosfamid, iguratimod, imatinib, imexon, infliximab, inotuzumab-ozogamicin, interferon alfa-2, interferon beta-1a, interferon beta-1b, interferon gamma-1, ipilimumab, irofulven, isatuximab, ispinesib, itacitinib, ixazomib, lapatinib, laquinimod, laromustin, ld-aminopterin, leflunomid, lenalidomid, lenvatinib, letrozol, levamisol, levocabastin, liponsyre, lirilumab, lonafamib, lumiliximab, masitinib, mavrilimumab, melphalan, mercaptopurin, methotrexat, methoxsalen, methylprednison, milatuzumab, mitoxantron, mizoribin, mocetinostat, monalizumab, mosunetuzumab, motesanibdiphosphat, moxetumomab-pasudotox, muromonab-CD3, mycophenolatmofetil, mycophenolsyre, namilumab, natalizumab, navitoclax, neihulizumab, nerispirdin, neurovax, niraparib, nivolumab, obatoclaxmesylat, obinutuzumab, oblimersen-natrium, ocrelizumab, ofatumumab, olokizumab, opicinumab, oprelvekin, osimertinib, otelixizumab, oxaliplatin, oxcarbazepin, ozanimod, paclitaxel, pacritinib, palifermin, panobinostat, pazopanib, peficitinib, pegfilgrastim, peginterferon beta-1a, pegsunercept (peg stnf-ri), pembrolizumab, pemetrexed, penclomedin, pentostatin, perifosin, pevonedistat, pexidartinib, picoplatin, pidilizumab, pivanex, pixantron, pleneva, plovameracetat, polatuzumab-vedotin, pomalidomid, ponatinib, ponesimod,

prednisone/prednisolon, pyroxamid, ravulizimab-cwvz, recombinant il-12, relatlimab, rhigf-1, rhigm22, rigosertib, rilonacept, rituximab, ruxolitinib, sarilumab, secukinumab, selumetinib, simvastatin, sintilimab, siplizumab, siponimod, sirolimus, sirukumab, sitravatinib, sonidegib, sorafenib, sotrastaurinacetat, sunitinib, sunphenon- epigallocatechin-gallat, tabalumab, tacrolimus, talabostatmesylat, talacotuzumab, tanespimycin, tegafur/gimeracil/oteracil, temozolomid, temsirolimus, tenalisib, terameprocol, teriflunomid, thalidomid, thiarabin, thiotepa, tipifarnib, tirabrutinib, tislelizumab, tivozanib, tocilizumab, tofacitinib, tregalizumab, tremelimumab, treosulfan, ublituximab, umbralisib, upadacitinib, urelumab, ustekinumab, varlilumab, vatelizumab, vedolizumab, veliparib, veltuzumab, venetoclax, vinblastin, vincristin, vinorelbinditartrat, visilizumab, vismodegib, vistusertib, voriconazol, vorinostat, vosaroxin, ziv-aflibercept, 6,8-bis(benzylthio)octansyre, eller en hvilken som helst kombination deraf;

(ii) det immunsuppressive lægemiddel omfatter natalizumab, interferon beta-1a, interferon beta-1b, glatirameracetat, peginterferon beta-1a, teriflunomid, fingolimod, dimethylfumarat, alemtuzumab, mitoxantron, rituximab, daclizumab, ocrelizumab, diroximelfumarat eller siponimod eller en hvilken som helst kombination deraf; eller

(iii) det immunsuppressive lægemiddel omfatter adalimumab, alemtuzumab, alemtuzumab, azathioprin, belimumab, bevacizumab, bortezomib, eculizumab, leflunomid, brentuximab vedotin, cetuximab, cyclophosphamid, dimethylfumarat, efalizumab, fingolimod, fludarabin, fumarsyre, imatinib, infliximab, methotrexat, mycophenolatmofetil, natalizumab, daclizumab, rituximab, vedolizumab, ruxolitinib eller ocrelizumab.

4. Immunsuppressivt lægemiddel til anvendelse ifølge et hvilket som helst foregående krav, hvor det immunsuppressive lægemiddel omfatter natalizumab, eventuelt hvor:

(i) natalizumab administreres ved intravenøs infusion;

(ii) ca. 100 mg til ca. 500 mg natalizumab administreres;

(iii) ca. 100 mg til ca. 500 mg natalizumab administreres hver fjerde uge;

(iv) ca. 100 mg til ca. 500 mg natalizumab administreres ved intravenøs infusion hver fjerde uge;

(v) ca. 100 mg til ca. 500 mg natalizumab administreres ved intravenøs infusion hver sjette uge;

(vi) ca. 100 mg til ca. 500 mg natalizumab administreres ved intravenøs infusion hver ottende uge; eller

(vii) natalizumab administreres subkutant.

5. Immunsuppressivt lægemiddel til anvendelse ifølge et hvilket som helst foregående krav, hvor individets genomiske sekvenser bestemmes ved en genetisk test, der omfatter påvisning af den genomiske variation i en polynukleinsyreprøve opnået fra individet, eventuelt hvor:

(i) den genetiske test omfatter analyse af et helgenom eller helexomet fra individet;

(ii) den genetiske test omfatter analyse af nukleinsyreinformationer, der allerede er opnået for et helgenom eller et helexom fra individet;

(iii) den genetiske test omfatter analyse af nukleinsyreinformationer der allerede er opnået for et helgenom eller et helexom fra individet, hvor nukleinsyreinformationerne opnås fra en analyse *in silico*;

(iv) individet er et menneske; eller

(v) polynukleinsyreprøven omfatter en polynukleinsyre fra blod, spyt, urin, serum, tårer, hud, væv eller hår fra individet.

6. Immunsuppressivt lægemiddel til anvendelse ifølge krav 5, hvor den genetiske test omfatter microarray-analyse, PCR, sekvensering, nukleinsyrehybridisering, eller en hvilken som helst kombination deraf, eventuelt hvor:

(i) microarray-analysen vælges fra gruppen bestående af en komparativ genomisk hydridiserings- (CGH) array-analyse og en SNP-array-analyse, eller

(ii) sekvenseringen vælges fra gruppen bestående af massiv parallel signatursekvensering (MPSS), polony-sekvensering, 454 pyrosekvensering, Illumina-sekvensering, Illumina (Solexa)-sekvensering ved hjælp af 10X Genomics-biblioteksprøparat, SOLiD-sekvensering, ionhalvleder-sekvensering, DNA-nanokuglesekvensering, heliskop-enkeltmolekyle-sekvensering, enkeltmolekyle-realtids-sekvensering (SMRT), RNAP-sekvensering, nanopore DNA-sekvensering, sekvensering ved hydridisering og mikrofluidisk Sanger-sekvensering.

7. Immunsuppressivt lægemiddel til anvendelse ifølge et hvilket som helst foregående krav, hvor den genomiske variation er en heterozygot SNV eller en homozygot SNV.

8. Immunsuppressivt lægemiddel til anvendelse ifølge et hvilket som helst foregående krav, hvor:

(a) det immunsuppressive lægemiddel til anvendelse endvidere omfatter et middel, der reducerer en JCV-virusbelastning hos individet; eller

(b) det immunsuppressive lægemiddel er beregnet til administration før eller i forbindelse med et middel, der reducerer en virusbelastning hos individet,

eventuelt hvor:

(i) det immunsuppressive lægemiddel administreres efter reduktion af virusbelastningen; eller

(ii) midlet, der reduceres virusbelastningen, er et middel målrettet JCV.

9. Fremgangsmåde til reduktion af en risiko for, at et individ udvikler progressiv multifokal leukoencephalopati (PML), omfattende

(a) test af et individ for tilstedeværelsen af en genomisk variation valgt fra gruppen bestående af: chr9:137779251, G>A og chr1:160769595, AG>A, hvor den genomiske variations kromosomposition defineres i forhold til UCSC hg19;

(b) bestemmelse af, at individet har den genomiske variation og

(c) frarådelse af administration af immunsuppressivt lægemiddel til individet, der blev bestemt til at have den genomiske variation, hvor det immunsuppressive lægemiddel omfatter natalizumab, interferon beta-1a, interferon beta-1b, glatirameracetat, peginterferon beta-1a, teriflunomid, fingolimod, dimethylfumarat, alemtuzumab, mitoxantron, rituximab, daclizumab, ocrelizumab, diroximelfumarat eller siponimod,

eventuelt hvor:

(i) rådgivningen omfatter rådgivning om, at administration af immunsuppressivt lægemiddel kontraindiceres;

(ii) rådgivningen omfatter rådgivning om, at administration af immunsuppressivt lægemiddel øger individets risiko for at udvikle multifokal leukoencephalopati (PML);

(iii) rådgivningen omfatter rådgivning om, at administration af immunsuppressivt lægemiddel er en faktor, der øger individets risiko for at udvikle progressiv multifokal leukoencephalopati (PML); eller

(iv) hvor det immunsuppressive lægemiddel omfatter natalizumab.

10. Fremgangsmåde til identificering af et individ som havende en risiko for at udvikle progressiv multifokal leukoencephalopati (PML) omfattende:

(a) analyse af en polynukleinsyreprøve fra individet for en genomisk variation valgt fra en gruppe bestående af: chr9:137779251, G>A og chr1:160769595, AG>A, hvor den genomiske variations kromosomposition defineres i forhold til UCSC hg19;

(b) konstatering af tilstedeværelsen af den genetisk variation i polynukleinsyreprøven og

(c) identificering af individet som havende en høj risiko for at udvikle PML;

hvor individet er immunsupprimeret,

eventuelt hvor individet har HIV, har fået en organtransplantation eller lider af en tilstand valgt blandt en cancer, en hæmatologisk malignitet, en autoimmun sygdom eller idiopatisk CD4+ lymfocytopeni (ICL).

11. Immunsuppressiv lægemiddel til anvendelse ifølge et hvilket som helst af krav 1 til 8, eller fremgangsmåde ifølge krav 9 eller 10, og som endvidere omfatter analyse for tilstedeværelse af JCV i en biologisk prøve fra individet, hvor analysen for tilstedeværelse af JCV omfatter en JCV-antistoftest eller en CSF-IgM-oligoklonalbåndtest,

eventuelt hvor:

analysen for tilstedeværelse af JCV omfatter test af individet med JCV-antistoftesten, hvor JCV-antistoftesten påviser eller ikke påviser en tilstedeværelse af JCV, eller

JCV-antistoftesten omfatter at bringe et JCV-påvisningsreagens i kontakt med en biologisk prøve fra individet, eventuelt hvor JCV-påvisningsreagenset vælges fra gruppen bestående af et anti-JCV-antistof, en specifik JCV-primer og kombinationer deraf.

DRAWINGS

Drawing

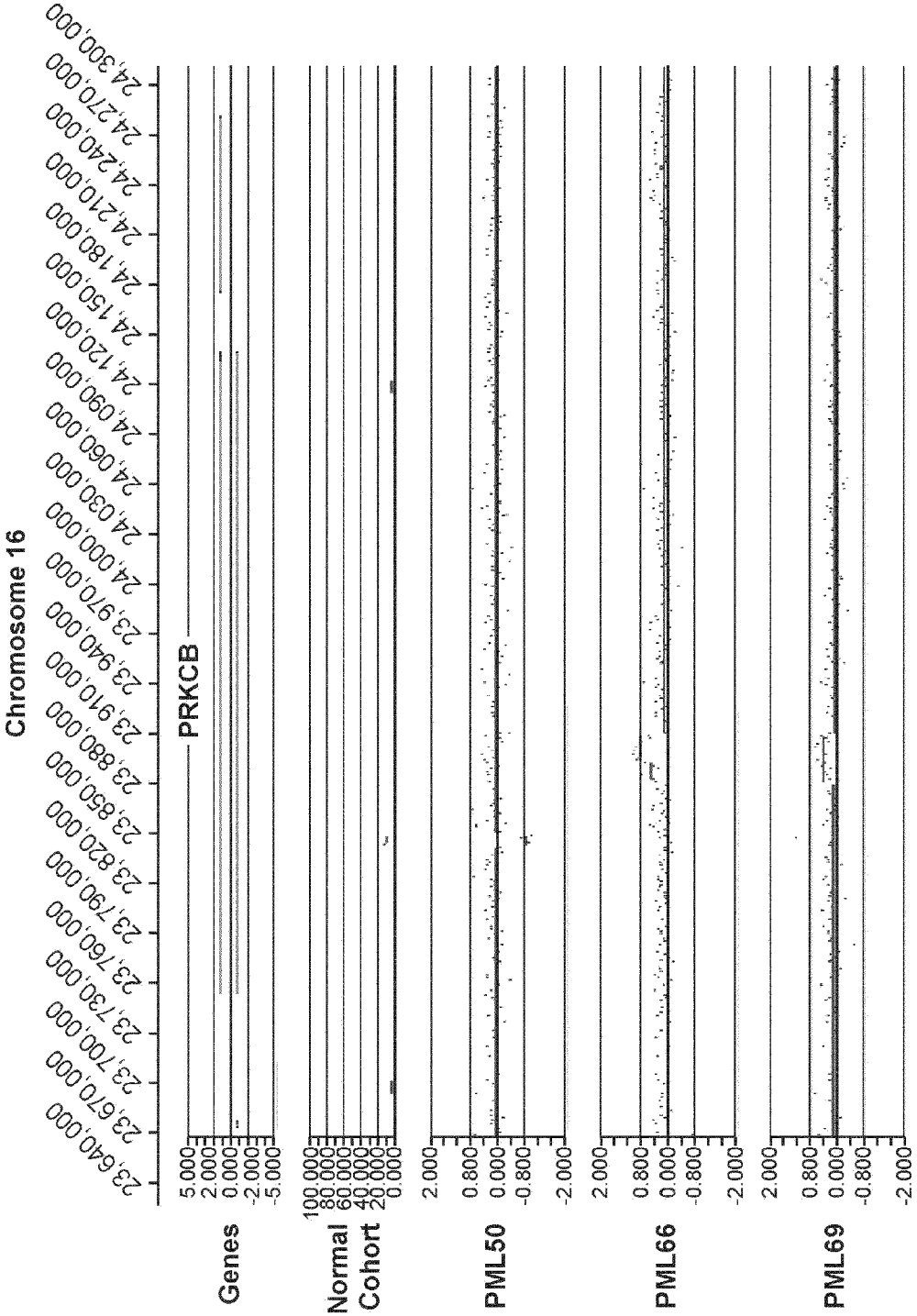


FIG. 1

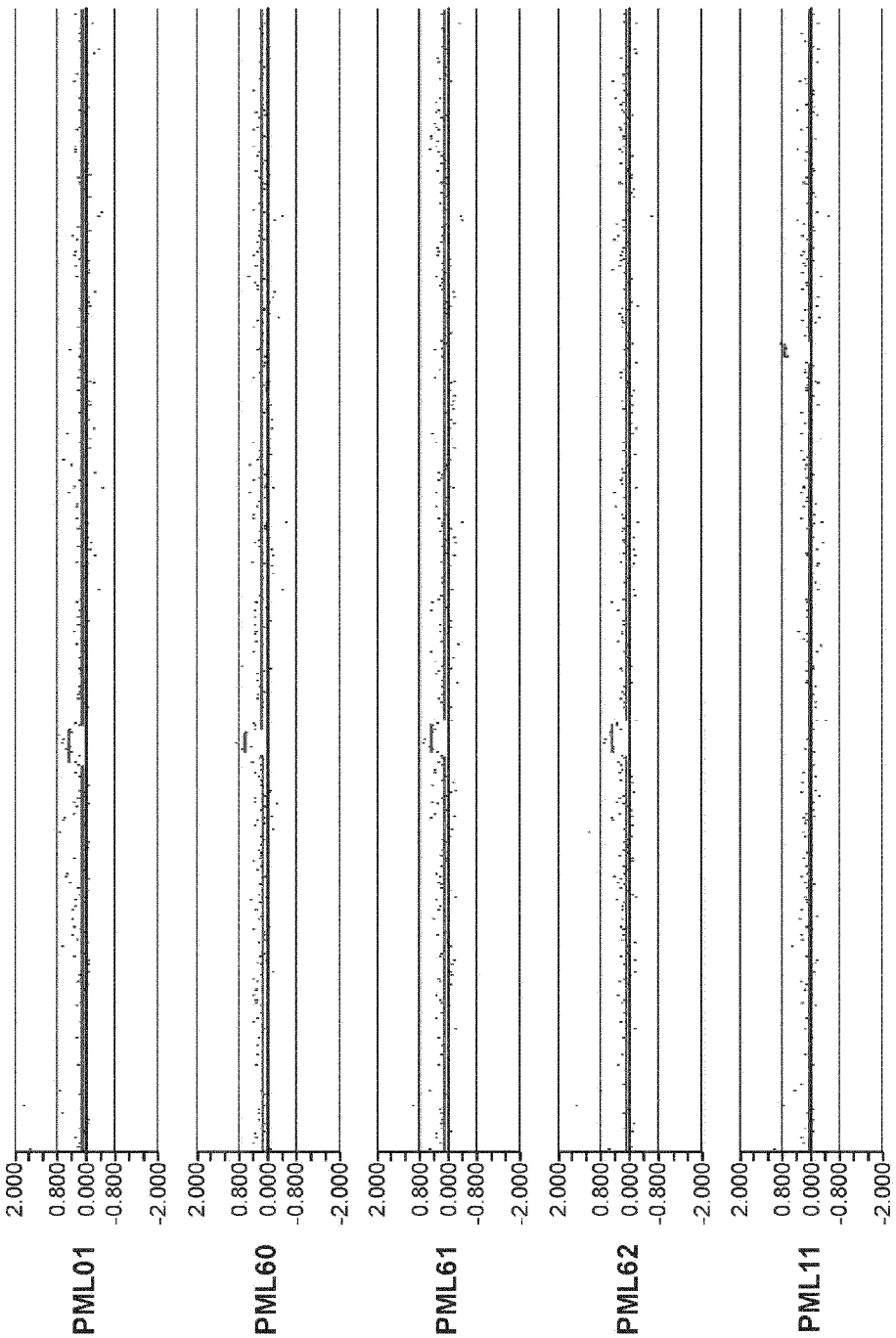


FIG. 1 (Cont.)

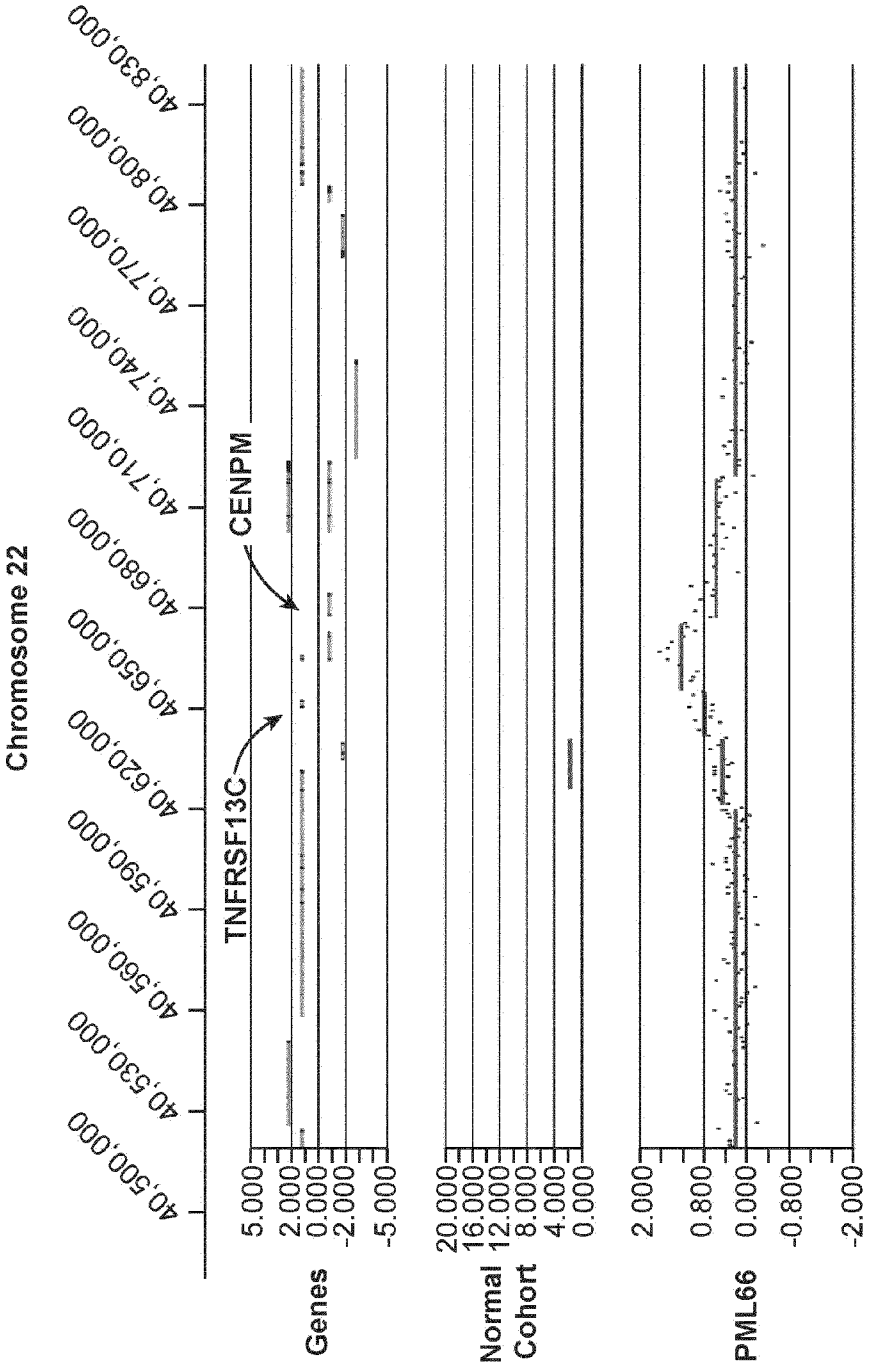

Chromosome 22
40,500,000
40,530,000
40,560,000
40,590,000
40,620,000
40,650,000
40,680,000
40,710,000
40,740,000
40,770,000
40,800,000
40,830,000
TNFRSF13C
CENPM
Genes
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
20.000
16.000
12.000
8.000
4.000
0.000
PML66
2.000
0.800
0.000
-0.800
-2.000

FIG. 2 (Cont. 1)

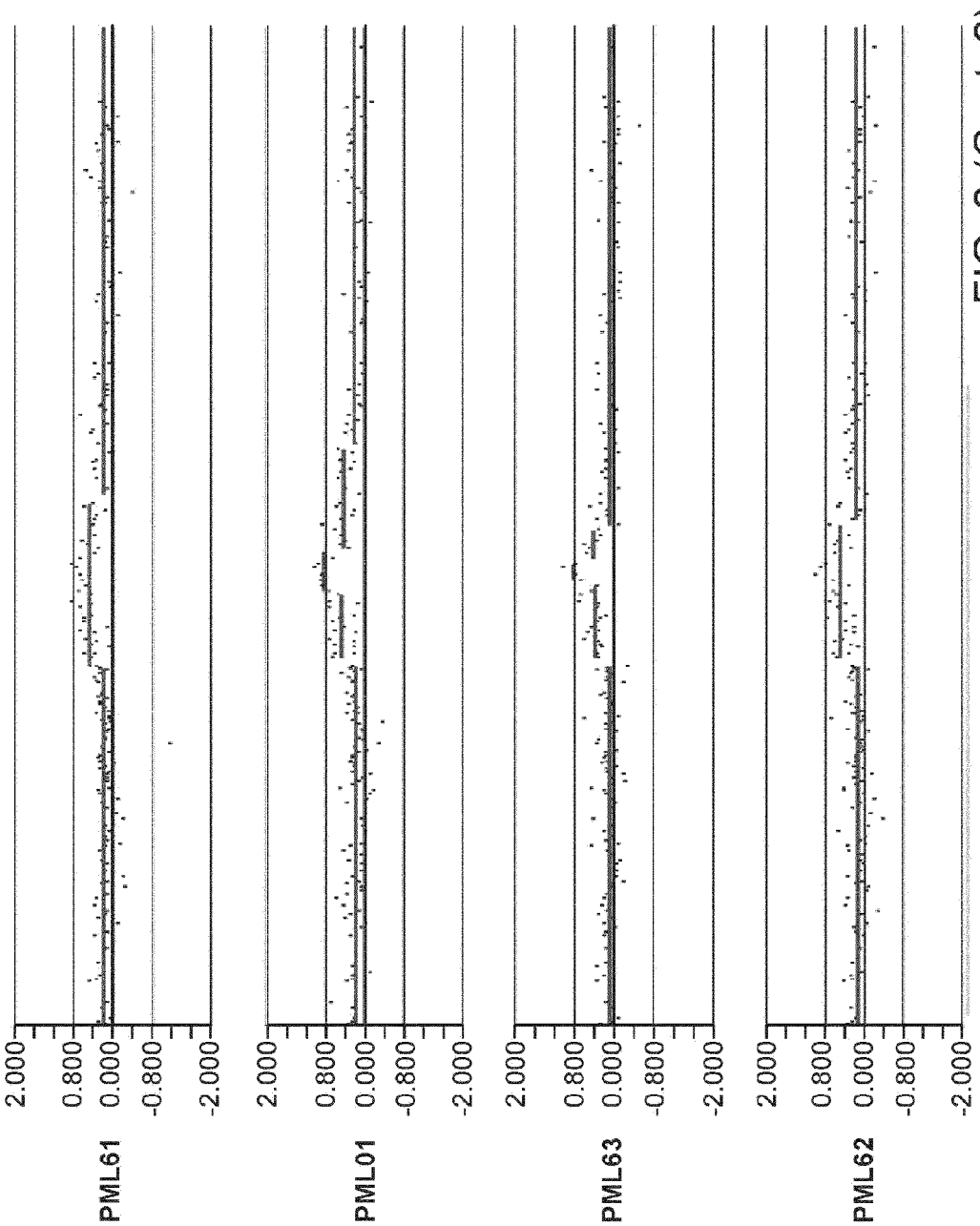


FIG. 2 (Cont. 2)

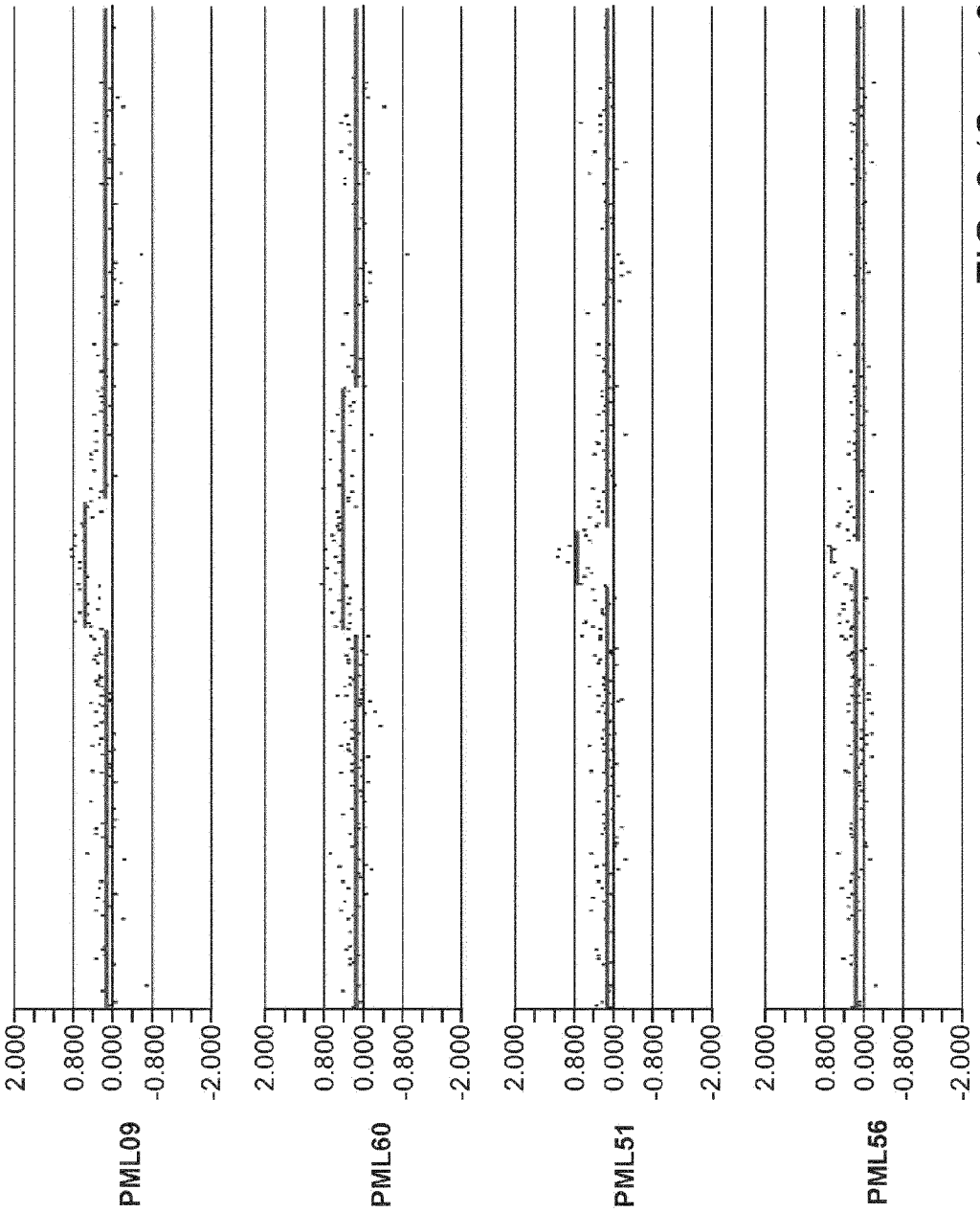


FIG. 2 (Cont. 3)

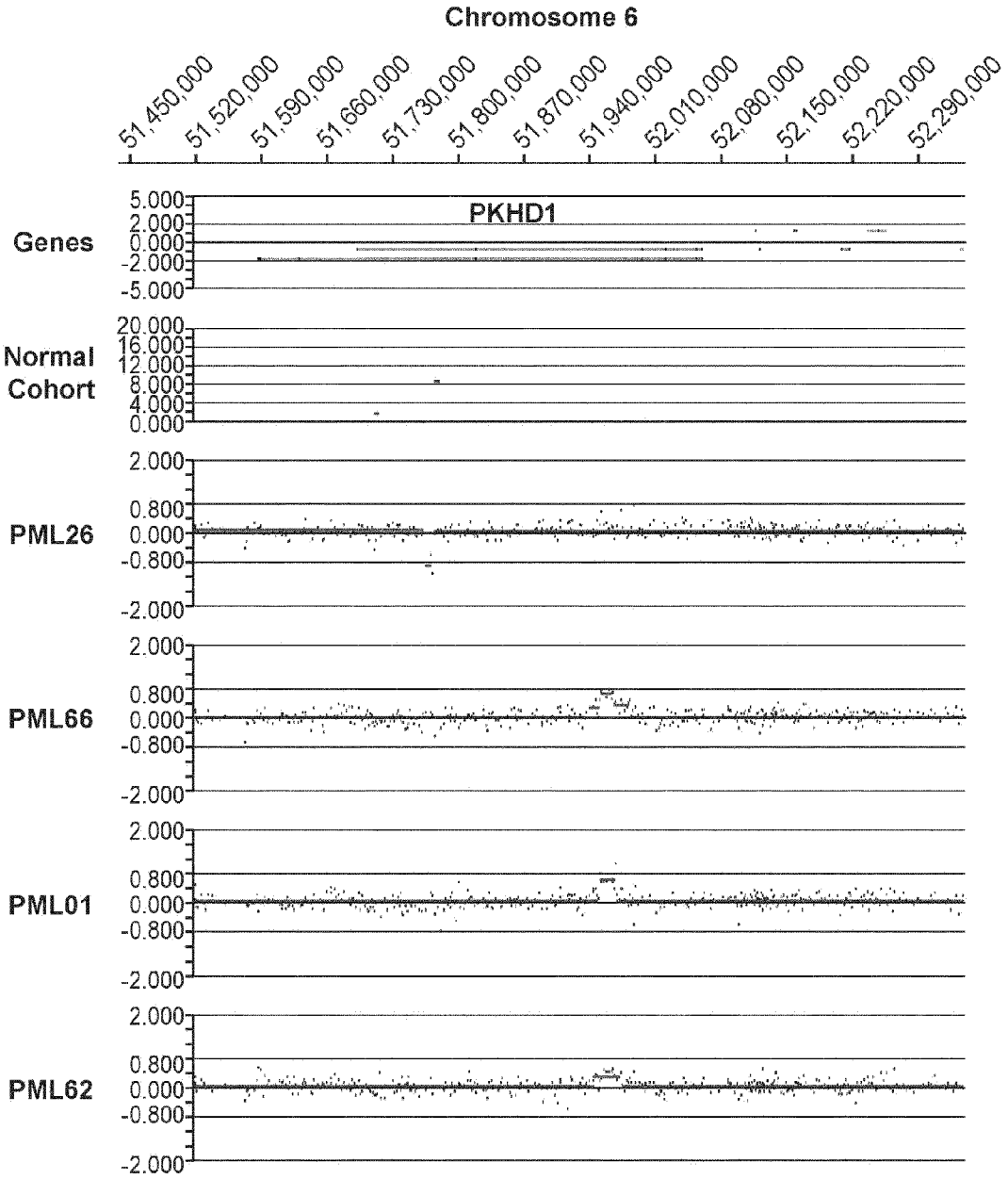

Chromosome 6
51,450,000
51,520,000
51,590,000
51,660,000
51,730,000
51,800,000
51,870,000
51,940,000
52,010,000
52,080,000
52,150,000
52,220,000
52,290,000
Genes
PKHD1
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
20.000
16.000
12.000
8.000
4.000
0.000
PML26
PML66
PML01
PML62
2.000
0.800
0.000
-0.800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

FIG. 3

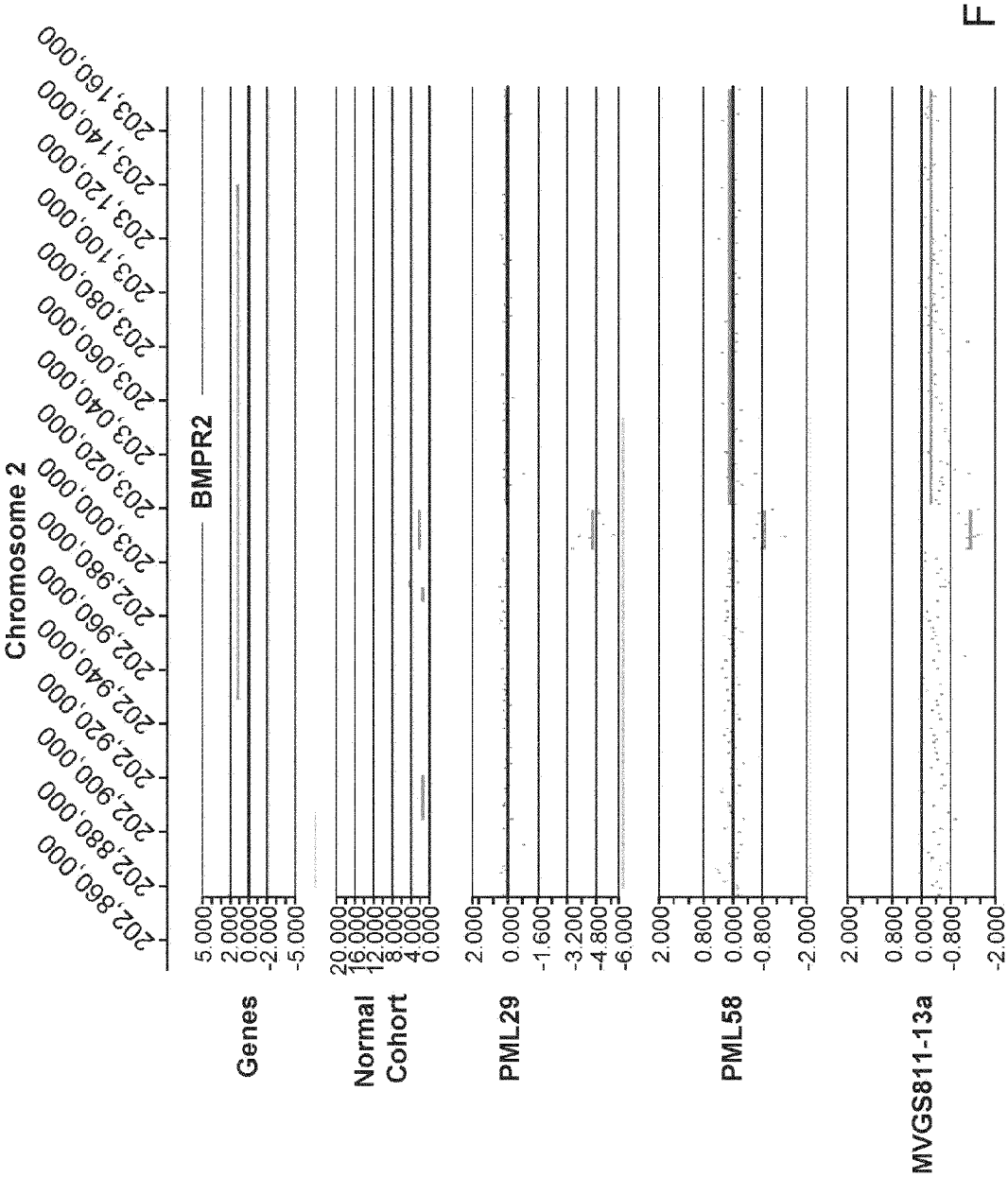

Chromosome 2
202,860,000
202,880,000
202,900,000
202,920,000
202,940,000
202,960,000
202,980,000
203,000,000
203,020,000
203,040,000
203,060,000
203,080,000
203,100,000
203,120,000
203,140,000
203,160,000
Genes
BMPR2
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
20.000
16.000
12.000
8.000
4.000
0.000
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

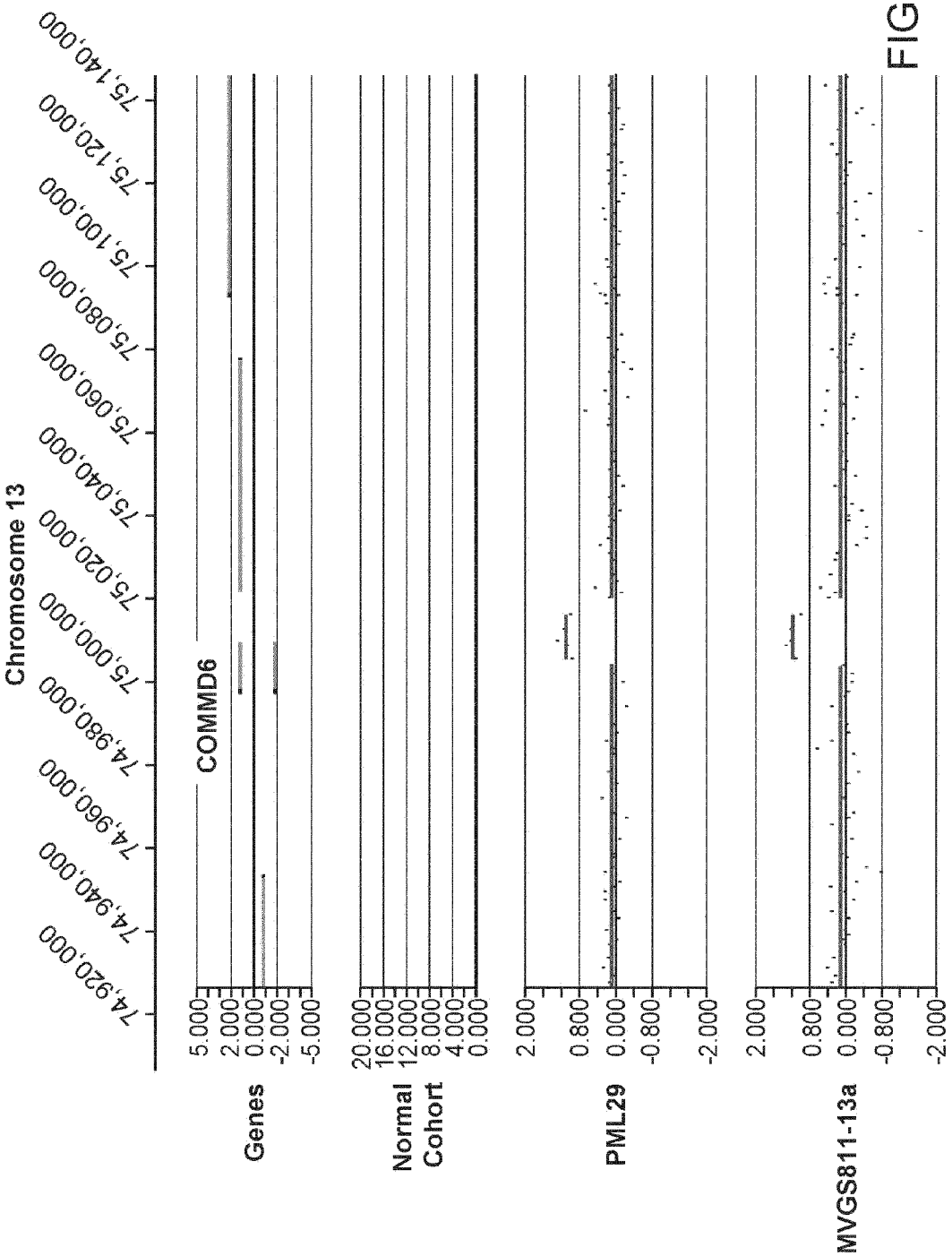

Chromosome 13
74,920,000
74,940,000
74,960,000
74,980,000
75,000,000
75,020,000
75,040,000
75,060,000
75,080,000
75,100,000
75,120,000
75,140,000
Genes
COMMD6
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
20.000
16.000
12.000
8.000
4.000
0.000
PML29
2.000
0.800
0.000
-0.800
-2.000
MVGS811-13a
2.000
0.800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

FIG. 5

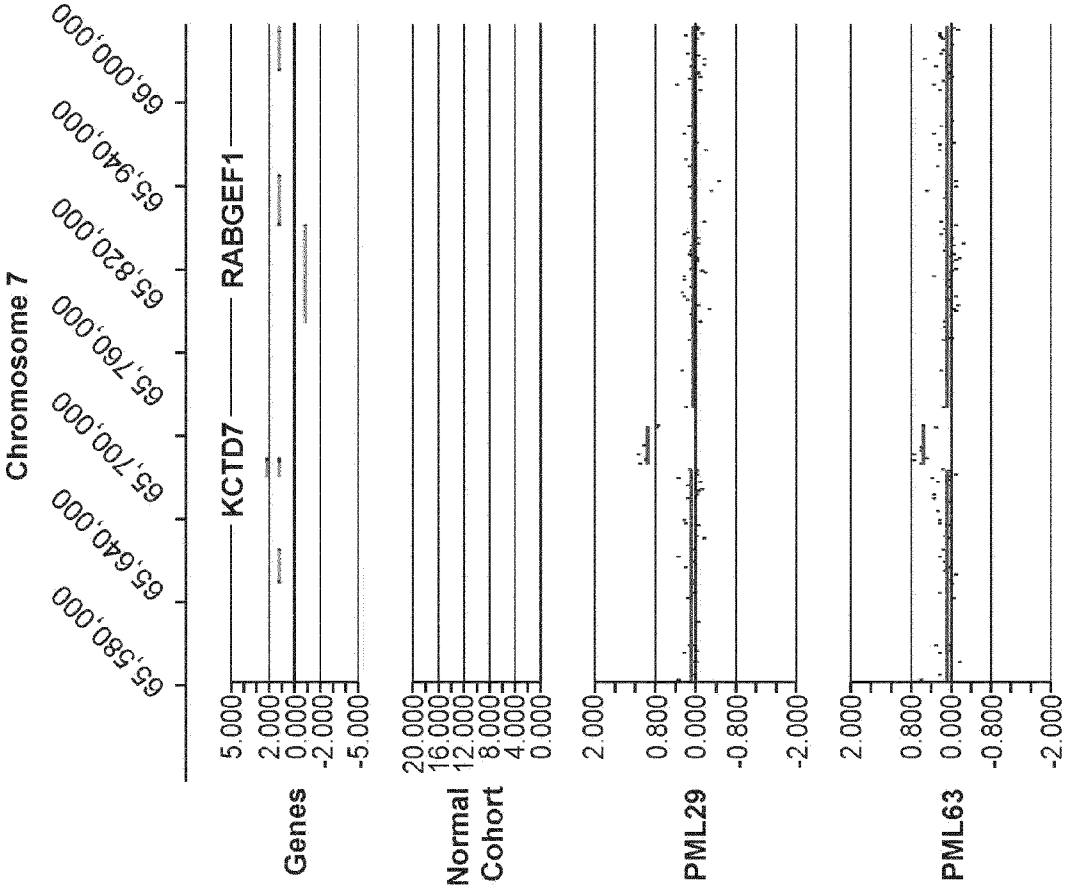

Chromosome 7
65,580,000
65,640,000
65,700,000
65,760,000
65,820,000
65,940,000
66,000,000
KCTD7
RABGEF1
Genes
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
20.000
16.000
12.000
8.000
4.000
0.000
PML29
2.000
0.800
0.000
-0.800
-2.000
PML63
2.000
0.800
0.000
-0.800
-2.000

FIG. 6

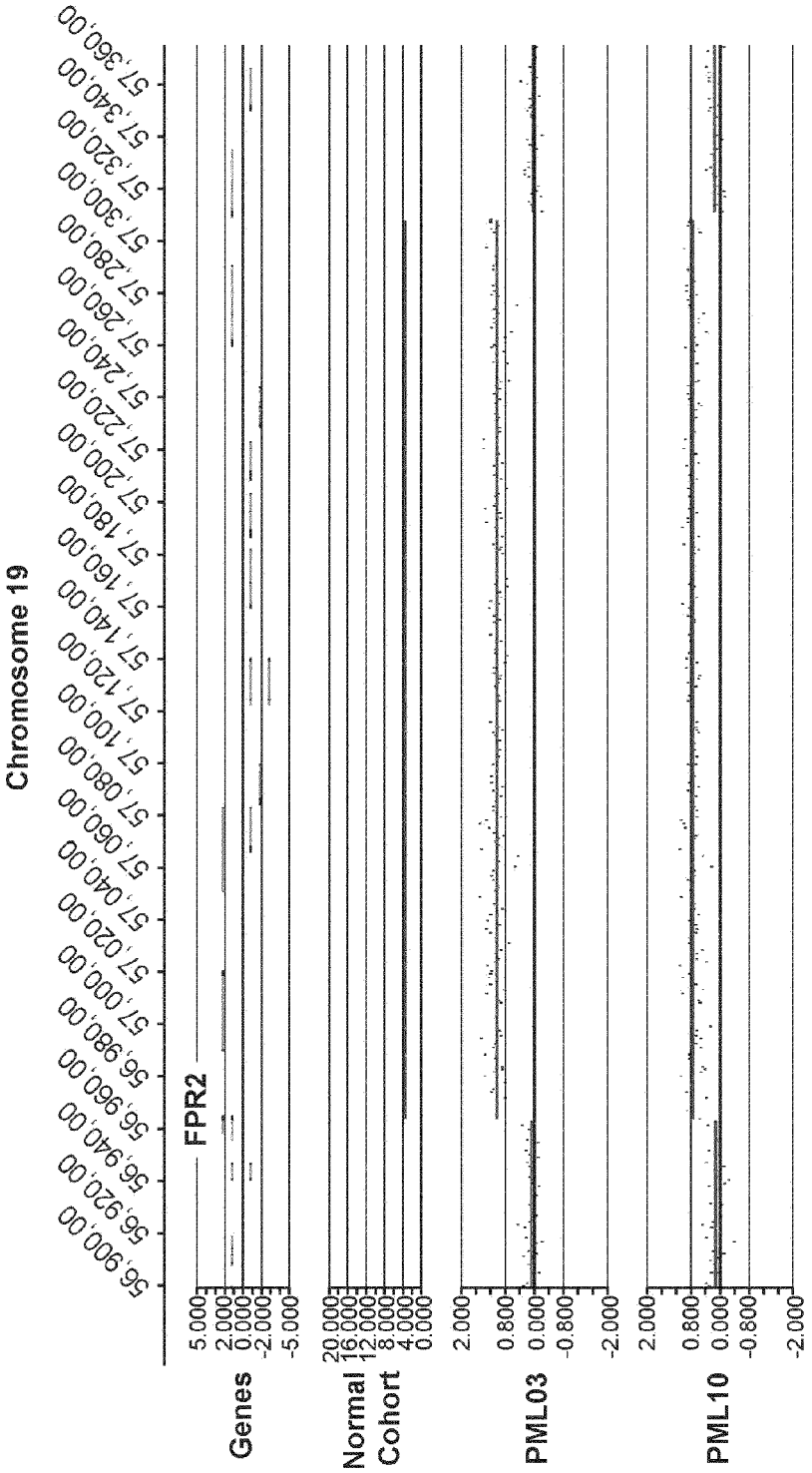

Chromosome 19
56,900,00
56,920,00
56,940,00
56,960,00
56,980,00
57,000,00
57,020,00
57,040,00
57,060,00
57,080,00
57,100,00
57,120,00
57,140,00
57,160,00
57,180,00
57,200,00
57,220,00
57,240,00
57,260,00
57,280,00
57,300,00
57,320,00
57,340,00
57,360,00
FPR2
Genes
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
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

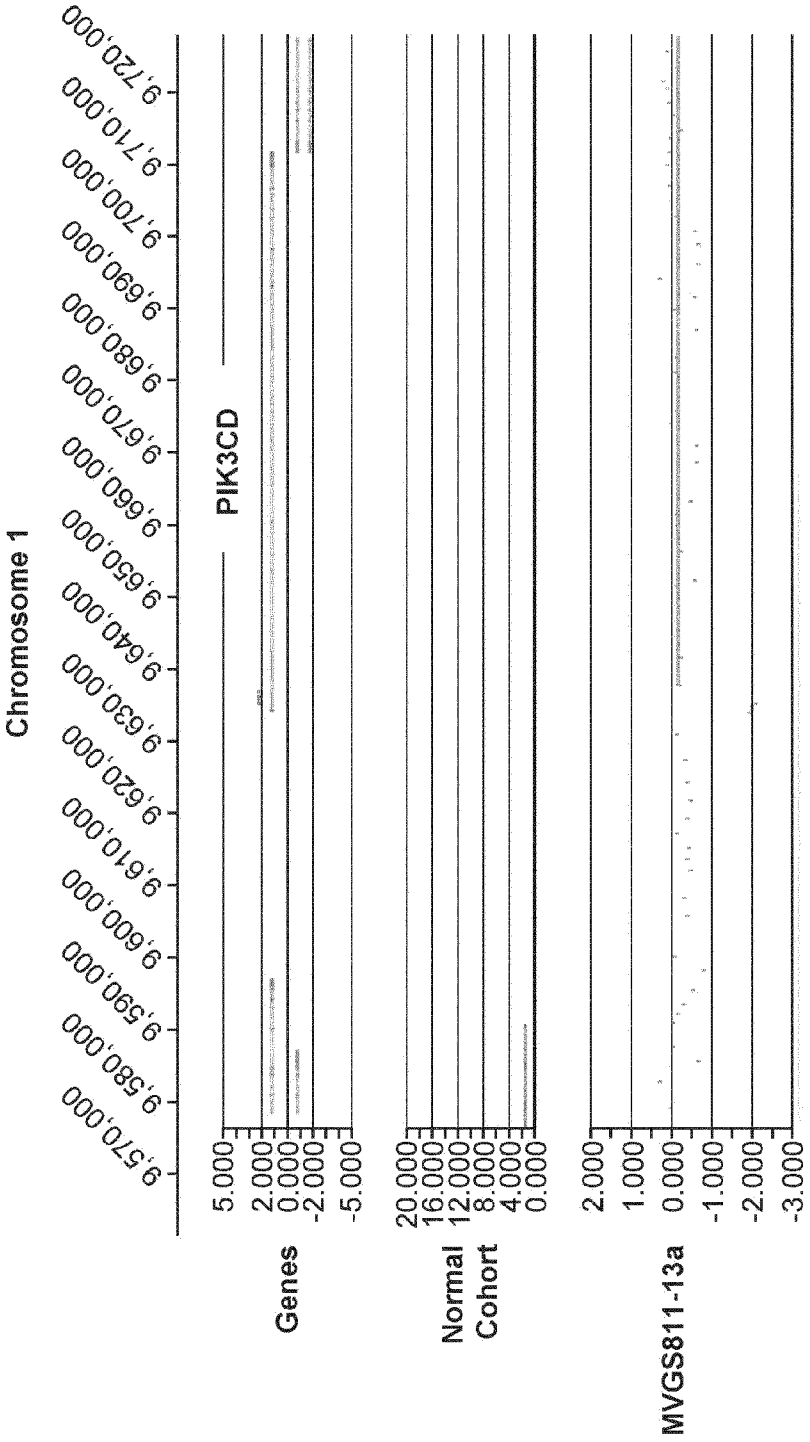

Chromosome 1
9,570,000
9,580,000
9,590,000
9,600,000
9,610,000
9,620,000
9,630,000
9,640,000
9,650,000
9,660,000
9,670,000
9,680,000
9,690,000
9,700,000
9,710,000
9,720,000
PIK3CD
Genes
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
20.000
16.000
12.000
8.000
4.000
0.000
MVGS811-13a
2.000
1.000
0.000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

FIG. 8

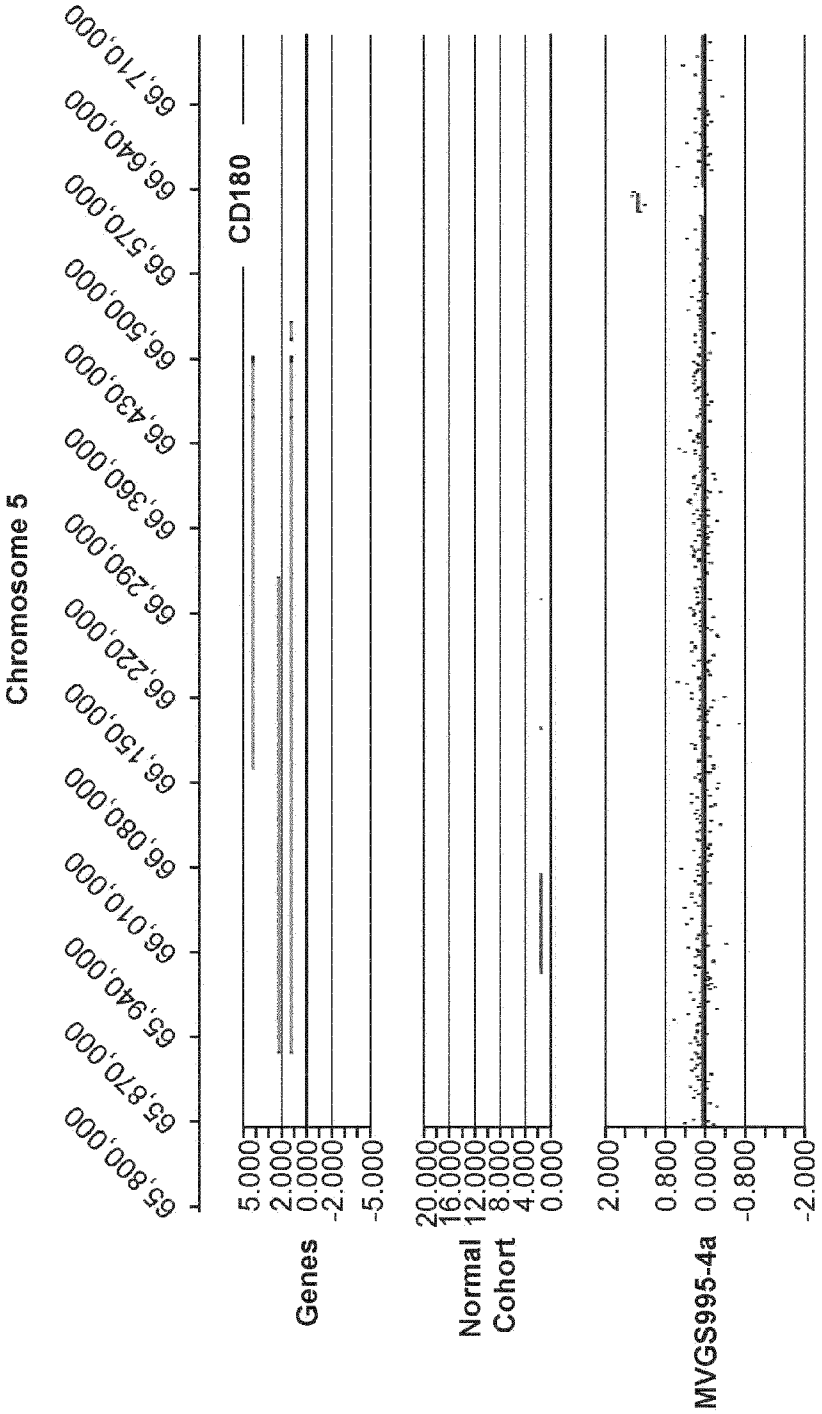

Chromosome 5
65,800,000
65,870,000
65,940,000
66,010,000
66,080,000
66,150,000
66,220,000
66,290,000
66,360,000
66,430,000
66,500,000
66,570,000
66,640,000
66,710,000
CD180
Genes
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
20.000
16.000
12.000
8.000
4.000
0.000
MVGS995-4a
2.000
0.800
0.000
-0.800
-2.000

FIG. 9

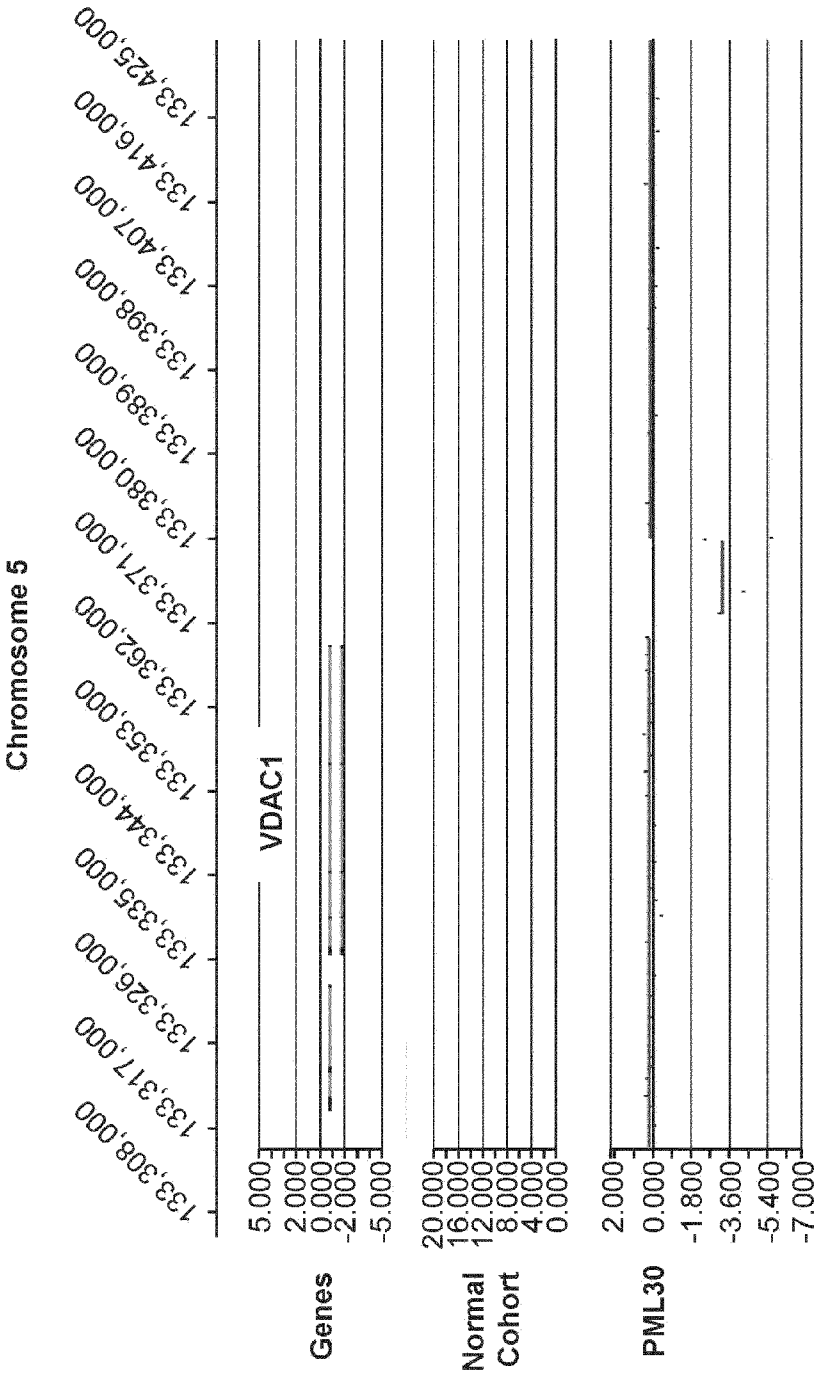

Chromosome 5
133,308,000
133,317,000
133,326,000
133,335,000
133,344,000
133,353,000
133,362,000
133,371,000
133,380,000
133,389,000
133,398,000
133,407,000
133,416,000
133,425,000
Genes
VDAC1
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
20.000
16.000
12.000
8.000
4.000
0.000
PML30
2.000
0.000
-1.800
-3.600
-5.400
-7.000

FIG. 10

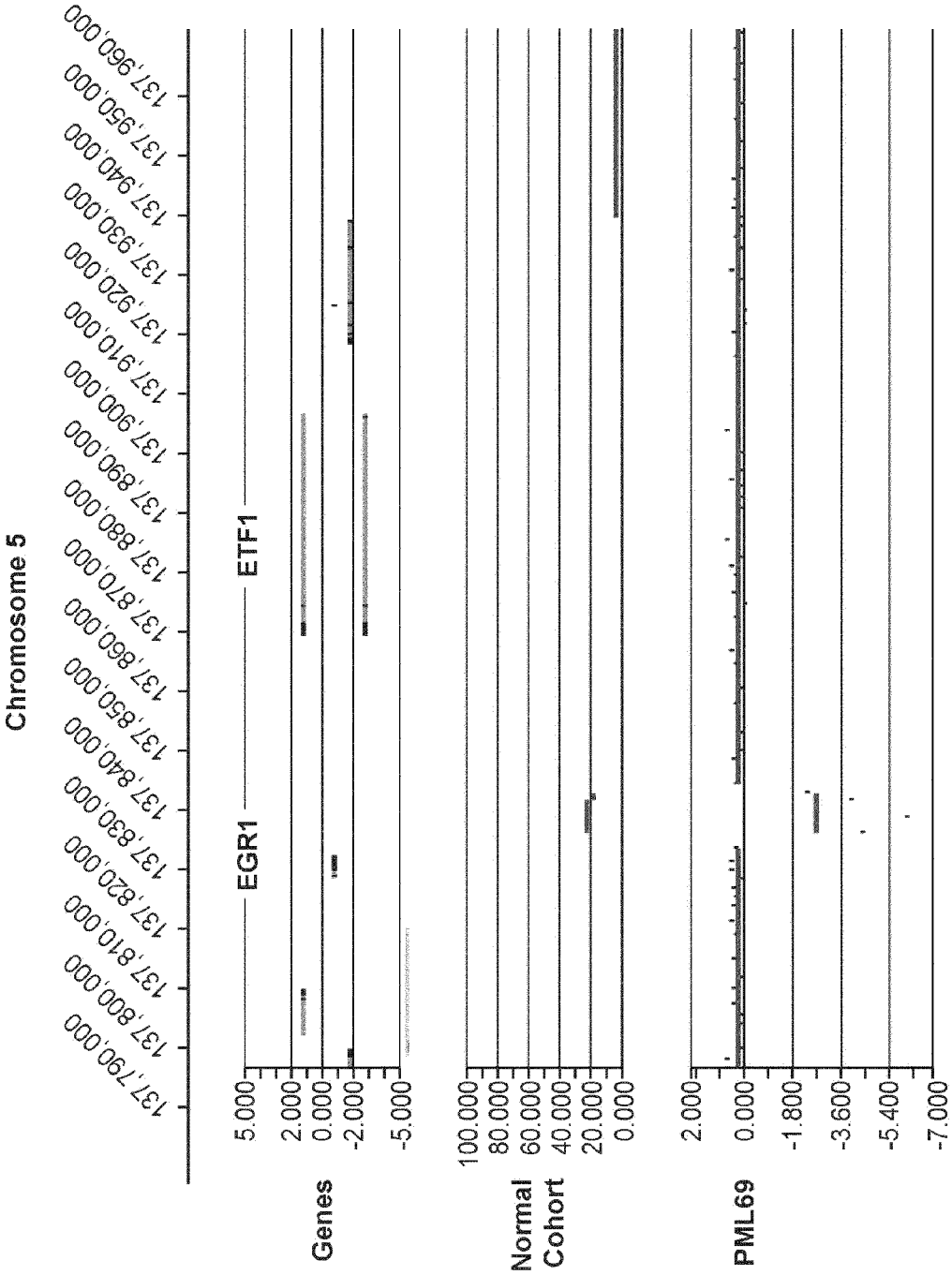

Chromosome 5
137,790,000
137,800,000
137,810,000
137,820,000
137,830,000
137,840,000
137,850,000
137,860,000
137,870,000
137,880,000
137,890,000
137,900,000
137,910,000
137,920,000
137,930,000
137,940,000
137,950,000
137,960,000
EGR1
ETF1
Genes
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
100.000
80.000
60.000
40.000
20.000
0.000
PML69
2.000
0.000
-1.800
-3.600
-5.400
-7.000

FIG. 11

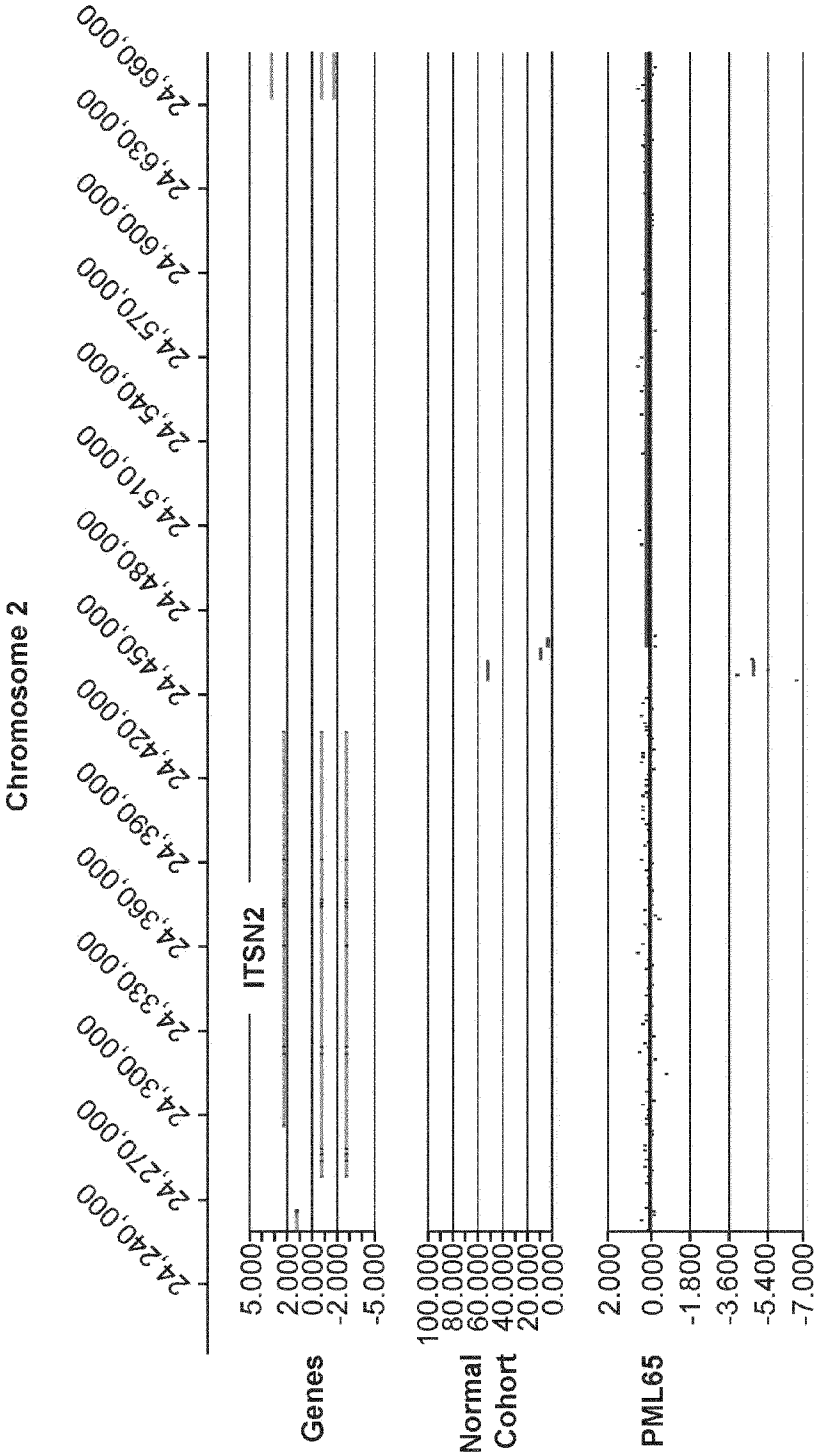

Chromosome 2
24,240,000
24,270,000
24,300,000
24,330,000
24,360,000
24,390,000
24,420,000
24,450,000
24,480,000
24,510,000
24,540,000
24,570,000
24,600,000
24,630,000
24,660,000
ITSN2
Genes
5.000
2.000
0.000
-2.000
-5.000
Normal Cohort
100.000
80.000
60.000
40.000
20.000
0.000
PML65
2.000
0.000
-1.800
-3.600
-5.400
-7.000

FIG. 12

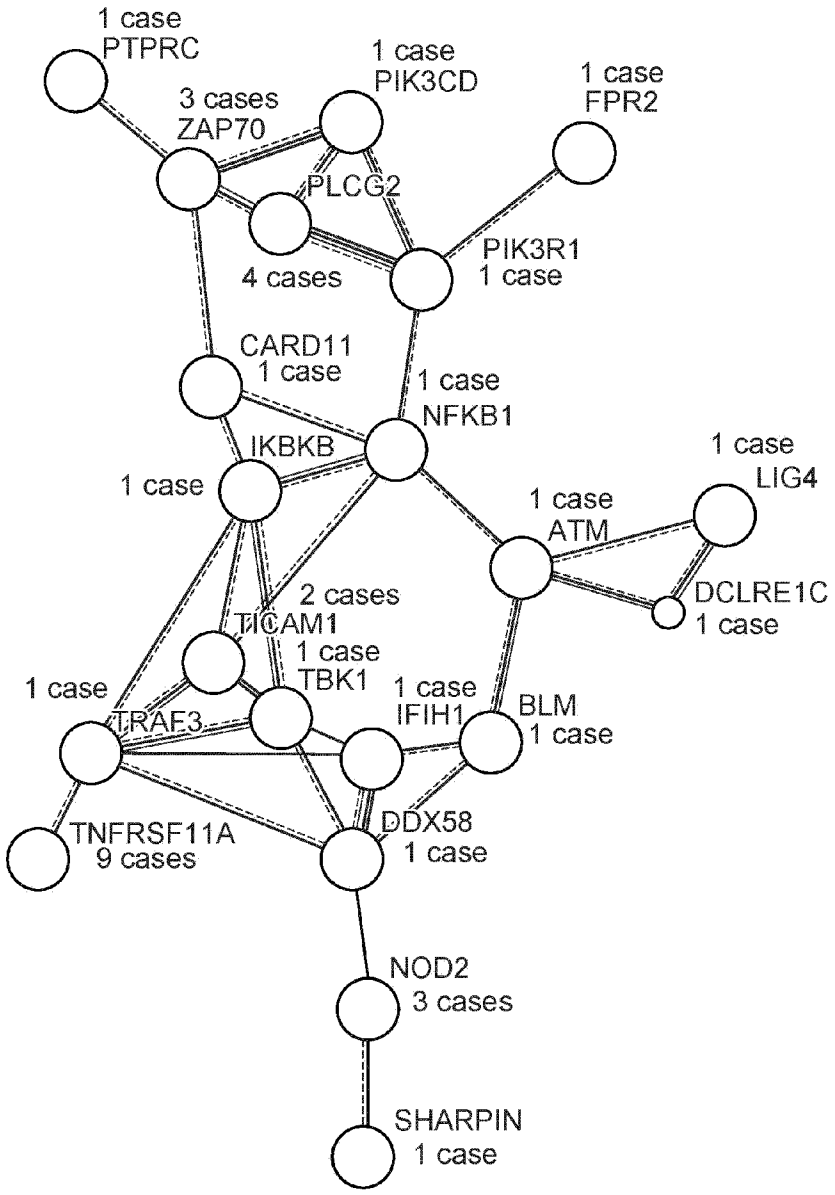

1 case
PTPRC
1 case
PIK3CD
1 case
FPR2
3 cases
ZAP70
PLCG2
4 cases
PIK3R1
1 case
CARD11
1 case
1 case
NFKB1
IKBKB
1 case
1 case
LIG4
1 case
ATM
2 cases
TICAM1
DCLRE1C
1 case
1 case
TBK1
1 case
1 case
IFIH1
BLM
1 case
TRAF3
TNFRSF11A
9 cases
DDX58
1 case
NOD2
3 cases
SHARPIN
1 case

FIG. 13

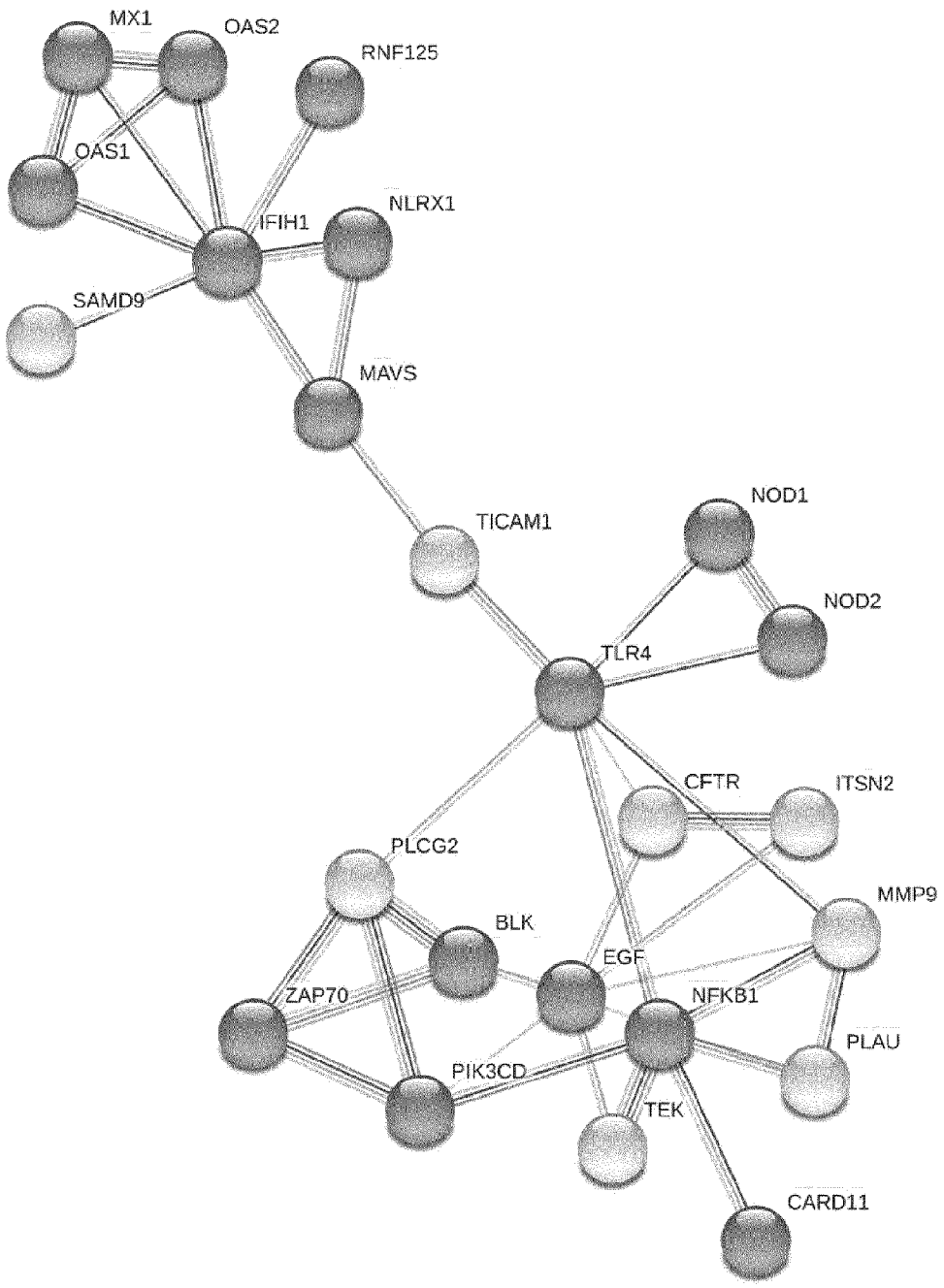

MX1
OAS2
RNF125
OAS1
NLRX1
IFIH1
SAMD9
MAVS
TICAM1
NOD1
NOD2
TLR4
CFTR
ITSN2
PLCG2
MMP9
BLK
EGF
ZAP70
NFKB1
PLAU
PIK3CD
TEK
CARD11

FIG. 14

SEKVENSLISTE

Sekvenslisten er udeladt af skriftet og kan hentes fra det Europæiske Patent Register.

The Sequence Listing was omitted from the document and can be downloaded from the European Patent Register.