



(51) International Patent Classification:

C07H 21/04 (2006.01) C12Q 1/6876 (2018.01)

C12Q 1/68 (2018.01)

(21) International Application Number:

PCT/US2019/050061

(22) International Filing Date:

06 September 2019 (06.09.2019)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/728,263 07 September 2018 (07.09.2018) US

62/741,434 04 October 2018 (04.10.2018) US

62/741,437 04 October 2018 (04.10.2018) US

62/741,439 04 October 2018 (04.10.2018) US

62/741,805 05 October 2018 (05.10.2018) US

62/741,807 05 October 2018 (05.10.2018) US

(71) Applicant: JUNEAU BIOSCIENCES, L.L.C. [US/US];

2749 East Parleys Way, Suite 300, Salt Lake City, Utah 84109 (US).

(72) Inventors: ALBERTSEN, Hans; 2749 East Parleys Way,

Suite 300, Salt Lake City, Utah 84109 (US). CHETTIER,

Rakesh N.; 2749 East Parleys Way, Suite 300, Salt Lake

City, Utah 84109 (US). WARD, Kenneth; 2749 East Par-

leys Way, Suite 300, Salt Lake City, Utah 84109 (US). AR-

GYLE, VeeAnn; 2749 East Parleys Way, Suite 300, Salt

Lake City, Utah 84109 (US).

(74) Agent: ARONIN, Caren P.; WILSON SONSINI

GOODRICH & ROSATI, 650 Page Mill Road, Palo Alto,

California 94304 (US).

(81) Designated States (unless otherwise indicated, for every

kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every

kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

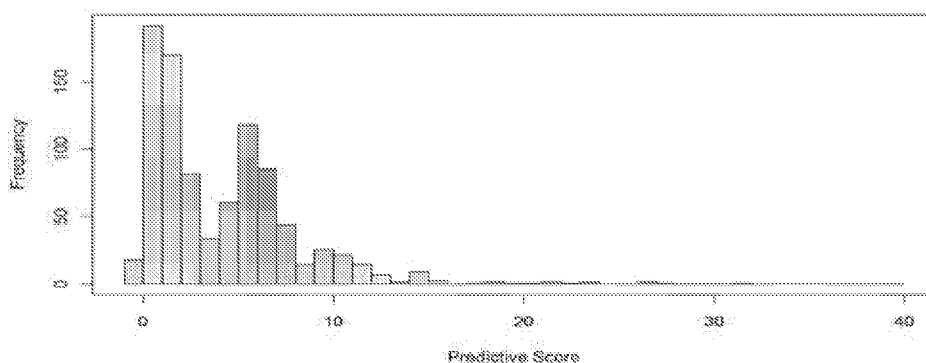
- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))
- as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))

Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: METHODS OF USING GENETIC MARKERS ASSOCIATED WITH ENDOMETRIOSIS

FIG. 1A Endometriosis Subjects



(57) Abstract: Disclosed herein are methods of using genetic markers associated with endometriosis, for example via a computer-implemented program to predict risk of developing endometriosis, and methods of preventing or treating endometriosis or a symptom thereof.



METHODS OF USING GENETIC MARKERS ASSOCIATED WITH ENDOMETRIOSIS

CROSS REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 62/728,263 filed September 7, 2018, U.S. Provisional Application No. 62/741,434 filed October 4, 2018, U.S. Provisional Application No. 62/741,805 filed October 5, 2018, U.S. Provisional Application No. 62/741,437 filed October 4, 2018, U.S. Provisional Application No. 62/741,807 filed October 5, 2018, and U.S. Provisional Application No. 62/741,439 filed October 4, 2018, each of which are incorporated by reference herein in their entirety.

BRIEF SUMMARY

[0002] The methods and systems described herein provide an approach for sequencing a nucleic acid sample using high throughput methods to detect genetic variants. These methods provide improved methods in the field of diagnosis, assessment and treatment of endometriosis. For example, disclosed herein is the use of nanopore sequencing to detect one or more genetic variants in a nucleic acid sample, wherein the one or more genetic variants are listed in **Table 1**, **Table 2** or **Table 3**.

INCORPORATION BY REFERENCE

[0003] All publications, patents, and patent applications mentioned, disclosed or referenced in this specification are herein incorporated by reference in their entirety and to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

[0004] **FIG. 1A-1B** is a set of bar charts showing distribution of predictive score using 775 rare variants among 917 endometriosis subjects and 917 controls generated through simulation using the ExAc published frequencies (All rare variants are assumed to be independent).

[0005] **FIG. 2** is a boxplot of the predictive score across the clinical subtypes of endometriosis. Endoscore is uniform across the severity of endometriosis.

[0006] **FIG. 3** is a pie chart showing diverse pathways implicated by these 729 genes. No pathway reaches statistical significance, but multiple genes implicated in the Wnt, cadherin, integrin, and inflammation mediated by cytokine signaling pathways.

[0007] **FIG. 4** is a diagram showing three experimental design strategies. Sequencing nuclear families can help identify Mendelian segregation, whereas relative pairs can help uncover distant relationships with IBD. Unrelated individuals are typically studied to identify common variants with small effects.

[0008] **FIG. 5** is a diagram showing a nuclear family with an IGF2 mutation on the left and an extended pedigree with a LONP1 mutation to the right.

[0009] **FIG. 6** is a diagram of mutation patterns cis/ trans/ haplotypes.

[0010] **FIG. 7** is a bar chart showing example of results: genes implicated in GWAS (genome-wide association studies) meta-analyses.

[0011] **FIG. 8A-8C** is a set of diagrams showing striking excess of pathogenic mutations ($p < 10^{-16}$).

[0012] **FIG. 9** is a set of charts showing examples of FN1 and GREB1 in which multiple damaging mutations were found.

[0013] **FIG. 10** is a diagram showing a computer-based system that may be programmed or otherwise configured to implement methods provided herein.

[0014] **FIG. 11** is a diagram showing a method and system as disclosed herein.

[0015] **FIG. 12** shows the whole exome sequencing method used in Example 9.

[0016] **FIG. 13** shows the sample population of Example 9 of 137 women with surgically confirmed endometriosis and a common ancestor born in 1608.

[0017] **FIG. 14** shows a common ancestor in GenDB 15-17 generations ago.

[0018] **FIG. 15** shows a three generation family with 7 women affected with endometriosis is shown in **FIG. 15A** with a brief clinical description of their endometriosis-related symptoms tabularized in **FIG. 15B**. in addition, patient 1 has been diagnosed with 14 additional co-morbidities including: Crohn's disease, interstitial cystitis, urinary bladder diverticulum, bronchial asthma, osteoporosis, multinodular goiter, cardiovascular disease, gastroesophageal reflux disease, malignant tumor of urinary bladder, Barrett's esophagus, lupus erythematosus, ankylosing spondylitis, multiple sclerosis, and bilateral cataract.

[0019] **FIG. 16** shows the chromosomal position and characteristics of the genetic variants surrounding the hemizygous deletions is shown to the left, and the genotypes for each of the seven affected women is shown to the right. Bold borders indicate the extent of the deletion and the individual that carries the deletions. Thin borders indicate possible carriers of the deletion.

[0020] **FIG. 17** shows results of Example 11 including number and percentage of matched probands.

[0021] **FIG. 18** shows the materials and methods of Example 11.

[0022] FIG. 19 shows percentage of affected subjects in both the index pedigree and unrelated pedigrees.

[0023] FIG. 20 shows the rate of surgically diagnosed endometriosis.

DETAILED DESCRIPTION

[0024] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of the ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the compositions or unit doses herein, some methods and materials are now described. Unless mentioned otherwise, the techniques employed or contemplated herein are standard methodologies. The materials, methods and examples are illustrative only and not limiting.

[0025] The details of one or more inventive instances are set forth in the accompanying drawings, the claims, and the description herein. Other features, objects, and advantages of the inventive instances disclosed and contemplated herein can be combined with any other instance unless explicitly excluded.

[0026] In some of many aspects, the disclosure provides methods of using genetic markers associated with endometriosis, for example via a computer-implemented program to predict risk of developing endometriosis, and methods of preventing or treating endometriosis or a symptom thereof. The methods disclosed herein can prevent or cancel an invasive procedure, such as a laparoscopy, that would otherwise have been performed on a subject but for the results, for example a (negative) diagnosis/prognosis, from the methods disclosed herein performed on the subject.

[0027] In some cases, genetic markers disclosed herein can be used for early diagnosis and prognosis of endometriosis, as well as early clinical intervention to mitigate progression of the disease. The use of these genetic markers can allow selection of subjects for clinical trials involving novel treatment methods. In some instances, genetic markers disclosed herein can be used to predict endometriosis and endometriosis progression, for example in treatment decisions for individuals who are recognized as having endometriosis. In some instances, genetic markers disclosed herein can enable prognosis of endometriosis in much larger populations compared with the populations which can currently be evaluated by using existing risk factors and biomarkers.

[0028] In some cases, disclosed herein is a method for endometriosis diagnosis/prognosis that can utilize detection of endometriosis associated biomarkers such as single nucleotide polymorphisms (SNPs), insertion deletion polymorphisms (indels), damaging mutation variants,

loss of function variants, synonymous mutation variants, nonsynonymous mutation variants, nonsense mutations, recessive markers, splicing/splice-site variants, frameshift mutations, insertions, deletions, genomic rearrangements, stop-gain, stop-loss, Rare Variants (RVs), some of which are identified in **Tables 1-4** (or diagnostically and predicatively functionally comparable biomarkers). In some instances, the method can comprise using a statistical assessment method such as Multi Dimensional Scaling analysis (MDS), logistic regression, machine learning, or Bayesian analysis.

[0029] Some of the variants listed in **Table 1** can be splicing variants, for example TMED3(NM_007364:exon1:c.168+1G>A), NM_001276480:c.-160+1G>A, KCNK6(NM_004823:exon2:c.323-1G>A), RGP4(NM_182588:exon19:c.2606-1G>T), NM_001001891:exon18:c.1988+1G>A, NM_001882:exon3:c.176-2->C. The NM number indicates that a particular GenBank cDNA reference sequence was used for reference. The "c" indicates that the nucleotide number which follows is based on coding DNA sequence. The numbers provide the position of the mutation in the DNA. For instance, 168+1G>A means one base after (+1) the 168th coding nucleotide at the end of the exon is mutated from a G to an A. Likewise for NM_182588:exon19:c.2606-1G>T, one base before (-1) the 2606th coding nucleotide. NM_001882:exon3:c.176-2->C involves an insertion of a C.

[0030] In some cases, disclosed herein is a treatment method to a subject determined to have or be predisposed to endometriosis. In some instances, the method can comprise administering to the subject a hormone therapy or an assisted reproductive technology therapy. In some instances, the method can comprise administering to the subject a therapy that at least partially compensates for endometriosis, prevents or reduces the severity of endometriosis that the subject would otherwise develop, or prevents endometriosis related complications, cancers, or associated disorders.

[0031] In some cases, provided herein is identification of new variants such as SNPs or indels, unique combinations of such variants, and haplotypes of variants that are associated with endometriosis and related pathologies. In some instances, the polymorphisms disclosed herein can be directly useful as targets for the design of diagnostic reagents and the development of therapeutic agents for use in the diagnosis and treatment of endometriosis and related pathologies. Based on the identification of variants associated with endometriosis, the disclosure can provide methods of detecting these variants as well as the design and preparation of detection reagents needed to accomplish this task. Provided herein are novel variants in genetic sequences involved in endometriosis, methods of detecting these variants in a test sample, methods of identifying individuals who have an altered risk of developing endometriosis and for

suggesting treatment options for endometriosis based on the presence of a variant(s) disclosed herein or its encoded product and methods of identifying individuals who are more or less likely to respond to a treatment.

[0032] In some cases, provided herein are variants such as SNPs and indels associated with endometriosis, nucleic acid molecules containing variants, methods and reagents for the detection of the variants disclosed herein, uses of these variants for the development of detection reagents, and assays or kits that utilize such reagents. In some instances, the variants disclosed herein can be useful for diagnosing, screening for, and evaluating predisposition to endometriosis and progression of endometriosis. In some instances, the variants can be useful in the determining individual subject treatment plans and design of clinical trials of devices for possible use in the treatment of endometriosis. In some instances, the variants and their encoded products can be useful targets for the development of therapeutic agents. In some instances, the variants combined with other non-genetic clinical factors can be useful for diagnosing, screening, evaluating predisposition to endometriosis, assessing risk of progression of endometriosis, determining individual subject treatment plans and design of clinical trials of devices for possible use in the treatment of endometriosis. In some instances, the variants can be useful in the selection of recipients for an oral contraceptive type therapeutic.

[0033] Definitions

[0034] Unless otherwise indicated, open terms for example “contain,” “containing,” “include,” “including,” and the like mean comprising.

[0035] The singular forms “a”, “an”, and “the” are used herein to include plural references unless the context clearly dictates otherwise. Accordingly, unless the contrary is indicated, the numerical parameters set forth in this application are approximations that may vary depending upon the desired properties sought to be obtained by the present invention.

[0036] Unless otherwise indicated, some instances herein contemplate numerical ranges. When a numerical range is provided, unless otherwise indicated, the range includes the range endpoints. Unless otherwise indicated, numerical ranges include all values and subranges therein as if explicitly written out. Unless otherwise indicated, any numerical ranges and/or values herein, following or not following the term “about,” can be at 85-115% (i.e., plus or minus 15%) of the numerical ranges and/or values.

[0037] As used herein, "endometriosis" can refer to any nonmalignant disorder in which functioning endometrial tissue is present in a location in the body other than the endometrium of the uterus, i.e. outside the uterine cavity or is present within the myometrium of the uterus. For purposes herein it also includes conditions, such as adenomyosis/adenomyoma, that exhibit

myometrial tissue in the lesions. Endometriosis can include endometriosis externa, endometrioma, adenomyosis, adenomyomas, adenomyotic nodules of the uterosacral ligaments, endometriotic nodules other than of the uterosacral ligaments, autoimmune endometriosis, mild endometriosis, moderate endometriosis, severe endometriosis, superficial (peritoneal) endometriosis, deep (invasive) endometriosis, ovarian endometriosis, endometriosis-related cancers, and/or "endometriosis-associated conditions". Unless stated otherwise, the term endometriosis is used herein to describe any of these conditions.

[0038] As used herein, "treatment" includes one or more of: reducing the frequency and/or severity of symptoms, elimination of symptoms and/or their underlying cause, and improvement or remediation of damage. For example, treatment of endometriosis includes, for example, relieving the pain experienced by a woman suffering from endometriosis, and/or causing the regression or disappearance of endometriotic lesions.

[0039] As used herein, a "therapeutic" can include a medical device, a pharmaceutical composition, a medical procedure, or any combination thereof. In some embodiments, a medical device may comprise a spinal brace. In some embodiments a medical device may comprise an artificial disc device. A medical device may comprise a surgical implant. A pharmaceutical composition may comprise a muscle relaxant, an anti-depressant, a steroid, an opioid, a cannabis-based therapeutic, acetaminophen, a non-steroidal anti-inflammatory, a neuropathic agent, a cannabis, a progestin, a progesterone, or any combination thereof. A neuropathic agent may comprise gabapentin. A non-steroidal anti-inflammatory may comprise naproxen, ibuprofen, a COX-2 inhibitor, or any combination thereof. A pharmaceutical composition may comprises a biologic agent, cellular therapy, regenerative medicine therapy, a tissue engineering approach, a stem cell transplantation or any combination thereof. A medical procedure may comprise an epidural injection (such as a steroid injection), acupuncture, exercise, physical therapy, an ultrasound, a radiofrequency ablation, a surgical therapy, a chiropractic manipulation, an osteopathic manipulation, or any combination thereof. A therapeutic can include a regenerative therapy such as a protein, a stem cell, a cord blood cell, an umbilical cord tissue, a tissue, or any combination thereof. A therapeutic can include cannabis. A therapeutic can include a biosimilar.

[0040] "Haplotype" can mean a combination of genotypes on the same chromosome occurring in a linkage disequilibrium block. Haplotypes serve as markers for linkage disequilibrium blocks, and at the same time provide information about the arrangement of genotypes within the blocks. Typing of only certain variants which serve as tags can, therefore, reveal all genotypes for variants located within a block. Thus, the use of haplotypes greatly facilitates identification of candidate genes associated with diseases and drug sensitivity.

[0041] "Linkage disequilibrium" or "LD" can mean that a particular combination of alleles (alternative nucleotides) or genetic variants for example at two or more different SNP (or RV) sites are non-randomly co-inherited (i.e., the combination of alleles at the different SNP (or RV)

sites occurs more or less frequently in a population than the separate frequencies of occurrence of each allele or the frequency of a random formation of haplotypes from alleles in a given population). The term "LD" can differ from "linkage," which describes the association of two or more loci on a chromosome with limited recombination between them. LD can also be used to refer to any non-random genetic association between allele(s) at two or more different SNP (or RV) sites. In some instances, when a genetic marker (e.g. SNP or RV) is identified as the genetic marker associated with a disease (in this instance endometriosis), it can be the minor allele (MA) of the particular genetic marker that is associated with the disease. In some instances, if the Odds Ratio (OR) of the MA is greater than 1.0, the MA of the genetic marker (in this instance the endometriosis associated genetic marker) can be correlated with an increased risk of endometriosis in a case subject as compared to a control subject and can be considered a causative marker (C), and if the OR of the MA less than 1.0, the MA of the genetic marker can be correlated with a decreased risk of endometriosis in a case subject as compared to a control subject and can be considered a protective marker (P). "Linkage disequilibrium block" or "LD block" can mean a region of the genome that contains multiple variants located in proximity to each other and that are transmitted as a block.

[0042] Biological samples obtained from individuals (e.g., human subjects) may be any sample from which a genetic material (e.g., nucleic acid sample) may be derived. Samples/Genetic materials may be from buccal swabs, saliva, blood, hair, nail, skin, cell, or any other type of tissue sample. In some instances, the genetic material (e.g., nucleic acid sample) comprises mRNA, cDNA, genomic DNA, or PCR amplified products produced therefrom, or any combination thereof. In some instances, the genetic material (e.g., nucleic acid sample) comprises PCR amplified nucleic acids produced from cDNA or mRNA. In some instances, the genetic material (e.g., nucleic acid sample) comprises PCR amplified nucleic acids produced from genomic DNA.

[0043] As used herein, the term "cell-free" or "cell free" can refer to the condition of the nucleic acid sequence as it appeared in the body before the sample is obtained from the body. For example, circulating cell-free nucleic acid sequences in a sample may have originated as cell-free nucleic acid sequences circulating in the bloodstream of the human body. In contrast, nucleic acid sequences that are extracted from a solid tissue, such as a biopsy, are generally not considered to be "cell-free." In some cases, cell-free DNA may comprise fetal DNA, maternal DNA, or a combination thereof. In some cases, cell-free DNA may comprise DNA fragments released into a blood plasma. In some cases, the cell-free DNA may comprise circulating tumor DNA. In some cases, cell-free DNA may comprise circulating DNA indicative of a tissue origin,

a disease or a condition. A cell-free nucleic acid sequence may be isolated from a blood sample. A cell-free nucleic acid sequence may be isolated from a plasma sample. A cell-free nucleic acid sequence may comprise a complementary DNA (cDNA). In some cases, one or more cDNAs may form a cDNA library.

[0044] Analysis of Rare and Private Mutations in Sequenced Endometriosis Genes

[0045] In some cases, the disclosure provides an analysis to evaluate a coding region of a gene as a component of a genetic diagnostic or predictive test for endometriosis. In some instances, the analysis can comprise one or more of the approaches disclosed herein.

[0046] In some instances, the analysis can comprise performing DNA variant search on the next generation sequencing output file using a standard software designed for this purpose, for example Life Technologies TMAP algorithm with their default parameter settings, and Life Technologies Torrent Variant Caller software. ANNOVAR can be used to classify coding variants as synonymous, missense, frameshift, splicing, stop-gain, or stop-loss. Variants can be considered “loss-of-function” if the variant causes a stop-loss, stop-gain, splicing, or frameshift insertion or deletion).

[0047] In some instances, the analysis can comprise evaluating prediction of an effect of each variant on protein function in silico using a variety of different software algorithms: Polyphen 2, Sift, Mutation Accessor, Mutation Taster, FATHMM, LRT, MetaLR, or any combination thereof. Missense variants can be deemed "damaging" if they are predicted to be damaging by at least one of the seven algorithms tested.

[0048] In some instances, the analysis can comprise searching population databases (e.g., gnomAD) and proprietary endometriosis allele frequency databases for the prevalence of any loss of function or damaging mutations identified by these analyses. The log of the odds ratio can be used to weight the marker when the variant has been previously observed in the reference databases. When a damaging variant or loss of function variant has never been reported in the reference databases, a default odds ratio of 10 can be used to weight the finding.

[0049] In some instances, the analysis can comprise incorporating findings into the Risk Score as with the other low-frequency alleles. Risk Score = Summation [$\log(\text{OR}) \times \text{Count}$], where count equals the number of low frequency alleles detected at each endometriosis associated locus. Risk scores can be converted to probability using a nomogram based on confirmed diagnoses.

[0050] In some instances, the methods of the disclosure can provide a high sensitivity of detecting gene mutations and diagnosing endometriosis that is greater than 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 95.5%, 96%, 96.5%, 97%, 97.5%, 98%,

98.5%, 99%, 99.5% or more. In some instances, the methods disclosed herein can provide a high specificity of detecting and classifying gene mutations and endometriosis, for example, greater than 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 95.5%, 96%, 96.5%, 97%, 97.5%, 98%, 98.5%, 99%, 99.5% or more. In some instances, a nominal specificity for the method disclosed herein can be greater than or equal to 70%. In some instances, a nominal Negative Predictive Value (NPV) for the method disclosed herein can be greater than or equal to 95%. In some instances, a NPV for the method disclosed herein can be about 95%, 95.5%, 96%, 96.5%, 97%, 97.5%, 98%, 98.5%, 99%, 99.5% or more. In some instances, a nominal Positive Predictive Value (PPV) for the method disclosed herein can be greater than or equal to 95%. In some instances, a PPV for the method disclosed herein can be about 95%, 95.5%, 96%, 96.5%, 97%, 97.5%, 98%, 98.5%, 99%, 99.5% or more. In some instances, the accuracy of the methods disclosed herein in diagnosing endometriosis can be greater than 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 95.5%, 96%, 96.5%, 97%, 97.5%, 98%, 98.5%, 99%, 99.5% or more.

[0051] Computer Implemented Methods

[0052] In some aspects, the disclosure provides methods for analysis of gene sequence data associated software and computer systems. The method, for example being computer implemented, can enable a clinical geneticist or other healthcare technician to sift through vast amounts of gene sequence data, to identify potential disease-causing genomic variants. In some cases, the gene sequence data is from a patient who may be suspected of having a genetic disorder such as endometriosis.

[0053] In some cases, provided herein is a method for identifying a genetic disorder such as endometriosis or predicting a risk thereof in an individual, or identifying a genetic variant that is causative of a phenotype in an individual. In some instances, the method can comprise determining gene sequence for a patient suspected of having a genetic disorder, identifying sequence variants, annotating the identified variants based on one or more criteria, and filtering or searching the variants at least partially based on the annotations, to thereby identify potential disease-causing variants.

[0054] In some instances, the gene sequence is obtained by use of a sequencing instrument, or alternatively, gene sequence data is obtained from another source, such as for example, a commercial sequencing service provider. Gene sequence can be chromosomal sequence, cDNA sequence, or any nucleotide sequence information that allows for detection of genetic disease. Generally, the amount of sequence information is such that computational tools are required for data analysis. For example, the sequence data may represent at least half of the individual's

genomic or cDNA sequence (e.g., of a representative cell population or tissue), or the individuals entire genomic or cDNA sequence. In various embodiments, the sequence data comprises the nucleotide sequence for at least 1 million base pairs, at least 10 million base pairs, or at least 50 million base pairs. In certain embodiments, the DNA sequence is the individual's exome sequence or full exonic sequence component (i.e., the exome; sequence for each of the exons in each of the known genes in the entire genome). In some embodiments, the source of genomic DNA or cDNA may be any suitable source, and may be a sample particularly indicative of a disease or phenotype of interest, including blood cells (e.g, PBMCs, or a T-cell or B-cell population). In certain embodiments, the source of the sample is a tissue or sample that is potentially malignant.

[0055] In some instances, whole genome sequence can comprise the entire sequence (including all chromosomes) of an individual's germline genome. In some embodiments, the concatenated length for a whole genome sequence is approximately 3.2 Gbases or 3.2 billion nucleotides.

[0056] The term “subject,” as used herein, may be any animal or living organism. Animals can be mammals, such as humans, non-human primates, rodents such as mice and rats, dogs, cats, pigs, sheep, rabbits, and others. A subject may be a dog. A subject may be a human. Animals can be fish, reptiles, or others. Animals can be neonatal, infant, adolescent, or adult animals. Humans can be more than about: 1, 2, 5, 10, 20, 30, 40, 50, 60, 65, 70, 75, or about 80 years of age. The subject may have or be suspected of having a condition or a disease, such as endometriosis or related condition. The subject may be a patient, such as a patient being treated for a condition or a disease, such as a patient suffering from endometriosis. The subject may be predisposed to a risk of developing a condition or a disease such as endometriosis. The subject may be in remission from a condition or a disease, such as a patient recovering from endometriosis. The subject may be healthy. The subject may be a subject in need thereof. The subject may be a female subject or a male subject.

[0057] The term “sequencing” as used herein, may comprise high-throughput sequencing, next-gen sequencing, Maxam-Gilbert sequencing, massively parallel signature sequencing, Polony sequencing, 454 pyrosequencing, pH sequencing, Sanger sequencing (chain termination), Illumina sequencing, SOLiD sequencing, Ion Torrent semiconductor sequencing, DNA nanoball sequencing, Heliscope single molecule sequencing, single molecule real time (SMRT) sequencing, nanopore sequencing, shot gun sequencing, RNA sequencing, Enigma sequencing, sequencing-by-hybridization, sequencing-by-ligation, or any combination thereof. The sequencing output data may be subject to quality controls, including filtering for quality (e.g., confidence) of base reads. Exemplary sequencing systems include 454 pyrosequencing (454 Life Sciences), Illumina (Solexa) sequencing, SOLiD (Applied Biosystems), and Ion Torrent Systems' pH sequencing system. In some cases, a nucleic acid of a sample may be sequenced without an associated label or tag. In some cases, a nucleic acid of a sample may be sequenced, the nucleic acid of which may have a label or tag associated with it.

[0058] In some instances, the gene sequence may be determined by any suitable method. For example, the gene sequence may be a cDNA sequence determined by clonal amplification (e.g., emulsion PCR) and sequencing. Base calling may be conducted based on any available method, including Sanger sequencing (chain termination), pH sequencing, pyrosequencing, sequencing-by-hybridization, sequencing-by-ligation, etc. The sequencing output data may be subject to quality controls, including filtering for quality (e.g., confidence) of base reads. Exemplary sequencing systems include 454 pyrosequencing (454 Life Sciences), Illumina (Solexa) sequencing, SOLiD (Applied Biosystems), and Ion Torrent Systems' pH sequencing system. In some embodiment, sequencing can be performed by nanopore sequencing. For example, Oxford nanopore sequencing.

[0059] Nanopores may be used to sequence, a sample, a small portion (such as one full gene or a portion of one gene), a substantial portion (such as multiple genes or multiple chromosomes), or the entire genomic sequence of an individual. Nanopore sequencing technology may be commercially available or under development from Sequenom (San Diego, Calif.), Illumina (San Diego, Calif.), Oxford Nanopore Technologies LTD (Kidlington, United Kingdom), and Agilent Laboratories (Santa Clara, Calif.). Nanopore sequencing methods and apparatus are have been described in the art and for example are provided in U.S. Pat. No. 5,795,782, herein incorporated by reference in its entirety.

[0060] Nanopore sequencing can use electrophoresis to transport a sample through a pore. A nanopore system may contain an electrolytic solution such that when a constant electric field is applied, an electric current can be observed in the system. The magnitude of the electric current density across a nanopore surface may depend on the nanopore's dimensions and the composition of the sample that is occupying the nanopore. During nanopore sequencing, when a sample approaches and or goes through the nanopore, the samples cause characteristic changes in electric current density across nanopore surfaces, these characteristic changes in the electric current enables identification of the sample. Nanopores used herein may be solid-state nanopores, protein nanopores, or hybrid nanopores comprising protein nanopores or organic nanotubes such as carbon or graphene nanotubes, configured in a solid-state membrane, or like framework. In some embodiments, nanopore sequencing can be biological, a solid state nanopore or a hybrid biological/solid state nanopore.

[0061] In some instances, a biological nanopore can comprise transmembrane proteins that may be embedded in lipid membranes. In some embodiments, a nanopore described herein may comprise alpha hemolysin. In some embodiments, a nanopore described herein may comprise mycobacterium smegmatis porin.

[0062] Solid state nanopores do not incorporate proteins into their systems. Instead, solid state nanopore technology uses various metal or metal alloy substrates with nanometer sized pores that allow samples to pass through. Solid state nanopores may be fabricated in a variety of materials including but not limited to, silicon nitride (Si_3N_4), silicon dioxide (SiO_2), and the like. In some instances, nanopore sequencing may comprise use of tunneling current, wherein a measurement of electron tunneling through bases as sample (ssDNA) translocates through the nanopore is obtained. In some embodiments, a nanopore system can have solid state pores with single walled carbon nanotubes across the diameter of the pore. In some embodiments, nanoelectrodes may be used on a nanopore system described herein. In some embodiments, fluorescence can be used with nanopores, for example solid state nanopores and fluorescence. For example, In such a system the fluorescence sequencing method converts each base of a sample into a characteristic representation of multiple nucleotides which bind to a fluorescent probe strand-forming dsDNA (where the sample comprises DNA). Where a two color system is used, each base is identified by two separate fluorescence, and will therefore be converted into two specific sequences. Probes may consist of a fluorophore and quencher at the start and end of each sequence, respectively. Each fluorophore may be extinguished by the quencher at the end of the preceding sequence. When the dsDNA is translocating through a solid state nanopore, the probe strand may be stripped off, and the upstream fluorophore will fluoresce.

[0063] In some embodiments, a 1-100 nm channel or aperture may be formed through a solid substrate, usually a planar substrate, such as a membrane, through which an analyte, such as single stranded DNA, is induced to translocate. In other embodiments, a 2-50 nm channel or aperture is formed through a substrate; and in still other embodiments, a 2-30 nm, or a 2-20 nm, or a 3-30 nm, or a 3-20 nm, or a 3-10 nm channel or aperture if formed through a substrate.

[0064] In some embodiments, nanopores used in connection with the methods and devices of the invention are provided in the form of arrays, such as an array of clusters of nanopores, which may be disposed regularly on a planar surface. In some embodiments, clusters are each in a separate resolution limited area so that optical signals from nanopores of different clusters are distinguishable by the optical detection system employed, but optical signals from nanopores within the same cluster cannot necessarily be assigned to a specific nanopore within such cluster by the optical detection system employed.

[0065] In some instances, the gene sequence may be mapped with one or more reference sequences to identify sequence variants. For example, the base reads are mapped against a reference sequence, which in various embodiments is presumed to be a “normal” non-disease sequence. The DNS sequence derived from the Human Genome Project is generally used as a

“premier” reference sequence. A number of mapping applications are known, and include TMAP, BWA, GSMAPPER, ELAND, MOSAIK, and MAQ. Various other alignment tools are known, and could also be implemented to map the base reads.

[0066] In some cases, based on the sequence alignments, and mapping results, sequence variants can be identified. Types of variants may include insertions, deletions, indels (a colocalized insertion and deletion), damaging mutation variants, loss of function variants, synonymous mutation variants, nonsynonymous mutation variants, nonsense mutations, recessive markers, splicing/splice-site variants, frameshift mutation, insertions, deletions, genomic rearrangements, stop-gain, stop-loss, Rare Variants (RVs), translocations, inversions, and substitutions. While the type of variants analyzed is not limited, the most numerous of the variant types will be single nucleotide substitutions, for which a wealth of data is currently available. In various embodiments, comparison of the test sequence with the reference sequence will produce at least 500 variants, at least 1000 variants, at least 3,000 variants, at least 5,000 variants, at least 10,000 variants, at least 20,000 variants, or at least 50,000 variants, but in some embodiments, will produce at least 1 million variants, at least 2 million variants, at least 3 million variants, at least 4 million variants, or at least 10 million variants. The tools provided herein enable the user to navigate the vast amounts of genetic data to identify potentially disease-causing variants.

[0067] In some cases, a wealth of data can be extracted for the identified variants, including one or more of conservation scores, genic/genomic location, zygosity, SNP ID, Polyphen, FATHMM, LRT, Mutation Accessor, and SIFT predictions, splice site predictions, amino acid properties, disease associations, annotations for known variants, variant or allele frequency data, and gene annotations. Data may be calculated and/or extracted from one or more internal or external databases. Since certain categories of annotations (e.g., amino acid properties/PolyPhen and SIFT data) are dependent on a nature of the region of the genome in which they are contained (e.g., whether a variant is contained within a region translated to give rise to an amino acid sequence in a resultant protein), these annotations can be carried out for each known transcript. Exemplary external databases include OMIM (Online Mendelian Inheritance in Man), HGMD (The Human Gene Mutation Database), PubMed, PolyPhen, SIFT, SpliceSite, reference genome databases, the University of California Santa Cruz (UCSC) genome database, CLINVAR database, the BioBase biological databases, the dbSNP Short Genetic Variations database, the Rat Genome Database (RGD), and/or the like. Various other databases may be employed for extracting data on identified variants. Variant information may be further stored in a central data repository, and the data extracted for future sequence analyses.

[0068] In some instances, variants may be tagged by the user with additional descriptive information to aid subsequent analysis. For example, confidence in the existence of the variant can be recorded as confirmed, preliminary, or sequence artifact. Certain sequencing technologies have a tendency to produce certain types of sequence artifacts, and the method herein can allow such suspected artifacts to be recorded. The variants may be further tagged in basic categories of benign, pathogenic, or unknown, or as potentially of interest.

[0069] In some instances, queries can be run to identify variants meeting certain criteria, or variant report pages can be browsed by chromosomal position or by gene, the latter allowing researchers to focus on only those variations that exist in a particular set of genes of interest. In some embodiments, the user selects only variants with well-documented and published disease associations (e.g., by filtering based on HGMD or other disease annotation). Alternatively, the user can filter for variants not previously associated with disease, but of a type likely to be deleterious, such as those introducing frameshifts, non-synonymous substitutions (predicted by Polyphen or SIFT), or premature terminations. Further, the user can exclude from analysis those variants believed to be neutral (based on their frequency of occurrence in studies populations), for example, through exclusion of variants in dbSNP. Additional exclusion criteria include mode of inheritance (e.g., heterozygosity), depth of coverage, and quality score.

[0070] In certain embodiments, base calling is carried out to extract the sequence of the sequencing reads from an image file produced by an instrument scanner. Following base calling and base quality trimming/filtering, the reads are mapped against a reference sequence (assumed to be normal for the phenotype under analysis) to identify variations (variants) between the two with the assumption that one or more of these differences will be associated with phenotype of the individual whose DNA is under analysis. Subsequently, each variant is annotated with data that can be used to determine the likelihood that that particular variant is associated with the phenotype under analysis. The analysis may be fully or partially automated as described in detail below, and may include use of a central repository for data storage and analysis, and to present the data to analysts and clinical geneticists in a format that makes identification of variants with a high likelihood of being associated with the phenotypic difference more efficient and effective.

[0071] In some embodiments, a user can be provided with the ability to run cross sample queries where the variants from multiple samples are interrogated simultaneously. In such embodiments, for example, a user can build a query to return data on only those variants that are exactly shared across a user defined group of samples. This can be useful for family based analyses where the same variant is believed to be associated with disease in each of the affected family members. For another example, the user can also build a query to return only those variants that are present

in genes where the gene contains at least one, but not necessarily the same, variant. This can be useful where a group of individuals with disease are not related (the variants associated with the disease are not necessary exactly the same, but result in a common alteration in normal function). For yet another example, the user can specify to ignore genes containing variants in a user defined group of samples. This can be useful to exclude polymorphisms (variants believed or confirmed not to be associated with disease) where the user has access to a user defined group of control individuals who are believed to not have the disease associated variant. For each of these queries a user can additionally filter the variants by specifying any or all of the previously discussed filters on top of the cross sample analyses. This allows a user to identify variants matching these criteria, which are shared between or segregated amongst samples.

[0072] For example, a variant analysis system can be implemented locally, or implemented using a host device and a network or cloud computing. For example, the variant analysis system can be software stored in memory of a personal computing device (PC) and implemented by a processor of the PC. In such embodiments, for example, the PC can download the software from a host device and/or install the software using any suitable device such as a compact disc (CD).

[0073] The method may employ a computer-readable medium, or non-transitory processor-readable medium. Some embodiments described herein relate to a computer storage product with a non-transitory computer-readable medium (also can be referred to as a non-transitory processor-readable medium) having instructions or computer code thereon for performing various computer-implemented operations. The computer-readable medium (or processor-readable medium) is non-transitory in the sense that it does not include transitory propagating signals per se (e.g., a propagating electromagnetic wave carrying information on a transmission medium such as space or a cable). The media and computer code (also can be referred to as code) may be those designed and constructed for the specific purpose or purposes. Examples of non-transitory computer-readable media include, but are not limited to: magnetic storage media such as hard disks, floppy disks, and magnetic tape; optical storage media such as Compact Disc/Digital Video Discs (CD/DVDs), Compact Disc-Read Only Memories (CD-ROMs), and holographic devices; magneto-optical storage media such as optical disks; carrier wave signal processing modules; and hardware devices that are specially configured to store and execute program code, such as Application-Specific Integrated Circuits (ASICs), Programmable Logic Devices (PLDs), Read-Only Memory (ROM) and Random-Access Memory (RAM) devices.

[0074] Examples of computer code can include, but are not limited to, micro-code or micro-instructions, machine instructions, such as produced by a compiler, code used to produce a web service, and files containing higher-level instructions that are executed by a computer using an

interpreter. For example, embodiments may be implemented using Python, Java, C++, or other programming languages (e.g., object-oriented programming languages) and development tools. Additional examples of computer code can include, but are not limited to, control signals, encrypted code, and compressed code.

[0075] In some cases, variants provided herein may be “provided” in a variety of mediums to facilitate use thereof. As used in this section, “provided” can refer to a manufacture, other than an isolated nucleic acid molecule, that contains variant information of the disclosure. Such a manufacture provides the variant information in a form that allows a skilled artisan to examine the manufacture using means not directly applicable to examining the variants or a subset thereof as they exist in nature or in purified form. The variant information that may be provided in such a form includes any of the variant information provided by the disclosure such as, for example, polymorphic nucleic acid and/or amino acid sequence information, information about observed variant alleles, alternative codons, populations, allele frequencies, variant types, and/or affected proteins, or any other information provided herein.

[0076] In some instances, the variants can be recorded on a computer readable medium. As used herein, “computer readable medium” can refer to any medium that can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled artisan can readily appreciate how any of the presently known computer readable media can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the disclosure. One such medium is provided with the present application, namely, the present application contains computer readable medium (CD-R) that has nucleic acid sequences (and encoded protein sequences) containing variants provided/recorded thereon in ASCII text format in a Sequence Listing along with accompanying Tables that contain detailed variant and sequence information.

[0077] As used herein, “recorded” can refer to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the variant information of the disclosure. A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide or amino acid sequence of the disclosure. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide/amino acid sequence

information of the disclosure on computer readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, represented in the form of an ASCII file, or stored in a database application, such as OB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g., text file or database) in order to obtain computer readable medium having recorded thereon the variant information of the disclosure.

[0078] By providing the variants in computer readable form, a skilled artisan can access the variant information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. Examples of publicly available computer software include BLAST and BLAZE search algorithms.

[0079] In some cases, the disclosure can provide systems, particularly computer-based systems, which contain the variant information described herein. Such systems may be designed to store and/or analyze information on, for example, a large number of variant positions, or information on variant genotypes from a large number of individuals. The variant information of the disclosure represents a valuable information source. The variant information of the disclosure stored/analyzed in a computer-based system may be used for such computer-intensive applications as determining or analyzing variant allele frequencies in a population, mapping endometriosis genes, genotype-phenotype association studies, grouping variants into haplotypes, correlating variant haplotypes with response to particular treatments or for various other bioinformatic, pharmacogenomic or drug development.

[0080] As used herein, "a computer-based system" can refer to the hardware means, software means, and data storage means used to analyze the variant information of the disclosure. The minimum hardware means of the computer-based systems of the disclosure typically comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the disclosure. Such a system can be changed into a system of the disclosure by utilizing the variant information provided on the CD-R, or a subset thereof, without any experimentation.

[0081] As stated above, the computer-based systems can comprise a data storage means having stored therein variants of the disclosure and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" can refer to memory which can store variant information of the disclosure, or a memory access means

which can access manufactures having recorded thereon the variant information of the disclosure.

[0082] As used herein, "search means" can refer to one or more programs or algorithms that are implemented on the computer-based system to identify or analyze variants in a target sequence based on the variant information stored within the data storage means. Search means can be used to determine which nucleotide is present at a particular variant position in the target sequence. As used herein, a "target sequence" can be any DNA sequence containing the variant position(s) to be searched or queried.

[0083] A variety of structural formats for the input and output means can be used to input and output the information in the computer-based systems of the disclosure. An exemplary format for an output means is a display that depicts the presence or absence of specified nucleotides (alleles) at particular variant positions of interest. Such presentation can provide a rapid, binary scoring system for many variants simultaneously.

[0084] In some cases, the disclosure provides computer-based systems that are programmed to implement methods of the disclosure. **FIG. 10** shows a computer system **101** that can be programmed or configured for endometriosis diagnosis. The computer system **101** can regulate various aspects of detection of genetic variants associated with endometriosis of the disclosure. The computer system **101** can be an electronic device of a user or a computer system that is remotely located with respect to the electronic device. The electronic device can be a mobile electronic device.

[0085] The computer system **101** includes a central processing unit (CPU, also "processor" and "computer processor" herein) **105**, which can be a single core or multi core processor, or a plurality of processors for parallel processing. The computer system **101** also includes memory or memory location **110** (e.g., random-access memory, read-only memory, flash memory), electronic storage unit **115** (e.g., hard disk), communication interface **120** (e.g., network adapter) for communicating with one or more other systems, and peripheral devices **125**, such as cache, other memory, data storage and/or electronic display adapters. The memory **110**, storage unit **115**, interface **120** and peripheral devices **125** are in communication with the CPU **105** through a communication bus (solid lines), such as a motherboard. The storage unit **115** can be a data storage unit (or data repository) for storing data. The computer system **101** can be operatively coupled to a computer network ("network") **130** with the aid of the communication interface **120**. The network **130** can be the Internet, an internet and/or extranet, or an intranet and/or extranet that is in communication with the Internet. The network **130** in some cases is a telecommunication and/or data network. The network **130** can include one or more computer

servers, which can enable distributed computing, such as cloud computing. The network **130**, in some cases with the aid of the computer system **101**, can implement a peer-to-peer network, which may enable devices coupled to the computer system **101** to behave as a client or a server.

[0086] The CPU **105** can execute a sequence of machine-readable instructions, which can be embodied in a program or software. The instructions may be stored in a memory location, such as the memory **110**. The instructions can be directed to the CPU **105**, which can subsequently program or otherwise configure the CPU **105** to implement methods of the disclosure.

Examples of operations performed by the CPU **105** can include fetch, decode, execute, and writeback.

[0087] The CPU **105** can be part of a circuit, such as an integrated circuit. One or more other components of the system **101** can be included in the circuit. In some cases, the circuit is an application specific integrated circuit (ASIC).

[0088] The storage unit **115** can store files, such as drivers, libraries and saved programs. The storage unit **115** can store user data, e.g., user preferences and user programs. The computer system **101** in some cases can include one or more additional data storage units that are external to the computer system **101**, such as located on a remote server that is in communication with the computer system **101** through an intranet or the Internet.

[0089] The computer system **101** can communicate with one or more remote computer systems through the network **130**. For instance, the computer system **101** can communicate with a remote computer system of a user. Examples of remote computer systems include personal computers (e.g., portable PC), slate or tablet PC's (e.g., Apple® iPad, Samsung® Galaxy Tab), telephones, Smart phones (e.g., Apple® iPhone, Android-enabled device, Blackberry®), or personal digital assistants. The user can access the computer system **101** via the network **130**.

[0090] Methods as described herein can be implemented by way of machine (e.g., computer processor) executable code stored on an electronic storage location of the computer system **101**, such as, for example, on the memory **110** or electronic storage unit **115**. The machine executable or machine readable code can be provided in the form of software. During use, the code can be executed by the processor **105**. In some cases, the code can be retrieved from the storage unit **115** and stored on the memory **110** for ready access by the processor **105**. In some situations, the electronic storage unit **115** can be precluded, and machine-executable instructions are stored on memory **110**.

[0091] The code can be pre-compiled and configured for use with a machine having a processor adapted to execute the code, or can be compiled during runtime. The code can be supplied in a

programming language that can be selected to enable the code to execute in a pre-compiled or as-compiled fashion.

[0092] Aspects of the systems and methods provided herein, such as the computer system **101**, can be embodied in programming. Various aspects of the technology may be thought of as “products” or “articles of manufacture” typically in the form of machine (or processor) executable code and/or associated data that is carried on or embodied in a type of machine readable medium. Machine-executable code can be stored on an electronic storage unit, such as memory (e.g., read-only memory, random-access memory, flash memory) or a hard disk. “Storage” type media can include any or all of the tangible memory of the computers, processors or the like, or associated modules thereof, such as various semiconductor memories, tape drives, disk drives and the like, which may provide non-transitory storage at any time for the software programming. All or portions of the software may at times be communicated through the Internet or various other telecommunication networks. Such communications, for example, may enable loading of the software from one computer or processor into another, for example, from a management server or host computer into the computer platform of an application server. Thus, another type of media that may bear the software elements includes optical, electrical and electromagnetic waves, such as used across physical interfaces between local devices, through wired and optical landline networks and over various air-links. The physical elements that carry such waves, such as wired or wireless links, optical links or the like, also may be considered as media bearing the software. As used herein, unless restricted to non-transitory, tangible “storage” media, terms such as computer or machine “readable medium” refer to any medium that participates in providing instructions to a processor for execution.

[0093] Hence, a machine readable medium, such as computer-executable code, may take many forms, including but not limited to, a tangible storage medium, a carrier wave medium or physical transmission medium. Non-volatile storage media include, for example, optical or magnetic disks, such as any of the storage devices in any computer(s) or the like, such as may be used to implement the databases, etc. shown in the drawings. Volatile storage media include dynamic memory, such as main memory of such a computer platform. Tangible transmission media include coaxial cables; copper wire and fiber optics, including the wires that comprise a bus within a computer system. Carrier-wave transmission media may take the form of electric or electromagnetic signals, or acoustic or light waves such as those generated during radio frequency (RF) and infrared (IR) data communications. Common forms of computer-readable media therefore include for example: a floppy disk, a flexible disk, hard disk, magnetic tape, any other magnetic medium, a CD-ROM, DVD or DVD-ROM, any other optical medium, punch

cards paper tape, any other physical storage medium with patterns of holes, a RAM, a ROM, a PROM and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave transporting data or instructions, cables or links transporting such a carrier wave, or any other medium from which a computer may read programming code and/or data. Many of these forms of computer readable media may be involved in carrying one or more sequences of one or more instructions to a processor for execution.

[0094] The computer system **101** can include or be in communication with an electronic display **135** that comprises a user interface (UI) **140** for providing, for example a monitor. Examples of UI's include, without limitation, a graphical user interface (GUI) and web-based user interface.

[0095] Methods and systems of the disclosure can be implemented by way of one or more algorithms. An algorithm can be implemented by way of software upon execution by the central processing unit **105**. The algorithm can, for example, Polyphen 2, Sift, Mutation Accessor, Mutation Taster, FATHMM, LRT, MetaLR, or any combination thereof.

[0096] In some cases, as shown in **FIG. 11**, a sample **202** containing a genetic material may be obtained from a subject **201**, such as a human subject. A sample **202** may be subjected to one or more methods as described herein, such as performing an assay. In some cases, an assay may comprise sequencing (such as nanopore sequencing), genotyping, hybridization, amplification, labeling, or any combination thereof. One or more results from a method may be input into a processor **204**. One or more input parameters such as a sample identification, subject identification, sample type, a reference, or other information may be input into a processor **204**. One or more metrics from an assay may be input into a processor **204** such that the processor may produce a result, such as a diagnosis of endometriosis or a recommendation for a treatment. A processor may send a result, an input parameter, a metric, a reference, or any combination thereof to a display **205**, such as a visual display or graphical user interface. A processor **204** may (i) send a result, an input parameter, a metric, or any combination thereof to a server **207**, (ii) receive a result, an input parameter, a metric, or any combination thereof from a server **207**, (iii) or a combination thereof.

[0097] Methods of Detection of Variants

[0098] In some aspects, the disclosure provides methods to detect variants, e.g, detecting a genetic variant in a panel comprising two or more genetic variants defining a minor allele disclosed herein (e.g., in **Table 1**). In some instances, the detecting comprises, DNA sequencing, hybridization with a complementary probe, an oligonucleotide ligation assay, a PCR-based assay, or any combination thereof. In some instances, the panel comprises at least: 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 75, 80, 90, 100, 150, 200, 250, 300,

350, 400, 450, 500, or more genetic variants defining minor alleles disclosed herein (e.g., in **Table 1**). In some instances, the genetic variant to detect or detected has an odds ratio (OR) of at least: 0.1, 1, 1.5, 2, 5, 10, 20, 50, 100, 127, 130, 140, 150, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000, or more. In some embodiments, the OR is at least 127. In some instances, the panel to detect further comprises one or more protein damaging or loss of function variants in one or more genes selected from the group consisting of *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR*, and any combinations thereof. In some instances, the panel further comprises one or more additional variants defining a minor allele listed in **Table 4**.

[0099] In some cases, variants of the disclosure may include single nucleotide polymorphisms (SNPs), insertion deletion polymorphisms (indels), damaging mutation variants, loss of function variants, synonymous mutation variants, nonsynonymous mutation variants, nonsense mutations, recessive markers, splicing/splice-site variants, frameshift mutation, insertions, deletions, genomic rearrangements, stop-gain, stop-loss, Rare Variants (RVs), translocations, inversions, and substitutions.

[0100] Variants for example SNPs are usually preceded and followed by highly conserved sequences that vary in less than 1/100 or 1/1000 members of the population. An individual may be homozygous or heterozygous for an allele at each SNP position. A SNP may, in some instances, be referred to as a "cSNP" to denote that the nucleotide sequence containing the SNP is an amino acid "coding" sequence. A SNP may arise from a substitution of one nucleotide for another at the polymorphic site. Substitutions can be transitions or transversions. A transition is the replacement of one purine nucleotide by another purine nucleotide, or one pyrimidine by another pyrimidine. A transversion is the replacement of a purine by a pyrimidine, or vice versa.

[0101] A synonymous codon change, or silent mutation is one that does not result in a change of amino acid due to the degeneracy of the genetic code. A substitution that changes a codon coding for one amino acid to a codon coding for a different amino acid (i.e., a non-synonymous codon change) is referred to as a missense mutation. A nonsense mutation results in a type of non-synonymous codon change in which a stop codon is formed, thereby leading to premature termination of a polypeptide chain and a truncated protein. A read-through mutation is another type of non-synonymous codon change that causes the destruction of a stop codon, thereby resulting in an extended polypeptide product. An indel that occur in a coding DNA segment gives rise to a frameshift mutation.

[0102] Causative variants are those that produce alterations in gene expression or in the structure and/or function of a gene product, and therefore are predictive of a possible clinical

phenotype. One such class includes SNPs falling within regions of genes encoding a polypeptide product, i.e. cSNPs. These SNPs may result in an alteration of the amino acid sequence of the polypeptide product (i.e., non-synonymous codon changes) and give rise to the expression of a defective or other variant protein. Furthermore, in the case of nonsense mutations, a SNP may lead to premature termination of a polypeptide product. Such variant products can result in a pathological condition, e.g., genetic endometriosis.

[0103] An association study of a variant and a specific disorder involves determining the presence or frequency of the variant allele in biological samples from individuals with the disorder of interest, such as endometriosis, and comparing the information to that of controls (i.e., individuals who do not have the disorder; controls may be also referred to as "healthy" or "normal" individuals) who are for example of similar age and race. The appropriate selection of patients and controls is important to the success of variant association studies. Therefore, a pool of individuals with well-characterized phenotypes is extremely desirable.

[0104] A variant may be screened in tissue samples or any biological sample obtained from an affected individual, and compared to control samples, and selected for its increased (or decreased) occurrence in a specific pathological condition, such as pathologies related to endometriosis. Once a statistically significant association is established between one or more variant(s) and a pathological condition (or other phenotype) of interest, then the region around the variant can optionally be thoroughly screened to identify the causative genetic locus/sequence(s) (e.g., causative variant/mutation, gene, regulatory region, etc.) that influences the pathological condition or phenotype. Association studies may be conducted within the general population and are not limited to studies performed on related individuals in affected families (linkage studies). For diagnostic and prognostic purposes, if a particular variant site is found to be useful for diagnosing a disease, such as endometriosis, other variant sites which are in LD with this variant site would also be expected to be useful for diagnosing the condition. Linkage disequilibrium is described in the human genome as blocks of variants along a chromosome segment that do not segregate independently (i.e., that are non-randomly co-inherited). The starting (5' end) and ending (3' end) of these blocks can vary depending on the criteria used for linkage disequilibrium in a given database, such as the value of D' or r^2 used to determine linkage disequilibrium.

[0105] In some instances, variants can be identified in a study using a whole-genome case-control approach to identify single nucleotide polymorphisms that were closely associated with the development of endometriosis, as well as variants found to be in linkage disequilibrium with (i.e., within the same linkage disequilibrium block as) the endometriosis-associated variants,

which can provide haplotypes (i.e., groups of variants that are co-inherited) to be readily inferred. Thus, the disclosure provides individual variants associated with endometriosis, as well as combinations of variants and haplotypes in genetic regions associated with endometriosis, methods of detecting these polymorphisms in a test sample, methods of determining the risk of an individual of having or developing endometriosis and for clinical sub-classification of endometriosis.

[0106] In some cases, the disclosure provides variants associated with endometriosis, as well as variants that were previously known in the art, but were not previously known to be associated with endometriosis. Accordingly, the disclosure provides novel compositions and methods based on the variants disclosed herein, and also provides novel methods of using the known but previously unassociated variants in methods relating to endometriosis (e.g., for diagnosing endometriosis. etc.).

[0107] In some instances, particular variant alleles of the disclosure can be associated with either an increased risk of having or developing endometriosis, or a decreased risk of having or developing endometriosis. Variant alleles that are associated with a decreased risk may be referred to as "protective" alleles, and variant alleles that are associated with an increased risk may be referred to as "susceptibility" alleles, "risk factors", or "high-risk" alleles. Thus, whereas certain variants can be assayed to determine whether an individual possesses a variant allele that is indicative of an increased risk of having or developing endometriosis (i.e., a susceptibility allele), other variants can be assayed to determine whether an individual possesses a variant allele that is indicative of a decreased risk of having or developing endometriosis (i.e., a protective allele). Similarly, particular variant alleles of the disclosure can be associated with either an increased or decreased likelihood of responding to a particular treatment. The term "altered" may be used herein to encompass either of these two possibilities (e.g., an increased or a decreased risk/likelihood).

[0108] In some instances, nucleic acid molecules may be double-stranded molecules and that reference to a particular site on one strand refers, as well, to the corresponding site on a complementary strand. In defining a variant position, variant allele, or nucleotide sequence, reference to an adenine, a thymine (uridine), a cytosine, or a guanine at a particular site on one strand of a nucleic acid molecule also defines the complementary thymine (uridine), adenine, guanine, or cytosine (respectively) at the corresponding site on a complementary strand of the nucleic acid molecule. Thus, reference may be made to either strand in order to refer to a particular variant position, variant allele, or nucleotide sequence. Probes and primers may be designed to hybridize to either strand and variant genotyping methods disclosed herein may

generally target either strand. Throughout the specification, in identifying a variant position, reference is generally made to the forward or "sense" strand, solely for the purpose of convenience. Since endogenous nucleic acid sequences exist in the form of a double helix (a duplex comprising two complementary nucleic acid strands), it is understood that the variants disclosed herein will have counterpart nucleic acid sequences and variants associated with the complementary "reverse" or "antisense" nucleic acid strand. Such complementary nucleic acid sequences, and the complementary variants present in those sequences, are also included within the scope of the disclosure.

[0109] Disclosed herein are methods for detecting genetic variants in a nucleic acid sample. The method can comprise sequencing a nucleic acid sample obtained from a subject having endometriosis or suspected of having endometriosis using a high throughput method. The high throughput method can comprise nanopore sequencing. The method can comprise detecting one or more genetic variants in a nucleic acid sample, wherein the one or more genetic variants are listed in **Table 1**, **Table 2** or **Table 3**. The nucleic acid sample can comprise RNA. The RNA can comprise mRNA. The nucleic acid sample can comprise DNA. The DNA can comprise cDNA, genomic DNA, sheared DNA, cell free DNA, fragmented DNA, or PCR amplified products produced therefrom, or any combination thereof. The one or more genetic variants can comprise a genetic variant defining a minor allele. The one or more genetic variants can comprise at least about: 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 500, or more genetic variants defining minor alleles. The detection of the one or more genetic variants can have an odds ratio (OR) for endometriosis of at least about: 1.5, 2, 5, 10, 20, 50, 100, or more. The one or more genetic variants can comprise a synonymous mutation, a non-synonymous mutation, a stop-gain mutation, a nonsense mutation, an insertion, a deletion, a splice-site variant, a frameshift mutation, or any combination thereof. The one or more genetic variants can comprise a protein damaging mutation. The one or more genetic variants can comprise a protein damaging or loss of function variant in one or more genes selected from the group consisting of *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR*, and any combinations thereof. The one or more genetic variants can be comprised in *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR* or a combination thereof. The method can comprise detecting one or more additional variants defining a minor allele listed in **Table 4**. The one or more genetic variants can be identified based on a predictive computer algorithm. The one or more genetic variants can be identified based on reference to a database. The method can further comprise identifying a subject as having endometriosis or being at risk of developing endometriosis. The method can comprise

identifying a subject as having endometriosis or being at risk of developing endometriosis with a specificity of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%. The method can comprise identifying a subject as having endometriosis or being at risk of developing endometriosis with a sensitivity of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%. The method can comprise identifying a subject as having endometriosis or being at risk of developing endometriosis with an accuracy of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%. The method can comprise identifying a subject as having endometriosis. The subject can be asymptomatic for endometriosis. In some cases, the subject can have endometriosis and be asymptomatic. The subject can be symptomatic for endometriosis. The subject can be identified as being at risk of developing endometriosis. The method can further comprise administering a therapeutic to a subject. The therapeutic can comprise hormonal therapy, an advanced reproductive technology therapy, a pain managing medication, or any combination thereof. The therapeutic can comprise hormonal contraceptives, gonadotropin-releasing hormone (Gn-RH) agonists, gonadotropin-releasing hormone (Gn-RH) antagonists, progestin, danazol, or any combination thereof. The therapeutic can comprise a pain medication. The pain medication can comprise a nonsteroidal anti-inflammatory drug (NSAID), ibuprofen, naproxen, an opioid, a cannabis-based therapeutic, or any combination thereof. In some cases, the one or more genetic variants are listed in **Table 1**. In some cases, the one or more genetic variants are listed in **Table 2**. In some cases, the one or more genetic variants are listed in **Table 3**. The method can further comprise identifying a subject as having endometriosis-associated infertility or being at risk of developing endometriosis-associated infertility. The method can further comprise administering assisted reproductive technology therapy to a subject. The assisted reproductive technology therapy can comprise *in vitro* fertilization, gamete intrafallopian transfer, or any combination thereof. The method can further comprise administering, intrauterine insemination or ovulation induction. A subject described herein can be a mammal. The mammal can be a human. Nanopore sequencing can be performed with a biological nanopore, a solid state nanopore, or a hybrid nanopore. Methods disclosed herein can detect 1, 5, 10, 15, 20, 30, 50, 60, 100, 80, 90, 100, 200 or more variants disclosed herein. Genetic variants detected herein can indicate endometriosis or a risk of developing endometriosis. In some embodiments, one or more genetic variant listed in **Table 1** are the only genetic variants detected. In some embodiments, one or more genetic variants listed in **Table 2** are the only genetic variant detected. In some embodiments, one or more genetic variants listed in **Table 3** are the only genetic variant detected. In some embodiments, one or more genetic variant listed in

Table 4 are the only genetic variant detected. In other embodiments, one or more genetic variants are detected from two or more of **Table 1**, **Table 2**, **Table 3** and **Table 4**.

[0110] Genotyping Methods

[0111] In some cases, the process of determining which specific nucleotide (i.e., allele) is present at each of one or more variant positions, such as a variant position in a nucleic acid molecule characterized by a variant, is referred to as variant genotyping. The disclosure provides methods of variant genotyping, such as for use in screening for endometriosis or related pathologies, or determining predisposition thereto, or determining responsiveness to a form of treatment, or in genome mapping or variant association analysis, etc.

[0112] Nucleic acid samples can be genotyped to determine which allele(s) is/are present at any given genetic region (e.g., variant position) of interest by methods well known in the art. The neighboring sequence can be used to design variant detection reagents such as oligonucleotide probes, which may optionally be implemented in a kit format. Common variant genotyping methods include, but are not limited to, TaqMan assays, molecular beacon assays, nucleic acid arrays, allele-specific primer extension, allele-specific PCR, arrayed primer extension, homogeneous primer extension assays, primer extension with detection by mass spectrometry, mass spectrometry with or with monoisotopic dNTPs (pyrosequencing, multiplex primer extension sorted on genetic arrays, ligation with rolling circle amplification, homogeneous ligation, OLA, multiplex ligation reaction sorted on genetic arrays, restriction-fragment length polymorphism, single base extension-tag assays, and the Invader assay. Such methods may be used in combination with detection mechanisms such as, for example, luminescence or chemiluminescence detection, fluorescence detection, time-resolved fluorescence detection, fluorescence resonance energy transfer, fluorescence polarization, mass spectrometry, electrospray mass spectrometry, and electrical detection.

[0113] Various methods for detecting polymorphisms can include, but are not limited to, methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA duplexes, comparison of the electrophoretic mobility of variant and wild type nucleic acid molecules, and assaying the movement of polymorphic or wild-type fragments in polyacrylamide gels containing a gradient of denaturant using denaturing gradient gel electrophoresis (DGGE). Sequence variations at specific locations can also be assessed by nuclease protection assays such as RNase and SI protection or chemical cleavage methods.

[0114] In some instances, a variant genotyping can be performed using the TaqMan assay, which is also known as the 5' nuclease assay. The TaqMan assay detects the accumulation of a specific amplified product during PCR. The TaqMan assay utilizes an oligonucleotide probe

labeled with a fluorescent reporter dye and a quencher dye. The reporter dye is excited by irradiation at an appropriate wavelength, it transfers energy to the quencher dye in the same probe via a process called fluorescence resonance energy transfer (FRET). When attached to the probe, the excited reporter dye does not emit a signal. The proximity of the quencher dye to the reporter dye in the intact probe maintains a reduced fluorescence for the reporter. The reporter dye and quencher dye may be at the 5' most and the 3' most ends, respectively, or vice versa. Alternatively, the reporter dye may be at the 5' or 3' most end while the quencher dye is attached to an internal nucleotide, or vice versa. In yet another embodiment, both the reporter and the quencher may be attached to internal nucleotides at a distance from each other such that fluorescence of the reporter is reduced. During PCR, the 5' nuclease activity of DNA polymerase cleaves the probe, thereby separating the reporter dye and the quencher dye and resulting in increased fluorescence of the reporter. Accumulation of PCR product is detected directly by monitoring the increase in fluorescence of the reporter dye. The DNA polymerase cleaves the probe between the reporter dye and the quencher dye only if the probe hybridizes to the target variant-containing template which is amplified during PCR, and the probe is designed to hybridize to the target variant site only if a particular variant allele is present. TaqMan primer and probe sequences can readily be determined using the variant and associated nucleic acid sequence information provided herein. A number of computer programs, such as Primer Express (Applied Biosystems, Foster City, Calif.), can be used to rapidly obtain optimal primer/probe sets. It will be apparent to one of skill in the art that such primers and probes for detecting the variants of the disclosure are useful in diagnostic assays for endometriosis and related pathologies, and can be readily incorporated into a kit format. The disclosure also includes modifications of the Taqman assay well known in the art such as the use of Molecular Beacon probes and other variant formats.

[0115] In some instances, a method for genotyping the variants can be the use of two oligonucleotide probes in an OLA. In this method, one probe hybridizes to a segment of a target nucleic acid with its 3' most end aligned with the variant site. A second probe hybridizes to an adjacent segment of the target nucleic acid molecule directly 3' to the first probe. The two juxtaposed probes hybridize to the target nucleic acid molecule, and are ligated in the presence of a linking agent such as a ligase if there is perfect complementarity between the 3' most nucleotide of the first probe with the variant site. If there is a mismatch, ligation would not occur. After the reaction, the ligated probes are separated from the target nucleic acid molecule, and detected as indicators of the presence of a variant.

[0116] In some instances, a method for variant genotyping is based on mass spectrometry. Mass spectrometry takes advantage of the unique mass of each of the four nucleotides of DNA. variants can be unambiguously genotyped by mass spectrometry by measuring the differences in the mass of nucleic acids having alternative variant alleles. MALDI-TOF (Matrix Assisted Laser Desorption Ionization-Time of Flight) mass spectrometry technology is exemplary for extremely precise determinations of molecular mass, such as variants. Numerous approaches to variant analysis have been developed based on mass spectrometry. Exemplary mass spectrometry-based methods of variant genotyping include primer extension assays, which can also be utilized in combination with other approaches, such as traditional gel-based formats and microarrays.

[0117] In some instances, a method for genotyping the variants of the disclosure is the use of electrospray mass spectrometry for direct analysis of an amplified nucleic acid. In this method, in one aspect, an amplified nucleic acid product may be isotopically enriched in an isotope of oxygen (O), carbon (C), nitrogen (N) or any combination of those elements. In an exemplary embodiment the amplified nucleic acid is isotopically enriched to a level of greater than 99.9% in the elements of O¹⁶, C¹² and N¹⁴. The amplified isotopically enriched product can then be analyzed by electrospray mass spectrometry to determine the nucleic acid composition and the corresponding variant genotyping. Isotopically enriched amplified products result in a corresponding increase in sensitivity and accuracy in the mass spectrum. In another aspect of this method an amplified nucleic acid that is not isotopically enriched can also have composition and variant genotype determined by electrospray mass spectrometry.

[0118] In some instances, variants can be scored by direct DNA sequencing. The nucleic acid sequences of the disclosure enable one of ordinary skill in the art to readily design sequencing primers for such automated sequencing procedures. Commercial instrumentation, such as the Applied Biosystems 377, 3100, 3700, 3730, and 3730.times.1 DNA Analyzers (Foster City, Calif.), is commonly used in the art for automated sequencing.

[0119] Variant genotyping can include the steps of, for example, collecting a biological sample from a human subject (e.g., sample of tissues, cells, fluids, secretions, etc.), isolating nucleic acids (e.g., genomic DNA, mRNA or both) from the cells of the sample, contacting the nucleic acids with one or more primers which specifically hybridize to a region of the isolated nucleic acid containing a target variant under conditions such that hybridization and amplification of the target nucleic acid region occurs, and determining the nucleotide present at the variant position of interest, or, in some assays, detecting the presence or absence of an amplification product (assays can be designed so that hybridization and/or amplification will only occur if a particular variant allele is present or absent). In some assays, the size of the amplification product is

detected and compared to the length of a control sample; for example, deletions and insertions can be detected by a change in size of the amplified product compared to a normal genotype.

[0120] In some instances, a variant genotyping can be used in applications that include, but are not limited to, variant-endometriosis association analysis, endometriosis predisposition screening, endometriosis diagnosis, endometriosis prognosis, endometriosis progression monitoring, determining therapeutic strategies based on an individual's genotype, and stratifying a patient population for clinical trials for a treatment such as minimally invasive device for the treatment of endometriosis.

[0121] Analysis of Genetic Association Between Variants and Phenotypic Traits

[0122] In some cases, genotyping for endometriosis diagnosis, endometriosis predisposition screening, endometriosis prognosis and endometriosis treatment and other uses described herein, can rely on initially establishing a genetic association between one or more specific variants and the particular phenotypic traits of interest.

[0123] In some instances, in a genetic association study, the cause of interest to be tested is a certain allele or a variant or a combination of alleles or a haplotype from several variants. Thus, tissue specimens (e.g., saliva) from the sampled individuals may be collected and genomic DNA genotyped for the variant(s) of interest. In addition to the phenotypic trait of interest, other information such as demographic (e.g., age, gender, ethnicity, etc.), clinical, and environmental information that may influence the outcome of the trait can be collected to further characterize and define the sample set. Specifically, in an endometriosis genetic association study, clinical information such as body mass index, age and diet may be collected. In many cases, these factors are known to be associated with diseases and/or variant allele frequencies. There are likely gene-environment and/or gene-gene interactions as well. Analysis methods to address gene-environment and gene-gene interactions (for example, the effects of the presence of both susceptibility alleles at two different genes can be greater than the effects of the individual alleles at two genes combined) are discussed below.

[0124] In some instances, after all the relevant phenotypic and genotypic information has been obtained, statistical analyses are carried out to determine if there is any significant correlation between the presence of an allele or a genotype with the phenotypic characteristics of an individual. For example, data inspection and cleaning are first performed before carrying out statistical tests for genetic association. Epidemiological and clinical data of the samples can be summarized by descriptive statistics with tables and graphs. Data validation is for example performed to check for data completion, inconsistent entries, and outliers. Chi-squared tests may then be used to check for significant differences between cases and controls for discrete and

continuous variables, respectively. To ensure genotyping quality, Hardy-Weinberg disequilibrium tests can be performed on cases and controls separately. Significant deviation from Hardy-Weinberg equilibrium (HWE) in both cases and controls for individual markers can be indicative of genotyping errors. If HWE is violated in a majority of markers, it is indicative of population substructure that should be further investigated. Moreover, Hardy-Weinberg disequilibrium in cases only can indicate genetic association of the markers with the disease of interest.

[0125] In some instances, to test whether an allele of a single variant is associated with the case or control status of a phenotypic trait, one skilled in the art can compare allele frequencies in cases and controls. Standard chi-squared tests and Fisher exact tests can be carried out on a 2 x 2 table (2 variant alleles x 2 outcomes in the categorical trait of interest). To test whether genotypes of a variant are associated, chi-squared tests can be carried out on a 3 x 2 table (3 genotypes x 2 outcomes). Score tests are also carried out for genotypic association to contrast the three genotypic frequencies (major homozygotes, heterozygotes and minor homozygotes) in cases and controls, and to look for trends using 3 different modes of inheritance, namely dominant (with contrast coefficients 2, -1, -1), additive (with contrast coefficients 1, 0, -1) and recessive (with contrast coefficients 1, 1, -2). Odds ratios for minor versus major alleles, and odds ratios for heterozygote and homozygote variants versus the wild type genotypes are calculated with the desired confidence limits, usually 95%. In the present study a software algorithm, PLINK, has been applied to automate the calculation of Hardy-Weinberg equilibrium, chi-square, p-values and odds-ratios for very large numbers of variants and Case-Control individuals simultaneously.

[0126] In some instances, in order to control for confounding effects and to test for interactions a stepwise multiple logistic regression analysis using statistical packages such as SAS or R may be performed. Logistic regression is a model-building technique in which the best fitting and most parsimonious model is built to describe the relation between the dichotomous outcome (for instance, getting a certain endometriosis or not) and a set of independent variables (for instance, genotypes of different associated genes, and the associated demographic and environmental factors). The most common model is one in which the logit transformation of the odds ratios is expressed as a linear combination of the variables (main effects) and their cross-product terms (interactions). To test whether a certain variable or interaction is significantly associated with the outcome, coefficients in the model are first estimated and then tested for statistical significance of their departure from zero.

[0127] In some instances, in addition to performing association tests one marker at a time, haplotype association analysis may also be performed to study a number of markers that are closely linked together. Haplotype association tests can have better power than genotypic or allelic association tests when the tested markers are not the disease-causing mutations themselves but are in linkage disequilibrium with such mutations. The test will even be more powerful if the endometriosis is indeed caused by a combination of alleles on a haplotype. In order to perform haplotype association effectively, marker-marker linkage disequilibrium measures, both D' and r^2 , are typically calculated for the markers within a gene to elucidate the haplotype structure. Variants within a gene can be organized in block pattern, and a high degree of linkage disequilibrium exists within blocks and very little linkage disequilibrium exists between blocks. Haplotype association with the endometriosis status can be performed using such blocks once they have been elucidated.

[0128] Haplotype association tests can be carried out in a similar fashion as the allelic and genotypic association tests. Each haplotype in a gene is analogous to an allele in a multi-allelic marker. One skilled in the art can either compare the haplotype frequencies in cases and controls or test genetic association with different pairs of haplotypes. Score tests can be done on haplotypes using the program "haplo.score". In that method, haplotypes are first inferred by EM algorithm and score tests are carried out with a generalized linear model (GLM) framework that allows the adjustment of other factors.

[0129] In some instances, an important decision in the performance of genetic association tests is the determination of the significance level at which significant association can be declared when the p-value of the tests reaches that level. In an exploratory analysis where positive hits will be followed up in subsequent confirmatory testing, an unadjusted p-value <0.1 (a significance level on the lenient side) may be used for generating hypotheses for significant association of a variant with certain phenotypic characteristics of a endometriosis. It is exemplary that a p-value <0.05 (a significance level traditionally used in the art) is achieved in order for a variant to be considered to have an association with a endometriosis. It is more exemplary that a p-value <0.01 (a significance level on the stringent side) is achieved for an association to be declared. Permutation tests to control for the false discovery rates, FDR, can further be employed. Such methods to control for multiplicity would be exemplary when the tests are dependent and controlling for false discovery rates is sufficient as opposed to controlling for the experiment-wise error rates.

[0130] In some instances, since both genotyping and endometriosis status classification can involve errors, sensitivity analyses may be performed to see how odds ratios and p-values would change upon various estimates on genotyping and endometriosis classification error rates.

[0131] Once individual risk factors, genetic or non-genetic, have been found for the predisposition to endometriosis, the next step can be to set up a classification/prediction scheme to predict the category (for instance, endometriosis or no endometriosis) that an individual will be in depending on his genotypes of associated variants and other non-genetic risk factors. Logistic regression for discrete trait and linear regression for continuous trait are standard techniques for such tasks. Moreover, other techniques can also be used for setting up classification. Such techniques include, but are not limited to, MART, CART, neural network, and discriminant analyses that are suitable for use in comparing the performance of different methods.

[0132] Endometriosis Diagnosis and Predisposition Screening

[0133] In some cases, information on association/correlation between genotypes and endometriosis-related phenotypes can be exploited in several ways. For example, in the case of a highly statistically significant association between one or more variants with predisposition to a disease for which treatment is available, detection of such a genotype pattern in an individual may justify particular treatment, or at least the institution of regular monitoring of the individual. In the case of a weaker but still statistically significant association between a variant and a human disease, immediate therapeutic intervention or monitoring may not be justified after detecting the susceptibility allele or variant.

[0134] The variants disclosed herein may contribute to endometriosis in an individual in different ways. Some polymorphisms occur within a protein coding sequence and contribute to endometriosis phenotype by affecting protein structure. Other polymorphisms occur in noncoding regions but may exert phenotypic effects indirectly via influence on, for example, replication, transcription, and/or translation. A single variant may affect more than one phenotypic trait. Likewise, a single phenotypic trait may be affected by multiple variants in different genes.

[0135] The variants disclosed herein may contribute to endometriosis in an individual in different ways. Some polymorphisms occur within a protein coding sequence and contribute to endometriosis phenotype by affecting protein structure. Other polymorphisms occur in noncoding regions but may exert phenotypic effects indirectly via influence on, for example, replication, transcription, and/or translation. A single variant may affect more than one

phenotypic trait. Likewise, a single phenotypic trait may be affected by multiple variants in different genes.

[0136] Haplotypes can be particularly useful in that, for example, fewer variants can be genotyped to determine if a particular genomic region harbors a locus that influences a particular phenotype, such as in linkage disequilibrium-based variant association analysis.

[0137] Linkage disequilibrium (LD) can refer to the co-inheritance of alleles (e.g., alternative nucleotides) at two or more different variant sites at frequencies greater than would be expected from the separate frequencies of occurrence of each allele in a given population. The expected frequency of co-occurrence of two alleles that are inherited independently is the frequency of the first allele multiplied by the frequency of the second allele. Alleles that co-occur at expected frequencies are said to be in "linkage equilibrium". In contrast, LD can refer to any non-random genetic association between allele(s) at two or more different variant sites, which is generally due to the physical proximity of the two loci along a chromosome. LD can occur when two or more variant sites are in close physical proximity to each other on a given chromosome and therefore alleles at these variant sites will tend to remain unseparated for multiple generations with the consequence that a particular nucleotide (allele) at one variant site will show a non-random association with a particular nucleotide (allele) at a different variant site located nearby. Hence, genotyping one of the variant sites will give almost the same information as genotyping the other variant site that is in LD.

[0138] For diagnostic purposes, if a particular variant site is found to be useful for diagnosing endometriosis, then the skilled artisan would recognize that other variant sites which are in LD with this variant site would also be useful for diagnosing the condition. Various degrees of LD can be encountered between two or more variants with the result being that some variants are more closely associated (i.e., in stronger LD) than others. Furthermore, the physical distance over which LD extends along a chromosome differs between different regions of the genome, and therefore the degree of physical separation between two or more variant sites necessary for LD to occur can differ between different regions of the genome.

[0139] For diagnostic applications, polymorphisms (e.g., variants and/or haplotypes) that are not the actual disease-causing (causative) polymorphisms, but are in LD with such causative polymorphisms, are also useful. In such instances, the genotype of the polymorphism(s) that is/are in LD with the causative polymorphism is predictive of the genotype of the causative polymorphism and, consequently, predictive of the phenotype (e.g., endometriosis) that is influenced by the causative variant(s). Thus, polymorphic markers that are in LD with causative

polymorphisms are useful as diagnostic markers, and are particularly useful when the actual causative polymorphism(s) is/are unknown.

[0140] The contribution or association of particular variants and/or variant haplotypes with endometriosis phenotypes, such as endometriosis, can enable the variants of the disclosure to be used to develop superior diagnostic tests capable of identifying individuals who express a detectable trait, such as endometriosis, as the result of a specific genotype, or individuals whose genotype places them at an increased or decreased risk of developing a detectable trait at a subsequent time as compared to individuals who do not have that genotype. As described herein, diagnostics may be based on a single variant or a group of variants. In some instances, combined detection of a plurality of variations, for example about 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 24, 25, 30, 32, 35, 40, 45, 48, 50, 55, 60, 64, 70, 75, 80, 85, 80, 96, 100, or any other number in-between, or more, of the variants provided herein can increase the probability of an accurate diagnosis. To further increase the accuracy of diagnosis or predisposition screening, analysis of the variants of the disclosure can be combined with that of other polymorphisms or other risk factors of endometriosis, such as gender and age.

[0141] In some instances, the method herein can indicate a certain increased (or decreased) degree or likelihood of developing the endometriosis based on statistically significant association results. This information can be valuable to initiate earlier preventive treatments or to allow an individual carrying one or more significant variants or variant haplotypes to regularly scheduled physical exams to monitor for the appearance or change of their endometriosis in order to identify and begin treatment of the endometriosis at an early stage.

[0142] The diagnostic techniques herein may employ a variety of methodologies to determine whether a test subject has a variant or a variant pattern associated with an increased or decreased risk of developing a detectable trait or whether the individual suffers from a detectable trait as a result of a particular polymorphism/mutation, including, for example, methods which enable the analysis of individual chromosomes for haplotyping, family studies, single sperm DNA analysis, or somatic hybrids. The trait analyzed using the diagnostics of the disclosure may be any detectable trait that is commonly observed in pathologies and disorders related to endometriosis.

[0143] Another aspect of the disclosure relates to a method of determining whether an individual is at risk (or less at risk) of developing one or more traits or whether an individual expresses one or more traits as a consequence of possessing a particular trait-causing or trait-influencing allele. These methods generally involve obtaining a nucleic acid sample from an individual and assaying the nucleic acid sample to determine which nucleotide(s) is/are present at one or more variant positions, wherein the assayed nucleotide(s) is/are indicative of an

increased or decreased risk of developing the trait or indicative that the individual expresses the trait as a result of possessing a particular trait-causing or trait-influencing allele.

[0144] The variants herein can be used to identify novel therapeutic targets for endometriosis. For example, genes containing the disease-associated variants ("variant genes") or their products, as well as genes or their products that are directly or indirectly regulated by or interacting with these variant genes or their products, can be targeted for the development of therapeutics that, for example, treat the endometriosis or prevent or delay endometriosis onset. The therapeutics may be composed of, for example, small molecules, proteins, protein fragments or peptides, antibodies, nucleic acids, or their derivatives or mimetics which modulate the functions or levels of the target genes or gene products.

[0145] The variants/haplotypes herein can be useful for improving many different aspects of the drug development process. For example, individuals can be selected for clinical trials based on their variant genotype. Individuals with variant genotypes that indicate that they are most likely to respond to or most likely to benefit from a device or a drug can be included in the trials and those individuals whose variant genotypes indicate that they are less likely to or would not respond to a device or a drug, or suffer adverse reactions, can be eliminated from the clinical trials. This not only improves the safety of clinical trials, but also will enhance the chances that the trial will demonstrate statistically significant efficacy. Furthermore, the variants of the disclosure may explain why certain previously developed devices or drugs performed poorly in clinical trials and may help identify a subset of the population that would benefit from a drug that had previously performed poorly in clinical trials, thereby "rescuing" previously developed therapeutic treatment methods or drugs, and enabling the methods or drug to be made available to a particular endometriosis patient population that can benefit from it.

[0146] Detection Kits and Systems

[0147] In some instances, based on a variant such as SNP or indels and associated sequence information disclosed herein, detection reagents can be developed and used to assay any variant of the disclosure individually or in combination, and such detection reagents can be readily incorporated into one of the established kit or system formats which are well known in the art. The terms "kits" and "systems" can refer to such things as combinations of multiple variant detection reagents, or one or more variant detection reagents in combination with one or more other types of elements or components (e.g., other types of biochemical reagents, containers, packages such as packaging intended for commercial sale, substrates to which variant detection reagents are attached, electronic hardware components, etc.). Accordingly, the disclosure further provides variant detection kits and systems, including but not limited to, packaged probe and

primer sets (e.g., TaqMan probe/primer sets), arrays/microarrays of nucleic acid molecules, and beads that contain one or more probes, primers, or other detection reagents for detecting one or more variants of the disclosure. The kits/systems can optionally include various electronic hardware components; for example, arrays ("DNA chips") and microfluidic systems ("lab-on-a-chip" systems) provided by various manufacturers typically comprise hardware components. Other kits/systems (e.g., probe/primer sets) may not include electronic hardware components, but may be comprised of, for example, one or more variant detection reagents (along with, optionally, other biochemical reagents) packaged in one or more containers.

[0148] In some instances, provided herein is a kit comprising one or more variant detection agents, and methods for detecting the variants disclosed herein by employing detection reagents and optionally a questionnaire of non-genetic clinical factors. In some instances, provided herein is a method of identifying an individual having an increased or decreased risk of developing endometriosis by detecting the presence or absence of a variant allele disclosed herein. In some instances, provided herein is a method for diagnosis of endometriosis by detecting the presence or absence of a variant allele disclosed herein is provided. In some instances, provided herein is a method for predicting endometriosis sub-classification by detecting the presence or absence of a variant allele. In some instances, the questionnaire would be completed by a medical professional based on medical history physical exam or other clinical findings. In some instances, the questionnaire would include any other non-genetic clinical factors known to be associated with the risk of developing endometriosis. In some instances, a reagent for detecting a variant in the context of its naturally-occurring flanking nucleotide sequences (which can be, e.g., either DNA or mRNA) is provided. In some instances, the reagent may be in the form of a hybridization probe or an amplification primer that is useful in the specific detection of a variant of interest. In some instances, a variant can be a genetic polymorphism having a Minor Allele Frequency (MAF) of at least 1% in a population (such as for instance the Caucasian population or the CEU population) and an RV is understood to be a genetic polymorphism having a Minor Allele Frequency (MAF) of less than 1% in a population (such as for instance the Caucasian population or the CEU population).

[0149] In some instances, a detection kit can contain one or more detection reagents and other components (e.g., a buffer, enzymes such as DNA polymerases or ligases, chain extension nucleotides such as deoxynucleotide triphosphates, and in the case of Sanger-type DNA sequencing reactions, chain terminating nucleotides, positive control sequences, negative control sequences, and the like) necessary to carry out an assay or reaction, such as amplification and/or detection of a variant-containing nucleic acid molecule. A kit may further contain means for

determining the amount of a target nucleic acid, and means for comparing the amount with a standard, and can comprise instructions for using the kit to detect the variant-containing nucleic acid molecule of interest. In one embodiment of the disclosure, kits are provided which contain the necessary reagents to carry out one or more assays to detect one or more variants disclosed herein. In an exemplary embodiment of the disclosure, the detection kits/systems can be in the form of nucleic acid arrays, or compartmentalized kits, including microfluidic/lab-on-a-chip systems.

[0150] In some instances, variant detection kits/systems may contain, for example, one or more probes, or pairs of probes, that hybridize to a nucleic acid molecule at or near each target variant position. Multiple pairs of allele-specific probes may be included in the kit/system to simultaneously assay large numbers of variants, at least one of which is a variant of the disclosure. In some kits/systems, the allele-specific probes are immobilized to a substrate such as an array or bead. For example, the same substrate can comprise allele-specific probes for detecting at least 1; 10; 100; 1000; 10,000; 100,000; 500,000 (or any other number in-between) or substantially all of the variants disclosed herein.

[0151] The terms "arrays," "microarrays," and "DNA chips" are used herein interchangeably to refer to an array of distinct polynucleotides affixed to a substrate, such as glass, plastic, paper, nylon or other type of membrane, filter, chip, or any other suitable solid support. The polynucleotides can be synthesized directly on the substrate, or synthesized separate from the substrate and then affixed to the substrate.

[0152] In some instances, any number of probes, such as allele-specific probes, may be implemented in an array, and each probe or pair of probes can hybridize to a different variant position. In the case of polynucleotide probes, they can be synthesized at designated areas (or synthesized separately and then affixed to designated areas) on a substrate using a light-directed chemical process. Each DNA chip can contain, for example, thousands to millions of individual synthetic polynucleotide probes arranged in a grid-like pattern and miniaturized (e.g., to the size of a dime). For example, probes are attached to a solid support in an ordered, addressable array.

[0153] In some instances, a microarray can be composed of a large number of unique, single-stranded polynucleotides fixed to a solid support. Typical polynucleotides are for example about 6-60 nucleotides in length, more for example about 15-30 nucleotides in length, and most for example about 18-25 nucleotides in length. For certain types of microarrays or other detection kits/systems, it may be suitable to use oligonucleotides that are only about 7-20 nucleotides in length. In other types of arrays, such as arrays used in conjunction with chemiluminescent detection technology, exemplary probe lengths can be, for example, about 15-80 nucleotides in

length, for example about 50-70 nucleotides in length, more for example about 55-65 nucleotides in length, and most for example about 60 nucleotides in length. The microarray or detection kit can contain polynucleotides that cover the known 5' or 3' sequence of the target variant site, sequential polynucleotides that cover the full-length sequence of a gene/transcript; or unique polynucleotides selected from particular areas along the length of a target gene/transcript sequence, particularly areas corresponding to one or more variants disclosed herein. Polynucleotides used in the microarray or detection kit can be specific to a variant or variants of interest (e.g., specific to a particular SNP allele at a target SNP site, or specific to particular SNP alleles at multiple different SNP sites), or specific to a polymorphic gene/transcript or genes/transcripts of interest.

[0154] In some instances, hybridization assays based on polynucleotide arrays rely on the differences in hybridization stability of the probes to perfectly matched and mismatched target sequence variants. For variant genotyping, it is generally suitable that stringency conditions used in hybridization assays are high enough such that nucleic acid molecules that differ from one another at as little as a single variant position can be differentiated (e.g., typical variant hybridization assays are designed so that hybridization will occur only if one particular nucleotide is present at a variant position, but will not occur if an alternative nucleotide is present at that variant position). Such high stringency conditions may be suitable when using, for example, nucleic acid arrays of allele-specific probes for variant detection. In some instances, the arrays are used in conjunction with chemiluminescent detection technology.

[0155] In some instances, a nucleic acid array can comprise an array of probes of about 15-25 nucleotides in length. In further embodiments, a nucleic acid array can comprise any number of probes, in which at least one probe is capable of detecting one or more variants disclosed herein and/or at least one probe comprises a fragment of one of the sequences selected from the group consisting of those disclosed herein, and sequences complementary thereto, said fragment comprising at least about 8 consecutive nucleotides, for example 10, 12, 15, 16, 18, 20, more for example 22, 25, 30, 40, 47, 50, 55, 60, 65, 70, 80, 90, 100, or more consecutive nucleotides (or any other number in-between) and containing (or being complementary to) a variant. In some embodiments, the nucleotide complementary to the variant site is within 5, 4, 3, 2, or 1 nucleotide from the center of the probe, more for example at the center of said probe.

[0156] In some instances, using such arrays or other kits/systems, the disclosure provides methods of identifying the variants disclosed herein in a test sample. Such methods typically involve incubating a test sample of nucleic acids with an array comprising one or more probes corresponding to at least one variant position of the disclosure, and assaying for binding of a

nucleic acid from the test sample with one or more of the probes. Conditions for incubating a variant detection reagent (or a kit/system that employs one or more such variant detection reagents) with a test sample vary. Incubation conditions depend on such factors as the format employed in the assay, the detection methods employed, and the type and nature of the detection reagents used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization, amplification and array assay formats can readily be adapted to detect the variants disclosed herein.

[0157] In some instances, a detection kit/system may include components that are used to prepare nucleic acids from a test sample for the subsequent amplification and/or detection of a variant-containing nucleic acid molecule. Such sample preparation components can be used to produce nucleic acid extracts, including DNA and/or RNA, extracts from any bodily fluids. In an exemplary embodiment of the disclosure, the bodily fluid is blood, saliva or buccal swabs. The test samples used in the above-described methods will vary based on such factors as the assay format, nature of the detection method, and the specific tissues, cells or extracts used as the test sample to be assayed. Methods of preparing nucleic acids are well known in the art and can be readily adapted to obtain a sample that is compatible with the system utilized. In some instances, in addition to reagents for preparation of nucleic acids and reagents for detection of one of the variants of this disclosure, the kit may include a questionnaire inquiring about non-genetic clinical factors such as age, gender, or any other non-genetic clinical factors known to be associated with endometriosis.

[0158] In some instances, a form of kit can be a compartmentalized kit. A compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include, for example, small glass containers, plastic containers, strips of plastic, glass or paper, or arraying material such as silica. Such containers allow one to efficiently transfer reagents from one compartment to another compartment such that the test samples and reagents are not cross-contaminated, or from one container to another vessel not included in the kit, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another or to another vessel. Such containers may include, for example, one or more containers which will accept the test sample, one or more containers which contain at least one probe or other variant detection reagent for detecting one or more variants of the disclosure, one or more containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and one or more containers which contain the reagents used to reveal the presence of the bound probe or other variant detection reagents. The kit can optionally further comprise compartments and/or reagents for, for example, nucleic acid amplification or other enzymatic

reactions such as primer extension reactions, hybridization, ligation, electrophoresis (for example capillary electrophoresis), mass spectrometry, and/or laser-induced fluorescent detection. The kit may also include instructions for using the kit. In such microfluidic devices, the containers may be referred to as, for example, microfluidic "compartments", "chambers", or "channels".

[0159] In some instances, microfluidic devices, which may also be referred to as "lab-on-a-chip" systems, biomedical micro-electro-mechanical systems (bioMEMs), or multicomponent integrated systems, are exemplary kits/systems of the disclosure for analyzing variants. Such systems miniaturize and compartmentalize processes such as probe/target hybridization, nucleic acid amplification, and capillary electrophoresis reactions in a single functional device. Such microfluidic devices typically utilize detection reagents in at least one aspect of the system, and such detection reagents may be used to detect one or more variants of the disclosure. One example of a microfluidic system is the integration of PCR amplification and capillary electrophoresis in chips. Exemplary microfluidic systems comprise a pattern of microchannels designed onto a glass, silicon, quartz, or plastic wafer included on a microchip. The movements of the samples may be controlled by electric, electroosmotic or hydrostatic forces applied across different areas of the microchip to create functional microscopic valves and pumps with no moving parts. Varying the voltage can be used as a means to control the liquid flow at intersections between the micro-machined channels and to change the liquid flow rate for pumping across different sections of the microchip. In some instances, for genotyping variants, a microfluidic system may integrate, for example, nucleic acid amplification, primer extension, capillary electrophoresis, and a detection method such as laser induced fluorescence detection.

[0160] Detection Kits and Systems

[0161] In some instances, based on a variant, detection reagents can be developed and used to assay any variant of the disclosure individually or in combination, and such detection reagents can be readily incorporated into one of the established kit or system formats which are well known in the art. The terms "kits" and "systems" can refer to such things as combinations of multiple variant detection reagents, or one or more variant detection reagents in combination with one or more other types of elements or components (e.g., other types of biochemical reagents, containers, packages such as packaging intended for commercial sale, substrates to which variant detection reagents are attached, electronic hardware components, etc.). Accordingly, the disclosure further provides variant detection kits and systems, including but not limited to, packaged probe and primer sets (e.g., TaqMan probe/primer sets), arrays/microarrays of nucleic acid molecules, and beads that contain one or more probes, primers, or other detection reagents for detecting one or more variants of the disclosure. The kits/systems can optionally include various electronic hardware components; for example, arrays ("DNA chips") and microfluidic systems ("lab-on-a-chip" systems) provided by various manufacturers may comprise

hardware components. Other kits/systems (e.g., probe/primer sets) may not include electronic hardware components, but may be comprised of, for example, one or more variant detection reagents (along with, optionally, other biochemical reagents) packaged in one or more containers.

[0162] Methods of Treatment

[0163] In some aspects, disclosed herein is a method of treating a select subject in need thereof. The use of these genetic markers can allow selection of subjects for clinical trials involving novel treatment methods. In some cases, genetic markers disclosed herein can be used for early diagnosis and prognosis of endometriosis, as well as early clinical intervention to mitigate progression of the disease. In some instances, genetic markers disclosed herein can be used to predict endometriosis and endometriosis progression, for example in treatment decisions for individuals who are recognized as having endometriosis.

[0164] In some cases, a treatment disclosed herein includes one or more of: reducing the frequency and/or severity of symptoms, elimination of symptoms and/or their underlying cause, and improvement or remediation of damage. For example, treatment of endometriosis includes, relieving the pain experienced by a woman suffering from endometriosis, and/or causing the regression or disappearance of endometriotic lesions.

[0001] In some cases, the treatment can be an advanced reproductive technology therapy such as in vitro fertilization (IVF); a hormonal treatment; progestogen; progestin; an oral contraceptive; a hormonal contraceptive; danocrine; gentrinone; a gonadotrophin releasing hormone agonist; Lupron; danazol; an aromatase inhibitor; pentoxifylline; surgical treatment; laparoscopy; cauterization; or cystectomy. In some instances, the progestogen can be progesterone, desogestrel, etonogestrel, gestodene, levonorgestrel, medroxyprogesterone, norethisterone, norgestimate, megestrol, megestrol acetate, norgestrel, a pharmaceutically acceptable salt thereof (e.g., acetate), or any combination thereof. In some instances, a therapeutic used herein is selected from progestins, estrogens, antiestrogens, and antiprogestins, for example micronized danazol in a micro- or nanoparticulate formulation.

[0002] In some cases, a method of treatment disclosed herein comprises direct administration into or within an endometriotic lesion in a subject suffering from endometriosis of a pharmaceutical composition comprising a therapeutic disclosed herein. In some instances, the therapeutic is micronized in a suspension, e.g., non-oil based suspension. In some embodiments, the suspension comprises water, sodium sulfate, a quaternary ammonium wetting agent, glycerol, propylene glycol, polyethylene glycol, polypropylene glycol, a hydrophilic colloid, or any combination thereof.

[0165] The term “effective amount,” as used herein, can refer to a sufficient amount of a therapeutic being administered which relieve to some extent one or more of the symptoms of the

disease or condition being treated. The result can be reduction and/or alleviation of the signs, symptoms, or causes of a disease, or any other desired alteration of a biological system. A therapeutic can be administered for prophylactic, enhancing, and/or therapeutic treatments. An appropriate “effective” amount in any individual case can be determined using techniques, such as a dose escalation study.

[0166] A treatment can comprise administering a therapeutic to a subject, intralesionally, transvaginally, intravenously, subcutaneously, intramuscularly, by inhalation, dermally, intra-articular injection, orally, intrathecally, transdermally, intranasally, via a peritoneal route, or directly onto or into a lesion/site, e.g., via endoscopically, open surgical administration, or injection route of application. In some instances, intralesional administration can mean administration into or within a pathological area. Administration can be effected by injection into a lesion and/or by instillation into a pre-existing cavity, such as in endometrioma. With reference to treatments for endometriosis provided herein, intralesional administration can refer to treatment within endometriotic tissue or a cyst formed by such tissue, such as by injection into a cyst. In some instances, intralesional administration can include administration into tissue in such close proximity to the endometriotic tissue such that the progestogen acts directly on the endometriotic tissue. In some instances, intralesional administration may or may not include administration to tissue remote from the endometriotic tissue that the progestogen acts on the endometriotic tissue through systemic circulation. In some instances, intralesional administration or delivery includes transvaginal, endoscopic or open surgical administration including, but are not limited to, via laparotomy. In some instances, transvaginal administration can refer to all procedures, including drug delivery, performed through the vagina, including intravaginal delivery and transvaginal sonography (ultrasonography through the vagina).

[0167] In some instances, administration is by injection into the endometriotic tissue or into a cyst formed by such tissue; or into tissue immediately surrounding the endometriotic tissue in such proximity that the progestogen acts directly on the endometriotic tissue. In some embodiments, the tissue is visualized, for example laparoscopically or by ultrasound, and the progestogen is administered by intralesional (intracystic) injection by, for example direct visualization under ultrasound guidance or by any other suitable methods. A suitable amount of the therapeutic, e.g., progestogen expressed in terms of progesterone of about 1-2 gm per lesion/cyst, can be applied. Precise quantity generally is determined on case to case basis, depending upon parameters, such as the size of the endometriotic tissue mass, the mode of the administration, and the number and time intervals between treatments.

[0168] In some instances, methods herein can comprise intralesional delivery of the medicaments into the lesion. Intralesional delivery includes, for example, transvaginal, endoscopic or open surgical administration including via laparotomy. Delivery can be effected, for example, through a needle or needle like device by injection or a similar injectable or syringe-like device that can be delivered into the lesion, such as transvaginally, endoscopically or by open surgical administration including via laparotomy. In some embodiments, the method includes intravaginal and transvaginal delivery. For intravaginal/transvaginal delivery an ultrasound probe can be used to guide delivery of the needle from the vagina into lesions such as endometriomas and utero sacral nodules. Under ultrasound guidance the needle tip is placed in the lesion, the contents of the lesion aspirated if necessary and the formulation is injected into the lesion. In an exemplary delivery system a 17 to 20 gauge needle can be used for injection of the drug. Such system can be used for intralesional delivery including, but not limited to, transvaginal, endoscopic or open surgical administration including via laparotomy. For treatment of endometrioma 17 or 18 gauge needles are used under ultrasound guidance for aspiration of the thick contents of the lesion and delivery of the formulation. The length of the needle used depends on the depth of the lesion. Pre-loaded syringes and other administration systems, which obviate the need for reloading the drug can be used.

[0169] In some cases, a therapeutic (e.g., an active agent) used herein can be a solution, a suspension, liquid, a paste, aqueous, non-aqueous fluid, semi-solids, colloid, gel, lotion, cream, solid (e.g., tablet, powder, pellet, particulate, capsule, packet), or any combination thereof. In some instances, a therapeutic disclosed herein is formulated as a dosage form of tablet, capsule, gel, lollipop, parenteral, intraspinal infusion, inhalation, spray, aerosol, transdermal patch, iontophoresis transport, absorbing gel, liquid, liquid tannate, suppositories, injection, I.V. drip, or a combination thereof to treat subjects. In some instances, the active agents are formulated as single oral dosage form such as a tablet, capsule, cachet, soft gelatin capsule, hard gelatin capsule, extended release capsule, tannate tablet, oral disintegrating tablet, multi-layer tablet, effervescent tablet, bead, liquid, oral suspension, chewable lozenge, oral solution, lozenge, lollipop, oral syrup, sterile packaged powder including pharmaceutically-acceptable excipients, other oral dosage forms, or a combination thereof. In some instances, a therapeutic of the disclosure herein can be administered using one or more different dosage forms which are further disclosed herein. In some instances, therapeutics disclosed herein are provided in modified release dosage forms (such as immediate release, controlled release, or both),

[0170] The methods, compositions, and kits of this disclosure can comprise a method to prevent, treat, arrest, reverse, or ameliorate the symptoms of a condition of a subject, e.g., a patient. A

subject can be, for example, an elderly adult, adult, adolescent, pre-adolescence, teenager, or child. A subject can be, for example, 10-50 years old, 10-40 years old, 10-30 years old, 10-25 years old, 10-21 years old, 10-18 years old, 10-16 years old, 18-25 years old, or 16-34 years old. The subject can be a female mammal, e.g., a female human being. In some instances, the human subject can be asymptomatic for endometriosis.

[0171] Treatment can be provided to the subject before clinical onset of disease. Treatment can be provided to the subject after clinical onset of disease. Treatment can be provided to the subject after 1 day, 1 week, 6 months, 12 months, or 2 years or more after clinical onset of the disease. Treatment may be provided to the subject for more than 1 day, 1 week, 1 month, 6 months, 12 months, 2 years or more after clinical onset of disease. Treatment may be provided to the subject for less than 1 day, 1 week, 1 month, 6 months, 12 months, or 2 years after clinical onset of the disease. Treatment can also include treating a human in a clinical trial.

[0172] A treatment, e.g., administration of a therapeutic, can occur 1, 2, 3, 4, 5, 6, 7, or 8 times daily. A treatment, e.g., administration of a therapeutic, can occur 1, 2, 3, 4, 5, 6, or 7 times weekly. A treatment, e.g., administration of a therapeutic, can occur 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10 times monthly. A treatment, e.g., administration of a therapeutic, can occur 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 times yearly. In some instances, therapeutics disclosed herein are administered to a subject at about every 4 to about 6 hours, about every 12 hours, about every 24 hours, about every 48 hours, or more often. In some instances, therapeutics disclosed herein can be administered once, twice, three times, four times, five times, six times, seven times, eight times, or more often daily. In some instances, a dosage form disclosed herein provides an effective plasma concentration of an active agent at from about 1 minute to about 20 minutes after administration, such as about: 2 min, 3 min, 4 min, 5 min, 6 min, 7 min, 8 min, 9 min, 10 min, 11 min, 12 min, 13 min, 14 min, 15 min, 16 min, 17 min, 18min, 19 min, 20 min, 21 min, 22 min, 23min, 24 min, 25 min. In some instances, a dosage form of the disclosure herein provides an effective plasma concentration of an active agent at from about 20 minutes to about 24 hours after administration, such as about 20 minutes, 30 minutes, 40 minutes, 50 minutes, 1hr, 1.2 hrs, 1.4hrs, 1.6 hrs, 1.8 hrs, 2 hrs, 2.2 hrs, 2.4 hrs, 2.6 hrs, 2.8 hrs, 3 hrs, 3.2 hrs, 3.4 hrs, 3.6 hrs, 3.8 hrs, 4 hrs, 5 hrs, 6 hrs, 7 hrs, 8 hrs, 9 hrs, 10 hrs, 11 hrs, 12 hrs, 13 hrs, 14 hrs, 15 hrs, 16 hrs, 17 hrs, 18 hrs, 19 hrs, 20 hrs, 21 hrs, 22 hrs, 23 hrs, or 24 hrs following administration. In some instances, an active agent can be present in an effective plasma concentration in a subject for about 4 to about 6 hours, about 12 hours, about 24 hour, or 1 day to 30 days, including but not limited to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30 days.

[0173] In some instances, a therapeutic (e.g., an active agent) is administered to a subject in a dosage of about 0.01 mg to about 500 mg per day, *e.g.*, about 1-50 mg/day for an average person. In some embodiments, the daily dosage is from about 0.01 mg to about 5 mg, about 1 to about 10 mg, about 5 mg to about 20 mg, about 10 mg to about 50 mg, about 20 mg to about 100 mg, about 50 mg to about 150 mg, about 100 mg to about 250 mg, about 150 mg to about 300 mg, or about 250 mg to about 500 mg.

[0174] In some instances, each administration of a therapeutic (e.g., an active agent) is in an amount of about: 0.1-5 mg, 0.1-10 mg, 1-5 mg, 1-10 mg, 1-20 mg, 10-20 mg, 10-30 mg, 10-40 mg, 10-50 mg, 20-30 mg, 20-40 mg, 20-50 mg, 25-50 mg, 30-40 mg, 30-50 mg, 30-60 mg, 40-50 mg, 40-60 mg, 50-60 mg, 50-75 mg, 60-80 mg, 75-100 mg, or 80-100 mg, for example: about 0.5 mg, about 1 mg, about 1.5 mg, about 2 mg, about 2.5 mg, about 3 mg, about 3.5 mg, about 4 mg, about 4.5 mg, about 5 mg, about 5.5 mg, about 6 mg, about 6.5 mg, about 7 mg, about 7.5 mg, about 8 mg, about 8.5 mg, about 9 mg, about 9.5 mg, about 10 mg, about 10.5 mg, about 11 mg, about 11.5 mg, about 12 mg, about 12.5 mg, about 13 mg, about 13.5 mg, about 14 mg, about 14.5 mg, about 15 mg, about 15.5 mg, about 16 mg, about 16.5 mg, about 17 mg, about 17.5 mg, about 18 mg, about 18.5 mg, about 19 mg, about 19.5 mg, about 20 mg, about 22.5 mg, about 25 mg, about 27.5 mg, about 30 mg, about 32.5 mg, about 35 mg, about 37.5 mg, about 40 mg, about 42.5 mg, about 45 mg, about 47.5 mg, about 50 mg, about 55 mg, about 60 mg, about 65 mg, about 70 mg, about 75 mg, about 80 mg, about 85 mg, about 90 mg, about 95 mg, or about 100 mg.

[0175] In some instances, a therapeutic (e.g., an active agent) is administered to a subject in a dosage of about 0.01 g to about 100 g per day, *e.g.*, about 1-10 g/day for an average person. In some embodiments, the daily dosage is from about 0.01 g to about 5 g, about 1 to about 10 g, about 5 g to about 20 g, about 10 g to about 50 g, about 20 g to about 100 g, or about 50 g to about 100 g.

[0176] In some instances, each administration of a therapeutic (e.g., an active agent) is in an amount of about: 0.01-1 g, 0.1-5 g, 0.1-10 g, 1-5 g, 1-10 g, 1-20 g, 10-20 g, 10-30 g, 10-40 g, 10-50 g, 20-30 g, 20-40 g, 20-50 g, 25-50 g, 30-40 g, 30-50 g, 30-60 g, 40-50 g, 40-60 g, 50-60 g, 50-75 g, 60-80 g, 75-100 g, or 80-100 g, for example: about 0.5 g, about 1 g, about 1.5 g, about 2 g, about 2.5 g, about 3 g, about 3.5 g, about 4 g, about 4.5 g, about 5 g, about 5.5 g, about 6 g, about 6.5 g, about 7 g, about 7.5 g, about 8 g, about 8.5 g, about 9 g, about 9.5 g, about 10 g, about 10.5 g, about 11 g, about 11.5 g, about 12 g, about 12.5 g, about 13 g, about 13.5 g, about 14 g, about 14.5 g, about 15 g, about 15.5 g, about 16 g, about 16.5 g, about 17 g, about 17.5 g, about 18 g, about 18.5 g, about 19 g, about 19.5 g, about 20 g, about 22.5 g, about

25 g, about 27.5 g, about 30 g, about 32.5 g, about 35 g, about 37.5 g, about 40 g, about 42.5 g, about 45 g, about 47.5 g, about 50 g, about 55 g, about 60 g, about 65 g, about 70 g, about 75 g, about 80 g, about 85 g, about 90 g, about 95 g, or about 100 g.

[0177] In some instances, a therapeutic (e.g., in a liquid) administered to a subject having an active agent concentration of about: 0.01-0.1, 0.1-1, 1-10, 1-20, 5-30, 5-40, 5-50, 10-20, 10-25, 10-30, 10-40, 10-50, 15-20, 15-25, 15-30, 15-40, 15-50, 20-30, 20-40, 20-50, 20-100, 30-40, 30-50, 30-60, 30-70, 30-80, 30-90, 30-100, 40-50, 40-60, 40-70, 40-80, 40-90, 40-100, 50-60, 50-70, 50-80, 50-90, 50-100, 50-150, 50-200, 50-300, 100-300, 100-400, 100-500, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, or 1000 μ M, or any combination thereof.

[0178] In some cases, a therapeutic can comprise one or more active agents, administered to a subject at least about: 0.001 mg, 0.01 mg, 0.1 mg, 0.2 mg, 0.3 mg, 0.4 mg, 0.5 mg, 0.6 mg, 0.7 mg, 0.8 mg, 0.9 mg, 1 mg, 1.5 mg, 2 mg, 2.5 mg, 3 mg, 3.5 mg, 4 mg, 4.5 mg, 5 mg, 5.5 mg, 6 mg, 6.5 mg, 7 mg, 7.5 mg, 8 mg, 8.5 mg, 9 mg, 9.5 mg, or 10 mg, or per kg body weight of a subject in need thereof. The therapeutic may comprise a total dose of one or more active agents administered at about 0.1 to about 10.0 mg, for example, about 0.1-10.0 mg, about 0.1-9.0 mg, about 0.1-8.0 mg, about 0.1-7.0 mg, about 0.1-6.0 mg, about 0.1-5.0 mg, about 0.1-4.0 mg, about 0.1-3.0 mg, about 0.1-2.0 mg, about 0.1-1.0 mg, about 0.1-0.5 mg, about 0.2-10.0 mg, about 0.2-9.0 mg, about 0.2-8.0 mg, about 0.2-7.0 mg, about 0.2-6.0 mg, about 0.2-5.0 mg, about 0.2-4.0 mg, about 0.2-3.0 mg, about 0.2-2.0 mg, about 0.2-1.0 mg, about 0.2-0.5 mg, about 0.5-10.0 mg, about 0.5-9.0 mg, about 0.5-8.0 mg, about 0.5-7.0 mg, about 0.5-6.0 mg, about 0.5-5.0 mg, about 0.5-4.0 mg, about 0.5-3.0 mg, about 0.5-2.0 mg, about 0.5-1.0 mg, about 1.0-10.0 mg, about 1.0-5.0 mg, about 1.0-4.0 mg, about 1.0-3.0 mg, about 1.0-2.0 mg, about 2.0-10.0 mg, about 2.0-9.0 mg, about 2.0-8.0 mg, about 2.0-7.0 mg, about 2.0-6.0 mg, about 2.0-5.0 mg, about 2.0-4.0 mg, about 2.0-3.0 mg, about 5.0-10.0 mg, about 5.0-9.0 mg, about 5.0-8.0 mg, about 5.0-7.0 mg, about 5.0-6.0 mg, about 6.0-10.0 mg, about 6.0-9.0 mg, about 6.0-8.0 mg, about 6.0-7.0 mg, about 7.0-10.0 mg, about 7.0-9.0 mg, about 7.0-8.0 mg, about 8.0-10.0 mg, about 8.0-9.0 mg, or about 9.0-10.0 mg, or per kg body weight of a subject in need thereof.

[0179] In some cases, a method of treatment disclosed herein comprises administering a therapeutic. In some instances, the method comprises administering a therapeutic includes one or more of the following steps: a) obtaining a genetic material sample of a human female subject, b) identifying in the genetic material of the subject a genetic marker having an association with endometriosis, c) assessing the subject's risk of endometriosis or risk of

endometriosis progression, d) identifying the subject as having an altered risk of endometriosis or an altered risk of endometriosis progression, e) administering to the subject a therapeutic, or any combination thereof.

[0180] In some instances, the subject may be endometriosis presymptomatic or the subject may exhibit endometriosis symptoms. In some instances, the assessment of risk may include non-genetic clinical factors. In some instances, the therapeutic is adapted to the specific subject so as to be a proper and effective amount of therapeutic for the subject. In some instances, the administration of the therapeutic may comprise multiple sequential instances of administration of the therapeutic and that such sequence instances may occur over an extended period of time or may occur on an indefinite on-going basis. In some instances, the therapeutic may be a gene or protein based therapy adapted to the specific needs of a select patient.

[0181] *Hormonal Therapy*

[0182] In some cases, a treatment method herein comprises supplementing the body with a hormone thereof such as a steroid hormone, for example a method of preventing endometriosis comprising administering a hormonal therapy to a human subject having at least one genetic variant defining a minor allele disclosed herein, e.g., listed in **Table 1**. In some instances, the hormone can be progestin, progestogen, progesterone, desogestrel, etonogestrel, gestodene, levonorgestrel, medroxyprogesterone, norethisterone, norgestimate, megestrol, megestrol acetate, norgestrel, a pharmaceutically acceptable salt thereof (e.g., acetate), or any combination thereof. In some instances, a therapeutic used herein is selected from progestins, estrogens, antiestrogens, and antiprogestins, for example micronized danazol in a micro- or nanoparticulate formulation. Methods and therapeutics presented herein can utilize an active agent in a freebase, salt, hydrate, polymorph, isomer, diastereomer, prodrug, metabolite, ion pair complex, or chelate form. An active agent can be formed using a pharmaceutically acceptable non-toxic acid or base, including an inorganic acid or base, or an organic acid or base. In some instances, an active agent that can be utilized in connection with the methods and compositions presented herein is a pharmaceutically acceptable salt derived from acids including, but not limited to, the following: acetic, alginic, anthranilic, benzenesulfonic, benzoic, camphorsulfonic, citric, ethenesulfonic, formic, fumaric, furoic, galacturonic, gluconic, glucuronic, glutamic, glycolic, hydrobromic, hydrochloric, isethionic, lactic, maleic, malic, mandelic, methanesulfonic, mucic, nitric, pamoic, pantothenic, phenylacetic, phosphoric, propionic, salicylic, stearic, succinic, sulfanilic, sulfuric, tartaric acid, or p-toluenesulfonic acid. For further description of pharmaceutically acceptable salts that can be used in the methods described herein see, for

example, S.M. Barge et al., “Pharmaceutical Salts,” 1977, J. Pharm. Sci. 66:1-19, which is incorporated herein by reference in its entirety.

[0183] In some instances, the therapeutic may take the form of a testosterone or a modified testosterone such as Danazol. In some instances, the therapeutic can be a hormonal treatment therapeutic which may be administered alone or in combination with a gene therapy. For instance, the therapeutic may be an estrogen containing composition, a progesterone containing composition, a progestin containing composition, a gonadotropin releasing-hormone (GnRH) agonist, a gonadotropin releasing-hormone (GnRH) antagonist, or other ovulation suppression composition, or a combination thereof. In some instances, the GnRH agonist may take the form of a GnRH agonist in combination with a patient specific substantially low dose of estrogen, progestin, or tibolone via an add-back administration. In some instances, in such add-back therapy, the dosage of estrogen, progestin, or tibolone is relatively small so as to not reduce the effectiveness of the GnRH agonist. In some instances, the therapeutic is an oral contraceptive (OC). In some instances, the OC is in a pill form that is comprised at least partially of estrogen, progesterone, or a combination thereof. In some instances, the progesterone component may be any of Desogestrel, Drospirenone, Ethynodiol, Levonorgestrel, Norethindrone, Norgestimate, and Norgestrel, and the estrogen component may further be any of Mestranol, Estradiol, and Ethinyl. In some instances, the OC may be any commercially available OC including ALESSE, APRI, ARANELLE, AVIANE, BREVICON, CAMILA, CESIA, CRYSELLE, CYCLESSA, DEMULEN, DESOGEN, ENPRESSE, ERRIN, ESTROSTEP, JOLIVETTE, JUNEL, KARIVA, LEENA, LESSINA, LEVLEN, LEVORA, LOESTRIN, LUTERA, MICROGESTIN, MICRONOR, MIRCETTE, MODICON, MONONESSA, NECON, NORA, NORDETTE, NORINYL, NOR-QD, NORTREL, OGESTREL, ORTHO-CEPT, ORTHO-CYCLEN, ORTHO-NOVUM, ORTHO-TRI-CYCLEN, OVCON, OVRAL, OVRETTE, PORTIA, PREVIFEM, RECLIPSEN, SOLIA, SPRINTec, TRINESSA, TRI-NORINYL, TRIPHASIL, TRIVORA, VELIVET, YASMIN, AND ZOVIA (the preceding names are the registered trademarks of the respective providers).

[0184] *Assisted Reproductive technology Therapy*

[0185] In some cases, a method herein can comprise administering to a select subject assisted reproductive technology therapy (ART), for example a method of treating endometriosis-associated infertility comprising administering ART to a select human subject having at least one genetic variant defining a minor allele disclosed herein, e.g., listed in **Table 2**. In some instances, ART can comprise *in vitro* fertilization (IVF), embryo transfer (ET), fertility medication, intracytoplasmic sperm injection (ICSI), cryopreservation, or any combination

thereof. In some instances, ART can comprise surgically removing eggs from a woman's ovaries, combining them with sperm in the laboratory, and returning them to the woman's body or donating them to another woman.

[0186] In some cases, assisted reproductive technology therapy can comprises all treatments or procedures that include the handling of human eggs or embryos to help a woman become pregnant. For example, in vitro fertilization (IVF), gamete intrafallopian transfer (GIFT), zygote intrafallopian transfer (ZIFT), tubal embryo transfer, gg and embryo cryopreservation, egg and embryo donation and gestational surrogacy.

[0187] In some instances, the in vitro fertilization (IVF) procedure can provide for a live birth event following the IVF procedure. In some instances, a method herein provides a probability of a live birth event occurring resulting from the first or subsequent in vitro fertilization cycle based at least in part on items of information from the female subjects.

[0188] In some instances, the IVF can comprise ovulation induction, utilizing fertility medication can comprise agents that stimulate the development of follicles in the ovary. Examples are gonadotropins and gonadotropin releasing hormone.

[0189] In some instances, IVF can comprise transvaginal ovum retrieval (OVR), which can be a process whereby a small needle is inserted through the back of the vagina and guided via ultrasound into the ovarian follicles to collect the fluid that contains the eggs.

[0190] In some instances, IVF can comprise embryo transfer, which can be the step in the process whereby one or several embryos are placed into the uterus of the female with the intent to establish a pregnancy.

[0191] In some instances, IVF can comprise assisted zona hatching (AZH), which can be performed shortly before the embryo is transferred to the uterus. A small opening can be made in the outer layer surrounding the egg in order to help the embryo hatch out and aid in the implantation process of the growing embryo.

[0192] In some instances, IVF can comprise artificial insemination, for example intrauterine insemination, intracervical insemination, intrauterine tuboperitoneal insemination, intratubal insemination, or any combination thereof.

[0193] In some instances, IVF can comprise intracytoplasmic sperm injection (ICSI), which can be beneficial in the case of male factor infertility where sperm counts are very low or failed fertilization occurred with previous IVF attempt(s). The ICSI procedure can involve a single sperm carefully injected into the center of an egg using a microneedle. With ICSI, only one sperm per egg is needed. Without ICSI, one may need between 50,000 and 100,000. In some embodiments, this method can be employed when donor sperm is used.

[0194] In some instances, IVF can comprise autologous endometrial coculture, which can be a possible treatment for patients who have failed previous IVF attempts or who have poor embryo quality. The patient's fertilized eggs can be placed on top of a layer of cells from the patient's own uterine lining, creating a more natural environment for embryo development.

[0195] In some instances, IVF can comprise zygote intrafallopian transfer (ZIFT), in which egg cells can be removed from the woman's ovaries and fertilized in the laboratory; the resulting zygote can be then placed into the fallopian tube.

[0196] In some instances, IVF can comprise cytoplasmic transfer, in which the contents of a fertile egg from a donor can be injected into the infertile egg of the patient along with the sperm.

[0197] In some instances, IVF can comprise egg donors, which are resources for women with no eggs due to surgery, chemotherapy, or genetic causes; or with poor egg quality, previously unsuccessful IVF cycles or advanced maternal age. In the egg donor process, eggs can be retrieved from a donor's ovaries, fertilized in the laboratory with the sperm from the recipient's partner, and the resulting healthy embryos can be returned to the recipient's uterus.

[0198] In some instances, IVF can comprise sperm donation, which may provide the source for the sperm used in IVF procedures where the male partner produces no sperm or has an inheritable disease, or where the woman being treated has no male partner.

[0199] In some instances, IVF can comprise preimplantation genetic diagnosis (PGD), which can involve the use of genetic screening mechanisms such as fluorescent in-situ hybridization (FISH) or comparative genomic hybridization (CGH) to help identify genetically abnormal embryos and improve healthy outcomes.

[0200] In some instances, IVF can comprise embryo splitting can be used for twinning to increase the number of available embryos.

[0201] In some instances, ART can comprise gamete intrafallopian transfer (GIFT), in which a mixture of sperm and eggs can be placed directly into a woman's fallopian tubes using laparoscopy following a transvaginal ovum retrieval.

[0202] In some instances, ART can comprise reproductive surgery, treating e.g. fallopian tube obstruction and vas deferens obstruction, or reversing a vasectomy by a reverse vasectomy. In surgical sperm retrieval (SSR) the reproductive urologist can obtain sperm from the vas deferens, epididymis or directly from the testis in a short outpatient procedure. By cryopreservation, eggs, sperm and reproductive tissue can be preserved for later IVF.

[0203] In some instances, a subject to treat can be a pre-in vitro fertilization (pre-IVF) procedure patient. In certain embodiments, the items of information relating to preselected patient variables for determining the probability of a live birth event for a pre-IVF procedure patient may include

age, diminished ovarian reserve, 3 follicle stimulating hormone (FSH) level, body mass index, polycystic ovarian disease, season, unexplained female infertility, number of spontaneous miscarriages, year, other causes of female infertility, number of previous pregnancies, number of previous term deliveries, endometriosis, tubal disease, tubal ligation, male infertility, uterine fibroids, hydrosalpinx, and male infertility causes.

[0204] In some instances, a subject to treat can be a pre-surgical (pre-OR) procedure patient (pre-OR is also referred to herein as pre-oocyte retrieval). In certain embodiments, the items of information relating to preselected patient variables for determining the probability of a live birth event for a pre-OR procedure patient may include age, endometrial thickness, total number of oocytes, total amount of gonatropins administered, number of total motile sperm after wash, number of total motile sperm before wash, day 3 follicle stimulating hormone (FSH) level, body mass index, sperm collection, age of spouse, season number of spontaneous miscarriages, unexplained female infertility, number of previous term deliveries, year, number of previous pregnancies, other causes of female infertility, endometriosis, male infertility, tubal ligation, polycystic ovarian disease, tubal disease, sperm from donor, hydrosalpinx, uterine fibroids, and male infertility causes.

[0205] In some instances, a subject to treat can be a post-in vitro fertilization (post-IVF) procedure patient. In certain embodiments, the items of information relating to preselected patient variables for determining the probability of a live birth event for a post-IVF procedure patient may include blastocyst development rate, total number of embryos, total amount of gonatropins administered, endometrial thickness, flare protocol, average number of cells per embryo, type of catheter used, percentage of 8-cell embryos transferred, day 3 follicle stimulating hormone (FSH) level, body mass index, number of motile sperm before wash, number of motile sperm after wash, average grade of embryos, day of embryo transfer, season, number of spontaneous miscarriages, number of previous term deliveries, oral contraceptive pills, sperm collection, percent of unfertilized eggs, number of embryos arrested at 4-cell stage, compaction on day 3 after transfer, percent of normal fertilization, percent of abnormally fertilized eggs, percent of normal and mature oocytes, number of previous pregnancies, year, polycystic ovarian disease, unexplained female infertility, tubal disease, male infertility only, male infertility causes, endometriosis, other causes of female infertility, uterine fibroids, tubal ligation, sperm from donor, hydrosalpinx, performance of ICSI, or assisted hatching.

[0206] *Pain Managing Medications*

[0207] In some cases, a method disclosed herein can comprise administering a pain medication to a select subject, for example to a human subject having at least one genetic variant defining a

minor allele listed in **Table 3**. In some instances, the pain medication comprises a nonsteroidal anti-inflammatory drug (NSAID), ibuprofen, naproxen, acetaminophen, an opioid, a cannabis-based therapeutic, or any combination thereof.

[0208] In some instances, the pain medication described herein can comprise an NSAID, for example amoxiprin, benorilate, choline magnesium salicylate, diflunisal, faislamine, methyl salicylate, magnesium salicylate, diclofenac, aceclofenac, acemetacin, bromfenac, etodolac, indometacin, nabumetone, sulindac, tolmetin, ibuprofen, carprofen, fenbuprofen, flubiprofen, ketaprofen, ketorolac, loxoprofen, naproxen, suprofen, mefenamic acid, meclofenamic acid, piroxicam, lomoxicam, meloxicam, tenoxicam, phenylbutazone, azapropazone, metamizole, oxyphenbutazone, or sulfinprazone, or a pharmaceutically acceptable salt thereof.

[0209] In some instances, the pain medication described herein can comprise an opioid analgesic, for example hydrocodone, oxycodone, morphine, diamorphine, codeine, pethidine, alfentanil, buprenorphine, butorphanol, dezocine, fentanyl, hydromorphone, levomethadyl acetate, levorphanol, meperidine, methadone, morphine sulfate, nalbuphine, oxymorphone, pentazocine, propoxyphene, remifentanyl, sufentanyl, or tramadol, or a pharmaceutically acceptable salt thereof.

[0210] In some instances, the pain medication described herein can comprise a cannabis-based therapeutic such as a cannabinoid for the treatment, reduction or prevention of pain. Exemplary cannabinoid for the treatment of pain include, without limitation, nabilone, dronabinol (THC), cannabidiol (CBD), cannabinol (CBN), cannabichromene (CBC), cannabigerol (CBG), tetrahydrocannabivarin (THCV), tetrahydrocannabinolic acid (THCA), cannabidivarin (CBDV), cannadidiolic acid (CBDA), ajulemic acid, dexanabinol, cannabinor, HU 308, HU 331, and a pharmaceutically acceptable salt thereof.

[0211] Specific Embodiments

[0212] A number of methods and systems are disclosed herein. Specific exemplary embodiments of these methods and systems are disclosed below.

[0213] Section 1 of Specific Embodiments

[0214] Embodiment 1. A method comprising: hybridizing a nucleic acid probe to a nucleic acid sample from a human subject suspected of having or developing endometriosis; and detecting a genetic variant in a panel comprising two or more genetic variants defining a minor allele listed in **Table 1**.

[0215] Embodiment 2. The method of embodiment 1, wherein the nucleic acid sample comprises mRNA, cDNA, genomic DNA, or PCR amplified products produced therefrom, or any combination thereof.

[0216] Embodiment 3. The method of embodiment 1 or 2, wherein the nucleic acid sample comprises PCR amplified nucleic acids produced from cDNA or mRNA.

[0217] Embodiment 4. The method of embodiment 1 or 2, wherein the nucleic acid sample comprises PCR amplified nucleic acids produced from genomic DNA.

[0218] Embodiment 5. The method of any one of embodiments 1-4, wherein the nucleic acid probe is a sequencing primer.

[0219] Embodiment 6. The method of any one of embodiments 1-4, wherein the nucleic acid probe is an allele specific probe.

[0220] Embodiment 7. The method of any one of embodiments 1-6, wherein the detecting comprises DNA sequencing, hybridization with a complementary probe, an oligonucleotide ligation assay, a PCR-based assay, or any combination thereof.

[0221] Embodiment 8. The method of any one of embodiments 1-7, wherein the panel comprises at least: 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 500, or more genetic variants defining minor alleles listed in **Table 1**.

[0222] Embodiment 9. The method of any one of embodiments 1-8, wherein the genetic variant has an odds ratio (OR) of at least: 1.5, 2, 5, 10, 20, 50, 100, or more.

[0223] Embodiment 10. The method of any one of embodiments 1-9, wherein the genetic variant comprises a synonymous mutation, a non-synonymous mutation, a nonsense mutation, an insertion, a deletion, a splice-site variant, a frameshift mutation, or any combination thereof.

[0224] Embodiment 11. The method of any one of embodiments 1-9, wherein the genetic variant comprises a protein damaging mutation.

[0225] Embodiment 12. The method of any one of embodiments 1-10, wherein the panel further comprises one or more protein damaging or loss of function variants in one or more genes selected from the group consisting of *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR*, and any combinations thereof.

[0226] Embodiment 13. The method of embodiment 12, further comprising sequencing the one or more genes to identify the one or more protein damaging or loss of function variants.

[0227] Embodiment 14. The method of embodiment 13, wherein the one or more protein damaging or loss of function variants are identified based on a predictive computer algorithm.

[0228] Embodiment 15. The method of embodiment 13 or 14, wherein the one or more protein damaging or loss of function variants are identified based on reference to a database.

[0229] Embodiment 16. The method of any one of embodiments 12-15, wherein the one or more protein damaging or loss of function variants comprise a stop-gain mutation, a splice-site mutation, a frameshift mutation, a missense mutation, or any combination thereof.

[0230] Embodiment 17. The method of any one of embodiments 1-16, wherein the panel further comprises one or more additional variants defining a minor allele listed in **Table 4**.

[0231] Embodiment 18. The method of any one of embodiments 1-17, wherein the panel is capable of identifying human subjects as having or being at risk of developing endometriosis with a specificity of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

[0232] Embodiment 19. The method of any one of embodiments 1-18, wherein the panel is capable of identifying human subjects as having or being at risk of developing endometriosis with a sensitivity of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

[0233] Embodiment 20. The method of any one of embodiments 1-19, wherein the panel is capable of identifying human subjects as having or being at risk of developing endometriosis with an accuracy of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

[0234] Embodiment 21. The method of any one of embodiments 1-20, further comprising administering a therapeutic to the human subject.

[0235] Embodiment 22. The method of embodiment 21, wherein the therapeutic comprises hormonal therapy, an advanced reproductive technology therapy, a pain managing medication, or any combination thereof.

[0236] Embodiment 23. The method of embodiment 21, wherein the therapeutic comprises hormonal contraceptives, gonadotropin-releasing hormone (Gn-RH) agonists, gonadotropin-releasing hormone (Gn-RH) antagonists, progestin, danazol, or any combination thereof.

[0237] Embodiment 24. The method of any one of embodiments 1-23, wherein the human subject is asymptomatic for endometriosis.

[0238] Embodiment 25. The method of any one of embodiments 1-24, wherein the human subject is a teenager.

[0239] Embodiment 26. A method comprising detecting one or more genetic variants defining a minor allele listed in **Table 1** in genetic material from a human subject suspected of having or developing endometriosis.

[0240] Embodiment 27. The method of embodiment 26, wherein the genetic material comprises mRNA, cDNA, genomic DNA, or PCR amplified products produced therefrom, or any combination thereof.

[0241] Embodiment 28. The method of embodiment 26 or 27, wherein the detecting comprises DNA sequencing, hybridization with a complementary probe, an oligonucleotide ligation assay, a PCR-based assay, of any combination thereof.

[0242] Embodiment 29. The method of any one of embodiments 26-28, wherein the detecting comprises hybridizing a nucleic acid probe to the genetic material.

[0243] Embodiment 30. The method of any one of embodiments 26-29, wherein the detecting comprises testing for the presence or absence of at least: 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 150, 250, or 500 genetic variants defining a minor allele listed in **Table 1**.

[0244] Embodiment 31. The method of any one of embodiments 26-30, wherein the one or more genetic variants have an odds ratio (OR) of at least: 1.5, 2, 5, 10, 20, 50, 100, or more.

[0245] Embodiment 32. The method of any one of embodiments 26-31, further comprising administering a therapeutic to the human subject.

[0246] Embodiment 33. A method comprising: sequencing one or more genes selected from the group consisting of *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR*, and any combinations thereof to identify one or more protein damaging or loss of function variants in a human subject suspected of having or developing endometriosis; and administering an endometriosis therapy to the human subject.

[0247] Embodiment 34. The method of embodiment 33, wherein the one or more protein damaging or loss of function variants are identified based on a predictive computer algorithm, reference to a database, or a combination thereof.

[0248] Embodiment 35. The method of embodiment 33 or 34, wherein the one or more protein damaging or loss of function variants comprise a stop-gain mutation, a splice-site mutation, a frameshift mutation, a missense mutation, or any combination thereof.

[0249] Embodiment 36. The method of any one of embodiments 33-35, wherein the endometriosis therapy comprises a hormonal therapy, an assisted reproductive technology therapy, a pain medication, or any combination thereof.

[0250] Embodiment 37. A method of preventing endometriosis comprising administering a hormonal therapy to a human subject having at least one genetic variant defining a minor allele listed in **Table 1**.

[0251] Embodiment 38. The method of embodiment 37, wherein the hormonal therapy comprises administration of hormonal contraceptives, gonadotropin-releasing hormone (Gn-RH) agonists, gonadotropin-releasing hormone (Gn-RH) antagonists, progestin, danazol, or any combination thereof.

[0252] Embodiment 39. A method of treating endometriosis-associated infertility comprising administering an assisted reproductive technology therapy to a human subject having at least one genetic variant defining a minor allele listed in **Table 2**.

[0253] Embodiment 40. The method of embodiment 39, wherein the assisted reproductive technology therapy comprises in vitro fertilization, gamete intrafallopian transfer, or any combination thereof.

[0254] The method can further comprise administering, intrauterine insemination or ovulation induction.

[0255] Embodiment 41. A method comprising administering a pain medication to a human subject having at least one genetic variant defining a minor allele listed in **Table 3**.

[0256] Embodiment 42. The method of embodiment 41, wherein the pain medication comprises a nonsteroidal anti-inflammatory drug (NSAID), ibuprofen, naproxen, an opioid, a cannabis-based therapeutic, or any combination thereof.

[0257] Embodiment 43. The method of any one of embodiment 37-42, further comprising detecting the at least one genetic variant in a genetic material from the human subject.

[0258] Embodiment 44. The method of embodiment 43, wherein the detecting comprises DNA sequencing, hybridization with a complementary probe, an oligonucleotide ligation assay, a PCR-based assay, or any combination thereof.

[0259] Embodiment 45. The method of embodiment 43, wherein the detecting comprises hybridizing a nucleic acid probe to the genetic material.

[0260] Embodiment 46. The method of embodiment 45, wherein the nucleic acid probe is a sequencing primer or an allele-specific probe.

[0261] Embodiment 47. The method of any one of embodiments 37-46, wherein the at least one genetic variant has an odds ratio (OR) of at least: 1.5, 2, 5, 10, 20, 50, 100, or more.

[0262] Embodiment 48. The method of any one of embodiments 37-47, wherein the at least one genetic variant comprises a synonymous mutation, a non-synonymous mutation, a nonsense mutation, an insertion, a deletion, a splice-site variant, a frameshift mutation, or any combination thereof.

[0263] *Section 2 of Specific Embodiments*

[0264] Embodiment 1. A method comprising: (a) sequencing or genotyping a nucleic acid sample obtained from a subject having endometriosis, suspected of having endometriosis, or suspected of having a risk of developing endometriosis using a high throughput method; and (b) detecting one or more genetic variants in said nucleic acid sample, wherein said one or more genetic variants are listed in Table 1, Table 2 or Table 3.

[0265] Embodiment 2. The method of embodiment 1, wherein said high throughput method comprises nanopore sequencing.

[0266] Embodiment 3. The method of embodiment 1 or 2, wherein said nucleic acid sample comprises RNA.

[0267] Embodiment 4. The method of embodiment 3, wherein said RNA comprises mRNA.

[0268] Embodiment 5. The method of embodiment 1 or 2, wherein said nucleic acid sample comprises DNA.

[0269] Embodiment 6. The method of embodiment 5, wherein said DNA comprises cDNA, genomic DNA, sheared DNA, cell free DNA, fragmented DNA, or PCR amplified products produced therefrom, or any combination thereof.

[0270] Embodiment 7. The method of embodiment 5, wherein said DNA comprises DNA from an endometriosis lesion or peritoneal fluid.

[0271] Embodiment 8. The method of any one of embodiments 1-7, wherein said one or more genetic variants comprise a genetic variant defining a minor allele.

[0272] Embodiment 9. The method of any one of embodiments 1-7, wherein said one or more genetic variants comprise at least about: 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 500, or more genetic variants defining minor alleles.

[0273] Embodiment 10. The method of any one of embodiments 1-9, wherein detection of said one or more genetic variants has an odds ratio (OR) for endometriosis of at least about: 1.5, 2, 5, 10, 20, 50, 100, or more.

[0274] Embodiment 11. The method of any one of embodiments 1-10, wherein said one or more genetic variants comprise a synonymous mutation, a non-synonymous mutation, a stop-gain mutation, a nonsense mutation, an insertion, a deletion, a splice-site variant, a frameshift mutation, or any combination thereof.

[0275] Embodiment 12. The method of any one of embodiments 1-11, wherein said one or more genetic variants comprise a protein damaging mutation.

[0276] Embodiment 13. The method of any one of embodiments 12, wherein said one or more genetic variants further comprise a protein damaging or loss of function variant in one or more genes selected from the group consisting of *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR*, and any combinations thereof.

[0277] Embodiment 14. The method of any one of embodiments 1-12, wherein said one or more genetic variants are comprised in *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR* or a combination thereof.

[0278] Embodiment 15. The method of any one of embodiments 1-13, further comprising detecting one or more additional variants defining a minor allele listed in Table 4.

[0279] Embodiment 16. The method of any one of embodiment 1-15, wherein said one or more genetic variants are identified or weighted based on a predictive mathematical or computer programmed algorithm.

[0280] Embodiment 17. The method of any one of embodiments 1-16, wherein said one or more genetic variants are identified based on reference to a database.

[0281] Embodiment 18. The method of any one of embodiments 1-17, further comprising identifying said subject as having endometriosis or being at risk of developing endometriosis.

[0282] Embodiment 19. The method of embodiment 18, wherein said identifying said subject as having endometriosis or being at risk of developing endometriosis is with a specificity of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

[0283] Embodiment 20. The method of any one of embodiments 18-19, wherein said identifying said subject as having endometriosis or being at risk of developing endometriosis is with a sensitivity of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

[0284] Embodiment 21. The method of any one of embodiments 18-20, wherein said identifying said subject as having endometriosis or being at risk of developing endometriosis is with an accuracy of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.

[0285] Embodiment 22. The method of any one of embodiments 18-21, wherein said subject is identified as having endometriosis.

[0286] Embodiment 23. The method of embodiment 22, wherein said subject is asymptomatic for endometriosis.

[0287] Embodiment 24. The method of embodiment 22, wherein said subject is symptomatic for endometriosis.

[0288] Embodiment 25. The method of any one of embodiments 18-21, wherein said subject is identified as being at risk of developing endometriosis.

[0289] Embodiment 26. The method of any one of embodiments 1-25, further comprising administering a therapeutic to said subject.

[0290] Embodiment 27. The method of embodiment 26, wherein said therapeutic comprises hormonal therapy, an advanced reproductive technology therapy, a pain managing medication, or any combination thereof.

[0291] Embodiment 28. The method of embodiment 26, wherein said therapeutic comprises hormonal contraceptives, gonadotropin-releasing hormone (Gn-RH) agonists, gonadotropin-releasing hormone (Gn-RH) antagonists, progestin, danazol, or any combination thereof.

[0292] Embodiment 29. The method of any one of embodiments 26-28, wherein said therapeutic comprises a pain medication.

[0293] Embodiment 30. The method of embodiment 29, wherein said pain medication comprises a nonsteroidal anti-inflammatory drug (NSAID), ibuprofen, naproxen, an opioid, a cannabis-based therapeutic, or any combination thereof.

[0294] Embodiment 31. The method of any one of embodiments 1-26, wherein said one or more genetic variants are listed in Table 1.

[0295] Embodiment 32. The method of any one of embodiments 1-26, wherein said one or more genetic variants are listed in Table 2.

[0296] Embodiment 33. The method of any one of embodiments 1-26, wherein said one or more genetic variants are listed in Table 3.

[0297] Embodiment 34. The method of any one of embodiments 1-33, further comprising identifying said subject as having endometriosis-associated infertility or being at risk of developing endometriosis-associated infertility.

[0298] Embodiment 35. The method of embodiment 34, further comprising administering assisted reproductive technology therapy to said subject.

[0299] Embodiment 36. The method of embodiment 35, wherein said assisted reproductive technology therapy comprises *in vitro* fertilization, gamete intrafallopian transfer, or any combination thereof.

[0300] Embodiment 37. The method of embodiment 34, further comprising administering intrauterine insemination or ovulation induction.

[0301] Embodiment 38. The method of any one of embodiments 1- 37, wherein said subject is a mammal.

[0302] Embodiment 39. The method of embodiment 38, wherein said mammal is a human.

[0303] Embodiment 40. The method of any one of embodiments 2-39, wherein said nanopore sequencing is performed with a biological nanopore, a solid state nanopore, or a hybrid nanopore.

[0304] Embodiment 41. The method of any one of embodiments 1-40, wherein said one or more genetic variants further comprise a mutation in *SEPT10*, *TNFRSF6B*, *UGT2B28*, *USP17L2* or any combination thereof.

[0305] Embodiment 42. The method of embodiment 41, wherein said one or more genetic variants comprise a mutation in *SEPT10* and wherein said mutation comprises a missense mutation.

[0306] Embodiment 43. The method of embodiment 41, wherein said one or more genetic variants comprise a mutation in *TNFRSF6B* and wherein said mutation comprises a homozygous or hemizygous mutation.

[0307] Embodiment 44. The method of embodiment 41, wherein said one or more genetic variants comprise a mutation in *UGT2B28* or *USP17L2* and wherein said mutation comprises a hemizygous deletion.

- [0308]** Embodiment 45. The method of any one of embodiments 1-44, wherein the one or more variants are identified based on a predictive computer algorithm.
- [0309]** Embodiment 46. The method of embodiment 45, wherein said predictive computer algorithm is Polyphen 2, Sift, Mutation Accessor, Mutation Taster, FATHMM, LRT, or MetaLR.
- [0310]** Embodiment 47. The method of any one of embodiments 1-46, further comprising administering a hormonal therapy to said subject.
- [0311]** Embodiment 48. The method of embodiment 47, wherein the hormonal therapy comprises administration of hormonal contraceptives, gonadotropin-releasing hormone (GnRH) agonists, gonadotropin-releasing hormone (GnRH) antagonists, progestin, danazol, or any combination thereof.
- [0312]** Embodiment 49. The method of any one of embodiments 1-46, further comprising administering to the subject an assisted reproductive therapy.
- [0313]** Embodiment 50. The method of embodiment 49, wherein the assisted reproductive therapy comprises in vitro fertilization, intrauterine insemination, ovulation induction, gamete intrafallopian transfer, or any combination thereof.
- [0314]** Embodiment 51. The method of any one of embodiments 1-46, further comprising administering to the subject a pain medication.
- [0315]** Embodiment 52. The method of embodiment 51, wherein the pain medication comprises a nonsteroidal anti-inflammatory drug (NSAID), ibuprofen, naproxen, an opioid, a cannabis-based therapeutic, or any combination thereof.
- [0316]** Embodiment 53. The method of any one of embodiments 1-46, further comprising administering a therapeutic to the subject.
- [0317]** Embodiment 54. The method of embodiment 53, wherein the therapeutic comprises a regenerative therapy, a medical device, a pharmaceutical composition, a medical procedure, or any combination thereof.
- [0318]** Embodiment 55. The method of embodiment 53, wherein the therapeutic comprises a non-steroidal anti-inflammatory, a hormone treatment, a dietary supplement, a cannabis-derived therapeutic or any combination thereof.
- [0319]** Embodiment 56. The method of embodiment 53, wherein the therapeutic comprises the pharmaceutical composition, and wherein the pharmaceutical composition comprises an at least partially hemp-derived therapeutic, an at least partially cannabis-derived therapeutic, a cannabidiol (CBD) oil derived therapeutic, or any combination thereof.

[0320] Embodiment 57. The method of embodiment 53, wherein the therapeutic comprises the medical procedure, and wherein the medical procedure comprises a laparoscopy, a laser ablation procedure, a hysterectomy or any combination thereof.

[0321] Embodiment 58. The method of embodiment 53, wherein the therapeutic comprises the regenerative therapy, and wherein the regenerative therapy comprises a stem cell, a cord blood cell, a Wharton's jelly, an umbilical cord tissue, a tissue, or any combination thereof.

[0322] Embodiment 59. The method of embodiment 53, wherein the therapeutic comprises the pharmaceutical composition, and wherein the pharmaceutical composition comprises cannabis, cannabidiol oil, hemp, or any combination thereof.

[0323] Embodiment 60. The method of embodiment 53, wherein the therapeutic comprises the pharmaceutical composition, and wherein the pharmaceutical composition is formulated in a unit dose.

[0324] Embodiment 61. The method of embodiment 53, wherein the therapeutic comprises hormonal therapy, an advanced reproductive therapy, a pain managing medication, or any combination thereof.

[0325] Embodiment 62. The method of embodiment 53, wherein the therapeutic comprises a hormonal contraceptive, gonadotropin-releasing hormone (GnRH) agonist, gonadotropin-releasing hormone (GnRH) antagonist, progestin, danazol, or any combination thereof.

[0326] Embodiment 63. The method of any one of embodiments 1-62, wherein the subject is asymptomatic for endometriosis.

[0327] Embodiment 64. A kit comprising: one or more probes for detecting one or more genetic variants of Table 1, Table 2, Table 3, or any combination thereof in a sample.

[0328] Embodiment 65. The kit of embodiment 64, further comprising a control sample.

[0329] Embodiment 66. The kit of embodiment 64, wherein the control sample comprises one or more genetic variants of Table 1, Table 2, Table 3, or any combination thereof.

[0330] Embodiment 67. The kit of any one of embodiments 64-66, wherein the one or more probes comprise a hybridization probe or amplification primer.

[0331] Embodiment 68. The kit of any one of embodiments 64-67, wherein the one or more probes is configured to associate with a solid support.

[0332] Embodiment 69. The kit of any one of embodiments 64-68, wherein the kit further comprises instructions for use and wherein the instructions for use comprise high stringent hybridization conditions.

[0333] Embodiment 70. The kit of any one of embodiments 64-69, wherein the one or more probes is configured to hybridize to a target region of a nucleic acid of the sample, wherein the target region comprises one or more genetic variants.

[0334] Embodiment 71. A system comprising: (a) a computer processor configured to receive sequencing data obtained from assaying a sample, wherein the computer processor is configured to identify a presence or an absence of one or more genetic variants of Table 1, Table 2, Table 3 or any combination thereof in the sample, and (b) a graphical user interface configured to display a report comprising the identification of the presence or the absence of the one or more genetic variants in the sample.

[0335] Embodiment 72. The system of embodiment 71, wherein the computer processor comprises a trained algorithm.

[0336] Embodiment 73. The system of embodiment 71 or 72, wherein the computer processor communicates a result.

[0337] Embodiment 74. The system of embodiment 73, wherein the result comprises an identification of the presence or the absence of one or more genetic variants in the sample.

[0338] Embodiment 75. A method comprising: (a) sequencing or genotyping a nucleic acid sample obtained from a subject having endometriosis, suspected of having endometriosis, or suspected of having a risk of developing endometriosis using a high throughput method; and (b) detecting a genetic variant in said nucleic acid sample, wherein said genetic variant comprises a mutation in *SEPT10*, *TNFRSF6B*, *UGT2B28*, *USP17L2* or any combination thereof.

[0339] Embodiment 76. The method of embodiment 75, wherein said genetic variant is a mutation in *SEPT10* and wherein said mutation comprises a missense mutation.

[0340] Embodiment 77. The method of embodiment 75, wherein said genetic variant is a mutation in *TNFRSF6B* and wherein said mutation comprises a homozygous or hemizygous mutation.

[0341] Embodiment 78. The method of embodiment 75, wherein said genetic variant is a mutation in *UGT2B28* or *USP17L2* and wherein said mutation comprises a hemizygous deletion.

[0342] Embodiment 79. The method of embodiment 75, wherein said high throughput method comprises nanopore sequencing.

EXAMPLES

Example 1. Low-Frequency, Damaging Mutations in Hundreds of Genes Are Risk Factors For Endometriosis.

[0343] This study performed exome-wide association analysis for rare low frequency mutations in the women with endometriosis. Rare exome variants associated with endometriosis were searched using an exome genotyping array and confirmatory whole exome sequencing (WES).

[0344] Consent and Medical Review

[0345] All subjects and controls were provided written informed consent in accordance with study protocols approved by Quorum Review IRB (Seattle, WA 98101). Trained OB/GYN clinicians performed the medical record review and clinical assessment of each patient.

[0346] Methods

[0347] Illumina Exome Human BeadChip. 1518 Caucasian patients with surgically confirmed endometriosis were tested for more than 200,000 rare non-synonymous variants (minor allele frequency <0.005). Allele frequencies were compared to the population datasets (genotyping dataset UK Michigan ($n=50,000$) and publicly available sequencing dataset Exac ($n=33,000$)).

[0348] Affymetrix Axiom Custom Chip. 1888 Caucasian patients with surgically confirmed endometriosis were tested for more than 700,000 variants. Allele frequencies were compared to the population sequencing dataset Exac ($n=33,000$). Replication was performed on 530 endometriosis subjects with whole exome sequencing data. Association testing was performed using Fisher's exact test. Nominal threshold was selected for significance ($p<0.05$). Panther software was used to test gene ontologies. A predictive score (E) was estimated for each subject as follows: $E=\sum \log(L95OR_j)*C_j$, in which C is a count of risk allele, L95OR is a lower limit of 95% CI of an odds ratio, and j is 1,2,3 ... n, wherein n is the number of the associated variants.

[0349] Results

[0350] 775 rare variants associated with endometriosis were identified, 561 of which were identified using Illumina Exome Beadchip, and 214 of which were identified using Affymetrix Axiom Custom Chip. **FIG. 1A-1B to FIG. 3** illustrate the results. Multiple low-frequency coding variants can be important in the genetic architecture of endometriosis. The relative risk of having endometriosis is significantly higher in women with multiple damaging variants, suggesting that they may serve as useful predictive or diagnostic markers. Genes involved with Wnt, cadherin, integrin, and inflammation mediated by cytokine signaling pathways are enriched, but trends did not reach significance.

Example 2. Genetic Variation Underlying the Clinical Heterogeneity of Endometriosis.

[0351] The study investigated whether two of the typical symptoms- pain and infertility may be linked to distinct genetic factors. A pool of 2818 non-synonymous SNP markers were selected

to classify markers associated with pain or infertility patients. In one group, cases were included that reported pain as their primary symptom but not infertility (n=727), and in the other group, cases were included with infertility as their primary symptom with only minimal or no pain (n=138). SNPs were then evaluated for significant variation between the two groups.

[0352] Methods

[0353] Genotyping. The samples were genotyped on a custom designed microarray using the Affymetrix Axiom platform per the manufacturer's instructions.

[0354] Statistical Analysis. Differences in allele frequencies between the two cohorts were tested for each SNP by a 1-degree-of-freedom Corchran-Armitage Trend test.

[0355] Ethnicity. Subjects were confirmed Caucasian ethnicity using principal component analysis.

[0356] Population Controls. The marker frequencies were compared to population control dataset of European Ethnicity (n=33,000; ExAc Database) to associate the marker to the respective group.

[0357] Consent and Medical Review

[0358] All subjects were provided written informed consent in accordance with study protocols approved by Quorum Review IRB (Seattle, WA 98101). Trained OB/GYN clinicians performed the medical record review and clinical assessment of each patient. Inclusion criteria in the endometriosis case population in the study were surgically confirmed diagnosis of endometriosis.

[0359] Results

[0360] The analysis identified nine SNP variants with differential prevalence between pelvic pain patients and infertility patients as shown in **Table 5**.

SNP	Gene	Chr	Pos	AA change	Allele Frequency			CPP vs. INF	
					ExAC	CPP	INF	P _{trend}	OR
Genes associated with chronic pain									
rs172562	TBX18	6	85,473,758	G48R	0.5706	0.4805	0.5766	0.0024	1.47
rs12339210	WHRN /DFNB31	9	117,170,241	P562A	0.1636	0.1007	0.1606	0.0040	1.69
rs35471617	COL21A1	6	56,033,094	T343M	0.1274	0.0639	0.1159	0.0021	1.92
rs72899872	LPR1B	2	141,232,800	A3178T	0.0127	0	0.0109	0.0001	∞
Genes associated with infertility									
rs8139422	CRELD2	22	50,315,363	D182E	0.0313	0.0282	0.0616	0.0040	2.27
rs78214713	OR51Q1	11	5,444,040	L204F	0.0066	0.0089	0.029	0.0259	3.33

rs7597367	<i>SCLY</i>	2	238,973,062	K60E	0.0006	0	<i>0.0073</i>	0.0011	∞
rs35880972	<i>BIRC8</i>	19	53,793,162	A156T	0.0004	0	<i>0.0072</i>	0.0012	∞
rs34505126	<i>BMP3</i>	4	81,967,240	T222M	0.0006	0	<i>0.0072</i>	0.0012	∞

[0361] Table 5 summarizes the results from a comparison of endometriosis associated variants with significantly different allele frequencies between patients with pelvic pain or infertility. ExAc can refer to frequencies reported by the ExAc consortium. CPP can refer to chronic pelvic pain and INF to infertility. Italic font indicates frequencies deviant from the general population.

[0362] The analysis identified five genes (*CRELD2*, *OR51Q1*, *SCLY*, *BIRC8*, *BMP3*) associated with infertility and four genes (*TBX18*, *WHRN*, *COL21A1*, *LRP1B*) associated with chronic pain. There was a sufficient power (>0.8) to detect markers with OR greater than 1.5 at significance level of 0.05. A review of the function of the genes identified can implicate several of the genes in both the pain and infertility pathways. Both *WHRN* and *TBX18* which show differential allele frequencies in patients with pelvic pain have been shown to be linked to pain-pathways. Mutations in *WHRN* have been linked to deafness and mechano- and thermo-sensitive deficiencies and can stabilize the paranodal region and axonal cytoskeleton in myelinated axons. *TBX18* is an important development regulator of the pericardium, prostate, nephrons, urogenital tubes, and seminiferous tubules and mutations in *TBX18* have been linked to pain in the chest, back, and flank. Conversely, *CRELD2* which show differential allele frequencies in infertility patients is linked with fertility. *CRELD2* is expressed in Oviductal epithelial cells in a manner that is very strongly correlated with the menstrual cycle and suggestive of an important reproductive role.

[0363] Pain and infertility can be two common but distinct clinical symptoms of endometriosis. In the present study, 9 non-synonymous variants were identified from a broad group of endometriosis associated variants that show distinct association with only one of the two symptoms and thus are suggestive of genetic classification of clinical subgroups of endometriosis.

Example 3. Novel High-Risk Damaging Mutations discovered in Familial Endometriosis.

[0364] Whole exome sequencing (WES) was used in endometriosis families to determine if inherited, rare, high-risk protein coding variants contribute to endometriosis. Endometriosis is a complex disease with underlying genetic and environmental factors. Array-based genotyping platforms are well suited for GWA studies detecting association with common variants (minor

allele frequencies >3-5%), whereas sequencing is required to detect rare and low-frequency protein coding variants. Subjects with familial endometriosis tend to carry a higher burden of genetic variants; families can be less likely to have potentially confounding (population stratification) effects. Studying genetic variants located on the same DNA strand (haplotypes) can help resolve the inheritance pattern of a disease variant by determining if two individuals who carry the same genetic variant have inherited the variant via shared recent ancestry (same haplotype) or whether their variants are derived from two independent mutation events (different haplotypes).

[0365] Methods

[0366] WES was performed on 489 women with familial endometriosis and 530 unrelated women (confirmed with identity-by-descent test) with endometriosis. WES was also performed using Ion Proton Instrument (**FIG. 4**) and AmpliSeq Exome Capture kit. All missense and protein truncating variants with a MAF<1% in ExAc database (Broad Institute) were considered for downstream analysis. Variant frequencies were compared with population frequency in ExAc database (n=33,000) using Fisher's exact test (exac. broadinstitute.org). Several software packages were used to predict whether the identified mutation would damage the encoded protein.

[0367] Consent and Medical Review

[0368] All subjects were provided written informed consent in accordance with study protocols approved by Quorum Review IRB (Seattle, WA 98101). Inclusion criteria were surgically confirmed diagnosis.

[0369] Results

[0370] This study identified 4 protein damaging variants significantly more prevalent in familial endometriosis. The 4 high-risk variants also pass genome-wide significance as shown in **Table 6** below. Association was verified for all but the BRD9 variant in the cohort of unrelated endometriosis patient.

Table 6. Four genes with low-frequency damaging mutations showing association to endometriosis.

Gene	Index mutation					Gene burden			
	AA _{change}	Endo _{Frq}	Exac _{Frq}	P	OR	Endo _{Frq}	Exac _{Frq}	P	OR
<i>LONP1</i>	splice	0.0028	Not seen	4.2×10^{-19}	Inf	0.0302	0.0199	2.6×10^{-2}	1.5[1-2]
<i>IGF2</i>	Q33X	0.0048	0.0009	3.0×10^{-10}	15[8-27]	0.0085	0.0014	3.0×10^{-5}	6[3-12]
<i>BRD9</i>	K39R	0.0009	0.0017	5.6×10^{-9}	10[5-21]	0.0057	0.0101	2.1×10^{-1}	0.6[0.3-1.3]
<i>SNAP91</i>	T555A	0.0106	0.0050	1.1×10^{-8}	5[3-8]	0.0179	0.0045	1.3×10^{-6}	4[2-6]

[0371] *LONP1* (Lon protease) is a nuclear encoded protease in the mitochondria responsible for the degradation of misfolded proteins. *LONP1* is expressed in endometrium and endometrial cancer, and affects endothelial mesenchymal transition in a dose dependent manner. Using a Genealogy database (GenDB) a shared ancestor ~ 13 generations ago was identified. All affected individuals shown with *LONP1* variant in **FIG. 5** share identical haplotype of ~140kb which is concordant with a single shared ancestor 11-15 generations in the past.

[0372] *IGF2* (Insulin-like growth factor 2) has previously been implicated in endometriosis in Korean women. The IGF axis has been implicated in growth regulation of endometriosis. In blood, *IGF2* is an imprinted gene expressed only from the paternal haplotype.

[0373] *SNAP91* (Synaptosome Associated Protein 91) and *BRD9* (Bromodomain Containing 9) are novel endometriosis candidates but little is known about their function.

[0374] This study identified low-frequency damaging protein mutations segregating in families with endometriosis. *IGF2* is the second implicated gene identified associated with endometriosis after *NLRP2*. Only 50 imprinted genes are known in humans to date suggesting imprinting plays a role in endometriosis. *LONP1* and *IGF2* regulate EMT in the pathogenesis of endometriosis.

Example 4. *CCDC168* and *MUC12* Show Recessive Effects in Women with Endometriosis.

[0375] Compound heterozygosity help identify genes involved in endometriosis. Whole Exome Sequencing (WES) was used on samples from 1,385 participants.

[0376] Samples

[0377] 1019 Endometriosis samples were sequenced, 530 of which were for discovery, 301 of which were for replication, and 188 of which were related (2nd cousin or closer). 366 control samples were sequenced.

[0378] Variant and Gene selection

[0379] Protein-altering variants in discovery w frequency <1% in ExAC. 3039 genes were found individuals with 2+ variants per gene in the discovery set and thus can possibly be recessive genes. **FIG. 6** illustrates mutation patterns cis/ trans/ haplotypes. Excess burden analysis of samples with 2+ protein-altering variants. Discovery (530 Endo vs 366 Ctl)- two genes with excess burden, $P_{\text{Fisher}} < 0.001$. Replication (301 Endo vs 366 Ctl)- both genes replicate, $P_{\text{Fisher}} < 0.05$.

[0380] Results

[0381] *CCDC168* and *MUC12* show significant excess variant count in endometriosis. Sample counts with rare protein-altering variants ($\text{ExAC}_{\text{freq}} < 1\%$)

Table 7. Variant count of *CCDC168*

95 Unique variants	2+	0-1
Cases	31	988
Controls	0	366
gnomAD (0.05)	1	365

Table 8. Variant count of *MUC12*

82 Unique variants	2+	0-1
Cases	47	970
Controls	1	365
gnomAD (0.14)	7	359

[0382] The variant counts of 2+ include all homozygotes, hemizygotes, and compound heterozygotes (cis and trans). Both genes show significant excess in endometriosis samples with 2+ hits also when compared with gnomAD.

[0383] The two novel genes, *CCDC168* and *MUC12*, have large recessive effects in endometriosis and can be biologically relevant in endometriosis. 7.6% of endometriosis patients can have compound heterozygote mutations with 4-30 fold excess compared with control populations.

[0384] *CCDC168* is coiled-coil domain containing 168. *CCDC168* can be differentially expressed in malignancies. Antibody staining can show prominent staining in various epithelial tissues. In some instances, *CCDC168* is only present in placental animals (those with endometrium).

[0385] *MUC12* is a transmembrane mucin expressed across many epithelial tissues including colon, pancreas, prostate or uterus. In some instances, transmembrane mucins are single-stranded proteins undergo proteolytic cleavage splitting TM and EC domains, lubricate epithelial surfaces, bind ligands, regulate epithelial wound healing, and/or extracellular domain detach with excess force (intracellular signaling and EMT). In some instances, a transmembrane mucin disclosed herein is *MUC1*, *MUC4*, *MUC12*, or *MUC16*. The extra cellular domain of *MUC16* can be cancer antigen 125 (CA125), an important marker of ovarian cancer and endometriosis.

Example 5. Rare Synonymous Mutations Show Strong Association with Endometriosis

[0386] The study is to determine if rare synonymous variants might contribute to the genetic risk for developing endometriosis. Synonymous and non-synonymous DNA variants can occur within the protein-coding part of a gene. Synonymous variants do not affect the amino-acid sequence, and non-synonymous variants do affect the amino-acid sequence, due to the redundancy in the genetic code. GWAS intergenic SNP variants may be determined from eQTL fine mapping, and rare non-synonymous variants may be determined from Whole Exome Sequencing.

[0387] Methods

[0388] Whole exome sequencing was performed on 1,077 study participants with surgically diagnosed endometriosis. Saliva DNA underwent AmpliSeq sequencing on an Ion Proton, and sequence was assembled using the Torrent software. Variant frequencies were compared to frequencies in gnomAD, which was used as reference for population-wide variant frequencies. Synonymous variants with a minor allele frequency <0.01 in the general population were considered. Fisher's Exact test was used to calculate association statistics. PANTHER database was used for GO (Gene Ontology) term enrichment analysis.

[0389] Results

[0390] 114,877 synonymous rare variants were identified among patients. 648 synonymous variants passed the nominal significance threshold ($p < 0.05$) across 617 genes. **Table 9** shows five variants strongly associated with endometriosis that pass the genome-wide significance threshold of $p \leq 5 \times 10^{-8}$.

Table 9. Five strongly associated synonymous variants

Gene	Chr	Position	P	OR	Nucl change	Amino Acid
<i>KRTAP5-1</i>	11	1,606,402	2.0×10^{-11}	43	C78T	S26S
<i>GPR137</i>	11	64,051,889	6.7×10^{-15}	49	G51A	G17G
<i>UBC</i>	12	125,398,297	1.5×10^{-33}	94	T21C	T7T
<i>ADAMTS7</i>	15	79,058,944	2.5×10^{-11}	11	T3309A	A1103A
<i>SYNE1</i>	6	152,457,795	6.7×10^{-8}	5	G25617A	E8539E

[0391] 17 genes have 2-or-more rare synonymous disease associated variants were found with only one expected by chance ($p < 0.001$): *ABCC5*, *ANK3*, *ATP8B4*, *CCDC147*, *CELSRI*, *DNAH3*, *EML6*, *HERC2*, *ITGA2*, *KIF23*, *LAMA5*, *PKD1*, *SLC22A20*, *SSPO*, *TENM2*, *TUBGCP2*, *VPS18*. GO-term analysis show significant enrichment of a single GO term:

“*cytoskeletal structure and regulation*” (OR=13.4). Rare intronic splice-junction variants were considered among the 17 genes, and 5 variants in *CCDC147*, *LAMA5*, and *SSPO* may affect the risk-burden.

[0392] This is the first time that rare synonymous variants may have been implicated in endometriosis. The genes may carry these mutations that are enriched for cytoskeletal function. Go-term and functional analysis implicate cytoskeletal regulation in the genetic predisposition of endometriosis. These variants may prove useful in developing a non-invasive test for endometriosis.

Example 6. Large Effect Mutations in Endometriosis Genes Implicated by GWAS.

[0393] Genome-wide association studies (GWAS) implicate several chromosomal regions as genetic risk factors for endometriosis. These regions have been "tagged" by polymorphic markers located between genes or in non-coding introns. Sequenced were the exons of 16 genes in GWAS regions to search for causative mutations, i.e., to find gene mutations responsible for the association observed in 16 genes implicated by endometriosis GWAS.

[0394] Methods

[0395] AmpliSeq sequencing on Ion Protons was conducted on DNA samples from 1,019 women with confirmed endometriosis. After sequence assembly using Torrent software, variant annotation was performed using ANNOVAR (hg19 reference). Frequencies of coding variants were compared against a large reference dataset (sequence data from 63,369 non-Finnish Europeans in gnomAD). Variants were found using Torrent Variant Caller (UCSC hg19). Association statistics were calculated using Fisher's Exact test; linkage disequilibrium statistics were calculated using LDlink. Cases: n=1,019 European women with confirmed endometriosis. Controls: n=63,369 non-Finnish Europeans in gnomAD).

[0396] Results

[0397] 571 variants were detected; 333 of these alter an amino acid in the encoded protein and 234 low-frequency (MAF<1%), missense mutations are predicted to be pathogenic (in-silico). Likely pathologic variants are uncommon in the reference data (which contains women with endometriosis and males carrying risk factors); but the identified variants were often seen in multiple endometriosis patients. The excess of pathogenic mutations in cases was striking ($p < 10^{-16}$). 4 mutations (see **Table 10**) have high odds ratios for endometriosis with p values well below a multiple testing threshold ($p \leq 9 \times 10^{-5}$). Mutations predicted to shorten the encoded protein (loss of function) were also detected (2 splicing changes, and 7 “stop” mutations). Stop mutations (seen in five genes: *GREB1*, *NFE2L3*, *FNI*, *SYNE1* and *VEZT*) were more prevalent

in the endometriosis cohort compared to the population data ($p=1.7 \times 10^{-13}$). There is no measureable linkage disequilibrium between any of the new variants and tagging GWAS markers. **FIG. 7** to **FIG. 9** further illustrate the results.

Table 10. Mutations with p values below multiple correction threshold. Inf means that the variant was not observed in the control cohort.

Gene	Protein change	Control Frequency	Endometriosis Frequency	p(fisher)	Odds Ratio [L95-U05]
<i>FNI</i>	p.V527M	Not seen	0.00147	4.03E-06	Inf.
<i>NFE2L3</i>	p.I233V	Not seen	0.00147	4.03E-06	Inf.
<i>SYNE1</i>	p.E8539E	0.00206	0.00785	1.11E-05	3.84
<i>VEZT</i>	p.P712S	0.00005	0.00196	1.23E-05	41.50

[0398] This is the first comprehensive study of coding mutations in all 16 GWAS candidate genes. Coding variants may not explain the association observed in GWAS studies, thus regulatory mutations outside of the coding regions are likely to be involved. The mutations having large effects confirm an important role for these genes in the pathogenesis of endometriosis.

[0399] Example 7. Detailed Methods for Detection of Low Frequency Variants

[0400] Medical Review.

[0401] The inclusion criteria in the endometriosis case population in the present study were surgically confirmed diagnosis of endometriosis with laparoscopy being the preferred method. Trained OB/GYN clinicians performed the medical record review and clinical assessment of each individual patient. Patients were considered to be affected if they had biopsy-proven lesions or if operative reports revealed unambiguous gross lesions. Patients were further categorized by severity, clinical history of pelvic pain, infertility, dyspareunia or dysmenorrhea and family history. Patients were grouped into one of three classes of severity: mild, moderate or severe, following the general guidelines set forth by ASRM. This analysis compared cases with 100% prevalence of endometriosis to controls with the population prevalence of endometriosis (5–10%).

[0402] DNA Extraction.

[0403] Saliva samples were collected using the Oragene 300 saliva collection kit (DNA Genotek; Ottawa, Ontario, Canada) and DNA was extracted using an automated extraction instrument, AutoPure LS (Qiagen; Valencia, CA), and manufacturer's reagents and protocols.

DNA quality was evaluated by calculation absorbance ratio OD260/OD280, and DNA quantification was measured using PicoGreenH (Life Technologies; Grand Island, NY).

[0404] Microarray Genotyping.

[0405] The discovery set of 2019 endometriosis cases and 25476 population controls were genotyped using the Illumina Human OmniExpress Chip (Illumina; San Diego, CA) according to protocols provided by the manufacture. An additional 905 endometriosis cases were genotyped on a custom designed microarray using the Affymetrix GeneTitan platform according to the manufacturer's instructions.

[0406] Sample Quality Control.

[0407] Samples were excluded from the analysis if they missed any of the following quality thresholds:

- a) Evidence of familial relationship closer than 3rd-degree ($\pi^2 > 0.2$) using genome-wide Identity-By-State (IBS) estimation implemented in PLINK
- b) Samples with missing genotypes > 0.02
- c) Samples with non-European admixture > 0.05 as determined by ADMIXTURE

[0408] SNP Quality Control.

[0409] SNPs were excluded from the analysis if they missed any of the following quality thresholds:

- a) SNPs from copy number variant regions or regions with adjacent SNPs
- b) SNPs failing Hardy-Weinberg Equilibrium (HWE) $P \leq 10^{-3}$
- c) SNPs with minor allele frequency (MAF) ≤ 0.01 in the control population
- d) SNP call rate $\leq 98\%$

[0410] Admixture.

[0411] ADMIXTURE (ver. 1.22) was used to estimate the individual ancestry proportion. The software estimates the relative admixture proportions of a given number of a priori defined ancestral groups contributing to the genome of each individual. The POPRES dataset (Nelson MR et al. 2008) was used as a reference group to create a supervised set of 9 ancestral clusters. Seven of them belong to the European subgroups along with African and Asian groups. Since POPRES dataset utilized Affymetrix 5.0 chip, 105,079 autosomal SNPs that overlapped with the Illumina OmniExpress dataset were used. Among the 105,079 SNPs, a subset of 33,067 SNPs was selected that showed greater genetic variation (absolute difference in frequency) among the 9 reference groups. The pair-wise autosomal genetic distance determined by Fixation Index (F_{ST}) using 33,067 SNPs was calculated for the 9 reference groups as listed in POPRES dataset.

Subsequently, a conditional test was used to estimate the admixture proportions in the unknown samples as described by Alexander et al. (2009).

[0412] Principal Component Analysis (PCA).

[0413] PCA was applied to account for population stratification among the European subgroups. The previously identified 33,067 SNPs were selected to infer the axes of variation using EIGENSTRAT . Only the top 10 eigenvectors were analyzed. Most of the variance among the European populations was observed in the first and second eigenvector. The first eigenvector accounts for the east-west European geographical variation while the second accounts for the north-south component. Only the top 10 eigenvectors showed population differences using Anova statistics ($p < 0.01$). The PCA adjusted Armitage trend P-values were calculated using the top 10 eigenvectors as covariates.

[0414] Association Analysis.

[0415] After the quality of all data was confirmed for accuracy, genetic association was determined using the whole-genome association analysis toolset, PLINK (ver. 1.07) .Differences in allele frequencies between endometriosis patients and population controls were tested for each SNP by a 1 degrees of freedom Cochran-Armitage Trend test. The allelic odds ratios were calculated with a confidence interval of 95%. SNPs that passed the quality control parameters were prioritized using the PCA adjusted cochran-Armitage trend test P-values. The combined/metaanalysis of different datasets was performed using Cochran-Mantel-Hanszel method as well as using Cochran-Armitage Trend test. Breslow Day test was used to determine between-cluster heterogeneity in the odds ratio for the disease/SNP association.

[0416] Software Used.

[0417] PLINK (version 1.07). R (version 2.15.0). EIGENSTRAT (version 3.0).

[0418] Example 8. Detailed Methods for Gene Sequencing and Detection of Low-Frequency Damaging Variants

[0419] DNA extraction and Genotyping.

[0420] DNA used in the present study was extracted from blood or saliva using standard extraction methods. Genotyping was performed using the Illumina HumanExome (Illumina, San Diego, CA) according to protocols provided by the manufactures.

[0421] Sample and SNP Quality Control

[0422] The discovery set of 1518 cases were genotyped using the Illumina Human Exome Chip (Illumina; San Diego, CA) per protocols provided by the manufacture.

[0423] Samples were excluded from the analysis if they missed any of the following quality thresholds:

- a) Evidence of familial relationship closer than 3rd-degree ($\hat{\pi} > 0.2$) using genome-wide Identity-By-State (IBS) estimation implemented in PLINK.
- b) Samples with missing genotypes > 0.02
- c) Samples with non-European admixture > 0.05 as determined by ADMIXTURE

[0424] SNPs were excluded from the analysis if they missed any of the following quality thresholds:

- a) SNPs with Illumina GenTrain Score < 0.65
- b) SNPs from copy number variant regions or regions with adjacent SNPs
- c) SNP call rate $\leq 98\%$

[0425] Exome Sequencing and Variant Discovery

[0426] Whole exome sequencing (WES) was performed on 2400 endometriosis cohort using Ion Proton Instrument as per the manufacturer's protocol (Life Technologies, Carlsbad CA) using their AmpliSeq Exome Capture Kit. Sequence alignment and variant calling was performed against the reference human genome (UCSC hg19 version). The variant discovery was performed using Life Technologies TMAP algorithm with their default parameter settings, and Life Technologies Torrent Variant Caller was used to discover variants. The variants identified from the Torrent Variant Caller were taken further for downstream analysis. The variants included were single nucleotide variants, short insertions, or deletions. Variant annotation was performed using ANNOVAR. The coding variants were classified as missense, frameshift, splicing, stop-gain, or stop-loss. Variants were considered "loss-of-function" if they caused a stop-gain, splicing, or frame-shift insertion or deletion. Prediction of protein function was evaluated in silico using seven different algorithms (Polyphen 2, Sift, Mutation Accessor, Mutation Taster, FATHMM, LRT, and MetaLR. Missense variants were deemed "damaging missense" if they were predicted damaging by at least one of the seven algorithms tested. The genes that harbor these variants were also checked against the published "FLAGS" gene list (Shyr C et al. 2014) to understand whether the gene is frequently mutated in humans.

[0427] Low Frequency Variants

[0428] Variants that pass the population control frequency (gnomAD) of $MAF < 1\%$ were called "low frequency variants". These variants were analyzed to test for association using Fisher's Exact Test. The low frequency variants were prioritized based on their Fisher's p value.

[0429] Gene Burden

[0430] The genetic burden was calculated for each gene by collapsing/combining all low frequency variants identified through WES. Fisher's Exact Test was used to determine excess gene burden in endometriosis subjects compared to the control population counts as observed in gnomAD database by generating 2×2 table per gene for the number of reference and alternative alleles. The genes were then prioritized based on their Fisher's p value.

[0431] Example 9. Whole exome sequencing identifies markers of endometriosis.

[0432] Twin and family studies show that heritability for endometriosis may be high, yet the GWAS markers and copy number variants identified explain about 5% of the heritability. Multigenerational pedigrees can be used to identify variants/genes with large effects in complex diseases. A large endometriosis family spanning 19 generations with 218 women with surgically confirmed disease was used for this study. For endometriosis, one cannot assume that all distant relatives share a single causative mutation. However, segregation analyses suggest that autosomal major gene effects may be likely. Referring to **FIG. 12**, whole exome sequencing (WES) was performed on 137 women with surgically confirmed endometriosis having a common ancestor born in 1608. The WES was utilized to search for pathogenic mutations. All coding variants were evaluated. Variants may be deemed damaging if they were predicted to be damaging “in-silico” by at least one algorithm of Polyphen 2, Sift, Mutation Accessor, Mutation Taster, FATHMM, LRT, or MetaLR. The excess variant burden in this family was compared against population controls (e.g. exome sequencing data for the non-Finnish European cohort in the gnomAD database, wherein n = 55,860).

[0433] A damaging missense variant (p.T166I) in SEPT10 gene seen in about 0.18% of controls was seen in a 20 fold excess (3.6%) in the large endometriosis family ($p=1.98 \times 10^{-10}$; Odds Ratio (OR) = 20.6 [10.8=39.30]. SEPTIN10 is a cytoskeletal protein that may have GTP-binding and GTPase activity. Referring to **FIG. 13**, septins may polymerize into heterooligomeric protein complexes that form filaments, and can associate with cellular membranes, actin filaments and microtubules.

[0434] Referring to **FIG. 14**, a TNFRSF6B haplotype (also called DcR3) which spans 41kb was present in 76 of the 137 subjects with surgically confirmed endometriosis ($p=1.7 \times 10^{-8}$; Odds Ratio (OR) = 2.11); 21 of the affected women were homozygous or hemizygous. These results suggest mutations in one or more genes - TNFRSF6B and SEPT10 – may be diagnostic for predicting a risk of developing or confirming a presence of endometriosis.

[0435] Example 10. Whole Exome Sequencing Identifies Markers of Endometriosis.

[0436] Endometriosis affects from about 6% to about 10% of women during their reproductive years with symptoms including pelvic pain, dyspareunia, dysmenorrhea, infertility, or any

combination thereof. Twin and family studies show that heritability for endometriosis may be high; yet common genetic variations identified by genome-wide association studies only explain about 5% of the heritability. It is possible that rare and recent familial mutations, not detectable by GWAS, may be responsible for part of the missing heritability. Next generation sequencing gives the opportunity to look for less-common variants with large effects. In this study, we used whole exome sequencing (WES) to identify inherited deletion variants in a three generation family of seven affected women with surgically confirmed endometriosis.

[0437] Exome sequencing was performed using the AmpliSeq technology on Ion Proton platform (Thermo Fisher, Inc) according to manufacturers instructions. Variants were determined using Ion Proton protocol and confirmed using the GATK (Genome Analysis Toolkit) pipeline. Segmental deletions were identified by observing three-or-more homozygous neighboring variants in the matriarch that failed to segregate in a Mendelian manner in her daughters/grand-daughters.

[0438] A three-generation family with seven affected members is shown in **FIG. 15** together with notable symptoms tabulated to the right of the pedigree.

[0439] Case 1 was the first individual in the family to be diagnosed with endometriosis and underwent surgical hysterectomy at age 32 due to stage IV bilateral ovarian endometriosis. Her mother (not shown in the pedigree) had four children with no gynecological problems. However, her three daughters (cases 2-4) and three grand-daughters (cases 5-7) all have been surgically diagnosed with endometriosis. In addition to endometriosis, case 1 has been diagnosed with 14 other morbidities including Crohn's disease, interstitial cystitis, bronchial asthma, cardiovascular diseases, lupus erythematosus and multiple sclerosis, all of which have been positively associated with endometriosis.

[0440] Referring to **FIG. 16**, we identified approximately 20,000 exonic variants in each of the seven individuals, and almost 34,000 variants combined across the pedigree, which is in line with expectations. IBD and segregation analysis confirmed all individual relationships and the overall pedigree structure. We identified two hemizygous deletions segregating in this three-generation family. A deletion was found in UGT2B28 (UDP Glucuronosyltransferase Family 2 Member B28), spanning seven informative sequence variants across at least 14kb, a deletion in USP17L2 (Ubiquitin Specific Peptidase 17-Like Family Member 2) spans three informative variants across at least 2kb. Both deletions are present in the affected grand-mother and segregate in as many as four and five of her descendants respectively. Further in-depth analysis of normally segregating rare variants is ongoing.

[0441] These results implicate UGT2B28 and USP17L2 in the pathogenesis of endometriosis. UGT2B28 is phase II detoxification gene involved in glucuronidation of many substrates including steroid hormones and lipid-soluble drugs. USP17L2 is a deubiquitinase that regulates key cellular processes like proliferation, migration and apoptosis through the activation of small GTPases like RAC1A, CDC42 and RHOA, and the regulation of adherence junctions. USP17L2 plays a central role in the regulation of the transcription factors SNAIL, SLUG and TWIST which are key gate-keepers of epithelial-to-mesenchymal transition (EMT). Dosage dependent loss of USP17L2 may affect mesothelial integrity and may increase the risk for developing endometriosis.

[0442] Example 11. Segregation Analysis of Families: Dominant Genes Contribute to Pathogenesis of Endometriosis.

[0443] The largest endometriosis family reported to date: spanning 19 generations with 218 women with surgically confirmed endometriosis were utilized in this study. An autosomal major gene effect may be likely in this family. Risk of endometriosis in 123 smaller families with probands selected from the same time period and unrelated to the index pedigree were examined. A genetic segregation analysis was performed to identify large pedigrees with familial endometriosis, observe segregation patterns for surgical endometriosis, and compare segregation patterns with dominant patterns seen in the 137 subject index pedigree.

[0444] Referring to **FIG. 18**, over 1900 women from Intermountain West US with surgically confirmed endometriosis were included in the study. A 3-generation pedigree was obtained for each affected woman. A genealogy database (GenDB) was utilized to find common ancestors linking one or more grandparents of probands. Probands with no genealogy available were eliminated. Probands with a birth date between 1960 and 1995 were selected arriving at an index pedigree of 89 subjects having surgically confirmed endometriosis and unrelated pedigree of 123 subjects having surgically confirmed endometriosis. Percentage of affected subjects in both the index pedigree and unrelated pedigrees is shown in **FIG. 19** and the rate of surgically diagnosed endometriosis is shown in **FIG. 20**.

[0445] Referring to **FIG. 17**, prevalence of endometriosis in close relatives is much higher than the 2-3% prevalence of surgically diagnosed endometriosis in the general population. The rates observed are also higher than expected with multifactorial polygenic inheritance.

[0446] Analysis of larger number of families confirms that autosomal dominant, high penetrance risk alleles for endometriosis segregate in families. The heritability of endometriosis may be higher than estimated by older twin studies.

[0447] Table 1. Variants associated with endometriosis. Inf means that the variant was not observed in the control cohort.

Chr	Position	Reference Allele	Alternate Allele/ Minor Allele	Gene	Amino Acid position	Case MAF	Control MAF	p value	OR[L 95-U95]	Context Sequence	SEQ ID NO
chr 1	1139216	G	A	<i>TNFRSF18</i>	p.R175C	0.00686	0.00437	2.97E-02	1.57[1.07-2.31]	CCTGGGGAG GGGCTGGCT GCGGTCGGT GGCCCCGGA GGAC[G/A]G CCAGGCTCA CACCCACAG GTCTCCCAG CCGCCCTTCTC	SEQ ID NO: 1
chr 1	1452592	T	C	<i>ATAD3A</i>	p.W110R	0.00735	0.00038	2.93E-22	19.24[11.09-33.38]	GCTGGAAGC CCTGAGCCT GCTGCACAC ACTAGTCTG GGCA[T/C]GG AGTCTCTGC CGTGCCGGA GCCGTGCAG ACACAGGAGCG	SEQ ID NO: 2
chr 1	3703589	C	T	<i>LRRC47</i>	p.V301M	0.00662	0.00412	2.53E-02	1.61[1.09-2.38]	ACGTGCAGG ACCCTGAGC AGCAGCCGG CCGGCATCT CCCA[C/T]GT CCTGCTCCTC CCCATCACC ACCTTCCCG CCTCTGCTTC	SEQ ID NO: 3
chr 1	9083112	G	T	<i>SLC2A7</i>	p.T59N	0.00613	0.00361	1.69E-02	1.7[1.14-2.55]	GAGCTTCCC GTCCATGAA TGTTGCGTG TCGCTCAAA GTAG[G/T]TT TCGTTGTAA AATGACTTG AAGACCTGG AAAACATTGCC	SEQ ID NO: 4
chr 1	10529326	A	G	<i>DFFA</i>	p.I69T	0.00760	0.00520	4.67E-02	1.46[1.02-2.1]	ATTGGAAGG TAGACACAG AAAGTAATC GTCATCATC CACT[A/G]TG GTGCCATCC TCTGCCAGG ACCAGGGTG ACTGGTGTCAG	SEQ ID NO: 5

chr 1	119833 52	C	T	<i>KLAA2</i> <i>013</i>	p.E410 K	0.005 88	0.003 42	1.44 E- 02	1.72[1.14- 2.61]	ATCTGCTGG ACGGAGGAC AGCCGCCCC GGCCACAGG TTCT[C/T]GG CGTGCATGG TGCGGTGCC CGCTGAAGC AGTGATCTT CA	SEQ ID NO: 6
chr 1	128559 96	A	G	<i>PRAM</i> <i>EF1</i>	p.N42 6D	0.005 39	0.003 31	3.74 E- 02	1.63[1.06- 2.52]	TCCTGCCCC GAGGAGAGT TTGAATTCCT TGGTTCGTG TC[A/G]ATTG GGAGATCTT CACCCCACT TCGGGCTGA GCTGATGTG	SEQ ID NO: 7
chr 1	128560 79	C	T	<i>PRAM</i> <i>EF1</i>	p.G45 3G	0.014 22	0.003 01	6.69 E- 20	4.78[3.6- 6.33]	CACCTAGGG AAGTCAGGC AGCCCAAGA GGATCTTCA TTGG[C/T]CC CACCCCTG CCCTTCCTGT GGCTCATCA CCGTCTGAG G	SEQ ID NO: 8
chr 1	136692 76	C	T	<i>PRAM</i> <i>EF14</i>	p.E352 K	0.006 86	0.000 03	5.37 E- 35	201. 46[6 1.22- 662. 92]	TGGGAGTAG TGGATCTGA CAGCCCTCC AAGATGAGG GTTT[C/T]GA GAGAGGCAG CAATTTTCTC TAGCAGAGC TCCGAGGGG T	SEQ ID NO: 9
chr 1	159869 77	A	T	<i>RSC1A</i> <i>1</i>	p.N20 5I	0.005 205	0.002 931	2.92 E- 02	1.78[1- 2.94]	AACATAGGG GACCTTGAG CTTCCTGAA GAAAGGCAA CAGA[A/T]TC AACACAAAA TTGTTGATTT GGAAGCTAC GATGAAAGG A	SEQ ID NO: 10
chr 1	176033 40	C	T	<i>PADI3</i>	p.H50 8H	0.009 07	0.006 19	2.64 E- 02	1.47[1.05- 2.05]	CCTGCTTCA AGCTCTTCC AGGAAAAGC AGAAGTGTG GCCA[C/T]GG GAGGGCCCT CCTGTTCCA GGGGGTTGT TGGTGGGTA AC	SEQ ID NO: 11

chr 1	194511 76	A	T	<i>UBR4</i>	p.A31 49A	0.011 27	0.008 21	4.27 E- 02	1.38[1.02- 1.86]	TTTCTGTTAG AAGCTGAGT ATAGGCCTC AAACACATC AGC[A/T]GCA TGACCCTGG GAGAAGAAA ATTTGCATG AGAACCTGT G	SEQ ID NO: 12
chr 1	195040 62	T	C	<i>UBR4</i>	p.M84 4V	0.011 27	0.008 19	4.24 E- 02	1.38[1.03- 1.86]	CGAGCCAAG ATAAGCGGC ACGAAGCGC ATCTGAGCA TCCA[T/C]GT TGACGCTCA ACTCCTGGA TGATCTGGA CAAAAAGCG AC	SEQ ID NO: 13
chr 1	195458 93	G	A	<i>EMC1</i>	p.Y96 1Y	0.011 52	0.008 23	2.81 E- 02	1.4[.05- 1.89]	TGGCAAAAA CCAGGCCAA AGAGGACGC TGCTGATTA ACAC[G/A]TA GTCATAGTC ATCCTTCAG AACGTCAAA CTGCTTGGA TG	SEQ ID NO: 14
chr 1	204428 78	C	T	<i>PLA2G 2D</i>	p.G45 S	0.009 80	0.006 95	3.60 E- 02	1.41[1.03- 1.95]	TCTTTGGGTT GGCCTCTGC CACCTAGTC CGCAGTGAC AGC[C/T]GTA GGGCCAGTA GGAGAGGAT GGGCATTTT CCCAGTCAC T	SEQ ID NO: 15
chr 1	238455 89	A	T	<i>E2F2</i>	p.A25 7A	0.011 52	0.008 56	4.78 E- 02	1.35[1.01- 1.81]	CCTTGACGG CAATCACTG TCTGCTCCTT AAAGTTGCC AAC[A/T]GCA CGGATATCC TGGTAAGTC ACATAGGCC AGCGTAGGG C	SEQ ID NO: 16
chr 1	244881 31	C	T	<i>IFNLR I</i>	p.E137 E	0.009 31	0.006 39	3.42 E- 02	1.46[1.04- 2.05]	TGCAGGGGG GCAGCTGGT ACGTGGCAT TGGCACTCA GGAT[C/T]TC CTCCGTCTG GGTGAGCAC CAGGACAGG TGGGGCCGG CT	SEQ ID NO: 17

chr 1	266088 83	A	G	<i>UBXN 11</i>	p.G49 0G	0.008 82	0.000 20	5.67 E- 34	44.6 2[23. 65- 84.2]	CGGGACTGG GGCCGGGAC CGGGACCGG GACTGGGGC CGGG[A/G]CC GGGACCGGG ACAGGGACC AGGACTGAA TTTCAGGCT GG	SEQ ID NO: 18
chr 1	266714 13	G	C	<i>AIMIL</i>	p.P579 R	0.018 63	0.000 00	5.40 E- 89	Inf	TGAGGCAGC AGGAGCACC AGGGCCCTT CACAACCTC TTTT[G/C]GG GTGGTGGAC AAGGCAGCA GGAGCACCA GACCCCTGC AC	SEQ ID NO: 19
chr 1	266716 25	A	G	<i>AIMIL</i>	p.S508 S	0.025 98	0.000 21	4.40 E- 88	127. 41[5 5.94- 290. 2]	CAGGAGCAC TGGACCCCT GCACCACCT CCTTCTGGG TGGG[A/G]G ATGAGGCAG CAGGAGCAC CAGGGCCCT TCACGACCT CTT	SEQ ID NO: 20
chr 1	276743 34	G	A	<i>SYTL1</i>	p.A12 6T	0.005 88	0.000 00	1.34 E- 35	Inf	CCCAGGAGA CCAGGCTCC AGGCCACGA CAGGGAGGC TGAG[G/A]CT GCTGTGAAA GAGAAGGAA GAGGGGCCA GAGCCCAGG TG	SEQ ID NO: 21
chr 1	289319 01	T	A	<i>TAF12</i>	p.T145 S	0.006 86	0.004 65	4.77 E- 02	1.48[1.01- 2.17]	CCAGGGCTC TGGCATTTC CTCACCTGTT TGTGAGCTT CTG[T/A]GGT GCAAGCTTT TTTGTAGGG TCGGATTTCT TCAGAGCCA	SEQ ID NO: 22
chr 1	294477 92	G	A	<i>TMEM 200B</i>	p.C183 C	0.005 64	0.002 91	4.90 E- 03	1.94[1.27- 2.97]	TACCCCGAC GCGGGGACG GGTCCCAGA TTTCTGGCTC TGC[G/A]CAG CCTACGGCT CGGGGACTC CTAGGGCCG GGGCTGGGA A	SEQ ID NO: 23

chr 1	314096 34	A	G	<i>PUM1</i>	p.A10 97A	0.005 39	0.003 29	3.69 E- 02	1.64[1.07- 2.53]	GGGGACCGT CGTTCATGG TGCACACCT CATCGATGA GCAC[A/G]GC GCGCTCCGT ACGTGAGGC GTGAGTAAC ACACTTCTC CA	SEQ ID NO: 24
chr 1	353707 38	C	A	<i>DLGA P3</i>	p.G83 W	0.013 24	0.000 04	1.13 E- 63	301. 95[9 4.37- 966. 13]	TGGCCAGGG TACATCCTG GGGAAGGTG CTGCTACCC CCCC[C/A]AA CCCCGGCCC CCGCTGGCC CTCCCTCAG GGCCTACCG AC	SEQ ID NO: 25
chr 1	405332 89	C	G	<i>CAPI</i>	p.C236 W	0.029 90	0.000 00	5.89 E- 176	Inf	GACCCTCTG CCGATCAT GTCCTCCTCC CCCTCCACC ATG[C/G]CCC CCTCCTCCCC CAGTCTCTA CCATTTTCAT GCTCATATG	SEQ ID NO: 26
chr 1	407023 42	A	G	<i>RLF</i>	p.T656 T	0.011 03	0.007 87	3.11 E- 02	1.41[1.04- 1.9]	TGAATGACC AAGCCAAAG GAGAGTCTC ATGAATATG TCAC[A/G]TT CAGCAAATT AGAAGATTG CCACCTGCA AGACAGAGA TT	SEQ ID NO: 27
chr 1	409289 69	A	G	<i>ZFP69 B</i>	p.Q43 8R	0.006 86	0.004 12	1.30 E- 02	1.67[1.14- 2.45]	AGTAAAACC TTCAGCCAT AGTACATAC CTAACTCAA CACC[A/G]GA GAACTCATA CTGGAGAAA GACCATATA AATGTAAGG AA	SEQ ID NO: 28
chr 1	476914 81	G	C	<i>TAL1</i>	p.A27 G	0.005 15	0.000 01	6.06 E- 28	460. 35[6 1.91- 3423 .16]	CTCCTTGGC GACGCCGTT CAGCAGGAC CAGGTGCGG GGGG[G/C]CC ATGCTGGCC TCGGCCGCG TCCCGTCCCT CTAGCTGGG G	SEQ ID NO: 29

chr 1	477168 89	C	T	<i>STIL</i>	p.T126 2T	0.009 56	0.006 79	4.22 E- 02	1.41[1.02- 1.95]	AGAAGGTGC CTACTGAAT TCATGCTATT CATCTGCTTT AG[C/T]GTTT CAGAAGGTT GCAAACCTT CAGGAAAAA TTGTAATGT	SEQ ID NO: 30
chr 1	556436 58	T	C	<i>USP24</i>	p.T158 A	0.007 11	0.004 71	4.65 E- 02	1.51[1.03- 2.22]	GTTCTAAGG TCTGAAAAC TTACCAAGT CTTGCTAGG TAGG[T/C]AG ATGCCACA GGCATTTC CTAGTGATT CTTCTCGCTT G	SEQ ID NO: 31
chr 1	953304 40	C	T	<i>SLC44 A3</i>	p.N42 4N	0.006 62	0.004 23	2.78 E- 02	1.57[1.06- 2.32]	TGGTGAGGA TTCCGAGAA TCATTGTCAT GTACATGCA AAA[C/T]GCA CTGAAAGAA CAGGTAAGG CTACCTCCT GATACACAG C	SEQ ID NO: 32
chr 1	109792 751	T	C	<i>CELSR 2</i>	p.L17P	0.009 80	0.000 47	2.93 E- 32	21.0 6[13. 61- 32.5 9]	CCGGCCACC GGCGTCCCC CTCCCAACG CCGCCGCCG CCGC[T/C]GC TGCTGCTGTT GCTGCTGCT GCTGCCGCC GCCACTATT G	SEQ ID NO: 33
chr 1	110302 392	A	T	<i>EPS8L 3</i>	p.F55I	0.006 13	0.003 20	3.20 E- 03	1.92[1.28- 2.88]	AAGTCTTGG CTCCACACC CGGCCCTGT GCATCCATC TCGA[A/T]CA GCTTCTGCA AGGCATCCT CGGGCCCCT GGACTCTCT GA	SEQ ID NO: 34
chr 1	117122 350	T	C	<i>IGSF3</i>	p.K10 20E	0.025 25	0.000 00	1.05 E- 150	Inf	CTTTCCTCTT CCTGTTCTTC CAGGCCAGG GCTGCTCCTT T[T/C]CCCC CAGCTTTAG TCCTCAGGG AATACCAGG CCACAGCG	SEQ ID NO: 35

chr 1	120054 192	G	T	<i>HSD3 B1</i>	p.R71I	0.010 54	0.007 17	1.85 E- 02	1.48[1.08- 2.01]	GTGCTGGAA GGAGACATT CTGGATGAG CCATTCCTG AAGA[G/T]A GCCTGCCAG GACGTCTCG GTCATCATC CACACCGCC TGT	SEQ ID NO: 36
chr 1	144856 852	C	T	<i>PDE4 DIP</i>	p.A21 05A	0.009 07	0.005 92	1.71 E- 02	1.54[1.1- 2.14]	TTACCTCTGT GCCTTGGGC TTCAAGGCC AGGGAAGCT GCA[C/T]GCT GATCTCACA AGAGACACT ATCTTTTGA CCAGCAGCT	SEQ ID NO: 37
chr 1	144912 191	G	T	<i>PDE4 DIP</i>	p.P695 H	0.005 15	0.002 16	5.35 E- 04	2.39[1.53- 3.74]	ACAGCCAGT GGGGGTAAC TTCAGCTTGT TGGTTAGAG ATG[G/T]GTG CTTGGGACA TCAGGGAGT CTCTCCCTCC TAAATATTG	SEQ ID NO: 38
chr 1	144930 977	A	C	<i>PDE4 DIP</i>	p.S244 S	0.007 35	0.004 27	7.25 E- 03	1.73[1.19- 2.5]	CTTTCTGTTG TGGAGGGCT AGCCTGGAC GCTTGCATC CAA[A/C]GAT TCCACAGAG GAACCAGGC GTCTCTTCCT CCATGCTTT	SEQ ID NO: 39
chr 1	145537 513	C	A	<i>ITGA1 0</i>	p.S841 R	0.009 31	0.006 22	2.01 E- 02	1.5[.08- 2.09]	CAACTCTGG AGAACAGAA AGGAAAATG CTTACAATA CGAG[C/A]CT GAGTCTCAT CTTCTCTAG AAACCTCCA CCTGGCCAG TC	SEQ ID NO: 40
chr 1	149897 906	G	A	<i>SF3B4</i>	p.P245 P	0.007 84	0.005 17	2.66 E- 02	1.52[1.06- 2.18]	GGGGTATCC CAGGTGGGA GGGCTCCAG GAGGTGGCA CTGG[G/A]GG TGGGAAGGA GCCAGGAGG AGGCATGCC TATAGAGGA AA	SEQ ID NO: 41

chr 1	152080 275	C	T	<i>TCHH</i>	p.E180 6E	0.010 54	0.000 00	2.67 E- 63	Inf	TTCCGTCAC GCTGTTGGG GGCGCAGCT GCTGTTCTTC CCT[C/T]TCC TGGCGTAGC TGTTCTCCT CGCGGAATT TTCTGTCAG	SEQ ID NO: 42
chr 1	152082 449	T	C	<i>TCHH</i>	p.K10 82E	0.013 24	0.000 46	1.95 E- 48	28.9 5[19. 46- 43.0 5]	CTCAGCAGC TGCTCTTCCT CCTGCTGCA GCTCCTCTTC CT[T/C]CCGA TATTGCCTCT CCAGCTCCT GGCGCCTTC TCGTCTCC	SEQ ID NO: 43
chr 1	152083 327	G	T	<i>TCHH</i>	p.P789 Q	0.010 29	0.000 00	1.16 E- 61	Inf	CTCCTCGGC CCTCAGCTG CCTCTCCCG CTGCTCCCG CAAT[G/T]GG GGCCTGGCC GACAGCCTC TGACGGCCC CTCTCGCTCT T	SEQ ID NO: 44
chr 1	152083 829	G	T	<i>TCHH</i>	p.R622 S	0.019 36	0.000 00	1.65 E- 115	Inf	TTCAGCAGC TGCTGGCGC CTCTCTTCCT CCGGCTCCT CGC[G/T]CTT CAGCCGCTG CTCGCGCCT CTCCTCCTGC TCGAGTCTC	SEQ ID NO: 45
chr 1	152084 213	C	G	<i>TCHH</i>	p.E494 Q	0.014 71	0.000 09	4.56 E- 70	164. 52[7 5.16- 360. 14]	AGTTGCTGC TCGCGCCTC TCCTGCTGCT CGCGCCTCT CCT[C/G]CTC CTCGAGCTT CAGCCAACG TTCGCGCCT CTCCTCCTCC	SEQ ID NO: 46
chr 1	152325 166	G	C	<i>FLG2</i>	p.T169 9R	0.007 11	0.000 01	1.95 E- 41	799. 16[1 08.8 4- 5868 .08]	TAATCCATG ATGATAGTG GGCATGTCT AGTGGTATC TCCT[G/C]TC TGTCATGA GTAGTTCCA TGTCTCTCA GGAACATG GA	SEQ ID NO: 47

chr 1	156011 387	G	A	<i>UBQL</i> <i>N4</i>	p.P514 P	0.005 15	0.003 16	3.43 E- 02	1.63[1.05- 2.54]	CTGTTGGAG AAGATGTGG CTGGCGTGG CTGGTGAGG AAGT[G/A]G GGGCCTCGG GCGTAGACC CTGCGTTGC TGCCTGCTG AGG	SEQ ID NO: 48
chr 1	156046 473	T	C	<i>MEX3</i> <i>A</i>	p.G48 5G	0.005 15	0.002 83	1.14 E- 02	1.82[1.17- 2.84]	CGCAGATGC GTACTGCAC ACTCCATGC AGAACAGGT TGTG[T/C]CC GCAGGGCAC AAGGGCGGC AGTCACTTC GCTCTCAA GC	SEQ ID NO: 49
chr 1	156438 664	C	T	<i>MEF2</i> <i>D</i>	p.Q38 5Q	0.010 05	0.000 01	2.23 E- 58	1107 .97[52.3 7- 8056 .7]	GTTGCGGCT GCTGAGGCT GCTGTGGCT GTGGCTGCT GTGG[C/T]TG CGGTGGCTG CTGCTGTGG AGGCTGTGG CTGCTGCGG CT	SEQ ID NO: 50
chr 1	156521 547	C	T	<i>IQGAP</i> <i>3</i>	p.A56 2T	0.005 88	0.003 68	3.53 E- 02	1.6[.06- 2.42]	TGCCTTTTGG CTGCCACAA GGAGGAGAT GGTACCGAG GGG[C/T]GAC AGGGAGGCT GACATCATC TAGGCCAGC TGCAGGAAG C	SEQ ID NO: 51
chr 1	156779 118	G	A	<i>SH2D2</i> <i>A</i>	p.G29 3G	0.006 37	0.003 76	1.36 E- 02	1.7[.14- 2.53]	CCACATAGA TGTTGCTGG GGGCTTCCC CAGGGCTGC CCCG[G/A]CC CATGGCATA GAAAGCTAT GGGTTTCATC AGGCTCATT GT	SEQ ID NO: 52
chr 1	157069 134	G	A	<i>ETV3L</i>	p.S32L	0.012 25	0.008 99	3.56 E- 02	1.37[1.03- 1.82]	GATGAAGTG CCACAGCTG GATCTGCCG GGAGCCTGG GGAC[G/A]A CTCGGCTTT GTAGGCCCA ATCAGGGAA GGCCAACCC TGG	SEQ ID NO: 53

chr 1	157738 309	G	T	<i>FCRL2</i>	p.L260 M	0.005 21	0.001 94	6.20 E- 04	2.69[1.51- 4.48]	TATTTGCCG GCATCACTC TCTTTCACA GCTGGGATC TCCA[G/T]CT CTGCTGACA GGGAACGCT GGGTTTTCTT TCCCATACT G	SEQ ID NO: 54
chr 1	158669 772	G	C	<i>OR6K2</i>	p.A22 4G	0.005 39	0.000 01	3.16 E- 31	596. 28[8 0.36- 4424 .77]	TGTGCGGCG GCCTCCAGC TGAATGAAT ACGTAGAAT TACA[G/C]CC ACAATACCA TCGTAGGAC ATGAAGATG AGCATCACA GC	SEQ ID NO: 55
chr 1	161336 289	A	G	<i>Clorf1 92</i>	p.Y10 Y	0.005 64	0.003 41	2.89 E- 02	1.66[1.09- 2.53]	GAGACCACT TCTGCAGAT ACTTGATG AGAAAGCCT TTTC[A/G]TA CTGTGGAGA GAAAGATAA GTAGCCCTA TGAGACTTC AA	SEQ ID NO: 56
chr 1	161476 227	C	T	<i>FCGR 2A</i>	p.S69S	0.005 15	0.003 20	4.84 E- 02	1.61[1.04- 2.5]	CTGTGACTC TGACATGCC AGGGGGCTC GCAGCCCTG AGAG[C/T]GA CTCCATTCA GTGGTTCCA CAATGGGAA TCTCATTCCC A	SEQ ID NO: 57
chr 1	161641 237	G	A	<i>FCGR 2B</i>	p.Q63 Q	0.010 78	0.003 40	2.83 E- 10	3.19[2.33- 4.37]	CTGTGCTGA AACTCGAGC CCCAGTGGA TCAACGTGC TCCA[G/A]GA GGACTCTGT GACTCTGAC ATGCCGGGG GACTCACAG CC	SEQ ID NO: 58
chr 1	169697 268	A	G	<i>SELE</i>	p.L404 L	0.005 15	0.003 10	3.16 E- 02	1.67[1.07- 2.59]	TCCCCTGTG GGGCCACAT TGGAGCCTT TTGGATCCC TTCA[A/G]CA CAAAACCCT GCTCACAGG AGAACTCAC AGCTGGACC CA	SEQ ID NO: 59

chr 1	170115 300	G	C	<i>METT L11B</i>	p.D18 H	0.000 74	0.000 74	1.00 E+0 0	1[0.3 1- 3.22]	GGGAGCCCA TTTTGCCTTT AGATCCCGC TGGCAGAAG ACC[G/C]ACG ATGAACTCT GTAGACATA GCATGTCTTT TATCCTTCA	SEQ ID NO: 60
chr 1	170129 701	T	C	<i>METT L11B</i>	p.M66 T	0.008 82	0.006 84	1.44 E- 01	1.29[0.92- 1.82]	AAATTGTAC GCTTTAACA AGCCAAGTC ATCAATGGT GAGA[T/C]GC AGTTCTATG CCAGAGCTA AACTTTTCTA CCAAGAAGT A	SEQ ID NO: 61
chr 1	170136 876	T	C	<i>METT L11B</i>	p.L277 P	0.010 78	0.010 87	1.00 E+0 0	0.99[0.73- 1.35]	GGCTTCCCA GAGCATTGC ATCCCCGTG TGGATGTTC GCAC[T/C]GC ACAGCGACA GACACTCCT GAAAAAGCA GTGGGAATG AA	SEQ ID NO: 62
chr 1	176563 779	G	A	<i>PAPP A2</i>	p.V34 7M	0.008 09	0.005 37	2.96 E- 02	1.51[1.06- 2.15]	GCGGGATGC TCGCTTCTTC TTCTCCCTCT GCACCGACC GC[G/A]TGAA GAAAGCCAC CATCTTGATT AGCCACAGT CGCTACCA	SEQ ID NO: 63
chr 1	176833 427	T	C	<i>ASTNI</i>	p.E129 3G	0.006 62	0.003 85	1.03 E- 02	1.72[1.17- 2.54]	TCATTCTGG CAGCAGCTC CCTGGCCTT ATGGTGCTA GATC[T/C]CT TTGCTGTCCC CATAGTCGT TGTAGGGGA TACTCAGGG T	SEQ ID NO: 64
chr 1	176833 480	C	T	<i>ASTNI</i>	p.T127 5T	0.006 13	0.004 04	4.58 E- 02	1.52[1.02- 2.28]	CATAGTCGT TGTAGGGGA TACTCAGGG TCTGCTCCTC ACA[C/T]GTC TTCCTGAGG TCCCGGCTG AGCTCCGCC CAGTCAAGT C	SEQ ID NO: 65

chr 1	176852 074	T	G	<i>ASTN1</i>	p.M10 95L	0.006 13	0.003 99	4.39 E- 02	1.54[1.03- 2.3]	GAGATGGTG GTGAGCTGC TTGTCCGGC ACCTGAGAT GGCA[T/G]TG CACAAGGAG ACTTTGCTCC AGAGATGAT GTCGTCCAC A	SEQ ID NO: 66
chr 1	186276 268	G	A	<i>PRG4</i>	p.E473 K	0.006 62	0.000 00	3.12 E- 39	Inf	TACACCCAC CACTCCCAA GGAGCCTGC ACCCACCAC CAAG[G/A]A GCCTGCACC CACCCTCC CAAAGAGCC TGCACCCAC TGC	SEQ ID NO: 67
chr 1	198222 215	C	G	<i>NEK7</i>	p.R35 G	0.012 25	0.008 67	2.08 E- 02	1.42[1.07- 1.89]	CTTACGACC GGATATGGG CTATAATAC ATTAGCCAA CTTT[C/G]GA ATAGAAAAG AAAATTGGT CGCGGACAA TTTAGTGAA GT	SEQ ID NO: 68
chr 1	201178 688	A	G	<i>IGFNI</i>	p.E155 6G	0.009 80	0.000 00	6.26 E- 47	Inf	GGGAGTAAG GCAGGTTTT ACGGATGGT TTAGGAGGT TCTG[A/G]AG AAATGGGGT CAGTGAATA AGGCAGGTT ATAGGAAGG AT	SEQ ID NO: 69
chr 1	201180 217	A	G	<i>IGFNI</i>	p.N20 66D	0.008 58	0.000 02	6.77 E- 40	476. 2[65. 22- 3476 .77]	TAGGGATGG TTTAGGGAG TTCTGTAGA AATGGGGTC AGTG[A/G]AT GAGGCAGGT TATAGGAAG GATTTAGGG GCTCCTAAG GG	SEQ ID NO: 70
chr 1	203194 834	C	T	<i>CHIT1</i>	p.E74 K	0.006 62	0.003 80	9.72 E- 03	1.74[1.18- 2.58]	CACATCTTCT TCAGGCCAT TGAACCTCT GGTAGAGAG TCT[C/T]GTC ATTCCACTC AGTGGTGCT CAGCTGGTG GTTGGTCAT G	SEQ ID NO: 71

chr 1	203691 612	A	G	<i>ATP2B</i> 4	p.K94 0R	0.005 15	0.002 55	4.39 E- 03	2.02[1.3- 3.15]	ACTTAACCT CCAGTGCTT CTCCTCTCCC CACTAGGTG AGA[A/G]ATT CTTTGATATT GATAGTGGG AGGAAGGCA CCTCTACAT	SEQ ID NO: 72
chr 1	204923 359	G	A	<i>NFAS</i> C	p.D81 N	0.005 64	0.000 00	3.59 E- 34	Inf	CCACTGGAC ACGAAACAG CAGATTCTT CAACATCGC CAAG[G/A]A CCCCCGGGT GTCCATGAG GAGGAGGTC TGGGACCCT GGT	SEQ ID NO: 73
chr 1	204923 461	C	T	<i>NFAS</i> C	p.R115 C	0.005 39	0.000 00	1.05 E- 32	Inf	GCGGCCGGA GGAATATGA GGGGGAATA TCAGTGCTT CGCC[C/T]GC AACAAATTT GGCACGGCC CTGTCCAAT AGGATCCGC CT	SEQ ID NO: 74
chr 1	206658 569	G	A	<i>IKBKE</i>	p.T514 T	0.010 05	0.006 84	2.08 E- 02	1.47[1.07- 2.02]	AGCTAGCGG AGGTCCTCT CCAGATGCT CCCAAATA TCAC[G/A]GA GACCCAGGA GAGCCTGAG CAGCCTGAA CCGGGAGCT GG	SEQ ID NO: 75
chr 1	222712 108	G	T	<i>HHIPL</i> 2	p.L487 M	0.010 78	0.006 33	1.35 E- 03	1.71[1.26- 2.32]	ACTGACTTC CCCCTGCA TGGCCATAA GCATAGATT GGCA[G/T]AA CATCATCTG TCCAGGAGA GAGGAAAGA GAGTGAGTG TC	SEQ ID NO: 76
chr 1	227843 024	T	A	<i>ZNF67</i> 8	p.F413 Y	0.009 56	0.000 01	1.18 E- 55	1063 [146. 01- 7739 .02]	GGAGAGAAA CCCTACAAA TGTGAAGAA TGTGGCAGA ACCT[T/A]TA CTCAATTCTC AAACCTCAC TCAGCATAA AAGAATTCA T	SEQ ID NO: 77

chr 1	231057 248	C	T	<i>TTC13</i>	p.G55 3D	0.012 99	0.000 00	5.31 E- 75	Inf	TTTCTCAAA ATATTCTAG GTATCTCAT GTTGATCAC CTGA[C/T]CC CTATAAGGC AAAAATAAT AAAATTAAG AATATTTT T	SEQ ID NO: 78
chr 1	236144 919	G	T	<i>NID1</i>	p.S107 3S	0.005 15	0.002 68	8.73 E- 03	1.93[1.24- 3]	AGAGATGCA CACACATAT TTACACAAA GATACCCTC TCAC[G/T]GA ATCCGTTAC AATGCCTCT GGGATTCAC CAAGTCAGT CT	SEQ ID NO: 79
chr 1	236433 208	T	G	<i>ERO1L B</i>	p.K63 N	0.005 88	0.003 65	3.47 E- 02	1.62[1.07- 2.44]	ACCTTACCTT GTAATAACG AAAATAGTC TCTCTCTTGC AA[T/G]TTTT TTATTTTGGG GAAGATTTT GTAGGTATT GAAGTTAT	SEQ ID NO: 80
chr 1	246907 410	A	G	<i>SCCP DH</i>	p.I183 V	0.005 21	0.002 73	2.17 E- 02	1.91[1.08- 3.17]	TCTTTTAGGT ACTTTGACT GCTGTGGAA AGTTTCCTG ACT[A/G]TAC ATTCAGGAC CTGAGGTTG GTTTTTTGGT TTGTCTTGT	SEQ ID NO: 81
chr 1	248436 265	G	A	<i>OR2T3 3</i>	p.N28 4N	0.008 33	0.002 45	5.54 E- 09	3.43[2.4- 4.9]	CTCCCTTCAC CTCACTGTTC TTCACACTG TAGATGAGG GG[G/A]TTTA GTAAAGGGG TGAACATAG TATAGAAGG CTGACACAA	SEQ ID NO: 82
chr 10	592504 7	G	A	<i>ANKR D16</i>	p.F257 F	0.005 39	0.003 25	2.59 E- 02	1.66[1.08- 2.56]	TGGCTCTCA CATCTACAT CGACGCCAA GTTTCAGAGA CCAA[G/A]A ATCGGATGG CTTCGTCCTG CCCTGTGAC AGCTGCCCT GT	SEQ ID NO: 83

chr 10	597922 2	C	G	<i>FBXO 18</i>	p.A96 3A	0.005 64	0.003 16	1.08 E- 02	1.79[1.17- 2.73]	AGCGCACTG TGGAGAACA TCGTACTGC CCCGGCATG AGGC[C/G]CT GCTCTTCCTC GTCTTCTGA GGACAAGGC GCACGTTCT C	SEQ ID NO: 84
chr 10	777195 8	C	T	<i>ITIH2</i>	p.N44 1N	0.006 37	0.003 96	2.30 E- 02	1.61[1.09- 2.4]	AACTAAAAC TGTCAAAAA TTCAGAAAA ACGTTAAGG AGAA[C/T]AT CCAAGACAA TATCTCCTTG TTCAGTTTG GGCATGGGA T	SEQ ID NO: 85
chr 10	210975 46	G	A	<i>NEBL</i>	p.S885 F	0.006 37	0.004 13	3.59 E- 02	1.55[1.04- 2.3]	TGACCTGTC GTCTCCGAG ACCTGTACC GAAAGTACT GCTG[G/A]AA TGGGATCGA GACCAGTGT CGCCTATAG TGA CTGCC TT	SEQ ID NO: 86
chr 10	345587 15	C	T	<i>PARD 3</i>	p.G10 17R	0.005 15	0.002 97	1.96 E- 02	1.74[1.12- 2.7]	CTAGCGTTG AGAGCCATG GAACCTTCA TAAGAAGAA ACTC[C/T]CC CATACATTA ACTCATCAT CACAGCCAA ATGTCCGAT GA	SEQ ID NO: 87
chr 10	353221 99	C	T	<i>CUL2</i>	p.M34 8I	0.009 778	0.004 78	5.26 E- 04	2.06[1.37- 2.98]	TACCATGCA CTTCCAAAA CTGACTCCA CAAATAGTG TTGG[C/T]AT CTAAAAATG AAATATAAG TACAAAACC ACATTTTAA GA	SEQ ID NO: 88
chr 10	454730 44	C	G	<i>C10orf 10</i>	p.M14 5I	0.019 36	0.000 01	7.99 E- 114	2197 .65[3 05.6 9- 1579 9.34]	CAGGCATCC TGGCTTCAC AGAGCCTCC CTCTGGGGG CCCC[C/G]AT GGGCTTGCT GCTGTCCAT CTGTCTATGT GGACCCCA A	SEQ ID NO: 89

chr 10	469992 09	G	A	<i>GPRIN</i> 2	p.R110 Q	0.007 84	0.005 38	3.98 E- 02	1.46[1.02- 2.09]	AATGTGTCC ACCATGGGC GGCAGTGAC CTGTGTGCG CTGC[G/A]GG CCCCTAGTG CTGCTGCTA TGCAGAGGA GCCATTTCAG AC	SEQ ID NO: 90
chr 10	469993 90	A	G	<i>GPRIN</i> 2	p.A17 0A	0.010 29	0.003 47	4.09 E- 09	2.99[2.17- 4.12]	AGCCAGGTG GTACTTCTG GCCAGGGTG GCCAGGCCC CTGC[A/G]GG CCTGGAAAG GGACCTGGC TCCTGAGGA TGAGACTTC TA	SEQ ID NO: 91
chr 10	470872 99	G	C	<i>LOC10</i> <i>099675</i> 8	p.L172 L	0.006 62	0.003 53	3.19 E- 03	1.88[1.27- 2.78]	GGATTGTGC TCATCTGGG TCATTGCCT GTGTCCTCTC CCT[G/C]CCC TTCCTGGCC AACAGCATC CTGGAGAAT GTCTTCCAC A	SEQ ID NO: 92
chr 10	518279 00	A	G	<i>FAM2</i> <i>1A</i>	p.P13P	0.022 30	0.002 48	5.96 E- 49	9.18[7.2- 11.7]	TGCAGATGA ACCGGACGA CCCCGACC AGGAGCTGG CGCC[A/G]GC GTCGGAGCC CGTGTGGGA GCGGCCGTG GTCGGTGGA GG	SEQ ID NO: 93
chr 10	734648 12	G	A	<i>CDH2</i> 3	p.E960 K	0.008 133	0.004 201	2.90 E- 03	1.94[1.24- 2.91]	GGTGGTCAC CACCACCGA GCTGGACCG CGAGCGCAT CGCG[G/A]A GTACCAGCT GCGGGTGGT GGCCAGTGA TGCAGGCAC GCC	SEQ ID NO: 94
chr 10	750106 35	G	C	<i>MRPS</i> 16	p.T130 R	0.008 458	0.004 873	8.85 E- 03	1.74[1.13- 2.59]	CTAAAGTCA GCTCATTTAT GTTTCTGTA GCCTCTGTA TCT[G/C]TAG CTTCTGCATC TGTTTTCTGA GAAGCTAAC AGGACTTC	SEQ ID NO: 95

chr 10	795887 06	G	A	<i>DLG5</i>	p.A74 1A	0.007 35	0.004 35	8.01 E- 03	1.69[1.17- 2.45]	GGGACCCCTT CTTTAGCGG CAGGGCTTC CAGGCAGCA CAGC[G/A]GC AGCATACAC TCCATTCTCC AGACTGATG CCACTGTCT G	SEQ ID NO: 96
chr 10	995312 84	C	T	<i>SFRP5</i>	p.D10 3N	0.010 05	0.006 89	2.15 E- 02	1.46[1.07- 2.01]	CAGACGGGC GCAAAGAGC GAGCACAGG AAGACCTGC GTAT[C/T]CG AGTGGCAGC GCTTGGCCA GCAGCGGCA GCCAGCTGC TC	SEQ ID NO: 97
chr 10	999696 50	A	G	<i>R3HC CIL</i>	p.L593 L	0.006 86	0.003 60	2.22 E- 03	1.91[1.3- 2.81]	TGTTTAACG ATGATGGTG ACTGCCTGG ATCCACGTC TTCT[A/G]CA AGAGGTATG TTTAATTGA AATTGCTTG ATGCTTAGT TA	SEQ ID NO: 98
chr 10	102770 315	A	G	<i>PDZD 7</i>	p.R777 R	0.011 03	0.000 09	2.35 E- 44	126. 17[4 5.36- 350. 99]	ACTTGCCCTT GACCCCGGC TGCTGCGGC TGCGGCTGC GGCT[A/G]CG GCTGCGGCT ACGGCTCTG AGCCCGGCC CCGGATCTG GC	SEQ ID NO: 99
chr 10	104230 587	G	A	<i>TMEM 180</i>	p.T139 T	0.010 54	0.007 62	4.44 E- 02	1.39[1.02- 1.89]	AGTTCTTGCT GTGCCTGTG CCTCTATGA TGGCTTCCT GAC[G/A]CTC GTGGACCTG CACCACCAT GCCTTGCTG GCCGACCTG G	SEQ ID NO: 100
chr 10	125780 760	G	C	<i>CHST1 5</i>	p.P453 P	0.008 58	0.000 01	3.19 E- 47	793. 53[1 08.6 9- 5793 .56]	GCTCCTTCTG CCAGGGGCC AGCTCGGGG GGTACGGGG GGG[G/C]GG GGTACACAC AGGCATGGC GTTGTTGAG GGTGTGTT GT	SEQ ID NO: 101

chr 10	135106 137	G	A	<i>TUBG CP2</i>	p.H36 0H	0.005 39	0.003 26	2.61 E- 02	1.66[1.08- 2.55]	CCTGCGCCT GGCTGTCCC CTGTGTAGC TGAAGCTCC TGTC[G/A]TG GAGCAGGCT CAGCGTGGA CCCCCAAG ACATTCGCC TT	SEQ ID NO: 102
chr 10	135368 906	G	C	<i>SYCE1</i>	p.V28 9V	0.008 09	0.005 37	2.96 E- 02	1.51[1.06- 2.15]	GGCCAGCCT CTTCCTCTTG TGTCCTCTG GGCTTGGGC AGG[G/C]ACT TGCATTCCA TGCTTTTCCA GCTCTTCCTT CAGCCTGG	SEQ ID NO: 103
chr 11	394511	C	T	<i>PKP3</i>	p.A73 A	0.006 86	0.000 00	6.27 E- 11	Inf	AGCCGCGGC ACAACGGGG CCGCTGAGC CCGAGCCTG AGGC[C/T]GA GACTGCCAG AGGTAGGCG GTGGGGACA GCGGCGGGG AT	SEQ ID NO: 104
chr 11	610300	A	G	<i>PHRF 1</i>	p.S145 5G	0.006 86	0.003 57	2.18 E- 03	1.93[1.3- 2.85]	CACAGGGGT CAGGCAGGT GTTCTCCGA GCTGCCCTTT CCC[A/G]GTC ACGTGCTTC CGGAACCCG GGTTCCCAG ACACAGACC C	SEQ ID NO: 105
chr 11	614967	C	G	<i>IRF7</i>	p.R88 T	0.005 88	0.000 00	4.94 E- 32	Inf	GCGCTCCGC AGTCTCAGC CTCGGGGGG CGGGCCACC TCCC[C/G]TG CTGCTAGGC GGCCACCTG CCGCGGGCC ACAGCCCAG GC	SEQ ID NO: 106
chr 11	764414	A	G	<i>TALD O1</i>	p.K32 1R	0.006 13	0.003 59	1.66 E- 02	1.71[1.14- 2.56]	CTCTCTGAC GGGATCCGC AAGTTTGCC GCTGATGCA GTGA[A/G]GC TGGAGCGGA TGCTGACAG TGAGTGTTG TGTGTGGGT AC	SEQ ID NO: 107

chr 11	101685 4	G	A	<i>MUC6</i>	p.P198 3S	0.011 27	0.000 00	1.29 E- 67	Inf	GGATAGGTA GTGGTGGTC TGGAAGGAT GTTGCAGTC ATAG[G/A]AC CTGTGGAAG AGAAGGGAC TGCTCCCTGT AGGTGGGGA G	SEQ ID NO: 108
chr 11	101708 5	G	A	<i>MUC6</i>	p.P190 6S	0.007 84	0.001 74	3.28 E- 11	4.53[3.11- 6.59]	GGTAGGGAT GTAGAAGTT TTGGCCGTG CTAAATGAG CTTG[G/A]GG ATTGGCTGG TCCCACTGG TGGTCGGTG TCATTGGTG GG	SEQ ID NO: 109
chr 11	101754 3	G	A	<i>MUC6</i>	p.T175 3I	0.025 25	0.000 00	8.09 E- 151	Inf	GGTAGAAGT TGAGGTGAC TTCAGGATG GTGTGTGGA GGAA[G/A]T GTGTGAATG TAGGGATGT AGAGGTTTT GGCCGTGCT AAA	SEQ ID NO: 110
chr 11	101776 1	T	C	<i>MUC6</i>	p.Q16 80Q	0.009 80	0.000 05	1.12 E- 51	180. 29[7 6.39- 425. 47]	GGGATGTAG AGGTTTTGG CTGTGTTTA ATGAGCTCA GGGC[T/C]TG GCTGGTCCC GCTGGTGGT CAGCGTCAT TGTTGGCGC TG	SEQ ID NO: 111
chr 11	101778 5	C	T	<i>MUC6</i>	p.T167 2T	0.009 80	0.000 36	1.86 E- 36	27.7 8[17. 86- 43.2 4]	TTAATGAGC TCAGGGCTT GGCTGGTCC CGCTGGTGG TCAG[C/T]GT CATTGTTGG CGCTGTGTG GGTGGACCC TGTGGCCTT GA	SEQ ID NO: 112
chr 11	101791 2	G	A	<i>MUC6</i>	p.T163 0I	0.014 95	0.000 29	6.50 E- 49	51.6 5[26. 44- 100. 88]	GGCAGAAGT GGCCATCTG TGCATGGGT AGGGGTGAT GACT[G/A]TG TGAGTACTT GGAGTCACC AAAGAGGTG GAGAAAGGT GG	SEQ ID NO: 113

chr 11	101797 4	C	G	<i>MUC6</i>	p.Q16 09H	0.007 60	0.000 49	2.56 E- 23	15.7 2[10. 08- 24.5 1]	AAGAGGTGG AGAAAGGTG GAACGTGAG TGGGAAGTG TGGT[C/G]TG AGGGTGTGA TGGGGTTGG ATAGGTAGT GGTGGTCTT GA	SEQ ID NO: 114
chr 11	102362 2	G	A	<i>MUC6</i>	p.T113 8M	0.009 80	0.007 03	4.58 E- 02	1.4[1 .02- 1.92]	GGCCTCCTG TGTGTACTG GTACTCGCC ATGGCCGTC CTGC[G/A]TG TGC GTGTTG TAGAAGCCG CAGTAGATG GCTGGGAGG AA	SEQ ID NO: 115
chr 11	109353 7	A	C	<i>MUC2</i>	p.K17 86Q	0.007 11	0.000 08	4.33 E- 27	94.8 1[28. 87- 311. 37]	CACCACTAC GATGACCCC AACCCCAAC ACCCACCAG CACA[A/C]AG AGTACAACC GTGACACCC ATCACCACC ACA ACTACG GT	SEQ ID NO: 116
chr 11	126418 7	C	T	<i>MUC5 B</i>	p.T202 6M	0.006 62	0.003 93	1.15 E- 02	1.69[1.14- 2.49]	ACTCCAGAG ACTGCCCAC ACCTCCACA GTGCTTACC GCCA[C/T]GG CCACCACAA CTGGGGCCA CCGGCTCTG TGGCCACCC CC	SEQ ID NO: 117
chr 11	126996 9	G	A	<i>MUC5 B</i>	p.T395 3T	0.006 86	0.004 53	4.28 E- 02	1.52[1.03- 2.24]	CCAGTGGTA CTCCCCCAT CACTGATCA CCACGGCCA CTAC[G/A]AT CACGGCCAC CGGCTCCAC CACCAACCC CTCCTCAAC TC	SEQ ID NO: 118
chr 11	127131 3	A	G	<i>MUC5 B</i>	p.T440 1T	0.014 95	0.000 00	2.18 E- 89	Inf	CGACCTGGA TCCTCACAG AGCTGACCA CAGCAGCCA CTAC[A/G]AC TGCAGCCAC TGGCCCCAC GGCCACCCC GTCCTCCAC CC	SEQ ID NO: 119

chr 11	160615 0	G	A	<i>KRTA</i> <i>P5-1</i>	p.G11 0G	0.005 39	0.000 00	6.34 E- 31	Inf	CACAGCCGG AACCACAGC CACCCCTTGG ATCCCCCAC AAGA[G/A]C CACAGCCCC CCTTGGAGC CCCCACAGG AGCCACAAC CCC	SEQ ID NO: 120
chr 11	160640 2	G	A	<i>KRTA</i> <i>P5-1</i>	p.S26S	0.004 64	0.000 10	2.01 E- 11	42.7 7[16. 27- 112. 48]	AGCCAGAAC CTCCACAGC CAGAGCCAC AGCCCCCAC AGCC[G/A]G AGCCACAGC CCCCACAGC CGGAGCCAC AGCCCCCAC AGC	SEQ ID NO: 121
chr 11	161943 0	A	G	<i>KRTA</i> <i>P5-2</i>	p.C17 C	0.012 25	0.000 01	1.27 E- 71	1373 .66[1 89.7 1- 9946 .24]	AGCCCCCAC AGCCAGAGC CACAACCCC CACAGCTGG AGCC[A/G]CA GCCCCCACA GCCGGAGCC ACAGCCTCT GGAGCAGCC AC	SEQ ID NO: 122
chr 11	162916 3	G	A	<i>KRTA</i> <i>P5-3</i>	p.C151 C	0.010 29	0.000 01	5.33 E- 58	1023 .61[1 40.8 5- 7439 .08]	AGCAGGGCT TACAGCAGC TGGACTGGG AGCAGCTGG GCTT[G/A]CA GCAGCTGGA CTGGCAGCA GGATGACCC ACAGCCTGA GG	SEQ ID NO: 123
chr 11	162936 4	C	A	<i>KRTA</i> <i>P5-3</i>	p.K84 N	0.013 48	0.000 00	1.22 E- 80	Inf	AGCAGCAGA CGGGCACAC AGCAGCTGG AGCCACAGC CCCC[C/A]TT GGAGCCTCC ACAGGAGCC ACAGCCCCC CTTGCAGCC CC	SEQ ID NO: 124
chr 11	164288 0	A	G	<i>KRTA</i> <i>P5-4</i>	p.S148 S	0.011 27	0.000 00	1.28 E- 67	Inf	TACAGCAGC TGGACTGGC AGCAGGATG ACCCACAGC CTGA[A/G]GA GAAGCAGCA GGGCTTACA GCAGCTGCA CTGGGAGCA GC	SEQ ID NO: 125

chr 11	165135 9	A	G	<i>KRTA</i> <i>P5-5</i>	p.R97 G	0.027 94	0.000 00	1.04 E- 166	Inf	CTGTGGCAA AGGGGGCTG TGGCTCTTG CGGGGGCTC CAAG[A/G]G AGGCTGTGT CTCCTGTGG GGTGTCCAA GGGGGCCTG TGG	SEQ ID NO: 126
chr 11	216143 0	G	A	<i>IGF2</i>	p.Q33 X	0.016 68	0.000 90	9.89 E- 16	19.8[11.5- 34.2]	CGTCTAAGT AGCTCGCCT TTGCGGCCC ACCCAAAAT ATCT[G/A]GA TAATGGTTA CCCCGTCCT CAGTGC GTT GGA CTTGCA TA	SEQ ID NO: 127
chr 11	438911 0	G	A	<i>OR52B</i> <i>4</i>	p.T139 I	0.005 21	0.002 91	2.82 E- 02	1.79[1.01- 2.96]	CAGAGAGAC AGTCACACA AATTTTCTTG ATCAGAGCA TTT[G/A]TAA GAATGGTGG TGTACCTCA GTGGGTAGC ATATGGCAA T	SEQ ID NO: 128
chr 11	544404 0	C	T	<i>OR51</i> <i>Q1</i>	p.L204 F	0.008 58	0.005 50	1.36 E- 02	1.57[1.11- 2.2]	CTGTGCTGA CATCAGGCT CAACAGCTG GTATGGATT TGCT[C/T]TT GCCTTGCTC ATTATTATC GTGGATCCT CTGCTCATT GT	SEQ ID NO: 129
chr 11	691328 1	T	C	<i>OR2D</i> <i>2</i>	p.S151 G	0.008 133	0.004 873	1.80 E- 02	1.67[1.07- 2.5]	AGTATGAAG GTGGTGTCT ACCACAGAC ACCAGAATG CCAC[T/C]GG TCCATGATC CTGTTGCCA GCTGGACAC ACACTTTCC AG	SEQ ID NO: 130
chr 11	694291 5	C	T	<i>OR2D</i> <i>3</i>	p.S228 F	0.014 71	0.010 03	5.32 E- 03	1.47[1.13- 1.92]	ATCTTTTCAA TGGGCGTGG TAATCCTCCT GGCCCCTGT CT[C/T]CCTG ATTCTTGGTT CTTATTGGA ATATTATCTC CACTGTT	SEQ ID NO: 131

chr 11	122463 55	G	A	<i>MICAL</i> 2	p.R559 Q	0.008 33	0.005 56	2.51 E- 02	1.51[.06- 2.13]	CGCAGTGGG TTGGCCCTG TGTGCCATC ATCCACCGC TTCC[G/A]GC CTGAGCTCA TGTGAGTCT GGGGCCCAG GCTGGCCCC TG	SEQ ID NO: 132
chr 11	341650 53	G	A	<i>NAT10</i>	p.A98 3T	0.008 133	0.003 762	5.80 E- 04	2.17[1.39- 3.26]	TGAAGAGTG GAATGAAGT TTTGAACAA AGCTGGGCC GAAC[G/A]CC TCGATCATC AGCCTGAAA AGGTGAGGG CCCAGGGTC TG	SEQ ID NO: 133
chr 11	354560 85	T	A	<i>PAMR</i> 1	p.D53 4V	0.007 60	0.005 11	3.47 E- 02	1.49[1.04- 2.14]	CAAGCCCTC TCTTACCTGT AGGCTCTGG ATGGTCTTCT CA[T/A]CCCCG GTCATCATC CCGGTAGAA TTTCCCCAA AACAACTTT	SEQ ID NO: 134
chr 11	474696 31	G	T	<i>RAPSN</i>	p.N88 K	0.005 15	0.002 63	5.29 E- 03	1.96[1.26- 3.06]	TCTTGTGAA ACTCGCACA GCTTCTCGTT GCTGCGTGC CAG[G/T]TTC AGGTAGCTC TCCAGGAGG AAGTCGGCA TCCTCCAGC T	SEQ ID NO: 135
chr 11	619595 31	A	C	<i>SCGB1</i> D1	p.N20 T	0.005 15	0.002 33	1.60 E- 03	2.21[1.42- 3.46]	TCCTTACAC AAATTATAT TTTTATCTT TTGCTCCAG CAA[A/C]TGC AGTGGTCTG CCAAGCTCT TGGTTCTGA AATCACAGG C	SEQ ID NO: 136
chr 11	622880 14	G	A	<i>AHNA</i> K	p.P462 5P	0.007 60	0.004 94	2.38 E- 02	1.54[1.07- 2.22]	GGACATCAA TGTCCACTTT GGGGTCCCT GATGTCAAC TTC[G/A]GGG CCCTTGAGG TCGCCTTCC ACTTTGGGC AGAGAAATG T	SEQ ID NO: 137

chr 11	624339 12	C	T	<i>METT</i> <i>L12</i>	p.R38 W	0.005 21	0.002 72	2.15 E- 02	1.92[1.08- 3.18]	ACTGGCTGA TAGTTGCCT GGCGGACCG CTGTCTCTG GGAT[C/T]GG CTGCATGCC CAGCCTCGT TTGGGCACT GTCCCCACC TT	SEQ ID NO: 138
chr 11	624443 84	C	G	<i>UBXN</i> <i>1</i>	p.E249 Q	0.012 25	0.008 88	3.42 E- 02	1.38[1.04- 1.84]	CTGAGCAAT TGCACAGGG TCCTGGCCC CCACCTAGT TCCT[C/G]CC CACGGTGGA GCTCCACAT AGAGCCTCA CAGCTGCCA GC	SEQ ID NO: 139
chr 11	627608 00	C	T	<i>SLC22</i> <i>A8</i>	p.R422 Q	0.005 88	0.003 38	1.35 E- 02	1.75[1.15- 2.64]	GGCCTTTTCC ACCTCTGGC TCCTGCTTTG GCTTCTTTGC C[C/T]GCAGG GACCTAGGG ACAGAGAGC TAAGGAAAA GCCCTGGG	SEQ ID NO: 140
chr 11	634874 75	G	C	<i>RTN3</i>	p.D50 1H	0.010 54	0.007 68	4.56 E- 02	1.38[1.01- 1.87]	ATTGGGAGA AATCACAGA AGCTGATAG TTCTGGTGA GTCT[G/C]AT GACACAGTA ATAGAGGAC ATCACAGCA GATACATCA TT	SEQ ID NO: 141
chr 11	636815 04	C	T	<i>RCOR</i> <i>2</i>	p.T271 T	0.009 31	0.004 07	1.13 E- 05	2.3[.65- 3.2]	GGAGCGTGA GGTTGGCAA GGTCCGGGC TTCCTGACA CTGC[C/T]GT GAGGCCTTC AGGGCTCAG GTACATGCC CTTGGGTGG GC	SEQ ID NO: 142
chr 11	640518 89	G	A	<i>GPR13</i> <i>7</i>	p.G17 G	0.006 04	0.000 10	6.66 E- 15	48.8 3[20. 22- 117. 93]	CTGTGAGGA CAAGATGTT ACGTAGTCA AGGCACAGC TGGG[G/A]CC AACGGTGGC CCTGGAAGG CAGAGGCAG GTACCCCTG GC	SEQ ID NO: 143

chr 11	640832 93	G	T	<i>ESRR4</i>	p.R376 L	0.018 87	0.000 67	4.17 E- 69	28.9[20.8 7- 40]	GAAGCCGGC CGGGCTGGC CCCGGAGGG GGTGCTGAG CGGC[G/T]GC GGGCGGGCA GGCTGCTGC TCACGCTAC CGCTCCTCC GC	SEQ ID NO: 144
chr 11	640833 00	G	A	<i>ESRR4</i>	p.A37 8A	0.016 91	0.000 63	5.66 E- 61	27.1 7[19. 38- 38.0 8]	GCCGGGCTG GCCCCGGAG GGGGTGCTG AGCGGCGGC GGGC[G/A]G GCAGGCTGC TGCTCACGC TACCGCTCC TCCGCCAGA CAG	SEQ ID NO: 145
chr 11	649850 72	G	A	<i>SLC22 A20</i>	p.A18 4A	0.005 15	0.003 20	4.82 E- 02	1.61[1.04- 2.51]	GGTCCTACC TGCAGCTGG CAGCTTCGG GGGCCGCCA CAGC[G/A]TA TTTCAGCTCC TTCAGTGCC TATTGCGTCT TCCGGTTCC	SEQ ID NO: 146
chr 11	724060 46	C	T	<i>ARAP1</i>	p.V12 25I	0.005 15	0.002 47	2.50 E- 03	2.09[1.34- 3.25]	CAAGCCCAG CGTCACCCA CCTGCCTCCT CCCTCTCGTT GA[C/T]CTCA AAGCAGGTC CAATAGTCC TTCTCCCTGA TGCCCACG	SEQ ID NO: 147
chr 11	738439 93	C	T	<i>C2CD 3</i>	p.R371 R	0.009 31	0.006 62	4.99 E- 02	1.41[1.02- 1.96]	CAGTTGAAG GGAGGAGGT GATCTTCAA TGTGGTCTTT AAA[C/T]CGA TTCCTAGAA AAGGCTCTG ATCCTAAGG TGTGGA AAA A	SEQ ID NO: 148
chr 11	740535 73	G	A	<i>PGM2 L1</i>	p.T522 I	0.005 15	0.002 51	4.03 E- 03	2.06[1.32- 3.21]	ATATCCAGT GGTAACGTC CCGTACATG CAATATAGC AAAT[G/A]TT CCACAAAAT TTTGGATATT CTTTTGGAG AATCAAAAT T	SEQ ID NO: 149

chr 11	747175 37	A	T	NEU3	p.X46 2Y	0.006 62	0.004 13	2.55 E- 02	1.61[1.09- 2.37]	CCAGCCCTG GTAGGAACC CAAGCCAAT TCAAAAGCA ATTA[A/T]TT GGCTTAGGA CCCAATTTC CATAGATGC AAATGGCAG TT	SEQ ID NO: 150
chr 11	755093 32	C	T	DGAT 2	p.F247 F	0.012 25	0.000 00	1.70 E- 73	Inf	ACTCCTTTG GAGAGAATG AAGTGTACA AGCAGGTGA TCTT[C/T]GA GGAGGGCTC CTGGGGCCG ATGGGTCCA GAAGAAGTT CC	SEQ ID NO: 151
chr 11	755093 41	C	T	DGAT 2	p.G25 0G	0.018 14	0.000 00	1.61 E- 108	Inf	GAGAGAATG AAGTGTACA AGCAGGTGA TCTTCGAGG AGGG[C/T]TC CTGGGGCCG ATGGGTCCA GAAGAAGTT CCAGAAATA CA	SEQ ID NO: 152
chr 11	768348 87	C	A	CAPN 5	p.L632 I	0.007 60	0.004 56	9.41 E- 03	1.67[1.16- 2.41]	GCAGCCCAG CAACCTGCC AGGCACTGT GGCCGTGCA CATT[C/A]TC AGCAGCACC TCCCTCATG GCTGTCTGA CACCTGCCC AC	SEQ ID NO: 153
chr 11	828797 61	C	T	PCF11	p.P795 L	0.007 84	0.005 34	3.85 E- 02	1.47[1.03- 2.11]	GGACCTCCC ACACCAGCT TCTCTTCGGT TTGATGGGT CAC[C/T]AGG ACAAATGGG GGGAGGAGG CCCTTTGAG ATTTGAGGG G	SEQ ID NO: 154
chr 11	896073 39	C	T	TRIM6 4B	p.E205 K	0.008 82	0.003 16	3.95 E- 07	2.81[1.96- 4.02]	ATTCTCACTT GACTGTCTT GTAGTTGTT GGAAAAGCT CTT[C/T]TGC TTCTCTTCC AGTGCCTGC AGATGCCGT TGCTCCTCC	SEQ ID NO: 155

chr 11	947598 63	G	A	<i>KDM4 E</i>	p.C381 Y	0.008 09	0.003 74	1.56 E- 04	2.17[1.5- 3.14]	GCTCTGGGC CTGAGGCTT CTCCCAAAC CTCACAGCC CAGT[G/A]TC CCACACAGC CTGTGTCCTC AGGGCACTG TTACAACCC A	SEQ ID NO: 156
chr 11	961175 37	A	C	<i>CCDC 82</i>	p.D12 5E	0.006 62	0.002 24	2.43 E- 06	2.96[1.99- 4.42]	TGTTGAGAT CATTATCCTC TTGACTTAA ATGTTTTTCC TG[A/C]TCTT GTAAGTCAA TATTCTATG TTTGATTTTG TTCGTTT	SEQ ID NO: 157
chr 11	107381 630	G	T	<i>ALKB H8</i>	p.H47 4N	0.006 62	0.004 08	2.37 E- 02	1.63[1.09- 2.43]	AGAGAAAGA AAGACCATA CTTACTGCT GTTGCAAAA TGAT[G/T]AA TAACAGCAA TGGAGATGC AGGCATCAC AAGACCCAC TG	SEQ ID NO: 158
chr 11	114451 010	T	C	<i>NXPE4</i>	p.I31V	0.005 39	0.003 33	3.81 E- 02	1.62[1.05- 2.5]	ACAGGATTC CATGTGTTTC TCCAGACAT GCCCACTGG GGA[T/C]TGT GGATGTCAT TCCAAACTT GCATTTCTCT TTCATTGCA	SEQ ID NO: 159
chr 11	116744 648	A	G	<i>SIK3</i>	p.L518 L	0.005 39	0.003 38	3.98 E- 02	1.6[1.04- 2.46]	TGTC TAGGT ACCTTG TAC TCAAGTTGC CCGTTTGGT TGCA[A/G]GT TTTGCATAG GCAACAGGT TGTGCATGA AGTTCACAT TA	SEQ ID NO: 160
chr 11	117054 496	G	A	<i>SIDT2</i>	p.R235 H	0.005 39	0.003 13	2.25 E- 02	1.73[1.12- 2.66]	ATGATGATG AAGAAGATA TTTATCATCA TCATCCTGC AGC[G/A]CA AAGACTTCC CCAGCAACA GCTTTTATGT GGTGGTGGT G	SEQ ID NO: 161

chr 11	117057 334	C	T	<i>SIDT2</i>	p.R333 X	0.005 64	0.000 01	3.63 E- 31	533. 81[7 2.07- 3953 .67]	ATGCAGGCA GAAGAAGAA GACCCTGCT GGTGGCCAT TGAC[C/T]GA GCCTGCCCCA GAAAGCGGT ACCTCCAGG GGGCCTGGG TG	SEQ ID NO: 162
chr 11	118516 274	G	A	<i>PHLD B1</i>	p.A11 08T	0.005 39	0.002 19	2.98 E- 04	2.47[1.59- 3.82]	CCTGCCTGC GGGGCGGGA GCGTGCGGA GGAGGGTGA GCAC[G/A]CC TATGATACG CTGAGTCTG GAGAGCTCT GACAGCATG GA	SEQ ID NO: 163
chr 11	118850 225	C	G	<i>FOXR I</i>	p.A15 3G	0.005 15	0.000 02	7.54 E- 29	287. 7[67. 44- 1227 .4]	GACAGCTCC TCTATGGCT CTCCCATCC CCTCACAAA AGGG[C/G]CC CCCTCCAGA GTCGGAGGC TTCGGCAAG CCAGCAGCC AG	SEQ ID NO: 164
chr 11	120188 060	T	A	<i>POU2 F3</i>	p.F422 I	0.018 87	0.000 01	5.03 E- 111	2148 .21[2 98.7 1- 1544 8.88]	TCAAAATAA CTCCAAAGC AGCAGTGAA CTCCGCCTC CAGT[T/A]TT AACTCTTCA GGGTAAGGT GAAGGGGAC GGTGCAGAG AC	SEQ ID NO: 165
chr 11	123476 177	C	T	<i>GRAM D1B</i>	p.A29 5A	0.006 37	0.003 64	8.40 E- 03	1.76[1.18- 2.62]	TCACCAACA GCACACTAA CATCCACAG GGAGCAGTG AGGC[C/T]CC CGTCTCGGT ATGGGCAGT CAGCCTTTG ACTTCTACC CC	SEQ ID NO: 166
chr 11	124266 390	A	G	<i>OR8B3</i>	p.P286 P	0.009 31	0.003 09	1.44 E- 08	3.04[2.17- 4.25]	GTGCAACTT TGACATCCT TGTTTCCTCA AACTGTAGA TGAG[A/G]G GATTGAGCA TGGGCACCA CATTAGTGT AGAAAACAG AAG	SEQ ID NO: 167

chr 11	124620 746	G	T	<i>VSIG2</i>	p.N97 K	0.005 15	0.000 02	6.91 E- 29	288. 96[6 7.73- 1232 .76]	CGTCAGTCA GTTTCAGTG TGGCCACCC CCACTGTGG GGGG[G/T]TT CTGAAGCAG GCTGACCCG CTTTGACTTA GAACCAGTT G	SEQ ID NO: 168
chr 12	368928	T	C	<i>SLC6A 13</i>	p.P97P	0.008 82	0.005 15	2.84 E- 03	1.72[1.23- 2.41]	CCTCTAAGC GTCCTCCTA CCTCCAGAA TTCTATACAT CTA[T/C]GGG ACTCCCCAG AGGGGCCGT AAGTGCAGG AGATGGAAG T	SEQ ID NO: 169
chr 12	704483 8	C	T	<i>ATN1</i>	p.Y13 6Y	0.011 03	0.007 57	1.73 E- 02	1.46[1.08- 1.98]	ATATCGACC AGGACAACC GAAGCACGT CCCCCAGTA TCTA[C/T]AG CCCTGGAAG TGTGGAGAA TGACTCTGA CTCATCTTCT G	SEQ ID NO: 170
chr 12	109594 16	C	A	<i>TAS2R 8</i>	p.R55I	0.007 84	0.004 77	1.11 E- 02	1.65[1.15- 2.36]	TACAATGCC ATTTACAAC CATTACACT GATCAAACA AATT[C/A]TG GCGATAACT AAATTGGTA AGGATGTAG TCAACTGTG GA	SEQ ID NO: 171
chr 12	114617 69	G	T	<i>PRB4</i>	p.P50T	0.026 72	0.006 86	4.29 E- 29	3.98[3.23- 4.9]	TGTGGGGGT GGTCCTTGT GGCTTTTCT GGAGGAGGT GGGG[G/T]AC GTTGGGGCT GGTTTCCTCC TTGTGGGCG TCGTCTTCT	SEQ ID NO: 172
chr 12	130615 83	A	G	<i>GPRC 5A</i>	p.I134 V	0.013 48	0.009 32	1.04 E- 02	1.45[1.11- 1.91]	CAAGCTCGT CCGGGGGAG GAAGCCCCCT TTCCCTGTTG GTG[A/G]TTC TGGGTCTGG CCGTGGGCT TCAGCCTAG TCCAGGATG T	SEQ ID NO: 173

chr 12	152623 59	C	T	<i>RERG</i>	p.V95 V	0.009 80	0.006 58	1.84 E- 02	1.49[1.09- 2.06]	TGGGCTTTTT GATCTCATC TAGGATGTT CTTAAGTGG CAG[C/T]ACT TCCTCAAAA CTTCCTCGGT CAGTAATGT CGTAGACCA	SEQ ID NO: 174
chr 12	482402 33	G	A	<i>VDR</i>	p.A35 3A	0.005 64	0.003 30	1.83 E- 02	1.71[1.11- 2.64]	GGAAGTTGA TGAGGGGCT CAATCAGCT CCAGGCTGT GTCC[G/A]GC TGTGAGAGA CAATGGCCA GGTACTGCG GGCAGAGCT GA	SEQ ID NO: 175
chr 12	494255 75	C	T	<i>KMT2 D</i>	p.V43 05I	0.005 39	0.002 57	1.83 E- 03	2.1[.36- 3.25]	TTTGGCTCTT GAGGGCTGG ATGGTGGAG GTGTGGGAT GGA[C/T]AGG GCCAAGGAC TGGTCCTGT AGATAAGGC TCCTGGTGG G	SEQ ID NO: 176
chr 12	504801 02	G	T	<i>SMAR CD1</i>	p.Q11 2H	0.007 807	0.004 359	8.42 E- 03	1.8[.14- 2.71]	CCCGCAAGA GACCTGCCC CTCAGCAGA TCCAGCAGG TCCA[G/T]CA GCAGGCGGT CCAAAATCG AAACCACAA GTAAGATGA TC	SEQ ID NO: 177
chr 12	507457 92	G	A	<i>FAM1 86A</i>	p.A16 08V	0.005 88	0.000 56	3.41 E- 14	10.5 6[6.1 9- 18.0 2]	CTGGGCCTG CTGAGGGGT GAGAGGGAT CCCCTGAGC CTGC[G/A]CC TGCTGAGGG GTGAGAGGG ATCCCCAGT TCCTGCGCC TG	SEQ ID NO: 178
chr 12	507468 36	A	G	<i>FAM1 86A</i>	p.V12 60A	0.025 74	0.000 07	7.89 E- 110	399. 44[1 26.7 2- 1259 .11]	CTGGGCCTG CTGAGGAGT AAGAGGGAT CCCCAGTTC CTGA[A/G]CC TGCTTAGGG GTGAGAGTG ATTCCGAGA GCCTGCGCC TG	SEQ ID NO: 179

chr 12	507481 69	T	G	<i>FAM1</i> 86A	p.K81 6Q	0.005 21	0.002 65	1.29 E- 02	1.97[1.11- 3.26]	TCTTGCAAA TATTGCTCCT GCCTTTGTTT TTCCTTCTCC T[T/G]GTGGT CTTTCTGTAC TGTTGAGAC TGTTGGAAT ATCTCTT	SEQ ID NO: 180
chr 12	529608 23	C	A	<i>KRT74</i>	p.G50 7V	0.005 21	0.002 49	9.47 E- 03	2.09[1.18- 3.47]	GGCTGGGGT GCTCTTGCC CTGGGTGTC CTTGAGGTC TCCC[C/A]CT CGCGCCTCT GTGGTCTTG GTCTGCCCCG CTCTGGGTG CT	SEQ ID NO: 181
chr 12	529620 50	G	A	<i>KRT74</i>	p.R420 W	0.008 33	0.005 55	2.50 E- 02	1.51[1.06- 2.13]	AGTTTCAGG CTCATGAGC TCCTGGTAC TCGCGCAGC ATCC[G/A]CG CCAGCTCCT CCTTGGCCT GGTGCAGGG CGCCCTCCA GC	SEQ ID NO: 182
chr 12	534481 14	G	A	<i>TENCI</i>	p.T13 T	0.005 39	0.003 01	1.35 E- 02	1.79[1.16- 2.77]	TCATGGAGC GGCGCTGGG ACTTAGACC TCACCTACG TGAC[G/A]GA GCGCATCTT GGCCGCCGC CTTCCCCGC GCGGCCCGA TG	SEQ ID NO: 183
chr 12	535169 93	C	T	<i>SOAT2</i>	p.V45 5V	0.010 54	0.007 32	2.52 E- 02	1.45[1.06- 1.97]	TGGGGTTCT TCTATCCCGT CATGCTGAT ACTCTTCCTT GT[C/T]ATTG GAGGTGAGC TGGTCTCTGT GCCACTGGA AGGGAGCC	SEQ ID NO: 184
chr 12	537144 30	G	T	<i>AAAS</i>	p.T57 N	0.009 31	0.000 00	5.15 E- 56	Inf	GATGAAGGC AGTTCTTGT GCCATGGTC CAGCCTTCC AGGG[G/T]TC TTAGGGGA TCCTTTGTCA GTTGTAGGA CAGGAAGAT T	SEQ ID NO: 185

chr 12	558464 89	C	A	<i>OR6C2</i>	p.L164 L	0.005 15	0.002 87	1.70 E- 02	1.8[1 .16- 2.8]	TGATGATCA TTGTTCCACC ACTTAGCTT AGGCCTCCA GCT[C/A]GAA TTCTGTGACT CCAATGCCA TTGATCATTT TAGCTGTG	SEQ ID NO: 186
chr 12	563509 77	C	G	<i>PMEL</i>	p.E370 D	0.005 88	0.002 57	5.00 E- 04	2.3[1 .51- 3.49]	CCTCTGAAA CTGGCACCT TCTCAGGTG TCATACCTG TGCT[C/G]TC TGCAGTTGG CATCTGCAC AGGTGCAGT GCTTATGAC TT	SEQ ID NO: 187
chr 12	570092 16	G	A	<i>BAZ2A</i>	p.N10 6N	0.007 35	0.004 97	4.23 E- 02	1.48[1 1.02- 2.14]	GTCCCCCG AGAAGTGGG AGAGAAGGG GTGGGTCCT TGAG[G/A]TT GCTGCCAGG ATTGGCAGA TGGGTACTG TGAGTAGTT CC	SEQ ID NO: 188
chr 12	575693 39	G	A	<i>LRP1</i>	p.G12 15E	0.008 58	0.005 26	8.35 E- 03	1.64[1 1.16- 2.3]	GAAGGCATT GTGTGTTCTT GCCCTCTGG GCATGGAGC TGG[G/A]GCC CGACAACCA CACCTGCCA GATCCAGAG CTACTGTGC C	SEQ ID NO: 189
chr 12	667251 38	G	A	<i>HELB</i>	p.G95 9S	0.005 88	0.003 56	2.31 E- 02	1.66[1 1.1- 2.51]	TCGTTTGAA ACATTTCTTG CAAAGTAAG CTCTCCTCTA GC[G/A]GCGC ACCTCCAGC AGATTTTCC GTCCCCACG GAAGAGCTC	SEQ ID NO: 190
chr 12	856951 06	C	T	<i>ALXI</i>	p.N27 8N	0.009 56	0.000 00	1.86 E- 57	Inf	TTTCAAACC ACCAGAACC AGTTCAGCC ACGTGCCCC TCAA[C/T]AA TTTTTTCCT GACTCTCTTC TTACTGGGG CAACCAATG	SEQ ID NO: 191

chr 12	899169 68	G	A	<i>POCI B- GALN T4</i>	p.I450I	0.006 62	0.004 28	3.83 E- 02	1.55[1.05- 2.29]	GGTTGTTGT CAGGAGAAT TATAATCTA AACATTTCAG ACGA[G/A]AT CCCTCTACT GCGAATAGC CCCATGCCA GCCTGGTCT AT	SEQ ID NO: 192
chr 12	956942 43	C	T	<i>VEZT</i>	p.P712 S	0.001 96	0.000 05	1.23 E- 05	41.5[11.7- 147]	TGAACCACA AGCAGATGG AAGTGGTCT GACCACTGC CCCT[C/T]CA ACTCCCAGG GACTCATT CAGCCCTCC ATTAAGCAG AG	SEQ ID NO: 193
chr 12	104144 426	C	T	<i>STAB2</i>	p.P217 OS	0.006 13	0.003 66	1.77 E- 02	1.68[1.12- 2.52]	CTATGTCGG AGATGGGCT GAACTGTGA GCCGGAGCA GCTG[C/T]CC ATTGACCGC TGCTTACAG GACAATGGG CAGTGCCAT GC	SEQ ID NO: 194
chr 12	108920 173	G	A	<i>SART3</i>	p.D69 ID	0.005 64	0.003 48	3.10 E- 02	1.62[1.06- 2.47]	TGATGCTGT CCTTGCTGCT GTCGTGCAG CACCTGGG CAT[G/A]TCC CTCTTCAGG GAGGCTGCC TTCTCCTTCT GCTTCGAAG	SEQ ID NO: 195
chr 12	111317 855	T	C	<i>CCDC 63</i>	p.L172 S	0.007 11	0.004 78	4.98 E- 02	1.49[1.03- 2.17]	CTCCAGCAC TGCCTGTTG ATGGAGAAG AAAACCATG AACT[T/C]GG CCATTGAGC AATCTTCTC AGGCCTATG AGCAGAGGT GG	SEQ ID NO: 196
chr 12	119594 354	C	T	<i>SRRM 4</i>	p.S529 S	0.013 48	0.000 00	4.82 E- 80	Inf	CCATCCCCT ACTATCGGC CCAGCCCCT CCTCATCCG GCAG[C/T]CT CAGCAGCAC CTCCTCCTG GTACAGCAG CAGCAGTAG CC	SEQ ID NO: 197

chr 12	122361 711	C	T	<i>WDR6</i> 6	p.R188 W	0.012 25	0.008 07	5.94 E- 03	1.53[1.15- 2.03]	TGAAAGGCA GCCCTCAGG AGAGCTTGA GGAGAAAAC CGAC[C/T]GG ATGCCCCAA GATGAACTG GGACAAGAA AGAAGGGAC TT	SEQ ID NO: 198
chr 12	122404 946	C	T	<i>WDR6</i> 6	p.R860 C	0.012 01	0.008 07	1.00 E- 02	1.49[1.12- 1.99]	ACAAGTCCT CCCAGTGAG AAGCATGGC GGAGCTACA GAAA[C/T]GC TACTTGGTG TTTATTAAC AGAGACAAG GTAACAGCG CT	SEQ ID NO: 199
chr 12	122676 056	A	G	<i>LRRC4</i> 3	p.Y15 9C	0.005 39	0.002 69	3.52 E- 03	2.01[1.3- 3.1]	CCCGAAGGC CCTTTCATCA CTTACAAC ATTACGTGA CCT[A/G]TGA TTTTGTGAA AGATGAAGA AGGCGAAAT GAATGAGTC C	SEQ ID NO: 200
chr 12	123706 313	T	G	<i>MPHO</i> <i>SPH9</i>	p.S160 R	0.006 51	0.000 46	8.30 E- 15	14.3 6[7.8 - 25.7 8]	GTGGATTCA GGATAATGG ATAACAGAT TCATTTCTCT CAC[T/G]GCT TAGAGAAAA AAAACCCAT TTGACTTTCC GAAGATACT	SEQ ID NO: 201
chr 12	124364 285	C	T	<i>DNAH</i> 10	p.H27 39H	0.007 35	0.004 64	1.93 E- 02	1.59[1.1- 2.3]	GGGATCCCA TATTGTTTGG AGACTTCCA GATGGCTCT GCA[C/T]GAA GGAGAACCA CGCATTTAT GAAGACATC CAGGACTAC G	SEQ ID NO: 202
chr 12	125396 833	G	A	<i>UBC</i>	p.D49 5D	0.028 92	0.012 95	4.07 E- 08	2.27[1.69- 3.06]	CATCTTCCA GCTGTTTCCC AGCAAAGAT CAACCTCTG CTG[G/A]TCA GGAGGGATG CCTTCCTTGT CTTGGATCTT TGCCTTGA	SEQ ID NO: 203

chr 12	125397 541	T	C	UBC	p.Q25 9Q	0.005 15	0.000 07	1.03 E- 24	71.9 8[31. 86- 162. 59]	AGATCAACC TCTGCTGGT CAGGAGGAA TGCCTTCCTT GTC[T/C]TGG ATCTTTGCTT TGACGTTCT CGATAGTGT CACTGGGCT	SEQ ID NO: 204
chr 12	125398 297	A	G	UBC	p.T7T	0.012 53	0.000 10	1.46 E- 33	94.0 3[44. 17- 200. 19]	CACTGGGCT CAACCTCGA GGGTGATGG TCTTACCAG TCAG[A/G]GT CTTCACGAA GATCTGCAT TGTCTAACA AAAAAGCCA AA	SEQ ID NO: 205
chr 12	132625 260	G	A	DDX5 1	p.S487 S	0.022 30	0.000 01	2.59 E- 131	2540 .86[3 54- 1823 7.12]	CCAGGACCA GGTGCAGGA CGACCAGCG GCTTAGAGC TGAG[G/A]CT GCAGGGCAC GTAGTGGTG CTACAGGGA CGGCAGGGG GT	SEQ ID NO: 206
chr 13	368717 82	G	T	CCDC 169	p.V25 V	0.006 37	0.003 72	1.29 E- 02	1.72[1.14- 2.59]	GGGACCCCA CACCGCGCC GCCCCCGGA CTCACTTCTT GCG[G/T]ACT TCTTCCAGC AACTGCTGT TTCAGGCGG TTGGTGCTC A	SEQ ID NO: 207
chr 13	423521 71	T	C	VWA8	p.M76 7V	0.005 88	0.003 63	2.56 E- 02	1.62[1.07- 2.45]	ACCAATAAT AAGTGTCT CCAAGGAGA AAGTCTTTC AGCA[T/C]AT CTTCCATCA CTATCACAT GCTAGAGAA AAAGGAACT AG	SEQ ID NO: 208
chr 13	492817 85	T	A	CYSLT R2	p.L278 I	0.016 93	0.001 66	1.09 E- 30	10.3 7[7.3 2- 14.5]	CACACTGAG GACCGTCCA CTTGACGAC ATGGAAAGT GGGT[T/A]TA TGCAAAGAC AGACTGCAT AAAGCTTTG GTTATCACA CT	SEQ ID NO: 209

chr 13	763816 79	T	C	<i>LMO7</i>	p.H18 7H	0.008 82	0.004 66	7.04 E- 04	1.9[.36- 2.67]	TCCAAACAT ACTCTGATG ACATCTTGT CTTCTGAAA CACA[T/C]AC CAAAATTGA TCCCACTTCT GGCCCAAGG CTCATAACC C	SEQ ID NO: 210
chr 13	995404 20	G	T	<i>DOCK 9</i>	p.P679 T	0.008 33	0.000 00	3.06 E- 49	Inf	CGTAGGTGA ACATATATT AAAAAAAAA CAAACCTTA AGGG[G/T]CT GAGAGTCTT CCTCATCTG AATCTTTGA ATTCAATGC AA	SEQ ID NO: 211
chr 13	103382 057	T	C	<i>CCDC 168</i>	p.K69 97R	0.000 25	0.000 02	1.26 E- 01	14.3 1[0.8 9- 228. 77]	TTTTCTTTCA GAATAGAAG TTGATATCG TCATGATGA GGT[T/C]TTG ATGCTGATT TATGTTTGCT TTGGAAACA ATCCAATCT	SEQ ID NO: 212
chr 13	103382 483	G	A	<i>CCDC 168</i>	p.T685 5I	0.000 49	0.000 22	2.60 E- 01	2.18[0.49- 9.67]	TCTATATTTTC CTGCTTTTGT GGGACTTAC AGGAAGGTG GT[G/A]TAAT AATTAAGGT TTCCTTTCTG CACTCTCTA GTACAATG	SEQ ID NO: 213
chr 13	103382 660	A	G	<i>CCDC 168</i>	p.V67 96A	0.009 56	0.008 43	4.27 E- 01	1.13[0.82- 1.57]	TTCTGATTCC TGACTTAAA TAAGAGTTG GCTTCCAGA AAC[A/G]CAC ATTCTCACT CTCACTTACT TCAAGACAT GAACACTC	SEQ ID NO: 214
chr 13	103382 700	C	T	<i>CCDC 168</i>	p.E678 3K	0.000 25	0.000 05	2.43 E- 01	4.63[0.48- 44.5 2]	ACACATTCC TCACTCTCA CTTACTTCA AGACATGAA CACT[C/T]GT CCAAGTCAG CTGGACTCT CAATATCTG TCTGAATAT CA	SEQ ID NO: 215

chr 13	103383 228	C	T	CCDC 168	p.E660 7K	0.000 25	0.000 04	1.87 E- 01	7[0.6 3- 77.2 1]	TATTGTAAA TCAAGATCT ATTTGATGG AGAGATTTC TCCT[C/T]AG AAAGTAACA AAATTCTGT TTTGTGTTT TGGTCCTGT G	SEQ ID NO: 216
chr 13	103383 339	T	C	CCDC 168	p.R657 0G	0.000 49	0.000 20	2.19 E- 01	2.5[0 .55- 11.2 7]	TTCTTTCTCT CATGAGCAC TGGTCATTG CATAAGATT CTC[T/C]TAC AATTCTGGG AAAGGCTTT CATTTGTATC TCCAATGTT	SEQ ID NO: 217
chr 13	103383 524	T	G	CCDC 168	p.E650 8A	0.002 70	0.002 81	1.00 E+0 0	0.96[0.52- 1.77]	ATTTTCTAGC TTATTAATA CTCTGTAGC TTTGTGATTG TC[T/G]CCTC ACTGTCAC TGAAACATC ACAATCAG TGTCTTCAT	SEQ ID NO: 218
chr 13	103383 666	A	C	CCDC 168	p.S646 1A	0.000 25	0.000 04	1.89 E- 01	6.91[0.63- 76.1 9]	GTCCCTTCTA GAGACATAA AGTTCATTG TTTTATGTCT AG[A/C]ATAG AACCTCCAA CTGTTATCTT TTGAAATAG TCCCTTTT	SEQ ID NO: 219
chr 13	103383 792	G	A	CCDC 168	p.H64 19Y	0.002 94	0.000 30	1.49 E- 07	9.81[4.68- 20.5 6]	ATCAGATTC AGTTGTATTT CAAGTGCTT TTGACTCTA AAT[G/A]ACT AGTAAGCTT ATTTTTTTCT TTGGGAGTA AACTGTTCT	SEQ ID NO: 220
chr 13	103383 812	T	G	CCDC 168	p.E641 2A	0.000 25	0.000 00	6.73 E- 02	Inf[NaN -Inf]	AAGTGCTTT TGACTCTAA ATGACTAGT AAGCTTATT TTTT[T/G]CT TTGGGAGTA AACTGTTCT AAAAGGGAT TTGTGCTGC GT	SEQ ID NO: 221

chr 13	103383 951	C	T	CCDC 168	p.D63 66N	0.001 72	0.002 82	2.75 E- 01	0.61[0.28- 1.3]	AAGTCGTCA GGCTTATAG GCTTGTATG TTATCTAGTT TAT[C/T]AGA AGAAACTTT GTCTTGGAT CATATTTTAA ACCTGGGAC	SEQ ID NO: 222
chr 13	103384 070	C	T	CCDC 168	p.S632 6N	0.000 25	0.000 12	4.28 E- 01	1.98[0.24- 16.0 6]	ATGTTCTGC ATTTGTACT GTCTGCAAC TATTTTGACT TCG[C/T]TAC TTTAACTTG AGGCGGTAT GGGCACAGT TCCTGGGAA	SEQ ID NO: 223
chr 13	103384 712	G	A	CCDC 168	p.T611 2M	0.021 32	0.024 94	1.58 E- 01	0.85[0.68- 1.06]	ATACTCTAA TTTCTTTCTA TTGCTTGGT GTACCACGC CCC[G/A]TGA TATTAAGCA TCTGTGGAA TTGGGTGAT TCTGGATTTT	SEQ ID NO: 224
chr 13	103385 064	T	C	CCDC 168	p.K59 95E	0.003 43	0.004 24	5.30 E- 01	0.81[0.47- 1.39]	GGGTGTGCA CTACTGCTT GTGTCCATT CTTCCTCTCT CCT[T/C]CTC CAGATTGGC AGTCCTGGC CTGTGTCAT CTCTGTTTTTC	SEQ ID NO: 225
chr 13	103385 294	G	A	CCDC 168	p.P591 8L	0.000 25	0.000 00	6.72 E- 02	Inf[NaN -Inf]	TGATTGAAA TTGAAAAGT CCAGGGAGG GAATAGGGA CTTC[G/A]GA AGAAATTCC AGAACACCT TCCTCTTGTT CTGAAATGA G	SEQ ID NO: 226
chr 13	103385 340	C	A	CCDC 168	p.A59 03S	0.000 25	0.000 12	4.26 E- 01	1.99[0.24- 16.1 6]	AATTCCAGA ACACCTTCC TCTTGTCTG AAATGAGCA ATG[C/A]CTG CTTCCTTCCC CCTTTTGCA GGGTCAATC TCTGTCATA	SEQ ID NO: 227

chr 13	103385 520	C	T	CCDC 168	p.G58 43R	0.000 25	0.000 02	1.31 E- 01	13.7 5[0.8 6- 219. 86]	GGAAACTTA GAAAGGATA GTGTTCGTC CTGGTCTTGT GCC[C/T]ATG TTCACACCG TCGGATCAC TTGCTTTTTC ATGACAATA	SEQ ID NO: 228
chr 13	103385 654	G	T	CCDC 168	p.S579 8Y	0.000 25	0.000 28	1.00 E+0 0	0.86[0.11- 6.5]	TTTGAGTGA TCCCTTTGTC TGTTGGTGCT AACACTTTG GGA[G/T]AA AACATTTTG CTGATTCTAT CATTACTTTG TCCATCTTC	SEQ ID NO: 229
chr 13	103386 222	C	T	CCDC 168	p.V56 09I	0.000 25	0.000 00	6.74 E- 02	Inf[NaN -Inf]	GCCTCTGGG CGGGGCACA TACTGTTCTG CTTGCTTAA CAA[C/T]GTT TTTATCAAC GCCTTCAAC TGAGTCTCT ATTTGTTATT	SEQ ID NO: 230
chr 13	103387 002	C	T	CCDC 168	p.V53 49I	0.000 25	0.000 07	2.98 E- 01	3.42[0.38- 30.5 6]	TGCTTTTCAT TTTTAACATC TTTTGGGAT ATCACCAAC GA[C/T]GGAC TCTCTATGTA CAGTCTCCC CTATGTGTG ATATTCTC	SEQ ID NO: 231
chr 13	103387 043	C	T	CCDC 168	p.R533 5Q	0.002 70	0.004 28	1.64 E- 01	0.63[0.34- 1.15]	GGACTCTCT ATGTACAGT CTCCCCTAT GTGTGATAT TCTC[C/T]GC AAAATAGGT CTTTTAAGTC TTAGCATTTT ATTACCTAA	SEQ ID NO: 232
chr 13	103387 196	G	A	CCDC 168	p.P528 4L	0.020 10	0.017 80	2.99 E- 01	1.13[0.9- 1.42]	TTCACCTTCA CATTCCTGC ACCTTCTCTT CCTGATGTTT G[G/A]GGAA TATTAAGAT GCTTACTATT TGCACGTCA TCCTCTTC	SEQ ID NO: 233

chr 13	103387 313	C	A	CCDC 168	p.G52 45V	0.000 49	0.000 00	4.64 E- 03	Inf[NaN -Inf]	GATTAAAAT ATCACCAGC AATTGGCCT TATACATGT GCCT[C/A]CC TCAGTATCT GGTGATACC TGGAGTTT ACTAGGGA AA	SEQ ID NO: 234
chr 13	103387 767	C	T	CCDC 168	p.V50 94M	0.000 49	0.000 07	5.68 E- 02	6.92[1.27- 37.8 1]	GACCGTGAC TGTGGGAGA GACACTTTT GCAATTCTT ATCA[C/T]GT TCTCCTGTCC TTCTGTTGTA TCAAACCTTA AGATATGGT	SEQ ID NO: 235
chr 13	103388 015	C	G	CCDC 168	p.G50 11A	0.035 78	0.034 78	7.24 E- 01	1.03[0.87- 1.22]	TTTGTCTTCC ATATCTATT TGAGTCCAC CTTTCTCTTC T[C/G]CCTGT GCTGTGGGT TGCACTGGT CCTTTTGAGT TGCTTAA	SEQ ID NO: 236
chr 13	103388 343	A	T	CCDC 168	p.L490 2M	0.000 25	0.000 02	1.30 E- 01	13.8 3[0.8 7- 221. 2]	CCATTGCAT AGAAGTGCA AGTGGGAGT GCCTCTGCC CTCA[A/T]AT GTATCCTTTT GGGGAGTAT TCTACCTTCC CTGCCTTCT	SEQ ID NO: 237
chr 13	103388 378	C	T	CCDC 168	p.G48 90D	0.002 45	0.003 50	3.31 E- 01	0.7[0 .37- 1.32]	CCTCAAATG TATCCTTTTG GGGAGTATT CTACCTTCCC TG[C/T]CTTC TATTTTACT CTGTCCTTTG CCTCTTTATA TGGCAT	SEQ ID NO: 238
chr 13	103388 877	G	A	CCDC 168	p.P472 4S	0.002 94	0.003 48	6.78 E- 01	0.85[0.47- 1.52]	GTTTGCCTTG AAGGCAATG ATTCTGGA TCTCAAGAT GTG[G/A]CAT AAAGCTTCT TGTTATTCGT GGTTCACCT TCCTCTTCT	SEQ ID NO: 239

chr 13	103388 880	T	C	CCDC 168	p.M47 23V	0.043 14	0.041 03	5.14 E- 01	1.05[0.9- 1.23]	TGCCTTGAA GGCAATGAT TCCTGGATC TCAAGATGT GGCA[T/C]AA AGCTTCTTGT TATTCGTGG TTCACCTTCC TCTTCTTTT	SEQ ID NO: 240
chr 13	103389 072	G	A	CCDC 168	p.P465 9S	0.001 96	0.000 25	6.03 E- 05	7.86[3.3- 18.7 5]	TTCACCTGC AGTTCCTTTG TTTTTAGTAT ATGGGAAAG GG[G/A]TGAT TTCTCTGCCT TTACAGCTA TGTA CTGG GATGCATT	SEQ ID NO: 241
chr 13	103389 164	T	G	CCDC 168	p.K46 28T	0.004 41	0.002 76	6.72 E- 02	1.6[0 .98- 2.61]	TGAAATATT TGCTTTATCC TTTTGGATCT GGGCCATGT AT[T/G]TTGT TCTGTTTGA ATCACCTGT GATATCATT CAAATATGA	SEQ ID NO: 242
chr 13	103389 306	G	A	CCDC 168	p.R458 1X	0.000 49	0.000 23	2.68 E- 01	2.13[0.48- 9.44]	GATCTTGTT ACTCCTTGTT CCTCTTTTTT GCCTGCTGT TC[G/A]TTTG TCTAATTAC AGTGAGATA GAGAAGGTA TTGTCAGA	SEQ ID NO: 243
chr 13	103389 321	A	G	CCDC 168	p.C457 6R	0.000 98	0.000 11	3.09 E- 03	9.25[2.61- 32.7 9]	TGTTCTCTT TTTTGCCTGC TGTTCTTTG TCTAATTAC [A/G]GTGAG ATAGAGAAG GTATTGTCA GAAACACAT CCAGTTCA	SEQ ID NO: 244
chr 13	103389 594	C	A	CCDC 168	p.V44 85L	0.000 25	0.000 04	1.89 E- 01	6.93[0.63- 76.4]	TTGTATTCTT GTACTGTTTT TACATCATTT GAGCTATCC A[C/A]CCCAA AAGACTTTG TATGTGCTA TTTTCCCTGC ATCAAAT	SEQ ID NO: 245

chr 13	103389 656	A	G	CCDC 168	p.L446 4S	0.002 45	0.001 36	8.43 E- 02	1.8[0 .93- 3.48]	TATTTTCCCT GCATCAAAT GATTTCTGCT GCCTTAGTT GC[A/G]AAGT AGCAGATTT TATTATTCCT TGTAAGTCT TCCTCTCC	SEQ ID NO: 246
chr 13	103389 867	C	T	CCDC 168	p.E439 4K	0.000 25	0.000 02	1.30 E- 01	13.8 7[0.8 7- 221. 8]	TGTTGCTCTT CAGTTTCTCC ATCCCTGTTC CCTTGCTCCT [C/T]ACCTTC TCCGTCCTCT TTCCCTTGCT CCTGGCCTT CTCCA	SEQ ID NO: 247
chr 13	103389 885	T	G	CCDC 168	p.K43 88Q	0.011 27	0.014 80	7.81 E- 02	0.76[0.56- 1.02]	CCATCCCTG TTCCCTTGCT CCTCACCTTC TCCGTCCTCT T[T/G]CCCTT GCTCCTGGC CTTCTCCATC CCTTTTCCCT GGCTCT	SEQ ID NO: 248
chr 13	103390 083	C	T	CCDC 168	p.G43 22S	0.004 90	0.003 86	2.99 E- 01	1.27[0.8- 2.01]	ATGTAATCT TTTGCTTTTT GTACTTCAC TTGCGCTAT CAC[C/T]CTC ACTGGGCAC CCCATTGCT TTTTTCCCTG TCTCTGAT	SEQ ID NO: 249
chr 13	103390 086	C	T	CCDC 168	p.E432 1K	0.012 99	0.010 98	2.45 E- 01	1.19[0.89- 1.57]	TAATCTTTTG CTTTTTGTAC TTCACTTGC GCTATCACC CT[C/T]ACTG GGCACCCCA TTTGCTTTTT TCCCTGTCTC TGATGAT	SEQ ID NO: 250
chr 13	103390 173	G	C	CCDC 168	p.Q42 92E	0.000 74	0.000 69	7.59 E- 01	1.07[0.33- 3.46]	TGCCTTGGTT GTAAAATAC CAGGTCTGA TTATTCCTTG TT[G/C]GTCT TCCTCTCCTT CTATTCTTGT GTCCAATAT ATAATGG	SEQ ID NO: 251

chr 13	103390 257	C	A	CCDC 168	p.E426 4X	0.000 25	0.000 07	2.94 E- 01	3.47[0.39- 31.0 8]	AGAGAAGAA TTGGAAGGC AAATATAGG AACAGAACT CTTT[C/A]CT GTTCAATCTT GTCTCCATC CATTTTCCCT TGCTCTATG	SEQ ID NO: 252
chr 13	103390 322	T	C	CCDC 168	p.E424 2G	0.006 86	0.009 27	1.25 E- 01	0.74[0.5- 1.08]	TTTCCCTTGC TCTATGCCT ACTCCATCT GCTTTCTGTT GC[T/C]CTTC AACTTCGTG ATCCATTTTC CCTTGCTCTT TGTCTTC	SEQ ID NO: 253
chr 13	103390 332	C	T	CCDC 168	p.E423 9K	0.000 49	0.000 36	6.59 E- 01	1.37[0.32- 5.87]	TCTATGCCT ACTCCATCT GCTTTCTGTT GCTCTTCAA CTT[C/T]GTG ATCCATTTTC CCTTGCTCTT TGTCTTCTCT ATCAACC	SEQ ID NO: 254
chr 13	103390 626	T	C	CCDC 168	p.I414 1V	0.000 98	0.000 36	7.53 E- 02	2.76[0.94- 8.07]	TGTTGCATG TAATCTTTTG CTTTTTGTAC TTTGATTGTG A[T/C]ATCAC CCTTACTGG CCACTCCAT CTGCTTTTTC CCCTGCC	SEQ ID NO: 255
chr 13	103390 701	A	T	CCDC 168	p.Y41 16N	0.004 90	0.004 21	5.32 E- 01	1.17[0.74- 1.84]	CCTGCCTCT GATGATTTTT GGTGTGATA GTTCTGGAA GAT[A/T]GTA TCTTGTTATT TCAGTGACA TACTCTGCTT TTTCTCTC	SEQ ID NO: 256
chr 13	103390 938	A	T	CCDC 168	p.L403 7M	0.000 25	0.000 41	1.00 E+0 0	0.6[0 .08- 4.42]	GCCCTAATT TTTTCCATTT TTTGCCCTCTG TTCTTTTTGC A[A/T]TATAG ATTCTAGGG CCTTTTTTAC ACTGTTTGA GATATTA	SEQ ID NO: 257

chr 13	103391 300	G	A	CCDC 168	p.P391 6L	0.000 25	0.000 22	6.02 E- 01	1.14[0.15- 8.74]	TTTTTCCAAA GCCTTTTCCA CTCTGTCTTT GTCTTTCTGC [G/A]GCATAT GTTTTGCTTT TTCAATACT GCTTAAACT ATCATC	SEQ ID NO: 258
chr 13	103391 357	T	A	CCDC 168	p.K38 97I	0.000 49	0.000 14	1.45 E- 01	3.42[0.73- 16.1 3]	TTCAATACT GCTTAAACT ATCATCAAT TGGCTGCTC ACAT[T/A]TT TCCATTGTAT CTGATAATT CCTGCTGTG TTGATGATG A	SEQ ID NO: 259
chr 13	103392 113	C	G	CCDC 168	p.G36 45A	0.000 25	0.000 29	1.00 E+0 0	0.86[0.11- 6.47]	TATGTGTTGT TTTGTACTTT TAACATTAC TTGAGATCA CC[C/G]CATC AATTGTTTCT TTATTCAATT TGAAGTGAG GTAAAGA	SEQ ID NO: 260
chr 13	103392 562	C	A	CCDC 168	p.M34 95I	0.021 08	0.026 02	5.76 E- 02	0.81[0.65- 1]	TTGATATTA AATCAAAGA CCTGTACCC CATCTGATG ATTT[C/A]AT TCCTTTTGA AATAAGAGA CTTGCAATT TTATAGTTT	SEQ ID NO: 261
chr 13	103392 735	G	C	CCDC 168	p.P343 8A	0.000 25	0.000 04	1.90 E- 01	6.88[0.62- 75.8 8]	ATAGTGCTT AGCTGATCT GCAGAAAAC AAGTCTAGT CCTG[G/C]TG TCCGGCTTG ATAAATTAC CTCCTTCTGA TAATGCTTC C	SEQ ID NO: 262
chr 13	103392 741	G	A	CCDC 168	p.R343 6W	0.008 82	0.008 75	9.31 E- 01	1.01[0.72- 1.42]	CTTAGCTGA TCTGCAGAA AACAAGTCT AGTCCTGGT GTCC[G/A]GC TTGATAAAT TACCTCCTTC TGATAATGC TTCCTTTTCC	SEQ ID NO: 263

chr 13	103393 330	A	T	CCDC 168	p.D32 39E	0.001 23	0.000 21	4.33 E- 03	5.77[2.03- 16.3 8]	CTTTAATATT CAAATGTAT TCCTTCTGA ACATGGAGG TTG[A/T]TCC ACCGGAATA CCTACTTCAT GTGATGCTT TCTCTACCA	SEQ ID NO: 264
chr 13	103393 337	G	A	CCDC 168	p.P323 7L	0.000 25	0.000 16	5.03 E- 01	1.54[0.19- 12.1 3]	ATTCAAATG TATTCCTTCT GAACATGGA GGTTGATCC ACC[G/A]GA ATACCTACT TCATGTGAT GCTTTCTCTA CCATTGGGC T	SEQ ID NO: 265
chr 13	103393 383	C	G	CCDC 168	p.V32 22L	0.000 25	0.000 02	1.31 E- 01	13.7 9[0.8 6- 220. 58]	CCTACTTCAT GTGATGCTT TCTCTACCAT TGGGCTTAG AA[C/G]TTTT GAACTCATG ATTTCTTCTG CTGAGCCTT CTTTCTTG	SEQ ID NO: 266
chr 13	103393 580	T	C	CCDC 168	p.Q31 56R	0.000 49	0.000 20	2.20 E- 01	2.49[0.55- 11.2 6]	TTTCTGTCTA TTTGATTTTA ATGTAATAT CCAACCTTG AT[T/C]GCTC TTTTCCCCAA AGATTTTCA TTGAAACTT TCAGAGAT	SEQ ID NO: 267
chr 13	103393 731	C	T	CCDC 168	p.V31 06M	0.000 25	0.000 59	7.28 E- 01	0.41[0.06- 3.03]	TCAGAATCC AGAATACTT TCGGGAACA TGATCTGGA TTCA[C/T]CT GTTCTTTCTG CTCTGCAGG CACTTTGTG CTGTACCTCT	SEQ ID NO: 268
chr 13	103394 336	A	G	CCDC 168	p.M29 04T	0.000 25	0.000 05	2.44 E- 01	4.6[0 .48- 44.1 9]	TTCTCTAATA TCTTGTTCCT GTTTTCTAA GAATGCTGG AC[A/G]TATC AGTACAACC TGACAATGA CCTTTGCATT TCTTTTAG	SEQ ID NO: 269

chr 13	103394 421	T	C	CCDC 168	p.K28 76E	0.003 43	0.004 04	6.99 E- 01	0.85[0.49- 1.46]	TTCTCCAGCT TTGGCTGTG GAAGAATGC ATGTCCTGT CTT[T/C]TGG CTTGTCTTTC TCCATTTTAA CTTCTGTAA GCTTTTAA	SEQ ID NO: 270
chr 13	103394 544	G	A	CCDC 168	p.Q28 35X	0.001 72	0.001 05	2.15 E- 01	1.63[0.74- 3.57]	ACTCGATGT ACTGCATTTT TACTCAGCT GGAATGACT TCT[G/A]CTG CTGGATGTT ACCTCTCAG TTCTTTTAA TTGCTTGCA	SEQ ID NO: 271
chr 13	103395 359	T	G	CCDC 168	p.K25 63T	0.002 94	0.003 69	5.88 E- 01	0.8[0 .44- 1.43]	TTTGTTTTTT TCTATTTTAA CATTTTTTTC TGAATTCCC T[T/G]TGTA ATCTGACTTT TTGAGAAAA AAGTTTCTC CCAAAAG	SEQ ID NO: 272
chr 13	103395 425	C	T	CCDC 168	p.R254 1H	0.001 72	0.001 34	5.07 E- 01	1.28[0.59- 2.77]	AGTTTCTCCC AAAAGCACA TCCTCTGATT TACCAAGAT GA[C/T]GATC CTTTCTAAG ATATGTGTTT GCCATGAAG TTTTCTGC	SEQ ID NO: 273
chr 13	103395 789	G	C	CCDC 168	p.L242 0V	0.001 23	0.001 28	1.00 E+0 0	0.96[0.39- 2.38]	TGCCACATT GCTTTCAGTT TGGTTTTTAA ATTGGATTC AA[G/C]TTTC TTCCTATGTT TTGTAGTAA ACTGCCAC TGATTTTA	SEQ ID NO: 274
chr 13	103396 163	T	C	CCDC 168	p.K22 95R	0.000 25	0.000 11	3.90 E- 01	2.28[0.27- 18.9 4]	CTGTGAAAT TGACGACTT CTTTTCCTTC ATAGTTAAA CAT[T/C]TGG CATTGAATA TAATTTCTTT TTCTGATAA CTGTGCTGT	SEQ ID NO: 275

chr 13	103396 628	C	T	CCDC 168	p.R214 0Q	0.003 68	0.005 37	1.77 E- 01	0.68[0.41- 1.15]	ACTCATACT TTTCTTGCCT ATAAACTCT AATGTATAG CTC[C/T]GGC TTTCATATTC AGATGACAT GAGGCTGGA GAAATCTAA	SEQ ID NO: 276
chr 13	103397 030	C	T	CCDC 168	p.R200 6H	0.000 49	0.000 14	1.43 E- 01	3.46[0.73- 16.3]	TTTGCAAGG GTCAGGATC TTTCATTTGA TGTGTACTG AAA[C/T]GGA GGTGTGAC TATAGCATG GAACTGATT CTGTTAACA T	SEQ ID NO: 277
chr 13	103397 280	C	T	CCDC 168	p.D19 23N	0.000 74	0.000 53	4.85 E- 01	1.39[0.42- 4.55]	CCTTTACCTG AATTGTGCT GTTCCCCCA TACATTTCCCT AT[C/T]AGTT GGTACACCA CGTTTTATTG CACCAGTTA AAACTTCA	SEQ ID NO: 278
chr 13	103397 387	T	G	CCDC 168	p.Q18 87P	0.021 08	0.026 03	5.77 E- 02	0.81[0.65- 1]	AGGAAGAAG TTTTGAATTT ACTGTACAT ATTGTGCCA TTT[T/G]GGG TCTGGAGGC ATTTCTTTGT CTCCTCTCTT TGTATTGG	SEQ ID NO: 279
chr 13	103398 023	G	A	CCDC 168	p.A16 75V	0.000 25	0.000 00	6.79 E- 02	Inf[NaN -Inf]	TTTAGGTGT AGATAAAGC AGGCATGCA GGAACCAAA AATC[G/A]CT GTCTCTTTCT TTTCAGTAC CACCAGCCT GTTCCCTTTG	SEQ ID NO: 280
chr 13	103398 261	T	C	CCDC 168	p.T159 6A	0.001 96	0.001 21	1.74 E- 01	1.62[0.78- 3.37]	GTTTGTGTA AAATGTGTT TGTGGTTGT ACCTGAATA TTTG[T/C]AC TTCCTGGTTG GTTCAGTTC CTCATCTGA TTTGACAAG C	SEQ ID NO: 281

chr 13	103398 339	C	T	CCDC 168	p.D15 70N	0.000 25	0.000 37	1.00 E+0 0	0.66[0.09- 4.91]	AGCTCATTA TCCTTCTGAT ATGCATTGA GTATTAAGC CAT[C/T]GCT GTTCTCCAG AGCCTGTAA AGCTTTGGG AGGTGGAAT C	SEQ ID NO: 282
chr 13	103398 453	C	A	CCDC 168	p.G15 32C	0.000 49	0.000 05	3.93 E- 02	9.28[1.55- 55.5 3]	GTTTCGTTG GCTTTTTGTA GTTCTTCAG CTTCTAAAG GAC[C/A]CAT TTGGAGACT AGTCTCTAA AGTAGTTTG TTCAAAACC T	SEQ ID NO: 283
chr 13	103399 313	G	A	CCDC 168	p.T124 5I	0.010 05	0.011 47	4.45 E- 01	0.88[0.64- 1.2]	AGATAGTTC CATTATGGG AGAAACAAC AGACTCAAT AATA[G/A]TT TCTGTGAAT GGGATTGGT TGATGCATT TCTTTCTCTG T	SEQ ID NO: 284
chr 13	103399 553	A	G	CCDC 168	p.I116 5T	0.000 49	0.000 99	4.36 E- 01	0.5[0 .12- 2.04]	TTCTTCCCTT TCAATTTGG GATTCCTCTT GGACTAGCT TG[A/G]TATG ACTGTGATT CTCTGCATTT AATCTGCTA TACATTCT	SEQ ID NO: 285
chr 13	103399 573	A	T	CCDC 168	p.N11 58K	0.000 98	0.000 34	6.72 E- 02	2.89[0.98- 8.48]	ATTCTCTTG GACTAGCTT GATATGACT GTGATTCTCT GC[A/T]TTTA ATCTGCTAT ACATTCTAG TATTAGGCA AAATAGACA	SEQ ID NO: 286
chr 13	103399 761	G	T	CCDC 168	p.P109 6T	0.006 37	0.007 35	5.66 E- 01	0.87[0.58- 1.29]	GTACCACAT ATATTAATA TAAGGCATC AGTGAGATT GCTG[G/T]CT TCTTTACTTT CATAATTAC ATATTTGAC ACTGAGTAC A	SEQ ID NO: 287

chr 13	103399 848	A	G	CCDC 168	p.Y10 67H	0.000 25	0.000 04	1.89 E- 01	6.91[0.63- 76.1 9]	GTTTCTGAT AATTTTTTTT TAATTTTCCTG CCTTTTAAA AT[A/G]TGGT AAAGTAAGC AAGTGGTTA TTGAAAGAC CCCAGGGCA	SEQ ID NO: 288
chr 13	103399 943	G	A	CCDC 168	p.T103 5M	0.000 25	0.000 07	2.94 E- 01	3.47[0.39- 31.0 6]	TCTTTTACAA TCTTCCTTTT CTTCTGCAA TATGACTAT CC[G/A]TTGT CTTTTGGAG GTTTCCACC AAATGGGAC ACTATACTC	SEQ ID NO: 289
chr 13	103400 048	T	A	CCDC 168	p.D10 00V	0.000 49	0.000 11	9.71 E- 02	4.61[0.93- 22.8 4]	AACTGGCAA GTTCTCTGG CATTGTAAG TGGATTCTTT GGA[T/A]CTC CGGCACTCT CTCTGTCTGT AGGTCTATC TGTGCTTTG	SEQ ID NO: 290
chr 13	103400 198	T	G	CCDC 168	p.K95 0T	0.001 47	0.000 21	8.40 E- 04	6.95[2.61- 18.5 3]	AAGAGTTTG TGGTTGGAC TTCTTGCTCT TTATTTGGG GCT[T/G]TAC TACTTCCTG AACTGATCT GTTCCATTG GAATTTGAC	SEQ ID NO: 291
chr 13	103400 532	C	G	CCDC 168	p.D83 9H	0.000 98	0.000 55	2.95 E- 01	1.78[0.63- 5.05]	AGTTGAGAA ATGGTAGTG TAAGTGGCA CTGTGAAAT GCAT[C/G]AG ACGTTTCTTT ATCTTGATG CATATTTGTT ATGTTACTT	SEQ ID NO: 292
chr 13	103400 781	C	A	CCDC 168	p.D75 6Y	0.000 25	0.000 00	6.77 E- 02	Inf[NaN -Inf]	AAACCGACA TTTGACAAC TCCAGAACA AGTTCCAAA AAAT[C/A]TT TTTGTTTCTG TGTATTTTCC CTTGGAAG CACCTTTGC	SEQ ID NO: 293

chr 13	103400 792	T	C	CCDC 168	p.Q75 2R	0.000 25	0.000 07	2.95 E- 01	3.45[0.39- 30.8 4]	TGACAACTC CAGAACAAG TTCCAAAAA ATCTTTTGT TTC[T/C]GTG TATTTTCCCT TGGAAAGCA CCTTTGCGTT TTTGGTGT	SEQ ID NO: 294
chr 13	103400 825	T	A	CCDC 168	p.K74 1I	0.000 25	0.000 07	2.95 E- 01	3.45[0.39- 30.9 1]	TTGTTTCTGT GTATTTTCCC TTGGAAAGC ACCTTTGCG TT[T/A]TTGG TGTAAGTGT TGGTAACTC CTCTCCATT GAAAGTTG	SEQ ID NO: 295
chr 13	103400 847	C	A	CCDC 168	p.E734 X	0.000 74	0.000 34	1.82 E- 01	2.18[0.65- 7.38]	GGAAAGCAC CTTTGCGTTT TTGGTGTAC TGGTTGGTA ACT[C/A]CTC TCCATTTGA AAGTTGAAG ATGGGAATT TTCTGAACTT	SEQ ID NO: 296
chr 13	103401 291	C	G	CCDC 168	p.E586 Q	0.000 25	0.000 07	2.96 E- 01	3.43[0.38- 30.7 1]	ATTCTGTCT CCTCAAGAG GACCTGCAT AATTGATTTT CT[C/G]TGTA TCTGGTGAC TTATTTTGCT TCTGCAGAA AATGTCCA	SEQ ID NO: 297
chr 13	103401 480	T	C	CCDC 168	p.N52 3D	0.000 98	0.001 54	5.28 E- 01	0.64[0.23- 1.74]	ATATCTTTCC TTTCATGTA ATTCTTTCTT CTCAGTGTT AT[T/C]CTTG CATCCTAAC TCATTCCTAT TTTTTAAAGT GTGACAT	SEQ ID NO: 298
chr 13	103401 929	A	G	CCDC 168	p.V37 3A	0.001 47	0.001 45	8.33 E- 01	1.01[0.44- 2.33]	CAGGCCCTT TACTGAATA TTTTGCCTCA ACAATTGAT GGA[A/G]CTT CAACAAAAT GTTGGTTCCT ATCCAGATC TTGGGACTG	SEQ ID NO: 299

chr 13	103402 542	A	G	CCDC 168	p.Y16 9H	0.000 25	0.000 21	5.95 E- 01	1.16[0.15- 8.89]	TGCTCTGTAT GGCTTAGAC ACGTTTCCTC TACTTCTGA AT[A/G]AAAC AATGGCAAA GATGAGCTG ATTCCATTG AAGATGGC	SEQ ID NO: 300
chr 13	103402 547	A	G	CCDC 168	p.L167 S	0.000 25	0.000 30	1.00 E+0 0	0.82[0.11- 6.13]	TGTATGGCT TAGACACGT TTCCTCTACT TCTGAATAA AAC[A/G]ATG GCAAAGATG AGCTGATTC CATTTGAAG ATGGCACAT G	SEQ ID NO: 301
chr 13	103402 638	A	G	CCDC 168	p.W13 7R	0.000 25	0.000 80	3.71 E- 01	0.31[0.04- 2.22]	GAGGGACTT ACTTGATCTT CACTTTCCT AGTACCTGA CC[A/G]TAGT ATTTACCGT GAGAATAAA ATTCTATCTT CAAAGTTA	SEQ ID NO: 302
chr 13	103411 167	G	A	CCDC 168	p.A39 V	0.000 49	0.000 04	2.46 E- 02	13.9 1[1.9 6- 98.8 1]	TATCTCAAA AATAATTCC TAGTAAAT TATAAAGAA AATT[G/A]CC ACCCAATCA TTTTGAATA ATCCAGGAC TCTAGAAAG TC	SEQ ID NO: 303
chr 13	103514 444	C	T	BIVM- ERCC 5	p.H76 9H	0.007 84	0.005 31	3.78 E- 02	1.48[1.04- 2.12]	AAGTGGATT CAGAGTCTC TTCCTTCTTC CAGCAAAAT GCA[C/T]GGC ATGTCTTTTG ACGTGAAGT CATCTCCAT GTGAAAAAC	SEQ ID NO: 304
chr 13	103701 648	A	G	SLC10 A2	p.F304 L	0.005 64	0.003 50	3.18 E- 02	1.61[1.06- 2.46]	ATCATGAAA TGGGATTGG CATGATTCC TTACATCCT AAGA[A/G]T ATTGCGGCA AAGGCGAGC TGAAAAATG CTGTAGATG AGC	SEQ ID NO: 305

chr 13	110864 264	C	T	<i>COL4 A1</i>	p.E131 E	0.010 29	0.006 37	3.86 E- 03	1.62[1.19- 2.22]	CAGCGAAAC CAGGCAAGC CAGGAGGCC CGAGCGGCC CTCT[C/T]TC CCCCTGGGG AGACAGCAG AGCATCATT CATACGCAC TG	SEQ ID NO: 306
chr 13	113201 864	C	T	<i>TUBG CP3</i>	p.R413 H	0.011 52	0.000 80	1.08 E- 30	14.5[9.66- 21.7 5]	GGGAAAGAC GCGCGTGGG AAAGACGTG CATGGGAAA GTCG[C/T]GC GTGGGAAAG TCGCGCGTG GGAAAGTCG CGCGTGGGA AA	SEQ ID NO: 307
chr 13	114175 013	G	A	<i>TMCO 3</i>	p.P436 P	0.012 01	0.008 69	3.24 E- 02	1.39[1.04- 1.85]	CGCAGGACG TGCAGCTCG GGCTCTTCA TGGCCGTCA TGCC[G/A]AC TCTCATACA GGCGGGCGC CAGTGCATC TTCTAGGTA AA	SEQ ID NO: 308
chr 14	212161 36	G	A	<i>EDDM 3A</i>	p.V13 3I	0.007 35	0.004 56	1.38 E- 02	1.62[1.12- 2.34]	CTTCAGCTA CATTGAATT CCATTGTGG CGTAGATGG ATAT[G/A]TT GATAACATA GAAGACCTG AGGATTATA GAACCTATC AG	SEQ ID NO: 309
chr 14	233538 96	G	A	<i>REM2</i>	p.T39 T	0.009 07	0.004 52	1.55 E- 04	2.02[1.44- 2.82]	TTTCTTTGCC CTCCCATTTT ATTTTAGAA GCAGATGCC AC[G/A]CTAC TAAAGAAGT CAGAGAAAC TGTTGGCAG AGTTGGACC	SEQ ID NO: 310
chr 14	244643 24	C	T	<i>DHRS 4L2</i>	p.T29 T	0.008 33	0.001 13	3.09 E- 17	7.44[5.09- 10.8 9]	CTGCTGTCA ACCCTTTCTT TGGAAGCCT AATGGATGT CAC[C/T]GAG GAGGTGTGG GACAAGGTG AGAGGGGAT TAAAGAAGC G	SEQ ID NO: 311

chr 14	247723 73	C	T	<i>NOP9</i>	p.R413 C	0.007 482	0.004 658	3.19 E- 02	1.61[1.01- 2.45]	GGGCCACCC AGGGGTAGT CATTGCCCT GGTGGGGGC CTGT[C/T]GC AGAGTTGGG GCCTACCAA GCCAAGGTC CTACAGCTC TT	SEQ ID NO: 312
chr 14	449751 03	G	A	<i>FSCB</i>	p.P363 L	0.010 29	0.000 00	7.71 E- 62	Inf	AGGAGACTT TTCAGCTGG TGGAGGCAG AATTCAGC AGGA[G/A]G CTCTTCTGA AGGGGACTC TTCAGCTGA TGGAGGCAG AAT	SEQ ID NO: 313
chr 14	449751 15	G	A	<i>FSCB</i>	p.P359 L	0.024 51	0.000 01	1.52 E- 144	2806 .41[3 91.3 8- 2012 3.7]	AGCTGGTGG AGGCAGAAT TTCAGCAGG AGGCTCTTC TGAA[G/A]G GGACTCTTC AGCTGATGG AGGCAGAAT TTCAGCCAG AAG	SEQ ID NO: 314
chr 14	505810 11	A	C	<i>VCPK MT</i>	p.Y18 8D	0.010 05	0.006 79	2.01 E- 02	1.48[1.08- 2.04]	ACTACAAAG ATAATAGAG TACTTAATA CTTACCTCA AAAT[A/C]TT TTTTCTCAAT TTCTGGATTT TTCCCCATTG TTCGTTGT	SEQ ID NO: 315
chr 14	524954 81	C	T	<i>NID2</i>	p.R830 Q	0.005 15	0.003 20	4.83 E- 02	1.61[1.04- 2.51]	GATGCAAGT ATGCCGGTC ATCTGCAAA CTCATAACC ACTC[C/T]GG CACTCACAC CTGTAGCTT CCAGGCAAG TTGATACAT AC	SEQ ID NO: 316
chr 14	524963 99	T	C	<i>NID2</i>	p.D75 6G	0.011 03	0.007 71	2.33 E- 02	1.44[1.06- 1.94]	CATGTGGCT CCCATCATA GCAAGGATT CCCCGGAGT GGGG[T/C]CT GAATCCTCT GCATGAGTA GAGGGGAAA TAAAAGCAC AA	SEQ ID NO: 317

chr 14	525096 01	C	T	<i>NID2</i>	p.R493 K	0.011 76	0.008 72	4.93 E- 02	1.35[1.01- 1.81]	AGTGGCATA GTCCGTGCA GAAGGCATG CCGGGAGCA TTGT[C/T]TG TGGTTGTGTT CACAGGTTT CCTTGTTGG CAGCATTAT A	SEQ ID NO: 318
chr 14	609218 36	T	G	<i>Cl4orf 39</i>	p.E462 D	0.006 86	0.004 52	4.35 E- 02	1.52[1.04- 2.23]	TAAGAAAAG AAAGTCCAG GGGATTCTT TTTCTGTTTG AAC[T/G]TCA GGTACTGCA TTTCTATTTC TGTTACTGA GAAATAAGA	SEQ ID NO: 319
chr 14	622448 54	C	T	<i>SNAP C1</i>	p.T253 M	0.005 21	0.003 03	4.52 E- 02	1.72[0.97- 2.84]	AATGATGGA GAAGAAAAA ATGGAAGGA AATTCACAA GAAA[C/T]GG AGGTCAGAA AACTTTGCA ATTCATATT ATGTGTGGC TG	SEQ ID NO: 320
chr 14	695216 37	C	T	<i>DCAF 5</i>	p.R589 H	0.006 86	0.003 88	7.18 E- 03	1.78[1.21- 2.6]	TGGGGCACT GGGCTTGTC TTCTCGGGTT GTCTTCTGTC GG[C/T]GCCG CATGGCATT CCGCTGCCA GGTAGAGGC TCGGCGTTC	SEQ ID NO: 321
chr 14	704189 85	C	T	<i>SMOC 1</i>	p.P77L	0.005 39	0.003 36	3.93 E- 02	1.61[1.04- 2.47]	GAGTCCATG TGTGAGTAC CAGCGAGCC AAGTGCCGA GACC[C/T]GA CCCTGGGCG TGGTGCATC GAGGTAGAT GCAAAGGTG AG	SEQ ID NO: 322
chr 14	751512 52	C	T	<i>AREL1</i>	p.V50 M	0.007 157	0.004 135	1.58 E- 02	1.74[1.08- 2.67]	GAGACTTTG CAAGACCGG GGATCCAGG TAATTTCCCC GCA[C/T]GTA GTCATAAAT AGTCCGGTC CCCTCGGCG CTCGCGGTC C	SEQ ID NO: 323

chr 14	860881 77	C	A	<i>FLRT2</i>	p.L107 I	0.006 13	0.003 82	2.85 E- 02	1.61[1.07- 2.41]	CTACCTGTA TGGCAACCA ACTGGACGA ATCCCCCAT GAAC[C/A]TT CCCAAGAAT GTCAGAGTT CTCCATTGTC AGGAAAACA A	SEQ ID NO: 324
chr 14	888929 32	C	T	<i>SPATA 7</i>	p.R211 R	0.005 39	0.003 41	4.13 E- 02	1.59[1.03- 2.44]	CTGAACTCT TTTCTAACCA AACAATTGC CATTCACCTC CTCG[C/T]AC TTTAAAAAC AGAAGCAAA ATCTTTCCTG TCACAGTAT C	SEQ ID NO: 325
chr 14	891108 01	T	C	<i>EML5</i>	p.V13 61V	0.009 56	0.006 63	3.12 E- 02	1.45[1.05- 2]	AGTGAGTTT TCCTTACCTC TATAGGTCT CTTTTTCTTG CC[T/C]ACAT TGTTTGTCTG GAGTTTCTCT GGCTGTGGT GGGGCCC	SEQ ID NO: 326
chr 14	101004 386	A	G	<i>BEGAI N</i>	p.F568 L	0.005 88	0.000 01	2.13 E- 33	607. 53[8 2.17- 4491 .9]	CTGTCCTTGC GGCTCAGCC CCGAGCCAC CAGTCCGCG GAA[A/G]GG CCTGCTGGG GGCTGAGGC GGGCGGCAG GATGCATTT CC	SEQ ID NO: 327
chr 14	103593 342	T	A	<i>TNFAI P2</i>	p.V79 E	0.009 80	0.000 00	1.74 E- 07	Inf	GTGGGCTGG GGCCGGGGC TGACGCGGC TTTCCCGGC GCAG[T/A]GG AGGAGCTGA AGGCGGCGC TGGAGCGCG GGCAGCTGG AG	SEQ ID NO: 328
chr 14	105415 242	C	T	<i>AHNA K2</i>	p.K21 82K	0.011 27	0.000 26	4.90 E- 47	43.0 7[27. 03- 68.6 2]	GGTCCCCCT GCATGGAGG GGAGACTCA TGTCGGCCT CCAC[C/T]TT GGGTGGAGA CACATCCAC CGAGGCCTC GATGGACTT GC	SEQ ID NO: 329

chr 14	105415 333	T	C	<i>AHNA</i> <i>K2</i>	p.K21 52R	0.019 61	0.000 94	7.30 E- 63	21.3 6[15. 62- 29.1 9]	CACCCCAAA CGACGGCAT CTTGAACCTT GGGCATTTT GAAC[T/C]TG CTGTCTTTGG TAGTCAGGT CCTTGTTGG CCAGGGTCA G	SEQ ID NO: 330
chr 14	105415 752	A	T	<i>AHNA</i> <i>K2</i>	p.D20 12E	0.005 64	0.003 25	1.74 E- 02	1.74[1.14- 2.65]	AGGGGAGAC TCACGTCGG CCTCCACCTT GGGTGCAGG CAC[A/T]TCC ACCGAGGCC TCGATGGAC CTCCCTGGG GCCGATACC C	SEQ ID NO: 331
chr 14	105418 170	G	C	<i>AHNA</i> <i>K2</i>	p.L120 6L	0.008 82	0.001 82	3.11 E- 13	4.88[3.41- 6.97]	GGTCAGCGG AAGGGGGCT GAATGCTGA GGTCAGTGG TCTT[G/C]AG GTCCCCCTG CATGGAGGG GAGACTCAC GTCGGCCTC CA	SEQ ID NO: 332
chr 15	315155 19	G	A	<i>LOC28</i> <i>3710</i>	p.L124 F	0.011 52	0.000 00	2.03 E- 51	Inf	TGGGATCAG TGCGGCCTG TCGTCTGCT GTTGTCATG TGGA[G/A]CT CAGCAAACG GTGGGAGTC CTAGGGGAC AACATACAC AG	SEQ ID NO: 333
chr 15	387768 33	T	A	<i>FAM9</i> <i>8B</i>	p.G42 5G	0.007 35	0.000 12	7.32 E- 27	61.2 9[23. 77- 158. 06]	ATCCATATG GAGGAGGTG GTGGTGTTG GTGGTGTTG GTGG[T/A]GG AGGAGGTGG ATATAGAAG ATACTAAAA ACTATAAAA AT	SEQ ID NO: 334
chr 15	418623 46	G	A	<i>TYRO3</i>	p.T458 T	0.008 33	0.005 24	1.15 E- 02	1.6[.13- 2.26]	CCCTGGCCC TCATCCTGCT TCGAAAGAG ACGGAAGA GAC[G/A]CG GTTTGGGTA AGGGGATGG GGATGTGGA GGGAGAGGC AG	SEQ ID NO: 335

chr 15	436533 05	C	T	<i>ZSCAN</i> 29	p.R842 Q	0.005 39	0.003 41	4.15 E- 02	1.58[1.03- 2.44]	AGGGGCTTA CTTGGGAGC TGACTGTGT CAGAAGCTT TTCC[C/T]GT GCATGGATT TCTCCGTGCT TATTAAGGG CAGAGCTTT T	SEQ ID NO: 336
chr 15	484704 30	G	T	<i>MYEF</i> 2	p.A2E	0.026 23	0.000 00	2.09 E- 39	Inf	GCCACCAGT GGCCCCGGG CACCTCGGC CTTGTTGGC GTCC[G/T]CC ATCCCGCCG CCGCTGCCT CCGCCTCGG CCGCCTGAG CT	SEQ ID NO: 337
chr 15	525107 96	A	G	<i>MYO5</i> C	p.L129 2L	0.005 15	0.003 08	3.12 E- 02	1.67[1.08- 2.6]	TTACACTTG ACTTCACTTT CAGTTTCAA ATTGTTTCTT CA[A/G]GTGG TCACTGGCC TCCTGCATTT CTTGAATCTT ATCAATC	SEQ ID NO: 338
chr 15	651578 74	G	A	<i>PLEK</i> <i>HO2</i>	p.S420 S	0.010 78	0.007 36	1.60 E- 02	1.47[1.08- 1.99]	AACGGCTAT ATCGGGCCC AGCTGGAGG TGAAGGTGG CCTC[G/A]GA ACAGACGGA GAAACTGTT GAACAAGGT GCTGGGCAG TG	SEQ ID NO: 339
chr 15	720235 02	G	A	<i>THSD4</i>	p.V52 6M	0.005 53	0.003 02	2.01 E- 02	1.83[1.05- 2.99]	GATACACCA GCAGCCAAA CCCAGGCGT GCACTACGA GTAC[G/A]TG ATCATGGGG ACCAACGCC ATCAGCCCC CAGGTGCCA CC	SEQ ID NO: 340
chr 15	721922 05	C	G	<i>MYO9</i> A	p.R109 8P	0.005 21	0.002 75	2.23 E- 02	1.89[1.07- 3.13]	GTAATCTCT CCATTTCTGC TGGATAACG ATGGCTGCA GCC[C/G]GTA ACTCCAAGT ACCGCTGCC TCTCTAAGT GAGCACGCC A	SEQ ID NO: 341

chr 15	725021 15	T	C	<i>PKM</i>	p.N15 5S	0.005 64	0.003 52	3.24 E- 02	1.61[1.05- 2.45]	CACCACCTT GCAGATGTT CTTGTAGTC CAGCCACAG GATG[T/C]TC TCGTCACAC TTTTCCATGT AGGCGTTAT CCAGCGTGA T	SEQ ID NO: 342
chr 15	725136 12	T	A	<i>PKM</i>	p.T36S	0.017 16	0.011 53	2.57 E- 03	1.51[.16- 1.93]	CTTGGCCTC ACTAGCAAA GACCGCTCA GAGCTGAAT ACGG[T/A]GT GCCCTGGAG AGCTGCACA AGGATTAAG GAAAAAGCT GA	SEQ ID NO: 343
chr 15	759815 11	C	A	<i>CSPG4</i>	p.G63 2V	0.005 15	0.000 00	7.77 E- 31	Inf	TCCATCGCT GACCCGGAA CGTCAAGTC CTGTGCAGG ACCA[C/A]CG CGGTGGACA TAGACTAGG CTGCCGGCC TCCAACCTCC CG	SEQ ID NO: 344
chr 15	759820 53	A	G	<i>CSPG4</i>	p.H45 1H	0.006 86	0.004 53	4.39 E- 02	1.52[1.04- 2.23]	TGCGCAGCT CAGCCTCCA TCAGGTCCA GCGTGGGCT GCAC[A/G]TG CCTCCACTC AAGCCAGGC TGTGCCCCC CTCGGCCAC CA	SEQ ID NO: 345
chr 15	784613 24	C	T	<i>IDH3A</i>	p.R360 C	0.006 86	0.003 96	7.88 E- 03	1.74[1.18- 2.55]	AGGCAATGC AAAATGCTC AGACTTCAC AGAGGAAAT CTGT[C/T]GC CGAGTAAAA GATTTAGAT TAACACTTC TACAACTGG CA	SEQ ID NO: 346
chr 15	790589 44	A	T	<i>ADAM TS7</i>	p.A11 03A	0.007 89	0.000 80	2.49 E- 11	10.5 6[6.0 4- 18.4 9]	GAGGCTCTG TGGCAGGCA CGGGGCTAC CCGTGGAGG GCGC[A/T]GC AGGATGGCT GTGTGGTGG GGGTGTCCG GTCCCCTGT CC	SEQ ID NO: 347

chr 15	796037 60	G	A	<i>TMED</i> 3	TMED 3(NM _0073 64:exo n1:c.1 68+1G >A)	0.006 86	0.004 47	3.23 E- 02	1.54[1.05- 2.26]	GGAGGTGGA GCAGGGCGT GAAGTTCTC CCTGGATTA CCAG[G/A]TG AGGCCGGGC GCCCCGGCAG CGCTCCCTTC TCCCTCCACT	SEQ ID NO: 348
chr 15	891697 38	G	A	<i>AEN</i>	p.G10 0R	0.006 62	0.004 20	2.72 E- 02	1.58[1.07- 2.33]	TGGATCTGG CAGTGCCCC ATGCAGCAG AAGGCCTGC TCCC[G/A]GG AAAGCCTCA GGGCCCTTG CCCAGCAAG TGTGTGGCT AT	SEQ ID NO: 349
chr 15	102346 082	C	T	<i>OR4F6</i>	p.R54 C	0.005 88	0.003 63	2.56 E- 02	1.62[1.07- 2.45]	GGGAAATCT CCTCATTGT GCTAACTGT GACCTCTGA CCCT[C/T]GT TTACAGTCC CCCATGTAC TTCCTGCTG GCCAACCTT TC	SEQ ID NO: 350
chr 16	315001	C	T	<i>ITFG3</i>	p.R547 W	0.005 39	0.003 35	3.86 E- 02	1.62[1.05- 2.49]	AGACAGTGA CCAAGCCAT CAGGGACCG GTTCTCCCG GCTG[C/T]GG TACCAGAGT GAGGCGTAG AGGCACGCC AGCCAGAGC CT	SEQ ID NO: 351
chr 16	863362	C	G	<i>PRR25</i>	p.P237 R	0.020 34	0.000 00	1.86 E- 108	Inf	GACATCCCC TCTGCTATTG CTGCGGGAC CGGCAAGGA CGC[C/G]GGA CCGACACGG CCTCCCCAT CCCTGGGTC CACCCCGAC T	SEQ ID NO: 352
chr 16	225857 5	G	A	<i>MLST8</i>	p.G27 5S	0.005 39	0.002 90	7.86 E- 03	1.86[1.21- 2.88]	GAGCGGCAA CCCCGGGGA GTCCTCCCG CGGCTGGAT GTGG[G/A]GC TGCGCCTTCT CGGGGGACT CCCAGTACA TCGTCACTG G	SEQ ID NO: 353

chr 16	228764 9	A	C	<i>DNAS EIL2</i>	p.D19 7A	0.005 64	0.003 51	3.19 E- 02	1.61[1.06- 2.46]	TACGACGTG TACCTGGAC GTGATCGAC AAGTGGGGC ACCG[A/C]CG TAAGCCCAC CCCTCGGTC CCGGGGTCC CTGCAGGCG CG	SEQ ID NO: 354
chr 16	236959 2	C	T	<i>ABCA3</i>	p.R288 K	0.014 46	0.009 32	1.71 E- 03	1.56[1.2- 2.03]	GAGTGTGG GGAGCCAAA GCGGGCAGT CACCTTCAG CCTC[C/T]TT TCCTTCTCCT GCACGACAG CACGGGCAA TGGTGAGCG C	SEQ ID NO: 355
chr 16	284851 5	G	T	<i>PRSS4 1</i>	p.A10 A	0.016 67	0.000 00	6.97 E- 73	Inf	GAGAGGAGG CCATGGGCG CGCGCGGGG CGCTGCTGC TGGC[G/T]CT GCTGCTGGC TCGGGCTGG ACTCGGGAA GCCGGGTGA GC	SEQ ID NO: 356
chr 16	363905 8	C	T	<i>SLX4</i>	p.P152 7P	0.005 15	0.002 53	4.17 E- 03	2.04[1.31- 3.18]	CTTCGGGCT TCTGAGCTC CACCAGCGC TTGGCATCT GGGC[C/T]GG AGGAGGGGT CTCTGGAGG CCTCTGCTCT TCCCCGTCC C	SEQ ID NO: 357
chr 16	363937 8	T	A	<i>SLX4</i>	p.I142 1F	0.011 76	0.001 07	8.31 E- 30	11.0 9[7.9 - 15.5 6]	GAGAGGGGC TCCATGTGC CAGCAGCAG TCGTCAATT GGAA[T/A]TG GGGGGTCAC TGTCCAGTG GGGGGCTTC TGTTGGCCT GA	SEQ ID NO: 358
chr 16	364081 5	C	G	<i>SLX4</i>	p.E942 Q	0.005 39	0.002 53	1.53 E- 03	2.14[1.39- 3.31]	TGGCCAAGC GCCTCCTCT GGCGCCTCC TGCTCAGGG GCCT[C/G]TG CTCCCCGTG CCCCTGAGT GCTGGCCCT GGGGTGGCG GG	SEQ ID NO: 359

chr 16	370719 1	G	A	<i>DNAS E1</i>	p.V18 5I	0.008 33	0.004 95	4.80 E- 03	1.69[1.19- 2.39]	CGCATGTGCC CAGGGCCAC AGGCAGCGT TTCCTGGTA GGAC[G/A]TC ATGTTGATG GGCGACTTC AATGCGGGC TGCAGCTAT GT	SEQ ID NO: 360
chr 16	373608 5	C	T	<i>TRAP1</i>	p.R128 H	0.005 15	0.002 70	9.00 E- 03	1.91[1.22- 2.97]	CATTTCTGG CAGTGCTTG GCCGTCAGA CACCAGTTT GTGA[C/T]GC AGTTTTTCCA AGGCATCGC TGGCATTGG AGATCAGCT C	SEQ ID NO: 361
chr 16	491077 7	A	G	<i>UBN1</i>	p.R262 G	0.024 02	0.000 01	1.26 E- 141	2748 .75[83.2 6- 1971 4.18]	GCTAAAGAA ATTTTCAGAA AGAGAAAGA GGCTCAGAA AAAA[A/G]G GGAGGAGGA GCATAAGCC TGTTGCGGT CCCATCAGC GGA	SEQ ID NO: 362
chr 16	209965 25	G	A	<i>DNAH 3</i>	p.D25 13D	0.006 62	0.004 39	4.15 E- 02	1.51[1.03- 2.23]	CGATGTCAG CCTTCTCGTC AGCAGGGAA GATGTTAGG CAC[G/A]TCA CCTGTGTTT AGAAGCATG TTGATGTCCT CCACGAATG	SEQ ID NO: 363
chr 16	209965 88	G	A	<i>DNAH 3</i>	p.A24 92A	0.007 11	0.004 23	1.02 E- 02	1.68[1.16- 2.45]	TGATGTCCT CCACGAATG ATTCATCCTT GATCTGGTT GTC[G/A]GCG AAGAGGAAC ACGGTGCTC TTGGTGGCC ACACCGACC T	SEQ ID NO: 364
chr 16	217476 33	A	C	<i>OTOA</i>	p.T706 P	0.007 35	0.000 10	5.19 E- 35	75.1 3[37. 62- 150. 01]	CCTTCTGCA AGCAGCTTC CAAGATGGC CAGGACCCT GCCC[A/C]CT AAAGAATTC CTCTGGGCT GTCTTTCAGT CTGTTCGGA A	SEQ ID NO: 365

chr 16	217476 39	G	T	<i>OTOA</i>	p.E708 X	0.007 35	0.000 02	1.12 E- 41	413. 18[9 8.71- 1729 .48]	GCAAGCAGC TTCCAAGAT GGCCAGGAC CCTGCCAC TAAA[G/T]AA TTCCTCTGG GCTGTCTTTC AGTCTGTTC GGAACAGCA G	SEQ ID NO: 366
chr 16	217476 62	G	A	<i>OTOA</i>	p.Q71 5Q	0.007 35	0.000 05	5.17 E- 38	136. 27[5 6.69- 327. 58]	GGACCCTGC CCACTAAAG AATTCCTCT GGGCTGTCT TTCA[G/A]TC TGTTCCGAA CAGCAGTGA TAAGATCCC CAGCTATGA CC	SEQ ID NO: 367
chr 16	289438 83	C	G	<i>CD19</i>	p.P102 R	0.019 36	0.000 00	1.99 E- 114	Inf	CAACAGATG GGGGGCTTC TACCTGTGC CAGCCGGGG CCCC[C/G]CT CTGAGAAGG CCTGGCAGC CTGGCTGGA CAGTCAATG TG	SEQ ID NO: 368
chr 16	289962 27	G	C	<i>LAT</i>	p.L15F	0.017 89	0.000 00	7.89 E- 80	Inf	AGGCCACGG CTGCCAGCT GGCAGGTGG CTGTCCCCG TCTT[G/C]GG GGGGGCCAG CAGACCTT GGTGAGTGC CTGGGGTGG CT	SEQ ID NO: 369
chr 16	307932 73	C	G	<i>ZNF62</i> 9	p.Q79 2H	0.014 22	0.000 01	4.59 E- 78	1291 .24[1 78.8 1- 9324 .5]	CTGCCTCTG GAGGGGGGT CCTCGGGAT TGGGGGGTT TTTC[C/G]TG GGTGTGGGT TTCTTGGTGC CGGGTGAGG GCCACGCGG T	SEQ ID NO: 370
chr 16	307942 06	G	T	<i>ZNF62</i> 9	p.T481 T	0.022 55	0.000 00	7.33 E- 134	Inf	AGCTCTTGC CGCACTCGG GGCACTTGT AGGGCTTCT CGCC[G/T]GT GTGCGTGCG GCGGTGCTG GATAAGGTG GGAGCTGCG GA	SEQ ID NO: 371

chr 16	620552 38	G	A	<i>CDH8</i>	p.P24S	0.015 93	0.000 04	1.88 E- 89	451. 14[1 64.2 8- 1238 .87]	ACTTGAGAC TGATTCATC GGAGCCATG TAAATGCAA GGGG[G/A]A AGAGTAATC CATAATATT ATTAATGGA GTCCAGAGA TCC	SEQ ID NO: 372
chr 16	672368 72	C	T	<i>ELMO 3</i>	p.T600 M	0.006 86	0.004 65	4.80 E- 02	1.48[1.01- 2.16]	CTGATCCGC CAGCAGCGC TTGCTCCGC CTCTGTGAG GGGA[C/T]GC TCTTCCGCA AGATCAGCA GCCGGCGGC GCCAGGGTC TC	SEQ ID NO: 373
chr 16	689615 76	C	T	<i>TANG 06</i>	p.R745 C	0.008 82	0.006 09	3.26 E- 02	1.45[1.04- 2.03]	ATACCCTGA TCCGGTCAT CCAAGAACT CGCTGTTGA TCTC[C/T]GC ATCACCATC TCTACCCAT GGAGCCTTT GCCACTGAG GC	SEQ ID NO: 374
chr 16	705088 51	A	G	<i>FUK</i>	p.T772 A	0.009 31	0.006 05	1.39 E- 02	1.54[1.11- 2.15]	TGAGCTGTG GCTGGCGGT GGGGCCTCG GCAGGATGA GATG[A/G]CT GTGAAGATA GTGTGCCGG TGCCTGGCT GACCTGCCG GA	SEQ ID NO: 375
chr 16	708947 71	C	T	<i>HYDIN</i>	p.P393 7P	0.025 98	0.000 04	4.43 E- 89	656. 67[9 1.63- 4706 .3]	GGCAGATGG GCAAGGTGC TCCGCCCTTT TGCTACCAG GAC[C/T]GGA CCTTGCTCTC CAGGTGGCA GGTTGGGAA TCCTGAGAG	SEQ ID NO: 376
chr 16	708970 62	C	T	<i>HYDIN</i>	p.R383 2H	0.005 39	0.000 00	1.11 E- 32	Inf	TGAGGTATC TTCTGAGAC CCAGCTGAA TTCCAGCTG GACA[C/T]GT CCTGAATTA ATCACATCG AACCTGCAA ATCGATCAG GG	SEQ ID NO: 377

chr 16	709350 93	C	T	<i>HYDIN</i>	p.R295 4R	0.005 88	0.002 00	9.20 E- 06	2.95[1.93- 4.51]	AGGCCACAG GCAGGAGCG TGACATTGC GGAGAAGAA CTAC[C/T]CT GGATTCTG TCTGCAGAG ACAAAAGGA AAGTTGCAA TT	SEQ ID NO: 378
chr 16	709550 79	G	A	<i>HYDIN</i>	p.I240 0I	0.017 40	0.000 00	2.60 E- 94	Inf	TCTCAGACA TTGTTTGTTC CCTAACAGA TATTTTCCTT TC[G/A]ATTG TCTCCATCTT GACATCCAC TTTGGTGAG CGGAGGAA	SEQ ID NO: 379
chr 16	709960 23	G	A	<i>HYDIN</i>	p.S193 6L	0.006 37	0.001 27	1.77 E- 09	5.05[3.17- 8.05]	CGATGTCCT CTTTGTGCTA TTGGAGGTT CCCTGATCT GAT[G/A]AG GTTATATCTT CCTCTTCTGC CAGGTAGCA AAGGATGAA	SEQ ID NO: 380
chr 16	711012 11	G	A	<i>HYDIN</i>	p.A71 3V	0.005 88	0.000 06	2.94 E- 29	93.9 2[40. 44- 218. 1]	AGAGCAAGC TGGGGAGCA ATACCTTGC TGTAATTAA GAGC[G/A]CC AGCACCTCT TCTCCGATG CCCTCCACG TCCACCACG AG	SEQ ID NO: 381
chr 16	851007 97	C	T	<i>KLAA0 513</i>	p.D40 D	0.005 15	0.001 61	1.06 E- 05	3.2[2 .03- 5.03]	CACCCCCTG TGCTGCAGG ACGGCGATG GCTCCCTGG GGGA[C/T]GG TGCATCAGA GAGTGAGAC CACTGAGTC TGCGGACAG TG	SEQ ID NO: 382
chr 16	887197 26	T	C	<i>MVD</i>	p.K36 8K	0.011 03	0.007 34	1.19 E- 02	1.51[1.11- 2.04]	CTGGGTGAG CCCCAGGCC TCACCTGAG TGACAATGA TGTA[T/C]TT GACCCACC GGGGGTCCG CTCCATGGC CAGCGCAGC CT	SEQ ID NO: 383

chr 17	168770 7	C	T	<i>SMYD</i> 4	p.V64 5I	0.006 86	0.003 53	2.00 E- 03	1.95[1.33- 2.87]	TGTAGGTCC TGTAACCGA GAGACCAGG TGGTCCCTG CTGA[C/T]GG CGGATTCTG CACAAGATC TGCTGCCAC AGCGCAGCA CG	SEQ ID NO: 384
chr 17	227571 9	G	C	<i>SGSM</i> 2	p.R530 R	0.005 88	0.000 01	8.71 E- 33	571. 47[7 7.29- 4225 .3]	TGTCCGGCGC TGGTGCACC ATAGCGTTA TCCCACCTG ACCG[G/C]CC CCCGGGGGC CTCCGCGGG CCTCACCAA GGACGTGTG GA	SEQ ID NO: 385
chr 17	319577 7	A	T	<i>OR3A1</i>	p.F34I	0.011 76	0.006 79	5.13 E- 04	1.74[1.3- 2.33]	CTGAGGTTG CCCCTGACC GTGACCAGG TAGGCAAAG AGGA[A/T]G AGCACAAAG ACAACCTGGC TGCAGCCCT GGCGCCTCC AGC	SEQ ID NO: 386
chr 17	722237 4	G	A	<i>NEUR</i> <i>L4</i>	p.L122 5F	0.006 13	0.003 57	1.18 E- 02	1.72[1.15- 2.58]	CCTGGTCCT GTTCTTCTC TCTGGCTCCT ACTCACCTT GA[G/A]ACC GTTGTGGAA GACCCACG GCCCCGAG CAGCCAGGC T	SEQ ID NO: 387
chr 17	819320 3	G	A	<i>RANG</i> <i>RF</i>	p.Q17 0Q	0.005 88	0.003 31	9.03 E- 03	1.78[1.18- 2.69]	ATCTGTAC CTGCACCCT GGAGCCTGG GTGACTTTG AACA[G/A]CT GGTGACCAG TCTGACCCTT CACGATCCT AACATCTTT G	SEQ ID NO: 388
chr 17	117846 88	C	T	<i>DNAH</i> 9	p.A35 88A	0.008 09	0.005 58	4.28 E- 02	1.45[1.02- 2.07]	TCACCGTGA CCAGGGATG GCCTGGAGG ACCAGTTGC TGGC[C/T]GC TGTGGTCAG CATGGAGAG GCCAGACTT GGAGCAGCT GA	SEQ ID NO: 389

chr 17	142048 68	C	T	<i>HS3ST</i> <i>3B1</i>	p.C11 C	0.005 39	0.002 71	3.72 E- 03	2[1.2 9- 3.1]	GGCAGCGCA TGGGGCAGC GCCTGAGTG GCGGCAGAT CTTG[C/T]CT CGATGTCCC CGGCCGGCT CCTACCGCA GCCGCCGCC GC	SEQ ID NO: 390
chr 17	171844 95	C	T	<i>COPS3</i>	p.A2A	0.008 09	0.005 56	4.22 E- 02	1.46[1.03- 2.07]	AGAGCTGTC GGACACTGT TCACGAACT GCTCCAGGG CAGA[C/T]GC CATGTTTTCC CCCGGGCGG CCCGAGCGG CGAAGGCAG C	SEQ ID NO: 391
chr 17	188746 89	C	T	<i>FAM8</i> <i>3G</i>	p.D81 9N	0.011 03	0.000 01	1.70 E- 63	1177 .01[62.2 1- 8540 .78]	TGGCTCCAG GCTGGGACA TGCTGCTAG GGGTCTTTG CGGT[C/T]CC GGGGGGCTT GAGCCCTCC GTTTAGAAT CCGATGAGG CC	SEQ ID NO: 392
chr 17	212039 61	G	A	<i>MAP2</i> <i>K3</i>	p.M90 I	0.009 56	0.004 22	9.33 E- 06	2.28[1.64- 3.16]	TGGTAGAGA AGGTGCGGC ACGCCCAGA GCGGCACCA TCAT[G/A]GC CGTGAAGGT GAGCAGGGC CTGGAGGCA GCTGGGAGG GC	SEQ ID NO: 393
chr 17	212154 98	C	G	<i>MAP2</i> <i>K3</i>	p.T273 T	0.005 88	0.002 88	1.78 E- 03	2.05[1.35- 3.11]	AGATGGCCA TCCTGCGGT TCCCTTACG AGTCCTGGG GGAC[C/G]CC GTTCCAGCA GCTGAAGCA GGTGGTGGG GGAGCCGTC CC	SEQ ID NO: 394
chr 17	213186 71	G	A	<i>KCNJ1</i> <i>8</i>	p.R6Q	0.012 04	0.002 75	2.45 E- 12	4.42[3.02- 6.32]	AGCCAGGGT CCCCCAACC CCCGGGATG ACCGCGGCC AGCC[G/A]G GCCAACCCC TACAGCATC GTGTCATCG GAGGAGGAC GGG	SEQ ID NO: 395

chr 17	213188 26	G	A	<i>KCNJ1</i> 2	p.A58 T	0.017 16	0.000 35	4.39 E- 73	49.8] 33.5 1- 74.0 1]	CCGCTTCGT CAAGAAGAA TGGCCAGTG CAACATTGA GTTC[G/A]CC AACATGGAC GAGAAGTCA CAGCGCTAC CTGGCTGAC AT	SEQ ID NO: 396
chr 17	213197 92	G	A	<i>KCNJ1</i> 2	p.E380 K	0.010 05	0.000 32	9.65 E- 39	31.5 6[20. 08- 49.6]	GTTCCTGCT GCCCAGCGC CAACTCCTT CTGCTACGA GAAC[G/A]A GCTGGCCTT CCTGAGCCG TGACGAGGA GGATGAGGC GGA	SEQ ID NO: 397
chr 17	275807 75	G	A	<i>CRYB</i> <i>A1</i>	p.G15 9S	0.006 37	0.004 16	4.80 E- 02	1.53] 1.03- 2.28]	CCCCTCCTTG CAAGCCATG GGCTGGTTC AACAACGAA GTC[G/A]GCT CCATGAAGA TACAAAGTG GGGCGTAAG TACAAAAAC A	SEQ ID NO: 398
chr 17	276138 38	T	C	<i>NUFIP</i> 2	p.T392 A	0.006 37	0.004 10	3.46 E- 02	1.56] 1.05- 2.32]	GCTGACATA GGGACCTGG GATAAGCGA CTTGATGAT TGGG[T/C]CT GAGTTTCCC CGGTAGATG ATGAAGATG ATGAAGATG AA	SEQ ID NO: 399
chr 17	368296 76	A	C	<i>C17orf</i> 96	p.M35 8R	0.020 34	0.000 00	2.25 E- 80	Inf	GAATTTGAG GCCAGGGGG CTCAGGGAC AGCGGGACC CCCC[A/C]TC TGCCACCTC CACAGCGGG TGGGCGGGC GGGGGCTTA GA	SEQ ID NO: 400
chr 17	389534 72	G	C	<i>KRT28</i>	p.P251 R	0.019 36	0.000 01	7.38 E- 114	2199 .94[3 06.0 1- 1581 5.81]	CGCTCGCAT GTTGTTCAA CAAAACCGC GAGGTCTAC CCCC[G/C]GG GCCGCGTTC ATCTCCACG TTCACGTTG CCCCAGCC GC	SEQ ID NO: 401

chr 17	391908 97	A	G	<i>KRTA</i> <i>P1-3</i>	p.S59S	0.006 86	0.000 00	1.98 E- 41	Inf	GCTGGCAGC AGCTGGTCT CACAGCAGC TTGGCTGGC AGCA[A/G]CT GGAGCTGCA GGTCCCACT AGTTGAGAA GCTAGGAAA TC	SEQ ID NO: 402
chr 17	392743 71	C	T	<i>KRTA</i> <i>P4-11</i>	p.R66 H	0.005 15	0.000 04	3.87 E- 27	143. 18[4 9.13- 417. 29]	GCAGCTGGG GCGACAGCA GCTGGAGAT GCAGCATCT GGGG[C/T]GG CAGCAGGTG GGCTGGCAG CACACAGAC TGGCAGCAC TG	SEQ ID NO: 403
chr 17	392744 26	T	A	<i>KRTA</i> <i>P4-11</i>	p.S48 C	0.015 20	0.000 00	1.04 E- 90	Inf	TGGCAGCAC ACAGACTGG CAGCACTGG GGCCTGCAG CAGC[T/A]GG ACACACAGC AGCTGGGGC GACAGTAGG TGGTCCTGC AG	SEQ ID NO: 404
chr 17	392744 35	A	T	<i>KRTA</i> <i>P4-11</i>	p.C45 S	0.005 64	0.000 00	3.82 E- 34	Inf	ACAGACTGG CAGCACTGG GGCCTGCAG CAGCTGGAC ACAC[A/T]GC AGCTGGGGC GACAGTAGG TGGTCCTGC AGCAGGTGG TC	SEQ ID NO: 405
chr 17	392744 37	C	T	<i>KRTA</i> <i>P4-11</i>	p.C44 Y	0.005 15	0.000 00	3.08 E- 31	Inf	AGACTGGCA GCACTGGGG CCTGCAGCA GCTGGACAC ACAG[C/T]AG CTGGGGCGA CAGTAGGTG GTCCTGCAG CAGGTGGTC TC	SEQ ID NO: 406
chr 17	393166 23	C	T	<i>KRTA</i> <i>P4-4</i>	p.R107 R	0.011 52	0.000 00	4.62 E- 69	Inf	AGCAGGTGG GCTGGCAGC ACACAGACT GGCAGCACT GGGG[C/T]CT GCAGCAGCT GGGGCGGCA GCAGGTGGT CCTACAGCA GG	SEQ ID NO: 407

chr 17	393462 01	A	C	<i>KRTA</i> <i>P9-1</i>	p.T21 T	0.005 15	0.000 01	6.33 E- 30	577. 37[7 7.65- 4293 .3]	GCTGTCAGC CTACATGCT GCAGGACCA CCTGCTGCA GGAC[A/C]AC CTGCTGGAA GCCCACCAC TGTGACCAC CTGCAGCAG CA	SEQ ID NO: 408
chr 17	393465 75	A	C	<i>KRTA</i> <i>P9-1</i>	p.N14 6T	0.024 51	0.000 00	2.06 E- 133	Inf	TGCTGCCAG CCTACCTGC TGCCAGCCC ACCTGCTGC AGGA[A/C]C ACCTCTTGC CAGCCCACC TGCTGTGGG TCCAGCTGC TGC	SEQ ID NO: 409
chr 17	422392 92	A	G	<i>C17orf</i> <i>53</i>	p.S645 G	0.007 11	0.004 60	2.63 E- 02	1.55[1.06- 2.25]	ACTCCTGAG TGAGCTTCC TGAAGACTT CTTCTGTGG GACC[A/G]GT AGTTGAGAC TGCCCCAAC GCAGGACAA CCCACCATG AG	SEQ ID NO: 410
chr 17	428829 73	G	A	<i>GJC1</i>	p.L71 L	0.008 09	0.004 66	3.63 E- 03	1.74[1.22- 2.48]	CCACCAGGA TGATCTGGA ACACCCAGA AGCGTACAT GGGA[G/A]A GAGGTGCAA ACGCATCAT AACAGACAT TCTCACAGC CCG	SEQ ID NO: 411
chr 17	439234 10	C	T	<i>SPPL2</i> <i>C</i>	p.L380 L	0.011 27	0.007 13	4.61 E- 03	1.59[1.18- 2.14]	TGTGCGGCT GCCCCACTCT CAAGAACTG CTCCTCCTTC CTG[C/T]TGG CCCTGCTGG CCTTTGATGT CTTCTTTGTC TTCGTCAC	SEQ ID NO: 412
chr 17	452145 23	A	C	<i>CDC2</i> <i>7</i>	p.N57 5K	0.022 55	0.000 00	2.16 E- 134	Inf	ATACGACTT TGTCTTTGTA CTTCATTACC ACTTACCAT GC[A/C]TTAT AATGTCTAG GATTGACTC TGATAGCAT TTCGAAAAC	SEQ ID NO: 413

chr 17	452146 54	C	T	CDC2 7	p.A53 2T	0.005 88	0.000 00	1.38 E- 35	Inf	AGAGTATAG GCATAAGCG TAATTTGGA TCAACTTGG ATAG[C/T]TC TCTGGAAGA ATTTAATTG CAATATCAT GTTCCCGTT GC	SEQ ID NO: 414
chr 17	452146 99	T	C	CDC2 7	p.S517 G	0.015 93	0.000 00	3.52 E- 95	Inf	TGGAAGAAT TTAATTGCA ATATCATGT TCCCGTTGC AGAC[T/C]GA AACAGTTCC CTGCAGCAC ACCAGGCCT TAAAAAAT GG	SEQ ID NO: 415
chr 17	452162 16	A	G	CDC2 7	p.Y47 0Y	0.008 58	0.000 00	1.83 E- 51	Inf	GCCAAAGTG TTGTAGAGT AGATCTCCA TGCCTTCAA CTCT[A/G]TA ATTCTCAAT CCTTCTAAC CTCTGAGAA TATTCTTTCA G	SEQ ID NO: 416
chr 17	452192 83	T	C	CDC2 7	p.Y43 5C	0.015 93	0.000 00	2.76 E- 95	Inf	ATAGGCCCT TCCAATTTG GCACAGTAC CCAACCAGT ATTG[T/C]AG TGGTGAGAA GGTAGATGG CTCAAAATA TTTATAGCTT C	SEQ ID NO: 417
chr 17	452292 61	A	G	CDC2 7	p.T266 T	0.028 92	0.000 03	3.43 E- 167	1102 .44[3 50.3 6- 3468 .91]	CTCGGCTAT TTCCACTCTG TGAGAAGAC AGACTTTGT TCC[A/G]GTT TGGCCGATT CTGGCAACA GACTGTAAA ACACGAAAA G	SEQ ID NO: 418
chr 17	452342 98	G	C	CDC2 7	p.L214 V	0.017 89	0.000 01	6.35 E- 105	2015 .93[2 80.1 3- 1450 7.61]	AAAGTATCT TGTTTGACTT ACCTTGGGG TTAATGGAC TAA[G/C]AGC TGCTGGTCC TCCTAATAA ACTTCGACC AGTTTTTGGT	SEQ ID NO: 419

chr 17	452493 72	T	G	<i>CDC2</i> 7	p.A54 A	0.005 64	0.000 17	1.90 E- 22	32.6 9[17. 63- 60.6 2]	TAGTACAAC TGTGTCCTTT CAAGAGTCT ATATGCTTT ATA[T/G]GCC TTTCCTGAG CGGTAATAA CAGGTTGCC AGTAAAAAC A	SEQ ID NO: 420
chr 17	486534 06	G	C	<i>CACNA1G</i>	p.G54 8A	0.010 05	0.000 00	1.36 E- 53	Inf	CCACCACCC TCGACGCCT GCCCTCTCC GGGGCCCCC CCTG[G/C]TG GCGCAGAGT CTGTGCACA GCTTCTACC ATGCCGACT GC	SEQ ID NO: 421
chr 17	559172 91	G	A	<i>MRPS</i> 23	p.H14 2H	0.012 50	0.007 59	1.02 E- 03	1.65[1.25- 2.2]	CACTCAAGT GTTTCGGATT TCCGGGAAA CGTGACTAC CTCC[G/A]TG TTGCTTAAA AGACCAGAT TTAAGTATC ACAGAGATG TT	SEQ ID NO: 422
chr 17	560566 07	C	T	<i>VEZFI</i>	p.Q34 8Q	0.010 29	0.000 57	2.60 E- 32	18.2[12.1 3- 27.3 2]	TCCCTGGCC AGCTTGTCA CATGTTGTT GTTGTTGTTG TTG[C/T]TGC TGCTGCTGC TGCTGCTGC TGCTGCTGC TGCTGCTTTT	SEQ ID NO: 423
chr 17	615685 77	C	T	<i>ACE</i>	p.T342 M	0.009 80	0.006 47	1.34 E- 02	1.52[1.1- 2.09]	CCCCAGTTT GGGCAGAAC TCCCTCTGCT TGCAGGGCT GGA[C/T]GCC CAGGAGGAT GTTTAAGGA GGCTGATGA TTTCTTCACC	SEQ ID NO: 424
chr 17	616837 83	T	C	<i>TACO</i> 1	p.H16 6H	0.006 86	0.003 75	4.17 E- 03	1.84[1.25- 2.69]	TATCTAACA GTAGCCACA AGTGCCAAG CAGACATTA GACA[T/C]AT CCTGAATAA GAATGGGTA AGTGTGCGT CTGGGAGGA GT	SEQ ID NO: 425

chr 17	620386 02	T	C	<i>SCN4A</i>	p.H59 9R	0.007 11	0.004 73	3.76 E- 02	1.5[1 .03- 2.19]	CACAGTGAG CACGTTGTC AAAGTGCTC CGTCATGGG GTAA[T/C]GT TCCATGGCC ATGAAGAGG GTGTTGAGC ACGATGCAG AT	SEQ ID NO: 426
chr 17	742881 47	G	A	<i>QRIC H2</i>	p.D72 1D	0.009 80	0.000 01	2.53 E- 57	1105 .65[1 51.9 6- 8044 .57]	AACCAGGCT GATCTGCAC CAGGTTGGA TCAAACCAC GCTG[G/A]TC CATTCCAGG TTGGACCAA ACCACGCTG ATCCACTCC AG	SEQ ID NO: 427
chr 17	742881 72	C	T	<i>QRIC H2</i>	p.R713 H	0.006 62	0.000 00	5.40 E- 40	Inf	GATCAAACC ACGCTGGTC CATTCCAGG TTGGACCAA ACCA[C/T]GC TGATCCACT CCAGGTTGC ACCAAACCA CGCTGATCC AC	SEQ ID NO: 428
chr 17	742882 02	C	T	<i>QRIC H2</i>	p.R703 H	0.017 89	0.000 04	8.11 E- 100	407[164. 39- 1007 .67]	GACCAAACC ACGCTGATC CACTCCAGG TTGCACCAA ACCA[C/T]GC TGATCCACT CCAGGTTGG ACCAAACCA CGCTGATCT GC	SEQ ID NO: 429
chr 17	742884 18	A	T	<i>QRIC H2</i>	p.V63 1D	0.009 31	0.000 00	5.96 E- 56	Inf	ACCACGCTG AACTGCACC AGGTTGCAC CAAACCACG CTGA[A/T]CT ATACCAGGT TGCACCAA CTACGCTGA ACTTCACCA GG	SEQ ID NO: 430
chr 17	742885 95	C	T	<i>QRIC H2</i>	p.R572 H	0.007 11	0.000 01	1.92 E- 41	799. 67[1 08.9 1- 5871 .76]	CAAACCACG CTGATGATC TGCACGAGG TTGTGCCAA ACCA[C/T]GC TGATCTACT CCAGGTTGG ACCAAACCA TGCTGAAC GC	SEQ ID NO: 431

chr 17	743831 09	T	C	<i>SPHK1</i>	p.R285 R	0.005 15	0.002 84	1.15 E- 02	1.82[1.17- 2.83]	GTCTGGGGG AGATGCGCT TCACTCTGG GCACCTTCC TGCG[T/C]CT GGCAGCCCT GCGCACCTA CCGCGGCCG ACTGGCCTA CC	SEQ ID NO: 432
chr 17	768883 19	G	A	<i>LOC10 065351 5</i>	p.G89 G	0.010 78	0.007 79	4.44 E- 02	1.39[1.02- 1.9]	TCCACAGCT TGGCATCCG CTCTTCTCTG CAGAGCGAG ATC[G/A]CCT TTGCCCCGG GCTTGTAGC AATTTGTGC TTTTCTCTCC	SEQ ID NO: 433
chr 17	792545 30	C	T	<i>SLC38 A10</i>	p.V16 9M	0.006 37	0.003 72	1.29 E- 02	1.72[1.15- 2.56]	CACTGCCCA CTGAAGAGG CCGTGCTTG AGAGAGGAG AGCA[C/T]GA TCTGCAGAG GGAGAGGGG AGAGAGCAC GGGGCAGGT CA	SEQ ID NO: 434
chr 17	796820 59	T	C	<i>SLC25 A10</i>	p.I57T	0.005 15	0.002 31	1.51 E- 03	2.24[1.43- 3.5]	ATGACGGGC ATGGCGCTG CGGGTGGTG CGTACCGAC GGCA[T/C]CC TGGCACTCT ACAGCGGCC TGAGCGCCT CGCTGTGCA GA	SEQ ID NO: 435
chr 17	798471 52	G	A	<i>ALYRE F</i>	p.R148 R	0.005 64	0.003 17	1.10 E- 02	1.78[1.17- 2.72]	GCTCAAAGT GCACGTCTG CTGTTCTTA AGCTGCGAC CAGA[G/A]C GATCATAGT GCACAGCCG CCTTCTTCAG CGTTCCAAA TT	SEQ ID NO: 436
chr 17	799545 45	G	A	<i>ASPS R1</i>	p.L252 L	0.018 38	0.000 01	1.09 E- 107	2063 .13[2 86.7 9- 1484 2.01]	CTGCCCCCTT TGTTCTTTTC TCGGGTGGG GGACAGAGA CT[G/A]GGGG GCCCTCCTG GGCCACGA GGCCTCTGA CATCATCTT	SEQ ID NO: 437

chr 17	805296 14	G	T	<i>FOXK 2</i>	p.P259 P	0.009 80	0.006 35	1.23 E- 02	1.55[1.12- 2.14]	GTTTTGTGTT TGTTTTTTAA ATACAGGAT GATTCAAAG CC[G/T]CCTT ACTCCTACG CGCAGCTGA TAGTTCAGG CGATTACGA	SEQ ID NO: 438
chr 17	808993 49	T	C	<i>TBCD</i>	p.L118 5P	0.011 27	0.006 98	3.07 E- 03	1.62[1.2- 2.19]	AACCGTCTG TGTCACCTT CTGGGCGTA CCCAGGCCC CAGC[T/C]GG TGCCCCAGG TAACCCTGT CACCTTCAC AGCATGAGG TG	SEQ ID NO: 439
chr 18	345222 3	T	C	<i>TGIF1</i>	p.P82P	0.006 13	0.000 67	1.05 E- 14	9.21[5.81- 14.5 9]	CGACCCCT CTGCGCTCC TGGGGTCTT CCTGCGCCC CCCC[T/C]CC TCCACCGGC GCGCTGCCC ACAGCCGCG TGCCCTCTCC C	SEQ ID NO: 440
chr 18	939652 4	C	T	<i>TWSG 1</i>	p.A15 7V	0.006 13	0.003 99	4.36 E- 02	1.54[1.03- 2.31]	CACCACCAG AATGTGTCT GTCCCCAGC AATAATGTT CACG[C/T]GC CTTATTCCA GTGACAAAG GTAAGTGCC AACAGTTGA CT	SEQ ID NO: 441
chr 18	988737 1	C	A	<i>TXND C2</i>	p.L232 I	0.010 29	0.000 10	4.36 E- 49	99.4 4[49. 86- 198. 33]	CAAGTCCCC AGAAGAAGC CATCCAGCC CAAGGAGGG TGAC[C/A]TC CCCAAGTCC CTAGAGGAA GCCATCCAG CCCAAGGAG GG	SEQ ID NO: 442
chr 18	125467 78	A	G	<i>SPIRE 1</i>	p.A46 A	0.005 15	0.002 84	1.16 E- 02	1.82[1.17- 2.83]	CTTCTTCTGC AGCCTCATA GCCCTCATC ATTGCTACC GTC[A/G]GCT TCCACCGTG TTGGCCATG TGATCGATA AGCTGCTCT A	SEQ ID NO: 443

chr 18	189642 86	G	A	<i>GREB IL</i>	p.E93 K	0.007 60	0.004 61	1.32 E- 02	1.65[1.14- 2.4]	CAATCTAAC AGTTAATGA AATGGAAGA TGATGAAGA CGAT[G/A]AA GAAATGTCT GATTCAAAC AGCCCACCA ATTCCCTATT C	SEQ ID NO: 444
chr 18	289343 74	A	G	<i>DSG1</i>	p.I739 V	0.011 03	0.007 18	8.43 E- 03	1.54[1.14- 2.09]	TGTAGGTTTC CCCTGCTGG CTCTGTGGG TTGTTGTAG CTTC[A/G]TT GGAGAAGAC CTGGATGAC AGCTTCTTG GATACCCTG GG	SEQ ID NO: 445
chr 18	337850 83	G	A	<i>MOCO S</i>	p.Q35 4Q	0.012 75	0.009 01	1.86 E- 02	1.42[1.07- 1.88]	GAATGGAGA ATATAAAGC AGCACACCT TCACCTTGG CTCA[G/A]TA TACCTACGT GGCCCTGTC CTCTCTCCA GTACCCCAA TG	SEQ ID NO: 446
chr 18	641789 22	C	A	<i>CDH1 9</i>	p.V48 7L	0.005 88	0.000 01	1.39 E- 33	618. 95[8 3.71- 4576 .34]	ATGGATTCA TCTCTATCCA CTGCACTGA TAGTCTGAA TTA[C/A]CTA AAAAAAAAAG GGGGATAGA TTTTTGTGT TGTTTGGAT	SEQ ID NO: 447
chr 18	721140 55	G	A	<i>FAM6 9C</i>	p.A22 1V	0.006 62	0.003 79	1.29 E- 02	1.75[1.16- 2.65]	TGCCCTGTG GTGGGGGCT GCCCCGCGC CAGGAACTC CACC[G/A]CG TAGAAGTGG CCGCAGGAA CCCAGCACG GGCAGCACG TG	SEQ ID NO: 448
chr 18	723467 01	T	C	<i>ZNF40 7</i>	p.G12 42G	0.008 33	0.005 35	1.66 E- 02	1.56[1.1- 2.21]	ATTGTGAGG GTGAAGGAG GAAACGCAG GAGACGGTG GAGG[T/C]GT TGTCCTCCA CAGACACCT GTGCCCTGT GACGCTCGA TG	SEQ ID NO: 449

chr 19	287703	G	A	<i>PPAP2 C</i>	p.R85 C	0.010 54	0.006 93	9.91 E- 03	1.53[1.12- 2.08]	ACCTTGTAT ACAGCAGCC ACGTAGTTG TTGAAGTCC GAGC[G/A]A GAATAGAGC CGGTCTGTG TACACCAGG TAGGCTTCC CCG	SEQ ID NO: 450
chr 19	474688	T	G	<i>ODF3 L2</i>	p.R20 R	0.012 25	0.006 67	1.39 E- 04	1.85[1.38- 2.47]	ACTTCCTCA GGCCGGTCT CCGGAATCT GGCCCTCCG TCAC[T/G]CG CCGGCCAAG GGGGGCTGT GGCCAGCCG TGGGGTGGA GT	SEQ ID NO: 451
chr 19	104374 7	C	T	<i>ABCA7</i>	p.L318 L	0.005 39	0.003 01	1.35 E- 02	1.79[1.16- 2.76]	GGGGGTGCT GTCCACAGG TGAACCGGA CCTTCGAGG AGCT[C/T]AC CCTGCTGAG GGATGTCCG GGAGGTGTG GGAGATGCT GG	SEQ ID NO: 452
chr 19	143033 0	C	T	<i>DAZA PI</i>	p.F280 F	0.007 11	0.000 00	5.26 E- 40	Inf	TGTCCACCC CTCCTGGAG GCTTTCCCCC TCCCCAGGG CTT[C/T]CCT CAGGGCTAC GGTGCCCCG CCACAGTTC AGTAAGTCT A	SEQ ID NO: 453
chr 19	145711 1	C	A	<i>APC2</i>	p.P359 Q	0.029 90	0.000 01	8.67 E- 162	2568 .75[3 58.8 6- 1838 7.26]	CGCGCCAAC GCGGCGCTG CACAACATC GTCTTCTCGC AGC[C/A]GG ACCAGGGCC TGGCGCGCA AGGAGATGC GCGTCCTGC AC	SEQ ID NO: 454
chr 19	162098 0	G	T	<i>TCF3</i>	p.P360 P	0.015 20	0.000 15	1.11 E- 64	103. 81[5 1.56- 209. 02]	TCCCCTCCCC CCAAAACCC TCACAGACC TGCCAGGCC CTG[G/T]GGG GAGCCCACG GGGGTAGAA GGGCTGGAC GAGAAGTTA T	SEQ ID NO: 455

chr 19	177540 8	C	G	<i>ONEC UT3</i>	p.G48 3G	0.006 13	0.000 00	1.74 E- 27	Inf	TGAACCGCT GGGCTGAGG AGCCCAGCA CGGCCCCCG GGGG[C/G]CC CGCCGGCGC CACGGCCAC TTTCTCAA GGCCTGAGG CG	SEQ ID NO: 456
chr 19	224844 5	A	G	<i>SF3A2</i>	p.N43 2S	0.012 25	0.000 00	3.76 E- 47	Inf	CCTGGGGTC CACCCCTCAG CCTCCGGGA GTTACCCCC TCAA[A/G]TC CTGGGGTGC ACCCCCCAA CTCCCATGC CCCCAATGC TG	SEQ ID NO: 457
chr 19	225042 3	A	G	<i>AMH</i>	p.Y16 7C	0.007 84	0.000 05	2.29 E- 36	171. 42[5 2.47- 560. 02]	GGAGGAGCT GGCCCCCA GAGCTGGCG CTGCTGGTG CTGT[A/G]CC CTGGGCCTG GCCCTGAGG TCACTGTGA CGAGGGCTG GG	SEQ ID NO: 458
chr 19	287732 0	C	T	<i>ZNF55 6</i>	p.R122 C	0.008 09	0.005 57	4.26 E- 02	1.46[1.02- 2.07]	AAGGGTGGA GAGACCATG TAAAAGCAG TAAAGGTAA TAAA[C/T]GT GGAAGAACC TTCAGAAAG ACTCGAAAT TGTAATCGT CA	SEQ ID NO: 459
chr 19	395944 4	G	A	<i>DAPK 3</i>	p.R340 R	0.007 84	0.005 07	2.43 E- 02	1.55[1.08- 2.24]	CCACGTCCT CGTGGCAGA GCCGCCGGC TGCGCTGCA GCTC[G/A]CG CAGGCCCTC CTCGGCCGC CGCCGCCTC CTCCAGCAC CT	SEQ ID NO: 460
chr 19	451121 3	C	G	<i>PLIN4</i>	p.S906 T	0.007 84	0.000 00	3.99 E- 45	Inf	ACTGCAGAC GGTGTCTTT GGTACCGGT CAGGACAGT CTTG[C/G]TG GTGTCCACG CCGGTCTGG ACAGTCCCT TTGGCCAAG TT	SEQ ID NO: 461

chr 19	451351 9	C	T	<i>PLIN4</i>	p.K13 7K	0.006 13	0.000 01	1.43 E- 35	681. 44[9 2.31- 5030 .26]	GGACAGCCT TCGAGGTGT CCAGACCCC CTTGGACGG CCCC[C/T]TT AGCCATGTC CATGGCCCC TGTGACCCC GCTGGACAC CA	SEQ ID NO: 462
chr 19	572022 9	C	T	<i>LONP 1</i>	NM_0 01276 480:c.- 160+1 G>A	0.012 76	not found	2.62 E- 24	Inf	CGCCGCGAA ACGCACGTG ACGCCCCGGC GCGTGCCCTC GGTA[C/T]CC GATGGGCGC GTGGCTCGA AACAGCCGC TTCAGGGAG CT	SEQ ID NO: 463
chr 19	583160 8	G	A	<i>FUT6</i>	p.T324 M	0.009 56	0.006 38	1.69 E- 02	1.5[1 .09- 2.08]	GCAGAAAGC GAGTGCCCA GCTGAAGGA GCGAGGCCG CAGC[G/A]TC TCCCGCCAG CGAAAGTAG CTCAGGTAG CGGGCGTGG TC	SEQ ID NO: 464
chr 19	813810 4	C	T	<i>FBN3</i>	p.V25 94I	0.007 60	0.004 59	9.73 E- 03	1.66[1.15- 2.39]	CCCCGAGG GCCTGATCA AAGTCAAAG CCAGAGGGG CAGA[C/T]GC AGCGGAAGC CACCAAGAG TGTTGCGAC AGGAGGCGC TC	SEQ ID NO: 465
chr 19	815480 2	G	A	<i>FBN3</i>	p.P207 6S	0.006 86	0.000 00	1.44 E- 40	Inf	TGGGTGAGG GGCTCACCT TCTCGGGAG TCATCCGGG CCTG[G/A]GA CTGCCCGT GGCCAAAGG GGCAGAGCT CCTGAAAGG CA	SEQ ID NO: 466
chr 19	837316 4	C	T	<i>CD320</i>	p.G4D	0.006 51	0.003 69	2.31 E- 02	1.77[1.07- 2.78]	CAGAGCCCC TGTTCCGCA CGCTCCAAC CTGCGCCAT CCAA[C/T]CG CCGCTCATG CTGTCCCCA CAGCGGCGC CGGCCACGC GC	SEQ ID NO: 467

chr 19	839896 1	A	G	<i>KANK</i> 3	p.D48 9D	0.022 30	0.000 00	1.53 E- 101	Inf	AGCTACCCG GGGGCTCGG CGCCACCGT TCTCGCTGTC GCC[A/G]TCG CTGTCTGCTG GCGTCCTCG CTGGAGGAG CTCTCGTAC C	SEQ ID NO: 468
chr 19	856436 6	G	T	<i>PRAM</i> 1	p.P109 Q	0.006 86	0.000 02	1.85 E- 38	373. 54[8 8.96- 1568 .57]	CAGTTTGGGA CGGCTTCTT GGGGAGGTC AGTGACCTC AGGC[G/T]GC GGGGGCTTC TTGGGGAGG TCAGTGACC TCAGGCGGC GG	SEQ ID NO: 469
chr 19	905982 7	A	G	<i>MUC1</i> 6	p.S920 7P	0.013 24	0.009 05	9.32 E- 03	1.47[1.11- 1.93]	GTATCTGTA GTGACTTCA GTGATGGCC AGTATTTCA GCTG[A/G]GG TGCTGCTCA AATTTGGGG GTGAACTGG TTTCAGGTTT T	SEQ ID NO: 470
chr 19	907288 6	G	C	<i>MUC1</i> 6	p.P485 4A	0.005 64	0.003 66	4.88 E- 02	1.54[1.01- 2.35]	ATGGTGGAG GTGGTAACA TTTGGAGAT GTGACTTTA GATG[G/C]CT CTGGGTAAG CTGAGACAG TAGAATGTG ATTCAAATG CT	SEQ ID NO: 471
chr 19	923755 4	G	A	<i>OR7G</i> 3	p.P25S	0.008 133	0.004 873	1.80 E- 02	1.67[1.07- 2.5]	GTGGCCAGG TACATGGAC AGGAACAGC ATGAAGAGG ATGG[G/A]CT GCAGCTCCG GATCCCCTG ACAATCCCA AGAGAAAGA AT	SEQ ID NO: 472
chr 19	114887 25	G	A	<i>EPOR</i>	p.P488 S	0.005 39	0.003 25	2.60 E- 02	1.66[1.08- 2.56]	GGCAGAGGC TCAGCGGCT GGGATAAGG CTGTTCTCAT AAG[G/A]GTT GGAGTAGGG GCCATCGGA TAAGCCCCC TTGGGCTCC C	SEQ ID NO: 473

chr 19	120606 27	A	G	<i>ZNF70</i> 0	p.Q59 6Q	0.005 15	0.000 82	4.11 E- 10	6.32[3.93- 10.1 6]	AAAGGACTC ACACTGGAG AGAAACCCT ATGAGTGTA AGCA[A/G]TG TGGGAAAGC CTTCAGTTGT GCCTCAAAC CTTCGAAAG C	SEQ ID NO: 474
chr 19	121556 97	A	G	<i>ZNF87</i> 8	p.C173 C	0.007 35	0.000 00	2.32 E- 44	Inf	GAACAGAAC TGGGAAAAC TGAATGCTT TCCCACACT GCTT[A/G]CA TTCATAGGG TTTTTTTGCA GAGTGGATT CTTTCATGTC	SEQ ID NO: 475
chr 19	125014 51	C	T	<i>ZNF79</i> 9	p.P587 P	0.005 15	0.000 04	1.69 E- 26	115. 34[4 3.47- 306. 02]	TGAGAGAAG CAAATGCTT TCCCACATT CCTTACATTC ATA[C/T]GGG TTCTCTCCAG TATGAGTTTT TTCATGTCCT TGAAGAA	SEQ ID NO: 476
chr 19	125411 41	C	T	<i>ZNF44</i> 3	p.P615 P	0.013 24	0.000 01	1.75 E- 77	1489 .83[2 06.0 5- 1077 1.89]	TGAGAGAAG CAAATGCTT TCCCACATT CCTTACATTC ATA[C/T]GGG TTCTCTCCAG TATGAGTTTT TTCATGTCCT TGAAGAA	SEQ ID NO: 477
chr 19	141045 56	G	C	<i>RFX1</i>	p.P34 A	0.006 37	0.000 00	9.57 E- 37	Inf	GCAGCGGTG GGTGGCTGC GGGGGCTGG GGTGCCGCT GGGG[G/C]TG GTGGCGGTG GCGGCTGGG GCTGGGCTT GTGGCGGGG CC	SEQ ID NO: 478
chr 19	153539 36	G	A	<i>BRD4</i>	p.P982 S	0.017 65	0.000 00	8.48 E- 60	Inf	CGTGGAGGG GGCTGATGC TGCTGCTGG GGTGGAGGC TGGG[G/A]CT GGGGTGGTG GGGGTGGTG GCGGCTGCT GCTGCAGCT GC	SEQ ID NO: 479

chr 19	162756 56	C	T	<i>CIB3</i>	p.G13 9R	0.007 84	0.000 09	3.10 E- 38	88.1 6[43. 31- 179. 45]	ACCTTCTCA CATACCAGG CTCACCTCCT CGGCACTCA GCC[C/T]CCC CCGCGTCAG TTTGGTCAC CGTCTGCTC CAGGTCCCA C	SEQ ID NO: 480
chr 19	170390 23	A	C	<i>CPAM D8</i>	p.S110 3A	0.005 88	0.003 53	2.24 E- 02	1.67[1.11- 2.53]	GGCCTCGGG AGGGTCCAG GCCACAATG ACAGACTCA TTGG[A/C]TG GCTCTGGAC CATGGCCAA CCTGGAAAA AGAAACCAA GG	SEQ ID NO: 481
chr 19	178816 68	G	A	<i>FCHO 1</i>	p.R186 Q	0.009 56	0.006 78	4.21 E- 02	1.41[1.02- 1.95]	GAGAGCCTG CGGCGCTCA GTGGAAAAA TACAACTCA GCCC[G/A]AG CTGACTTTG AGCAGAAGA TGCTGGACT CAGCCCTGG TA	SEQ ID NO: 482
chr 19	178889 54	A	G	<i>FCHO 1</i>	p.E423 G	0.006 62	0.004 38	4.13 E- 02	1.51[1.03- 2.23]	AGAAGCAGC CCTCTTGGC CTCACCTCT CTAGCTGTG CAG[A/G]GA GATTGCAGT CAGAGGAGC AGGTGTCCA AGAACCTCT TT	SEQ ID NO: 483
chr 19	197446 14	C	T	<i>GMIP</i>	p.E795 K	0.009 07	0.005 65	8.03 E- 03	1.61[1.15- 2.26]	AGGCCCTCT CCATAGCTG TGGGCCCAG TGGGTTCTT ACCT[C/T]GG TAGGTGTGG CCGTGGGAT GCTGCTCCA GGTACTGT GG	SEQ ID NO: 484
chr 19	202294 04	C	A	<i>ZNF90</i>	p.G34 7G	0.018 14	0.000 00	3.39 E- 108	Inf	TCCATACTG GAGAGAAAC CCTACAAAT GTGAAGAAT GTGG[C/A]AA AGCCTTCAG GCGCTCCTT AGTCCTTCG TACACATAA GA	SEQ ID NO: 485

chr 19	202295 72	C	A	ZNF90	p.G40 3G	0.008 82	0.000 00	2.64 E- 52	Inf	GTCATAGTG AAAAGAAAC CCTACAAAT GTGAAGAAT GTGG[C/A]AA AGCCTTCAA GCGCTCCTC AACACTTAC TATACATAA GA	SEQ ID NO: 486
chr 19	212400 06	T	C	ZNF43 0	p.F298 L	0.011 03	0.000 00	4.36 E- 66	Inf	TGGAGAGAA ACCCTACAG ATGTGAAGA ATGTGGCAA AACC[T/C]TT AACCGGTCC TCACACCTT ACTACACAT AAAAGAATT CA	SEQ ID NO: 487
chr 19	217194 40	T	A	ZNF42 9	p.H19 5Q	0.010 05	0.007 15	3.83 E- 02	1.41[1.03- 1.93]	TTTGCATGCT TTCACAAC AACTCAACA TAAGAAAAT TCA[T/A]ATT AGAGAGAAT ACCTACAGA TGTAAGAA TTTGGCAAT G	SEQ ID NO: 488
chr 19	221543 42	A	C	ZNF20 8	p.V11 65G	0.024 02	0.000 00	1.90 E- 142	Inf	AAAGCCTTT GCCACATTC TTCACATTG TAGGGTTTC TCT[A/C]CAG TATGAATTT CTTATGATA ACTAAGGGT TGAGGACCA	SEQ ID NO: 489
chr 19	221548 29	A	T	ZNF20 8	p.C100 3S	0.005 64	0.003 66	4.87 E- 02	1.54[1.01- 2.35]	AGGTTTGAT GACCAGTTG AAAGCTTTG CCACATTCTT CAC[A/T]TTT GTAGGGTTT CTCTCCAGT ATGAATTAC CTTATGTTTA	SEQ ID NO: 490
chr 19	221556 91	A	G	ZNF20 8	p.H71 5H	0.017 65	0.000 01	2.21 E- 102	1917 .14[2 66.3 5- 1379 9.28]	TTTTGCCAC ATTCTTCAC ATTTGTAGG GTTTCTCTCC AGT[A/G]TGA ATTCTCTTAT GTTCCATAA GGTTTGAGG ACCAGTTGA	SEQ ID NO: 491

chr 19	222719 18	G	A	ZNF25 7	p.E456 K	0.014 71	0.000 00	6.24 E- 88	Inf	GTCTTCATA CCTTATTCG ACATAAGAT AATTCATAC TGGA[G/A]A GAAACCCTA CAAATGTGA AGAGTGTGG CAAAGCCTT TAA	SEQ ID NO: 492
chr 19	222720 71	A	G	ZNF25 7	p.I507 V	0.016 42	0.000 02	8.46 E- 95	926. 98[2 27.0 4- 3784 .7]	CAAAGCCTT TAACCGGTC TTCACACCTT TCTCAACAT AAG[A/G]TA ATTCATACT GGAGAGAAA CCCTACAAA TGTGAAGAA TG	SEQ ID NO: 493
chr 19	228476 44	G	A	ZNF49 2	p.K39 IK	0.008 09	0.000 00	2.33 E- 48	Inf	CACACCTTA CTACACATA AGAGAATTC ATACTGGAG AGAA[G/A]C CCTACAAAT GTGAAGAAT GTGGCAAAG CTTTTAACCT AT	SEQ ID NO: 494
chr 19	351753 06	G	A	ZNF30 2	p.D12 2N	0.008 33	0.004 89	4.44 E- 03	1.71[1.21- 2.42]	ATTTTCAAA TTCTAATAA GAATTTGGA ATATACAGA ATGC[G/A]AC ACATTTAGA AGCACCTTT CATTCAAAG TCTACTCTTT C	SEQ ID NO: 495
chr 19	360024 05	T	C	DMKN	p.S276 G	0.006 37	0.001 07	1.79 E- 11	6.01[3.87- 9.32]	CTGCCACCA CTGCTGCCG CCACTGCTG CCGCCACTG CTGC[T/C]GC CACTGCTGC TGCCACCAC TGCTGCTGC CATTGTTGTT G	SEQ ID NO: 496
chr 19	383774 17	C	T	WDR8 7	p.E229 8E	0.009 07	0.000 00	5.77 E- 42	Inf	CCTCCTCCTT CCTTTCCTCC TCCTCCTCCC TTACCTCCTC [C/T]TCCTCC CTTTCCTCTT CTTCTCCTCT TTCCTCCTCC TCCT	SEQ ID NO: 497

chr 19	383792 29	C	T	<i>WDR8</i> 7	p.A16 94A	0.005 64	0.003 03	8.94 E- 03	1.86[1.21- 2.88]	ATTTCTTGGC CAGTTTCTTC CTTTTCTGGG CCAATTCTC [C/T]GCCTCC TGGCTTAGC TTCTCCCCTC TTTGGGCCA GTGTTT	SEQ ID NO: 498
chr 19	388172 32	G	A	<i>KCNK</i> 6	KCNK 6(NM _0048 23:exo n2:c.3 23- 1G>A)	0.006 86	0.000 06	1.69 E- 34	108. 13[4 7.2- 247. 68]	AAAAGAAAA AGATTTACC CTTTACTCTC TTTACTCCCC TA[G/A]GCTA TGGGTACAC AACGCCACT GACTGATGC GGGCAAGGC	SEQ ID NO: 499
chr 19	404084 20	G	A	<i>FCGB</i> P	p.S147 3S	0.006 37	0.000 01	6.61 E- 37	702. 38[9 5.29- 5177 .19]	AATCTTTCA AGGGACCCT GGGGATCCA CCAGCTTGT GGCA[G/A]G AGGACAGTG GCCCTGTGG GGCTGGAGA GGAGCCCAC AGA	SEQ ID NO: 500
chr 19	404086 85	T	A	<i>FCGB</i> P	p.Q13 85L	0.006 37	0.003 03	8.36 E- 04	2.11[1.41- 3.15]	CTTGGGGTC GCCGTTGTA GTTCCCACA CAGGCCACA CATC[T/A]GC TGGTAGTAG TTTCCGGGG ACGGTGACC CGCACATAG TA	SEQ ID NO: 501
chr 19	405805 06	A	T	<i>ZNF78</i> 0A	p.C615 S	0.006 62	0.004 24	2.82 E- 02	1.57[1.06- 2.31]	AGCTGGGTG GGAAGACTA AAAACCTTT CCACATTCC TTAC[A/T]TT CAAAGGGTT TCTCACCAG TATGCAATT TCTGATGTC GA	SEQ ID NO: 502
chr 19	413558 49	A	G	<i>CYP2A</i> 6	p.L73 L	0.005 88	0.002 32	1.18 E- 04	2.55[1.67- 3.88]	GCATCATGT CCACACAGC ACCACGACC CGCCGGGGC CCCA[A/G]GT GAATGGTGA ACACGGGGC CATAGCGCT CACTGATCT GA	SEQ ID NO: 503

chr 19	416339 27	A	G	<i>CYP2F1</i>	p.P472 P	0.008 09	0.004 41	1.79 E- 03	1.84[1.29- 2.62]	TGCAGCCGC TGGGTGCGC CCGAGGACA TCGACCTGA CCCC[A/G]CT CAGCTCAGG TCTTGCAA TTTGCCGCG GCCTTTCCA GC	SEQ ID NO: 504
chr 19	428553 73	C	T	<i>MEGF8</i>	p.P847 P	0.009 31	0.000 00	3.25 E- 47	Inf	TGGGGTTCT GACTCCTCT GCCCAACTG ACCCCCAGG ACCC[C/T]TT CTGTGAGTG GCATCAGAG CACCAGCCG CAAAGGGGA CG	SEQ ID NO: 505
chr 19	434117 38	C	T	<i>PSG6</i>	p.L325 L	0.005 39	0.001 47	8.18 E- 07	3.69[2.36- 5.76]	CTGGCCAC AGAGGAACA AAGGATACT CACAGAGGA CATT[C/T]AG GGTGACTGG GTTACTGCG GATGCCACC ATATCGGTC CC	SEQ ID NO: 506
chr 19	434117 42	G	A	<i>PSG6</i>	p.T324 I	0.005 39	0.001 17	2.40 E- 08	4.64[2.95- 7.29]	CCCACAGAG GAACAAAGG ATACTCACA GAGGACATT CAGG[G/A]TG ACTGGGTTA CTGCGGATG CCACCATAT CGGTCCCGT AT	SEQ ID NO: 507
chr 19	440651 67	C	T	<i>XRCC1</i>	p.E50 E	0.006 62	0.004 26	2.90 E- 02	1.56[1.06- 2.3]	CATCATTCC CAATGTCCA CACTGTGTA TCTGCTCCTC CTT[C/T]TCC AACTGTGGG CAGAGAGAG AGGCCACTG TCAGTGCCT G	SEQ ID NO: 508
chr 19	445006 77	A	T	<i>ZNF155</i>	p.Q22 3L	0.005 15	0.002 93	1.84 E- 02	1.76[1.13- 2.74]	GGCAAGGAA TTTAGTCAA AGCTCACAT CTGCAAACT CATC[A/T]GA GAGTCCACA CTGGAGAGA AACCATTCA AATGTGAGC AA	SEQ ID NO: 509

chr 19	448906 74	A	G	<i>ZNF28</i> 5	p.L578 P	0.008 82	0.005 64	1.46 E- 02	1.57[1.12- 2.2]	TTATAATGTT TCTCTCTGCT CATGTAGTC TTTGATGAG TC[A/G]GAAG GTCCTTTCCA CGCTCACAA TGTGTGTAC TGTGTCTC	SEQ ID NO: 510
chr 19	458987 43	A	G	<i>PPP1R</i> <i>13L</i>	p.P435 P	0.008 33	0.000 31	1.83 E- 22	26.8[12.4- 57.9 3]	CAGGGGGCC ATGTCTGTT GGGGATGCT GGGGGGCTG GGGT[A/G]G GGGTTTGGG GTTGGGTCT GGGGCTGTG GGGGCAGCT GGG	SEQ ID NO: 511
chr 19	461377 13	G	A	<i>EML2</i>	p.R213 X	0.006 62	0.000 00	1.59 E- 39	Inf	TCCCCGGTG GGCAGCAAA TAAAGGTTG GCCCGGCAG TCTC[G/A]GC CACGGTAGC CATAGCTGG AGCCACCCA GGGGCTGGT TA	SEQ ID NO: 512
chr 19	462154 95	G	C	<i>FBXO</i> 46	p.P420 R	0.005 88	0.000 01	3.35 E- 33	595. 7[80. 57- 4404 .4]	GCCGGGCGC AGTGGCCGG GGAGTCGGC CGGGGGTGG CTCC[G/C]GG GGCCCGTCC GGCCCGCGG TTCTGGAGA AAGAAGAGC TG	SEQ ID NO: 513
chr 19	463139 18	C	G	<i>RSPH6</i> A	p.A27 7A	0.006 13	0.003 90	3.10 E- 02	1.58[1.05- 2.36]	CCTGTTTCGC CTTCAGTGC CGCCTCCAC TCCGGGTGA ACAG[C/G]GC CTTCTGTTTC TCCGCCATC TTGTAGGTG GGCTGCATC T	SEQ ID NO: 514
chr 19	472042 07	C	T	<i>PRKD</i> 2	p.V32 4M	0.011 27	0.008 30	4.45 E- 02	1.36[1.01- 1.83]	TTGTCAGCC TCGCTGAAA TCGGTGGCC TCCTCCATC GGCA[C/T]AT CTGTGGGGA CGGAGGCAT CAGAGGGGT CTCCACCCA GT	SEQ ID NO: 515

chr 19	475752 94	A	G	<i>ZC3H4</i>	p.H62 9H	0.005 15	0.002 07	4.03 E- 04	2.49[1.58- 3.93]	CAGGGTGCA TGTCGGGT GCATGTCGG GGTGCATGT CAGG[A/G]TG CATTGGACC GCCCATTGG CCCTGGGGG TCCCATGTT GG	SEQ ID NO: 516
chr 19	486245 55	C	T	<i>LIG1</i>	p.V68 5M	0.013 24	0.009 24	1.28 E- 02	1.44[1.09- 1.89]	AGGTAGGCG CCGATCACC ACCAGGTCC AGGGTGTCA CCCA[C/T]GC CATCAAGGT AGTCCTTCTT CAGCTGGGA GAAGGGGAG G	SEQ ID NO: 517
chr 19	486433 12	G	A	<i>LIG1</i>	p.L304 F	0.005 21	0.002 65	1.28 E- 02	1.97[1.11- 3.26]	CCAAGCTCC AGGCCCTGC TGGGGTGGC CCAAGGTGG TTGA[G/A]GC TGAGGTAGA GGACAGGGA GGAGGTCTG GAGGCGACA GG	SEQ ID NO: 518
chr 19	499318 84	T	G	<i>GFY</i>	p.L456 V	0.006 37	0.001 66	1.97 E- 07	3.86[2.44- 6.11]	CCAGAGATG ACCACGCCC CTTTGCACC CACAGTTCT GCAT[T/G]TG GACGCCCCG AAAGACCCC TACGACCTC TACTTTTATG C	SEQ ID NO: 519
chr 19	515180 60	T	C	<i>KLK10</i>	p.N27 6S	0.013 97	0.000 03	4.10 E- 79	525. 15[64.3 9- 1677 .55]	CATAACATC TGGATCAGC TGGAGCGTA GCATCTGGA TCAG[T/C]TG GAGCGTATG ACTTTATTG ATCCAGGAC ATGTATTG CA	SEQ ID NO: 520
chr 19	516283 92	G	T	<i>SIGLEC9</i>	p.G54 V	0.009 31	0.000 00	5.23 E- 56	Inf	TGCTCCTTCT CCTACCCCT CGCATGGCT GGATTTACC CTG[G/T]CCC AGTAGTTCA TGGCTACTG GTTCCGGGA AGGGGCCAA T	SEQ ID NO: 521

chr 19	519197 82	C	A	LOC10 012908 3	p.C38 X	0.005 88	0.002 98	3.44 E- 03	1.98[1.3- 3.02]	GTGTGGACC AGACGCCAT TCCCATCCC CCTCCCAGG GCTG[C/A]GG CGGCATCCT GGGACCCCA CAGCTTCCT CTCCCTGGA TG	SEQ ID NO: 522
chr 19	519197 84	G	C	LOC10 012908 3	p.G39 A	0.005 88	0.002 97	3.33 E- 03	1.99[1.3- 3.04]	GTGGACCAG ACGCCATTC CCATCCCCC TCCCAGGGC TGCG[G/C]CG GCATCCTGG GACCCACAC GCTTCCTCTC CCTGGATGC T	SEQ ID NO: 523
chr 19	519198 40	G	A	LOC10 012908 3	p.A58 T	0.008 82	0.005 15	3.85 E- 03	1.72[1.23- 2.42]	CCACAGCTT CCTCTCCCTG GATGCTCCT GAGCTGGGA GCC[G/A]CTC ACTGTCCCA CTGGGCTCC TCCACCTCC CCACCCACC G	SEQ ID NO: 524
chr 19	528880 30	T	A	ZNF88 0	p.H39 9Q	0.018 63	0.000 18	1.62 E- 81	106. 63[5 7.96- 196. 18]	GCAAGGTCT TCAGGCACA AGTTTTGTCT AACCAATCA TCA[T/A]AGA ATGCACACG GGAGAGCAA CCTTACAAA TGTAATGAA T	SEQ ID NO: 525
chr 19	528880 34	A	G	ZNF88 0	p.M40 1V	0.018 87	0.000 19	5.01 E- 81	102. 4[55. 69- 188. 29]	GGTCTTCAG GCACAAGTT TTGTCTAAC CAATCATCA TAGA[A/G]TG CACACGGGA GAGCAACCT TACAAATGT AATGAATGT GG	SEQ ID NO: 526
chr 19	528880 36	G	T	ZNF88 0	p.M40 1I	0.019 85	0.000 20	1.04 E- 84	99.0 5[55. 1- 178. 05]	TCTTCAGGC ACAAGTTTT GTCTAACCA ATCATCATA GAAT[G/T]CA CACGGGAGA GCAACCTTA CAAATGTAA TGAATGTGG CA	SEQ ID NO: 527

chr 19	531165 14	C	T	<i>ZNF83</i>	p.G43 5E	0.007 482	0.004 537	2.91 E- 02	1.65[1.04- 2.52]	CCGATGATG TGCTAGGGA TGAGTTTAG ACCGAAGAC CTTC[C/T]CA CATTTCATTA CATTATATAA GCTTTTCTC CAGTATGAA T	SEQ ID NO: 528
chr 19	532689 31	G	A	<i>ZNF60</i> 0	p.P693 L	0.012 99	0.000 01	4.13 E- 76	1466 .88[2 02.8 1- 1060 9.54]	CTGCTTGCT AAAGGCTTT GCCACACTC ATTACACTT GTAA[G/A]GT TTCTCTCCAG TGTGAAGTC CAGTATGTT GTTTCAGGT G	SEQ ID NO: 529
chr 19	536445 48	C	T	<i>ZNF34</i> 7	p.K51 2K	0.007 35	0.000 03	3.82 E- 40	264. 49[8 0.69- 866. 98]	TTTGAGTGA AGACCTTGC CACATTCAT TACATTTGT AAGG[C/T]TT TTCTCCAGT ATGGATGAC CTGATGGGT AGTTAGGTT TG	SEQ ID NO: 530
chr 19	537931 62	C	T	<i>BIRC8</i>	p.A15 6T	0.000 25	0.000 00	3.71 E- 02	Inf[NaN -Inf]	GAAGTCTGA TTCAATTCAT TTTCTGTAGT GTCTTTCTGA G[C/T]GCTCA CTAGATCTG CAACAAGAA CCTCAAGCG TTTTATAG	SEQ ID NO: 531
chr 19	552392 37	C	T	<i>KIR3D</i> L3	p.H17 2H	0.009 80	0.000 01	1.52 E- 52	829. 79[1 14.0 5- 6037 .46]	GGATCACTG AGGACCCCT TGCGCCTCG TTGGACAGC TCCA[C/T]GA TGCGGGTTC CCAGGTCAA CTATTCCAT GGGTCCCAT GA	SEQ ID NO: 532
chr 19	552509 79	C	A	<i>KIR2D</i> L3	p.P21T	0.010 29	0.000 00	8.87 E- 55	Inf	ATCTTTCTTT CCAGGGTTC TTCTTGCTGC AGGGGGCCT GG[C/A]CACA TGAGGGTGA GTCCTTCTCC AAACCTTCG GGTGTTCAT	SEQ ID NO: 533

chr 19	552848 21	G	A	<i>KIR2D</i> <i>L1</i>	p.G36 D	0.005 64	0.002 50	7.72 E- 04	2.26[1.47- 3.49]	CTAGGAGTC CACAGAAAA CCTTCCCTCC TGGCCCACC CAG[G/A]TCG CCTGGTGAA ATCAGAAGA GACAGTCAT CCTGCAGTG T	SEQ ID NO: 534
chr 19	552867 67	G	T	<i>KIR2D</i> <i>L1</i>	p.G17 4V	0.007 84	0.002 17	4.86 E- 09	3.64[2.5- 5.28]	TCCAGGGAA GGGGAGGCC CATGAACGT AGGCTCCCT GCAG[G/T]GC CCAAGGTCA ACGGAACAT TCCAGGCTG ACTTTCCTCT G	SEQ ID NO: 535
chr 19	552951 21	A	G	<i>KIR2D</i> <i>L1</i>	p.T301 T	0.006 62	0.003 25	1.28 E- 03	2.04[1.37- 3.04]	CTCTCCAGG ACTCTGATG AACAAGACC CTCAGGAGG TGAC[A/G]TA CACACAGTT GAATCACTG CGTTTTTAC ACAGAGAAA AA	SEQ ID NO: 536
chr 19	553300 36	G	A	<i>KIR3D</i> <i>L1</i>	p.V11 3M	0.026 23	0.000 38	5.79 E- 118	69.9 5[48. 58- 100. 73]	CCCACACTC CCCCACTGG GTGGTCGGC ACCCAGCAA CCCC[G/A]TG GTGATCATG GTCACAGGT CAGAGGCTT TCCGTCTGG GC	SEQ ID NO: 537
chr 19	553330 23	C	T	<i>KIR3D</i> <i>L1</i>	p.P220 L	0.028 68	0.000 02	9.70 E- 164	1523 .42[3 76.4- 6165 .8]	AGAACCTCC CTGAGGAAA CTGCCTCTTC TCCTTCCAG GTC[C/T]ATA TGAGAAACC TTCTCTCTCA GCCCAGCCG GGCCCCAAG	SEQ ID NO: 538
chr 19	554941 21	T	G	<i>NLRP2</i>	p.I330 S	0.007 85	0.001 80	8.85 E- 04	4.3[2 .1- 8.8]	AGGGCCCTG AGGGACCTC CGGATCCTG GCGGAGGAG CCGA[T/G]CT ACATAAGGG TGGAGGGCT TCCTGGAGG AGGACAGGA GG	SEQ ID NO: 539

chr 19	560296 21	A	C	<i>SSC5D</i>	p.T132 6T	0.016 67	0.000 00	1.11 E- 80	Inf	CCACCACTA CTCCTGATC CCACCACGA CCCCTCACC CCAC[A/C]AC TCCTGACCC TTCCTCAAC CCCTGTCAT CACTACTGT GT	SEQ ID NO: 540
chr 19	564163 47	G	A	<i>NLRP1</i> 3	p.A86 0V	0.006 86	0.003 84	4.86 E- 03	1.79[1.22- 2.63]	CTCCAGTCT CTCTAAGGC ACACTTGGG GTGAGTCAG GGCC[G/A]CA CACAATAGC TTTATGCCAT CATCTTGGA GCCGATTAA A	SEQ ID NO: 541
chr 19	579108 59	T	G	<i>ZNF54</i> 8	p.F402 V	0.007 60	0.000 00	8.20 E- 46	Inf	TGGAGAAAG GCCTTATAA ATGCAGTGA ATGTGGGAA ATCA[T/G]TT AGGTACCAC TGCAGGCTC ATTAGACAC CAGAGAGTC CA	SEQ ID NO: 542
chr 19	581183 90	T	C	<i>ZNF53</i> 0	p.S499 S	0.005 64	0.000 00	3.59 E- 34	Inf	CTGGAGAAA GGCCTTATG AGTGCAGTG TATGTGGGA AATC[T/C]TT TATCCGAAA AACCCACCT CATTCGACA CCAGACTGT TC	SEQ ID NO: 543
chr 19	583862 84	T	C	<i>ZNF81</i> 4	p.A15 8A	0.017 16	0.009 32	3.97 E- 06	1.86[1.45- 2.37]	AGACAGATG ACTCCCCTG ACACATGCA ACTTACACC TCTT[T/C]GC AAACAACGC CTCCTCAAC ACTCCCTCT GTAGGGTTT CT	SEQ ID NO: 544
chr 19	584385 05	C	T	<i>ZNF41</i> 8	p.G34 8G	0.007 11	0.000 00	6.65 E- 43	Inf	GTTGATGTT GAATGAGAT TGCCCTTCTG AGTAAAACA TTT[C/T]CCA CATTCTTCAC ACTCATAAG GTCTTTCTCC AGTGTGAA	SEQ ID NO: 545

chr 19	587723 21	C	A	<i>ZNF54</i> 4	p.P117 T	0.005 53	0.002 866	1.59 E- 02	1.93[1.11- 3.15]	ATCCCCACCA CGTGGAAGT GTACAGGAG TGGACCGGA GGAG[C/A]C ACCCTCTTTG GTATTAGGA AAAGTGCAA GATCAGAGC AA	SEQ ID NO: 546
chr 2	141821 0	G	A	<i>TPO</i>	p.T10 T	0.009 31	0.005 53	3.81 E- 03	1.69[1.22- 2.35]	TTAATTTTAG AATGAGAGC GCTCGCTGT GCTGTCTGT CAC[G/A]CTG GTTATGGCC TGCACAGAA GCCTTCTTCC CCTTCATCT	SEQ ID NO: 547
chr 2	100450 15	A	T	<i>TAF1B</i>	p.K27 9X	0.005 39	0.000 03	2.18 E- 29	201. 64[6 0.33- 673. 95]	TCTTTTATTT CAGTCTTGG CCTGACTAC GAGGACATC TAC[A/T]AAA AAACAGTAG AAGTTGGAA CATTTTTAG ATTTGCCTC G	SEQ ID NO: 548
chr 2	117744 03	C	T	<i>GREB</i> 1	p.S171 3F	0.001 47	0.000 13	3.24 E- 03	11.7 5[3.3 7- 40.9 2]	TCCAGCAAG ACCCGGGCC AGCGAGGTG CAAGAGCCC TTCT[C/T]CC GCTGCCACG TGCACAAC TCATCATCCT GAACGTGGA C	SEQ ID NO: 549
chr 2	179980 01	C	T	<i>MSGN</i> 1	p.G72 G	0.005 39	0.002 93	8.49 E- 03	1.84[1.2- 2.84]	CTCCCTGTCC AGCTGTGGC TGGGCTGCC CTGTGAGCA CGG[C/T]GGG GCCAGCAGT GGGGCAGC GAAGGCTGC AGTGTCCGT G	SEQ ID NO: 550
chr 2	239295 01	C	T	<i>KLHL2</i> 9	p.C865 C	0.011 03	0.008 04	4.80 E- 02	1.38[1.01- 1.87]	TCCTCCCCC ACATGCCCT GCCCTGTGT TCAGACACG GCTG[C/T]GT CGTGATAAA GAAATATAT TCAAAGCGG CTGACATCA GC	SEQ ID NO: 551

chr 2	243023 58	G	A	<i>TP53I3</i>	p.R258 X	0.003 93	0.000 20	2.03 E- 04	16.4[5.5- 49.2]	TTGTCCCTA GACCTCAGC AAACTGGTG ATCAGACTT CCTC[G/A]CT TAAAAAGTA GCTTTGAAA ACAGGGGCC CATTGATGT CA	SEQ ID NO: 552
chr 2	249302 62	C	T	<i>NCOA I</i>	p.A64 1A	0.009 56	0.006 69	3.26 E- 02	1.43[1.04- 1.98]	AAACCAGTC ACAACTAG TGCAGCTTTT GACAACAAC TGC[C/T]GAA CAGCAGTTA CGGCATGCT GATATAGAC ACAAGCTGC A	SEQ ID NO: 553
chr 2	264151 98	G	A	<i>HADH A</i>	p.L661 L	0.010 05	0.007 13	3.77 E- 02	1.41[1.03- 1.94]	TAGCCACTC AAACGGACT TACACTTCA GACTTAGGA GGCA[G/A]CT TCAGACTCG CTAAAATAC TATCCATGT CAGAATTCA AA	SEQ ID NO: 554
chr 2	266633 49	C	T	<i>DRCI</i>	p.T331 I	0.005 856	0.003 08	1.29 E- 02	1.91[1.11- 3.07]	TACAACCTTG CAGGTGCTG AAGAAGAGA GATGAAGAA AGCA[C/T]AG TAATTAAAT CCCAGCAGA AGAGGAAGA TCAATCGGT AA	SEQ ID NO: 555
chr 2	268523 40	C	G	<i>CIB4</i>	p.G42 R	0.017 40	0.000 00	2.59 E- 95	Inf	ACCTGGTCC ATGGTGAGC GTTGCCTCCT TGTAAGTACT TCC[C/G]AGG AGGGCAGAG CTTCAGGAA GGTGTGATG GATGCTGAA A	SEQ ID NO: 556
chr 2	292460 48	G	A	<i>FAM1 79A</i>	p.V53 6V	0.005 64	0.003 66	4.86 E- 02	1.55[1.01- 2.36]	AGGTCCTCA CCGGGAAGC TGCACGACG TGTGCTTGG TGCT[G/A]AC TGGGGAGGT GAGGCCCCC CAGCCTGTG TGCTGTGCA TT	SEQ ID NO: 557

chr 2	315951 30	C	T	<i>XDH</i>	p.R607 Q	0.005 64	0.003 51	3.21 E- 02	1.61[1.06- 2.45]	TCACCTTGAT CTTGCGGTG GGCCCCGGGT GCTGGTGAC CAGC[C/T]GG AGAGACAGC TCATTCTCGT AGCGAGGAA TGTCGTCAC A	SEQ ID NO: 558
chr 2	322890 02	C	T	<i>SPAST</i>	p.P34P	0.021 81	0.000 01	3.17 E- 123	2161 .13[3 01.0 2- 1551 5.36]	CTCCCAGGC CTCCGCCCC CTTGCCCTGG CCCCCGCCC CTCC[C/T]GC CGCCGGGCC GGCCCCCTCC GCCCCAGTC GCCGCATAA GC	SEQ ID NO: 559
chr 2	489827 64	A	T	<i>LHCG R</i>	p.L16 Q	0.008 58	0.000 29	3.94 E- 28	29.6 3[15. 93- 55.1 2]	GAGCGCTC GCGCAGCGC TCGTGGCAG CGGCGGCTG CAGC[A/T]GC AGCAGCAGC TTCAGCAGC TGCAGCGCC GAGAACCGC TG	SEQ ID NO: 560
chr 2	624498 65	C	T	<i>B3GN T2</i>	p.N17 0N	0.008 82	0.006 00	3.07 E- 02	1.47[1.05- 2.06]	GAAGGCAAG CAATCCGGG AATCCTGGG GCCAAGAAA GCAA[C/T]GC AGGGAACCA AACGGTGGT GCGAGTCTT CCTGCTGGG CC	SEQ ID NO: 561
chr 2	743265 94	C	T	<i>TET3</i>	p.P115 3P	0.019 61	0.000 01	5.92 E- 115	2204 .62[3 06.7- 1584 7.05]	AGGTGCTCA CCGCCTTCC CCCGCGAGG TCCGACGCC TGCC[C/T]GA GCCTGCCAA GTCCTGCCG CCAGCGGCA GCTGGAAGC CA	SEQ ID NO: 562
chr 2	744793 68	G	A	<i>SLC4A 5</i>	p.S472 S	0.006 37	0.003 58	7.75 E- 03	1.78[1.2- 2.66]	CCCCGATTT CATGCATGG CTGGCATCT CTCCATCAT CCCC[G/A]CT GCTTGTTCC GCCGGCCCC GCCACTGCC AGCCCCGCC GC	SEQ ID NO: 563

chr 2	747513 92	G	C	<i>DQX1</i>	p.T158 T	0.005 88	0.003 53	2.25 E- 02	1.67[1.1- 2.52]	CCTCATCTA GTACCAGCA CGCCCCAGG CTCCAGTGC CTCG[G/C]GT CGAGGCCAC CTCCTGCAG AAGCAGCCT GTCCCAGCA GA	SEQ ID NO: 564
chr 2	868317 51	G	C	<i>RNF103</i>	p.L421 V	0.006 62	0.003 69	5.86 E- 03	1.8[.22- 2.66]	CTCTTCTTCT CAAAGTAAT CAATTAGTA AACCATGAC CAA[G/C]GTA TGTA CTGAG AAACAGGGC TGGGTGTGA AGAGTAAAA C	SEQ ID NO: 565
chr 2	959456 67	G	A	<i>PROM2</i>	p.G45 0D	0.005 64	0.000 01	8.10 E- 32	571. 38[7.15- 4231 .98]	CTATTCGTG GTGCTCTGC AACCTGCTG GGCCTCAAT CTGG[G/A]CA TCTGGGGCC TGCTGCCA GGGACGACC CCAGCCACC CA	SEQ ID NO: 566
chr 2	981282 58	G	C	<i>ANKRD36B</i>	p.L102 1L	0.007 35	0.000 00	4.61 E- 27	Inf	GTCTTTGCCT GCTCTCTCTT TGCTTCTCCA GTTTGGAAC G[G/C]AGCGT TGTGTTTCA TCTGTCAGA GCAGCAAGC TGTCCAC	SEQ ID NO: 567
chr 2	981283 13	G	A	<i>ANKRD36B</i>	p.T100 3M	0.017 16	0.000 07	7.10 E- 60	240. 02[5.84- 979. 16]	ATCTGTCTAG AGCAGCAAG CTGTCCACT ATAACAGGC TATC[G/A]TT TTTGCTAAT GTTTCCCAT TCCGTTTGTAG AGCCTTTTG	SEQ ID NO: 568
chr 2	996517 98	G	A	<i>TSGA10</i>	p.S503 S	0.005 21	0.001 39	1.76 E- 05	3.76[2.09- 6.31]	TAATACAGA GTTCCCTAG TAGAAGACA AATCTGCAA GAGC[G/A]G ACACTTTTTTC AAACTGAAC CTTCTGAAG CTCCTCTTCC A	SEQ ID NO: 569

chr 2	108486 338	G	T	<i>RGPD</i> 4	RGPD 4(NM _1825 88:exo n19:c. 2606- 1G>T)	0.025 25	0.000 38	1.47 E- 74	67.6[34.1 8- 133. 72]	ACTTTAACA GTGTTTCTT TCTTTTCTTT TTTTTTTTTT A[G/T]TTGCA ACTACTGGC CCTTCAGTA TATTATAGT CAGTCACC	SEQ ID NO: 570
chr 2	109347 813	T	G	<i>RANB</i> P2	p.L96 L	0.014 95	0.000 00	1.94 E- 89	Inf	ATTAGCGTT CAGTGGAAAT TAAACCCAA CACAAAAAG ATCT[T/G]GT GTTGAAGAT TGCAGAATT GCTTTGTAA AAATGATGT TA	SEQ ID NO: 571
chr 2	112922 601	C	G	<i>FBLN7</i>	p.P87 A	0.007 35	0.004 26	7.23 E- 03	1.73[1.19- 2.51]	TCCATCTCTC CTTACAGTTT CCTGCCCGG CTCTGAACA CC[C/G]CCGC AGACGGCAG AAAGTTTGG AAGCAAGTA CTTAGTGGA	SEQ ID NO: 572
chr 2	113940 187	G	A	<i>PSD4</i>	p.A52 T	0.022 55	0.000 01	6.94 E- 133	2577 .18[3 59.1- 1849 5.63]	CCATGAGGA TCCACCGGA GCCTTTCGA GGAGCAAAC CTGG[G/A]CC ACTGACCCT CCTGAACCT ACCAGACAA AATGTTCTC CC	SEQ ID NO: 573
chr 2	114500 349	C	T	<i>SLC35</i> F5	p.E224 K	0.009 07	0.006 35	4.53 E- 02	1.43[1.03- 1.99]	GCAGTAAGT TTCCCCACA GTTTTTCAGT ATGGATTCT TGTT[C/T]TT TCACAGGAT ATGACATGC GAGACAACT TTGCTTCCA AT	SEQ ID NO: 574
chr 2	132238 100	T	C	<i>TUBA3</i> D	p.A27 8A	0.007 35	0.004 75	2.79 E- 02	1.55[1.07- 2.25]	TCCACTTCCC CCTGGCCAC CTATGCCCC AGTCATCTC AGC[T/C]GAG AAGGCCTAC CACGAGCAG CTGTCTGTG GCCGAGATC A	SEQ ID NO: 575

chr 2	136418 868	A	G	<i>R3HD M1</i>	p.H59 6R	0.005 64	0.002 60	1.00 E- 03	2.18[1.42- 3.33]	TTATGATCCT AGATGCCAG CCTGTTATTG CGCTCCAGG CC[A/G]CTAT CACTCCAGC CAACCTCAG TATCGCCCA GTCCCTTCT	SEQ ID NO: 576
chr 2	141232 800	C	T	<i>LRP1B</i>	p.A31 78T	0.007 84	0.011 71	2.07 E- 02	0.67[0.47- 0.95]	GCCCAGTAG AGTCTACGA TTAACATAA TCTATTGTTA GTG[C/T]CAT AGGTCTAGA AATCTTGGT TTCTATGAC AACACTCTG A	SEQ ID NO: 577
chr 2	152982 745	T	C	<i>STAM2</i>	p.M39 2V	0.006 62	0.003 83	9.98 E- 03	1.73[1.17- 2.56]	ATAATTTAG AAAATGTTC TCAAAAAAC ATGCTCACC TGCA[T/C]TG GAACCCAG ATGATGCAG GTGGGTAAT GTGCTGGAG GG	SEQ ID NO: 578
chr 2	165984 284	C	T	<i>SCN3A</i>	p.V10 84I	0.012 25	0.007 22	7.31 E- 04	1.71[1.28- 2.27]	GGGTTGTTT ATGAATGAC ATATAATCA TTTTCATCGA TTA[C/T]GTA TTTTTCAACA CTGCTTCCA GTACCTACA CCACTGGTG	SEQ ID NO: 579
chr 2	171070 982	G	A	<i>MYO3 B</i>	p.G13 9R	0.005 205	0.003 108	4.93 E- 02	1.68[0.95- 2.77]	CCAGCGGTT GGATGAAGC AATGATCTC ATACATCTT GTAC[G/A]GG GCCCTCTTG GTAAGAACA TCTATCAAA TGGGGTATG AC	SEQ ID NO: 580
chr 2	178096 406	G	A	<i>NFE2L 2</i>	p.L286 F	0.005 64	0.003 04	6.39 E- 03	1.86[1.22- 2.84]	AGATCAGAA ACATCAATG GGCCCATTT AGAAGTTCA GAGA[G/A]T GAATGGCTT AAAGTAGCA GGTGAGGGC ATGCTGTTG CTG	SEQ ID NO: 581

chr 2	186661 602	A	G	<i>FSIP2</i>	p.R333 6G	0.006 86	0.003 99	1.12 E- 02	1.72[1.16- 2.56]	ATCGTGTTCT ACTAGAAAC AAAGTACAA GACCACAGA CCA[A/G]GG GAATCTAAC TTTGGTAGTT TTGATCAGA CCATGAAAG G	SEQ ID NO: 582
chr 2	186678 577	A	T	<i>FSIP2</i>	p.K68 00N	0.025 49	0.000 00	3.65 E- 151	Inf	TTTCTCCTAA GTCAACACT AAGCACGAG CAGCCTGAA AAA[A/T]TTT TTGTCATA AGTAAATGT TGTCAGACC ACAGCCAGT G	SEQ ID NO: 583
chr 2	187605 000	G	A	<i>FAM1 71B</i>	p.R95 H	0.007 11	0.004 51	2.40 E- 02	1.58[1.09- 2.3]	GTATTTATGT TGAAAGTCC AGGTGAATG ACATCATCA GTC[G/A]TCA GTACCTGAG CCAAGCAGT TGTAAGAGT GTTTGTA C	SEQ ID NO: 584
chr 2	209302 329	G	A	<i>PTH2R</i>	p.S82S	0.006 62	0.000 01	1.50 E- 38	743. 52[1 01.0 1- 5472 .96]	GACTCATTT GTTGGCCCA GAGGAACAG TGGGGA TATC[G/A]GC TGTTCCATG CCCTCCTTAT ATTTATGAC TTCAACCAT A	SEQ ID NO: 585
chr 2	211068 107	C	A	<i>ACAD L</i>	p.R311 M	0.007 11	0.002 04	4.63 E- 08	3.5[2 .37- 5.16]	AACTGTTTT GCCAAAAGC TTTTCTTTGT TTAACATAG TTC[C/A]TGG TTTCTTCAAA CATGAATTC ACTAGCTGA AATTGCCAC	SEQ ID NO: 586
chr 2	216285 492	C	T	<i>FNI</i>	p.V52 7M	0.001 47	not found	4.03 E- 06	Inf	ATGTGCCCC TCTTCATGA CGCTTGTTGG AATGTGTCG TTCA[C/T]AT TGTAAGTGA TGTCATCAA CAATGCACT GATCTGTTT AG	SEQ ID NO: 587

chr 2	233246 249	A	G	<i>ALPP</i>	p.E451 G	0.006 86	0.004 01	8.56 E- 03	1.71[1.17- 2.52]	AGCCCCGAG TATCGGCAG CAGTCAGCA GTGCCCCTG GACG[A/G]A GAGACCCAC GCAGGCGAG GACGTGGCG GTGTTGCGC CGC	SEQ ID NO: 588
chr 2	233498 515	C	G	<i>EFHD 1</i>	p.P34 R	0.010 05	0.000 00	2.59 E- 36	Inf	GAGAGTGGC CCCCAGCTG GCTCCCCTC GGCGCCCCA GCCC[C/G]GG AGCCCAAGC CCGAGCCCG AGCCTCCCG CCCGTGCGC CC	SEQ ID NO: 589
chr 2	234229 468	C	T	<i>SAG</i>	p.T125 M	0.005 88	0.003 32	1.25 E- 02	1.78[1.17- 2.7]	CTTAAAAAG CTGGGGAGC AACACGTAC CCCTTTCTCC TGA[C/T]GGT GGGTGACTC CTCCGGCCA GCCCTGCTT CCTTCACCC G	SEQ ID NO: 590
chr 2	237029 013	C	T	<i>AGAP 1</i>	p.C711 C	0.025 25	0.000 03	9.57 E- 145	943. 45[2 99.2 2- 2974 .8]	TGCTGGCAC ACGGCTCCC GGGACGAGG TGAACGAGA CCTG[C/T]GG GGAGGGAGA CGGCCGCAC GGCGCTGCA TCTGGCCTG CC	SEQ ID NO: 591
chr 2	238973 062	A	G	<i>SCLY</i>	p.K60 E	0.002 94	0.000 67	5.74 E- 05	4.37[2.37- 8.05]	AACGACTCC CCTGGAGCC AGAAGTTAT CCAGGCCAT GACC[A/G]A GGCCATGTG GGAAGCCTG GGGAAATCC CAGCAGCCC GTA	SEQ ID NO: 592
chr 2	240982 243	G	A	<i>PRR21</i>	p.R53 W	0.021 32	0.000 05	1.26 E- 112	480. 79[1 76.3 8- 1310 .53]	GGGTGAAGA GCCGTGGAT GAAGGGCCG TGGGTGAAG AGCC[G/A]TG GATGAAGGG CCATGGGTG AAGAGCCGT GGATGAAGG GC	SEQ ID NO: 593

chr 2	242154 318	G	A	<i>ANO7</i>	NM_0 01001 891:ex on18:c .1988+ 1G>A	0.005 89	0.000 80	3.42 E- 04	7.1[3 - 16.5]	GCAAGCAGG TCATCAACA ACATGCAGG AGGTCCTCA TCCC[G/A]TG AGTCCCCCA CTCCTCCCTG GGTGGCATC CAAGGACCG A	SEQ ID NO: 594
chr 2	242207 024	T	A	<i>HDLB P</i>	p.T14S	0.009 07	0.006 34	4.29 E- 02	1.43[1.02- 2.02]	ACCACACAC CTCTTAATG CTTACAAAA TGCATCATG ACAG[T/A]TG CTACAAAAA GCCAGCGGT CTCTCTCTGC AAGGTGCAT C	SEQ ID NO: 595
chr 2	242312 655	C	T	<i>FARP2</i>	p.H45 Y	0.008 82	0.006 12	4.12 E- 02	1.45[1.03- 2.03]	TGGGCAGAC TCTCTTGCCC AGAATGCAA GAGAAGCAC CTG[C/T]ACC TCAGAGTAA AGCTGCTGG ACAACACCA TGGAAATAT T	SEQ ID NO: 596
chr 20	314753 1	G	A	<i>LZTS3</i>	p.L93 L	0.009 56	0.006 19	1.14 E- 02	1.55[1.12- 2.14]	CACTGCCCC GCAGCTCAC CATTGAGGT AGAGGGAGT TGGC[G/A]AG ACCCTTGTC CTCTGAGGG GTAGCGGCC CGGCCTCTC CC	SEQ ID NO: 597
chr 20	468011 8	T	C	<i>PRNP</i>	p.S55P	0.005 64	0.000 02	1.14 E- 31	314. 81[7 4.2- 1335 .71]	GTGGCTGGG GGCAGCCCC ATGGTGGTG GCTGGGGAC AGCC[T/C]CA TGGTGGTGG CTGGGGTCA AGGAGGTGG CACCCACAG TC	SEQ ID NO: 598
chr 20	317569 87	C	T	<i>BPIFA 2</i>	p.G12 G	0.005 15	0.002 77	9.96 E- 03	1.86[1.2- 2.9]	AAAAGATGC TTCAGCTTTG GAAACTTGT TCTCCTGTGC GG[C/T]GTGC TCACTGGGA CCTCAGAGT CTCTTCTTGA CAATCTTG	SEQ ID NO: 599

chr 20	340785 17	G	A	CEP25 0	p.E881 K	0.010 78	0.007 88	4.80 E- 02	1.37[1.01- 1.86]	CTGGCACCA GCAGGAGCT GGCAAAGGC TCTGGAGAG CTTA[G/A]AA AGGGAAAAA ATGGAGCTG GAAATGAGG CTAAAGGAG CA	SEQ ID NO: 600
chr 20	341303 30	T	C	ERGIC 3	p.F76F	0.007 11	0.000 09	3.01 E- 34	79.9 3[38. 93- 164. 12]	CGCGGGGAG ATAAACTGA AGATCAACA TCGATGTAC TTTT[T/C]CC GCACATGCC TTGTGCCTGT GAGTACCTC ACCATGGGT G	SEQ ID NO: 601
chr 20	462798 39	G	A	NCOA 3	p.Q12 55Q	0.011 27	0.000 00	5.51 E- 65	Inf	GGGTGGCTA TGATGATGC AGCAGCAGC AGCAGCAGC AACA[G/A]C AGCAGCAGC AGCAGCAGC AGCAGCAAC AGCAACAGC AAC	SEQ ID NO: 602
chr 20	485033 06	G	A	SLC9A 8	p.S519 S	0.009 07	0.006 33	3.58 E- 02	1.44[1.03- 2]	GGCCGCCTT TCCTCCCTGC TCAGGGCAA CACTGTGGA GTC[G/A]GAG CACCTGTCTG GAGCTCACG GAGGAGGAG TACGAGGCC C	SEQ ID NO: 603
chr 20	491978 54	G	A	PTPNI	p.G30 8S	0.005 541	0.002 6	6.45 E- 03	2.14[1.23- 3.49]	CACTGAAGT TAGAAGTCG GGTCGTGGG GGGAAGTCT TCGA[G/A]GT GCCCAGGCT GCCTCCCCA GCCAAAGGG GAGCCGTCA CT	SEQ ID NO: 604
chr 20	609019 32	C	T	LAMA 5	p.V17 35M	0.011 27	0.007 93	2.52 E- 02	1.43[1.06- 1.92]	ACCCTGCCA CATCATCTC AGCTCCCTC ACCTGCAGC ACCA[C/T]AT CCGGCCTGC TCTCCATGG GGACAAAGA CATCTCCCC GC	SEQ ID NO: 605

chr 20	612963 67	C	A	<i>SLCO4 A1</i>	p.G40 1G	0.011 52	0.008 55	4.76 E- 02	1.35[1.01- 1.81]	TCTGCCTGG CCGGGGCCA CCGAGGCCA CTCTCATCA CCGG[C/A]AT GTCCACGTT CAGCCCCAA GTTCTTGGA GTCCCAGTT CA	SEQ ID NO: 606
chr 20	622005 87	C	T	<i>HELZ2</i>	p.S334 S	0.005 15	0.003 16	3.47 E- 02	1.63[1.05- 2.54]	GGTGCATCC TCTGCCGAT AGTTGGTTG GTGAGATGG GGCC[C/T]GA GGCCACGCT GCTGCGGTT GAACTCCAG GGCCAGGGC AG	SEQ ID NO: 607
chr 21	109429 55	T	G	<i>TPTE</i>	p.Q17 3P	0.005 88	0.000 41	9.09 E- 18	14.4 3[8.7 8- 23.7]	ACTTACCCG CCTTCTTATC AGCTTTTCA AGTTGTCTTT TT[T/G]GATG AAACAGATG AAAAATTCT TAACAGAAT AATAAGTCG	SEQ ID NO: 608
chr 21	109429 81	C	A	<i>TPTE</i>	p.L164 L	0.012 75	0.000 79	1.16 E- 38	16.3 9[11. 59- 23.1 7]	CAAGTTGTC TTTTTTGATG AAACAGATG AAAAATTCT TAA[C/A]AGA ATAATAAGT CGTAGAAGT CGAAGTAAA TGTGTCCAT C	SEQ ID NO: 609
chr 21	149827 21	A	G	<i>POTE D</i>	p.R58 G	0.022 79	0.000 11	8.43 E- 67	216. 23[5 3.26- 877. 86]	CACCTTCTGG AGACCACGA CGACTCCTTT ATGAAGATG CTC[A/G]GGA GCAAGATGG GCAAGTGTT GCCGCCACT GCTTCCCCT G	SEQ ID NO: 610
chr 21	349274 26	C	G	<i>SON</i>	p.R196 3R	0.008 33	0.000 00	2.93 E- 36	Inf	GCATTTCCC CAAGCCGCC GCAGCCGCA CCCCAGCC GCCG[C/G]AG CCGCACCCC CAGCCGCCG CAGCCGCAC CCCCAGCCG CC	SEQ ID NO: 611

chr 21	427708 96	G	A	<i>MX2</i>	p.G40 8R	0.010 05	0.006 66	1.46 E- 02	1.51[1.1- 2.08]	GGAGAGCCA CCAGAAGGC GACCGAGGA GCTGCGGCG TTGC[G/A]GG GCTGACATC CCCAGCCAG GAGGCCGAC AAGATGTTC TT	SEQ ID NO: 612
chr 21	434126 40	G	C	<i>ZBTB2 1</i>	p.A52 2G	0.007 60	0.005 10	3.45 E- 02	1.49[1.04- 2.15]	ACCAAATTC GTCTTTATTC AAATCAGAA TCTGGAAAA TCT[G/C]CAT CAAGGAGAG TAGGGCTTG AGCCTTCCT CAAAATTAT C	SEQ ID NO: 613
chr 21	456707 74	G	A	<i>DNMT 3L</i>	p.S276 S	0.024 75	0.000 01	1.25 E- 145	2810 .21[3 91.9 4- 2014 9]	GCACCAGAT TGTCACGA ACATCCAGA AGAAGGGCC TGGG[G/A]CT GCCTGGCTT GGGCCGTGC GTACTGCAG GAGCCGGTG GA	SEQ ID NO: 614
chr 21	457866 70	G	A	<i>TRPM 2</i>	p.V15 3M	0.008 33	0.005 61	3.32 E- 02	1.49[1.05- 2.11]	CCCGCAGTA CGTCCGAGT CTCCAGGA CACGCCCTC CAGC[G/A]TG ATCTACCAC CTCATGACC CAGCACTGG GGGCTGGAC GT	SEQ ID NO: 615
chr 21	459947 69	T	C	<i>KRTA P10-4</i>	p.P378 P	0.011 76	0.000 01	1.15 E- 68	1313 .63[1 81.2 8- 9519 .28]	GCCGCCCG TGTGCAGGC CCGCCTGCT GCGTGCCCG TCCC[T/C]TC CTGCTGTGC TCCACCTC CTCCTGCCA ACCCAGCTG CT	SEQ ID NO: 616
chr 21	459998 67	T	A	<i>KRTA P10-5</i>	p.T197 S	0.008 82	0.000 00	4.27 E- 53	Inf	CAGCAAGCC GGCTGACAG CTAGACTGC TGGCAGCAT GAAG[T/A]G GAAGCCCCA GAGCAGACG GGCACACAG CAGATGGGT TTG	SEQ ID NO: 617

chr 21	460000 42	G	A	<i>KRTA</i> <i>P10-5</i>	p.P138 P	0.026 47	0.000 00	3.02 E- 158	Inf	ATGAAGAGG AATCCTCAG AACAGGTGG GCACACAGC ACAC[G/A]G GCTTGCAGC AGACAGGCA CACAGCAGG ACTGCTGGC AGG	SEQ ID NO: 618
chr 21	460206 47	C	T	<i>KRTA</i> <i>P10-7</i>	p.C42 C	0.012 75	0.001 26	7.61 E- 31	10.2 4[7.4 3- 14.1 2]	CCGACTCCT GGCAGGTGG ACGACTGCC CAGAGAGCT GCTG[C/T]GA GCCCCCTG CTGCGCCCC CAGCTGCTG CGCCCCGGC CC	SEQ ID NO: 619
chr 21	460324 74	T	C	<i>KRTA</i> <i>P10-8</i>	p.S153 P	0.014 22	0.000 00	3.77 E- 85	Inf	TGGAGCTTC CTCCCCATG CTGCCAGCA GTCTAGCTG CCAG[T/C]CA GCTTGCTGC ACCTTCTCCC CATGCCAAC AGGCCTGCT G	SEQ ID NO: 620
chr 21	461174 08	T	C	<i>KRTA</i> <i>P10-12</i>	p.S98P	0.017 40	0.000 01	3.00 E- 102	1974 .74[2 74.3- 1421 6.51]	CTGCCAGCA GTCTAGCTG CCAGCCGGC TTGCTGCAC CTCC[T/C]CC CCCTGCCAG CAGGCCTGC TGCGTGCCC GTCTGCTGC AA	SEQ ID NO: 621
chr 21	461914 00	G	A	<i>UBE2</i> <i>G2</i>	p.P60P	0.008 33	0.005 68	3.46 E- 02	1.47[1.04- 2.08]	ACATTTTGG ACGCATCCA CGTTAGCTC CACTTTCGTC ATT[G/A]GGC TCTGAAAGA AAAGGGAAC ACCCTCCAT GTAAAAGGG A	SEQ ID NO: 622
chr 21	465964 59	G	A	<i>ADAR</i> <i>B1</i>	p.K28 1K	0.008 33	0.005 59	2.59 E- 02	1.5[.06- 2.12]	TCGTGGATG GTCAGTTCTT TGAAGGCTC GGGGAGAAA CAA[G/A]AA GCTTGCCAA GGCCCGGGC TGCGCAGTC TGCCCTGGC CG	SEQ ID NO: 623

chr 22	185627 34	T	C	<i>PEX26</i>	p.Y10 9H	0.005 21	0.002 87	2.61 E- 02	1.82[1.03- 3.01]	AATGGATCG GTGGCAAGA AGTCCTCTC CTGGGTCCT TCAG[T/C]AT TACCAGGTC CCTGAAAAG CTACCCCCC AAAGTCCTG GA	SEQ ID NO: 624
chr 22	240867 34	G	A	<i>ZNF70</i>	p.C198 C	0.013 48	0.000 01	4.80 E- 79	1525 .31[2 11.0 3- 1102 4.83]	TGAGGGCTG AGCTCTGGC GGAAGGCCT TCCCACACT CCCG[G/A]CA CTCGTAGGG CTTCTCCCCG GTGTGGATG ATCTGGTGC C	SEQ ID NO: 625
chr 22	250071 72	G	A	<i>GGT1</i>	p.A42 T	0.008 82	0.002 66	5.51 E- 09	3.34[2.34- 4.76]	AGCCTCCAA GGAACCTGA CAACCATGT GTACACCAG GGCT[G/A]CC GTGGCCGCG GATGCCAAG CAGTGCTCG AAGATTGGG AG	SEQ ID NO: 626
chr 22	250072 02	A	G	<i>GGT1</i>	p.K52 E	0.008 82	0.002 52	2.23 E- 09	3.52[2.45- 5.05]	CACCAGGGC TGCCGTGGC CGCGGATGC CAAGCAGTG CTCG[A/G]AG ATTGGGAGG TGAGCAGGG CAGGGCATG GGACATGGG CC	SEQ ID NO: 627
chr 22	268799 67	A	G	<i>SRRD</i>	p.R37 R	0.007 11	0.000 00	1.96 E- 08	Inf	CTCGACGGC CGCGGCGGA GGGAGGCGG CGCCCCGGG GGAG[A/G]G AGGCGGCGC CCCGGGGGA GAGAGGCGG CGCCCCGGG GCC	SEQ ID NO: 628
chr 22	299132 78	C	T	<i>THOC 5</i>	p.V52 3M	0.010 05	0.007 28	4.97 E- 02	1.38[1.01- 1.9]	ACTCCTTCA CCTACCATG TAATCCTCA TGGGCAACT GTCA[C/T]CC ATTTCACCA GGCGAGAGA CAACCTTGG CAGGGAAGA GG	SEQ ID NO: 629

chr 22	325904 48	C	T	<i>RFPL2</i>	p.R50 H	0.005 88	0.003 78	3.92 E- 02	1.56[1.03- 2.35]	GGGCCTTTT ATTGGTGAG ATTCCCACC TCCCACTGG GTCA[C/T]GC CCTCCACA CCCTCTAAC CTGATGAGG CTTTGATTTA A	SEQ ID NO: 630
chr 22	325904 71	G	A	<i>RFPL2</i>	p.I42I	0.005 88	0.003 01	3.59 E- 03	1.96[1.29- 2.97]	CACCTCCCA CTGGGTGAC GCCCTTCCA CACCTCTA ACCT[G/A]AT GAGGCTTTG ATTTAATTAT AACAGGGAA TTAGGTTTTT	SEQ ID NO: 631
chr 22	381203 59	C	G	<i>TRIOB P</i>	p.T599 R	0.008 58	0.000 01	4.23 E- 50	966. 49[1 32.3 8- 7056 .38]	AGAGCCTCC TCTCCCAAT AGAGCTACA CGAGACAAC CCCA[C/G]AA CATCCTGTG CCCAGCGGG ACAATCCCA GAGCCTCCA GA	SEQ ID NO: 632
chr 22	381208 24	C	T	<i>TRIOB P</i>	p.P754 L	0.021 08	0.000 01	3.86 E- 124	2405 .39[3 34.9 2- 1727 5.56]	CGAGACAAC CCCAGAACA TCCTGTGCC CAGCGGGAC AATC[C/T]CA GAGCCTCCT CTCCTAACA GAACCATCC AACAAGAGA AC	SEQ ID NO: 633
chr 22	381224 49	G	T	<i>TRIOB P</i>	p.G12 96W	0.026 23	0.000 00	4.02 E- 141	Inf	GGCCCAGAG ACAGCCAGG GCCCCAGGC GCAGTGCAG CAGC[G/T]GG GGCCGCACC CACAGCCCT GGCCGTGCA GAGGTGGAG CG	SEQ ID NO: 634
chr 22	425646 89	G	A	<i>TCF20</i>	p.S195 IS	0.015 44	0.000 00	1.36 E- 91	Inf	ACTGCCCCC CTCACCCCC GCTCCGACT GCTCTGTGC TGAG[G/A]CT GCCTTTCGC GGTCTTGTT TGCAAGGGG GGGAGAGGG C	SEQ ID NO: 635

chr 22	466578 81	T	C	<i>PKDR</i> <i>EJ</i>	p.R447 G	0.006 51	0.002 96	2.21 E- 03	2.21[1.33- 3.47]	ATGTGTGCT ATGGCTTTT GGTCCTTGG AGCACGTGG ACCC[T/C]CT TATCAGAAA ACGCTGTCC TAGAGTCCT TCCGAATCA CC	SEQ ID NO: 636
chr 22	503153 63	C	A	<i>CREL</i> <i>D2</i>	p.D18 2E	0.035 54	0.027 77	4.33 E- 03	1.29[1.09- 1.53]	ACATGGGGT ACCAGGGCC CGCTGTGCA CTGACTGCA TGGA[C/A]GG CTACTTCAG CTCGCTCCG GAACGAGAC CCACAGCAT CT	SEQ ID NO: 637
chr 22	507212 52	T	C	<i>PLXN</i> <i>B2</i>	p.M95 9V	0.009 31	0.006 34	2.76 E- 02	1.47[1.06- 2.04]	TTGGGCACG GGGGACCCC CCGTAGGAG ACCTCCAGA AGCA[T/C]CT GGCCCCGTG TCGCCTGGG GGCCAGTGA CACACTGGA GC	SEQ ID NO: 638
chr 3	126965 8	G	A	<i>CNTN</i> <i>6</i>	p.K11 3K	0.007 84	0.005 02	1.83 E- 02	1.57[1.1- 2.24]	GCCTGGCCA CCAATCTTCT GGGGACAAT TCTGAGTCG GAA[G/A]GC AAAGCTCCA ATTTGCATG TGAGTTTGG GGTAAATTT TG	SEQ ID NO: 639
chr 3	109768 31	C	T	<i>SLC6A</i> <i>11</i>	p.C564 C	0.005 15	0.003 17	3.50 E- 02	1.63[1.05- 2.53]	ATGGCATTG GCTGGCTCA TGGCCCTGT CCTCCATGC TCTG[C/T]AT CCCGCTCTG GATCTGCAT CACAGTGTG GAAGACGGA GG	SEQ ID NO: 640
chr 3	147246 64	C	T	<i>C3orf2</i> <i>0</i>	p.L26 L	0.009 80	0.006 11	5.92 E- 03	1.61[1.17- 2.22]	ACAGGTTTC AGCAGCAGT CCATCCACC TGCTGACGG AGCT[C/T]CT CAGACTGAA GATGAAGGC CATGGTGGA GTCTATGTC GG	SEQ ID NO: 641

chr 3	324094 08	C	T	<i>CMTM</i> 8	p.A12 2A	0.008 82	0.005 74	1.57 E- 02	1.54[1.1- 2.16]	TGTGCTTTA ACGGCAGTG CCTTCGTCTT GTACCTCTCT GC[C/T]GCTG TTGTAGATG CATCTTCCGT CTCCCCTGA GAGGGACA	SEQ ID NO: 642
chr 3	367800 80	C	T	<i>DCLK</i> 3	p.R24 Q	0.012 50	0.009 21	4.43 E- 02	1.36[1.02- 1.81]	TGGAGAAGG GGCACGGCT GTGCTGGGC CAGTGTGAG GGCC[C/T]GG GCTTTGTTG GGGTACAGT TCTTCTACA GCCACCTGA AT	SEQ ID NO: 643
chr 3	383476 80	C	T	<i>SLC22</i> <i>A14</i>	p.L55F	0.009 56	0.006 64	3.16 E- 02	1.44[1.04- 1.99]	GAGGGCTGT CCACACCAA GCAGGATGA CAAGTTTGC CAAC[C/T]TC CTGGATGCG GTGGGGGAG TTTGGCACA TTCCAGCAG AG	SEQ ID NO: 644
chr 3	386718 40	G	A	<i>SCN5A</i>	p.H11 8H	0.005 88	0.002 94	2.12 E- 03	2.01[1.33- 3.05]	ATGAGTGAA CCAGAATCT TCACAGCCG CTCTCCGGA TGGG[G/A]TG GAAGGGACT GAGGACATA CAAGGCGTT GGTGGCACT GA	SEQ ID NO: 645
chr 3	419493 48	G	A	<i>ULK4</i>	p.P391 S	0.008 58	0.005 74	2.71 E- 02	1.5[.06- 2.11]	TAGGAAGAA AATTTCCCA AGTCTGCTC ACCTTGGTC AGAG[G/A]A GAAGTCTTC TGTGGTGAA CAGTGAGTC ATATCCTCA CCA	SEQ ID NO: 646
chr 3	427750 60	G	A	<i>CCDC</i> <i>13</i>	p.R471 R	0.007 11	0.000 08	8.25 E- 35	88.7 2[41. 97- 187. 53]	CTGGGTCCT CCAGGAACT GGGTATAGG CAGGGCTGA CCTC[G/A]CG GCCACTGGA CCCCTCACC CACTCCTTTA TTCCGAAGA T	SEQ ID NO: 647

chr 3	455420 03	C	T	<i>LARS2</i>	p.A56 4A	0.006 86	0.003 41	1.03 E- 03	2.02[1.38- 2.97]	GGATGCCTG TGGATTTGT ACATTGGAG GGAAAGAAC ATGC[C/T]GT CATGCACTT GTTCTATGC AAGATTCTT TAGTCATTTT T	SEQ ID NO: 648
chr 3	460629 22	G	A	<i>XCR1</i>	p.S173 L	0.009 31	0.005 97	1.02 E- 02	1.57[1.13- 2.17]	GGAGGTGAG GTACCACGT GAGTTCGGA ATAATCACA GCCC[G/A]AA GAAAGCACC TTGTGGAAG ATGGTGTCTG AGGATGGAG GA	SEQ ID NO: 649
chr 3	464969 10	G	A	<i>LTF</i>	p.A17 4A	0.007 11	0.004 61	2.66 E- 02	1.55[1.06- 2.25]	GGTTGGGGA ACTGTCCTTT ATCTGCACC GGGAACACA GCT[G/A]GCT GAGAAGAAC CTGGCCACA GCTGTTAAA CACAGAGAA G	SEQ ID NO: 650
chr 3	495691 77	G	A	<i>DAG1</i>	p.V41 1V	0.006 37	0.002 96	7.27 E- 04	2.16[1.45- 3.22]	CTGGCCAGA TTCGCCCAA CGATGACCA TTCCTGGCT ATGT[G/A]GA GCCTACTGC AGTTGCTAC CCCTCCCAC AACCACCAC CA	SEQ ID NO: 651
chr 3	497288 70	A	G	<i>RNF12</i> 3	p.E32 G	0.009 56	0.006 43	2.22 E- 02	1.49[1.08- 2.06]	CTTTTCTCCC TTCTGACTTG TGGCTCAGG CATTGTGCA GG[A/G]GAA GCTGCTGAA TGACTIONCT GAACCGCAT CTTTTCTCT	SEQ ID NO: 652
chr 3	503345 40	C	T	<i>NAT6</i>	p.V14 II	0.008 09	0.005 29	2.17 E- 02	1.53[1.07- 2.18]	TGGTTCAGC ACCCGTGAC AGGCGGGCA TGGCCACC ACAA[C/T]GG GTGCTGCTT CAAGTGTGG GGTGGGGGC TTAGCAGCA TC	SEQ ID NO: 653

chr 3	520056 65	G	A	<i>ABHD 14B</i>	p.R8C	0.007 60	0.005 26	4.86 E- 02	1.45[1.01- 2.08]	AAGAAGAGG GCCTGGCCC TGCACCTGG ATGGTGCCC TCGC[G/A]CT GCTCCACGC TTGCTGCCA TGCCTGCTG CTGCTGTGC TG	SEQ ID NO: 654
chr 3	525408 42	C	T	<i>STAB1</i>	p.S655 S	0.006 62	0.004 12	2.53 E- 02	1.61[1.09- 2.38]	TGCCCCCGA CCATCCTGC CCATCCTGC CCAAGCACT GCAG[C/T]GA GGAGCAGCA CAAGATTGT GGCGGTGAG CCTCGCCTG CA	SEQ ID NO: 655
chr 3	757862 11	A	G	<i>ZNF71 7</i>	p.S855 P	0.005 15	0.000 06	1.25 E- 14	89.9 2[12. 09- 668. 7]	TTTTCTCCTG TGTGTGTTCT CTGATGTAT ACTGAGGCC TG[A/G]CTTC TGGGAGAAA GTTTTCTAC ATTCATTAC ATCTAAAG	SEQ ID NO: 656
chr 3	757869 42	C	A	<i>ZNF71 7</i>	p.R611 I	0.008 58	0.000 00	1.48 E- 41	Inf	ACATTCATT ACATTCATA GGGTCTTTC CCCTGTGTG AGTT[C/A]TC TTGTGTATCC CAAGGTTTA ACTTATTGA TAAAGGTTT T	SEQ ID NO: 657
chr 3	757872 58	G	T	<i>ZNF71 7</i>	p.P506 T	0.011 52	0.000 00	9.28 E- 35	Inf	TTACAGCGA AAGGTTTTT CCACATTCA TTGCATTCGT AGG[G/T]TTT TTCCCCTGTG TGAGTCCAT TGATGGATA GTGAGGAAT	SEQ ID NO: 658
chr 3	757875 46	G	C	<i>ZNF71 7</i>	p.L410 V	0.027 45	0.000 00	3.86 E- 62	Inf	TAGGGCTTT TCCCCTGTGT GAGTTCTAT GATGTATTG TGA[G/C]GTA TGACTTCTG GCTAAAGGT TTTTCCACAT TCACTACAC	SEQ ID NO: 659

chr 3	757881 58	G	C	<i>ZNF71</i> 7	p.L206 V	0.007 35	0.000 00	3.87 E- 34	Inf	TTGAAGGTT TTCCCTTGTT CATTACATT GAAAAGTCT GCA[G/C]CAG AGTTTGAAT CTTGTGATG CTGAGTAAG ATGTTTCATG A	SEQ ID NO: 660
chr 3	757882 92	T	A	<i>ZNF71</i> 7	p.D16 1V	0.006 13	0.000 42	7.04 E- 12	14.7 5[6.0 4- 35.9 7]	TGTCCTCCCC AGGCTTAAT AGGGAAAAG CATGTTCTG GCAA[T/A]CA TTAAACTGC CCAGGCTTC ATTCCTGAA CTGTTTCCAT T	SEQ ID NO: 661
chr 3	999985 31	G	C	<i>TBC1</i> <i>D23</i>	p.C31 S	0.008 09	0.005 59	4.33 E- 02	1.45[1.02- 2.06]	AGGGAAAAA GATCTTGAA GAAGCTCTG GAAGCAGGA GGTT[G/C]TG ATCTTGAAA CGTTGAGAA ATATAATTC AAGGAAGAC CG	SEQ ID NO: 662
chr 3	113052 314	G	C	<i>WDR5</i> 2	p.P118 5R	0.006 86	0.004 57	4.42 E- 02	1.5[.02- 2.23]	TTCCTCTTCC TTCTTGGA GCATTTATTC TCATGTGCT CA[G/C]GTAT CTTGAGTCT GGGGCTGTC TTCAGATTG AAATCTCC	SEQ ID NO: 663
chr 3	124578 212	C	G	<i>ITGB5</i>	p.E80 Q	0.009 07	0.006 25	3.42 E- 02	1.45[1.04- 2.03]	GGCAGGCTC CTCAGGACA TGGAAGCTG CTGGCTGGG CTCT[C/G]TA TCTCACCTCC ACAGCCATT TTTGACAAG GTTTGCCCTC	SEQ ID NO: 664
chr 3	124646 693	G	A	<i>MUC1</i> 3	p.T66I	0.006 37	0.004 03	3.25 E- 02	1.59[1.07- 2.36]	GGAGGAACT ATGTGTACT AATTATGGG GGGAGCAGG TGAA[G/A]TA GCTGTTGGG AAAGGTGTA TTTGCTGTG GTGCTAGCA GT	SEQ ID NO: 665

chr 3	129196 984	C	T	<i>IFT122</i>	p.R366 W	0.008 33	0.005 73	3.60 E- 02	1.46[1.03- 2.06]	CTATGAGTT GTATTCAGA GGACTTATC AGACATGCA TTAC[C/T]GG GTAAAGGAG AAGATTATC AAGAAGTTT GAGTGCAAC CT	SEQ ID NO: 666
chr 3	132198 097	G	A	<i>DNAJ C13</i>	p.R912 R	0.006 13	0.003 65	1.75 E- 02	1.68[1.12- 2.53]	ATTTATTTCA ATAGTGCAC AGATAAACT TGAACGAGA TAG[G/A]TTG ATTCTCTTCC TTAACAAGT TGATCCTTA ATAAGGTAC	SEQ ID NO: 667
chr 3	132247 160	T	G	<i>DNAJ C13</i>	p.L217 0W	0.006 86	0.004 09	1.27 E- 02	1.68[1.15- 2.47]	GCTCAGATT GTTAAAGCT CTCAAGGCA ATGACTCGA AGTT[T/G]GC AGTATGGAG AACAGGTGA GTCTGCATA GAGTCAACT TT	SEQ ID NO: 668
chr 3	136664 807	A	T	<i>NCK1</i>	p.S139 S	0.011 03	0.008 02	4.08 E- 02	1.38[1.02- 1.86]	AAGTGTTC ATGTGGTAC AGGCTCTTT ACCCATTCA GCTC[A/T]TC TAATGATGA AGAACTTAA TTTCGAGAA AGGAGATGT AA	SEQ ID NO: 669
chr 3	137849 808	G	T	<i>A4GN T</i>	p.P97P	0.008 82	0.005 83	2.16 E- 02	1.52[1.08- 2.13]	TTGCTGACA GGAAGGAAA AAGCTGGGT ATGTGGAGT TTGA[G/T]GG CATCGGTGT GGAATCAGT AAGACCCCT CATAAAGAA CA	SEQ ID NO: 670
chr 3	186953 913	C	T	<i>MASP I</i>	p.P582 P	0.009 07	0.005 90	1.70 E- 02	1.54[1.11- 2.15]	AGATGCCCC AGCCGGCCA CCAGGCCCA GCATGTGGG GGGC[C/T]GG GCCTTCAGG CTCAAGCCT TGGCAGGCA GACAGGCAT AA	SEQ ID NO: 671

chr 3	192980 784	C	T	<i>HRASL</i> S	p.S160 S	0.008 33	0.005 09	7.49 E- 03	1.64[1.16- 2.33]	AATTCTACTT TATAGATGG CATTCCTGC GTCCTTTAC AAG[C/T]GCC AAGTCTGTA TTCAGCAGT AAGGCCCTG GTGAAAATG C	SEQ ID NO: 672
chr 3	195306 289	A	G	<i>APOD</i>	p.F15S	0.009 31	0.005 89	9.46 E- 03	1.59[1.14- 2.2]	GCACTTCCC AAGATGAAA TGCTTGTCCT TCTGCCGCA CCG[A/G]AG AGGCCAGCC AGTGCGGAA AGCAGCAGC AGCAGCATC AC	SEQ ID NO: 673
chr 3	195505 772	C	G	<i>MUC4</i>	p.V42 27L	0.025 74	0.000 00	6.23 E- 146	Inf	GGGGTGGCG TGACCTGTG GATACTGAG GAAAGGCTG GTGA[C/G]AG GAAGAGGGG TGGCGTGAC CTGTGGATG CTGAGGAAG TG	SEQ ID NO: 674
chr 3	195508 178	G	C	<i>MUC4</i>	p.L342 5V	0.009 80	0.000 19	2.06 E- 37	51.1 6[26. 23- 99.7 9]	GCGTGACCG GTGGATGCT GAGGAAGTG CTGGTGACA GGAA[G/C]A GGGGTGGCG TGACCTGTG GATGCTGAG GAAGGGCTA GTG	SEQ ID NO: 675
chr 3	195508 194	T	G	<i>MUC4</i>	p.T341 9T	0.016 42	0.000 43	6.58 E- 58	38.6[24.1 9- 61.5 9]	CTGAGGAAG TGCTGGTGA CAGGAAGAG GGGTGGCGT GACC[T/G]GT GGATGCTGA GGAAGGGCT AGTGACAGG AAGAGGCAT GG	SEQ ID NO: 676
chr 3	195512 294	T	C	<i>MUC4</i>	p.S205 3G	0.015 20	0.000 23	2.79 E- 60	68.2 6[37. 51- 124. 21]	GGAAGAGGC GTGGTGTCA CCTGTGGAT ACTGAGGAA AGGC[T/C]GG TGACAGGAA GAGGGGTGT CCTGACCTG TGGATGCTG AG	SEQ ID NO: 677

chr 3	195512 316	C	G	<i>MUC4</i>	p.Q20 45H	0.011 27	0.000 35	1.51 E- 38	32.7 1[19. 15- 55.8 8]	TGGATACTG AGGAAAGGC TGGTGACAG GAAGAGGGG TGTC[C/G]TG ACCTGTGGA TGCTGAGGA AGTATCGGT GACAGGAAG CG	SEQ ID NO: 678
chr 3	195512 981	G	A	<i>MUC4</i>	p.P182 4S	0.011 52	0.000 03	3.64 E- 54	352[85.4 7- 1449 .66]	TCACCTGTG GATGCTGAG GAAGCGTCG GTGACAGGA AGAG[G/A]G GTGGTGTCA CCTGTGGAT GCTGAGGAA GGGCTGGTG ACA	SEQ ID NO: 679
chr 3	196214 336	C	T	<i>RNF16</i> 8	p.R164 R	0.023 77	0.000 06	6.42 E- 132	388. 53[1 80.3- 837. 21]	GTTCTCATC ACTTTTCAGT TGTTCTTCCA TCGCTCTTCG [C/T]CTTTTT CTGCCTGTCT TTTTTCCTCT TCTTCTCCT CTG	SEQ ID NO: 680
chr 3	196214 338	T	C	<i>RNF16</i> 8	p.R164 G	0.009 56	0.000 00	1.81 E- 57	Inf	TCCTCATCA CTTTTCAGTT GTTCTTCCAT CGCTCTTCG CC[T/C]TTTT TCTGCCTGTC TTTTTTCCTC TTCTTCCTCC TCTGCC	SEQ ID NO: 681
chr 4	265813	A	T	<i>ZNF73</i> 2	p.F277 Y	0.022 30	0.000 05	2.19 E- 124	492. 83[2 00.2- 1213 .19]	TGAGGATGA GGTAATGAT TTTGCCACA TTCTTCACAT GTG[A/T]AGG GTTTCTCTTC AGCATGAAT TCTCTTATGC TTAGTAAG	SEQ ID NO: 682
chr 4	265825	T	C	<i>ZNF73</i> 2	p.E273 G	0.011 52	0.000 00	2.01 E- 68	Inf	AATGATTTT GCCACATTC TTCACATGT GAAGGGTTT CTCT[T/C]CA GCATGAATT CTCTTATGCT TAGTAAGGG TTGAGGACC T	SEQ ID NO: 683

chr 4	265829	C	T	<i>ZNF73</i> 2	p.A27 2T	0.018 14	0.000 00	1.83 E- 107	Inf	ATTTTGCCA CATTCTTCAC ATGTGAAGG GTTTCTCTTC AG[C/T]ATGA ATTCTCTTAT GCTTAGTAA GGGTTGAGG ACCTATTA	SEQ ID NO: 684
chr 4	436337	G	A	<i>ZNF72</i> 1	p.P640 L	0.008 82	0.000 00	4.30 E- 53	Inf	TGATGGGGC AAAGGCTTT GCCACACTC TTCACATTTG TAA[G/A]GTT TCTCCCCAG TGTAATTTT CTTCTGTTGA TTCAGGTC	SEQ ID NO: 685
chr 4	436390	A	G	<i>ZNF72</i> 1	p.F622 F	0.005 88	0.000 01	3.07 E- 34	660. 68[8 9.36- 4884 .86]	TGTAAATTTT CTTCTGTTGA TTCAGGTCC GTGTACCAT AC[A/G]AAGT CTTTGCCAC ACTCTTCAC ATTTGTAAA GTTTCTCTC	SEQ ID NO: 686
chr 4	437293	A	G	<i>ZNF72</i> 1	p.Y32 1Y	0.013 73	0.000 13	1.88 E- 67	103. 33[5 8.4- 182. 84]	ATGTGTAGG GTTTCTCTCC AGTATGAAT TCTCCTATGT AC[A/G]TAAA GGTTTGCGG ACTGTCTAA AGGCTTTGC CACATACTT	SEQ ID NO: 687
chr 4	676125	G	C	<i>MFSD</i> 7	p.S434 R	0.007 11	0.004 18	9.52 E- 03	1.71[1.16- 2.51]	GGCGCCGGT ATGGGGTGT GGAAGAAGA CCGCCAGGA TGCA[G/C]CT GAAGAAGGT GCACAGGCC GGCCATCAG CAGCAGAGA CA	SEQ ID NO: 688
chr 4	138836 9	G	A	<i>CRIPA</i> K	p.A24 T	0.006 86	0.000 06	1.22 E- 34	109. 46[4 7.78- 250. 72]	GGAGTGCCC GCCTGCTCA CACGTGCCC ATGTGGAGT GCCC[G/A]CC TGCTCATGT GCCCATGTG GAGTGCCCG CCTGCTCAC AC	SEQ ID NO: 689

chr 4	138941 7	C	T	<i>CRIPA</i> <i>K</i>	p.P373 L	0.006 37	0.000 03	5.30 E- 35	238. 42[7 2.13- 788. 02]	GAGTGCCCG CCTGCTCAC ACACGTGCC CATGTGGAG TGCC[C/T]GC CTGCTCACA CGTGCCCAT GTGGAGTGC CTGCCTGCT CA	SEQ ID NO: 690
chr 4	180550 2	C	T	<i>FGFR</i> <i>3</i>	p.T338 T	0.007 35	0.003 89	1.52 E- 03	1.9[1 .31- 2.75]	CCTGCACA ACGTCACCT TTGAGGACG CCGGGGAGT ACAC[C/T]TG CCTGGCGGG CAATTCTATT GGGTTTTCTC ATCACTCTG	SEQ ID NO: 691
chr 4	341781 1	C	T	<i>RGS12</i>	p.A14 9V	0.010 29	0.006 78	1.19 E- 02	1.52[1.11- 2.08]	ATCGACAGC CAGGCCCAG CTAGCAGAC GACGTCCTC CGCG[C/T]AC CTCACCCAG ACATGTTCA AGGAGCAGC AGCTGCAGG TA	SEQ ID NO: 692
chr 4	351988 1	C	T	<i>LRPAP</i> <i>1</i>	p.D21 1N	0.005 15	0.003 19	4.80 E- 02	1.62[1.04- 2.51]	AGCTCCGTG TGCCTGCTG TGCAGGACG CTGCCCTTG ATGT[C/T]GC TCAGGTCCG AGGGGCTAA TGACGTTCT CGTGGATTT CT	SEQ ID NO: 693
chr 4	700663 6	G	C	<i>TBC1</i> <i>D14</i>	p.E166 Q	0.006 62	0.004 20	2.71 E- 02	1.58[1.07- 2.33]	AGCCAAGGA GAGGTGGCG GTCCCTTAG CACAGGAGG CTCT[G/C]AA GTGGAGAAC GAAGGTAGA ATGTCTTCTA AAACCAGCG G	SEQ ID NO: 694
chr 4	135457 02	C	G	<i>NKX3-</i> <i>2</i>	p.A11 3P	0.005 15	0.000 00	8.15 E- 28	Inf	CCGAGGCTC AAGGATCCC CCCGCAAGG CCGGCCCCG CTGG[C/G]CC CCCGCGCGT CCGCGCAGC GCCGCCTGC TCTCGTTCTC C	SEQ ID NO: 695

chr 4	165042 91	T	G	<i>LDB2</i>	p.N36 6T	0.020 83	0.000 01	1.23 E- 122	2373 .68[3 30.4 6- 1705 0.06]	CTGGTGCCG ATCATCTTAT TGGGAAGCC TGGGGTGGG GGG[T/G]TTT CTGATTTGG TCTCTTGAGT GGCGGGAGG TTTACTGTT	SEQ ID NO: 696
chr 4	577972 65	A	G	<i>REST</i>	p.I747 M	0.010 05	0.000 00	4.04 E- 60	Inf	CTCCTCCCAT GGAGGTGGT CCAGAAGGA GCCTGTTCA GAT[A/G]GA GCTGTCTCCT CCCATGGAG GTGGTCCAG AAGGAACCT G	SEQ ID NO: 697
chr 4	629360 92	C	A	<i>LPHN 3</i>	p.N12 92K	0.006 831	0.004 163	3.32 E- 02	1.65[1.01- 2.55]	GTGAACAGA ACAGGAATC TGATGAACA AGCTGGTGA ATAA[C/A]CT TGGCAGTGG AAGGGAAGA TGATGCCAT TGTCTTGGA TG	SEQ ID NO: 698
chr 4	694337 63	T	A	<i>UGT2 B17</i>	p.D14 7V	0.009 80	0.006 63	1.90 E- 02	1.48[1.08- 2.04]	CAGCTCACC ACAGGGATT AACGGCATC TGCCAGAAG GACA[T/A]CA AATTTTGAC TCTTGATGTT TTCTCATAA GTTTCTTGTT	SEQ ID NO: 699
chr 4	698747 38	T	C	<i>UGT2 B10</i>	p.T134 A	0.010 78	0.000 39	4.99 E- 40	27.9[18.3 1- 42.5 3]	AACAATGGA ATGCCCACC ATAGGGATC CCATGGTAG ATTG[T/C]CT CGTAGATGC CATTGGCTC CACCATGAG TTATAAAG CT	SEQ ID NO: 700
chr 4	698747 42	G	A	<i>UGT2 B10</i>	p.Y13 2Y	0.011 76	0.000 45	1.22 E- 42	26.1 7[17. 59- 38.9 4]	ATGGAATGC CCACCATAG GGATCCCAT GGTAGATTG TCTC[G/A]TA GATGCCATT GGCTCCACC ATGAGTTAT AAAAGCTCT GG	SEQ ID NO: 701

chr 4	712325 42	C	A	<i>SMR3</i> <i>A</i>	p.S79 Y	0.007 60	0.000 00	8.43 E- 46	Inf	CCCCTTTCTC CACCCATATG GTCCAGGGA GAATCCCAC CAT[C/A]CCC TCCTCCACC CTATGGTCC AGGGAGAAT TCAATCACA C	SEQ ID NO: 702
chr 4	723385 89	A	G	<i>SLC4A</i> <i>4</i>	p.K60 2R	0.007 11	0.003 50	7.37 E- 04	2.04[1.4- 2.98]	TCCTCTCTGA TTAGCTTCAT CTTTATCTAT GATGCTTTC A[A/G]GAAG ATGATCAAG CTTGCAAGT TACTACCCC ATCAACTCC	SEQ ID NO: 703
chr 4	772045 70	C	T	<i>FAM4</i> <i>7E</i>	p.R283 C	0.010 29	0.006 72	1.10 E- 02	1.54[1.12- 2.12]	TTAGTTCCTT GAGAATATG TATATCGGG AAGGAATGT AAA[C/T]GTG CATGTAATA AGACTCCTA TAAAACGAA CTCAAGCAT A	SEQ ID NO: 704
chr 4	797921 48	G	A	<i>BMP2</i> <i>K</i>	p.Q48 1Q	0.012 75	0.000 01	1.06 E- 73	1376 .85[1 90.3- 9961 .91]	AACAGCAAC AGCAGCAGC AGCAACAGC AACAGCAGC AGCA[G/A]C AGCAGCAGC AGCAGCACC ACCACCACC ACCACCACC ACC	SEQ ID NO: 705
chr 4	819672 40	C	T	<i>BMP3</i>	p.T222 M	0.000 49	0.000 62	1.00 E+0 0	0.79[0.19- 3.22]	GCCAAAGAA AATGAAGAG TTCCTCATA GGATTTAAC ATTA[C/T]GT CCAAGGGAC GCCAGCTGC CAAAGAGGA GGTTACCTTT T	SEQ ID NO: 706
chr 4	876662 25	A	G	<i>PTPN1</i> <i>3</i>	p.H86 5R	0.009 31	0.005 28	1.49 E- 03	1.77[1.27- 2.46]	AAGATATGC CAGTACCTG CTGCACCTC TGCTCTTACC AGC[A/G]TAA GTTCCAGCT ACAGATGAG AGCAAGACA GAGCAACCA A	SEQ ID NO: 707

chr 4	876722 35	G	T	<i>PTPN1</i> 3	p.D10 42Y	0.008 82	0.005 50	9.90 E- 03	1.61[1.15- 2.26]	GAGTTTAAA TAGAAGTCC TGAAAGGAG GAAACATGA ATCA[G/T]AC TCCTCATCC ATTGAAGAC CCTGGGCAA GCATATGTT CT	SEQ ID NO: 708
chr 4	877491 62	G	A	<i>SLC10</i> <i>A6</i>	p.H24 9Y	0.028 68	0.000 00	2.75 E- 171	Inf	AGTTTATGG ATAGTTTAA CTATACCTTT GCCAAGACT GGT[G/A]GGT AAAAAGTGC CAGCAGAAA ACCCGTGAC ATGGCCAAT C	SEQ ID NO: 709
chr 4	885375 52	C	T	<i>DSPP</i>	p.S124 6S	0.010 78	0.000 00	3.86 E- 52	Inf	AAAGCAGCG ACAGCAGTG ACAGCAGCG ATAGCAGTG ACAG[C/T]AG CAACAGCAG TGACAGCAG CGACAGCAG TGATAGCAG TG	SEQ ID NO: 710
chr 4	885375 58	C	T	<i>DSPP</i>	p.N12 48N	0.011 52	0.000 22	5.36 E- 43	54.1 4[28. 06- 104. 46]	GCGACAGCA GTGACAGCA GCGATAGCA GTGACAGCA GCAA[C/T]AG CAGTGACAG CAGCGACAG CAGTGATAG CAGTGACAG CA	SEQ ID NO: 711
chr 4	113303 632	A	G	<i>ALPK1</i>	p.Q67 R	0.011 76	0.007 35	2.92 E- 03	1.61[1.2- 2.15]	GCAAAGGAA ATGAAGTGG CCCTTCGTG CCTGAAAAG TGGC[A/G]GT ACAAACAAG CCGTGGGCC CAGAGGACA AAACAAACC TG	SEQ ID NO: 712
chr 4	115997 346	T	C	<i>NDST4</i>	p.I283 V	0.012 75	0.009 37	3.24 E- 02	1.37[1.03- 1.81]	AGCCTCTTC CCTGACAAG AAGGAGATG GCATCTATG AAGA[T/C]GA GCTTGTGCA GCCAAAAGT TCAAGTTGT TGCCAAAAA GT	SEQ ID NO: 713

chr 4	125592 869	G	A	<i>ANKR</i> <i>D50</i>	p.A52 1A	0.011 27	0.008 16	3.46 E- 02	1.39[1.03- 1.87]	CATTATCTA ATAATGTCC GAATGGAAT CCTCTCTTTC TAA[G/A]GCT TGTCGAACT ATGCATGAT GTGCGATCG TCTTCACTGT	SEQ ID NO: 714
chr 4	153690 727	G	A	<i>TIGD4</i>	p.T477 I	0.005 88	0.003 39	1.38 E- 02	1.74[1.15- 2.63]	ATCTTGACTT CTGAGAAAT TTTTTCAGA GTATCTAAA GCA[G/A]TTA TTGCCTCAG ATTTTGATG GTAAAGGGA GTTCAGTTC C	SEQ ID NO: 715
chr 4	165962 490	A	T	<i>TRIM6</i> <i>0</i>	p.E422 D	0.006 37	0.003 89	2.13 E- 02	1.64[1.11- 2.45]	TAGTAAAAC CCAGTAAAA TTGGTATTTT TCTGGACTA TGA[A/T]TTG GGTGATCTT TCCTTTTATA ATATGAATG ATAGGTCTA	SEQ ID NO: 716
chr 4	166300 524	T	C	<i>CPE</i>	p.F51L	0.005 15	0.000 00	4.59 E- 30	Inf	GAGGCGGCG CCGCGGGCT GCAGCAAGA GGACGGCAT CTCC[T/C]TC GAGTACCAC CGCTACCCC GAGCTGCGC GAGGCGCTC GT	SEQ ID NO: 717
chr 4	167656 074	A	T	<i>SPOC</i> <i>K3</i>	p.X31 7R	0.003 93	not found	5.64 E- 08	Inf	TTAGAAATG TAGAATTTA TTGATTTCA ACTGTCATC AATC[A/T]AA TGTATACAT CATGGTCAT CACCACCAT CATCATCAT CC	SEQ ID NO: 718
chr 4	170671 841	C	G	<i>C4orf2</i> <i>7</i>	p.G82 R	0.005 15	0.003 23	4.94 E- 02	1.6[.03- 2.48]	TTCTTCGTTT TATGTTTTC AGCAAGGAT ATCATAAGG AC[C/G]AACT AATTGAAGT CCAAGGCTT GCAGAAAGT GAATCTATA	SEQ ID NO: 719

chr 4	175898 879	T	C	<i>ADAM</i> 29	p.W73 5R	0.006 37	0.000 19	1.43 E- 25	33.5 3[18. 85- 59.6 3]	TCAGCGTCG ACCTCATGA GTTACCTCC CCAGAGTCA ACCT[T/C]GG GTGATGCCT TCCCAGAGT CAACCTCCT GTGACGCCT TC	SEQ ID NO: 720
chr 4	175898 947	C	T	<i>ADAM</i> 29	p.S757 S	0.006 62	0.000 52	7.49 E- 19	12.9 1[8.1 6- 20.4 2]	CTGTGACGC CTTCCCAGA GTCATCCTC AGGTGATGC CTTC[C/T]CA GAGTCAACC TCCTGTGAC ACCTCCCA GAGTCAACC TC	SEQ ID NO: 721
chr 4	177083 272	G	A	<i>WDR1</i> 7	p.D93 3N	0.006 86	0.004 05	1.23 E- 02	1.7[1 .16- 2.49]	GCACAAAGT CAGTAAAGA ACTGGCAGA ATGGTATTTT CAA[G/A]ATG GTCGAGCAG TACTAGCCG CATGTTGCC ATCTTGCCA T	SEQ ID NO: 722
chr 5	191718	C	G	<i>LRRC1</i> 4B	p.A22 G	0.008 82	0.000 00	8.74 E- 53	Inf	TTCATTTCTG CAGAAGCTC TGGTGTCCC ACCCCCAGG TGG[C/G]CCG GCAGAGCCT GGACAGCGT GGCCCCAAA CCTCTACCC A	SEQ ID NO: 723
chr 5	891400	T	C	<i>BRD9</i>	p.K39 R	0.000 90	0.001 70	5.63 E- 09	10[5 -21]	CCGTGTAC AGTGCTCCC TCTCTCGTT CCGCTTCTTC TC[T/C]TCCT GGCGGCAG AGTCAAGGG AGTGAGAAA GGCAGGAGT	SEQ ID NO: 724
chr 5	739660 2	T	G	<i>ADCY</i> 2	p.F65 V	0.008 58	0.000 00	3.47 E- 49	Inf	GCTCATCGT CATGGGCTC CTGCCTCGC CCTGCTCGC CGTC[T/G]TC TTCGCGCTC GGGCTGGTG AGTGGCCTC CCCGCGGGT CC	SEQ ID NO: 725

chr 5	369854 42	G	A	<i>NIPBL</i>	p.G72 0G	0.005 64	0.000 01	9.80 E- 33	628. 63[8 4.88- 4655 .98]	GTGAAAGCC GGCCTGAGA CTCCAAAAC AAAAGAGTG ATGG[G/A]CA TCCTGAAAC CCCCAAACA GAAGGGTGA TGGAAGGCC TG	SEQ ID NO: 726
chr 5	523473 66	A	C	<i>ITGA2</i>	p.T252 T	0.008 58	0.005 69	2.65 E- 02	1.51[1.07- 2.13]	CATCCCAGA CATCCCAAT ATGGTGGGG ACCTCACAA ACAC[A/C]TT CGGAGCAAT TCAATATGC AAGGTAAGT TTTGGTGCT AA	SEQ ID NO: 727
chr 5	550836 98	G	T	<i>DDX4</i>	p.A19 9S	0.005 39	0.000 01	2.43 E- 31	603. 79[8 1.37- 4480 .44]	GCAACTTAA CTTCTAGGC GGCTTTTCTC CTACCAATT TTG[G/T]CTC ATATGATGC ATGATGGAA TAACTGCCA GTCGTTTTA A	SEQ ID NO: 728
chr 5	708062 31	C	A	<i>BDP1</i>	p.G11 04G	0.008 09	0.000 00	1.09 E- 48	Inf	TGGAAGAAA CTGAAAGAG AAATATCCC CACAGGAAA ATGG[C/A]CT AGAGGAGGT TAAGCCTCT AGGTGAAAT GCAAACAGA TT	SEQ ID NO: 729
chr 5	715167 95	G	C	<i>MRPS 27</i>	p.Q39 6E	0.005 15	0.003 06	3.05 E- 02	1.68[1.08- 2.62]	GCTTTCTGA GCCTGGTAC TCCTGCTTCG CTTGCTCCCT CT[G/C]TTGC TGTTCTCTCT GGATCAACT GTACAAGGT CTAGATGC	SEQ ID NO: 730
chr 5	762495 03	G	A	<i>CRHB P</i>	p.P53P	0.010 29	0.007 39	4.12 E- 02	1.4[1 .02- 1.91]	TCAGCGCCA ACCTGAAGC GGGAGCTGG CTGGGGAGC AGCC[G/A]TA CCGCCGCGC TCTGCGTGA GTCGAGGCT GCCCCGCTC GC	SEQ ID NO: 731

chr 5	762498 52	A	AC	<i>CRHB</i> <i>P</i>	NM_0 01882: exon3: c.176- 2->C	0.006 87	not found	7.44 E- 12	223. 6[46. 4- 1077 .6]	GCTGCAGCC CGGGACTTA TTGCCCAT GCCCTCCTC CCCC[A/AC]G GGTGCCTGG ACATGCTGA GCCTCCAGG GCCAGTTCA CCT	SEQ ID NO: 732
chr 5	767606 20	C	T	<i>WDR4</i> <i>I</i>	p.G61 D	0.005 88	0.003 65	3.47 E- 02	1.62[1.07- 2.44]	ACACACCTG GGCATTCCA CACAACACT AATTCCATC ATCA[C/T]CA GCAGATGCA AATCTGGTT AGGGAGAAA GGGTCAAGA AA	SEQ ID NO: 733
chr 5	798548 26	A	G	<i>ANKR</i> <i>D34B</i>	p.V33 8A	0.005 39	0.003 28	2.71 E- 02	1.65[1.07- 2.54]	AAATATTGT CTGGTTAGA ATCTGGGTC CTGGTCAAC AGGG[A/G]CT TCAATGCAT TGCTGATTT CTTCTGAAA GATAAGATT G	SEQ ID NO: 734
chr 5	899698 80	A	G	<i>GPR98</i>	p.I164 7V	0.010 05	0.006 98	2.82 E- 02	1.44[1.05- 1.98]	GCTTAGTGC CTCTGGATA TTTATATTT TAGGTTCTG AAT[A/G]TAT ATGTTCTTG ATGATGATA TTCCTGAAC TTAATGAGT A	SEQ ID NO: 735
chr 5	899795 68	G	A	<i>GPR98</i>	p.D19 44N	0.009 31	0.005 69	6.02 E- 03	1.64[1.18- 2.28]	TATCACTGT GGAGATATT GCCTGACGA AGACCCAGA ACTG[G/A]AT AAGGCATTC TCTGTGTCA GTCCTCAGT GTTTCCAGT GG	SEQ ID NO: 736
chr 5	929210 20	C	T	<i>NR2F1</i>	p.H97 H	0.029 66	0.000 04	1.83 E- 169	838. 44[3 09.5 1- 2271 .28]	TCGAGTGCG TGGTGTGCG GGGACAAGT CGAGCGGCA AGCA[C/T]TA CGGCCAATT CACCTGCGA GGGCTGCAA AAGTTTCTTC A	SEQ ID NO: 737

chr 5	134002 614	G	C	<i>SEC24</i> <i>A</i>	p.A22 3P	0.013 48	0.000 01	5.18 E- 79	1523 .15[2 10.7 3- 1100 9.22]	TCATGGGGCC CCCTCCAGC TGGAGGCCC ACCCCCAGT GAGG[G/C]CC CTCACGCCC CTGACATCA TCATATAGA GATGTACCC CA	SEQ ID NO: 738
chr 5	137621 421	C	T	<i>CDC2</i> <i>5C</i>	p.R388 Q	0.006 13	0.003 76	2.71 E- 02	1.63[1.09- 2.45]	TCATGGGGCT CATGTCCTTC ACCAGAAGG GCAATCTGC TCC[C/T]GCA GCTGCCGCT CCCCTTCCTG CACTTTGCTC TGGCTTCG	SEQ ID NO: 739
chr 5	140209 170	G	A	<i>PCDH</i> <i>A6</i>	p.R498 R	0.006 62	0.004 29	3.84 E- 02	1.55[1.05- 2.28]	AGGAGAACG CGTGTTGT CCTACTCGC TGGTGGAGC GGCG[G/A]GT GGGCGAGCG CGCGTTGTC GAGCTACAT TTCGGTGCA CG	SEQ ID NO: 740
chr 5	140559 342	T	C	<i>PCDH</i> <i>B8</i>	p.L576 P	0.007 11	0.003 53	8.04 E- 04	2.02[1.38- 2.95]	CTGTACCCG CTGCAGAAT GGCTCCGCG CCCTGCACC GAGC[T/C]GG TGCCCCGGG CGGCCGAGC CGGGCTACC TGGTGACCA AG	SEQ ID NO: 741
chr 5	141336 635	G	A	<i>PCDH</i> <i>12</i>	p.T261 M	0.009 56	0.005 98	7.37 E- 03	1.6[.16- 2.22]	GCCTTGGTC AGGGTCTGT GGCGGTCAG TTTTATGAG AAGC[G/A]TA CCAGGTGCA GCATCTTCTT GGATTTCCTA GTGCCAGTG A	SEQ ID NO: 742
chr 5	141694 021	G	T	<i>SPRY4</i>	p.S218 Y	0.014 31	0.001 26	2.15 E- 28	11.5[7.94- 16.4 1]	GCAGTTGGA GCGGGAGCA GGAGCAGGG GTGGTCAGC GCAG[G/T]AG CCCTCATCG TCCTCATTCG TGCAGTGGT AGAAGATGC C	SEQ ID NO: 743

chr 5	148384 455	T	A	<i>SH3TC</i> 2	p.D12 29V	0.007 35	0.004 02	2.62 E- 03	1.84[1.27- 2.66]	GACCGCTGC TGCCAGGGC CAGAAGGAA GTACTCAGT GGCA[T/A]CA TGGGCATCC TAACCCCGT GGTATGGGG GCAAAGAAG AG	SEQ ID NO: 744
chr 5	149276 063	T	G	<i>PDE6</i> A	p.Q49 2H	0.019 52	0.001 76	8.09 E- 37	11.3 2[8.1 7- 15.5 5]	ATTTATTAAT TTCGTATTTA TCTGCATCT GGCAGCTCC GC[T/G]TGCT GTATAAGGA ATAGAGTCA GGTGATTAG GAAACATGA	SEQ ID NO: 745
chr 5	149301 253	G	A	<i>PDE6</i> A	p.P293 L	0.007 11	0.004 75	3.83 E- 02	1.5[1 .03- 2.18]	CCTGGGACC AGAGTAAGG TGGAACCTC ACCCATCAG AACC[G/A]GC CACACATCA AAAAATTCC TAGGAATGA GAAAAACAA TA	SEQ ID NO: 746
chr 5	149512 494	C	T	<i>PDGF</i> RB	p.V31 6M	0.006 86	0.004 19	1.88 E- 02	1.64[1.11- 2.43]	TCAGCAAAT TGTAGTGTG CCCACCTCT CCCAGGAGC CGCA[C/T]GT AGCCGCTCT CTGCAAGGG GTGACCGTC AGGGGCGGG GC	SEQ ID NO: 747
chr 5	150905 399	G	T	<i>FAT2</i>	p.P347 9Q	0.006 13	0.000 00	4.42 E- 37	Inf	CCTCCTGCTT AGGCCCTCA GCAGTCACC AGCCATCCA TCC[G/T]GGG TCACTCGGA AGGCAGAGC CGTTGTTCCC CTTGGTGAT	SEQ ID NO: 748
chr 5	167689 228	C	A	<i>TENM</i> 2	p.R257 1R	0.005 15	0.003 02	2.93 E- 02	1.71[1.1- 2.66]	CATCATTGG CAAAGGCAT CATGTTTGC CATCAAAGA AGGG[C/A]G GGTGACCAC GGGCGTGTC CAGCATCGC CAGCGAAGA TAG	SEQ ID NO: 749

chr 5	167881 032	A	T	<i>WWC1</i>	p.E862 V	0.011 76	0.000 00	5.03 E- 70	Inf	GAGAATGAG GCAGTAGCC GAGGAAGAG GAGGAGGAG GTGG[A/T]GG AGGAGGAGG GAGAAGAGG ATGTTTTTCAC CGAGAAAGC C	SEQ ID NO: 750
chr 5	168112 707	G	A	<i>SLIT3</i>	p.A11 80A	0.005 64	0.002 62	1.12 E- 03	2.16[1.41- 3.31]	AAGGGCAGG GCAGGGCGG GACACACCT GCAGGGAGA TGTT[G/A]GC CTGGGGTCG GACCTTGGC GGAGGCCAG TTCCACGTA GG	SEQ ID NO: 751
chr 5	171661 166	T	C	<i>UBTD 2</i>	p.A89 A	0.009 07	0.006 28	3.47 E- 02	1.45[1.04- 2.02]	CATGTGGTA ATGTTATGTT TGCACCATC AATGATTGC TTG[T/C]GCC AGTTCATGA TCATTGCTCT CAAAAGCAT GTGCAGCAG	SEQ ID NO: 752
chr 5	178139 385	C	T	<i>ZNF35 4A</i>	p.E498 E	0.020 34	0.002 38	8.40 E- 44	8.69[6.78- 11.1 4]	GATTACTAA GTGATGAGT TACACCTGA ATGTTTTCCC ACA[C/T]TCG TTACATTTAT AGGGTCTTT CTCCAGTAT GCATTCTCT	SEQ ID NO: 753
chr 5	178139 394	T	C	<i>ZNF35 4A</i>	p.K49 5K	0.020 34	0.002 84	4.38 E- 39	7.3[5 .72- 9.32]	GTGATGAGT TACACCTGA ATGTTTTCCC ACACTCGTT ACA[T/C]TTA TAGGGTCTT TCTCCAGTA TGCATTCTCT GATGTTGAA	SEQ ID NO: 754
chr 5	179192 341	A	G	<i>MAML 1</i>	p.T110 T	0.010 54	0.007 57	4.35 E- 02	1.4[1 .03- 1.9]	AAGTCATTC TTTTCAATGT TTTTACGCAT CTTCATGAT AC[A/G]GTTA AGAGGAATC TTGACAGCG CCACTTCCC CTCAGAATG	SEQ ID NO: 755

chr 5	179192 401	C	T	<i>MAML 1</i>	p.Y13 0Y	0.010 54	0.007 56	4.35 E- 02	1.4[1 .03- 1.9]	GCGCCACTT CCCCTCAGA ATGGCGATC AACAGAATG GCTA[C/T]GG GGACCTCTT TCCTGGGCA TAAGAAGAC TCGCCGGGA GG	SEQ ID NO: 756
chr 6	117684 82	C	T	<i>ADTR P</i>	p.T96 T	0.005 53	0.002 61	6.64 E- 03	2.12[1.22- 3.46]	CACATTTCT GTTAGATTA TGTAACACAT CTTTGAAAC TTAC[C/T]GT GGATACAGG AAAAGCCAG AGTGGTGAA AAGCAGGTC TC	SEQ ID NO: 757
chr 6	260322 17	C	T	<i>HIST1 H3B</i>	p.K24 K	0.007 11	0.000 00	8.79 E- 43	Inf	TTTTACGCC GCCGGTAGC CGGCGCGCT CTTGCGAGC AGC[C/T]TTG GTAGCCAGC TGCTTGCGT GGCGCTTTA CCGCCGGTG G	SEQ ID NO: 758
chr 6	294087 21	T	G	<i>OR10C 1</i>	p.M31 0R	0.009 07	0.005 59	7.68 E- 03	1.63[1.17- 2.27]	AAAGCTGCC CTAAAGAGA ACCATCCAG AAAACGGTG CCTA[T/G]GG AGATTTGAA AAGGGGGCG ATAGTGACT TCTGTGCAG TG	SEQ ID NO: 759
chr 6	300389 42	C	T	<i>RNF39</i>	p.L337 L	0.005 88	0.003 72	3.63 E- 02	1.59[1.04- 2.41]	GTACAATGC GGAGCGGAG CACGAGGGT CGCAGGTGC AGAA[C/T]AG CGGGAAGAT GCGCTCCCC CAGGGGGCC AGGCGCCTG GA	SEQ ID NO: 760
chr 6	306732 80	G	A	<i>MDC1</i>	p.A12 27V	0.011 76	0.000 04	4.12 E- 64	264. 73[1 05.3 3- 665. 33]	AGGGGTCTT GACAGAGGA TCTATTTTT CTCCCCCTA GTA[G/A]CCT GAGAGGTGG GTTGAGAGG TGACAGGTC GGTCGGTGG A	SEQ ID NO: 761

chr 6	309171 10	G	A	<i>DPCR</i> <i>1</i>	p.G29 0E	0.020 59	0.000 00	2.77 E- 100	Inf	GAGCTCACA CAATCTCTA GCAGAGCCT ACAGAACAT GGAG[G/A]A AGGACAGCC AATGAGAAC AACACACCA TCCCCAGCA GAG	SEQ ID NO: 762
chr 6	309174 17	T	C	<i>DPCR</i> <i>1</i>	p.T392 T	0.006 37	0.000 08	2.15 E- 25	78.3 3[27. 32- 224. 54]	AGCCTACAG AACATGGAG AAAGGACAG CCAATGAGA ACAC[T/C]AC ACCATCCCC AGCAGAGCC TACAGAACA TGGAGAAAG GA	SEQ ID NO: 763
chr 6	309178 57	A	G	<i>DPCR</i> <i>1</i>	p.E539 G	0.012 25	0.000 00	4.79 E- 60	Inf	ACCCCACTG GCCAATGAG AACACCACA CCATCCCCA GCAG[A/G]G CCTACAGAA AATAGAGAA AGGACAGCC AATGAGAAG ACC	SEQ ID NO: 764
chr 6	309181 60	G	A	<i>DPCR</i> <i>1</i>	p.G64 0E	0.005 64	0.000 13	6.87 E- 21	42.3 5[18. 16- 98.7 4]	GAAAGGACA GCCAATGAG AACACCACA CCATCCCCA GCAG[G/A]G CCTACAGAA AATAGAGAA ATGACAGCC AACGAGAAG ACC	SEQ ID NO: 765
chr 6	309207 55	A	C	<i>DPCR</i> <i>1</i>	p.Y13 48S	0.005 21	0.002 99	4.32 E- 02	1.75[0.98- 2.88]	GTTCTCATTC CTCCTTCTC ATCCCAATC ACAGGTCTC CT[A/C]TATG ATGCGGACA CGCCGCACA CTAACCCAG AACACCCAG	SEQ ID NO: 766
chr 6	309543 27	C	T	<i>MUC2</i> <i>1</i>	p.S125 S	0.013 24	0.000 00	5.07 E- 73	Inf	CAACCTCCA GTGGGGCCA GCACAGCCA CCAACCTCTG AGTC[C/T]AG CACACCCTC CAGTGGGGC CAGCACAGC CACCAACTC TG	SEQ ID NO: 767

chr 6	309544 39	A	G	<i>MUC2</i> <i>1</i>	p.S163 G	0.019 61	0.000 00	1.38 E- 116	Inf	AGCCACCAA CTCTGACTC CAGCACAAC CTCCAGTGA GGCC[A/G]GC ACAGCCACC AACTCTGAG TCCAGCACA ACCTCCAGT GG	SEQ ID NO: 768
chr 6	309956 35	C	T	<i>MUC2</i> <i>2</i>	p.S809 S	0.009 56	0.000 00	5.89 E- 43	Inf	CTACAGTTT CCACCACAG GCTTGAGAG CCACCACCA CTTC[C/T]AC TGAAGGCTC TGAGATGAC TACAGTCTC CACCACAGG TG	SEQ ID NO: 769
chr 6	316916 66	C	A	<i>C6orf2</i> <i>5</i>	p.G10 4G	0.005 39	0.002 60	2.87 E- 03	2.08[1.35- 3.22]	TCCGGCGGC TGGAGCTCC TCTTGAGCG CGGGGGACT CGGG[C/A]AC TTTTTTCTGC AAGGGCCGC CACGAGGAC GAGAGCCGT A	SEQ ID NO: 770
chr 6	317368 35	C	T	<i>VWA7</i>	p.R488 Q	0.005 39	0.002 54	1.64 E- 03	2.13[1.38- 3.29]	CAGGGCAGC CATGCTCTC CCCAACAAT GGCTGCCAC GTCT[C/T]GA ATGTGCTGG TCTTTGGTG AAGATCACCC TCTCCTCCTG A	SEQ ID NO: 771
chr 6	326342 82	A	G	<i>HLA-DQB1</i>	p.S35P	0.007 48	0.004 33	1.76 E- 02	1.73[1.09- 2.64]	TGGCGGCTC TGGAGAGCA GCTGCCCTG CACTTACCG GGAG[A/G]G TCTCTGCCCT CAGCCAGTA GGGAGCTCA GCATCGCCA GC	SEQ ID NO: 772
chr 6	327136 19	C	A	<i>HLA-DQA2</i>	p.P128 H	0.006 62	0.000 00	1.10 E- 39	Inf	GTCACAGTG TTTTCCAAGT TTCCTGTGA CGCTGGGTC AGC[C/A]CAA CACCTCAT CTGTCTTGTG GACAACATC TTTCCTCCT	SEQ ID NO: 773

chr 6	327140 58	T	G	<i>HLA-DQA2</i>	p.L219 V	0.020 59	0.000 01	4.06 E- 120	2275 .46[3 16.7 4- 1634 6.8]	GCCTGAGAT TCCAGCCCC TATGTCAGA GCTCACAGA GACT[T/G]TG GTCTGCGCC CTGGGGTTG TCTGTGGGC CTCATGGGC AT	SEQ ID NO: 774
chr 6	327141 08	C	G	<i>HLA-DQA2</i>	p.G23 5G	0.012 75	0.000 00	1.19 E- 75	Inf	CCCTGGGGT TGCTGTGG GCCTCATGG GCATTGTGG TGGG[C/G]AC TGTCTTCATC ATCCAAGGC CTGCGTTCA GTTGGTGCT T	SEQ ID NO: 775
chr 6	327141 11	T	C	<i>HLA-DQA2</i>	p.T236 T	0.012 50	0.000 00	3.37 E- 74	Inf	TGGGGTTGT CTGTGGGCC TCATGGGCA TTGTGGTGG GCAC[T/C]GT CTTCATCATC CAAGGCCTG CGTTCAGTT GGTGCTTCC A	SEQ ID NO: 776
chr 6	327141 17	C	G	<i>HLA-DQA2</i>	p.F238 L	0.016 91	0.000 00	4.00 E- 100	Inf	TGTCTGTGG GCCTCATGG GCATTGTGG TGGGCACTG TCTT[C/G]AT CATCCAAGG CCTGCGTTC AGTTGGTGC TTCCAGACA CC	SEQ ID NO: 777
chr 6	328200 00	C	A	<i>TAP1</i>	p.V30 4L	0.005 39	0.002 45	1.19 E- 03	2.21[1.43- 3.42]	TGCACGTGG CCCATGGTG TTGTTATAG ATCCCGTCA CCCA[C/A]GA ACTCCAGCA CTGCACTAT AAAGAACCC GGAAAAAAA GG	SEQ ID NO: 778
chr 6	333658 08	G	T	<i>KIFC1</i>	p.R5S	0.005 64	0.003 49	3.11 E- 02	1.62[1.06- 2.47]	CTCCTGGGT ATTGTCTTA AGGGTCTCT TTTCCCAAC AGAG[G/T]TC CCCCCTATT GGAAGTAAA GGGGAACAT AGAAGTGAA GA	SEQ ID NO: 779

chr 6	340039 28	C	T	<i>GRM4</i>	p.S520 S	0.005 39	0.003 40	4.08 E- 02	1.59[1.03- 2.45]	AGCTGATGC TCATCCCTA GTCCCAGGA AGATTCGGC GCAG[C/T]GA GCAGGTGCC AAGGTCGGG CTCAGCGAT CATGAGGAA GG	SEQ ID NO: 780
chr 6	357150 76	C	T	<i>ARMC 12</i>	p.I188I	0.005 15	0.002 81	1.08 E- 02	1.84[1.18- 2.86]	AGGAACACT CCATCAAAG TACTCGAAC TGATCTCCA CCAT[C/T]TG GGACACGGA ACTGCACAT TGCGGGCCT CAGACTCCT CA	SEQ ID NO: 781
chr 6	367100 50	T	A	<i>CPNE 5</i>	p.I593 F	0.006 62	0.000 00	5.82 E- 39	Inf	CCCAGGCCC CAGCCACCT GCCTGCTGA GACCAGGTT CAGA[T/A]GT GCGTGTGCA GGGGGGACG CAGGGGGCG TGCGGGCTG GG	SEQ ID NO: 782
chr 6	392828 16	G	A	<i>KCNK 16</i>	p.Q25 1X	0.007 84	0.004 49	4.28 E- 03	1.75[1.23- 2.51]	CCTCAGCTT CCCAGTCCT TTCTTGGAT ATGGGGAAG TCCT[G/A]GG GTGTGACTT GGACTCCTC TTGCTGCTGT AGAGCCTCT C	SEQ ID NO: 783
chr 6	441438 62	G	A	<i>CAPN 11</i>	p.A29 7T	0.005 21	0.002 82	2.43 E- 02	1.85[1.04- 3.06]	ACTGGAATC CATGACTGA CAAGATGCT GGTGAGAGG GCAC[G/A]CT TACTCTGTG ACTGGCCTT CAGGATGTG AGTCCTGAG AA	SEQ ID NO: 784
chr 6	466559 01	C	G	<i>TDRD 6</i>	p.A12 A	0.012 01	0.000 00	4.48 E- 58	Inf	TCAAGATGT GCTCGACGC CCGGAATGC CGGCGCCGG GGGC[C/G]TC GCTGGCCCT GCGGGTGTC CTTCGTGGA CGTGCATCC CG	SEQ ID NO: 785

chr 6	560330 94	G	A	<i>COL21A1</i>	p.T343 M	0.067 40	0.071 69	3.19 E- 01	0.94[0.83- 1.06]	TACTAAGAG ACGAATTTG GTGCCAGCC TTCATCAAA CAAC[G/A]TC TACAAAAAG AAAGTGTGG AAGATTCAT AAATAAAGC CC	SEQ ID NO: 786
chr 6	767318 54	G	A	<i>IMPG1</i>	p.N13 7N	0.010 78	0.007 76	3.74 E- 02	1.39[1.03- 1.89]	AACTCTAGG AACTTCTTA CTGTTGTAG GCATCTTGG TGTC[G/A]TT GAGTGTATT ATCGAGAAT TTCATTGAG GAGGGTGTC AT	SEQ ID NO: 787
chr 6	843032 30	T	C	<i>SNAP91</i>	p.T553 A	0.010 78	0.007 26	1.83 E- 02	1.49[1.09- 2.04]	AAATTACCA CCAAAGATA TCTAGAGCA GGAGGAGCA GTGG[T/C]GG CGGTGGCAG CGGAGGTGG TGGTAGTGG TGGTGGCAG CG	SEQ ID NO: 788
chr 6	854737 58	C	T	<i>TBX18</i>	p.G48 R	0.414 71	0.494 51	5.64 E- 23	0.72[0.68- 0.77]	GCGCCGCCG CCGCGGCTG CAGCCTCCG TCGTCCACG GCCC[C/T]CG CCGCCTCTT GGCGCCCAG TTTTCGCCGC TTCTTCTGA	SEQ ID NO: 789
chr 6	861950 33	G	A	<i>NT5E</i>	p.V27 8I	0.007 60	0.004 66	1.07 E- 02	1.64[1.14- 2.35]	ATTCATAGT CACTTCTGA TGATGGGCG GAAGGTTCC TGTA[G/A]TC CAGGCCTAT GCTTTTGGC AAATACCTA GGCTATCTG AA	SEQ ID NO: 790
chr 6	905721 38	G	A	<i>CASP8AP2</i>	p.G23 7D	0.005 39	0.003 19	2.39 E- 02	1.69[1.1- 2.61]	AATGGTGTT TGGTCACGT TCTCATTATC AGGTTGGCG AGG[G/A]TA GCTCAAATG AGGATAGTA GAAGAGGAA GAAAAGATA TT	SEQ ID NO: 791

chr 6	108882 487	A	T	<i>FOXO</i> 3	p.S26 C	0.005 39	0.000 26	2.37 E- 17	20.8 1[10. 92- 39.6 5]	TCCGCTCGA AGTGGAGCT GGACCCGGA GTTCGAGCC CCAG[A/T]GC CGTCCGCGA TCCTGTACG TGGCCCCTG CAAAGGCCG GA	SEQ ID NO: 792
chr 6	109867 190	T	C	<i>AK9</i>	p.E103 5E	0.023 28	0.000 08	6.49 E- 127	295. 82[1 49.2 3- 586. 43]	CGTTCTCAG AATCTTCCTC AAATTCAGG TCCCACTTTC TT[T/C]TCAG TTTTGAGTA GTAGTTTTTC TTGAAGAAC TTCTTCAA	SEQ ID NO: 793
chr 6	126073 212	T	G	<i>HEY2</i>	p.L74 L	0.005 88	0.003 72	3.66 E- 02	1.58[1.05- 2.39]	GGGATCGGA TAAATAACA GTTTATCTG AGTTGAGAA GACT[T/G]GT GCCAACTGC TTTTGAAAA ACAAGTAAG CTATCCCCTC C	SEQ ID NO: 794
chr 6	136597 174	G	A	<i>BCLA</i> <i>F1</i>	p.P497 S	0.005 64	0.002 29	2.13 E- 04	2.47[1.61- 3.78]	TCAAAGAGG TCTTTGAGCT TTTCAGACTT TACCTGCTC AG[G/A]TGAC TGAGTTTCTT TCTTTACTGT TATTCTTTCA GAATTT	SEQ ID NO: 795
chr 6	136597 456	C	A	<i>BCLA</i> <i>F1</i>	p.E403 X	0.008 58	0.004 71	1.21 E- 03	1.83[1.3- 2.58]	AGGACTGAC TTCCTGAAC TGCTATAA TCCTCTGTCT CCT[C/A]TGT GTCATCCCC TTCTGAATC ATTAAACTT TTGTTTCCA	SEQ ID NO: 796
chr 6	137814 936	G	T	<i>OLIG3</i>	p.I124I	0.024 51	0.000 01	1.52 E- 144	2806 .41[3 91.3 8- 2012 3.7]	TGAGCATGA GGATGTAGT TTCTGGCGA GCAGGAGTG TGGC[G/T]AT CTTGGAGAG CTTGCGCAC CGACGGCCC ATGCGCGTA GG	SEQ ID NO: 797

chr 6	139113 926	A	T	CCDC 28A	p.T271 S	0.008 14	0.002 15	7.18 E- 08	3.81[2.41- 5.78]	ACAAAAACT CCATTTGGC AGATGCACA AGATGTTCC AAAT[A/T]CT TCTGCTAGC TAAAATGAA ATGTAGTTT GCTTTCTTGT G	SEQ ID NO: 798
chr 6	152457 795	C	T	SYNE1	p.E853 9E	0.008 36	0.001 60	6.73 E- 08	5.3[3 .21- 8.74]	GGCACTGCA TCAGGGCAT CCTGCAGCA GGCCCCGCC ACTC[C/T]TC CAGCAGAGA GCACACTCG GTCCAGCG CCCATTCT CT	SEQ ID NO: 799
chr 6	155143 502	A	G	SCAF8	p.T629 A	0.005 64	0.003 54	3.35 E- 02	1.59[1.05- 2.43]	TCAGAGCCC AACTCCAGT TGAAAAGGA GACAGTGGT CACA[A/G]CC CAGGCAGAG GTTTTCCCTC CTCCTGTTGC TATGTTGCA	SEQ ID NO: 800
chr 6	158487 551	T	C	SYNJ2	p.M29 7T	0.009 31	0.005 94	9.88 E- 03	1.57[1.13- 2.18]	CAGTCCGAA TTCACAAAT TTCAAGCGG ATCCGGATT GCTA[T/C]GG GGACCTGGA ACGTGAACG GAGGAAAGC AGTTCCGGA GC	SEQ ID NO: 801
chr 6	167728 725	T	C	UNC9 3A	p.Y38 7H	0.007 35	0.000 97	1.30 E- 15	7.62[5.08- 11.4 4]	CGTTCTGTTT GAGAAGAGC AAGGAAGCT GCCTTCGCC AAT[T/C]ACC GCCTGTGGG AGGCCCTGG GCTTCGTCA TTGCCTTCG G	SEQ ID NO: 802
chr 7	331061	A	G	WI2- 2373II .2	p.K10 3K	0.015 69	0.000 04	1.78 E- 52	377. 58[5 2.37- 2722 .42]	GCGCGCAGG TGCCGCGGT CCGAGGGCC ACGAGAAGG GCAA[A/G]G GCAACTACT GGACGTTTCG CGGGCGGCT GCGAGTCGC TGC	SEQ ID NO: 803

chr 7	102700 7	G	A	<i>CYP2 W1</i>	p.R328 H	0.023 77	0.000 02	9.98 E- 133	1179 .78[2 90.7 9- 4786 .53]	CCACCCTTT GCCCCAGGC CGGGTGCAG GAGGAGCTA GACC[G/A]CG TGCTGGGCC CTGGGCGGA CTCCCCGGC TGGAGGACC AG	SEQ ID NO: 804
chr 7	102837 6	C	T	<i>CYP2 W1</i>	p.P464 L	0.010 78	0.007 56	2.76 E- 02	1.43[1.05- 1.94]	CTGCAGAGG TACCGCCTG CTGCCCCCG CCTGGCGTC AGTC[C/T]GG CCTCCCTGG ACACCACGC CCGCCCCGG CTTTTACCAT G	SEQ ID NO: 805
chr 7	178430 8	C	T	<i>ELFN1</i>	p.R26 C	0.006 13	0.003 97	4.27 E- 02	1.55[1.02- 2.34]	CGTGGCGGC CGCCACCCT GCTGCACGC TGGCGGCCT GGCC[C/T]GC GCAGACTGC TGGCTGATC GAGGGCGAC AAGGGCTTC GT	SEQ ID NO: 806
chr 7	225589 3	C	G	<i>MADI L1</i>	p.E236 D	0.007 35	0.004 54	1.35 E- 02	1.62[1.12- 2.35]	CCAGCTCAG ACTTCATGTT CTTCACAAT CGCTGCATC CTG[C/G]TCT TGCAGGGAC AGCTTCTGC TCCAGATCC TGATGGAGG C	SEQ ID NO: 807
chr 7	418545 7	G	A	<i>SDKI</i>	p.P144 4P	0.010 05	0.007 28	4.98 E- 02	1.38[1.01- 1.9]	GCGCCACAG TGAGGCAGT TCACAGCCA CCGACCTGG CCCC[G/A]GA GTCCGCATA CATCTTCAG GCTGTCCGC CAAGACGAG GC	SEQ ID NO: 808
chr 7	485690 4	T	C	<i>RADIL</i>	p.Y56 5C	0.009 07	0.006 17	2.59 E- 02	1.47[1.06- 2.06]	GTGCACCTT GGAGACATA GTAGACGCA CTGCTGGAA GGCG[T/C]AC AGCACCACC TCCTCCAGC ACCGCCATG GCCTCCTCG CT	SEQ ID NO: 809

chr 7	602682 7	G	C	<i>PMS2</i>	p.S523 S	0.006 13	0.003 70	1.88 E- 02	1.66[1.11- 2.49]	CCTGAGAGT CCACATGTT CCTGCGAGC CCCTGTCCC CTGG[G/C]GA GCTGGCCGC ATACTCGCT GCTGCAGTG ACTGCCCGT GT	SEQ ID NO: 810
chr 7	232218 11	A	G	<i>NUPL</i> 2	p.Q36 R	0.005 39	0.000 03	4.20 E- 29	195. 45[5 8.48- 653. 28]	CCCGGTGCT AGGGGTGCA GGAGGAGGA CGGCAGCAA CCGC[A/G]GC AGCAGCCTT CAGGTGACT CTCCTCTGA ATCCTCCGC GG	SEQ ID NO: 811
chr 7	262176 89	A	G	<i>NFE2L</i> 3	p.I233 V	0.001 47	not found	4.03 E- 06	Inf	GGAGAACTC ACTTCAGCA GAATGATGA TGATGAAAA CAAA[A/G]TA GCAGAGAAA CCTGACTGG GAGGCAGAA AAGACCACT GA	SEQ ID NO: 812
chr 7	309219 12	G	T	<i>FAM1</i> 88B	p.R696 S	0.006 62	0.000 02	2.24 E- 37	370. 83[8 8.15- 1559 .92]	GCCTGCAGC CGGGGCTCC TGCGTGACT GGAGGACTG AGAG[G/T]CT CTTTGACTTG TACTACTAC GATGGCCTG GCCAACCAG C	SEQ ID NO: 813
chr 7	379885 90	G	T	<i>EPDR</i> 1	p.G79 W	0.007 11	0.000 04	1.62 E- 37	159. 69[6 1.78- 412. 75]	CATTCTCA AAACTCCAC CTTTGAAGA CCAGTACTC CATC[G/T]GG GGCCTCAG GAGCAGATC ACCGTCCAG GAGTGGTCG GA	SEQ ID NO: 814
chr 7	420072 01	T	C	<i>GLI3</i>	p.I808 M	0.006 86	0.002 79	4.10 E- 05	2.47[1.67- 3.63]	CTGAGCAGA TGCATGGTC TGATGTAGA ACTACCAT TTCC[T/C]AT GAGAGGAGA GACCGCAGG GGCTTTAGG GGGTAGAAT GG	SEQ ID NO: 815

chr 7	441544 53	G	A	<i>POLD</i> 2	p.C447 C	0.006 13	0.004 02	4.50 E- 02	1.53[1.02- 2.29]	CATCGTCCT CTGCCCCGA AGCCCCGAGA AGCTGATGG GCTG[G/A]CA GGCCAGGCT GCGCAGGTT CACAAGGCA GGCGGTCTG CG	SEQ ID NO: 816
chr 7	451239 25	C	T	<i>NACA</i> D	p.K61 8K	0.005 15	0.000 11	6.20 E- 19	47.3 7[17. 85- 125. 69]	CTTCAGCCT GCTGGGACA CAATCGTGG CTGCAGCCA CAGG[C/T]TT TGGGGCTGA TGAGAGATC TGTGTCTTGT AGGGGCAGA G	SEQ ID NO: 817
chr 7	479255 20	C	T	<i>PKD1</i> LI	p.R990 Q	0.009 56	0.006 56	3.03 E- 02	1.46[1.06- 2.02]	TGAAGTGGC AGGTTGGCC AAGGGTCAC GGGTGAAGG TTCC[C/T]GT GAGAATGGT GTGGTCGTT GCATCAGGA TCTGCAGTG CC	SEQ ID NO: 818
chr 7	505717 55	C	A	<i>DDC</i>	p.M23 9I	0.005 39	0.003 34	3.84 E- 02	1.62[1.05- 2.49]	TGTCAAAGG AGCAGCATG TTGTGGTCC CCAGGGTGG CAAC[C/A]AT CTAGAGGGT AAAAAGCAG ACAGCCTTT TATTCCCCA GG	SEQ ID NO: 819
chr 7	506730 32	C	T	<i>GRB10</i>	p.P390 P	0.005 39	0.003 37	3.94 E- 02	1.61[1.04- 2.47]	AGGCGTGCC CCTCCTCCA GGGCTGCGC TCTGGGCCT CTGC[C/T]GG ATTCTCTATC ACGCGTCCT GTTTGCCCA GAAAAATCC A	SEQ ID NO: 820
chr 7	636803 38	C	A	<i>ZNF73</i> 5P	p.G30 3G	0.008 82	0.000 01	1.80 E- 51	989. 53[1 35.6 4- 7219 .03]	TTCATACTG GAGAGAGAC CCTACAAAT GTGAAGAAT GTGG[C/A]AA AGCCTTTAG CGTATCCTC AGCCCTCAT TTACCACAA GA	SEQ ID NO: 821

chr 7	638092 67	C	A	ZNF73 6	p.I342I	0.014 22	0.000 00	8.84 E- 84	Inf	GTAAACATA AGAGAATTC ATACTGGAG AGAAACCCT ACAT[C/A]TG TGAAGAATG TGGCAAAGC CTTACCCG CTCCTCAAC CC	SEQ ID NO: 822
chr 7	871606 45	G	A	ABCB1	p.L884 L	0.007 60	0.005 22	4.73 E- 02	1.46[1.02- 2.1]	CTCACCTTCC CAGAACCTT CTAGTTCTTT CTTATCTTTC A[G/A]TGCTT GTCCAGACA ACATTTTCAT TTCAACAAC TCCTGCT	SEQ ID NO: 823
chr 7	889655 53	C	T	ZNF80 4B	p.T108 6I	0.006 62	0.003 35	1.55 E- 03	1.98[1.34- 2.93]	TTCCCTGGT GCTTTTCCGT CTAATAAAT ATACTGGTG TGA[C/T]TGA TTCAACAGA GACCCAAGA AGACCAAAT AAATCTAGA C	SEQ ID NO: 824
chr 7	916030 56	C	T	AKAP9	p.S27L	0.005 39	0.002 82	6.73 E- 03	1.92[1.24- 2.96]	TTTTCTTAGC TTGCCCAGT TTCGACAAA GAAAAGCTC AGT[C/T]GGA TGGGCAGAG TCCTTCCAA GAAGCAGAA AAAAAAGAG A	SEQ ID NO: 825
chr 7	978223 61	G	A	LMTK 2	p.A86 2T	0.009 31	0.005 85	9.20 E- 03	1.6[.15- 2.22]	TGTCCCGGA GGACTGTCT CCACCAGGA CATCAGTCC AGAC[G/A]CT GTGACTGTC CCGGTTGAA ATTCTCTCA ACTGATGCC AG	SEQ ID NO: 826
chr 7	999995 51	T	C	ZCWP W1	p.R529 G	0.005 15	0.002 82	1.11 E- 02	1.83[1.17- 2.85]	CCTGGCTGG TCAGAATCT GAATTCCCT TGGCCTTCTT TCC[T/C]TCC CATTCTGGG TGCAGGAGG AGCTGTGGA TTTCCTGCCT	SEQ ID NO: 827

chr 7	100228 655	G	T	<i>TFR2</i>	p.A37 6D	0.006 62	0.004 19	2.67 E- 02	1.58[1.07- 2.34]	ATAAGGGGA GCCTAGGAG GCTCCCCTG CCATTCTTG GGGG[G/T]CC ACAGGGCCT TTGAGCTTC CTGGAGAGG AGGAAGGCA GA	SEQ ID NO: 828
chr 7	100633 938	G	A	<i>MUC1</i> 2	p.G32 S	0.005 88	0.006 17	9.17 E- 01	0.95[0.63- 1.44]	CTCTCAAAT CACAGGCTC AACAGTAAA CACCAGTAT TGGA[G/A]GT AATACAAC TCTGCATCC ACACCCAGT TCAAGCGAC CC	SEQ ID NO: 829
chr 7	100633 960	C	T	<i>MUC1</i> 2	p.T39I	0.000 25	0.000 03	1.85 E- 01	7.1[0 .64- 78.3 5]	GTAAACACC AGTATTGGA GGTAATACA ACTTCTGCA TCCA[C/T]AC CCAGTTCAA GCGACCCCT TTACCACCT TAGTGACTA T	SEQ ID NO: 830
chr 7	100634 145	G	A	<i>MUC1</i> 2	p.A10 1T	0.002 70	0.001 70	1.71 E- 01	1.59[0.85- 2.96]	CCCAGGTGC AACTGGAAC AACACTCTT CCCTTCCCA CTCT[G/A]CA ACCTCAGTT TTTGTTGGA GAACCTAAA ACCTCACCC AT	SEQ ID NO: 831
chr 7	100634 209	C	T	<i>MUC1</i> 2	p.T122 I	0.000 25	0.000 03	1.84 E- 01	7.15[0.65- 78.8 7]	CCTAAAACC TCACCCATC ACTTCAGCC TCAATGGAA ACAA[C/T]AG CGTTACCTG GCAGTACCA CAACAGCAG GCCTGAGTG AG	SEQ ID NO: 832
chr 7	100634 302	C	G	<i>MUC1</i> 2	p.P153 R	0.000 49	0.000 53	1.00 E+0 0	0.92[0.22- 3.85]	TTCTACAGT AGCCCCAGA TCACCAGAC AGAACACTC TCAC[C/G]TG CCCGCACGA CAAGCTCAG GCGTCAGTG AAAAATCAA CC	SEQ ID NO: 833

chr 7	100634 358	C	T	<i>MUC1</i> 2	p.P172 S	0.006 86	0.006 48	7.62 E- 01	1.06[0.72- 1.56]	CTCAGGCGT CAGTGA AAA ATCAACCAC CTCCACAG CCGA[C/T]CA GGCCCAACG CACACAATA GCGTTCCT GACAGTACC AC	SEQ ID NO: 834
chr 7	100634 374	C	A	<i>MUC1</i> 2	p.T177 K	0.000 25	0.000 24	1.00 E+0 0	1.02[0.13- 7.75]	AAATCAACC ACCTCCCAC AGCCGACCA GGCCCAACG CACA[C/A]AA TAGCGTTCC CTGACAGTA CCACCATGC CAGGCGTCA GT	SEQ ID NO: 835
chr 7	100634 386	C	T	<i>MUC1</i> 2	p.P181 L	0.001 96	0.002 37	7.38 E- 01	0.83[0.41- 1.69]	TCCCACAGC CGACCAGGC CCAACGCAC ACAATAGCG TTCC[C/T]TG ACAGTACCA CCATGCCAG GCGTCAGTC AGGAATCTA CA	SEQ ID NO: 836
chr 7	100634 440	T	G	<i>MUC1</i> 2	p.I199 S	0.000 25	0.000 17	5.26 E- 01	1.42[0.18- 11.1 3]	ATGCCAGGC GTCAGTCAG GAATCTACA GCTTCCCAC AGCA[T/G]CC CCGGCTCCA CAGACACAA CACTGTCCC CTGGCACTA CC	SEQ ID NO: 837
chr 7	100634 700	G	C	<i>MUC1</i> 2	p.D28 6H	0.005 39	0.005 46	1.00 E+0 0	0.99[0.64- 1.52]	GGGAGAACC TACCACCTT CCAGAGCTG GCCAAGCTC AAAG[G/C]A CACTTCGCC TGCACCTTCT GGTACCACA TCAGCCTTT GT	SEQ ID NO: 838
chr 7	100634 788	C	T	<i>MUC1</i> 2	p.T315 I	0.000 25	0.000 17	5.28 E- 01	1.42[0.18- 11.0 7]	TCTACA ACT TATCACAGC AGCCCCGAGC TCAACTCCA ACAA[C/T]CC ACTTTTCTGC CAGCTCCAC AACCTTGGG CCATAGTGA G	SEQ ID NO: 839

chr 7	100634 887	G	A	<i>MUC1</i> 2	p.R348 H	0.013 97	0.015 95	3.63 E- 01	0.87[0.67- 1.14]	AGCAGCCCA GTTGCAACT GCAACAACA CCCCACCT GCCC[G/A]CT CCGCGACCT CAGGCCATG TTGAAGAAT CTACAGCCT AC	SEQ ID NO: 840
chr 7	100635 034	A	T	<i>MUC1</i> 2	p.K39 7I	0.000 49	0.000 38	6.69 E- 01	1.3[0 .31- 5.53]	GAAGAATCA GCAACTTTC CACGGCAGC ACAACACAC ACAA[A/T]AT CTTCAACTC CTAGCACCA CAGCTGCCC TAGCACATA CA	SEQ ID NO: 841
chr 7	100635 052	C	G	<i>MUC1</i> 2	p.T403 S	0.000 25	0.000 00	6.53 E- 02	Inf[NaN -Inf]	CACGGCAGC ACAACACAC ACAAAATCT TCAACTCCT AGCA[C/G]CA CAGCTGCCC TAGCACATA CAAGCTACC ACAGCAGCC TG	SEQ ID NO: 842
chr 7	100635 091	T	C	<i>MUC1</i> 2	p.L416 P	0.000 49	0.000 55	1.00 E+0 0	0.89[0.21- 3.73]	ACCACAGCT GCCCTAGCA CATACAAGC TACCACAGC AGCC[T/C]GG GCTCAACTG AAACAACAC ACTTCCGTG ATAGCTCCA CA	SEQ ID NO: 843
chr 7	100635 236	C	G	<i>MUC1</i> 2	p.D46 4E	0.000 98	0.001 41	6.61 E- 01	0.7[0 .26- 1.9]	TCTTACCTGC CGGCTCTAC ACCCTCAGT TCTTGTGG AGA[C/G]TCG ACGCCCTCA CCCATCAGT TCAGGCTCA ATGGAAACC A	SEQ ID NO: 844
chr 7	100635 250	C	A	<i>MUC1</i> 2	p.P469 H	0.001 23	0.001 72	6.90 E- 01	0.71[0.29- 1.75]	TCTACACCC TCAGTTCTTG TTGGAGACT CGACGCCCT CAC[C/A]CAT CAGTTCAGG CTCAATGGA AACCACAGC GTTACCCGG C	SEQ ID NO: 845

chr 7	100635 267	A	C	<i>MUC1</i> 2	p.M47 5L	0.000 25	0.000 09	3.33 E- 01	2.86[0.33- 24.4 9]	TGTTGGAGA CTCGACGCC CTCACCCAT CAGTTCAGG CTCA[A/C]TG GAAACCACA GCGTTACCC GGCAGTACC ACAAAACCA GG	SEQ ID NO: 846
chr 7	100635 336	A	G	<i>MUC1</i> 2	p.S498 G	0.005 88	0.005 70	8.30 E- 01	1.03[0.68- 1.56]	CACAAAACC AGGCCTCAG TGAGAAATC TACCACTTTC TAC[A/G]GTA GCCCCAGAT CACCAGACA CAACACACT TACCTGCCA G	SEQ ID NO: 847
chr 7	100635 419	C	G	<i>MUC1</i> 2	p.H52 5Q	0.000 25	0.000 15	4.92 E- 01	1.59[0.2- 12.5 4]	TGACAAGCT CAGGCGTCA GTGAAGAAT CCACCACCT CCCA[C/G]AG CCGACCAGG CTCAACACA CACAACAGC ATTCCCTGG CA	SEQ ID NO: 848
chr 7	100635 442	C	A	<i>MUC1</i> 2	p.T533 K	0.000 25	0.000 02	1.26 E- 01	14.3 2[0.9 - 228. 9]	GAATCCACC ACCTCCCAC AGCCGACCA GGCTCAACA CACA[C/A]AA CAGCATTCC CTGGCAGTA CCACCATGC CAGGCCTCA GT	SEQ ID NO: 849
chr 7	100635 648	C	G	<i>MUC1</i> 2	p.L602 V	0.000 49	0.000 69	1.00 E+0 0	0.71[0.17- 2.95]	AACAACACT CTTACCTGA CAACACCAC AGCCTCAGG ACTC[C/G]TT GAAGCATCT ATGCCCGTC CACAGCAGC ACCAGATCG CC	SEQ ID NO: 850
chr 7	100635 652	A	C	<i>MUC1</i> 2	p.E603 A	0.001 47	0.000 84	1.73 E- 01	1.75[0.75- 4.09]	ACACTCTTA CCTGACAAC ACCACAGCC TCAGGACTC CTTG[A/C]AG CATCTATGC CCGTCCACA GCAGACCA GATCGCCAC AC	SEQ ID NO: 851

chr 7	100635 686	G	A	<i>MUC1</i> 2	p.S614 S	0.005 39	0.003 22	3.40 E- 02	1.68[1.08- 2.61]	TCCTTGAAG CATCTATGC CCGTCCACA GCAGCACCA GATC[G/A]CC ACACACAAC ACTGTCCCC TGCCGGCTC TACAACCCG TC	SEQ ID NO: 852
chr 7	100635 814	C	T	<i>MUC1</i> 2	p.P657 L	0.004 17	0.003 99	7.98 E- 01	1.04[0.64- 1.71]	AGGCCTGCA CCTCCTACT ACCACATCA GCCTTTGTTG AGC[C/T]ATC TACAACCTC CCACGGCAG CCCGAGCTC AATTCCAAC A	SEQ ID NO: 853
chr 7	100635 858	C	G	<i>MUC1</i> 2	p.H67 2D	0.000 25	0.000 15	4.91 E- 01	1.59[0.2- 12.5 5]	TACAACCTC CCACGGCAG CCCGAGCTC AATTCCAAC AACC[C/G]AC ATTTCTGCCC GCTCCACAA CCTCAGGCC TCGTTGAAG A	SEQ ID NO: 854
chr 7	100635 864	T	A	<i>MUC1</i> 2	p.S674 T	0.000 74	0.000 21	7.03 E- 02	3.58[1.01- 12.6 9]	CTCCCACGG CAGCCCGAG CTCAATTCC ACAACCCA CATT[T/A]CT GCCCCGCTCC ACAACCTCA GGCCTCGTT GAAGAATCT AC	SEQ ID NO: 855
chr 7	100635 871	G	A	<i>MUC1</i> 2	p.R676 H	0.000 25	0.000 03	1.84 E- 01	7.15[0.65- 78.8 7]	GGCAGCCCG AGCTCAATT CCAACAACC CACATTTCT GCCC[G/A]CT CCACAACCT CAGGCCTCG TTGAAGAAT CTACGACCT AC	SEQ ID NO: 856
chr 7	100635 880	C	A	<i>MUC1</i> 2	p.T679 N	0.004 66	0.003 65	2.87 E- 01	1.28[0.8- 2.04]	AGCTCAATT CCAACAACC CACATTTCT GCCCCGCTCC ACAA[C/A]CT CAGGCCTCG TTGAAGAAT CTACGACCT ACCACAGCA GC	SEQ ID NO: 857

chr 7	100635 928	C	G	<i>MUC1</i> 2	p.S695 X	0.000 25	0.000 00	3.71 E- 02	Inf[NaN -Inf]	CTCGTTGAA GAATCTACG ACCTACCAC AGCAGCCCG GGCT[C/G]AA CTCAAACAA TGCACTTCC CTGAAAGCG ACACAACCT CA	SEQ ID NO: 858
chr 7	100636 573	C	A	<i>MUC1</i> 2	p.S910 Y	0.005 64	0.016 13	8.63 E- 09	0.35[0.23- 0.53]	AGCACCACC ACCTCAGGC CCCAGTCAG GAATCAACA ACTT[C/A]CC ACAGCAGCT CAGGTTCAA CTGACACAG CACTGTCCC CT	SEQ ID NO: 859
chr 7	100636 765	G	A	<i>MUC1</i> 2	p.R974 H	0.000 49	0.000 03	2.64 E- 02	18.6 3[1.6 9- 205. 52]	GAAGCATCT ACACGCGTC CACAGCAGC ACTGGCTCA CCAC[G/A]CA CAACACTGT CCCCTGCCA GCTCCACAA GCCCTGGAC TT	SEQ ID NO: 860
chr 7	100636 831	C	G	<i>MUC1</i> 2	p.T996 S	0.000 49	0.000 23	2.81 E- 01	2.12[0.46- 9.8]	ACAAGCCCT GGACTTCAG GGAGAATCT ACTGCCTTC CAGA[C/G]CC ACCCAGCCT CAACTCACA CAACGCCTT CACCTCCTA GC	SEQ ID NO: 861
chr 7	100636 860	T	C	<i>MUC1</i> 2	p.S100 6P	0.005 15	0.006 58	3.48 E- 01	0.78[0.5- 1.22]	TGCCTTCCA GACCCACCC AGCCTCAAC TCACACAAC GCCT[T/C]CA CCTCCTAGC ACCGCAACA GCCCCTGTT GAAGAATCT AC	SEQ ID NO: 862
chr 7	100637 251	C	G	<i>MUC1</i> 2	p.P113 6R	0.006 37	0.000 03	5.46 E- 26	250. 3[33. 96- 1844 .95]	CTGGGCGTC GGTGAAGAA TCCACCACC TCCCGTAGC CAAC[C/G]AG GTTCTACTC ACTCAACAG TGTCACCTG CCAGCACCA CC	SEQ ID NO: 863

chr 7	100637 407	C	G	<i>MUC1</i> 2	p.T118 8S	0.001 47	0.001 07	4.55 E- 01	1.37[0.58- 3.23]	CACAGCACC ACAACCTCA GTTTCATGGT GAAGAGCCT ACAA[C/G]CT TCCACAGCC GGCCAGCCT CAACTCACA CAACACTGT TC	SEQ ID NO: 864
chr 7	100637 556	G	A	<i>MUC1</i> 2	p.G12 38S	0.008 82	0.011 19	1.98 E- 01	0.79[0.56- 1.11]	CCAAACAGG GTTACCTGC CACACTCAC AACCGCAGA CCTC[G/A]GT GAGGAATCA ACTACCTTTC CCAGCAGCT CAGGCTCAA C	SEQ ID NO: 865
chr 7	100637 902	C	T	<i>MUC1</i> 2	p.P135 3L	0.001 47	0.001 69	1.00 E+0 0	0.87[0.37- 2.04]	TTCCCTGAC AGCACCACC ACCTCAGAC CTCAGTCAG GAAC[C/T]TA CAACTTCCC ACAGCAGCC AAGGCTCAA CAGAGGCAA CA	SEQ ID NO: 866
chr 7	100638 584	C	G	<i>MUC1</i> 2	p.H15 80Q	0.006 13	0.000 00	1.36 E- 29	Inf	CGACAAGCT CAGGCGTCA GTGAAGAAT CCACCACCT CCCA[C/G]AG CCGACCAGG CTCAACGCA CACAACAGC ATTCCCTGG CA	SEQ ID NO: 867
chr 7	100638 673	G	T	<i>MUC1</i> 2	p.S161 0I	0.001 47	0.000 07	1.36 E- 05	21.2 8[6- 75.4 4]	ATGCCAGGC GTCAGTCAG GAATCTACA GCTTCCCAC AGCA[G/T]CC CAGGCTCCA CAGACACAA CATTGTCCC CTGGCAGTA CC	SEQ ID NO: 868
chr 7	100638 754	G	A	<i>MUC1</i> 2	p.S163 7N	0.000 49	0.000 03	2.36 E- 02	14.2 5[2.0 1- 101. 22]	ACAGCATCA TCCCTTGGTC CAGAATCTA CTACTTTCCA CA[G/A]CAGC CCAGGCTCC ACTGAAACA AACTCTTA CCTGACAAC	SEQ ID NO: 869

chr 7	100638 850	C	T	<i>MUC1</i> 2	p.S166 9L	0.000 25	0.000 22	6.13 E- 01	1.1[0 .14- 8.39]	CTCCTTGAA GCATCTACG CCCGTCCAC AGCAGCACT GGAT[C/T]GC CACACACAA CACTGTCCC CTGCCGGCT CTACAACAC GT	SEQ ID NO: 870
chr 7	100638 889	G	A	<i>MUC1</i> 2	p.R168 2H	0.001 23	0.000 72	2.33 E- 01	1.7[0 .67- 4.31]	TCGCCACAC ACAACACTG TCCCCTGCC GGCTCTACA ACAC[G/A]TC AGGGAGAAT CTACCACCT TCCAGAGCT GGCCAAGCT CA	SEQ ID NO: 871
chr 7	100638 919	G	A	<i>MUC1</i> 2	p.W16 92X	0.000 25	0.000 03	1.84 E- 01	7.15[0.65- 78.9]	TCTACAACA CGTCAGGGA GAATCTACC ACCTTCCAG AGCT[G/A]GC CAAGCTCAA AGGACACTA TGCCTGCAC CTCCTACTA CC	SEQ ID NO: 872
chr 7	100638 922	C	G	<i>MUC1</i> 2	p.P169 3R	0.000 25	0.000 00	6.53 E- 02	Inf[NaN -Inf]	ACAACACGT CAGGGAGAA TCTACCACC TTCCAGAGC TGGC[C/G]AA GCTCAAAGG ACACTATGC CTGCACCTC CTACTACCA CA	SEQ ID NO: 873
chr 7	100638 925	G	A	<i>MUC1</i> 2	p.S169 4N	0.000 49	0.000 33	6.45 E- 01	1.51[0.35- 6.47]	ACACGTCAG GGAGAATCT ACCACCTTC CAGAGCTGG CCAA[G/A]CT CAAAGGACA CTATGCCTG CACCTCCTA CTACCACAT CA	SEQ ID NO: 874
chr 7	100638 928	C	G	<i>MUC1</i> 2	p.S169 5X	0.000 25	0.000 00	3.71 E- 02	Inf[NaN -Inf]	CGTCAGGGA GAATCTACC ACCTTCCAG AGCTGGCCA AGCT[C/G]AA AGGACACTA TGCCTGCAC CTCCTACTA CCACATCAG CC	SEQ ID NO: 875

chr 7	100643 255	C	G	<i>MUC1</i> 2	p.H31 37Q	0.017 16	0.000 00	9.03 E- 84	Inf	CGACAAGCT CAGGCGTCA GTGAAGAAT CCACCACCT CCCA[C/G]AG CCGACCAGG CTCAACGCA CACAACAGC ATTCCTGG CA	SEQ ID NO: 876
chr 7	100643 388	G	A	<i>MUC1</i> 2	p.A31 82T	0.005 88	0.001 52	2.43 E- 07	3.89[2.47- 6.12]	AGGCTCCAC AGACACAAC ACTGTCCCC TGGCAGTAC CACA[G/A]CA TCATCCCTTG GTCCAGAAT CTACTACCTT CCACAGCGG	SEQ ID NO: 877
chr 7	100643 560	G	A	<i>MUC1</i> 2	p.R323 9H	0.003 92	0.000 09	2.27 E- 15	44.8 5[16. 42- 122. 48]	TCGCCACAC ACAACACTG TCCCCTGCC GGCTCTACA ACCC[G/A]TC AGGGAGAAT CTACCACCT TCCAGAGCT GGCCTAACT CG	SEQ ID NO: 878
chr 7	100643 574	A	G	<i>MUC1</i> 2	p.T324 4A	0.000 49	0.000 05	3.74 E- 02	9.54[1.59- 57.1 3]	ACTGTCCCC TGCCGGCTC TACAACCCG TCAGGGAGA ATCT[A/G]CC ACCTTCCAG AGCTGGCCT AACTCGAAG GACACTACC CC	SEQ ID NO: 879
chr 7	100643 737	C	T	<i>MUC1</i> 2	p.S329 8L	0.000 49	0.000 33	6.47 E- 01	1.48[0.34- 6.34]	TTTTCTGCCA GCTCCACAA CCTTGGGCC GTAGTGAGG AAT[C/T]GAC AACAGTCCA CAGCAGCCC AGTTGCAAC TGCAACAAC A	SEQ ID NO: 880
chr 7	100643 791	G	A	<i>MUC1</i> 2	p.R331 6H	0.008 09	0.000 22	1.03 E- 28	36.5 9[18. 88- 70.9]	AGCAGCCCA GTTGCAACT GCAACAACA CCCTCGCCT GCCC[G/A]CT CCACAACCT CAGGCCTCG TTGAAGAAT CTACGACCT AC	SEQ ID NO: 881

chr 7	100646 590	G	A	<i>MUC1</i> 2	p.S424 9N	0.000 74	0.000 54	4.92 E- 01	1.36[0.42- 4.44]	ACCATGCCA GGCGTCAGT CAGGAATCT ACAGCTTCC CACA[G/A]CA GCCCAGGCT CCACAGACA CAACACTGT CCCCTGGCA GT	SEQ ID NO: 882
chr 7	100646 712	A	C	<i>MUC1</i> 2	p.N42 90H	0.021 08	0.051 65	5.47 E- 22	0.4[0 .32- 0.49]	CAGCAGCCC AGGCTCCAC TGAAACAAC ACTCTTACCT GAC[A/C]ACA CCACAGCCT CAGGCCTCC TTGAAGCAT CTACACCCG T	SEQ ID NO: 883
chr 7	100646 749	C	G	<i>MUC1</i> 2	p.P430 2R	0.022 55	0.000 07	6.14 E- 92	313. 86[9 9.34- 991. 56]	GACAACACC ACAGCCTCA GGCCTCCTT GAAGCATCT ACAC[C/G]CG TCCACAGCA GCACTGGAT CGCCACACA CAACACTGT CC	SEQ ID NO: 884
chr 7	100646 809	G	A	<i>MUC1</i> 2	p.R432 2H	0.000 25	0.000 27	1.00 E+0 0	0.89[0.12- 6.73]	TCGCCACAC ACAACACTG TCCCCTGCC GGCTCTACA ACCC[G/A]TC AGGGAGAAT CTACCACCT TCCAGAGCT GGCCAAACT CG	SEQ ID NO: 885
chr 7	100646 973	C	T	<i>MUC1</i> 2	p.R437 7C	0.002 45	0.002 75	8.75 E- 01	0.89[0.47- 1.7]	TCCAACAAC CCACTTTTCT GCCAGCTCC ACAACATTG GGC[C/T]GTA GTGAGGAAT CGACAACAG TCCACAGCA GCCCAGTTG C	SEQ ID NO: 886
chr 7	100647 735	A	G	<i>MUC1</i> 2	p.R463 1G	0.000 25	0.000 00	6.60 E- 02	Inf[NaN -Inf]	CCCTGAAAG CTCCACAGC TTCAGGTCG TAGTGAAGA ATCA[A/G]GA ACTTCCCAC AGCAGCACA ACACACACA ATATCTTCA CC	SEQ ID NO: 887

chr 7	100647 774	C	G	<i>MUC1</i> 2	p.P464 4A	0.006 37	0.006 73	9.21 E- 01	0.95[0.64- 1.41]	AAGAACTTC CCACAGCAG CACAACACA CACAATATC TTCA[C/G]CT CCTAGCACC ACATCTGCC CTTGTTGAA GAACCTACC AG	SEQ ID NO: 888
chr 7	100647 976	C	G	<i>MUC1</i> 2	p.S471 1C	0.005 39	0.003 95	1.60 E- 01	1.37[0.88- 2.12]	TTACCTGCC CATTTTACTA CCTCAGGCC GCATTGCAG AAT[C/G]TAC CACCTTCTAT ATCTCTCCA GGCTCAATG GAAACAACA	SEQ ID NO: 889
chr 7	100647 988	A	G	<i>MUC1</i> 2	p.Y47 15C	0.000 49	0.000 03	2.36 E- 02	14.2 7[2.0 1- 101. 32]	TTTACTACCT CAGGCCGCA TTGCAGAAT CTACCACCT TCT[A/G]TAT CTCTCCAGG CTCAATGGA AACAACATT AGCCAGCAC T	SEQ ID NO: 890
chr 7	100648 044	C	G	<i>MUC1</i> 2	p.L473 4V	0.005 64	0.006 19	7.56 E- 01	0.91[0.6- 1.39]	AATGGAAAC AACATTAGC CAGCACTGC CACAACACC AGGC[C/G]TC AGTGCAAAA TCTACCATC CTTTACAGT AGCTCCAGA TC	SEQ ID NO: 891
chr 7	100648 148	C	G	<i>MUC1</i> 2	p.S476 8R	0.000 25	0.000 10	3.78 E- 01	2.38[0.29- 19.7 5]	CCAGCATGA CAAGCTCCA GCATCAGTG GAGAACCCA CCAG[C/G]TT GTATAGCCA AGCAGAGTC AACACACAC AACAGCGTT CC	SEQ ID NO: 892
chr 7	100648 183	C	T	<i>MUC1</i> 2	p.A47 80V	0.000 25	0.000 03	1.84 E- 01	7.12[0.65- 78.5 5]	ACCAGCTTG TATAGCCAA GCAGAGTCA ACACACACA ACAG[C/T]GT TCCCTGCCA GCACCACCA CCTCAGGCC TCAGTCAGG AA	SEQ ID NO: 893

chr 7	100649 758	G	T	<i>MUC1</i> 2	p.C498 8F	0.000 49	0.000 44	7.00 E- 01	1.12[0.27- 4.73]	CACGGTGAC TGCTGTGGA TTCTATCTCT CCACAGGGT TGT[G/T]CCA GGAAGGACA AATTTGGAA TGGAAAACA ATGCGTCTG T	SEQ ID NO: 894
chr 7	100649 815	G	C	<i>MUC1</i> 2	p.G50 07A	0.000 25	0.000 05	2.41 E- 01	4.67[0.49- 44.9 4]	TGGAATGGA AAACAATGC GTCTGTCCC CAAGGCTAC GTTG[G/C]TT ACCAGTGCT TGTCCCCTCT GGAATCCTT CCCTGTAGG T	SEQ ID NO: 895
chr 7	100649 847	C	T	<i>MUC1</i> 2	p.P501 8S	0.000 25	0.000 86	2.59 E- 01	0.29[0.04- 2.07]	CTACGTTGG TTACCAGTG CTTGTCCTCT CTGGAATCC TTC[C/T]CTG TAGGTAATG ACCTTTTCTG AGACCTGCA GCTCTTTGC	SEQ ID NO: 896
chr 7	100649 851	T	C	<i>MUC1</i> 2	p.V50 19A	0.000 74	0.000 24	9.98 E- 02	3[0.8 6- 10.4 6]	GTTGGTTAC CAGTGCTTG TCCCCTCTG GAATCCTTC CCTG[T/C]AG GTAATGACC TTTTCTGAG ACCTGCAGC TCTTTGCAG GC	SEQ ID NO: 897
chr 7	100651 921	C	T	<i>MUC1</i> 2	p.P502 2L	0.000 74	0.000 43	4.29 E- 01	1.69[0.51- 5.6]	GCTGTCTCA CGCATACCA TGGCCTTTTC CCACAGAAA CCC[C/T]GGA AAAAC TCAA CGCCACTTT AGGTATGAC AGTGAAAGT G	SEQ ID NO: 898
chr 7	100656 384	T	C	<i>MUC1</i> 2	p.L520 0P	0.000 25	0.000 02	1.29 E- 01	14.0 2[0.8 8- 224. 21]	AAGTGCACC AAAGGAACG AAGTCGCAA ATGAACTGT AACC[T/C]GG GCACATGTC AGCTGCAAC GCAGTGGCC CCCGCTGCC TG	SEQ ID NO: 899

chr 7	100657 247	T	C	<i>MUC1</i> 2	p.I523 1T	0.000 25	0.000 23	6.19 E- 01	1.08[0.14- 8.25]	AACACACAC TGGTACTGG GGAGAGACC TGTGAATTC AACA[T/C]CG CCAAGAGCC TCGTGTATG GGATCGTGG GGGCTGTGA TG	SEQ ID NO: 900
chr 7	100678 918	G	A	<i>MUC1</i> 7	p.P140 7P	0.018 14	0.000 02	2.01 E- 104	1009 .33[2 47.6 9- 4112 .96]	GAACCACTC CGTTAACAA GTATACCTG TCAGCACCA CGCC[G/A]GT AGTCAGTTC TGAGGCTAG CACCCCTTC AGCAACTCC TG	SEQ ID NO: 901
chr 7	100681 219	C	T	<i>MUC1</i> 7	p.A21 74A	0.012 99	0.000 00	8.18 E- 78	Inf	CTCCTTTAAC AAGTATGCC TGTCAGCAC CACAGTGGT GGC[C/T]AGT TCTGCAATC AGCACCCCTT TCAACAACCT CCTGTTGAC A	SEQ ID NO: 902
chr 7	100681 310	T	G	<i>MUC1</i> 7	p.S220 5A	0.006 37	0.000 00	2.25 E- 38	Inf	TGTGACCAA TTCTACTGA AGCCCGTTC ATCTCCTAC AACT[T/G]CT GAAGGTACC AGCATGCCA ACCTCAACT CCTAGTGAA GG	SEQ ID NO: 903
chr 7	100682 597	T	C	<i>MUC1</i> 7	p.S263 4P	0.007 11	0.001 44	4.22 E- 11	4.95[3.33- 7.38]	TACCAGCAT GCCAATCTC AACTCCTAG TGAAGTAAG TACT[T/C]CA TTAACAAGT ATACTTGTC AGCACCATG CCAGTGGCC AG	SEQ ID NO: 904
chr 7	100682 613	T	C	<i>MUC1</i> 7	p.L263 9P	0.006 86	0.000 48	2.08 E- 20	14.3 2[9.0 5- 22.6 5]	TCAACTCCT AGTGAAGTA AGTACTTCA TTAACAAGT ATAC[T/C]TG TCAGCACCA TGCCAGTGG CCAGTTCTG AGGCTAGCA CC	SEQ ID NO: 905

chr 7	102087 238	C	T	<i>ORAI2</i>	p.L168 L	0.011 27	0.006 23	2.80 E- 04	1.82[1.35- 2.46]	TGCTTGGCA TCCTACTCTT CCTGGCCGA GGTGGTGCT GCT[C/T]TGC TGGATCAAG TTCCTCCCCG TGGATGCCC GGCGCCAGC	SEQ ID NO: 906
chr 7	108112 902	A	G	<i>PNPL A8</i>	p.D76 4D	0.005 88	0.003 48	1.58 E- 02	1.69[1.12- 2.56]	ATGGAAGTC CTTCATACA TATCAGTTTT TAATTTTATC CA[A/G]TCAT TAATTTTCTG CAGAGTTGT TTTTTCTTGA CTTAATA	SEQ ID NO: 907
chr 7	111368 481	G	A	<i>DOCK 4</i>	p.P191 7L	0.009 31	0.005 38	2.33 E- 03	1.74[1.25- 2.42]	CGCGGGCGG CTCCGACGT GACGGGGAT GGAGAGGCT GTGA[G/A]GT AGCGGGACG GGGCGCCGC AGAGTCCGC TCGTAGACG CT	SEQ ID NO: 908
chr 7	117232 305	A	G	<i>CFTR</i>	p.E695 G	0.021 32	0.000 01	3.53 E- 125	2406 .22[3 35.0 8- 1727 9.21]	ACAAAAAAA CAATCTTTTA AACAGACTG GAGAGTTTG GGG[A/G]AA AAAGGAAGA ATTCTATTCT CAATCCAAT CAACTCTAT A	SEQ ID NO: 909
chr 7	123143 031	G	A	<i>IQUB</i>	p.P278 P	0.007 84	0.004 78	1.12 E- 02	1.65[1.15- 2.36]	ACATTACCT GCGTATCCC TACAAAATA TACTGAGTC TTTC[G/A]GG AATCCTTTTA GGTACAGTT TGTGTTCCA GCATTGTGA T	SEQ ID NO: 910
chr 7	141366 203	A	G	<i>KLAAL 147</i>	p.M23 5T	0.006 37	0.004 20	4.89 E- 02	1.52[1.02- 2.26]	GATGAGGAT CTGTTCTCCA AAGAACTTT ATAAACTGA GAC[A/G]TGC AGCCAGCTG GGTGTGTGA TCTGAAAAA ATTGAGGGG A	SEQ ID NO: 911

chr 7	141763 311	C	A	<i>MGAM</i>	p.P142 4T	0.012 99	0.009 45	2.68 E- 02	1.38[1.05- 1.82]	GAGGTATGT CTGTGTTTG GCATTTCTA GGATATGAA TGAA[C/A]CA TCAAGCTTC GTGAATGGG GCAGTTTCT CCAGGCTGC AG	SEQ ID NO: 912
chr 7	141794 442	C	T	<i>MGAM</i>	p.F154 7F	0.006 13	0.002 24	1.83 E- 05	2.75[1.82- 4.15]	CTGTGCTTCT CGTTGCAGG CATGATGGA GTTTCAGCCT CTT[C/T]GGC ATATCCTAT GTGAGTGTG CTTGGGATC CTCCTAAGC A	SEQ ID NO: 913
chr 7	150069 074	G	A	<i>REPIN 1</i>	p.K24 8K	0.009 56	0.000 00	1.71 E- 56	Inf	CCTTCCAGT GTGCCTGTT GTGGCAAGC GCTTCCGGC ACAA[G/A]CC CAACTTGAT CGCTCACCG CCGCGTGCA CACGGGCGA GC	SEQ ID NO: 914
chr 7	150738 005	C	T	<i>ABCB8</i>	p.G40 5G	0.005 64	0.002 97	5.50 E- 03	1.9[.24- 2.91]	TGCCCCCTG GCAAGATCG TGGCCCTCG TGGGCCAGT CTGG[C/T]GG AGGTAAGGG GAGCCCACC ACCTCTTCA CCCTCTGAC TC	SEQ ID NO: 915
chr 7	150840 440	A	T	<i>AGAP 3</i>	p.E431 D	0.005 15	0.002 79	1.04 E- 02	1.85[1.19- 2.88]	TCATGCCCT GATGGGCCT GTGGTTGCA GAGAGGAGA AGGA[A/T]CG CTGGATACG GGCCAAGTA TGAACAGAA GCTCTTCCTG G	SEQ ID NO: 916
chr 7	151078 993	C	T	<i>WDR8 6</i>	p.G31 3S	0.006 86	0.003 98	8.14 E- 03	1.73[1.18- 2.54]	GAGGGACCT ACCTGGATG CAGTTGATG ATGAATGTG TGGC[C/T]CC GGAACACCC TCCGCAGCT CTCCAGACT GCGCGTCGA AG	SEQ ID NO: 917

chr 7	151859 899	G	A	<i>KMT2 C</i>	p.S358 8L	0.005 64	0.003 57	4.53 E- 02	1.58[1.04- 2.41]	TTTTTCCTCT GGGATTATA TCAGAATAC AACTGAATG AGC[G/A]ATT GGGTTGATC CCGGATAAC TGTGTCCAT GGGTTATAG T	SEQ ID NO: 918
chr 8	623435	G	A	<i>ERICH I</i>	p.P306 L	0.027 21	0.000 02	7.40 E- 159	1561 .55[3 85.5 8- 6324 .12]	CTCCCCGGA GTCTGCACC CTCTTCCTCC CCAGCCCAT GTC[G/A]GGT CTTCCTCGCT GGCGTCCGC ACCGTCCTC CTCCCTGGT	SEQ ID NO: 919
chr 8	623519	A	C	<i>ERICH I</i>	p.I278 S	0.014 71	0.000 00	1.13 E- 87	Inf	TTTACCGTCT TCCTCCCCG GCCCCGTGTC AGGTCTTCC TCA[A/C]TGG TGCCACAC CGTCCTCCTC CCTGGCGTC TTTAACGTC	SEQ ID NO: 920
chr 8	623675	A	C	<i>ERICH I</i>	p.V22 6G	0.024 51	0.000 00	1.64 E- 145	Inf	CTCGCTAGC GTCCGCACC ATCTTCCTCC CTGGTATCTT TA[A/C]CGTC TTCCTCCCCG GCCAGTGTC GGGTCTTCC TCGCTGGT	SEQ ID NO: 921
chr 8	104660 31	A	C	<i>RP1L1</i>	p.D18 59E	0.005 15	0.000 00	1.36 E- 30	Inf	CTTCTGACTC TGGCTGGGC CTCCCCTTCA GCCTCCTGG GC[A/C]TCCC CTTCTGCCTC TGGGGCCTC TACACCTTCT GACTCTG	SEQ ID NO: 922
chr 8	171597 18	A	G	<i>MTMR 7</i>	p.M52 2T	0.005 15	0.002 79	1.04 E- 02	1.85[1.19- 2.88]	TAGTTCTTCC TCTAGCTGC TGAGTTTCTT CCTTCACTG CC[A/G]TTAG GTAATCTGT AACTGACTG TCGGGGCTG CATCCCCCTT	SEQ ID NO: 923

chr 8	180803 08	A	T	<i>NAT1</i>	p.D25 1V	0.005 88	0.003 35	1.32 E- 02	1.76[1.16- 2.66]	ACCCTCACC CATAGGAGA TTCAATTAT AAGGACAAT ACAG[A/T]TC TAATAGAGT TCAAGACTC TGAGTGAGG AAGAAATAG AA	SEQ ID NO: 924
chr 8	234289 24	C	G	<i>SLC25 A37</i>	p.T191 T	0.005 15	0.003 17	3.48 E- 02	1.63[1.05- 2.53]	ACCGGTCAG CAATCAGCT GCATCCGGA CGGTGTGGA GGAC[C/G]G AGGGGTTGG GGGCCTTCT ACCGGAGCT ACACCACGC AGC	SEQ ID NO: 925
chr 8	251746 10	C	T	<i>DOCK 5</i>	p.T469 M	0.010 78	0.007 86	4.77 E- 02	1.38[1.01- 1.87]	GACAAAGGG AAGAAGAAG ACGCCAAAG AATGTGGAG GTGA[C/T]GA TGTCTGTGC ACGATGAGG AGGGCAAGC TCTTGGAGG TG	SEQ ID NO: 926
chr 8	267219 90	C	T	<i>ADRA 1A</i>	p.R166 K	0.006 506	0.003 743	2.42 E- 02	1.74[1.05- 2.73]	GTTGATCTG GCAGATGGT CTCGTCCTC GGGGGCCCG CTGC[C/T]TC CAGCCGAAC AGGGGTCCA ATGGATATG ACCAGGGAG AG	SEQ ID NO: 927
chr 8	356480 09	G	A	<i>UNC5 D</i>	p.T930 T	0.009 07	0.005 08	1.25 E- 03	1.79[1.28- 2.5]	CCCTGGCCT GTGCCCTTG AAGAGATTG GGAGGACAC ACAC[G/A]A AACTCTCAA ACATTTTCAG AATCCCAGC TTGATGAAG CCG	SEQ ID NO: 928
chr 8	367933 75	T	C	<i>KCNU I</i>	p.N11 29N	0.005 64	0.002 53	8.09 E- 04	2.24[1.46- 3.43]	TATCATCTC AGATACCTT TAGGTGACA ATGCAAAAG AAAA[T/C]GA AAGGAAAAC TTCAGATGA GGTTTATGA TGAGGATCC CT	SEQ ID NO: 929

chr 8	376997 77	G	A	<i>GPR12</i> 4	p.K13 07K	0.005 64	0.000 00	9.89 E- 29	Inf	CGTACCCGC TCAACGCCG CCAGCCTAA ACGGCGCCC CCAA[G/A]G GGGGCAAGT ACGACGACG TCACCCTGA TGGGCGCGG AGG	SEQ ID NO: 930
chr 8	382600 50	C	T	<i>LETM</i> 2	p.A33 1V	0.006 13	0.003 94	4.19 E- 02	1.56[1.04- 2.34]	AAGTTCCAA CTCCATCCCT TACATTCTT TCAGATAAT TG[C/T]CAAG GAAGGGGTG ACAGCATTG AGTGTATCA GAACTACAG	SEQ ID NO: 931
chr 8	382657 55	C	T	<i>LETM</i> 2	p.T385 M	0.005 205	0.002 716	2.15 E- 02	1.92[1.08- 3.18]	GTTTTTTACG CCTAGACAC TCCAGGCCA AATCACAAA TGA[C/T]GGC CCAGAACAG CAAGGCTAG TTCAAAAGG AGCATAAAG G	SEQ ID NO: 932
chr 8	523208 32	G	C	<i>PXDNL</i>	p.L111 8V	0.007 482	0.004 863	4.90 E- 02	1.54[0.97- 2.34]	TAAGCCGCG GAGAAGAGC CTCTGGGTC AGCTCAGGA CTGA[G/C]AA GGTAGGAGG GTGCCCCGC ATTAGCAG CCACGCCAA AC	SEQ ID NO: 933
chr 8	550491 31	A	G	<i>MRPL</i> 15	p.R57 G	0.012 01	0.008 88	4.23 E- 02	1.36[1.02- 1.81]	GAGAAGAGG TAGAAAATG TGGCAGAGG CCATAAAGG AGAA[A/G]G GCAAAGAGG AACCCGGCC CCGCTTGGG CTTTGAGGG AGG	SEQ ID NO: 934
chr 8	813991 52	C	T	<i>ZBTB1</i> 0	p.S36L	0.018 87	0.014 49	3.39 E- 02	1.31[1.03- 1.67]	GGCGGCGGC TCCACGAAC AATAACGCT GGCGGGGAG GCCT[C/T]AG CTTGGCCTC CGCAGCCCC AGCCGAGAC AGCCCCCGC CG	SEQ ID NO: 935

chr 8	919530 77	G	A	<i>NECA B1</i>	p.A27 1T	0.007 16	0.004 12	1.54 E- 02	1.74[1.08- 2.68]	GATGTCTGT GATAGAAGA GGACCTGGA AGAATTCCA GCTC[G/A]CT CTGAAACAC TACGTGGAG AGTGCTTCC TCCCAAAGT GG	SEQ ID NO: 936
chr 8	947463 10	C	G	<i>RBM1 2B</i>	p.E777 Q	0.005 88	0.000 00	7.19 E- 34	Inf	GGCCGCTG AAATGCTCC TGGGGCGGT CTCCGGAAG TGCT[C/G]CG GGGGCGGGC GCCTGAAAT GCTCTGGGG GTGGCCGCC TG	SEQ ID NO: 937
chr 8	978921 19	G	A	<i>CPQ</i>	p.M24 5I	0.008 133	0.004 667	1.07 E- 02	1.75[1.12- 2.62]	CCTGTATTA CGGTGGAAG ATGCAGAAA TGATGTCAA GAAT[G/A]GC TTCTCATGG GATCAAAAT TGTCATTCA GCTAAAGAT GG	SEQ ID NO: 938
chr 8	989912 22	A	G	<i>MATN 2</i>	p.K35 6R	0.006 86	0.000 00	2.40 E- 41	Inf	CTTTGCCAG TGCCATGAA GGATTTGCT CTTAACCCA GATA[A/G]A AAAACGTGC ACAAGTAAG TTACACACA CATGCACAC ACA	SEQ ID NO: 939
chr 8	100832 259	A	G	<i>VPS13 B</i>	p.N29 68S	0.008 33	0.005 07	7.31 E- 03	1.65[1.17- 2.34]	ACTTTGTTG ATAGAACTT CTGCCCTGG GCCCTGCTT ATCA[A/G]TG AATCCAAAT GGGACCTCT GGCTATTTG AAGGAGAGA AA	SEQ ID NO: 940
chr 8	103573 042	G	A	<i>ODFI</i>	p.S228 N	0.005 64	0.000 00	4.61 E- 34	Inf	TGCAGCCCC TGCAACCCC TGCAACCCG TGCA[G/A]CC CATATGATC CTTGCAACC CGTGTTATC CCTGTGGAA GC	SEQ ID NO: 941

chr 8	104897 928	G	A	<i>RIMS2</i>	p.R175 R	0.005 64	0.003 56	4.50 E- 02	1.59[1.04- 2.42]	GGATCCATG CTGAAGTGT CCCGAGCAC GGCATGAGA GAAG[G/A]C ATAGTGATG TTTCTTTGGC AAATGCTGA TCTGGAAGA TT	SEQ ID NO: 942
chr 8	125711 789	A	G	<i>MTSS1</i>	p.A62 A	0.009 31	0.006 16	1.95 E- 02	1.52[1.09- 2.1]	CAGCCTCCA TCTGCTTACC ACGTGTGTT GGTGGCCAT GTC[A/G]GCC ACTTTCTGA AAGGCGTCC AAGAAGGCA GCTGCTGCT A	SEQ ID NO: 943
chr 8	144297 314	G	A	<i>GPIHB P1</i>	p.G15 9D	0.005 39	0.000 00	1.65 E- 32	Inf	GTCCAGGAC CCAACAGGC AAGGGGGCA GGCGGCCCC CGGG[G/A]C AGCTCCGAA ACTGTGGGC GCAGCCCTC CTGCTCAAC CTC	SEQ ID NO: 944
chr 8	144874 555	G	C	<i>SCRIB</i>	p.P145 0R	0.013 97	0.000 06	9.10 E- 60	229. 41[7 1.81- 732. 85]	AGCTTTGGC CGTCCGCAC CGGGGCGCC ACCTCCCAG GGGT[G/C]GG GGGGACGCC GGGCTCTGC CTGGGGAAG GGACAGGAC GT	SEQ ID NO: 945
chr 8	144940 621	C	T	<i>EPPK1</i>	p.A22 67A	0.008 09	0.001 03	2.32 E- 17	7.88[5.34- 11.6 3]	GCCTCAGGT TGCGCACGG GGTCGATGA CGAAGCCGG TGGC[C/T]GC CTGCGCCTC CAGCAGCAC CAGGGCCGT GCCGGGCCG CA	SEQ ID NO: 946
chr 8	144941 229	A	T	<i>EPPK1</i>	p.Y20 65N	0.006 13	0.003 82	2.84 E- 02	1.61[1.07- 2.41]	GTGTCTCTT GTGGGCGGC ACCTCTCCT GCAGCTCTC GGT[A/T]CGA GACCTTCTCT TGC GTGTTC GGGTCCACA AACCGTTTC	SEQ ID NO: 947

chr 8	144993 230	G	A	<i>PLEC</i>	p.L359 1L	0.008 82	0.006 04	3.15 E- 02	1.46[1.05- 2.05]	TGCTCCTCG GGGATCAGG TCCGACTGC ATCACCTCC CACA[G/A]G GACATGGTG GAGCCGCCG TGGCTGCCG CCGCCGGA ATG	SEQ ID NO: 948
chr 8	145736 853	C	G	<i>RECQ L4</i>	p.V11 96V	0.011 52	0.000 01	2.26 E- 67	1295 .85[1 78.7 5- 9394 .47]	GTCAGCGGG CCACCTGCA GGAGCTCTT CCGTGGCCA GGCC[C/G]AC CAGGGCATG GAAGCTCAG GTGCAGGTA TTTTCTCCAG A	SEQ ID NO: 949
chr 8	146157 265	C	T	<i>ZNF16</i>	p.S303 N	0.005 39	0.003 38	4.01 E- 02	1.6[1 .04- 2.46]	CATGTGAGA CTTTTGGTGC TTTTTAAGG CTCGAGTTC TGG[C/T]TGA AGGCTTTTC CACATTCAT TACACATAT AAGGCCTCT C	SEQ ID NO: 950
chr 9	411793 3	C	G	<i>GLIS3</i>	p.E360 D	0.008 133	0.004 527	6.65 E- 03	1.8[1 .15- 2.7]	GCTGGTCTGA TGTGGACCT TCTCGATGT GCCGCACGA GCTC[C/G]TC CTGCTGGTC GTACAGGGC GCTGCAGTC GATCCAGCG GC	SEQ ID NO: 951
chr 9	601362 4	C	T	<i>RANB P6</i>	p.D66 2N	0.006 831	0.002 903	5.81 E- 04	2.36[1.44- 3.68]	CTCTGCTGG TCTCCAAGA TTTACAAAT TGCCAGCCA TCAT[C/T]GT CACTCATAT TTTCCACATC CTGTGTGTCT AAGAGAGCA	SEQ ID NO: 952
chr 9	154230 04	C	T	<i>SNAP C3</i>	p.H43 Y	0.013 73	0.000 12	1.53 E- 64	117. 22[5 9.77- 229. 91]	TCCAGAGTA TGAGCTTCC CGAGCTAAA TACGCGCGC TTTC[C/T]AT GTGGGCGCC TTTGGGGAG CTGTGGCGG GGCCGTCTG CG	SEQ ID NO: 953

chr 9	190503 23	G	A	<i>RRAG A</i>	p.Q22 2Q	0.007 11	0.004 34	1.55 E- 02	1.64[1.13- 2.39]	CTACATTCTT GGTTATTTCC CACTACCAG TGCAAAGAG CA[G/A]CGCG ACGTCCACC GGTTTGAGA AGATCAGCA ACATCATCA	SEQ ID NO: 954
chr 9	337948 24	A	C	<i>PRSS3</i>	p.K12 T	0.007 35	0.004 13	6.12 E- 03	1.78[1.22- 2.62]	GACAGGATG CACATGAGA GAGACAAGT GGCTTCACA TTGA[A/C]GA AGGGGAGGA GTGCGCCAT TGGTTTTCCA TCCTCCAGA T	SEQ ID NO: 955
chr 9	337967 46	G	T	<i>PRSS3</i>	p.G10 6V	0.005 15	0.000 00	3.09 E- 31	Inf	CCCTACCAG GTGTCCCTG AATTCTGGC TCCCACTTCT GCG[G/T]TGG CTCCCTCATC AGCGAACAG TGGGTGGTA TCAGCAGCT	SEQ ID NO: 956
chr 9	356741 91	G	A	<i>CA9</i>	p.G79 R	0.014 46	0.001 37	1.18 E- 35	10.6 7[7.8 9- 14.4 3]	GCCCAGTGA AGAGGATTC ACCCAGAGA GGAGGATCC ACCC[G/A]GA GAGGAGGAT CTACCTGGA GAGGAGGAT CTACCTGGA GA	SEQ ID NO: 957
chr 9	358100 94	G	A	<i>SPAG8</i>	p.F433 F	0.005 88	0.003 53	2.25 E- 02	1.67[1.1- 2.52]	GAGACAAGG GTACTGGTG TTGAGAAGC TGCAGTTCTT CCG[G/A]AAT GGTGTGTCC AATGTCCTG ATGTTACTG ACACCCTGG A	SEQ ID NO: 958
chr 9	391092 17	C	T	<i>CNTN AP3</i>	p.A76 9T	0.022 55	0.000 02	4.31 E- 131	1284 .01[3 16.2 4- 5213 .39]	GGCCCCAGT GTATAAGCT GCTTCGGAA TGTGGTCGG CCTG[C/T]GT CTGTCATCA CAATCTGAG TGA CTGGCA GGTGCTCCT TT	SEQ ID NO: 959

chr 9	776135 39	A	G	<i>C9orf4</i> 1	p.D29 5D	0.009 80	0.006 81	3.33 E- 02	1.44[1.05- 1.99]	TTGATTTGG ACTTACTGC ATTCTGAAT AAATCTCTT GAAA[A/G]TC TCCTGCTGTC ATAGAAAAG TTAGAACCA GGAGGAAGA C	SEQ ID NO: 960
chr 9	845625 04	A	G	<i>SPATA</i> <i>31D3</i>	p.K77 9R	0.012 25	0.000 00	2.19 E- 72	Inf	GTGGGGAAT TATCAGGGA TGCAGCCAG GAGACTGCC CCAA[A/G]A AACCATCTC TTGCATGAT CCGGAGACA TCTTCAGAG GAG	SEQ ID NO: 961
chr 9	941725 07	C	T	<i>NFIL3</i>	p.M17 0I	0.005 88	0.003 69	3.57 E- 02	1.6[.06- 2.41]	GTGGAGAGT GTTTAATGA CAGAAATAC AACTACTTG ACAC[C/T]AT CGAGGGTTC GTGCTCGTC CACAAATGA ACTCACATT GG	SEQ ID NO: 962
chr 9	960518 69	G	A	<i>WNK2</i>	p.A16 48A	0.005 39	0.000 10	1.02 E- 22	56.0 7[23. 94- 131. 32]	GCGGGGGGG ACCTGGCCC TGCCCCCAG TGCCTAAGG AGGC[G/A]GT CTCAGGGCG TGTCAGCT GCCCCAGCC CTTGGTGAG TA	SEQ ID NO: 963
chr 9	960814 33	C	T	<i>C9orf1</i> 29	p.R130 H	0.010 54	0.007 41	2.68 E- 02	1.43[1.05- 1.94]	TGCCTGTGA ATCCCTTCCT TGTACATGG TGGTCAGTG GCA[C/T]GGA ATCCCAAT AGATTGTAT ATCTGAAGG AGAAAAATA A	SEQ ID NO: 964
chr 9	964390 19	C	A	<i>PHF2</i>	p.T992 T	0.022 30	0.001 14	8.06 E- 65	20.0 2[14. 5- 27.6 5]	CCTCCACCA CGCCAGCCT CTACCACCC CGGCCTCCA CCAC[C/A]CC GGCCTCCAC CAGCACGGC CAGCAGCCA GGCCTCGCA GG	SEQ ID NO: 965

chr 9	970809 53	A	C	<i>NUTM 2F</i>	p.S689 A	0.007 84	0.000 47	5.91 E- 24	16.6 9[10. 5- 26.5 2]	AAGAGAGGT CGCTTCTTG GACTTGCTG GCAGGAGAA GGTG[A/C]TG GGCTGAGGC CTCTTTTCTG AGCAGATGG AGACTGAAG A	SEQ ID NO: 966
chr 9	106889 571	C	T	<i>SMC2</i>	p.S867 L	0.005 15	0.003 16	3.42 E- 02	1.63[1.05- 2.54]	CCTCACCAC ATATTTTCTT TAATTTTTTT GTTTTAGGA GT[C/T]AGTA AATAAAGCT CAAGAAGAG GTGACCAAG CAAAAAGAG	SEQ ID NO: 967
chr 9	113562 589	T	C	<i>MUSK</i>	p.V55 8A	0.006 62	0.004 17	2.64 E- 02	1.59[1.08- 2.35]	GAAACTGAG ACTAACAGG GATGGTCTT TTGGTTCCA GGAG[T/C]GT GTGCTGTCG GGAAGCCAA TGTGCCTGC TCTTTGAAT AC	SEQ ID NO: 968
chr 9	117170 241	G	C	<i>DFNB 31</i>	p.P562 A	0.119 36	0.117 07	6.55 E- 01	1.02[0.93- 1.13]	AACCAAAGG GCCAGCCAG GGCCTTACC ACGGACACA TCTG[G/C]GA GGGCGTTGA TATTGCCCT GGACAGCCT CGCCAGTTT CC	SEQ ID NO: 969
chr 9	127623 742	G	A	<i>RPL35</i>	p.R32 R	0.011 76	0.008 52	3.12 E- 02	1.39[1.03- 1.85]	TAGAGAGCT TGGAGGCCG CACCGCCTG TCACTTTGG CGAC[G/A]CG CAGCTGGGA CAGCTCCAC CTTCAGGTC GTCCAGCTG TT	SEQ ID NO: 970
chr 9	131094 512	G	C	<i>COQ4</i>	p.E161 D	0.012 25	0.008 51	1.55 E- 02	1.44[1.09- 1.92]	ATGATGAGG AGCTAGCGT ATGTGATTC AGCGGTACC GGGA[G/C]GT GCACGACAT GCTTCACAC CCTGCTGGG GATGCCCCAC CA	SEQ ID NO: 971

chr 9	131258 331	G	C	<i>ODF2</i>	p.Q61 7H	0.007 84	0.000 00	2.84 E- 47	Inf	TAAACCAGT CTGTGTTTCCT GTCATTTTA GATCGAACA CCA[G/C]GGG GACAAGCTG GAGATGGCG AGAGAGAAA CATCAGGCT T	SEQ ID NO: 972
chr 9	132630 457	G	A	<i>USP20</i>	p.S288 S	0.005 64	0.003 05	9.34 E- 03	1.85[1.21- 2.83]	ACCGGAGCC CATCAGAAG ATGAGTTCT TGTCCTGTG ACTC[G/A]AG CAGTGACCG GGGTGAGGG TGACGGGCA GGGGCGTGG CG	SEQ ID NO: 973
chr 9	134353 141	G	A	<i>PRRC2 B</i>	p.E147 3K	0.005 15	0.003 03	2.96 E- 02	1.71[1.1- 2.65]	CTGGTTAAC AAGATCCTC TTTCCCTTAC AGATCCCCA GAC[G/A]AG GCCTTGCCT GGAGGTCTT AGTGGCTGC AGCAGTGGG AG	SEQ ID NO: 974
chr 9	135140 020	A	G	<i>SETX</i>	p.I254 7T	0.008 58	0.005 72	2.68 E- 02	1.5[.07- 2.12]	GGGTTGTGG ATCCCAAAG GAATATTCC TCCTTTGACC TCA[A/G]TGC CCATCCTCTT CAGCAGTCG TGGGTCCTG AAGTTGGTC	SEQ ID NO: 975
chr 9	136419 800	G	A	<i>ADAM TSL2</i>	p.G42 1S	0.023 28	0.000 00	1.28 E- 12	Inf	CGAGCAGGC CGGCGGCGG GGCCTGCGA GGGGCCCCC CAGG[G/A]G CAAGGGCTT CCGAGGTAA CCAGGAGGA GGGAGGCAT GAG	SEQ ID NO: 976
chr 9	137309 155	G	A	<i>RXR4</i>	p.M25 4I	0.006 13	0.003 79	2.76 E- 02	1.62[1.08- 2.43]	CCGTGGAGC CCAAGACCG AGACCTACG TGGAGGCAA ACAT[G/A]GG GCTGAACCC CAGCTCGGT GAGTTGCAG CCTGTGCAG GG	SEQ ID NO: 977

chr 9	139333 512	G	C	<i>INPP5 E</i>	p.G12 0G	0.007 11	0.000 02	1.78 E- 34	447. 13[6 0.89- 3283 .17]	TCAGGCAGG GCGGGGAGC AGCTGTGGG CGGGGGCCC CGGG[G/C]CC CTCGCTCTG CACTGAGCC CCTGGAGGG ACTGGTCCC AT	SEQ ID NO: 978
chr 9	139701 301	G	T	<i>CCDC 183</i>	p.M45 7I	0.005 856	0.003 603	4.82 E- 02	1.63[0.95- 2.61]	GCGAGGGGA AGCTCACGT ACCTGGCTG ACAGAGTGC AGAT[G/T]GT GTCCAGGAC CGAGGAGGT AGCCCCGGG CTGGGAGGA AC	SEQ ID NO: 979
chr 9	139752 023	A	T	<i>MAMD C4</i>	p.T771 S	0.009 07	0.006 39	4.61 E- 02	1.42[1.02- 1.98]	CTCGGGCCA TGCTGCCTG GGGCCCCC AACAGACCA TACC[A/T]CT GAGACAGCC CAAGGTATG GGGGCCTGG CAGGGGCAG GG	SEQ ID NO: 980
chr 9	140008 984	G	A	<i>DPP7</i>	p.Q38 X	0.005 15	0.000 00	4.86 E- 28	Inf	TTGTTGCCG AAGCGCTCG AAGTTGAAG TGGTCCAGA CGCT[G/A]CT GGAAGAAGC GCTCCTGGA AGCCGGGGT CCGGGGCCC TG	SEQ ID NO: 981
chr 9	140120 397	G	T	<i>CYSRT I</i>	p.A14 8A	0.011 03	0.000 00	2.82 E- 52	Inf	AGCGCCAGG CCGGACTGA CCTACGCTG GCCCTCCGC CCGC[G/T]GG GCGCGGGGA TGACATCGC CCACCACTG CTGCTGCTG CC	SEQ ID NO: 982
chr X	986397 4	C	CT	<i>SHRO OM2</i>	p.L676 fs	0.005 89	0.000 10	2.57 E- 07	61.9[12.5- 307. 1]	CTGGAGGGC CGGGTTGGG AGGTGGCAC CCAGGAAGG ACCC[C/CT]T CGCTGGCAC CTATAAAGA CCACCTGAA AGAGGCCCA AGC	SEQ ID NO: 983

chr X	100856 59	C	T	WWC3	p.H52 0H	0.006 13	0.003 94	4.13 E- 02	1.56[1.03- 2.36]	GGGACGAAG ACTTACCAG GCATGGCGG CCCTTCAGC CACA[C/T]GG GGTCCCCGG GGATGGGGA AGGGCCGCA CGAGCGAGG AC	SEQ ID NO: 984
chr X	349618 39	G	A	FAM4 7B	p.P297 P	0.005 88	0.000 01	6.33 E- 31	473. 89[6 4.09- 3503 .83]	GCCCCGAGC CTCCCCGAGA CTCGCGTAT CTCATCTCC ACCC[G/A]GA GCCTCCTGA GACTGGAGT GTCCCATCT CCGCCAGA GC	SEQ ID NO: 985
chr X	370279 59	C	G	FAM4 7C	p.D49 2E	0.006 86	0.000 00	5.71 E- 37	Inf	CAGAGAAGG ACGTATCTC ATCTCCGCC CAGAGCCTC CCGA[C/G]AC TGGAGTGTC CCATCTCTG CCCAGAGCC CCCCAAGAC AC	SEQ ID NO: 986
chr X	370287 70	C	T	FAM4 7C	p.R763 C	0.008 58	0.000 01	2.98 E- 45	692. 67[9 4.87- 5057 .22]	TCTCCGCCC AGAGCCTCT TGAGACTCG CGTATCTCA TCTC[C/T]GC CCGGAGCCT CCTGAGACT GGAGTGTCC CATCTCCAC CC	SEQ ID NO: 987
chr X	436286 23	G	A	MAOB	p.T426 T	0.008 82	0.000 00	6.54 E- 48	Inf	CAGCCCCCT CCATGTAGC CGCTCCAGT GTGTGGCAG TCTC[G/A]GT GCCTGCAAA GTAAATCCT GTCCACTGG CTGGCGTAG AA	SEQ ID NO: 988
chr X	474267 57	C	T	ARAF	p.A33 7A	0.010 05	0.007 11	3.68 E- 02	1.42[1.03- 1.95]	TTGGCACCG TGTTTCGAG GGCGGTGGC ATGGCGATG TGGC[C/T]GT GAAGGTGCT CAAGGTGTC CCAGCCCAC AGCTGAGCA GG	SEQ ID NO: 989

chr X	486648 50	C	T	<i>HDAC</i> 6	p.Y17 1Y	0.005 88	0.002 90	1.98 E- 03	2.04[1.34- 3.1]	ACATGAATG AGGGAGAAC TCCGTGTCCT AGCAGACAC CTA[C/T]GAC TCAGTTTATC TGCATCCGG TATGGATGA GAACTCTGC	SEQ ID NO: 990
chr X	491059 70	G	A	<i>CCDC</i> 22	p.D54 6N	0.008 58	0.005 80	3.56 E- 02	1.48[1.05- 2.09]	GCAGCCCAC TGATACCTTT GAGGTCCCT GTGTCTGGT CAG[G/A]ATG CCAAGAAGG ACGATGCTG TTCGGAAGG CCTATAAGT A	SEQ ID NO: 991
chr X	494559 76	C	T	<i>PAGE</i> 1	p.G56 G	0.008 82	0.005 92	2.89 E- 02	1.49[1.06- 2.1]	TTGGCTGAA CCAGTTCCT GGCTATCAG CTTCAGGCT CCTG[C/T]CC TTAAAGATA AAACAAAAT TATCATTTTA AGCAGCAAC A	SEQ ID NO: 992
chr X	531153 95	G	A	<i>TSPYL</i> 2	p.E607 E	0.009 07	0.006 06	2.37 E- 02	1.5[.07- 2.1]	AAGGCAGCG ATGATGACG ACAGAGACA TTGAGTACT ATGA[G/A]A AAGTTATTG AAGACTTTG ACAAGGATC AGGCTGACT ACG	SEQ ID NO: 993
chr X	562918 53	A	G	<i>KLF8</i>	p.I108 V	0.009 56	0.006 71	4.00 E- 02	1.43[1.03- 1.98]	CAAGGCTCC TCTCCAGCC TGCTAGCAT GCTACAAGC TCCA[A/G]TA CGTCCCCC AAGCCACAG TCTTCTCCCC AGACCCTTG T	SEQ ID NO: 994
chr X	708237 81	G	C	<i>ACRC</i>	p.K21 8N	0.005 88	0.000 18	1.40 E- 22	33.7 6[17. 45- 65.3 1]	CCGACGACA ACAGTGATG ATTTCGGATG TTCCCGACG ACAA[G/C]A GTGATGATT CGGATGTTC CCGACGACA GCAGTGATG ATT	SEQ ID NO: 995

chr X	738116 48	G	A	<i>RLIM</i>	p.S501 L	0.009 80	0.000 00	1.61 E- 52	Inf	ATGTCGACC CTCTCGCCT GGCACCTGA TGAGCCTGA TGAT[G/A]AG CTTCCTTCAT TACTGCCTTC AAATAAATC TGAGCTAGT	SEQ ID NO: 996
chr X	738116 95	A	G	<i>RLIM</i>	p.S485 S	0.010 29	0.000 23	6.36 E- 41	46.1 6[26. 25- 81.1 6]	CTTCATTACT GCCTTCAAA TAAATCTGA GCTAGTTTCT GA[A/G]CTTT CACCACCGG AACTGGAAC TAGGACTGG AACTGGAAC	SEQ ID NO: 997
chr X	738117 92	C	T	<i>RLIM</i>	p.S453 N	0.010 29	0.000 01	2.96 E- 54	825. 58[1 13.6- 5999 .93]	ACTCGAACT GGAACCTGGA ACTCGAACT GGAACCAGA ACTA[C/T]TA CCACCACCA GAACCTCCT CTTCCACTCC GTGACTCTG C	SEQ ID NO: 998
chr X	100507 675	G	T	<i>DRP2</i>	p.L571 L	0.011 76	0.008 56	3.77 E- 02	1.38[1.03- 1.85]	CCTGCTTCTT GACAGGCAG GGCCAGCAA AGGCAATAA GCT[G/T]CAC TACCCCATC ATGGAGTAT TACACACCG GTATGAAGC C	SEQ ID NO: 999
chr X	100524 197	C	T	<i>TAF7L</i>	p.R372 H	0.011 03	0.007 69	2.26 E- 02	1.44[1.06- 1.95]	TGTGGGCCA CGCCAATGG CTCTCCTCAC TTCTTCAGA AAA[C/T]GCT GCAACTGTT CCTGTAGGG AAATGAGCT GTAGGGAGA G	SEQ ID NO: 1000
chr X	100745 885	C	G	<i>ARMC X4</i>	p.A77 0G	0.008 33	0.000 00	8.99 E- 34	Inf	CAGGGTGAG GTCTTGCCCT GGTGCCAAA AATAAGGTC AAGG[C/G]C AATCTTAAT GCTGTGTCT AAGGCAGAA GCTGGGATG GGT	SEQ ID NO: 1001

chr X	100746 423	G	C	<i>ARMC X4</i>	p.Q94 9H	0.009 31	0.000 00	1.04 E- 38	Inf	CTAAGGCAG AGGCTGGGG CAGGCATAA TGGGCTCTG TCCA[G/C]GT CCAGGTTGT GGCCAGTTT TCAGGGTGA GGTCTTGCC TG	SEQ ID NO: 1002
chr X	101971 960	C	T	<i>ARMC X5- GPRA SP2</i>	p.S721 S	0.011 52	0.007 33	5.08 E- 03	1.58[1.17- 2.13]	TGACTATTG ACTATCACA CACTGATTG CCAACATA TGTC[C/T]GG GTTTCTCTCC TTATTAACC ACAGCCAAT GCGAGAACG A	SEQ ID NO: 1003
chr X	102754 916	C	T	<i>RAB40 A</i>	p.E257 K	0.008 33	0.001 96	5.24 E- 11	4.28[2.95- 6.22]	GTGCAGTTT TTGGGTGGG CTCTGGGGT GGGCAGACG ATCT[C/T]CA CTTTGCAGA GGCTGCTCT TGTGAGTGG AGCTGGTGG TG	SEQ ID NO: 1004
chr X	114425 545	G	A	<i>RBMX L3</i>	p.R514 Q	0.007 60	0.000 02	5.32 E- 32	323. 05[4 4.09- 2367 .01]	AGCGACCGC TACGGAGTA GGAGGCCAC TATGAGGAG AACC[G/A]A GGCCACTCT CTGGATGCC AACAGCGGA GGCCGTTCA CCC	SEQ ID NO: 1005
chr X	114426 551	C	T	<i>RBMX L3</i>	p.Y84 9Y	0.012 01	0.000 12	4.17 E- 46	101. 99[4 0.62- 256. 12]	ACGCCTACA GTGGGGGCC GTGACAGTT CCAGCAACA GTTA[C/T]GA CCGGAGCCA CCGCTATGG AGGAGGAGG CCACTACGA AG	SEQ ID NO: 1006
chr X	120008 980	G	C	<i>CT47B I</i>	p.P182 R	0.012 99	0.000 01	1.16 E- 68	1046 .3[14 4.66- 7567 .63]	CGACGCAGC CTCCTGGAT CAGGCCGAG GCCCTCGCC TTCT[G/C]GG GCTGCAGCC CCTGCACCC AGCCTCTGG GACAGCAGC AG	SEQ ID NO: 1007

chr X	124455 258	G	C	<i>LOC10 012952 0</i>	p.K43 0N	0.017 40	0.000 00	8.76 E- 72	Inf	ACAGCCACA GCATGAAGA AAGATCCAG TGATGCCCC AGAA[G/C]AT GGTCCCCCT GGGGGACAG CAACAGCCA CAGTCTGAA GA	SEQ ID NO: 1008
chr X	140993 751	A	G	<i>MAGE C1</i>	p.Q18 7Q	0.013 24	0.002 19	4.36 E- 16	6.11[3.92- 9.52]	CTTTAGTGA GTATTTTCCA GAGTTCCCC TGAGAGTAC TCA[A/G]AGT CCTTTTGAG GGTTTCCCC AGTCTCCAC TCCAGATTC	SEQ ID NO: 1009
chr X	140994 691	T	A	<i>MAGE C1</i>	p.C501 S	0.014 71	0.000 00	9.16 E- 80	Inf	CTCCTCCACT TTATTGAGT CTTTTCCAG AGTTCCCCT GAG[T/A]GTA CTCAAAGTA CTTTTGAGG GTTTTCCCCA GTCTCCTCT	SEQ ID NO: 1010
chr X	149100 775	C	T	<i>CXorf4 0B</i>	p.G15 5E	0.009 07	0.005 92	1.69 E- 02	1.54[1.1- 2.15]	AACATTCTT TTCAGGAGC CCACACTTG TCACACTTC ATGC[C/T]CC AAAGGGATC AGGTGCTCT GGGATGTCT ACCTGGAAT AC	SEQ ID NO: 1011
chr X	150908 168	G	T	<i>CNGA 2</i>	p.G11 3V	0.010 54	0.007 65	4.45 E- 02	1.38[1.01- 1.88]	GGGCCTGAA CTCCAGACT GTGACCACA CAGGAGGGG GATG[G/T]CA AAGGCGACA AGGATGGCG AGGACAAAG GCACCAAGT AC	SEQ ID NO: 1012
chr X	153295 986	C	T	<i>MECP 2</i>	p.K44 3K	0.018 87	0.000 00	3.45 E- 102	Inf	TGGCGGCGG TGGCAACCG CGGGCTGAG TCTTAGCTG GCTC[C/T]TT GGGGCAGCC GTCGCTCTC CAGTGAGCC TCCTCTGGG CA	SEQ ID NO: 1013

[0448] Table 2. Variants associated with infertility symptom of endometriosis

Chr	Position	Reference Allele	Alternate Allele/ Minor Allele	Gene	Amino Acid position	Chronic Pelvic Pain MAF	Infertility MAF	p value	OR[L95 -U95]	Context Sequence	SEQ ID NO
chr 11	5444040	C	T	<i>OR51Q1</i>	p.L204F	0.00894	0.02899	2.59E-02	0.30	CTGTGCTG ACATCAGG CTCAACAG CTGGTATG GATTTGCT[C/T]TTGCC TTGCTCAT TATTATCG TGGATCCT CTGCTCAT TGT	SEQ ID NO: 129
chr 19	53793162	C	T	<i>BIRC8</i>	p.A156T	0.00000	0.00725	1.16E-03	0.00	GAAGTCTG ATTCAATT CATTTTCT GTAGTGTC TTTCTGAG[C/T]GCTCA CTAGATCT GCAACAAG AACCTCAA GCGTTTTA TAG	SEQ ID NO: 531
chr 2	238973062	A	G	<i>SCLY</i>	p.K60E	0.00000	0.00730	1.11E-03	0.00	AACGACTC CCCTGGAG CCAGAAAGT TATCCAGG CCATGACC [A/G]AGGC CATGTGGG AAGCCTGG GGAAATCC CAGCAGCC CGTA	SEQ ID NO: 592
chr 22	50315363	C	A	<i>CRELD2</i>	p.D182E	0.02820	0.06159	4.03E-03	0.44	ACATGGGG TACCAGGG CCCGCTGT GCACTGAC TGCATGGA [C/A]GGCT ACTTCAGC TCGCTCCG GAACGAG ACCCACAG CATCT	SEQ ID NO: 637
chr 4	81967240	C	T	<i>BMP3</i>	p.T222M	0.00000	0.00725	1.16E-03	0.00	GCCAAAGA AAATGAAG AGTTCCTC ATAGGATT TAACATTA	SEQ ID NO: 706

										[C/T]GTCCA AGGGACGC CAGCTGCC AAAGAGG AGGTTACC TTTT	
--	--	--	--	--	--	--	--	--	--	---	--

[0449] Table 3. Variants associated with pelvic pain symptom of endometriosis

Chr	Position	Reference Allele	Alternate Allele/Minor Allele	Gene	Amino Acid position	Chronic Pelvic Pain MAF	Infertility MAF	p value	OR[L95-U95]	Context Sequence	SEQ ID NO
chr 2	141232800	C	T	<i>LRPIB</i>	p.A3178T	0.00000	0.01087	7.31E-05	0.00	GCCCAG TAGAGT CTACGA TTAACA TAATCT ATTGTT AGTG[C/T]CATA GGTCTA GAAATC TTGGTT TCTATG ACAACA CTCTGA	SEQ ID NO: 577
chr 6	56033094	G	A	<i>COL21A1</i>	p.T343M	0.06389	0.11590	2.12E-03	0.52	TACTAA GAGACG AATTTG GTGCCA GCCTTC ATCAAA CAAC[G/A]TCTA CAAAAA GAAAGT GTGGAA GATTCA TAAATA AAGCCC	SEQ ID NO: 786
chr 6	85473758	C	T	<i>TBX18</i>	p.G48R	0.48050	0.57660	2.41E-03	0.68	GCGCCG CCGCCG CGGCTG CAGCCT CCGTCG TCCACG GCCC[C/T]CGCC GCCTCT TCGGCG CCCAGT TTTCGC CGCTTC TTCTGA	SEQ ID NO: 789
chr 9	117170241	G	C	<i>DFNB31</i>	p.P562A	0.10070	0.16060	4.01E-03	0.59	AACCAA AGGGCC AGCCAG GGCCTT ACCACG	SEQ ID NO: 969

											GACACA TCTG[G/ C]GAGG GCGTTG ATATTG CCCTGG ACAGCC TCGCCA GTTTCC	
--	--	--	--	--	--	--	--	--	--	--	---	--

[0450] Table 4. Additional variants associated with endometriosis.

End ome trio sis pati ent Fre que ncy	Loc al pop ulat ion Co ntr ol Fre que ncy	gno mA D Fre que ncy	P (C his q test)	O R (o dd s R ati o)	L9 5 (lo we r lim it 95 % con fid enc e Int erv al)	U9 5 (up per lim it 95 % Co nfi den ce Int erv al)	C H R	SNP	Bas e Pai r Pos itio n	Mi nor Alle le	Ma jor Alle le	Context Sequence	SEQ ID NO
0.3 05 5	0.2 8	0.28 83	4.4 9E - 05	1.1 3	1.07	1.20	1	rs3 410 898 9	16, 08 2,1 27	C	T	GCATCAGGTATTTTTACCCACATT TACCCACCAGATTCT[T/C]GCTA TGAAGCCACAAGGGACAAACCTG GGTTGGCAACCCC	SEQ ID NO: 101 4
0.1 84 4	0.1 49 4	0.15 91	1.7 5E - 12	1.2 9	1.20	1.38	1	rs2 235 529	22, 45 0,4 87	T	C	AAGCATCTGTGCCCTAAAGCTG ATGGCGGCTCCTCCAGC[C/T]TTC TCTACCTGGTTCTGGTGTCCAGCC CTTGGA CTCCAGG	SEQ ID NO: 101 5
0.2 29 4	0.1 99 2	0.20 86	5.0 7E - 08	1.2 0	1.12	1.28	1	rs1 204 208 3	22, 47 2,7 32	A	G	CATGAGCCACCTTGCCTGGCCGG AAATTCTTAATGAGAAA[G/A]TCT CTTGGAGGAAATGCTCTTCTAAC TTCAAGAACAGCC	SEQ ID NO: 101 6
0.4 37 4	0.4 04 2	0.42 05	1.0 7E - 06	1.1 5	1.09	1.21	1	rs4 623 666	22, 48 0,3 12	G	A	ATCTTCAGCCTCCTACCAGCAAC TATGCACACAGAAGCCC[A/G]GC CGGTATCCCCACAGAGGCAGACG CCCCGGCACTGCCTT	SEQ ID NO: 101 7
0.1 12 6	0.0 96 37	0.09 91 5	9.4 3E - 05	1.1 9	1.09	1.30	1	rs1 206 112 4	97, 98 9,7 51	T	C	AGTTGAAACTCACAACTGCAGG AATATAGTCATTGGGGT[C/T]CCT TAGATGCAGAAAAGAAAATTAAC TACAGCGAGTTATG	SEQ ID NO: 101 8
0.3 21 6	0.3 48 7	0.33 88	3.6 5E - 05	0.8 9	0.84	0.94	2	rs2 349 415	49, 24 7,8 32	T	C	AAAAC TTATTTCATAAAAACAGG TGTCAGGCTGGATTTGA[T/C]CCA TTGGCTGTAGTTCAGTGACACTG TCCTAGATCGTGGA	SEQ ID NO: 101 9
0.0	0.0	0.	1.2	1.	1.	1.	2	rs1	98,	G	A	TCCGGGGAACACGATTCCACCCA	SEQ

95 59	77 47	08 62 5	4E - 06	2 6	15	38		702 577 8	63 7,5 04			TCAGTGGGTGCTAGGTC[A/G]AGG GTTCAGTTCTATGTCCTTCAGCAC TTATGAACTGAG	ID NO: 102 0
0.1 04 4	0.0 87 78	0. 09 06 2	2.5 5E - 05	1. 2 1	1. 11	1. 32	2	rs1 702 629 2	98, 67 7,1 64	A	G	GGATGAATGGAACTTGATTCTC TTAATACAGTCCACTTG[G/A]GCT CCATTGTCTTCACAGCAACCATT TGCTGGATTATT	SEQ ID NO: 102 1
0.4 03 6	0.3 74 4	0. 38 27	1.4 7E - 05	1. 1 3	1. 07	1. 20	2	rs7 555 03	13 5,1 44, 45 4	A	G	TATGCTTAGGAAATATGTATATA TGGGATATCTCAAAATA[A/G]GG AAAAGTTGGAGTGAAGATTAAAA TAGAAAATAACAAAA	SEQ ID NO: 102 2
0.1 66 2	0.1 88	0. 18 22	4.8 1E - 05	0. 8 6	0. 80	0. 93	2	rs1 017 799 6	21 9,7 46, 56 1	C	T	CTATGTGAATGTGACTGAAACAT ATCTGTGGGAGTGGGCT[T/C]GTG GGGAACCCTGTGTGTATGGGCAT CTATTCCTGGGGAT	SEQ ID NO: 102 3
0.2 85 2	0.2 59	0. 26 3	1.4 7E - 05	1. 1 4	1. 08	1. 21	2	rs3 882 08	22 5,9 38, 99 6	T	C	ACAGTTAATATTGACTGCTTTGTT CATTGATACATTCCCT[T/C]GACC TAGACCATTGCTGGGCACATAGT AGGCTCTCAGTAA	SEQ ID NO: 102 4
0.1 81 8	0.1 61 3	0. 16 95	5.2 8E - 05	1. 1 6	1. 08	1. 24	3	rs6 792 001	6,1 06, 25 1	A	G	CTATTGATTTTTGAGGTAGATATT GATGCAATTAGAGATA[A/G]GCTT TAGGAAGATCTTCCTGGAAGTGG TATATAAATAGTT	SEQ ID NO: 102 5
0.2 33 8	0.2 58	0. 25 84	6.2 6E - 05	0. 8 8	0. 82	0. 94	3	rs6 777 088	8,7 86, 48 7	G	A	CACCCTTCAGATCATAAAACAAT AGAATTTGAGAGCTGCG[A/G]CT ATAGCACTGCCACTAAGTCACTG TTGGCTTTAAGCAAG	SEQ ID NO: 102 6
0.1 51 3	0.1 74 4	0. 16 82	1.0 5E - 05	0. 8 4	0. 78	0. 91	3	rs4 293 672	25, 91 3,4 15	T	C	AATTGACACACTACTGAAAAGAA AAGAGAATTAGAACAAC[T/C]TG CCTGGAGTTAAAGTCCCTTAGTT AATGGATAAGTCACC	SEQ ID NO: 102 7
0.1 24 4	0.1 46	0. 13 44	9.2 1E - 06	0. 8 3	0. 77	0. 90	3	rs1 684 322 5	10 0,8 01, 25 7	G	T	TCTGGTGTCAATTAAGGAAGCAGG TTACAGGCCAGCATATC[T/G]TCA AATAGCTACACAGGTGTTAGAAC TGCATGGTCTTATA	SEQ ID NO: 102 8
0.1 40 5	0.1 22 6	0. 12 6	8.9 8E - 05	1. 1 7	1. 08	1. 27	3	rs4 680 277	15 6,2 45, 78 1	A	G	GTGCTAATTATCCAGAATCAGCT GCAGTTGCTACCATGGA[A/G]GTA ACCAGCTCTGCCCAGTGGGTTCT CCTGTGCCCTACAG	SEQ ID NO: 102 9
0.1 39 9	0.1 20 8	0. 12 59	2.7 8E - 05	1. 1 8	1. 09	1. 28	3	rs6 795 731	15 6,2 62, 46 0	T	C	TAGTGAAGAAAACATCATGCTGG TTATGTTACCATTTTTTC[C/T]CAGG CAACCAGGGTTATGGAAGAAAG GACTCATTAATGGC	SEQ ID NO: 103 0
0.2 68	0.2 98	0. 28	1.4 3E	0. 8	0. 81	0. 92	4	rs1 250	56, 00	A	C	GATGTGGTCATATGAAGGCTTGA CTGGGGCTGAAGAATAC[C/A]TTT	SEQ ID

3	8	8	- 06	6				509 6	6,1 02			CTGGTGTGACTCACTCACATGAC TATTGGCAAGAGAA	NO: 103 1
0.2 06 8	0.1 82 6	0. 19 07	6.9 6E - 06	1. 1 7	1. 09	1. 25	4	rs1 001 428 5	16 1,3 07, 97 2	A	G	CCTTGGAGAGTTCCTCCACTTCTC TCTGACAATTAATAATC[G/A]GTGT TTGCTGAGATTAGACATTTTTTTC TTCTCTGTTTAG	SEQ ID NO: 103 2
0.0 46 11	0.0 35 63	0. 03 23	5.5 0E - 05	1. 3 1	1. 15	1. 49	4	rs1 265 036 4	18 6,3 65, 99 8	A	G	TGGTGGTAGGGAGACCTTTTGGT GGTATTTGAATTAAACA[G/A]TAT CATTTTCTTTAAACCAACTCCAC AGACTACAAAAAT	SEQ ID NO: 103 3
0.0 54 81	0.0 40 1	0. 04 79	1.0 6E - 07	1. 3 9	1. 23	1. 57	4	rs4 611 976	18 8,9 90, 95 5	G	T	GTGTTGGTCGGTACAGTTCTAGA AGGAAAGCTCTGAGCTG[T/G]GC CCCTCTCTCCAGGTGGAATTAGA TTTTATATATTTCACT	SEQ ID NO: 103 4
0.3 72 7	0.3 46 6	0. 34 37	7.3 4E - 05	1. 1 2	1. 06	1. 19	5	rs4 128 741	76, 42 3,9 67	T	C	ATCCCCATTTCCTTTACAATTATA ATTGCCTCCATATTGT[C/T]CAAG GACCATAGTTACCACTTGACCCA GAGCCTCTCCCTT	SEQ ID NO: 103 5
0.4 17 3	0.3 83 7	0. 39 39	6.0 2E - 07	1. 1 5	1. 09	1. 22	5	rs1 252 105 8	76, 42 6,9 87	A	C	AGCTGTTCTCAGATACCAGACTG GAATAAACGAGAGACAT[C/A]TG GAGAAAGGAGACCTCTTCTATC CCAACAGGACTGTGT	SEQ ID NO: 103 6
0.1 80 7	0.1 56 6	0. 16 45	1.7 7E - 06	1. 1 9	1. 11	1. 28	6	rs6 456 259	19, 76 1,7 18	G	A	GCTCACCAAGCAAGATTCTCTC ATCCCCTGCCACTCCCT[A/G]TTT AATGCCTTTGTAAAACTGTAAT TTGGTGAATCCCAA	SEQ ID NO: 103 7
0.1 87 4	0.1 65 9	0. 16 15	2.8 8E - 05	1. 1 6	1. 08	1. 24	6	rs5 634 40	15 1,2 88, 99 1	C	T	GCTACTCTTTTCTTCCAAAATACT CTCTCCTCAGCAGCCA[T/C]AGAG ACTGAAACCTAATGAAGCCCTGT TGCCTTCTACTT	SEQ ID NO: 103 8
0.1 00 3	0.1 18	0. 12 62	6.9 5E - 05	0. 8 3	0. 76	0. 91	6	rs9 347 099	16 6,3 27, 88 6	T	G	TCATTGGGAGTTATGAGCACATT TCATAAACATAATTCCA[G/T]GGG TTCGCCTGTGATGACATCATTCT TTTACAAGGTTT	SEQ ID NO: 103 9
0.4 48 8	0.4 10 7	0. 41 52	2.0 1E - 08	1. 1 7	1. 11	1. 23	7	rs1 177 380 4	27, 20 6,6 88	G	T	CTCCCCCTGCCCCAATTCCTAAC AGAAAGCAGCGACTCC[T/G]AGA ACAGGGGTAATCAAATTCACGTG TGGATACTGTGCCT	SEQ ID NO: 104 0
0.1 70 4	0.1 91 6	0. 18 29	9.2 3E - 05	0. 8 7	0. 81	0. 93	7	rs1 153 519 1	37, 74 7,2 76	G	A	AGGAAAATAAATTATGGAGACAT TAAGTAAATTGCCCAAG[A/G]TG GCCAGCTAGTAAATAATAAAGG CAAGATTTTAGAGCC	SEQ ID NO: 104 1
0.2 47 9	0.2 24 6	0. 19 85	5.6 7E -	1. 1 4	1. 07	1. 21	8	rs1 734 224	60, 82 8,6	G	A	TAATGAATCTGAGTGGGATAGTG ATCAGAATAAGGAAGTA[A/G]GG CCAATAACATTTCTGGGTAACCT	SEQ ID NO:

			05					2	97			GCCATGAGCCAAGCA	104 2
0.0 61 99	0.0 79 25	0. 08	2.8 8E - 06	0. 7 7	0. 69	0. 86	9	rs9 695 167	10 6,1 69, 26 8	A	C	TTATAGTCCCAAGTAGTCAGAGA TGGACTGTATAATATGC[C/A]GGG CACAGGGCAAAACAAGAATGAG GGAAGTTGTTGACAG	SEQ ID NO: 104 3
0.3 57 9	0.3 91 9	0. 38 61	4.6 4E - 07	0. 8 7	0. 82	0. 92	1 0	rs1 125 314 1	5,4 22, 19 6	C	A	AGCTATCATTCCTCCAGTGTGAAC CTCAAGTCATCAGATTG[A/C]ATC TCCCCACCTGCCATTGTTTTATC ACCTACCAACACC	SEQ ID NO: 104 4
0.1 68 1	0.1 42 5	0. 13 27	1.6 2E - 07	1. 2 2	1. 13	1. 31	1 0	rs1 125 610 6	9,2 22, 22 8	C	A	TGAAATTGAAGTGGTGTATGA ATCACATATGATAGATT[A/C]GGC AATTGAGTTATATTTTATATCTG CTTATCTCTCTAA	SEQ ID NO: 104 5
0.4 00 8	0.3 73 4	0. 36 94	4.3 7E - 05	1. 1 2	1. 06	1. 19	1 3	rs7 997 707	46, 36 0,6 78	A	G	GGCTGGAGGTGCGAAAGACTCTAA TCTGTTTCACTGTTTAC[G/A]TGTT CAGTCAGTTCTCTCATTGGCAAA ATATTTATCTCAA	SEQ ID NO: 104 6
0.1 63 6	0.1 84 8	0. 17 26	7.4 9E - 05	0. 8 6	0. 80	0. 93	1 3	rs9 317 519	66, 13 7,5 62	C	T	TGTTAAGTTATTCCAATAATAAA ATGTCATCCATAGGTTA[T/C]TGT CACGTTTTAATATAAGACTTCTA ATCAAATTCCTGGG	SEQ ID NO: 104 7
0.1 58 9	0.1 39 5	0. 13 05	5.4 0E - 05	1. 1 7	1. 08	1. 26	1 3	rs3 362 37	11 0,4 96, 41 0	T	C	TGGCTTCTTCGCAACTTGCATAG AGGCTACCTCTGTGTCC[C/T]CTT ATGGCTCGATAGCTCATTCTTTT TATCCCCAAATAA	SEQ ID NO: 104 8
0.3 53 4	0.3 26 6	0. 32	3.8 0E - 05	1. 1 3	1. 07	1. 19	1 4	rs1 049 844 1	52, 54 4,2 24	G	A	ATAAACATAGTTATGCTTCATTA CTCTGGTACAGAAACCC[G/A]GTT CATTAGCCATTGAGAATGATTGT GATATCCAAAATGA	SEQ ID NO: 104 9
0.3 14 5	0.2 87 1	0. 28 55	1.3 6E - 05	1. 1 4	1. 07	1. 21	1 4	rs7 157 151	52, 57 1,5 83	T	C	TGTATCCAACCATGGGAAAAAGA CTTAGCTACATTGTATA[T/C]ATT TGATGAGTAACGTGTTTATAATA CAACAAAAAGTGAA	SEQ ID NO: 105 0
0.1 25 6	0.1 08 7	0. 11 31	9.9 4E - 05	1. 1 8	1. 08	1. 28	1 4	rs1 258 682 8	71, 18 6,5 13	T	C	TTGTGCTGCCTGAGAGGAGAGGG AGCATCTCACCATCTCC[C/T]GCC TTGGTATCTTTTATTCTTTAGGAC TCAGCTCAGGTTT	SEQ ID NO: 105 1
0.4 29 7	0.4 60 9	0. 45 72	5.7 3E - 06	0. 8 8	0. 83	0. 93	1 4	rs1 951 521	10 0,7 43, 42 1	G	A	AATAAGTGAAAGAACTAGCAGTG CAGCTAGTAAATCTAAC[G/A]TGG TTCTTTTTTGACAACTGACACCAG AACCTTAATCAT	SEQ ID NO: 105 2
0.3 16 7	0.3 43 6	0. 33 78	3.9 7E - 05	0. 8 9	0. 84	0. 94	1 5	rs7 181 230	40, 36 0,7 41	G	A	AAAAAACCTTACATTAGCATAA AATCTGTAACAGGAGTG[A/G]AA TGGAATAACAAGTTCTTGAGAG AACGAAATAATGTAA	SEQ ID NO: 105

													3
0.5 06 9	0.4 79 4	0. 47 46	7.2 8E - 05	1. 1 2	1. 06	1. 18	1 5	rs1 244 270 8	47, 14 4,3 86	C	T	TTGCCTTTAGGACAGGACTGTTCT TAGTCCTCTCCAGTTCT[T/C]ACTCT ATTGTAAAGTTTCTGAAAGTGCC TCAGGTATTTCA	SEQ ID NO: 105 4
0.4 95 5	0.4 66	0. 47 12	1.7 9E - 05	1. 1 3	1. 07	1. 19	1 6	rs1 085 243 2	66, 40 2,5 15	C	T	AGAATCTTAGGCTCATTTTGCCC ACATGGACCCATGACTG[T/C]TCC CTGTATCCTCTCTCTGCACCCCCT CAGTCACACTGAA	SEQ ID NO: 105 5
0.1 22 9	0.1 04 9	0. 10 56	2.6 0E - 05	1. 2 0	1. 10	1. 30	1 6	rs1 528 28	72, 12 3,8 86	T	C	CAGTGTCTACATCACTGACCTCT GTGGTATTTCTCTCTGC[T/C]TAT GACTGAGGGTAGAATCCTCTGGT CCTTTTTTCCCCAA	SEQ ID NO: 105 6
0.3 70 5	0.3 43	0. 34 88	2.6 9E - 05	1. 1 3	1. 07	1. 19	1 7	rs8 076 465	66, 51 3,0 25	A	G	GAGCCAGGTCATAGATGTAGCTT GTTTTGAAGTCAAGTGC[A/G]TTC CTGGAGATCCGGTTTTGAAATGG GTCACTGTAAGGTG	SEQ ID NO: 105 7
0.3 70 9	0.3 43 2	0. 34 75	2.4 5E - 05	1. 1 3	1. 07	1. 19	1 7	rs2 907 373	66, 53 3,6 55	A	G	CCCTTAGCTTGTCAAGTTAGCCTG GCCAGAGTCTGGGGCC[A/G]ACT GTTCCACTGGGGCCGTCGACTATG ACACTCTGCTGTCC	SEQ ID NO: 105 8
0.2 33 7	0.2 10 9	0. 20 7	6.3 1E - 05	1. 1 4	1. 07	1. 22	1 8	rs2 175 565	46, 07 9,8 52	G	A	GACGGTGAGGAGCGGGTGATGG GGTAATTCCCGGAATGCA[G/A]A CTGTAACCAGGGCAGTCAGAACA AGGATTGTTAACCTGC	SEQ ID NO: 105 9
0.3 78 8	0.3 52 5	0. 36 17	7.4 7E - 05	1. 1 2	1. 06	1. 18	1 8	rs3 900 176	74, 73 9,0 22	T	C	GTGAGTCGCCACTGTTGGCTTATT TTATGTATTTGCATCG[T/C]TCCC ATCTAAATGGGGATTCCCACT TCATAGGCCAGTA	SEQ ID NO: 106 0
0.0 71 72	0.0 57 86	0. 06 16 4	2.3 5E - 05	1. 2 6	1. 13	1. 40	2 0	rs6 110 759	15, 69 3,9 77	G	A	GTAATTATAAAGCAGCGGAATCT CCTGCTTTATGAACCTTT[A/G]GTT CTGGGCTTCAGCTCTGTATTAGTC TGTTCTCAGACTG	SEQ ID NO: 106 1
0.2 43 2	0.2 2	0. 23 04	5.6 7E - 05	1. 1 4	1. 07	1. 21	2 0	rs6 043 979	16, 45 1,6 42	C	A	AATTCTCAGATCCACCAGTGAGA CAGAAAACATAGGAGAC[A/C]GG AAAAGAAGAATCAATGGGAAG TGAAAAAAGACAGGG	SEQ ID NO: 106 2
0.0 27 77	0.0 19 14	0. 01 66 3	8.7 2E - 06	1. 4 6	1. 24	1. 73	2 1	rs1 170 282 6	41, 90 8,9 35	T	C	AAATGCTCCTAGAACTGCAAAAC ACCTAACTTATTCCAAA[C/T]TTT CCGGATGAAAAGGCAGAGGATTT TCTACTCCCATTTC	SEQ ID NO: 106 3
0.2 37 5	0.2 62 3	0. 24 8	4.6 8E - 05	0. 8 8	0. 82	0. 93	2 2	rs1 296 795	18, 02 1,7 60	G	A	TCTCTTTCCAGGTTAAATGTTGTT CATTGCGTCCTTTCCC[A/G]AAGA GTCTGTTCCCATAGAGAAGCATG GCACAAAGTGTGC	SEQ ID NO: 106

													4
0.0 77	0.0 94 21	0. 09 07 6	1.6 1E - 05	0. 8 0	0. 73	0. 89	2 2	rs7 364 90	45, 33 8,2 13	T	C	CAGCCGATGGGCTCTGCCAGATT CCTGATCCACAGTAGGA[C/T]CCT GGGGGCACCCTCTGCCCCGAGGAC CCTGGAACACACAG	SEQ ID NO: 106 5

[0451] While exemplary embodiments of the disclosure have been shown and described herein, it will be apparent to those skilled in the art that such embodiments are provided by way of example only. It is not intended that the disclosure be limited by the specific examples provided within the specification. While the disclosure has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the disclosure. Furthermore, it shall be understood that all embodiments of the disclosure are not limited to the specific depictions, configurations or relative proportions set forth herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the disclosure described herein may be employed in practicing the disclosure. It is therefore contemplated that the disclosure shall also cover any such alternatives, modifications, variations or equivalents. It is intended that the following claims define the scope of the disclosure and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED:

1. A method comprising: (a) sequencing or genotyping a nucleic acid sample obtained from a subject having endometriosis, suspected of having endometriosis, or suspected of having a risk of developing endometriosis using a high throughput method; and (b) detecting one or more genetic variants in said nucleic acid sample, wherein said one or more genetic variants are listed in Table 1, Table 2 or Table 3.
2. The method of claim 1, wherein said high throughput method comprises nanopore sequencing.
3. The method of claim 1 or 2, wherein said nucleic acid sample comprises RNA.
4. The method of claim 3, wherein said RNA comprises mRNA.
5. The method of claim 1 or 2, wherein said nucleic acid sample comprises DNA.
6. The method of claim 5, wherein said DNA comprises cDNA, genomic DNA, sheared DNA, cell free DNA, fragmented DNA, or PCR amplified products produced therefrom, or any combination thereof.
7. The method of claim 5, wherein said DNA comprises DNA from an endometriosis lesion or peritoneal fluid.
8. The method of any one of claims 1-7, wherein said one or more genetic variants comprise a genetic variant defining a minor allele.
9. The method of any one of claims 1-7, wherein said one or more genetic variants comprise at least about: 5, 10, 15, 20, 25, 50, 75, 100, 150, 200, 250, 500, or more genetic variants defining minor alleles.
10. The method of any one of claims 1-9, wherein detection of said one or more genetic variants has an odds ratio (OR) for endometriosis of at least about: 1.5, 2, 5, 10, 20, 50, 100, or more.
11. The method of any one of claims 1-10, wherein said one or more genetic variants comprise a synonymous mutation, a non-synonymous mutation, a stop-gain mutation, a nonsense mutation, an insertion, a deletion, a splice-site variant, a frameshift mutation, or any combination thereof.
12. The method of any one of claims 1-11, wherein said one or more genetic variants comprise a protein damaging mutation.
13. The method of any one of claims 12, wherein said one or more genetic variants further comprise a protein damaging or loss of function variant in one or more genes selected from the group consisting of *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR*, and any combinations thereof.

14. The method of any one of claims 1-12, wherein said one or more genetic variants are comprised in *GAT2*, *CCDC169*, *CASP8AP2*, *POU2F3*, *CD19*, *IGSF3*, *GLI3*, *PEX26*, *OLIG3*, *CIB4*, *NKX3-2*, *CFTR* or a combination thereof.
15. The method of any one of claims 1-13, further comprising detecting one or more additional variants defining a minor allele listed in Table 4.
16. The method of any one of claim 1-15, wherein said one or more genetic variants are identified or weighted based on a predictive mathematical or computer programmed algorithm.
17. The method of any one of claims 1-16, wherein said one or more genetic variants are identified based on reference to a database.
18. The method of any one of claims 1-17, further comprising identifying said subject as having endometriosis or being at risk of developing endometriosis.
19. The method of claim 18, wherein said identifying said subject as having endometriosis or being at risk of developing endometriosis is with a specificity of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.
20. The method of any one of claims 18-19, wherein said identifying said subject as having endometriosis or being at risk of developing endometriosis is with a sensitivity of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.
21. The method of any one of claims 18-20, wherein said identifying said subject as having endometriosis or being at risk of developing endometriosis is with an accuracy of at least: 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99%.
22. The method of any one of claims 18-21, wherein said subject is identified as having endometriosis.
23. The method of claim 22, wherein said subject is asymptomatic for endometriosis.
24. The method of claim 22, wherein said subject is symptomatic for endometriosis.
25. The method of any one of claims 18-21, wherein said subject is identified as being at risk of developing endometriosis.
26. The method of any one of claims 1-25, further comprising administering a therapeutic to said subject.
27. The method of claim 26, wherein said therapeutic comprises hormonal therapy, an advanced reproductive technology therapy, a pain managing medication, or any combination thereof.
28. The method of claim 26, wherein said therapeutic comprises hormonal contraceptives, gonadotropin-releasing hormone (Gn-RH) agonists, gonadotropin-releasing hormone (Gn-RH) antagonists, progestin, danazol, or any combination thereof.

29. The method of any one of claims 26-28, wherein said therapeutic comprises a pain medication.
30. The method of claim 29, wherein said pain medication comprises a nonsteroidal anti-inflammatory drug (NSAID), ibuprofen, naproxen, an opioid, a cannabis-based therapeutic, or any combination thereof.
31. The method of any one of claims 1-26, wherein said one or more genetic variants are listed in Table 1.
32. The method of any one of claims 1-26, wherein said one or more genetic variants are listed in Table 2.
33. The method of any one of claims 1-26, wherein said one or more genetic variants are listed in Table 3.
34. The method of any one of claims 1-33, further comprising identifying said subject as having endometriosis-associated infertility or being at risk of developing endometriosis-associated infertility.
35. The method of claim 34, further comprising administering assisted reproductive technology therapy to said subject.
36. The method of claim 35, wherein said assisted reproductive technology therapy comprises *in vitro* fertilization, gamete intrafallopian transfer, or any combination thereof.
37. The method of claim 34, further comprising administering intrauterine insemination or ovulation induction.
38. The method of any one of claims 1- 37, wherein said subject is a mammal.
39. The method of claim 38, wherein said mammal is a human.
40. The method of any one of claims 2-39, wherein said nanopore sequencing is performed with a biological nanopore, a solid state nanopore, or a hybrid nanopore.
41. The method of any one of claims 1-40, wherein said one or more genetic variants further comprise a mutation in *SEPT10*, *TNFRSF6B*, *UGT2B28*, *USP17L2* or any combination thereof.
42. The method of claim 41, wherein said one or more genetic variants comprise a mutation in *SEPT10* and wherein said mutation comprises a missense mutation.
43. The method of claim 41, wherein said one or more genetic variants comprise a mutation in *TNFRSF6B* and wherein said mutation comprises a homozygous or hemizygous mutation.
44. The method of claim 41, wherein said one or more genetic variants comprise a mutation in *UGT2B28* or *USP17L2* and wherein said mutation comprises a hemizygous deletion.

45. The method of any one of claims 1-44, wherein the one or more variants are identified based on a predictive computer algorithm.
46. The method of claim 45, wherein said predictive computer algorithm is Polyphen 2, Sift, Mutation Accessor, Mutation Taster, FATHMM, LRT, or MetaLR.
47. The method of any one of claims 1-46, further comprising administering a hormonal therapy to said subject.
48. The method of claim 47, wherein the hormonal therapy comprises administration of hormonal contraceptives, gonadotropin-releasing hormone (GnRH) agonists, gonadotropin-releasing hormone (GnRH) antagonists, progestin, danazol, or any combination thereof.
49. The method of any one of claims 1-46, further comprising administering to the subject an assisted reproductive therapy.
50. The method of claim 49, wherein the assisted reproductive therapy comprises in vitro fertilization, intrauterine insemination, ovulation induction, gamete intrafallopian transfer, or any combination thereof.
51. The method of any one of claims 1-46, further comprising administering to the subject a pain medication.
52. The method of claim 51, wherein the pain medication comprises a nonsteroidal anti-inflammatory drug (NSAID), ibuprofen, naproxen, an opioid, a cannabis-based therapeutic, or any combination thereof.
53. The method of any one of claims 1-46, further comprising administering a therapeutic to the subject.
54. The method of claim 53, wherein the therapeutic comprises a regenerative therapy, a medical device, a pharmaceutical composition, a medical procedure, or any combination thereof.
55. The method of claim 53, wherein the therapeutic comprises a non-steroidal anti-inflammatory, a hormone treatment, a dietary supplement, a cannabis-derived therapeutic or any combination thereof.
56. The method of claim 53, wherein the therapeutic comprises the pharmaceutical composition, and wherein the pharmaceutical composition comprises an at least partially hemp-derived therapeutic, an at least partially cannabis-derived therapeutic, a cannabidiol (CBD) oil derived therapeutic, or any combination thereof.
57. The method of claim 53, wherein the therapeutic comprises the medical procedure, and wherein the medical procedure comprises a laparoscopy, a laser ablation procedure, a hysterectomy or any combination thereof.

58. The method of claim 53, wherein the therapeutic comprises the regenerative therapy, and wherein the regenerative therapy comprises a stem cell, a cord blood cell, a Wharton's jelly, an umbilical cord tissue, a tissue, or any combination thereof.
59. The method of claim 53, wherein the therapeutic comprises the pharmaceutical composition, and wherein the pharmaceutical composition comprises cannabis, cannabidiol oil, hemp, or any combination thereof.
60. The method of claim 53, wherein the therapeutic comprises the pharmaceutical composition, and wherein the pharmaceutical composition is formulated in a unit dose.
61. The method of claim 53, wherein the therapeutic comprises hormonal therapy, an advanced reproductive therapy, a pain managing medication, or any combination thereof.
62. The method of claim 53, wherein the therapeutic comprises a hormonal contraceptive, gonadotropin-releasing hormone (GnRH) agonist, gonadotropin-releasing hormone (GnRH) antagonist, progestin, danazol, or any combination thereof.
63. The method of any one of claims 1-62, wherein the subject is asymptomatic for endometriosis.
64. A kit comprising: one or more probes for detecting one or more genetic variants of Table 1, Table 2, Table 3, or any combination thereof in a sample.
65. The kit of claim 64, further comprising a control sample.
66. The kit of claim 64, wherein the control sample comprises one or more genetic variants of Table 1, Table 2, Table 3, or any combination thereof.
67. The kit of any one of claims 64-66, wherein the one or more probes comprise a hybridization probe or amplification primer.
68. The kit of any one of claims 64-67, wherein the one or more probes is configured to associate with a solid support.
69. The kit of any one of claims 64-68, wherein the kit further comprises instructions for use and wherein the instructions for use comprise high stringent hybridization conditions.
70. The kit of any one of claims 64-69, wherein the one or more probes is configured to hybridize to a target region of a nucleic acid of the sample, wherein the target region comprises one or more genetic variants.
71. A system comprising: (a) a computer processor configured to receive sequencing data obtained from assaying a sample, wherein the computer processor is configured to identify a presence or an absence of one or more genetic variants of Table 1, Table 2, Table 3 or any combination thereof in the sample, and (b) a graphical user interface configured to display a

report comprising the identification of the presence or the absence of the one or more genetic variants in the sample.

72. The system of claim 71, wherein the computer processor comprises a trained algorithm.
73. The system of claim 71 or 72, wherein the computer processor communicates a result.
74. The system of claim 73, wherein the result comprises an identification of the presence or the absence of one or more genetic variants in the sample.
75. A method comprising: (a) sequencing or genotyping a nucleic acid sample obtained from a subject having endometriosis, suspected of having endometriosis, or suspected of having a risk of developing endometriosis using a high throughput method; and (b) detecting a genetic variant in said nucleic acid sample, wherein said genetic variant comprises a mutation in *SEPT10*, *TNFRSF6B*, *UGT2B28*, *USP17L2* or any combination thereof.
76. The method of claim 75, wherein said genetic variant is a mutation in *SEPT10* and wherein said mutation comprises a missense mutation.
77. The method of claim 75, wherein said genetic variant is a mutation in *TNFRSF6B* and wherein said mutation comprises a homozygous or hemizygous mutation.
78. The method of claim 75, wherein said genetic variant is a mutation in *UGT2B28* or *USP17L2* and wherein said mutation comprises a hemizygous deletion.
79. The method of claim 75, wherein said high throughput method comprises nanopore sequencing.

FIG. 1A **Endometriosis Subjects**

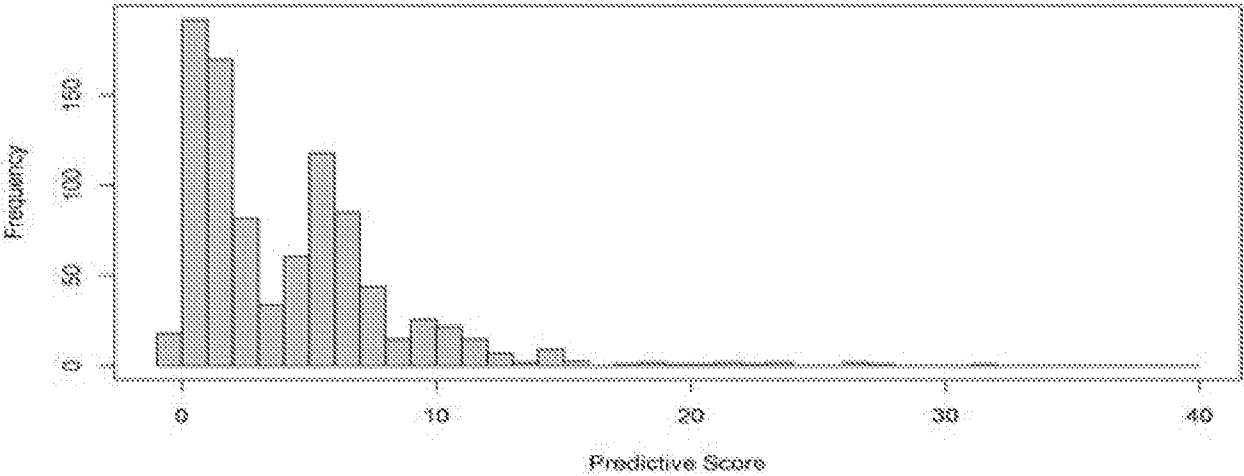


FIG. 1B **Control Subjects**

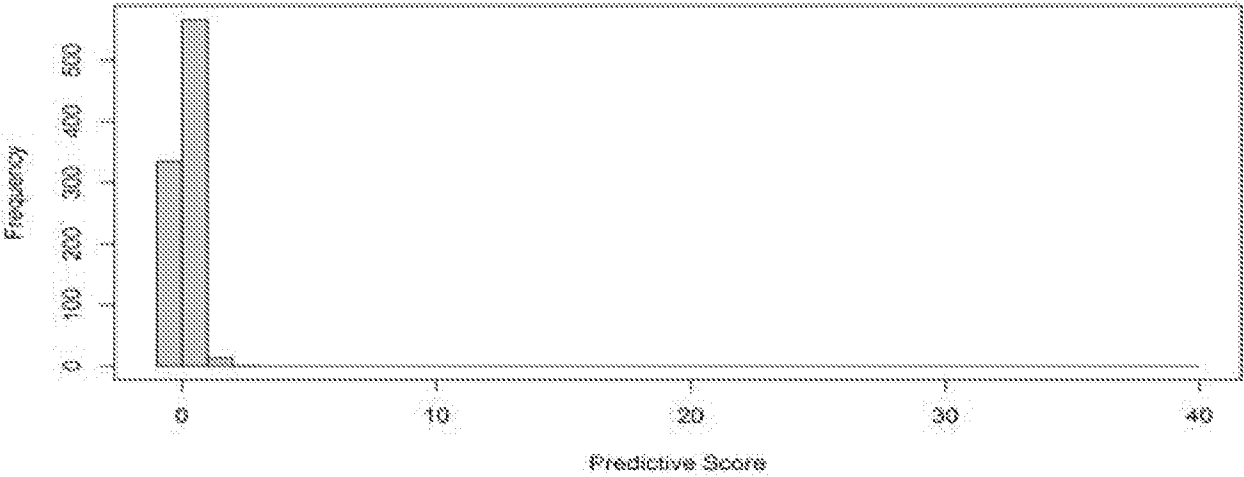
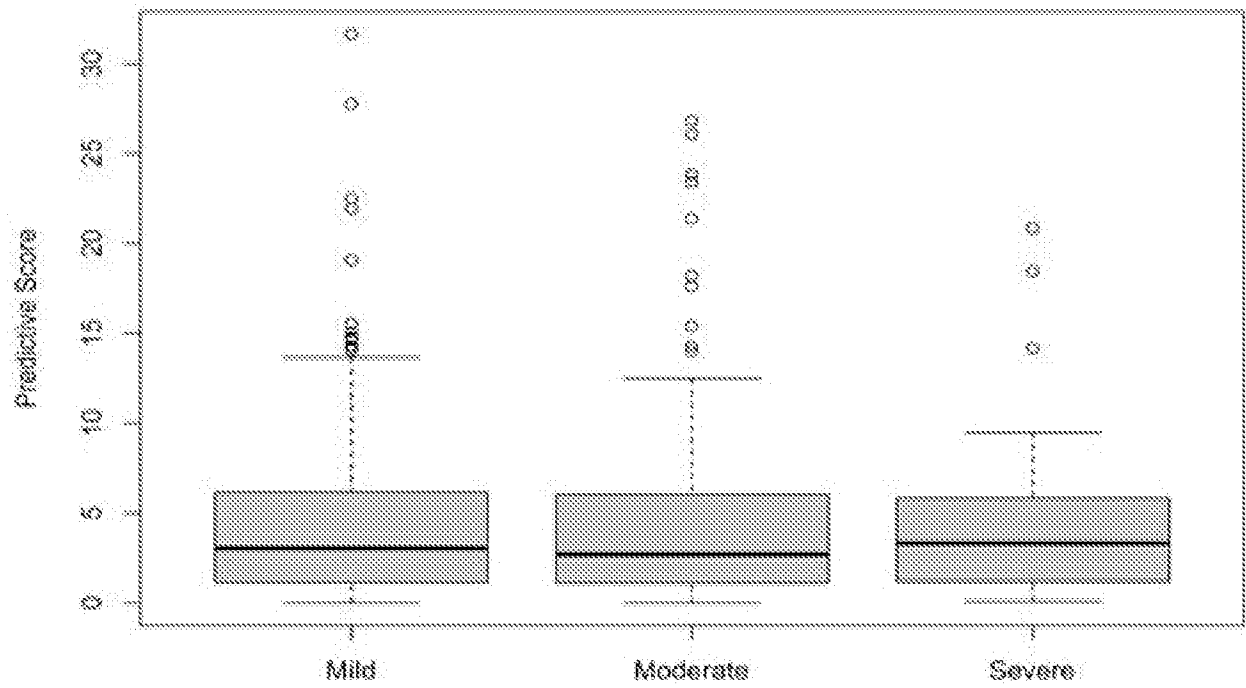
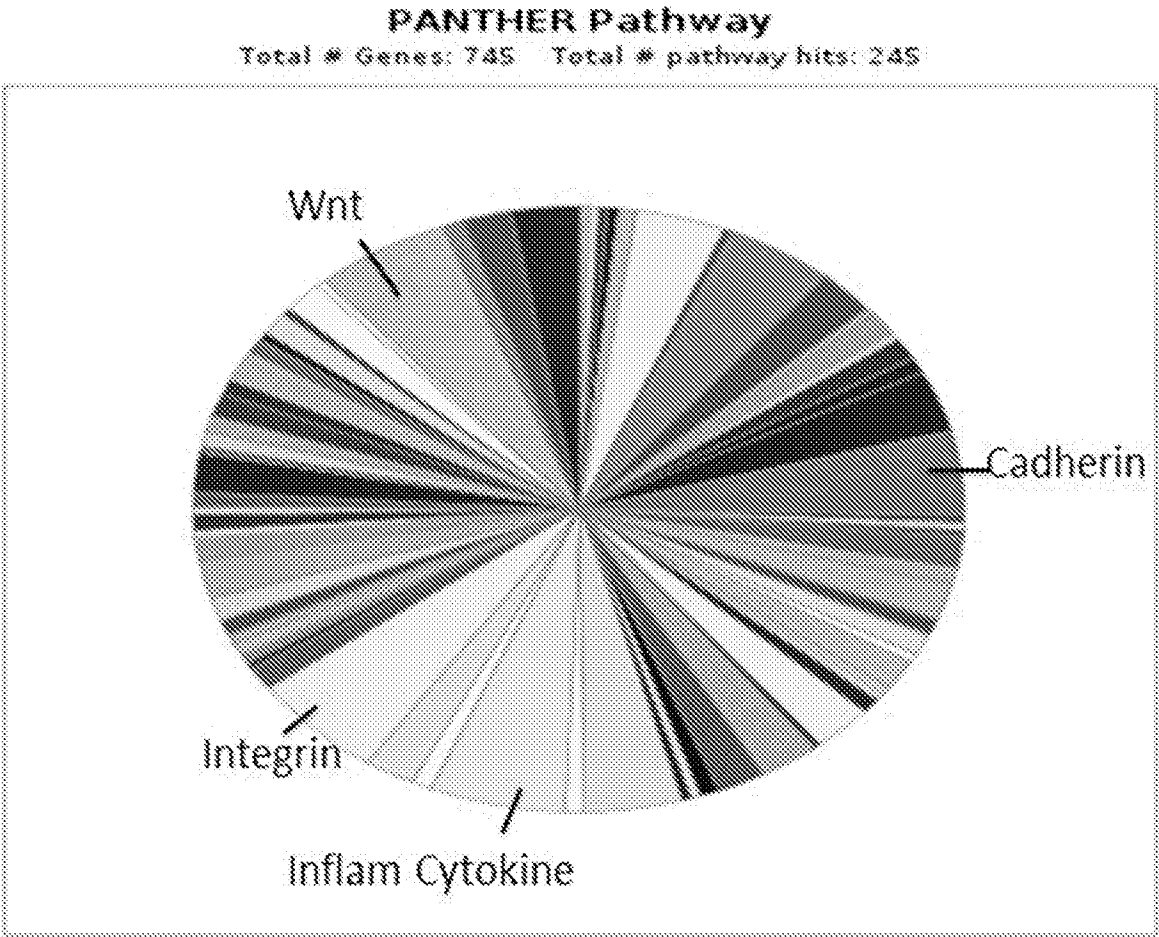


FIG. 2



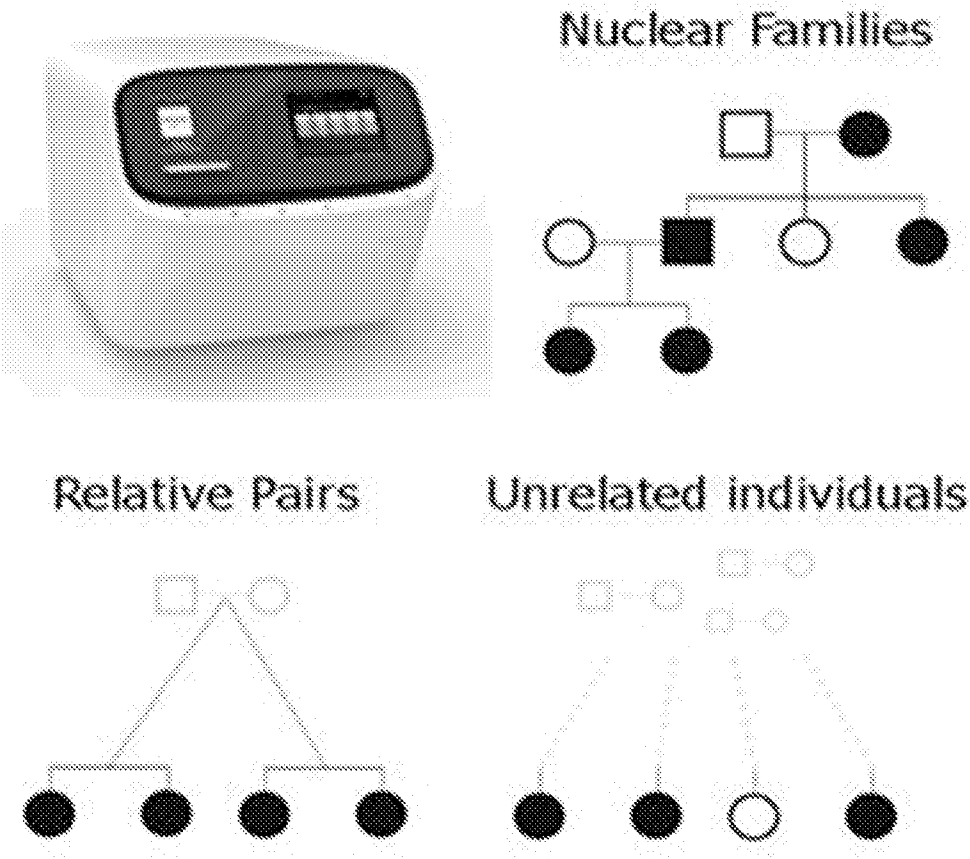
3/20

FIG. 3



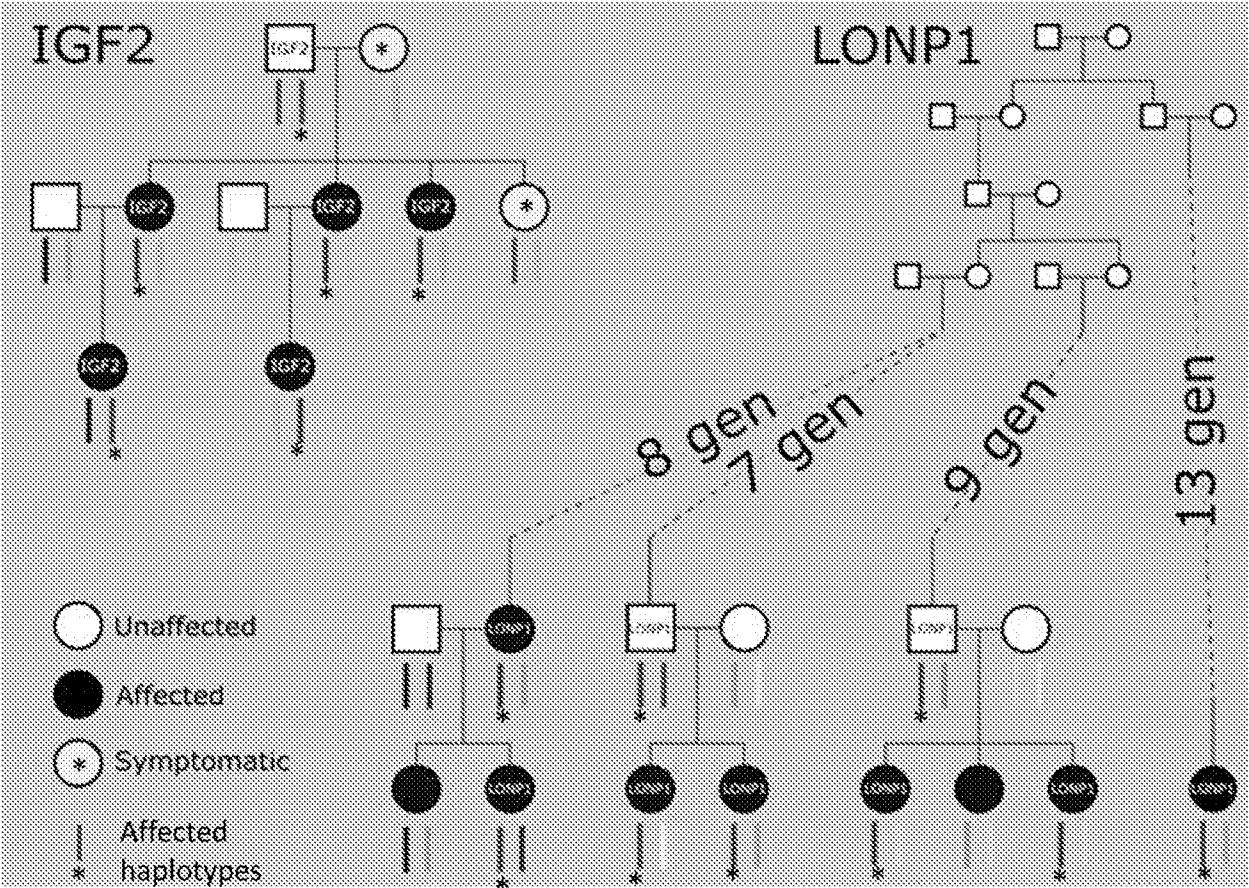
4/20

FIG. 4



5/20

FIG. 5



6/20

FIG. 6

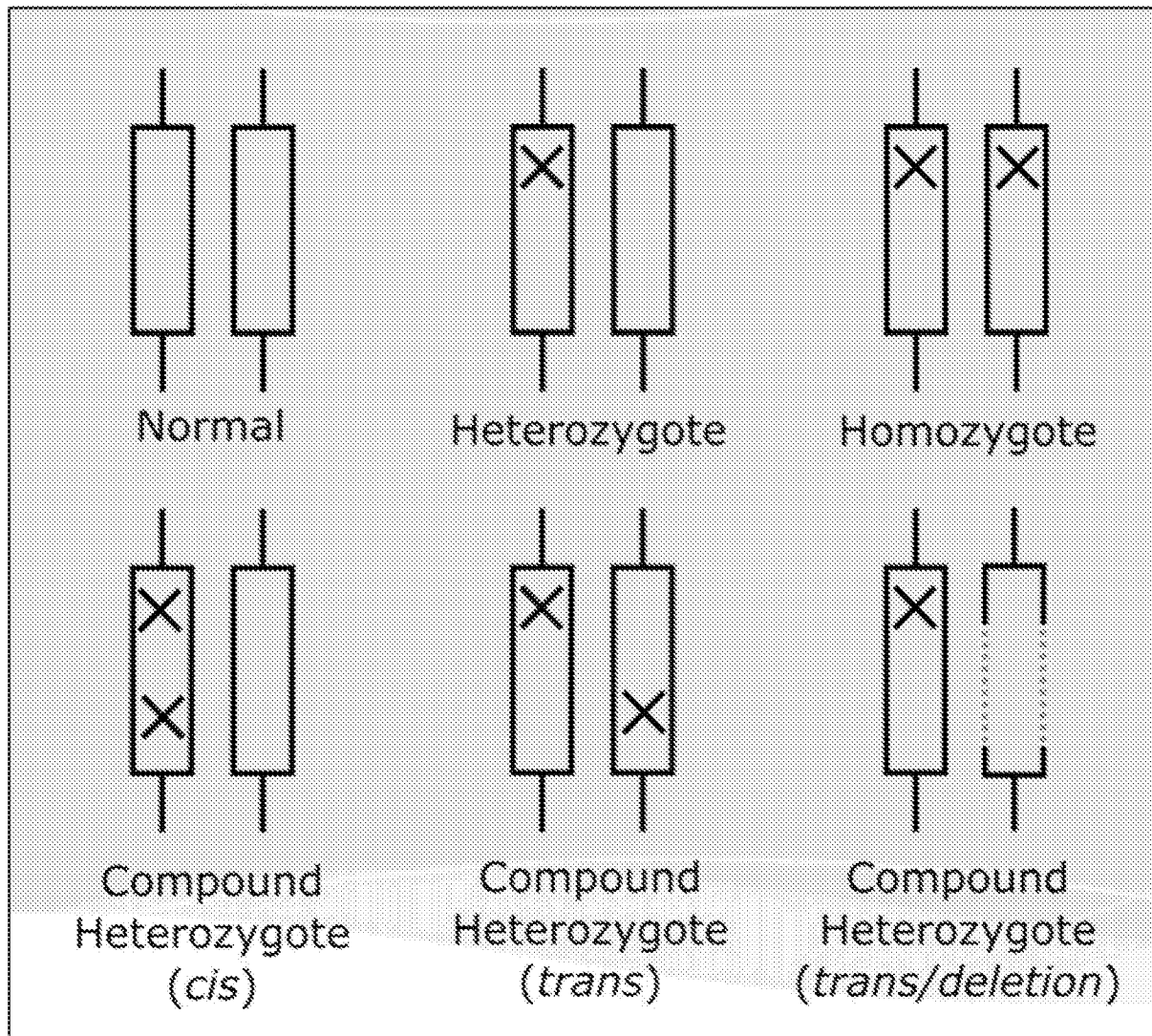


FIG. 7

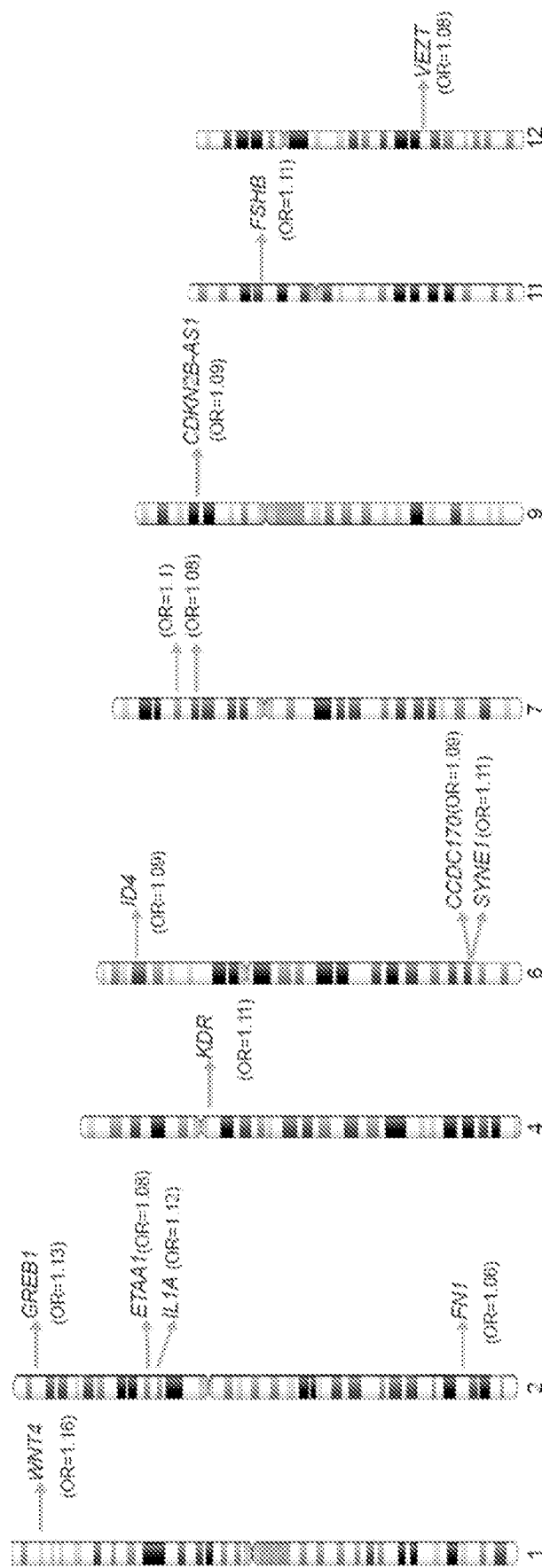


FIG. 8A

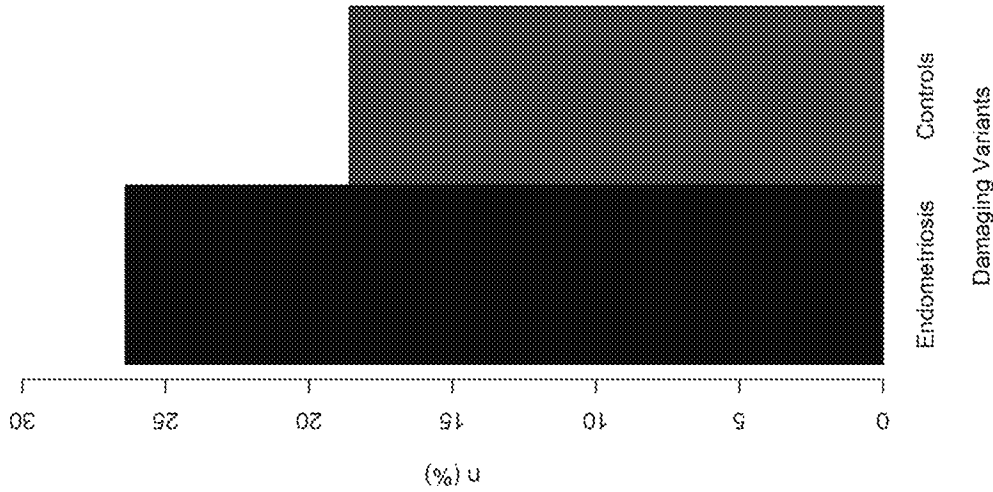


FIG. 8B

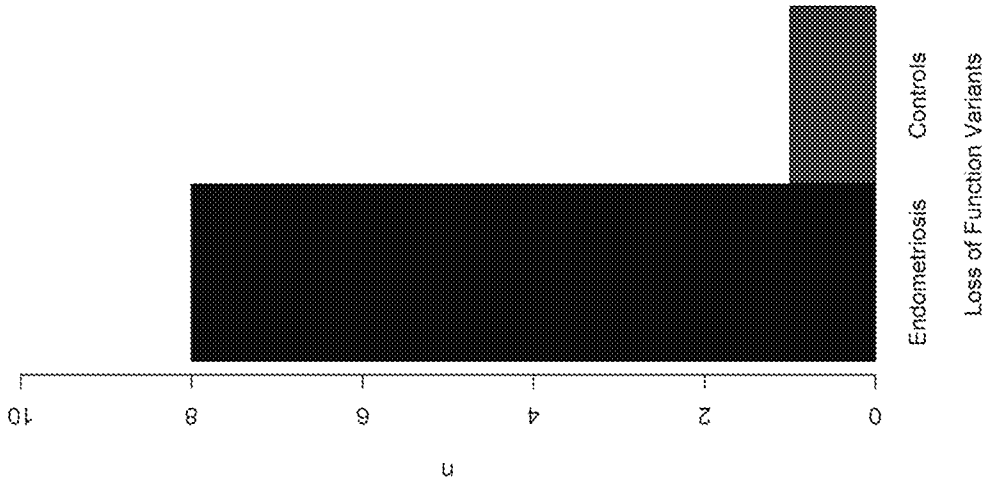
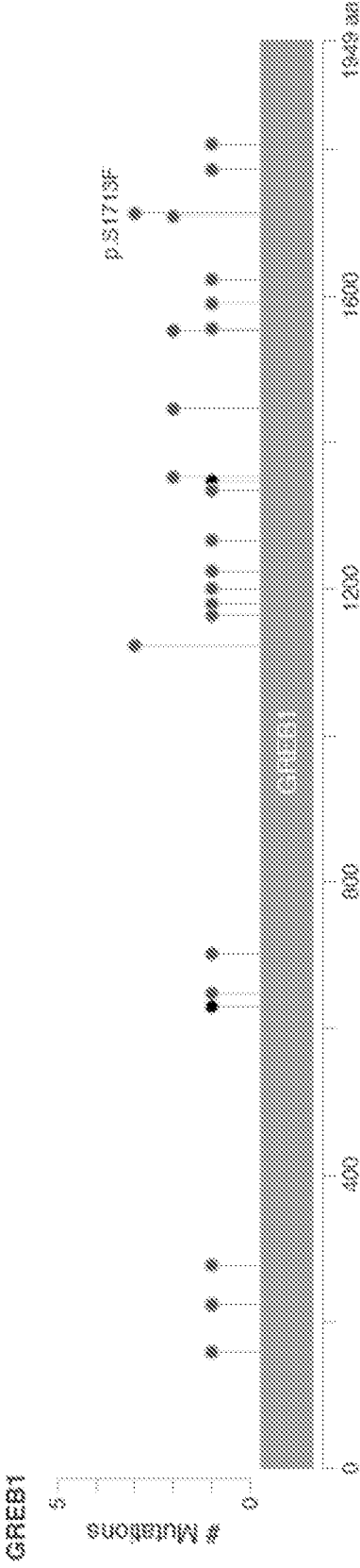
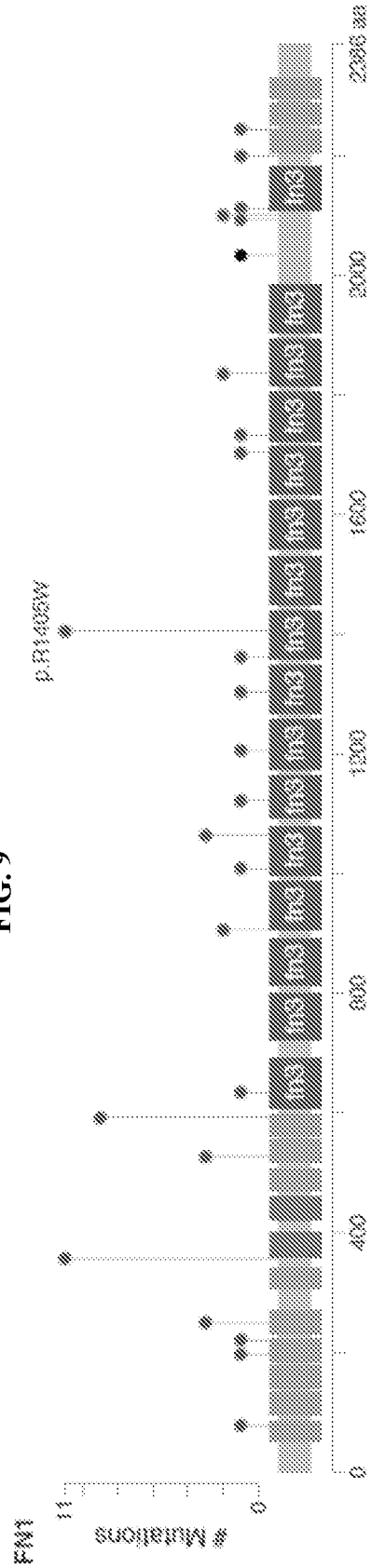


FIG. 8C

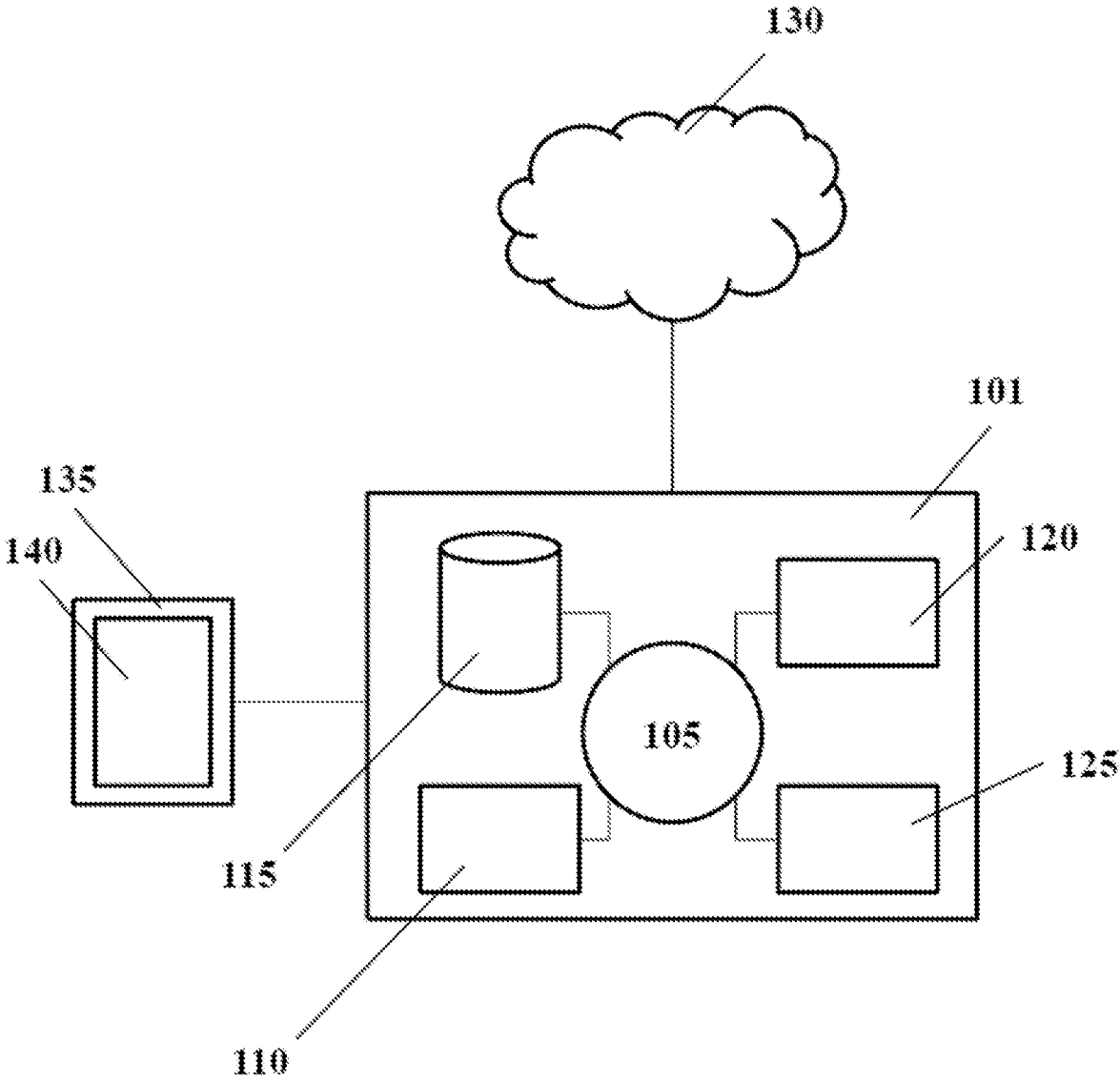
Gene	# damaging low frequency variants
CCDC170	4
ESR1	8
ETAA1	4
FN1	22
FSHB	3
GREB1	21
IL1A	2
KDR	10
NFE2L3	8
SYNE1	106
TNS3	24
VEZT	9
WNT4	1

FIG. 9



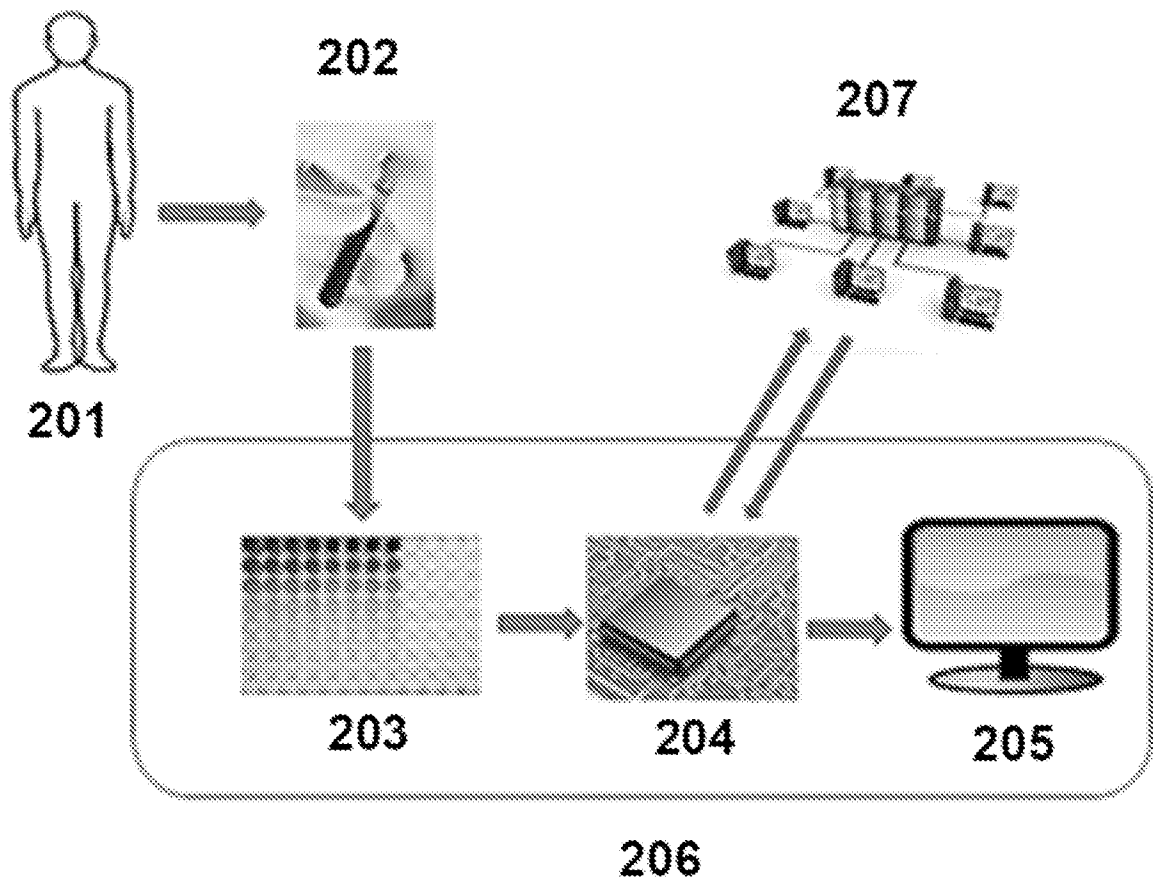
10/20

FIG. 10



11/20

FIG. 11



12/20

FIG. 12

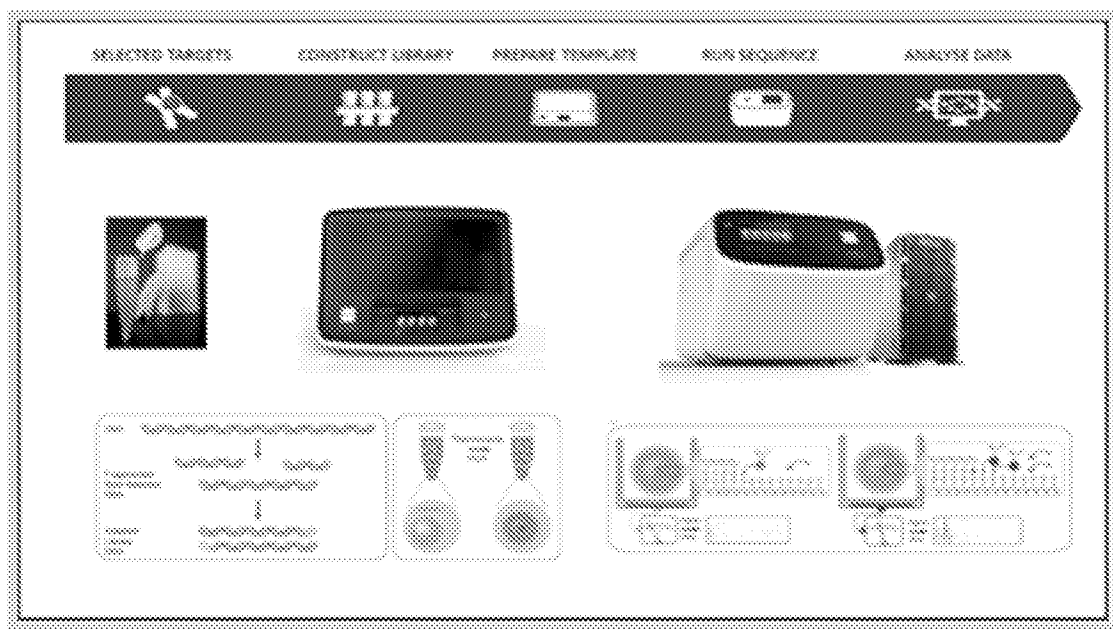
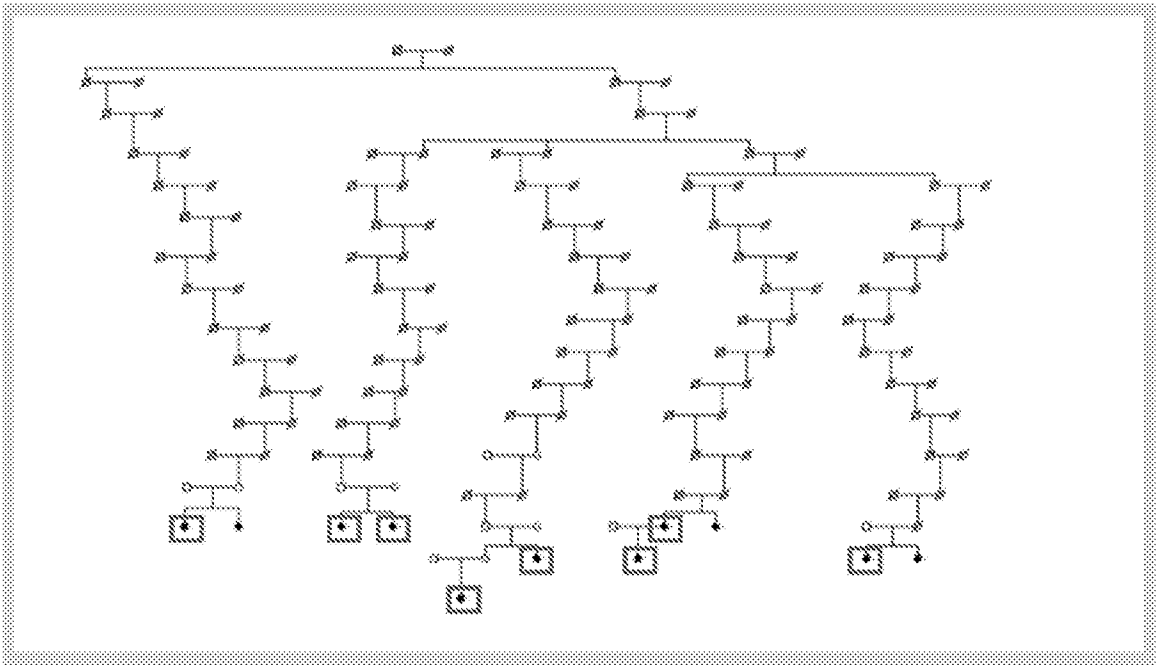
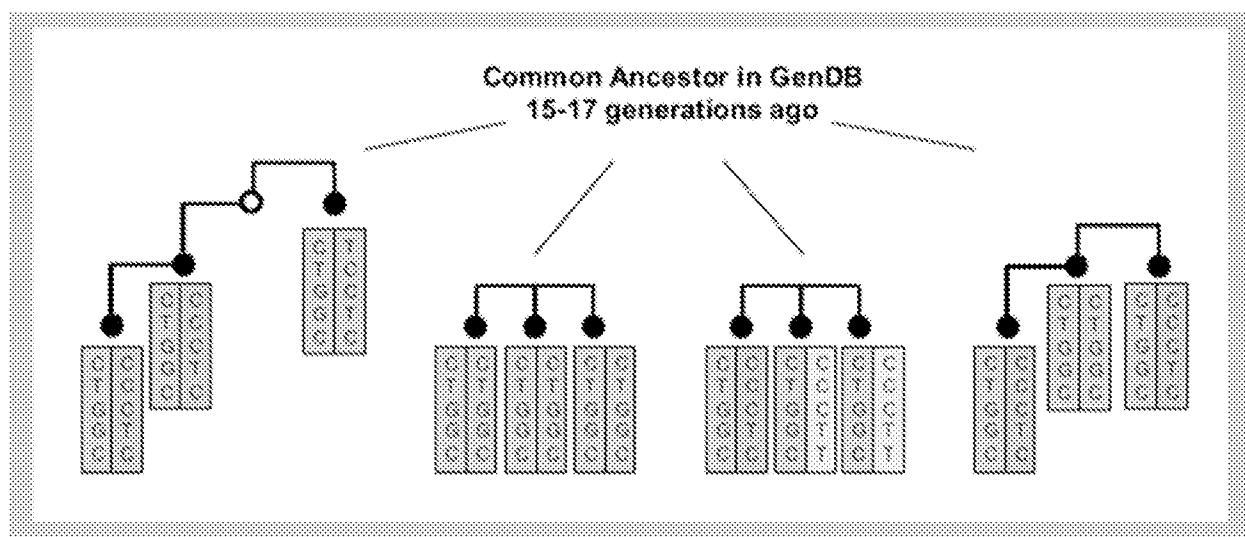


FIG. 13



14/20

FIG. 14



15/20

FIG. 15A

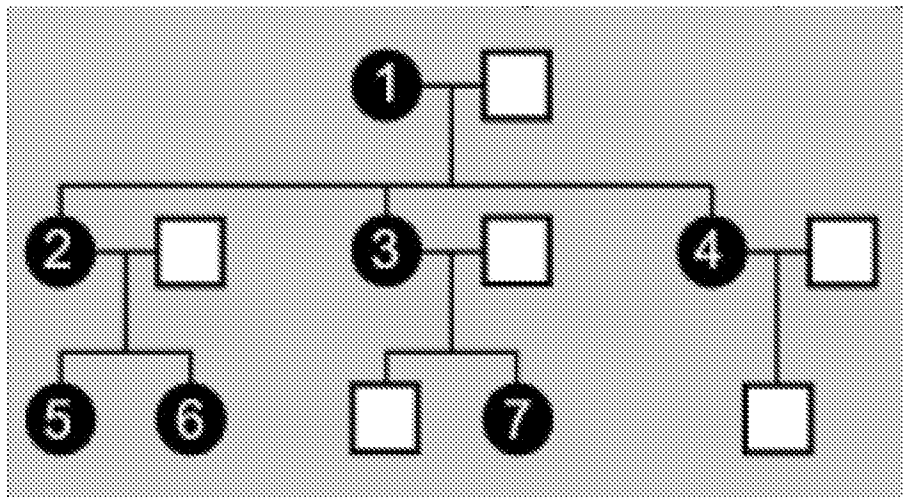


FIG. 15B

Case no.	Age	Symptoms	Stage	Diagnosis
1	65	PP, Dys	IV	TAH age 32
2	49	PP, Dys	IV	TAH age 33
3	46	PP	II	Endometrioma
4	40	Inf, Dys	II	Adenomyosis
5	32	Inf	III	Endometrioma
6	27	PP, Inf	II	Endometrioma
7	25	Inf	II	Endometrioma
Pevic Pain (PP), Dysmenohrea (Dys), Infertility (Inf), Total Abdominal Hysterectomy (TAH)				

16/20

FIG. 16

Chr	Pos	Ref	Alt	Gene	Nucl	SNP	Individual ID						
							1	2	3	4	5	6	7
4	69,972,949	C	G	UGT2B7	c.C1059G	rs4292394	0/1	1/1	0/0	1/1	1/1	1/1	1/1
4	70,079,838	T	C	UGT2B11	c.A603G	rs4694697	0/1	0/0	0/1	0/0	0/0	0/0	0/0
4	70,079,963	G	A	UGT2B11	c.C478T	rs72551399	0/0	0/0	0/0	0/0	0/1	0/1	0/0
4	70,146,230	G	A	UGT2B28	c.G12A	rs13139691	1/1	0/0	1/1	0/0	0/0	0/0	0/1
4	70,146,704	G	A	UGT2B28	c.G486A	rs7689398	1/1	0/0	0/1	0/0	0/0	0/0	0/0
4	70,146,804	G	C	UGT2B28	c.G586C	rs148987832	0/0	0/0	0/1	0/0	0/0	0/0	0/0
4	70,156,392	A	G	UGT2B28	c.A1173G	rs10013145	1/1	0/0	0/0	0/0	0/0	0/0	0/0
4	70,160,277	T	G	UGT2B28	c.T1340G	rs6843900	1/1	0/0	0/1	0/0	0/0	0/0	0/0
4	70,160,309	C	G	UGT2B28	c.C1372G	rs6828191	1/1	0/0	0/1	0/0	0/0	0/0	0/0
4	70,160,388	G	C	UGT2B28	c.G1401C	rs72552703	0/0	0/0	0/1	0/0	0/0	0/0	0/1
4	70,160,342	TG	CC	UGT2B28	c.1405_1406	rs796618077	0/0	0/0	0/1	0/0	0/0	0/0	0/0
4	70,346,546	GA	TT	UGT2B4	c.1374_1375	rs67904882	0/0	0/0	0/1	0/0	0/0	0/0	0/1
4	70,355,211	T	C	UGT2B4	c.A948G	rs1845555	1/1	1/1	1/1	1/1	1/1	1/1	0/1
8	11,706,581	T	G	CTSB	c.A42DC	rs13332	0/1	0/1	0/1	0/1	0/1	0/1	1/1
8	11,710,888	G	C	CTSB	c.C76G	rs12338	0/1	0/1	1/1	1/1	0/1	0/1	0/1
8	11,832,079	A	G	DEFB136	c.T3DC	rs10108075	0/0	0/1	0/1	0/1	0/0	0/0	0/1
8	11,994,716	C	A	USP17L2	c.G1554T	rs199935289	1/1	0/0	0/0	0/0	0/0	0/0	0/0
8	11,994,957	T	C	USP17L2	c.A1313G	rs12543578	1/1	0/0	0/0	0/0	1/1	1/1	0/0
8	11,995,062	G	A	USP17L2	c.C1208T	rs75807755	0/0	0/0	1/1	1/1	0/0	0/0	0/0
8	12,600,720	T	A	LONRF1	c.A793T	rs1139354	0/0	0/0	0/1	0/1	0/0	0/0	0/1
8	12,878,637	A	G	KIAA1456	c.A448G	rs528255	1/1	1/1	1/1	1/1	1/1	1/1	1/1
8	12,878,677	T	C	KIAA1456	c.T489C	rs522106	0/1	0/1	0/1	0/1	0/0	0/0	1/1

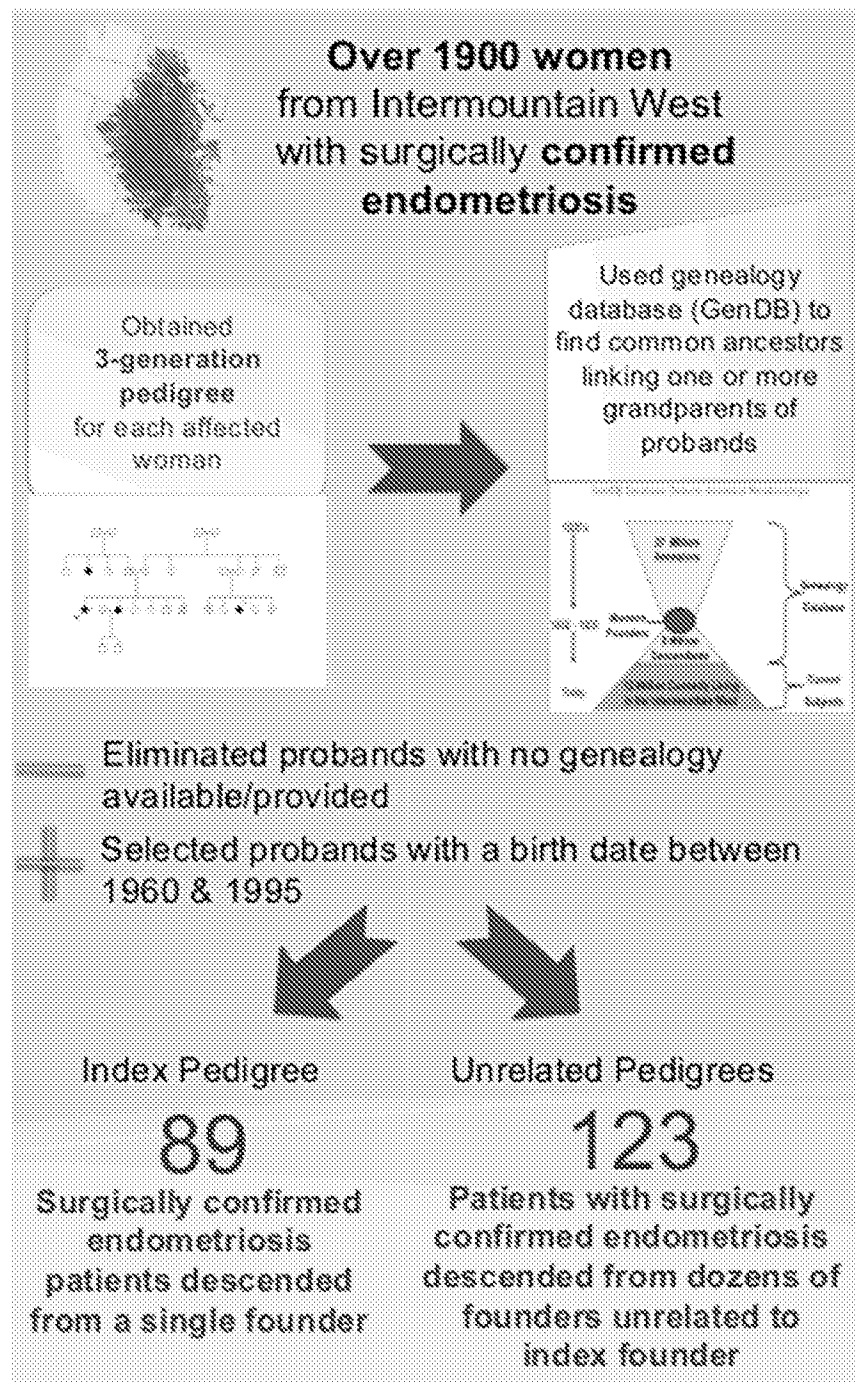
17/20

FIG. 17

	Index Endometriosis Pedigree: descendants of common ancestor born in 1608		Present Study: Matched Recent Probands	
	Number	%	Number	%
Confirmed affected and born between 1960-1995	89		123	
Sisters affected / total	1/3	33%	62/367	17%
Mothers affected / total	9/89	10%	29/123	24%
First cousins affected / total	21/82	26%	3/13	23%
Adult daughters affected / total	1/4	25%	1/8	12%

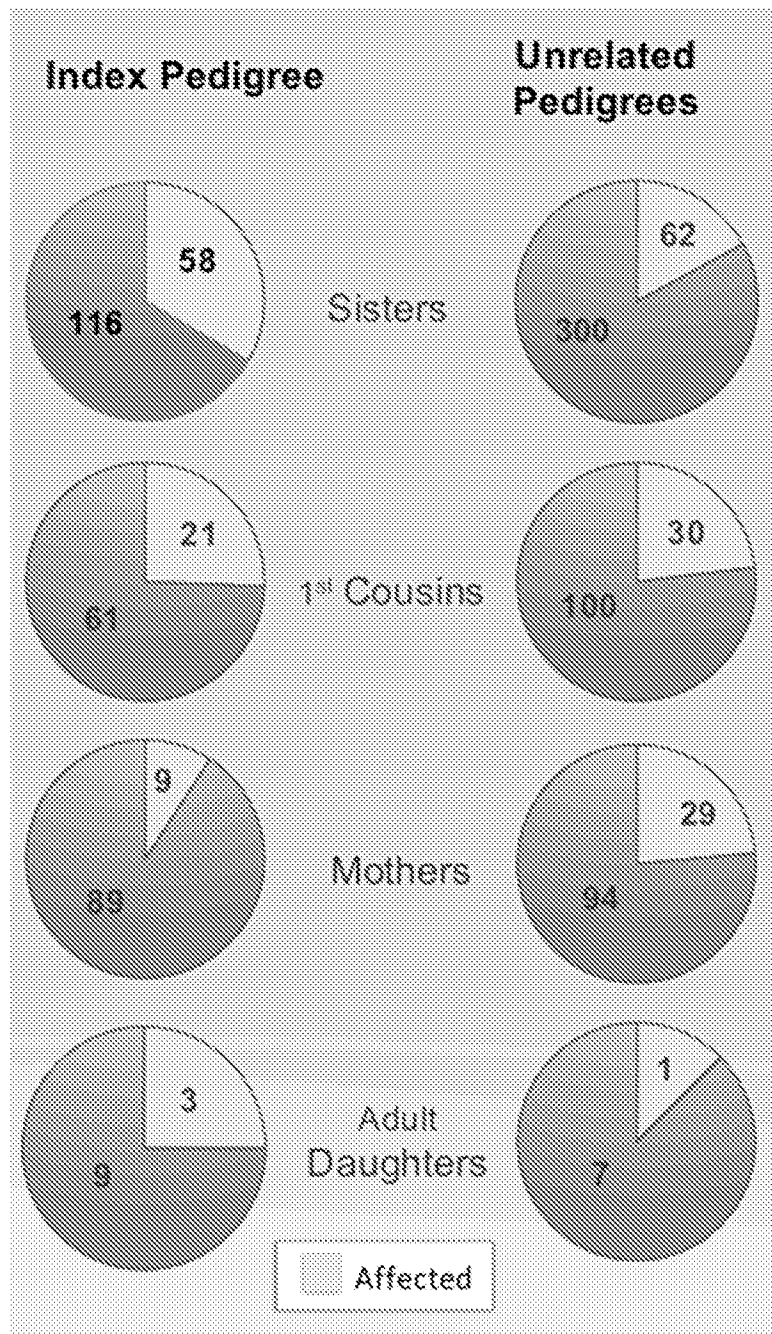
18/20

FIG. 18



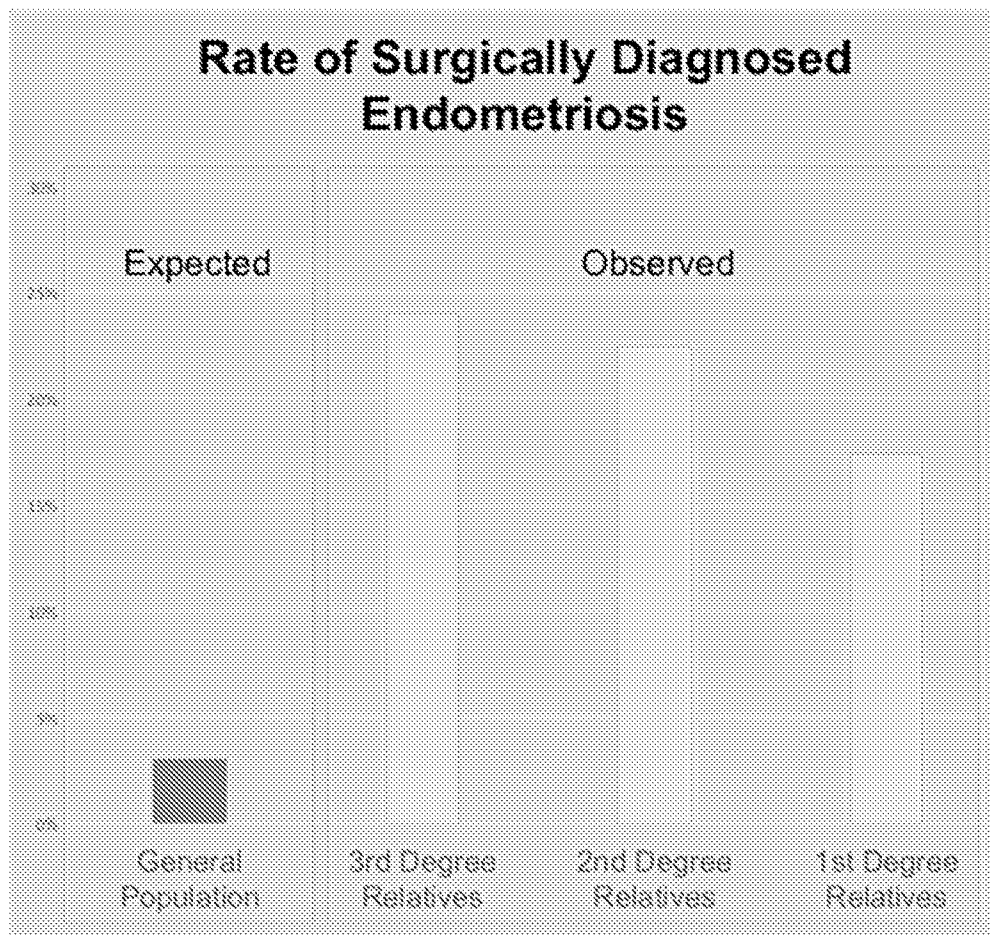
19/20

FIG. 19



20/20

FIG. 20



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 19/50061

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C07H 21/04, C12Q 1/68, C12Q 1/6876 (2020.01)

CPC - C07H 21/04, C12Q 1/6883, C12Q 2600/156

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GUO et al. Genomic alterations in the endometrium may be a proximate cause for endometriosis. European Journal of Obstetrics & Gynecology and Reproductive Biology, 10 September 2004, Vol 116, No 1, pg 89-99. Entire document, especially abstract; p. 90, col 1, para 2-3; p. 91, col 1, para 2-3; p. 96, col 2, para 1-2	1-7
Y	US 2006/0134663 A1 (HARKIN et al.) 22 June 2006 (22.06.2006) para [0066]-[0067]; [0153]; SEQ ID NO: 95904	1-7
Y	US 2015/0275298 A1 (STYLLI) 01 October 2015 (01.10.2015) para [0007]-[0009]; [0097]	2, (3-7)/2

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"D" document cited by the applicant in the international application

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

17 January 2020

Date of mailing of the international search report

05 FEB 2020

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Lee Young

Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 19/50061

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 8-63, 68-70
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

--continued on first extra sheet--

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-7 limited to R175C in TNFRSF18 (SEQ ID NO: 1)

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/US 19/50061

--continued from Box III: Observations where unity of invention is lacking--

Group I+, claims 1-7, 75-79, directed to a method comprising sequencing or genotyping a subject nucleic acid sample for a genetic variant related to endometriosis. The method will be searched to the extent that the genetic variant encompasses R175C in TNFRSF18 (SEQ ID NO: 1, nt 1131296 G>A), (note this is the first variant listed in Applicants' Table 1, specification p. 79, para [0447]). It is believed that claims 1-7 encompass this first named invention, and thus these claims will be searched without fee to the extent that the genetic variant encompasses R175C in TNFRSF18 (SEQ ID NO: 1, nt 1131296 G>A). Additional genetic variant(s) will be searched upon the payment of additional fees. Applicants must specify the claims that encompass any additionally elected genetic variant(s). Applicants must further indicate, if applicable, the claims which encompass the first named invention, if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched. An exemplary election would be the genetic variant comprising W110R in ATAD3A (SEQ ID NO: 2, nt 1452592 T>C), (claims 1-7).

[Note, Tables 1-3 do not comprise mutations in any of SEPT10, TNFRSF6B, UGT2B28, or USP17L2]

Group II, claims 64-67, directed to a kit comprising probes.

Group III, claims 71-74, directed to a system comprising a computer processor and a graphical user interface.

The inventions listed as Groups I+, II and III do not relate to a single special technical feature under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Special technical features

Group I+ has the special technical feature of sequencing or genotyping a nucleic acid sample obtained from a subject having endometriosis, that is not required by Group II or III.

Group II has the special technical feature of a kit comprising one or more probes, that is not required by Group I+ or III.

Group III has the special technical feature of a system comprising a computer processor and a graphical user interface, that is not required by Group I+ or II.

The inventions of Group I+ each include the special technical feature of a different genetic variant and is considered a distinct technical feature.

Common technical features

The inventions of Groups I+, II and III share the common technical feature of detecting one or more genetic variants in a nucleic acid sample from a subject. However, this shared technical feature is previously taught by US 2009/0307181 A1 to Colby et al., (hereinafter Colby).

Colby teaches a method comprising detecting one or more genetic variants in a nucleic acid sample from a subject (claim 1 "method of determining the predisposition or carrier status of an individual for two or more phenotypes related to pediatrics or reproduction comprising: (a) identifying by nucleic acid array or sequencing apparatus a set of genetic variants in an individual"; para [0037] "steps for obtaining a genetic analysis of a patient sample").

As the technical features were known in the art at the time of the invention, they cannot be considered special technical features that would otherwise unify the groups.

Therefore, Group I+, II and III inventions lack unity under PCT Rule 13 because they do not share the same or corresponding special technical feature.

NOTE, continuation of number 4 above: claims 8-63, 68-70 are held unsearchable because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).