Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 3212318 A1 2022/10/11

(21) 3 212 318

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) **Date de dépôt PCT/PCT Filing Date:** 2022/05/06
(87) **Date publication PCT/PCT Publication Date:** 2022/10/11
(85) **Entrée phase nationale/National Entry:** 2023/09/15
(86) **N° demande PCT/PCT Application No.:** CA 2022/050716
(87) **N° publication PCT/PCT Publication No.:** 2022/232946
(30) **Priorités/Priorities:** 2021/05/07 (US63/185,752); 2022/01/28 (US63/304,169)

(51) **Cl.Int./Int.Cl.** ***C12Q 1/6895*** (2018.01), ***A01H 6/28*** (2018.01)

(71) **Demandeur/Applicant:**
CANOPY GROWTH CORPORATION, CA

(72) **Inventeur/Inventor:**
DEOKAR, AMIT, CA

(74) **Agent:** BCF LLP

(54) **Titre : PLANTES DE CANNABIS PRESENTANT UNE RESISTANCE A L'OIDIUM ET LEURS PROCEDES D'OBTENTION**
(54) **Title: CANNABIS PLANTS EXHIBITING POWDERY MILDEW RESISTANCE AND METHODS FOR OBTAINING SAME**

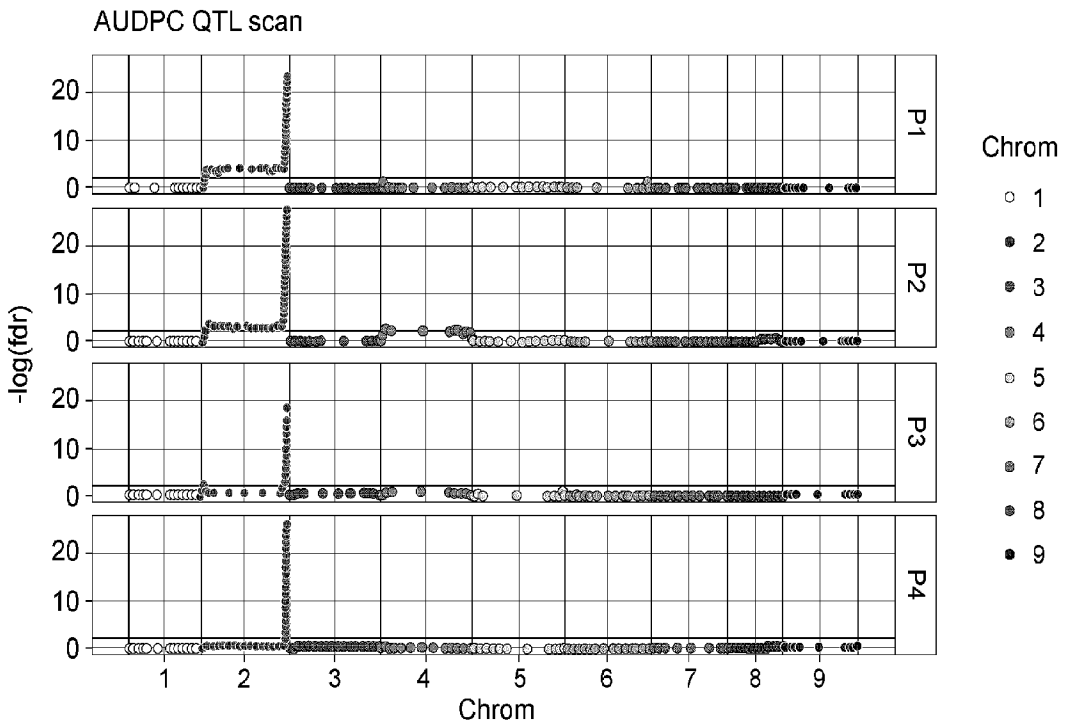

FIG. 1

(57) **Abrégé/Abstract:**
The present technology is in the field of plant breeding and genetics, particularly as it pertains to the genus Cannabis. More specifically, there are provided methods and compositions for producing a population of Cannabis plants with enhanced resistance to powdery mildew. The methods use the detection of molecular genetic markers linked to powdery mildew resistance loci to select for plants displaying an enhanced powdery mildew resistance phenotype.

Date Submitted: 2023/09/15

CA App. No.: 3212318

Abstract:

The present technology is in the field of plant breeding and genetics, particularly as it pertains to the genus Cannabis. More specifically, there are provided methods and compositions for producing a population of Cannabis plants with enhanced resistance to powdery mildew. The methods use the detection of molecular genetic markers linked to powdery mildew resistance loci to select for plants displaying an enhanced powdery mildew resistance phenotype.

CANNABIS PLANTS EXHIBITING POWDERY MILDEW RESISTANCE AND METHODS FOR OBTAINING SAME

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of and priority to U.S. provisional patent application No. 63/185,752, filed on May 7, 2021; and to U.S. provisional patent application No. 63/304,169, filed on January 28, 2022; the content of both of which is herein incorporated in entirety by reference.

FIELD OF TECHNOLOGY

[0002] The present technology generally relates to Cannabis plants exhibiting powdery mildew resistance as well as molecular markers for same and methods of producing same.

BACKGROUND

[0003] *Cannabis sativa* L. is a diploid (2n=20) annual species belonging to the genus Cannabis in the Cannabaceae family. *Cannabis sativa* is commonly known as cannabis, hemp, Indian hemp, marihuana, and marijuana. Cannabis has been cultivated throughout the centuries as a source of fiber and food, and for its therapeutic and recreational properties (Farnsworth, N.R., 1969). The pharmaceutical and recreational properties of this unique plant are associated with a unique class of terpenophenolic compounds known as cannabinoids. Cannabinoids interact with receptors of human and animal endocannabinoid systems and can lead to a plethora of potential medical and therapeutic effects (Di Marzo & Piscitelli, 2015). Cannabinoids are produced and stored in the glandular trichomes present on the surfaces of female inflorescences. Over 500 phytocannabinoids and non-cannabinoid constituents have been identified and/or isolated from *C. sativa* L, including the well-known psychoactive compound Δ9-tetrahydro- cannabinol (Δ9-THC) and non-psychoactive compounds such as cannabidiol (CBD) (an isomer of THC), cannabichromene (CBC) and cannabigerol (CBG) (ElSohly et al., 2017).

[0004] As the legal cannabis market size continues to grow, a consistent supply of high-quality cannabis products remains a priority. Of the many factors affect the yield, quality, and marketability of cannabis products, the levels of infection by destructive pathogens such as powdery mildew and contamination by agrochemicals used to control diseases are among the most important. General susceptibility of the commercial material, the pathogenicity of the fungus, large seasonal spore loads, and the lack of effective and approved disease management strategies make powdery mildew a major concern for commercial cannabis cultivation.

[0005] Cannabis powdery mildew is caused by biotrophic parasitic fungi such as the *Golovinomyces* species (Thompson et al., 2017; Pépin et al., 2018). Powdery mildew develops on leaves, stems, and flower buds at any stage of development. Early stages of infection can be observed on young leaves as white mycelium on the leaf surface, whereas in advanced stages of infection, profuse sporulation results in a powdery appearance on the leaf surface (Punja et al., 2019). Powdery mildew infection can severely damage a plant by limiting photosynthetic activities and reducing nutrient availability to the plant, causing premature leaf fall, and reducing overall vigor and potential yield (Scott and Punja, 2020).

[0006] Powdery mildew disease is largely managed by applications of chemical products or bio-control agents. However, development and growth of resistant cultivars will be a more sustainable, effective, and economical strategy against this disease. Similar approaches have been used successfully for a wide range of crops such as wheat, barley, and tomato (Ren et al., 2017; Bengtsson et al., 2017; Nekrasov et al., 2017). One of the major hurdles for using this approach in cannabis is a lack of natural resistance against the powdery mildew in cannabis, as well as a lack of knowledge of the genetic mechanisms for regulating powdery mildew disease resistance in cannabis.

[0007] Both qualitative and quantitative resistance against the powdery mildew disease has been reported in many plant species. In some plants, such as tomato nine monogenic resistance genes (*Ol-genes*) have been identified confirming resistance to Tomato powdery mildew (Seifi et al 2014), while in other species, such in wheat over 100 powdery mildew resistance quantitative trait loci (QTL) have been identified (Ren et al 2017). Resistance to powdery mildew in Hop (*Humulus lupulus* L. var lupulus) another member of the *Cannabaceae* plant family, has been reported as controlled by both qualitative and quantitative genetic control (Henning et al. 2017). Recently, Padgitt-Cobb and others (2020) identified a narrow genomic region mapped to hop linkage group 10. This region contained several putative R-genes and putative peroxidase-3 genes as a potential candidate for genes involved in resistance for powdery mildew in the hop.

[0008] Attempts to breed powdery mildew resistant cannabis cultivars have been greatly impeded by the lack of a natural source of resistance and a poor understanding of the inheritance of disease resistance, as well as the unavailability of robust genetic markers linked to the trait. The use of markers in cannabis breeding could not only reduce the cost of developing new varieties but may also increase the precision and efficiency of selection of powdery mildew resistant breeding lines in the breeding program, as well as reduce the number of years required to develop new and improved resistant varieties.

[0009] There is a need to identify genetic markers in Cannabis for powdery mildew resistance and to develop Cannabis strains exhibiting such resistance.

SUMMARY

[0010] According to various aspects, the present technology relates to a Cannabis plant, plant part, tissue or cell thereof, wherein the Cannabis plant, plant part, tissue or cell thereof has an enhanced resistance to powdery mildew, as well as molecular markers for same and methods of producing same.

[0011] According to various aspects, the present disclosure provides a method for creating a population of Cannabis plants with enhanced powdery mildew resistance, the method comprising: providing a first population of Cannabis plants; detecting the presence of a genetic marker that is genetically linked to a powdery mildew resistance locus in a defined linkage group, as described herein, by about 20 cM or less in the first population; selecting one or more Cannabis plants containing said marker from the first population of Cannabis plants; and producing a population of offspring from at least one of said selected Cannabis plants.

[0012] The present disclosure further provides a method wherein the genetic marker detected is genetically linked to the powdery mildew resistance locus on the defined linkage group by less than about 15 cM, or less than about 10 cM, or less than about 5 cM.

[0013] In some embodiments of the present technology, the powdery mildew resistance locus is located within a chromosome interval on Chromosome 2.

[0014] In some embodiments of the present technology, the powdery mildew resistance locus is located within a chromosome interval on Chromosome 2 at 83660977-84353662 bp.

[0015] In some embodiments of the present technology, the genetic marker for powdery mildew resistance is located within a chromosome interval on Chromosome 2.

[0016] In some embodiments of the present technology, the genetic marker for powdery mildew resistance is located within a chromosome interval on Chromosome 2 at 83660977-84353662 bp.

[0017] In some embodiments of the present technology, the genetic marker is selected from the markers set forth in FIGs. 4A-4E.

[0018] In some embodiments of the present technology, the genetic marker has a sequence selected from the sequences set forth in FIGs. 5A-5M. In some embodiments, the genetic marker has a sequence

selected from sequences having at least, greater than or about 75%, about 80%, about 85%, about 90%, about 95%, about 96%, about 97%, about 98% or about 99% sequence identity to sequences set forth in FIGs. 5A-5M.

[0019] In another embodiment, the present disclosure provides detecting a genetic marker located within a chromosome interval comprising and flanked by a first genetic marker and a second genetic marker, as described herein. In another embodiment, the present disclosure provides detecting a genetic marker located within a chromosome interval comprising and/or flanked by one or more genetic marker as described herein, e.g., by a DNA sequence comprising or consisting of the genetic marker. In some embodiments of the present technology, the powdery mildew resistance chromosome interval is located on Chromosome 2.

[0020] In some embodiments of the present technology, the powdery mildew resistance chromosome interval is located on Chromosome 2 at 83660977-84353662 bp.

[0021] In some embodiments of the present technology, the genetic marker is selected from the markers set forth in FIGs. 4A-4E.

[0022] In some embodiments of the present technology, the genetic marker has a sequence selected from the sequences set forth in FIGs. 5A-5M. In some embodiments, the genetic marker has a sequence selected from sequences having at least, greater than or about 75%, about 80%, about 85%, about 90%, about 95%, about 96%, about 97%, about 98% or about 99% sequence identity to sequences set forth in FIGs. 5A-5M.

[0023] The present disclosure also provides a method of creating a population of Cannabis plants comprising at least one allele associated with enhanced powdery mildew resistance comprising at least one sequence selected from the group of genetic markers described herein, the method comprising the steps of: genotyping a first population of Cannabis plants, said population containing at least one allele associated with enhanced powdery mildew resistance, the at least one allele associated with enhanced powdery mildew resistance comprising at least one sequence selected from the group consisting of sequences of genetic markers described herein; selecting from said first population one or more identified Cannabis plants containing said at least one allele associated with enhanced powdery mildew resistance comprising at least one sequence selected from the group consisting of sequences of genetic markers described herein; and producing from said selected Cannabis plants a second population, thereby creating a population of Cannabis plants comprising at least one allele associated with enhanced powdery mildew resistance

comprising at least one sequence selected from the group consisting of sequences of genetic markers described herein.

[0024] The present disclosure further provides a method for creating a population of Cannabis plants with enhanced powdery mildew resistance comprising: providing a first population of Cannabis plants; concurrently detecting the presence of at least one genetic marker that is genetically linked to one or more of the genetic markers for powdery mildew resistance described herein by about 20 cM or less in the first population; selecting one or more Cannabis plants containing said at least one markers from the first population of Cannabis plants; and producing a population of offspring from at least one of said selected Cannabis plants. In one embodiment, the at least one genetic marker detected is genetically linked to at least one of the genetic markers for powdery mildew resistance described herein by less than about 15 cM. In another embodiment, the at least one genetic marker detected is genetically linked to at least one of the genetic markers for powdery mildew resistance described herein by less than about 10 cM. In another embodiment, the at least one genetic marker detected is genetically linked to at least one of the genetic markers for powdery mildew resistance described herein by less than about 5 cM.

[0025] According to various aspects, the present disclosure provides a population of Cannabis plants with enhanced powdery mildew resistance.

[0026] According to various aspects, the present disclosure provides a population of Cannabis plants with one or more genetic markers for powdery mildew resistance described herein.

[0027] According to various aspects, the present disclosure provides a molecular marker, e.g., a genetic marker, for use in breeding and/or selecting Cannabis plants with enhanced powdery mildew resistance. In an embodiment, the molecular marker comprises or consists of a genetic marker as described herein, e.g., having the sequence thereof, or a portion of the sequence thereof. In an embodiment, the molecular marker comprises or consists of an isolated nucleic acid molecule having the sequence of a genetic marker as described herein, or a portion thereof, or a nucleotide sequence having at least, greater than or about 75% sequence identity thereto, or greater than or about 85% sequence identity to the complementary sequence thereto.

[0028] According to various aspects, the present technology relates to an isolated nucleic acid molecule for identifying a genetic marker for powdery mildew resistance as described herein.

[0029] According to various aspects, the present technology relates to an isolated polypeptide encoded by the isolated nucleic acid molecule as described herein.

[0030] According to various aspects, the present technology relates to an antibody that specifically binds the isolated polypeptide described herein.

[0031] According to various aspects, the present technology relates to an organism, tissue or cell comprising the isolated nucleic molecule or the isolated polypeptide or the genetic marker as described herein. In some instances, the organism, tissue or cell is a plant, a plant tissue, or a plant cell. In some further instances, the organism, tissue or cell is a Cannabis plant, a Cannabis tissue, or a Cannabis cell.

[0032] Other aspects and features of the present disclosure will become apparent to those ordinarily skilled in the art upon review of the following description of specific embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[0033] All features of embodiments which are described in this disclosure are not mutually exclusive and can be combined with one another. For example, elements of one embodiment can be utilized in the other embodiments without further mention. A detailed description of specific embodiments is provided herein below with reference to the accompanying drawings in which:

[0034] **FIG. 1** shows results from a genome-wide scan (Chromosomes 1-9) for each of 4 related pedigrees (P1, P2, P3, P4) sharing the same powdery mildew resistance/susceptibility phenotype.

[0035] **FIG. 2** is a chart of LOD score vs. Mbp, which shows an enlarged view of a QTL region identified on Chromosome 2 for each of the 4 pedigrees (P1, P2, P3, P4).

[0036] **FIG. 3** is a chart of phenotype (AUDPC) vs. genotype (Phase A or B on P5) at the QTL for each of the 4 pedigrees (P1, P2, P3, P4).

[0037] **FIGs. 4A-4E** show SNP positions and alleles for the two phases (Phase A and Phase B) for powdery mildew resistance genetic markers in the QTL region on Chromosome 2.

[0038] **FIGs. 5A-5M** shows the sequences of powdery mildew resistance genetic markers in the QTL region on Chromosome 2; sequences are shown in the 5' to 3' orientation from left to right, with the Phase A and Phase B alleles shown in square brackets in the center of the sequence as [A/B].

[0039] **FIG. 6** is a graph showing the results of allele specific T-ARMS for the 91K.chr2Ap83851294[T/C] SNP. Agarose gel analysis of 91K.chr2Ap83851294[T/C] SNP marker. The 203 and 150 bp amplicons corresponds to C and T allele of the 91K.chr2Ap83851294 SNP. All the PCR reactions except the no template control (NTC) showed a 302 bp PCR control amplicon. The resistant parent

(in duplicate 1 and 2) and seven F_1 plants shows heterozygous genotype (both 203 and 150bp amplicon corresponding to C and T allele), while the susceptible parent (in duplicate 1 and 2) and seven F_1 shows homozygous CC genotype corresponds to a 203 bp amplicon.

[0040] **FIG. 7** is a graph illustrating representative result (allele description plot) of KASP/PACE genotype assay for SNP marker 91kv2Chr2Ap84275859. Heterozygous individuals with allele 1/2 are represented in green, while the homozygous individual with allele 1 are represented in red

DETAILED DESCRIPTION

[0041] The present technology is explained in greater detail below. This description is not intended to be a detailed catalog of all the different ways in which the technology may be implemented, or all the features that may be added to the instant technology. For example, features illustrated with respect to one embodiment may be incorporated into other embodiments, and features illustrated with respect to a particular embodiment may be deleted from that embodiment. In addition, numerous variations and additions to the various embodiments suggested herein will be apparent to those skilled in the art in light of the instant disclosure in which variations and additions do not depart from the present technology. Hence, the following description is intended to illustrate some particular embodiments of the technology, and not to exhaustively specify all permutations, combinations and variations thereof.

Definitions

[0042] As used herein, the singular form "a," "an" and "the" include plural referents unless the context clearly dictates otherwise.

[0043] The recitation herein of numerical ranges by endpoints is intended to include all numbers subsumed within that range (e.g., a recitation of 1 to 5 includes 1, 1.25, 1.5, 1.75, 2, 2.45, 2.75, 3, 3.80, 4, 4.32, and 5).

[0044] The term "about" is used herein explicitly or not. Every quantity given herein is meant to refer to the actual given value, and it is also meant to refer to the approximation to such given value that would reasonably be inferred based on the ordinary skill in the art, including equivalents and approximations due to the experimental and/or measurement conditions for such given value. For example, the term "about" in the context of a given value or range refers to a value or range that is within 20%, preferably within 15%, more preferably within 10%, more preferably within 9%, more preferably within

8%, more preferably within 7%, more preferably within 6%, and more preferably within 5% of the given value or range.

[0045] The expression "and/or" where used herein is to be taken as specific disclosure of each of the two specified features or components with or without the other. For example, "A and/or B" is to be taken as specific disclosure of each of (i) A, (ii) B and (iii) A and B, just as if each is set out individually herein. The term "or" as used herein should in general be construed non-exclusively. For example, an embodiment of "a composition comprising A or B" would typically present an aspect with a composition comprising both A and B. "Or" should, however, be construed to exclude those aspects presented that cannot be combined without contradiction (e.g., a composition pH that is between 9 and 10 or between 7 and 8).

[0046] As used herein, the term "comprise" is used in its non-limiting sense to mean that items following the word are included, but items not specifically mentioned are not excluded.

[0047] As used herein, the term "Cannabis" refers to the genus of flowering plants in the family Cannabaceae regardless of species, subspecies, or subspecies variety classification. At present, there is no general consensus whether plants of genus Cannabis are comprised of a single or multiple species (McPartland & Guy, 2017). For example some describe Cannabis plants as a single species, *C. sativa* L., with multiple subspecies (Small & Cronquist, 1976)(McPartland & Small, 2020) while others classify Cannabis plants into multiple species, most commonly as *C. sativa* L. and *C. indica* Lam. and sometimes additionally as *C. ruderalis* Janisch. (Schultes et al., 1974), depending on multiple criteria including morphology, geographic origin, chemical content, and genetic measurements. Regardless, all plants of genus Cannabis can interbreed and produce fertile offspring (Small, 1972).

[0048] As used herein, the expression "powdery mildew" refers to a disease which affects a wide range of plants and is characterized by development of white mycelium or a powdery appearance on the leaf surface. Powdery mildew (also referred to as "PM") can be caused by many different fungal species in the genera *Erysiphe*, *Microsphaera*, *Phyllactinia*, *Podosphaera*, *Sphaerotheca*, and *Uncinula* (Braun and Cook, 2012). At present, there is no consensus on the species and races of powdery mildew pathogen affecting the Cannabis plant. Powdery mildew on cannabis has been reported to be caused by *Podosphaera macularis* (formerly *Sphaerotheca macularis*), *Leveillula taurica* (McPartland 1996), and *Golovinomyces* species including *G. cichoracearum* (Pépin et al 2018), *G. ambrosiae* (Wiseman et al., 2021) and *G. spadiceus* (Farinas and Peduto Hand 2020).

[0049] The term "strain" as used herein refers to different varieties of the plant genus *Cannabis*. For example, the term "strain" can refer to different pure or hybrid varieties of Cannabis plants. In some

instances, the Cannabis strain of the present technology can by a hybrid of two strains. Different Cannabis strains often exhibit distinct chemical compositions with characteristic levels of cannabinoids and terpenes, as well as other components. Differing cannabinoid and terpene profiles associated with different Cannabis strains can be useful e.g. for the treatment of different diseases, or for treating different subjects with the same disease.

[0050] As used herein, the term "cannabinoid" refers to a chemical compound belonging to a class of secondary compounds commonly found in plants of genus Cannabis, but also encompasses synthetic and semi-synthetic cannabinoids and any enantiomers thereof. In an embodiment, the cannabinoid is a compound found in a plant, e.g., a plant of genus Cannabis, and is sometimes referred to as a phytocannabinoid. In one embodiment, the cannabinoid is a compound found in a mammal, sometimes called an endocannabinoid. In one embodiment, the cannabinoid is made in a laboratory setting, sometimes called a synthetic cannabinoid. In one embodiment, the cannabinoid is derived or obtained from a natural source (e.g. plant) but is subsequently modified or derivatized in one or more different ways in a laboratory setting, sometimes called a semi-synthetic cannabinoid.

[0051] Synthetic cannabinoids and semi-synthetic cannabinoids encompass a variety of distinct chemical classes, for example and without limitation: the classical cannabinoids structurally related to THC, the non-classical cannabinoids (cannabimimetics) including the aminoalkylindoles, 1,5 diarylpyrazoles, quinolines, and arylsulfonamides as well as eicosanoids related to endocannabinoids.

[0052] In another embodiment, a cannabinoid is one of a class of diverse chemical compounds that may act on cannabinoid receptors such as CB_1 and CB_2 in cells that alter neurotransmitter release in the brain.

[0053] In many cases, a cannabinoid can be identified because its chemical name will include the text string "*cannabi*". However, there are a number of cannabinoids that do not use this nomenclature, such as for example those described herein.

[0054] As used herein, the expression "% by weight" is calculated based on dry weight of the total material.

[0055] Within the context of this disclosure, where reference is made to a particular cannabinoid, each of the acid and/or decarboxylated forms are contemplated as both single molecules and mixtures. In addition, salts of cannabinoids are also encompassed, such as salts of cannabinoid carboxylic acids. As well, any and all isomeric, enantiomeric, or optically active derivatives are also encompassed. In particular,

where appropriate, reference to a particular cannabinoid incudes both the "A Form" and the "B Form". For example, it is known that THCA has two isomers, THCA-A in which the carboxylic acid group is in the 1 position between the hydroxyl group and the carbon chain (A Form) and THCA-B in which the carboxylic acid group is in the 3 position following the carbon chain (B Form). Further, in some embodiments of the present disclosure, the cannabinoid is a cannabinoid dimer. The cannabinoid may be a dimer of the same cannabinoid (e.g. THC-THC) or different cannabinoids. In an embodiment of the present disclosure, the cannabinoid may be a dimer of THC, including for example Cannabisol.

[0056] In an embodiment, a cannabinoid may occur in its free form, or in the form of a salt; an acid addition salt of an ester; an amide; an enantiomer; an isomer; a tautomer; a prodrug; a derivative of an active agent of the present technology; different isomeric forms (for example, enantiomers and diastereoisomers), both in pure form and in admixture, including racemic mixtures; enol forms.

[0057] As used herein, the expressions "nucleic acid," "nucleic acid molecule," "oligonucleotide," and "polynucleotide" are each used herein to refer to a polymer of at least three nucleotides. In some embodiments, a nucleic acid comprises deoxyribonucleic acid (DNA). In some embodiments, a nucleic acid comprises ribonucleic acid (RNA). In some embodiments, a nucleic acid is single stranded. In some embodiments, a nucleic acid is double stranded. In some embodiments, a nucleic acid comprises both single and double stranded portions. Unless otherwise stated, the terms encompass nucleic acid-like structures with synthetic backbones, as well as amplification products. In some embodiments, nucleic acids of the present disclosure are linear nucleic acids.

[0058] As used herein, the term "gene" refers to a part of the genome that codes for a product (e.g., an RNA product and/or a polypeptide product). A "gene sequence" is a sequence that includes at least a portion of a gene (e.g., all or part of a gene) and/or regulatory elements associated with a gene. In some embodiments, a gene includes coding sequence; in some embodiments, a gene includes non-coding sequence. In some particular embodiments, a gene may include both coding (e.g., exonic) and non-coding (e.g., intronic) sequences. In some embodiments, a gene may include one or more regulatory elements (e.g., a promoter) that, for example, may control or impact one or more aspects of gene expression (e.g., cell-type-specific expression, inducible expression, etc.).

[0059] As used herein, the expression "coding sequence" refers to a sequence of a nucleic acid or its complement, or a part thereof, that: i) can be transcribed to an mRNA sequence that can be translated to produce a polypeptide or a fragment thereof; or ii) an mRNA sequence that can be translated to produce a

polypeptide or a fragment thereof. Coding sequences include exons in genomic DNA or immature primary RNA transcripts, which are joined together by the cell's biochemical machinery to provide a mature mRNA.

[0060] As used herein, the term "mutation" refers to a change introduced into a parental sequence, including, but not limited to, substitutions, insertions, deletions (including truncations). The consequences of a mutation include, but are not limited to, the creation of a new character, property, function, phenotype or trait not found in the protein encoded by the parental sequence, or the increase or reduction/elimination of an existing character, property, function, phenotype or trait not found in the protein encoded by the parental sequence.

[0061] The expression "degree or percentage of sequence identity" refers herein to the degree or percentage of sequence identity between two sequences after optimal alignment. Percentage of sequence identity (or degree of identity) is determined by comparing two aligned sequences over a comparison window, where the portion of the peptide or polynucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino-acid residue or nucleic acid base occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

[0062] As used herein, the term "isolated" refers to nucleic acids or polypeptides that have been separated from their native environment, including but not limited to virus, proteins, glycoproteins, peptide derivatives or fragments or polynucleotides. For example, the expression "isolated nucleic acid molecule" as used herein refers to a nucleic acid substantially free of cellular material or culture medium when produced by recombinant DNA techniques, or chemical precursors, or other chemicals when chemically synthesized. An isolated nucleic acid is also substantially free of sequences, which naturally flank the nucleic acid (i.e. sequences located at the 5′ and 3′ ends of the nucleic acid) from which the nucleic acid is derived.

[0063] Two nucleotide sequences or amino-acids are said to be "identical" if the sequence of nucleotide residues or amino-acids in the two sequences is the same when aligned for maximum correspondence as described below. Sequence comparisons between two (or more) peptides or polynucleotides are typically performed by comparing sequences of two optimally aligned sequences over a segment or "comparison window" to identify and compare local regions of sequence similarity. Optimal

alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman, Ad. App. Math 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, J. Mol. Biol. 48: 443 (1970), by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. (U.S.A.) 85: 2444 (1988), by computerized implementation of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wis.), or by visual inspection. Other alignment programs may also be used such as: "Multiple sequence alignment with hierarchical clustering", F. CORPET, 1988, Nucl. Acids Res., 16 (22), 10881-10890.

[0064] As used herein, the expression "conservative substitutions" refers to a substitution made in an amino acid sequence of a polypeptide without disrupting the structure or function of the polypeptide. Conservative amino acid substitutions may be accomplished by substituting amino acids with similar hydrophobicity, polarity, and R-chain length for one another. Additionally, by comparing aligned sequences of homologous proteins from different species, conservative amino acid substitutions may be identified by locating amino acid residues that have been mutated between species without altering the basic functions of the encoded proteins. Amino acid substitutions that are conservative are typically as follows: i) hydrophilic: Alanine (Ala) (A), Proline (Pro) (P), Glycine (Gly) (G), Glutamic acid (Glu) (E), Aspartic acid (Asp) (D), Glutamine (Gln) (Q), Asparagine (Asn) (N), Serine (Ser) (S), Threonine (Thr) (T); ii) Sulphydryl: Cysteine (Cys) (C); iii) Aliphatic: Valine (Val) (V), Isoleucine (Ile) (I), Leucine (Leu) (L), Methionine (Met) (M); iv) Basic: Lysine (Lys) (K), Arginine (Arg) (R), Histidine (His) (H); and v) Aromatic: Phenylalanine (Phe) (F), Tyrosine (Tyr) (Y), Tryptophan (Trp) (W).

[0065] An "expression system" as used herein refers to reagents and components (e.g. in a kit) and/or solutions comprising said reagents and components for recombinant protein expression, wherein the expression system is cell free and includes optionally translation competent extracts of whole cells and/or other translation machinery reagents or components optionally in a solution, said reagents and components optionally including RNA polymerase, one or more regulatory protein factors, one or more transcription factors, ribosomes, and tRNA, optionally supplemented with cofactors and nucleotides, and the specific gene template of interest. Chemical based expression systems are also included, optionally using unnaturally occurring amino acids. In some instances, the expression systems of the present technology are *in vitro* expression systems.

[0066] The expressions "transformed with", "transfected with", "transformation" and "transfection" are intended to encompass introduction of nucleic acid (e.g. a construct) into a cell by one of many possible techniques known in the art.

[0067] The term "primer" as used herein, typically refers to oligonucleotides that hybridize in a sequence specific manner to a complementary nucleic acid molecule (e.g., a nucleic acid molecule comprising a target sequence). In some embodiments, a primer will comprise a region of nucleotide sequence that hybridizes to at least 8, e.g., at least 10, at least 15, at least 20, at least 25, or 20 to 60 nucleotides of a target nucleic acid (i.e., will hybridize to a sequence of the target nucleic acid). In general, a primer sequence is identified as being either "complementary" (i.e., complementary to the coding or sense strand (+)), or "reverse complementary" (i.e., complementary to the anti-sense strand (−)). In some embodiments, the term "primer" may refer to an oligonucleotide that acts as a point of initiation of a template-directed synthesis using methods such as PCR (polymerase chain reaction) under appropriate conditions (e.g., in the presence of four different nucleotide triphosphates and a polymerization agent, such as DNA polymerase in an appropriate buffer solution containing any necessary reagents and at suitable temperature(s)). Such a template directed synthesis is also called "primer extension." For example, a primer pair may be designed to amplify a region of DNA using PCR. Such a pair will include a "forward primer" and a "reverse primer" that hybridize to complementary strands of a DNA molecule and that delimit a region to be synthesized and/or amplified.

[0068] As used herein, the expression "wild-type" refers to a typical or common form existing in nature; in some embodiments it is the most common form.

[0069] As used herein, "allele" generally refers to an alternative nucleic acid sequence at a particular locus; the length of an allele can be as small as 1 nucleotide base but is typically larger. For example, a first allele can occur on one chromosome, while a second allele occurs on a second homologous chromosome, e.g., as occurs for different chromosomes of a heterozygous individual, or between different homozygous or heterozygous individuals in a population. A "favorable allele" is the allele at a particular locus that confers, or contributes to, an agronomically desirable phenotype, or alternatively, is an allele that allows the identification of susceptible plants that can be removed from a breeding program or planting. A favorable allele of a marker is a marker allele that segregates with the favorable phenotype, or alternatively, segregates with susceptible plant phenotype, therefore providing the benefit of identifying disease prone plants. A favorable allelic form of a chromosome interval is a chromosome interval that includes a nucleotide sequence that contributes to superior agronomic performance at one or more genetic loci physically located on the chromosome interval.

[0070] "Allele frequency" refers to the frequency (proportion or percentage) at which an allele is present at a locus within an individual, within a line, or within a population of lines. For example, for an allele "A", diploid individuals of genotype "AA", "Aa", or "aa" have allele frequencies of 1.0, 0.5, or 0.0,

respectively. One can estimate the allele frequency within a line by averaging the allele frequencies of a sample of individuals from that line. Similarly, one can calculate the allele frequency within a population of lines by averaging the allele frequencies of lines that make up the population. For a population with a finite number of individuals or lines, an allele frequency can be expressed as a count of individuals or lines (or any other specified grouping) containing the allele. An allele positively correlates with a trait when it is linked to it and when presence of the allele is an indicator that the desired trait or trait form will occur in a plant comprising the allele. An allele negatively correlates with a trait when it is linked to it and when presence of the allele is an indicator that a desired trait or trait form will not occur in a plant comprising the allele.

[0071] "Crossed" or "cross" means to produce progeny via fertilization (e.g. cells, seeds or plants) and includes crosses between plants (sexual) and self fertilization (selfing).

[0072] "Gene" refers to a heritable sequence of DNA, i.e., a genomic sequence, with functional significance. The term "gene" can also be used to refer to, e.g., a cDNA and/or a mRNA encoded by a genomic sequence, as well as to that genomic sequence.

[0073] "Genotype" is the genetic constitution of an individual (or group of individuals) at one or more genetic loci, as contrasted with the observable trait (the phenotype). Genotype is defined by the allele(s) of one or more known loci that the individual has inherited from its parents. The term genotype can be used to refer to an individual's genetic constitution at a single locus, at multiple loci, or, more generally, the term genotype can be used to refer to an individual's genetic make-up for all the genes in its genome. A "haplotype" is the genotype of an individual at a plurality of genetic loci. Typically, the genetic loci described by a haplotype are physically and genetically linked, i.e., on the same chromosome interval. The terms "phenotype," or "phenotypic trait" or "trait" refers to one or more trait of an organism, i.e., the detectable characteristics of a cell or organism which can be influenced by genotype. The phenotype can be observable to the naked eye, or by any other means of evaluation known in the art, e.g., microscopy, biochemical analysis, genomic analysis, an assay for a particular disease resistance, etc. In some cases, a phenotype is directly controlled by a single gene or genetic locus, i.e., a "single gene trait." In other cases, a phenotype is the result of several genes.

[0074] "Germplasm" refers to genetic material of or from an individual (e.g., a plant), a group of individuals (e.g., a plant line, variety or family), or a clone derived from a line, variety, species, or culture. The germplasm can be part of an organism or cell, or can be separate from the organism or cell. In general, germplasm provides genetic material with a specific molecular makeup that provides a physical foundation

for some or all of the hereditary qualities of an organism or cell culture. As used herein, germplasm includes cells, seed or tissues from which new plants may be grown, or plant parts, such as leaves, stems, pollen, or cells that can be cultured into a whole plant.

[0075] "Plant" refers to a whole plant or any part thereof, such as a cell or tissue culture derived from a plant, comprising any of: whole plants, plant components or organs (e.g., leaves, stems, roots, etc,), plant tissues, seeds, plant cells, and/or progeny of the same. A plant cell is a biological cell of a plant, taken from a plant or derived through culture from a cell taken from a plant. As used herein, the expression "plant part" refers to any part of a plant including but not limited to the embryo, shoot, root, stem, seed, stipule, leaf, petal, flower bud, flower, ovule, bract, trichome, branch, petiole, internode, bark, pubescence, tiller, rhizome, frond, blade, ovule, pollen, stamen, and the like. The two main parts of plants grown in some sort of media, such as soil or vermiculite, are often referred to as the "above-ground" part, also often referred to as the "shoots", and the "below-ground" part, also often referred to as the "roots." Plant part may also include certain extracts such as kief or hash which includes Cannabis trichomes or glands.

[0076] "Polymorphism" means the presence of one or more variations in a population. A polymorphism may manifest as a variation in the nucleotide sequence of a nucleic acid or as a variation in the amino acid sequence of a protein. Polymorphisms include the presence of one or more variations of a nucleic acid sequence or nucleic acid feature at one or more loci in a population of one or more individuals. The variation may comprise but is not limited to one or more nucleotide base changes, the insertion of one or more nucleotides or the deletion of one or more nucleotides. A polymorphism may arise from random processes in nucleic acid replication, through mutagenesis, as a result of mobile genomic elements, from copy number variation and during the process of meiosis, such as unequal crossing over, genome duplication and chromosome breaks and fusions. The variation can be commonly found or may exist at low frequency within a population, the former having greater utility in general plant breeding and the latter may be associated with rare but important phenotypic variation. Useful polymorphisms may include single nucleotide polymorphisms (SNPs), insertions or deletions in DNA sequence (Indels), simple sequence repeats of DNA sequence (SSRs), a restriction fragment length polymorphism, and a tag SNP. A genetic marker, a gene, a DNA-derived sequence, a RNA-derived sequence, a promoter, a 5' untranslated region of a gene, a 3' untranslated region of a gene, microRNA, siRNA, a tolerance locus, a satellite marker, a transgene, mRNA, ds mRNA, a transcriptional profile, and a methylation pattern may also comprise polymorphisms. In addition, the presence, absence, or variation in copy number of the preceding may comprise polymorphisms.

[0077] A "population of plants" or "plant population" means a set comprising any number, generally more than one, of individuals, objects, or data from which samples are taken for evaluation, e.g. estimating QTL effects. Most commonly, the terms relate to a breeding population of plants from which members are selected and crossed to produce progeny in a breeding program. A population of plants can include the progeny of a single breeding cross or a plurality of breeding crosses, and can be either actual plants or plant derived material, or *in silica* representations of the plants. The population members need not be identical to the population members selected for use in subsequent cycles of analyses or those ultimately selected to obtain final progeny plants. Often, a plant population is derived from a single biparental cross, but may also derive from two or more crosses between the same or different parents. Although a population of plants may comprise any number of individuals, those of skill in the art will recognize that plant breeders commonly use population sizes ranging from one or two hundred individuals to several thousand, and that the highest performing 5-20% of a population is what is commonly selected to be used in subsequent crosses in order to improve the performance of subsequent generations of the population.

[0078] "Resistance allele" means the nucleic acid sequence associated with powdery mildew resistance or enhanced resistance to powdery mildew. Similarly, "resistance marker" or "resistance loci" means respectively the marker or loci associated with powdery mildew resistance or enhanced resistance to powdery mildew.

[0079] "Recombinant" in reference to a nucleic acid or polypeptide indicates that the material (e.g., a recombinant nucleic acid, gene, polynucleotide, polypeptide, etc.) has been altered by human intervention. The term recombinant can also refer to an organism that harbors recombinant material, e.g., a plant that comprises a recombinant nucleic acid is considered a recombinant plant.

[0080] "Resistance" or "enhanced resistance" in a plant to disease conditions (such as powdery mildew) is an indication that the plant is less affected by disease conditions with respect to yield, survivability and/or other relevant agronomic measures, compared to a less resistant, more "susceptible" plant. The terms "enhanced resistance" and "improved resistance" are used interchangeably. Resistance is a relative term, indicating that a "resistant" plant survives and/or produces better yields in disease conditions (e.g., powdery mildew conditions) compared to a different (less resistant) plant (e.g., a different Cannabis strain) grown in similar conditions. As used in the art, disease "resistance" is sometimes used interchangeably with disease "tolerance." One of skill in the art will appreciate that plant resistance to disease conditions varies widely, and can represent a spectrum of more-resistant or less-resistant phenotypes. However, by simple observation, one of skill in the art can generally determine the relative resistance or susceptibility of different plants, plant lines or plant families under disease conditions, and

furthermore, will also recognize the phenotypic gradations of "resistant." In an embodiment, a Cannabis plant with enhanced resistance to powdery mildew is a plant having enhanced powdery mildew resistance compared to a wild-type Cannabis plant. In another embodiment, a Cannabis plant with enhanced resistance to powdery mildew is a plant having enhanced powdery mildew resistance compared to a different, less resistant or more susceptible plant (e.g., a different Cannabis strain) grown in similar conditions.

[0081] As used herein, the term "chromosome interval" designates a contiguous linear span of genomic DNA that resides *in planta* on a single chromosome. The term also designates any and all genomic intervals defined by any of the markers set forth in this disclosure. The genetic elements located on a single chromosome interval are physically linked and the size of a chromosome interval is not particularly limited. In some aspects, the genetic elements located within a single chromosome interval are genetically linked, typically with a genetic recombination distance of, for example, less than or equal to 20 cM, or alternatively, less than or equal to 15 cM, or alternatively, less than or equal to 10 cM, or alternatively, less than or equal to 5 cM. That is, two genetic elements within a single chromosome interval undergo meiotic recombination at a frequency of less than or equal to about 20%, or less than or equal to15%, or less than or equal to 10%, or less than or equal to 5%, respectively.

[0082] The boundaries of a chromosome interval can be defined by genetic recombination distance or by markers. In one embodiment, the boundaries of a chromosome interval comprise markers. In another embodiment, the boundaries of a chromosome interval comprise markers that will be linked to the gene controlling the trait of interest, i.e., any marker that lies within a given interval, including the terminal markers that define the boundaries of the interval, and that can be used as a marker for the presence or absence of powdery mildew resistance. In one embodiment, the intervals described herein encompass marker clusters that co-segregate with powdery mildew resistance. The clustering of markers generally occurs in relatively small domains on the chromosomes, indicating the presence of a genetic locus controlling the trait of interest in those chromosome regions. The interval encompasses markers that map within the interval as well as the markers that define the terminal.

[0083] An interval described by the terminal markers that define the endpoints of the interval will include the terminal markers and any marker localizing within that chromosome domain, whether those markers are currently known or unknown. Although it is anticipated that one skilled in the art may describe additional polymorphic sites at marker loci in and around the markers identified herein, any marker within the chromosome intervals described herein that are associated with powdery mildew resistance fall within the scope of the present disclosure.

[0084] "Quantitative trait loci" or a "quantitative trait locus" (QTL) is a genetic domain that effects a phenotype that can be described in quantitative terms and can be assigned a "phenotypic value" which corresponds to a quantitative value for the phenotypic trait. A QTL can act through a single gene mechanism or by a polygenic mechanism. In some aspects, the disclosure provides QTL chromosome intervals, where a QTL (or multiple QTLs) that segregates with powdery mildew resistance is contained in those intervals. In one embodiment, the boundaries of chromosome intervals are drawn to encompass markers that will be linked to one or more QTL. In other words, the chromosome interval is drawn such that any marker that lies within that interval (including the terminal markers that define the boundaries of the interval) is genetically linked to the QTL. Each interval comprises at least one QTL, and furthermore, may indeed comprise more than one QTL. Close proximity of multiple QTL in the same interval may obfuscate the correlation of a particular marker with a particular QTL, as one marker may demonstrate linkage to more than one QTL. Conversely, e.g., if two markers in close proximity show co-segregation with the desired phenotypic trait, it is sometimes unclear if each of those markers identifies the same QTL or two different QTL. Regardless, knowledge of how many QTL are in a particular interval is not necessary to make or practice the present technology.

Chromosome intervals, QTLs and molecular markers associated with powdery mildew resistance

[0085] The present disclosure identifies QTLs and molecular markers linked to powdery mildew resistance. The present disclosure provides for strains with powdery mildew resistance. The present disclosure includes and provides for methods to introduce resistance into existing Cannabis strains as part of a powdery mildew management strategy. The present disclosure identifies loci associated with powdery mildew resistance, provides tightly linked single nucleotide polymorphisms (SNPs) and demonstrates that certain genetic markers are important for powdery mildew resistance.

[0086] As provided herein, powdery mildew resistance can be obtained using markers described herein, which tag certain genetic markers identified as linked to powdery mildew resistance. Such markers can be used in early generations to select for populations with resistance or used in later generations to characterize and prioritize which material to advance into yield testing trials. The markers provided herein also provide for the breeding of plants incorporating all known powdery mildew resistance alleles into a single strain. The present disclosure provides an expanded base of powdery mildew resistant material for breeding programs, and this larger base allows agronomically acceptable and high-yielding varieties with powdery mildew resistance.

[0087] In one embodiment, the present disclosure provides a plant comprising a nucleic acid molecule selected from the group consisting of sequences of genetic markers described herein and fragments thereof, and complements of both. In another embodiment, the present disclosure also provides a plant comprising the allele of at least one chromosome interval or genetic marker described herein, or fragments and complements thereof. The present disclosure also provides for any plant comprising any combination of one or more powdery mildew resistance loci linked to at least one marker selected from the group consisting of sequences of genetic markers described herein.

[0088] Locations in the Cannabis genome of QTLs and the chromosome intervals comprising markers closely linked to it are described herein. Genetic and physical map positions of markers and chromosome intervals associated with QTLs are described herein. In some embodiments, genetic map loci are represented in cM, with position zero being the first (most distal) marker known at the beginning of the chromosome on a Cannabis genetic map. As used herein, "cM" refers to the classical definition of a centimorgan (Edwards, J.H., 1919) wherein one cM is equal to a 1 % chance that a trait at one genetic locus will be separated from a trait at another locus due to crossing over in a single meiosis (meaning the traits cosegregate 99% of the time), and this definition is used herein to delineate map locations pertaining to this technology.

[0089] Thus, one skilled in the art can use the present technology to improve the efficiency of breeding for improved powdery mildew resistance in Cannabis by associating resistance phenotypes with genotypes at previously unknown resistance loci in the cannabis genome. Disclosed herein are chromosome intervals that comprise alleles responsible for phenotypic differences between powdery mildew resistant and powdery mildew susceptible cannabis lines. Example chromosome intervals are characterized by the genomic regions including and flanked by and including the markers described herein and comprise markers within or closely linked to (e.g., within 20 cM of) the loci described herein. The present technology also comprises other intervals whose borders fall between, or any interval closely linked to those intervals. Examples of markers useful for this purpose comprise the SNP markers described herein, or any marker that maps within the chromosome intervals described herein (including the termini of the intervals), or any marker linked to those markers. Such markers can be assayed simultaneously or sequentially in a single sample or population of samples. Accordingly, the markers and methods of the present disclosure can be utilized to guide the breeding of cannabis varieties with the desired complement (set) of allelic forms of chromosome intervals associated with superior agronomic performance (powdery mildew resistance, along with any other available markers for yield, other disease resistance, etc.). Any of the disclosed marker alleles can be introduced into a cannabis line using standard methods such as, without limitation, via introgression, by traditional breeding (or introduced via transformation, or both) to yield a Cannabis plant

with superior agronomic performance. The number of alleles associated with powdery mildew resistance that can be introduced or be present in a Cannabis plant of the present disclosure ranges from one to the number of alleles disclosed herein, each integer of which is incorporated herein as if explicitly recited.

[0090] In some embodiments, marker-assisted selection (MAS) using additional markers flanking either side of the QTLs described herein can provide further efficiency because an unlikely double recombination event would be needed to simultaneously break linkage between the locus and both markers. Moreover, using markers tightly flanking a locus, one skilled in the art of MAS can reduce linkage drag by more accurately selecting individuals that have less of the potentially deleterious donor parent DNA. Any marker linked to or among the chromosome intervals described herein could be useful and within the scope of the present technology.

[0091] Similarly, by identifying plants lacking the desired marker locus, susceptible or less resistant plants can be identified, and, e.g., eliminated from subsequent crosses. Similarly, these marker loci can be introgressed into any desired genomic background, germplasm, plant, line, variety, etc., as part of an overall MAS breeding program designed to enhance yield. The present disclosure also provides chromosome QTL intervals that find equal use in MAS to select plants that demonstrate powdery mildew resistance or enhanced powdery mildew resistance. Similarly, the QTL intervals can also be used to counterselect plants that are susceptible or have reduced resistance to powdery mildew.

[0092] In some embodiments, the present disclosure provides methods for selecting a Cannabis plant with enhanced powdery mildew resistance. These methods comprise detecting a powdery mildew resistant allele at a polymorphic locus in a chromosomal segment flanked by any two of marker loci described herein. In other embodiments, these methods comprise detecting a powdery mildew resistant allele at a polymorphic locus in a chromosomal segment flanked by any two of marker loci having the sequence of genetic markers described herein. In further embodiments, these methods comprise detecting a powdery mildew resistant haplotype in a chromosomal segment flanked by any two of marker loci described herein. In other embodiments, these methods comprise detecting a powdery mildew resistant allele at a polymorphic locus in a chromosomal segment flanked by any two of marker loci described herein. In further embodiments, these methods comprise detecting a powdery mildew resistant haplotype in a chromosomal segment flanked by any two of marker loci described herein. In other embodiments, these methods comprise detecting a powdery mildew resistant allele at a polymorphic locus in a chromosomal segment flanked by any two of marker loci described herein. In further embodiments, these methods comprise detecting a powdery mildew resistant haplotype in a chromosomal segment flanked by any two of marker loci described herein.

[0093] The present disclosure also extends to a method of making a progeny Cannabis plant. The method comprises crossing a first parent Cannabis plant with a second Cannabis plant and growing the female Cannabis plant under plant growth conditions to yield Cannabis plant progeny. Methods of crossing and growing Cannabis plants are well within the ability of those of ordinary skill in the art. Such Cannabis plant progeny can be assayed for alleles associated with powdery mildew resistance and, thereby, the desired progeny selected. Such progeny plants or seed can be sold commercially, used to provide cannabis for medicinal or recreational use, processed to obtain a desired constituent of the cannabis, or further utilized in subsequent rounds of breeding. At least one of the first or second Cannabis plants is a Cannabis plant of the present disclosure in that it comprises at least one of the allelic forms of the markers of the present disclosure, such that the progeny are capable of inheriting the allele.

[0094] In some embodiments, a method of the present disclosure is applied to at least one related Cannabis plant such as from progenitor or descendant lines in the subject Cannabis plants' pedigree such that inheritance of the desired resistance allele can be traced. The number of generations separating the Cannabis plants being subject to the methods of the present disclosure will generally be from 1 to 20, commonly 1 to 5, and typically 1, 2, or 3 generations of separation, and quite often a direct descendant or parent of the Cannabis plant will be subject to the method (i.e., one generation of separation). Thus, with this technology, one skilled in the art can detect the presence or absence of powdery mildew resistance genotypes in the genomes of Cannabis plants as part of a marker assisted selection program. In one embodiment, a breeder ascertains the genotype at one or more markers for a resistant parent, which contains a powdery mildew resistance allele, and the genotype at one or more markers for a susceptible parent, which lacks the resistance allele. For example, the markers of the present disclosure can be used in MAS in crosses by subjecting the segregating progeny to MAS to maintain resistance alleles. A breeder can then reliably track the inheritance of the resistance alleles through subsequent populations derived from crosses between the two parents by genotyping offspring with the markers used on the parents and comparing the genotypes at those markers with those of the parents. Depending on how tightly linked the marker alleles are with the trait, progeny that share genotypes with the resistant parent can be reliably predicted to express the resistant phenotype; progeny that share genotypes with the susceptible parent can be reliably predicted to express the susceptible phenotype. Thus, the laborious and inefficient process of manually phenotyping the progeny for powdery mildew resistance is avoided.

[0095] By providing the positions in the Cannabis genome of the intervals and the powdery mildew resistance associated markers within, this technology also allows one skilled in the art to identify other markers within the intervals disclosed herein or linked to the chromosome intervals disclosed herein.

[0096] Closely linked markers flanking the locus of interest that have alleles in linkage disequilibrium with a resistance allele at that locus may be effectively used to select for progeny plants with enhanced powdery mildew resistance. Thus, the markers described herein, as well as other markers genetically or physically mapped to the same chromosome interval, may be used to select for Cannabis plants with enhanced powdery mildew resistance. Typically, a set of these markers will be used (e.g., 2 or more, 3 or more, 4 or more, 5 or more, 6 or more, 7 or more, 8 or more, 9 or more, 10 or more) in the flanking region above the gene and a similar set in the flanking region below the gene. Optionally, as described above, a marker within the actual gene and/or locus may also be used. The parents and their progeny may be screened for these sets of markers, and the markers that are polymorphic between the two parents may be used for selection. In an introgression program, this allows for selection of the gene or locus genotype at the more proximal polymorphic markers and selection for the recurrent parent genotype at the more distal polymorphic markers. The choice of markers actually used to practice this technology is not particularly limited and can be any marker that maps within the chromosome intervals described herein, any marker closely linked (e.g., within 20 cM) to a marker in such chromosome intervals, or any marker selected from the markers described herein. Furthermore, since there are many different types of marker detection assays known in the art, it is not intended that the type of marker detection assay (e.g. RAPDs, RFLPs, SNPs, AFLPs, etc.) used to practice this technology be limited in any way.

[0097] Additional genetic markers can be used either in conjunction with the markers described herein or independently of the markers described herein to practice the methods of the present disclosure. Publicly available marker databases from which useful markers can be obtained can also be used.

Molecular Genetic Markers

[0098] As used herein, "marker," "genetic marker," "molecular marker," "marker nucleic acid," and "marker locus" refer to a nucleotide sequence or encoded product thereof (e.g., a protein) used as a point of reference when identifying a linked locus. A marker can be derived from genomic nucleotide sequence or from expressed nucleotide sequences (e.g., from a spliced RNA, a cDNA, etc.), or from an encoded polypeptide, and can be represented by one or more particular variant sequences, or by a consensus sequence. In another sense, a marker is an isolated variant or consensus of such a sequence. The term also refers to nucleic acid sequences complementary to or flanking the marker sequences, such as nucleic acids used as probes or primer pairs capable of amplifying the marker sequence. A "marker probe" is a nucleic acid sequence or molecule that can be used to identify the presence of a marker locus, e.g., a nucleic acid probe that is complementary to a marker locus sequence. Alternatively, in some aspects, a marker probe refers to a probe of any type that is able to distinguish (i.e., genotype) the particular allele that is present at

a marker locus. A "marker locus" is a locus that can be used to track the presence of a second linked locus, e.g., a linked locus that encodes or contributes to expression of a phenotypic trait. For example, a marker locus can be used to monitor segregation of alleles at a locus, such as a QTL, that are genetically or physically linked to the marker locus. Thus, a "marker allele," alternatively an "allele of a marker locus" is one of a plurality of polymorphic nucleotide sequences found at a marker locus in a population that is polymorphic for the marker locus.

[0099] As used herein, "marker" also refers to nucleic acid sequences complementary to the genomic sequences, such as nucleic acids used as probes. Markers corresponding to genetic polymorphisms between members of a population can be detected by methods well-established in the art. These include, for example and without limitation, PCR-based sequence specific amplification methods, detection of restriction fragment length polymorphisms (RFLP), detection of isozyme markers, detection of polynucleotide polymorphisms by allele specific hybridization (ASH), detection of amplified variable sequences of the plant genome, detection of self-sustained sequence replication, detection of simple sequence repeats (SSRs), detection of single nucleotide polymorphisms (SNPs), or detection of amplified fragment length polymorphisms (AFLPs). Well established methods are also known for the detection of expressed sequence tags (ESTs) and SSR markers derived from EST sequences and randomly amplified polymorphic DNA (RAPD).

[00100] A favorable allele of a marker is the allele of the marker that co-segregates with a desired phenotype (e.g., powdery mildew resistance). As used herein, a QTL marker has a minimum of one favorable allele, although it is possible that the marker might have two or more favorable alleles found in the population. Any favorable allele of that marker can be used advantageously for the identification and construction of powdery mildew resistant plant lines. Optionally, one, two, three or more favorable allele(s) of different markers are identified in, or introgressed into a plant, and can be selected for or against during MAS. Desirably, plants or germplasm are identified that have at least one such favorable allele that positively correlates with disease tolerance or improved disease tolerance. Alternatively, a marker allele that co-segregates with powdery mildew susceptibility also finds use with the technology, since that allele can be used to identify and counter select susceptible plants. Such an allele can be used for exclusionary purposes during breeding to identify alleles that negatively correlate with powdery mildew resistance, to eliminate susceptible plants or germplasm from subsequent rounds of breeding. In the present disclosure, favorable alleles confer powdery mildew resistance. The favorable alleles conferring resistance to powdery mildew may be referred to as "resistance alleles."

[00101] The more tightly linked a marker is with a DNA locus influencing a phenotype, the more reliable the marker is in MAS, as the likelihood of a recombination event unlinking the marker and the locus decreases. Markers containing the causal mutation for a trait, or that are within the coding sequence of a causative gene, are ideal as no recombination is expected between them and the sequence of DNA responsible for the phenotype.

[00102] Genetic markers are distinguishable from each other (as well as from the plurality of alleles of any one particular marker) on the basis of polynucleotide length and/or sequence. A number of Cannabis molecular markers are known in the art and are published or available from various sources. In general, any differentially inherited polymorphic trait (including a nucleic acid polymorphism) that segregates among progeny is a potential genetic marker.

[00103] In some embodiments of the disclosure, one or more marker alleles are selected for in a single plant or a population of plants. In these methods, plants are selected that contain favorable alleles from more than one powdery mildew resistance marker, or alternatively, favorable alleles from more than one powdery mildew resistance marker are introgressed into a desired germplasm. One of skill in the art recognizes that the identification of favorable marker alleles is germplasm-specific. The determination of which marker alleles correlate with powdery mildew resistance (or susceptibility) is determined for the particular germplasm under study. One of skill in the art recognizes that methods for identifying the favorable alleles are known in the art. Identification and use of such favorable alleles is within the scope of this disclosure. Furthermore, identification of favorable marker alleles in plant populations other than the populations used or described herein is within the scope of this disclosure.

Marker Detection

[00104] In some aspects, methods of the disclosure utilize an amplification step to detect/genotype a marker locus, but amplification is not always a requirement for marker detection (e.g., Southern blotting and RFLP detection may be used). Separate detection probes can also be omitted in amplification/detection methods, e.g., by performing a real time amplification reaction that detects product formation by modification of the relevant amplification primer upon incorporation into a product, incorporation of labeled nucleotides into an amplicon, or by monitoring changes in molecular rotation properties of amplicons as compared to unamplified precursors (e.g., by fluorescence polarization).

[00105] "Amplifying," in the context of nucleic acid amplification, is any process whereby additional copies of a selected nucleic acid (or a transcribed form thereof) are produced. In some embodiments, an amplification based marker technology is used wherein a primer or amplification primer pair is admixed

with genomic nucleic acid isolated from the first plant or germplasm, and wherein the primer or primer pair is complementary or partially complementary to at least a portion of the marker locus, and is capable of initiating DNA polymerization by a DNA polymerase using the plant genomic nucleic acid as a template. The primer or primer pair is extended in a DNA polymerization reaction having a DNA polymerase and a template genomic nucleic acid to generate at least one amplicon. In other embodiments, plant RNA is the template for the amplification reaction. In some embodiments, the QTL marker is an SNP type marker, and the detected allele is a SNP allele, and the method of detection is allele specific hybridization (ASH).

[00106] In general, the majority of genetic markers rely on one or more property of nucleic acids for their detection. Typical amplification methods include various polymerase-based replication methods, including the polymerase chain reaction (PCR), ligase mediated methods such as the ligase chain reaction (LCR) and RNA polymerase based amplification (e.g., by transcription) methods.

[00107] As used herein, an "amplicon" is an amplified nucleic acid, e.g., a nucleic acid that is produced by amplifying a template nucleic acid by any available amplification method (e.g., PCR, LCR, transcription, or the like).

[00108] As used herein, a "genomic nucleic acid" is a nucleic acid that corresponds in sequence to a heritable nucleic acid in a cell. Common examples include nuclear genomic DNA and amplicons thereof. A genomic nucleic acid is, in some cases, different from a spliced RNA, or a corresponding cDNA, in that the spliced RNA or cDNA is processed, e.g., by the splicing machinery, to remove introns. Genomic nucleic acids optionally comprise non-transcribed (e.g., chromosome structural sequences, promoter regions, enhancer regions, etc.) and/or non-translated sequences (e.g., introns), whereas spliced RNA/cDNA typically do not have non-transcribed sequences or introns.

[00109] As used herein, a "template nucleic acid" is a nucleic acid that serves as a template in an amplification reaction (e.g., a polymerase-based amplification reaction such as PCR, a ligase mediated amplification reaction such as LCR, a transcription reaction, or the like). A template nucleic acid can be genomic in origin, or alternatively, can be derived from expressed sequences, e.g., a cDNA or an EST. Details regarding the use of these and other amplification methods can be found in any of a variety of standard texts. Many available biology texts also have extended discussions regarding PCR and related amplification methods and one of skill in the art will appreciate that essentially any RNA can be converted into a double stranded DNA suitable for restriction digestion, PCR expansion and sequencing using reverse transcriptase and a polymerase.

[00110] PCR detection and quantification using dual-labeled fluorogenic oligonucleotide probes, commonly referred to as "TaqMan™" probes, can also be performed according to the present disclosure. These probes are typically composed of short (e.g., 20-25 base) oligodeoxynucleotides that are labeled with two different fluorescent dyes. On the 5' terminus of each probe is a reporter dye, and on the 3' terminus of each probe a quenching dye is found. The oligonucleotide probe sequence is complementary to an internal target sequence present in a PCR amplicon. When the probe is intact, energy transfer occurs between the two fluorophores and emission from the reporter is quenched by the quencher by FRET. During the extension phase of PCR, the probe is cleaved by 5' nuclease activity of the polymerase used in the reaction, thereby releasing the reporter from the oligonucleotide-quencher and producing an increase in reporter emission intensity. TaqMan™ probes are oligonucleotides that have a label and a quencher, where the label is released during amplification by the exonuclease action of the polymerase used in amplification, providing a real time measure of amplification during synthesis. A variety of TaqMan™ reagents are commercially available, e.g., from Applied Biosystems as well as from a variety of other vendors.

[00111] In one embodiment, the presence or absence of a molecular marker is determined simply through nucleotide sequencing of the polymorphic marker region. This method is readily adapted to high throughput analysis as are the other methods noted above, e.g., using available high throughput sequencing methods such as sequencing by hybridization.

[00112] In alternative embodiments, *in silica* methods can be used to detect the marker loci of interest. For example, the sequence of a nucleic acid comprising the marker locus of interest can be stored in a computer. The desired marker locus sequence or its homolog can be identified using an appropriate nucleic acid search algorithm as provided by, for example, in such readily available programs as BLAST, or even simple word processors.

[00113] While the exemplary markers provided in the figures and tables herein are SNP markers, any of the aforementioned marker types can be employed in the context of the disclosure to identify chromosome intervals encompassing genetic element(s) that contribute to superior agronomic performance (e.g., powdery mildew resistance or enhanced powdery mildew resistance).

Primers and Probes

[00114] In general, synthetic methods for making oligonucleotides, including probes, primers, molecular beacons, PNAs (peptide nucleic acids), LNAs (locked nucleic acids), etc., are known. For example, oligonucleotides can be synthesized chemically according to a solid phase phosphoramidite

triester method. Oligonucleotides, including modified oligonucleotides, can also be ordered from a variety of commercial sources.

[00115] Nucleic acid probes to the marker loci can be cloned and/or synthesized. Any suitable label can be used with a probe of the disclosure. Detectable labels suitable for use with nucleic acid probes include, for example, any composition detectable by spectroscopic, radioisotopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels include biotin for staining with labeled streptavidin conjugate, magnetic beads, fluorescent dyes, radio labels, enzymes, and colorimetric labels. Other labels include ligands which bind to antibodies labeled with fluorophores, chemiluminescent agents, and enzymes. A probe can also constitute radio labeled PCR primers that are used to generate a radio labeled amplicon.

[00116] It is not intended that the nucleic acid probes of the disclosure be limited to any particular size. In some embodiments, the molecular markers of the disclosure are detected using a suitable PCR-based detection method, where the size or sequence of the PCR amplicon is indicative of the absence or presence of the marker (e.g., a particular marker allele). In these types of methods, PCR primers are hybridized to the conserved regions flanking the polymorphic marker region. As used in the art, PCR primers used to amplify a molecular marker are sometimes termed "PCR markers" or simply "markers." It will be appreciated that suitable primers to be used with the technology can be designed using any suitable method. It is not intended that the technology be limited to any particular primer or primer pair. In some embodiments, the primers of the disclosure are radiolabelled, or labeled by any suitable means (e.g., using a non-radioactive fluorescent tag), to allow for rapid visualization of the different size amplicons following an amplification reaction without any additional labeling step or visualization step. In some embodiments, the primers are not labeled, and the amplicons are visualized following their size resolution, e.g., following agarose gel electrophoresis. In some embodiments, ethidium bromide staining of the PCR amplicons following size resolution allows visualization of the different size amplicons. It is not intended that the primers of the disclosure be limited to generating an amplicon of any particular size. For example, the primers used to amplify the marker loci and alleles herein are not limited to amplifying the entire region of the relevant locus. The primers can generate an amplicon of any suitable length that is longer or shorter than those disclosed herein. In some embodiments, marker amplification produces an amplicon at least 20 nucleotides in length, or alternatively, at least 50 nucleotides in length, or alternatively, at least 100 nucleotides in length, or alternatively, at least 200 nucleotides in length, or alternatively at least 500 nucleotides in length. Marker alleles in addition to those recited herein also find use with the present disclosure.

Linkage Analysis and QTL

[00117] As used herein, "linkage" or "genetic linkage" describes the degree with which one marker locus is "associated with" another marker locus or some other locus (for example, a powdery mildew resistance locus). For example, if locus A has genes "A" or "a" and locus B has genes "B" or "b" then a cross between parent 1 with AABB and parent 2 with aabb will produce four possible gametes where the genes are segregated into AB, Ab, aB and ab. The null expectation is that there will be independent equal segregation into each of the four possible genotypes, i.e. with no linkage 1/4 of the gametes will be of each genotype. Segregation of gametes into genotypes differing from 1/4 is attributed to linkage. As used herein, linkage can be between two markers, or alternatively between a marker and a phenotype. A marker locus can be associated with (linked to) a trait, e.g., a marker locus can be associated with powdery mildew resistance or enhanced powdery mildew resistance when the marker locus is in linkage disequilibrium (LD) with the resistance trait. The degree of linkage of a molecular marker to a phenotypic trait (e.g., a QTL) is measured, e.g., as a statistical probability of cosegregation of that molecular marker with the phenotype.

[00118] As used herein, the linkage relationship between a molecular marker and a phenotype is given as the statistical likelihood that the particular combination of a phenotype and the presence or absence of a particular marker allele is random. Thus, the lower the probability score, the greater the likelihood that a phenotype and a particular marker will cosegregate. In some embodiments, a probability score of 0.05 ($p=0.05$, or a 5% probability) of random assortment is considered a significant indication of co-segregation. However, the present disclosure is not limited to this particular standard, and an acceptable probability can be any probability of less than 50% ($p < 0.5$). For example, a significant probability can be less than 0.25, less than 0.20, less than 0.15, or less than 0.1. The phrase "closely linked," in the present disclosure, means that recombination between two linked loci occurs with a frequency of equal to or less than about 10% (i.e., they are separated on a genetic map by not more than 10 cM). In one embodiment, any marker of the disclosure is linked (genetically and physically) to any other marker that is at or less than 50 cM distant. In another embodiment, any marker of the disclosure is closely linked (genetically and physically) to any other marker that is in close proximity, e.g., at or less than 10 cM distant. Two closely linked markers on the same chromosome can be positioned 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0.75, 0.5 or 0.25 cM or less from each other.

[00119] Classical linkage analysis can be thought of as a statistical description of the relative frequencies of co-segregation of different traits. Linkage analysis is the well characterized descriptive framework of how traits are grouped together based upon the frequency with which they segregate together. That is, if two non-allelic traits are inherited together with a greater than random frequency, they are said

to be "linked." The frequency with which the traits are inherited together is the primary measure of how tightly the traits are linked, i.e., traits which are inherited together with a higher frequency are more closely linked than traits which are inherited together with lower (but still above random) frequency. The further apart on a chromosome the genes reside, the less likely they are to segregate together, because homologous chromosomes recombine during meiosis. Thus, the further apart on a chromosome the genes reside, the more likely it is that there will be a crossing over event during meiosis that will result in the marker and the DNA sequence responsible for the trait the marker is designed to track segregating separately into progeny. A common measure of linkage is the frequency with which traits cosegregate. This can be expressed as a percentage of cosegregation (recombination frequency) or, also commonly, in centiMorgans (cM).

[00120] Linkage analysis is used to determine which polymorphic marker allele demonstrates a statistical likelihood of co-segregation with the powdery mildew resistance phenotype (thus, a "resistance marker allele"). Following identification of a marker allele for co-segregation with the powdery mildew resistance phenotype, it is possible to use this marker for rapid, accurate screening of plant lines for the resistance allele without the need to grow the plants through their life cycle and await phenotypic evaluations, and furthermore, permits genetic selection for the particular resistance allele even when the molecular identity of the actual resistance QTL is unknown. Tissue samples can be taken, for example, from the endosperm, embryo, or mature/developing plant and screened with the appropriate molecular marker to rapidly determine which progeny contain the desired genetics. Linked markers also remove the impact of environmental factors that can often influence phenotypic expression. Because chromosomal distance is approximately proportional to the frequency of crossing over events between traits, there is an approximate physical distance that correlates with recombination frequency. Marker loci are themselves traits and can be assessed according to standard linkage analysis by tracking the marker loci during segregation. Thus, in the context of the present disclosure, one cM is equal to a 1 % chance that a marker locus will be separated from another locus (which can be any other trait, e.g., another marker locus, or another trait locus that encodes a QTL), due to crossing over in a single generation.

[00121] When referring to the relationship between two genetic elements, such as a genetic element contributing to resistance and a proximal marker, "coupling" phase linkage indicates the state where the "favorable" allele at the resistance locus is physically associated on the same chromosome strand as the "favorable" allele of the respective linked marker locus. In coupling phase, both favorable alleles are inherited together by progeny that inherit that chromosome strand. In "repulsion" phase linkage, the "favorable" allele at the locus of interest (e.g., a QTL for resistance) is physically linked with an "unfavorable" allele at the proximal marker locus, and the two "favorable" alleles are not inherited together (i.e., the two loci are "out of phase" with each other).

Quantitative Trait Loci

[00122] An allele of a QTL can comprise multiple genes or other genetic factors even within a contiguous genomic region or linkage group, such as a haplotype. As used herein, an allele of a powdery mildew resistance locus can encompass more than one gene or nucleotide sequence where each individual gene or nucleotide sequence is also capable of exhibiting allelic variation and where each gene or nucleotide sequence is also capable of eliciting a phenotypic effect on the quantitative trait in question. In an aspect of the present disclosure the allele of a QTL comprises one or more genes or nucleic acid sequences that are also capable of exhibiting allelic variation. The use of the term "an allele of a QTL" is thus not intended to exclude a QTL that comprises more than one gene or other genetic factors. Specifically, an "allele of a QTL" in the present technology can denote a haplotype within a haplotype window wherein a phenotype can be powdery mildew resistance. A haplotype window is a contiguous genomic region that can be defined, and tracked, with a set of one or more polymorphic markers wherein the polymorphisms indicate identity by descent. A haplotype within that window can be defined by the unique fingerprint of alleles at each marker. When all the alleles present at a given locus on a chromosome are the same, that plant is homozygous at that locus. If the alleles present at a given locus on a chromosome differ, that plant is heterozygous at that locus. Plants of the present disclosure may be homozygous or heterozygous at any particular resistance locus or for a particular polymorphic marker.

[00123] The principles of QTL analysis and statistical methods for calculating linkage between markers and useful QTL, or between any loci in a genome are well known in the art. Exemplary methods include penalized regression analysis, ridge regression, single point marker analysis, complex pedigree analysis, Bayesian MCMC, identity-by-descent analysis, interval mapping, composite interval mapping, and Haseman-Elston regression. QTL analyses are often performed with the help of a computer and specialized software available from a variety of public and commercial sources known to those of skill in the art.

[00124] In some embodiments of the present disclosure, a "LOD score" is used to indicate the likelihood that a marker is associated with a QTL. The LOD score essentially expresses how much more likely the data are to have arisen assuming the presence of a QTL than in its absence. The LOD threshold value for avoiding a false positive with a given confidence, say 95%, depends on the number of markers and the length of the genome. Graphs indicating LOD thresholds are set forth in Lander and Botstein, Genetics, 121:185-199 (1989), and further described by Arus and Moreno-Gonzalez, Plant Breeding, Hayward, Bosemark, Romagosa (eds.) Chapman & Hall, London, pp. 314-331 (1993). A $\log_{10}$ of an odds ratio (LOD) is calculated as: LOD=$\log_{10}$ MLE for the presence of a QTL (MLE given no linked QTL),

where MLE is a maximum likelihood estimate. As used herein, a nucleic acid marker is genetically linked to a QTL, where the marker nucleic acid molecule exhibits a LOD score of greater than 2.0, as judged by interval mapping, for powdery mildew resistance or partial resistance, preferably where the marker nucleic acid molecule exhibits a LOD score of greater than 3.0, as judged by interval mapping, for powdery mildew resistance or partial resistance, more preferably where the marker nucleic acid molecule exhibits a LOD score of greater than 3.5, as judged by interval mapping, for powdery mildew resistance or partial resistance, and even more preferably where the marker nucleic acid molecule exhibits a LOD score of about 4.0, as judged by interval mapping, for powdery mildew resistance or partial resistance based on maximum likelihood methods described by Lander and Botstein, Genetics, 121 :185-199 (1989), and implemented in the software package MAPMAKER/QTL (default parameters) (Lincoln and Lander, Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL, Whitehead Institute for Biomedical Research, Massachusetts (1990)).

Genetic Mapping

[00125] A "genetic map" is the relationship of genetic linkage among loci on one or more chromosomes (or linkage groups) within a given species, generally depicted in a diagrammatic or tabular form. "Genetic mapping" is the process of defining the linkage relationships of loci through the use of genetic markers, populations segregating for the markers, and standard genetic principles of recombination frequency. A "genetic map location" is a location on a genetic map relative to surrounding genetic markers on the same linkage group where a specified marker can be found within a given species. In contrast, a physical map of the genome refers to absolute distances (for example, measured in base pairs or isolated and overlapping contiguous genetic fragments, e.g., contigs). A physical map of the genome does not take into account the genetic behavior (e.g., recombination frequencies) between different points on the physical map. A "genetic recombination frequency" is the frequency of a crossing over event (recombination) between two genetic loci. Recombination frequency can be observed by following the segregation of markers and/or traits following meiosis. A genetic recombination frequency can be expressed in centimorgans (cM). In some cases, two different markers can have the same genetic map coordinates. In that case, the two markers are in such close proximity to each other that recombination occurs between them with such low frequency that it is undetected.

[00126] Genetic maps are graphical representations of genomes (or a portion of a genome such as a single chromosome) where the distances between markers are measured by the recombination frequencies between them. Plant breeders use genetic maps of molecular markers to increase breeding efficiency through Marker assisted selection (MAS), a process where selection for a trait of interest is not based on

the trait itself but rather on the genotype of a marker linked to the trait. A molecular marker that demonstrates reliable linkage with a phenotypic trait provides a useful tool for indirectly selecting the trait in a plant population, especially when accurate phenotyping is difficult, slow, or expensive.

[00127] In general, the closer two markers or genomic loci are on the genetic map, the closer they lie to one another on the physical map. A lack of precise proportionality between cM distances and physical distances can exist due to the fact that the likelihood of genetic recombination is not uniform throughout the genome; some chromosome regions are cross-over "hot spots," while other regions demonstrate only rare recombination events, if any.

[00128] Genetic mapping variability can also be observed between different populations of the same crop species. In spite of this variability in the genetic map that may occur between populations, genetic map and marker information derived from one population generally remains useful across multiple populations in identification of plants with desired traits, counter-selection of plants with undesirable traits and in guiding MAS. As one of skill in the art will recognize, recombination frequencies (and as a result, genetic map positions) in any particular population are not static. The genetic distances separating two markers (or a marker and a QTL) can vary depending on how the map positions are determined. For example, variables such as the parental mapping populations used, the software used in the marker mapping or QTL mapping, and the parameters input by the user of the mapping software can contribute to the QTL marker genetic map relationships. However, it is not intended that the technology be limited to any particular mapping populations, use of any particular software, or any particular set of software parameters to determine linkage of a particular marker or chromosome interval with the disease tolerance phenotype. It is well within the ability of one of ordinary skill in the art to extrapolate the novel features described herein to any gene pool or population of interest, and using any particular software and software parameters. Indeed, observations regarding genetic markers and chromosome intervals in populations in addition to those described herein are readily made using the teaching of the present disclosure.

Association Mapping

[00129] In one aspect, the present disclosure provides chromosome intervals, marker loci, germplasm for conducting genome-wide association mapping for powdery mildew resistance. Exemplary chromosome intervals and marker loci are provided herein. Smaller intervals defined by any two marker loci disclosed herein are also contemplated. Association or LD mapping techniques aim to identify genotype-phenotype associations that are significant. It is effective for fine mapping in outcrossing species where frequent recombination among heterozygotes can result in rapid LD decay. LD is non-random association of alleles

in a collection of individuals, reflecting the recombinational history of that region. Thus, LD decay averages can help determine the number of markers necessary for a genome-wide association study to generate a genetic map with a desired level of resolution.

[00130] Large populations are generally better for detecting recombination, while older populations are generally associated with higher levels of polymorphism, both of which contribute to accelerated LD decay. However, smaller effective population sizes tend to show slower LD decay, which can result in more extensive haplotype conservation. Understanding of the relationships between polymorphism and recombination is useful in developing strategies for efficiently extracting information from these resources. Association analyses compare the plants' phenotypic score with the genotypes at the various loci. Subsequently, any suitable Cannabis genetic map (for example, a composite map) can be used to help observe distribution of the identified QTL markers and/or QTL marker clustering using previously determined map locations of the markers.

Marker-Assisted Selection

[00131] "Introgression" refers to the transmission of a desired allele of a genetic locus from one genetic background to another. For example, introgression of a desired allele at a specified locus can be transmitted to at least one progeny via a sexual cross between two parents of the same species, where at least one of the parents has the desired allele in its genome. Alternatively, for example, transmission of an allele can occur by recombination between two donor genomes, e.g., in a fused protoplast, where at least one of the donor protoplasts has the desired allele in its genome. The desired allele can be, e.g., a selected allele of a marker, a QTL, a trans gene, or the like. In any case, offspring comprising the desired allele can be repeatedly backcrossed to a line having a desired genetic background and selected for the desired allele, to result in the allele becoming fixed in a selected genetic background.

[00132] A primary motivation for development of molecular markers in plant species is the potential for increased efficiency in plant breeding through marker assisted selection (MAS). Genetic markers are used to identify plants that contain a desired genotype at one or more loci, and that are expected to transfer the desired genotype, along with a desired phenotype to their progeny. Genetic markers can be used to identify plants containing a desired genotype at one locus, or at several unlinked or linked loci (e.g., a haplotype), and that would be expected to transfer the desired genotype, along with a desired phenotype to their progeny. The present disclosure provides the means to identify plants that are powdery mildew resistant, exhibit enhanced powdery mildew resistance or are susceptible to powdery mildew by identifying

plants having a specified allele that is linked to the phenotype or to one or more of the molecular markers described herein.

[00133] In general, MAS uses polymorphic markers that have been identified as having a significant likelihood of co-segregation with a resistance trait. Such markers are presumed to map near a gene or genes that give the plant its resistance phenotype, and are considered indicators for the desired trait, and are termed QTL markers. Plants are tested for the presence or absence of a desired allele in the QTL marker.

[00134] Identification of plants or germplasm that include a marker locus or marker loci linked to a resistance trait or traits provides a basis for performing marker assisted selection. Plants that comprise favorable markers or favorable alleles are selected for, while plants that comprise markers or alleles that are negatively correlated with tolerance can be selected against. Desired markers and/or alleles can be introgressed into plants having a desired genetic background to produce an introgressed resistance plant or germplasm. In some embodiments, it is contemplated that a plurality of resistance markers are sequentially or simultaneous selected and/or introgressed. The combinations of tolerance markers that are selected for in a single plant is not limited, and can include any combination of markers disclosed herein or any marker linked to the markers disclosed herein, or any markers located within the QTL intervals disclosed herein.

[00135] In some embodiments, the allele that is detected is a favorable allele that positively correlates with disease tolerance or improved disease tolerance. In the case where more than one marker is selected, an allele is selected for each of the markers; thus, two or more alleles are selected. In some embodiments, it can be the case that a marker locus will have more than one advantageous allele, and in that case, either allele can be selected. It will be appreciated that the ability to identify QTL marker loci alleles that correlate with powdery mildew resistance, enhanced powdery mildew resistance or susceptibility of a Cannabis plant to powdery mildew provides a method for selecting plants that have favorable marker loci as well. That is, any plant that is identified as comprising a desired marker locus (e.g., a marker allele that positively correlates with resistance) can be selected for, while plants that lack the locus, or that have a locus that negatively correlates with tolerance, can be selected against.

[00136] In some embodiments, a powdery mildew resistant first Cannabis plant or germplasm (the donor) can be crossed with a second Cannabis plant or germplasm (the recipient, depending on characteristics that are desired in the progeny) to create an introgressed Cannabis plant or germplasm as part of a breeding program designed to improve powdery mildew resistance of the recipient Cannabis plant or germplasm. In some aspects, the recipient plant can also contain one or more disease resistant loci, which can be qualitative or quantitative trait loci. In another aspect, the recipient plant can contain a transgene.

[00137] In some embodiments, the recipient Cannabis plant or germplasm will typically display increased resistance to powdery mildew conditions as compared to the first Cannabis plant or germplasm, while the introgressed Cannabis plant or germplasm will display an increased resistance to powdery mildew conditions as compared to the second plant or germplasm. An introgressed Cannabis plant or germplasm produced by these methods are also a feature of this technology.

[00138] MAS is a powerful shortcut to selecting for desired phenotypes and for introgressing desired traits into cultivars (e.g., introgressing desired traits into plant lines). MAS is easily adapted to high throughput molecular analysis methods that can quickly screen large numbers of plant or germplasm genetic material for the markers of interest and is much more cost effective than raising and observing plants for visible traits.

[00139] When a population is segregating for multiple loci affecting one or multiple traits, e.g., multiple loci involved in tolerance, or multiple loci each involved in tolerance or resistance to different diseases, the efficiency of MAS compared to phenotypic screening becomes even greater, because all of the loci can be evaluated in the lab together from a single sample of DNA.

Marker-Assisted Backcrossing

[00140] One application of MAS is to use the resistance or enhanced resistance markers to increase the efficiency of an introgression effort aimed at introducing a resistance QTL into a desired background. If the nucleic acids from a plant are positive for a desired genetic marker allele, the plant can be self-fertilized to create a true breeding line with the same genotype, or it can be crossed with a plant with the same marker or with other characteristics to create a sexually crossed hybrid generation.

[00141] Another use of MAS in plant breeding is to assist the recovery of the recurrent parent genotype by backcross breeding. Backcross breeding is the process of crossing a progeny back to one of its parents or parent lines. Backcrossing is usually done for the purpose of introgressing one or a few loci from a donor parent (e.g., a parent comprising desirable resistance marker loci) into an otherwise desirable genetic background from the recurrent parent. The more cycles of back crossing that are done, the greater the genetic contribution of the recurrent parent to the resulting introgressed variety. This is often necessary, because resistance plants may be otherwise undesirable, e.g., due to low yield, low fecundity, undesired chemotype, or the like. In contrast, strains which are the result of intensive breeding programs may have excellent yield, fecundity, favorable chemotype (e.g., THC and/or CBD levels) or the like, merely being deficient in one desired trait such as powdery mildew resistance.

[00142] Moreover, in another aspect, while maintaining the introduced markers associated with resistance, the genetic contribution of the plant providing resistance can be reduced by back-crossing or other suitable approaches. In one aspect, the nuclear genetic material derived from the donor material in the plant can be less than or about 50%, or less than or about 25%, or less than or about 13%, or less than or about 10%, or less than or about 7%, or less than or about 5%, or less than or about 3%, or less than or about 2% or less than or about 1%, as long as the recipient remains resistant to powdery mildew.

[00143] Genetic diversity is important for long term genetic gain in any breeding program. With limited diversity, genetic gain will eventually plateau when all of the favorable alleles have been fixed within the population. One objective is to incorporate diversity into a pool without losing the genetic gain that has already been made and with the minimum possible investment. MAS provides an indication of which genomic regions and which favorable alleles from the original ancestors have been selected for and conserved over time, facilitating efforts to incorporate favorable variation from other germplasm sources.

Methods of producing Cannabis plants resistant to powdery mildew

[00144] The present disclosure includes and provides for a method of producing Cannabis plants resistant to powdery mildew using molecular markers to select for one or more Cannabis plants containing at least one resistance loci described herein. Markers that are genetically linked to and can be used for selection of the resistance loci described herein are described in the present disclosure.

[00145] In some embodiments, the present disclosure provides methods for creating a population of Cannabis plants that are resistant or moderately resistant to powdery mildew, which methods comprise: (a) detecting in a first population of Cannabis plants or seeds the presence of at least one marker allele that is genetically linked to and within about 20, 15, 10, 5, 2.5, 1, 0.5, or 0.25 cM of at least one resistance loci described herein; (b) selecting a Cannabis plant or seed containing the genetically linked marker allele(s); and (c) producing a population of offspring from the selected Cannabis plant or seed. In some embodiments, the detection of the marker allele genetically linked to one or more powdery mildew resistance loci is performed concurrently, e.g., in a multiplexed reaction. In other embodiments, the detection of the marker allele genetically linked to one or more powdery mildew resistance loci is performed separately, e.g., in separate reaction.

[00146] In some embodiments of the disclosure, the method enables production of plants resistant to powdery mildew or having enhanced powdery mildew resistance, e.g., enhanced compared to a wild-type Cannabis plant. Systems, including automated systems for selecting plants that comprise a marker of interest, as described herein, and/or for correlating presence of the marker with powdery mildew resistance

are also a feature of the disclosure. These systems can include e.g. probes relevant to marker locus detection, detectors for detecting labels on the probes, appropriate fluid handling elements and temperature controllers that mix probes and templates and/or amplify templates and systems instructions that correlate label detection to the presence of a particular marker locus or allele.

[00147] In some embodiments, this technology could be used on any plant selected from the genus Cannabis. In one embodiment, the plant is selected from the species *Cannabis sativa*.

[00148] In some embodiments, the present disclosure provides a Cannabis plant to be assayed for resistance or susceptibility to powdery mildew by any method to determine whether a plant is resistant, susceptible, or whether it exhibits some degree of resistance or susceptibility. Populations of plants can be similarly characterized in this manner, or further characterized as segregating for the trait of powdery mildew resistance.

[00149] It is further understood that a Cannabis plant of the present disclosure may exhibit the characteristics of any chemotype. As used herein, the term "chemotype" refers to the cannabinoid chemical phenotype in individual Cannabis strains. In general, chemotype is primarily determined by, but not limited to, chemical ratios or predominance of CBD, THC, and CBG and/or their acid counterparts CBDA, THCA, and CBGA present in mature or semi-mature Cannabis flower. For example, Small and Beckstead assigned chemotypes based on ratios of THCA and CBDA: plants producing primarily THCA (Type I), CBDA (Type III) or both THCA and CBDA (Type II) (Small & Beckstead, 1973). Much rarer CBGA-dominant hemp plants were later identified as a new chemotype (Type IV) (Fournier et al., 1987). Cannabis with less than 0.3% total THC by dry weight is recognized as hemp. In some embodiments of the present technology, the Cannabis strain has a low-THC/high-CBD chemotype, e.g., having a tetrahydrocannabinol (THC) content of between about 0.05% and about 0.5% by weight, or between about 0.05% and about 0.25% by weight, or between 0.05% and about 0.1% by weight, and/or a cannabidiol (CBD) content of between about 0.01% and about 30% by weight, or between about 0.01% and about 25% by weight, or between about 0.01% and about 20% by weight, or between about 0,01% and about 15% by weight, or between about 0.01% and about 10% by weight, or between about 1% and about 30% by weight, or between about 1% and about 25% by weight, or between about 1% and about 20% by weight, or between about 1% and about 15% by weight, or between about 1% and about 10% by weight.

[00150] The present disclosure also provides for parts of the Cannabis plants of the present disclosure. Plant parts, without limitation, include seed, endosperm, ovule and pollen. In one embodiment of the present

disclosure, the plant part is a seed. In another embodiment of the present disclosure, the plant part is a plant cell.

[00151] In some embodiments, the Cannabis seed can be subjected to various treatments. For example, the seeds can be treated to improve germination by priming the seeds or by disinfection to protect against seed-born pathogens. In another embodiment, seeds can be coated with any available coating to improve, for example, plantability, seed emergence, and protection against seed-born pathogens. Seed coating can be any form of seed coating including, but not limited to, pelleting, film coating, and encrustments.

[00152] In another embodiment, the Cannabis plant can show a comparative powdery mildew resistance compared to a non-resistant control Cannabis plant. In this aspect, a control Cannabis plant will generally be genetically similar except for the powdery mildew resistance allele or alleles in question. Such plants can be grown under similar conditions with equivalent or near equivalent exposure to the powdery mildew.

[00153] In a further aspect, this disclosure provides processed products made from the disclosed Cannabis plants. Such products include, but are not limited to, meal, oil, plant extract, starch, or fermentation or digestion product.

[00154] In some embodiments, the present technology relates to an isolated nucleic acid molecule having at least about 75%, or at least about 80%, or at least about 85%, at least about 86%, or at least about 87%, or at least about 88%, or at least about 89%, or at least about 90%, or at least about 91%, or at least about 92%, or at least about 93%, or at least about 94%, or at least about 95%, or at least about 96%, or at least about 97%, or at least about 98%, or at least about 99% sequence identity to the sequence of a powdery mildew resistance marker described herein.

[00155] In some embodiments, the present technology relates to nucleic acid molecules that hybridize to the above disclosed sequences. Hybridization conditions may be stringent in that hybridization will occur if there is at least about a 96% or about a 97% sequence identity with the nucleic acid molecule described herein. The stringent conditions may include those used for known Southern hybridizations such as, for example, incubation overnight at 42°C. in a solution having 50% formamide, 5× SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5× Denhardt's solution, 10% dextran sulfate, and 20 mg/mL denatured, sheared salmon sperm DNA, following by washing the hybridization support in 0.1× SSC at about 65°C. Other known hybridization conditions are well known and are described in

Sambrook et al., Molecular Cloning: A Laboratory Manual, Third Edition, Cold Spring Harbor, N.Y. (2001).

[00156] In some embodiments, the isolated nucleic acid of the present technology comprises a nucleic acid sequence as set forth herein or a fragment thereof.

[00157] Fragments contemplated by the present technology, but are not limited to, fragments having a nucleic acid sequence as set forth in any one of the sequences described herein, as well as sequences with at least or about 85% or more sequence identity thereto.

[00158] In some embodiments, the isolated nucleic acid molecule of the present technology comprises at least and/or up to or about 15, at least and/or up to or about 20 at least and/or up to or about 25, at least and/or up to or about 30, at least and/or up to or about 40 at least and/or up to or about 50, at least and/or up to or about 60, at least and/or up to or about 70, at least and/or up to or about 80, at least and/or up to or about 90, at least and/or up to 100, at least or up to or about 200, at least or up to or about 300, at least or up or about 400, at least or up to or about 500, at least or up to or about 600, at least or up to or about 700, at least or up to or about 800, at least or up to or about 900, at least or up to or about 1000, at least or up to or about 1100, at least or up to or about 1200, at least or up to or about 1300, at least or up to or about 1400 or at least or up to or about 1500 or about 1600 contiguous nucleotides of a powdery mildew resistance marker described herein. For example, the nucleic acid molecule can be from 15 contiguous nucleotides up to 1500 contiguous nucleotides or any range or number of nucleotides there between.

[00159] The length of the nucleic acid molecule described above will depend on the intended use. For example, if the intended use is as a primer or probe, for example, for PCR amplification or for screening a library, the length of the nucleic acid molecule will be less than the full length sequence, such as a fragment of for example, about 15 to about 50 nucleotides, or at least about 15 nucleotides of the sequences described herein and/or its complement. In these embodiments, the primers or probes may be substantially identical to a highly conserved region of the nucleic acid sequence or may be substantially identical to either the 5′ or 3′ end of the DNA sequence. In some cases, these primers or probes may use universal bases in some positions so as to be 'substantially identical' but still provide flexibility in sequence recognition. Suitable primer and probe hybridization conditions are well known in the art.

[00160] In an embodiment, the nucleic acid molecule can be used as a primer and for example comprises the nucleic acid sequence as described herein.

[00161] In an embodiment, the nucleic acid is conjugated to and/or comprises a heterologous moiety, such as a unique tail, purification tag or detectable label. The unique tail can be a specific nucleic acid sequence. The nucleic acid can for example be end labelled (5′ or 3′) or the label can be incorporated randomly during synthesis.

[00162] In one embodiment, the present technology provides an isolated nucleic acid that encodes for the polypeptide having an amino acid sequence of a powdery mildew resistance marker described herein.

[00163] In some embodiments, the present technology relates to an isolated polypeptide having at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98% or at least about 99% identity to an amino acid sequence of a powdery mildew resistance marker described herein.

[00164] In some embodiments, the present technology relates to a construct or an *in vitro* expression system having an isolated nucleic acid molecule having at least, greater than or about 75% sequence identity to a powdery mildew resistance marker described herein. Accordingly, the present technology further relates to a method for preparing a construct or *in vitro* expression system including such a sequence, or a fragment thereof, for introduction of the sequence or partial sequence in a sense or anti-sense orientation, or a complement thereof, into a cell.

[00165] In some embodiments, the present technology also provides for organisms, tissues or cells such as Cannabis plants, Cannabis tissue and Cannabis cells having at least one powdery mildew resistance loci or genetic marker described herein.

[00166] In some embodiments, the present technology also provides for organisms, tissues or cells that comprise the nucleic acids and/or the polypeptides as defined herein. In some embodiments, the organisms, tissues or cells are plants, plant tissues or plant cells that exhibit enhanced powdery mildew resistance. In some instances, such plants are Cannabis plants and such plant tissues and plant cells are Cannabis tissue and Cannabis cells.

EXAMPLES

[00167] The examples below are given so as to illustrate the practice of various embodiments of the present disclosure. They are not intended to limit or define the entire scope of this disclosure. It should be appreciated that the disclosure is not limited to the particular embodiments described and illustrated herein

but includes all modifications and variations falling within the scope of the disclosure as defined in the appended embodiments.

Example 1 – Identification of novel sources of genetic resistance to powdery mildew from Cannabis strains and development of a segregating F1 population.

[00168] Initially, 40 commercial cannabis strains (Canopy Growth Corporation, Smiths Falls, ON, Canada) were screened to evaluate level of resistance against natural levels of powdery mildew infection. These cannabis strains were naturally inoculated by placing plants that were heavily infected with powdery mildew into the growth room and allowing the air circulation to disperse the spores. Plants were evaluated weekly starting in the 2nd week after plantlets were transferred to the room using a 0-4 scale, where 0 was no visible symptoms, 1= 1-10 percent, 2= 11-30 percent, 3= 31 to 60 percent, and 4= 61 to 100 percent of total leaf area of a whole plant was infected with powdery mildew. This initial screening identified twelve strains as susceptible (score = 4) and three strains with no visible PM symptoms (score = 0). Results of the initial PM screenings were further confirmed in a second replicated trial using artificial inoculations.

[00169] Based on the results of the two screenings, five strains were selected as parents for the development of mapping populations. A nested association mapping (NAM) population design was used to develop an F1 mapping population with the resistant strain as a common parent and pollen accepter, and the remaining four strains as susceptible parents and pollen donors. Parental strains were treated with Silver Thiosulfate (STS) to induce male flowers to generate the above indicated crosses. Over 200 seeds from each cross were germinated to develop a 200 F1 population, and altogether 800 F1 from all four crosses were generated. Five clones from each set of 800 F1 seedlings were generated for phenotyping.

Example 2 – Phenotyping the F1 population

[00170] Around 800 healthy F1 clones were grown using standard cannabis cultivation practices. Briefly, rooted clones were transferred to 6"X6" rockwool blocks and grown under non-inductive vegetative growth conditions (16 hours light, 8 hours night) for three weeks. Plants were then inoculated with powdery mildew spores and grown under 12 hours light and night for inductive reproductive growth conditions. After the initial visible symptoms of powdery mildew growth were detected (around two weeks after the inoculation), plants were scored for PM symptoms using the above 0-4 scale. A total of seven weekly observations were recorded on flowering stage plants. In addition, the Area Under the Disease Progress Curve (AUDPC) over 7 weeks was calculated as a measure of PM disease severity and used for QTL analysis along with the 0-4 PM rating.

Example 3 – Whole genome sequencing of Parental genomes

[00171] Genomic DNA of all five parental strains was sequenced using Illumina™ short reads sequencing technology and Pair-End (PE) and Mate-Pair (MP) and Linked-reads sequencing libraries (10X Genomics™ Chromium™). All five strains were sequenced with total depth (coverage) in the range of 190-204X (calculated based on estimated genome size of 1.6 Gbp).

[00172] The sequencing data was processed and assembled using NRGene's DeNovoMAGIC™ assembler application version 3.0. The resultant analysis generated fully assembled, unphased and fully phased genome of the five parental strains. The scaffold N50 for the genome assemblies ranged from 8.2-18.5 Mb for unphased and 0.8-3.4 Mb for the phased genome assemblies.

Example 4 – SNP genotyping

[00173] The bi-parental F1 mapping population was genotyped using skim-based genotyping by sequencing (skim GBS) which uses a low coverage, whole genome sequencing approach for whole genome genotyping (Golicz et al., 2015). Genomic DNA from the 800 F1 was extracted using NucleoSpin™ Plant II (Macherey-Nagel™). NGS libraries were prepared using RIPTIDE™ kit (iGenomX) and sequencing was done on Illumina™ NovaSeq6000 using 150 PE sequencing. Approximately 200 progenies from each of the four populations were sequenced to around 2.1X coverage of the diploid cannabis genome size of 1.6 Gbp.

[00174] SNP genotyping and genetic maps were constructed using NRGene's GenoMAGIC™ big-data toolkit. Briefly, for SNP calling, unique k-mers were extracted from the assembled scaffolds of the parental lines of each population. The k-mers were filtered according to their coverage in the sequencing data of the parental lines and the F1 population data. These k-mers were then detected in the data of the population samples. For imputation of the missing data, a HMM-based imputation model was used to fill in the missing data for each of the progeny samples using the available skim-sequencing data of the F1 population. Further inheritance pattern in the F1 population was determined and genetic maps were generated.

Example 5 - Linkage map creation and QTL identification

[00175] A QTL mapping study is performed to identify the genetic region associated with resistance/susceptibility to powdery mildew using the NRGene's GenoMAGIC™ software. The haplotype marker (SNP) segregation patterns in the nested population lines are used to construct a high resolution genetic linkage map in collaboration with NRGene, based on the Genetic Analysis of Clonal F1 and Double

cross (GACD) program described by Zhang and others (2015) or the Heterozygous Mapping Strategy (HetMappS) of Hyma and others (2015). Quantitative trait locus (QTL) analyses of powdery mildew resistance/susceptibility, yield and cannabinoid content is performed using MapQTL 6 (Kyazma, Wageningen, The Netherlands) using the multiple-QTL-mapping (MQM) approach (Jansen, 1993), which tests the presence of a QTL every 1 cM between pairs of adjacent markers. A genome-wide LOD score is estimated using 1000 permutations (Doerge and Churchill, 1996) at a significance level of $\alpha = 0.05$. In order to control background variation caused by genetic variation in other regions of the genome and detect QTL with higher resolution, automated cofactor selection is conducted, by which a set of markers throughout the linkage map are used as cofactors. The percentage of phenotypic variation explained by QTL (% Explained) and the additive effects of loci are calculated with the maximum likelihood method (Kao, 2000). Marker patterns in the parents and individuals in the mapping population are aligned and visualized by NRGene's proprietary software (the haplotype-based marker approach was presented at PAG2018 by Dr Ruth Wagner (Monsanto), https://pag.confex.com/pag/xxvi/meetingapp.cgi/Paper/31991).

[00176] QTL analysis was conducted for 4 related pedigrees (P1, P2, P3, P4) sharing the same phenotype (powdery mildew resistance/susceptibility score). All pedigrees shared a single parent designated as P5. The phenotype (powdery mildew resistance score) was a resistance score transformed to AUDPC (Area Under the Disease Progress Curve) over 7 weeks. A genome-wide AUDPC QTL scan for each of the 4 pedigrees identified a single QTL on Chromosome 2/Phase B of P5 (78892031-79717276 bp) responsible for an increase in the disease score (which means higher susceptibility). The associated allele was inherited from the P5 parent and corresponds to genome Phase B as defined by the genome assembly. Results of the AUDPC QTL scan are shown in FIGs. 1-3, as follows: FIG. 1 shows results from the genome-wide scan for each of the 4 pedigrees; FIG. 2 shows an enlarged view of the QTL region identified on Chromosome 2; and FIG. 3 shows a chart of phenotype vs. genotype at the QTL for each of the 4 pedigrees. The coordinates of the identified QTL chromosome intervals associated with powdery mildew susceptibility (on Phase B of P5) and powdery mildew resistance (on Phase A of P5) are given in Table 1.

Table 1. QTL chromosome intervals associated with powdery mildew resistance/susceptibility on Chromosome 2.

Phenotype	Associated QTL chromosome interval	CI start	CI end	Size (bp)
Powdery mildew resistance	On P5 Phase A	83660977	84353662	692685
Powdery mildew susceptibility	On P5 Phase B	78892031	79717276	825245

Example 6 - Gene-based marker development: SNP marker associated with the QTL for resistance to cannabis powdery mildew

[00177] Haplotype-markers in the linkage map and physical map created by the *de novo* genome assembly are used to align both maps and identify candidate genes underlying the QTLs for disease resistance, yield and cannabinoid content in cannabis. Amplicon sequencing is used to confirm or identify polymorphic regions, including SNPs, in candidate genes falling into the QTL regions. Gene-specific markers are developed for the genes using a three-primer temperature switch PCR (TSP) technique (Tabone et al., 2009; Hayden et al., 2009). The assay consists of a pair of a gene (locus)-specific primers (GS) and a nested allele-specific primer (AS) that is designed to assay for SNP. Forward and reverse GS primers are designed with melting temperature between 60–65°C (63°C optimum), to produce s PCR product size greater than 400 bp and located at least 100 bp from the SNP. A single allele-specific primer (forward) will be designed with the melting temperature between 43–48°C (45°C optimum) with its 3' end ending at the SNP.

[00178] SNP markers in the QTL region on Chromosome 2 identified in Example 5 above were extracted. 328 SNP markers were identified. The SNP positions and alleles for the two phases (Phase A and Phase B) are shown in FIGs. 4A-4E. The SNP sequences are shown in FIGs. 5A-5M, in which sequences are shown in the 5' to 3' orientation from left to right, with the A and B alleles shown in square brackets in the center of the sequence as [A/B].

Example 7 - Validation of powdery mildew linked SNP markers and Molecular market development for breeding

[00179] Molecular markers are developed to screen plants at an early stage in development for powdery mildew resistance. Validation of developed molecular markers is initiated. Marker assisted selection of resistance to powdery mildew using SNP markers is conducted.

Example 8 – Haplotype-based marker development: SNP marker associated with the QTL for resistance to cannabis powdery mildew

Gel based allele-specific PCR assays

[00180] To tag and select the QTL region in the F_1 generation, SNPs for marker development were selected if they are heterozygous (ab) in the maternal genome (in our cross resistant parent) and homozygous (aa) in the paternal genome (in our cross the susceptible parent). Allele-specific PCR primer were designed to genotype the SNPs located within the PM resistant QTL region on the chromosome 2.

Markers in the QTL interval, and related sequences are described in table/list of sequences. Primers were designed from the SNP flanking sequences using tetra primer-amplification refractory mutation system-based polymerase chain reaction (T-ARMS-PCR; see Table 2) based approach (Medrano and Oliveira et al., 2014). Briefly, primers are designed in a such a way that two allele-specific amplicons are generated using two pairs of primers, one pair (outer forward and inner reverse) producing an amplicon representing the allele 1 and the other pair (inner forward and outer reverse) producing an amplicon representing the allele 2. The two outer primers (outer forward and outer reverse) generate the outer fragment of the SNP locus and acts as an internal PCR control. The inner PCR primer pair was designed with a deliberate mismatch at position -2 from the 3′ end to increase allele specificity of inner primers. PCR amplification was carried out at 95°C for 2 min, followed by 35 cycles of denaturation at 95°C for 30 sec, annealing at 55-60°C for 30 and extension at 72°C for 1 min, and a final extension at 72°C for 5 min. The annealing temperature was optimised for each marker separately. The resulted PCR products are visualized on 1.5% agarose gel.

Table 2: T-ATMS-PCR primer sequences of 91K.chr2Ap83851294 [T/C] SNP

SNP primer Id	Primer sequences	Genotype pattern (bp)
91K.chr2Ap83851294.C.IF	5'- CCATATTTACCGAATGTATCGCTC -3'	302 bp (outer)
91K.chr2Ap83851294.T.IR	5'- TTAGACAGACAAATTTACAATAGTCTA -3'	203 bp (C allele)
91K.chr2Ap83851294.OF	5'- CTTACAATGCCACATTATTAGAGAA -3'	150 bp (T allele)
91K.chr2Ap83851294.OR	5'- ATGATACTTGGTCTGTAATCAATGA -3'	
T-ATMS-PCR primer sequences of 91K.chr2Ap83929791 [C/T] SNP		
91K.chr2Ap83929791.C.IF	5'- TTTACAACTTCCATAACTGTCATGTC -3'	316 bp (outer)
91K.chr2Ap83929791.T.IR	5'- CTTAGAAGGGTACTCTCTACTCAAACA -3'	200 bp (C allele)
91K.chr2Ap83929791.OF	5'- TGTTATTCTAATGAAATCTGATCCG -3'	169 bp (T allele)
91K.chr2Ap83929791.OR	5'- GAACTTTAGTAGCAATGTGAGTTGG -3'	

[00181] To confirm the utility of the PCR primer identifying the PM resistant and susceptible plants in the F_1 generation, known PM resistant and susceptible plants/cultivars along with the segregating F_1 progenies from the cross were genotyped using the 91K.chr2Ap83851294[T/C] and 91K.chr2Ap83929791[C/T] primers. Figure 6 shows the results of 91K.chr2Ap83851294[T/C] SNP markers on known PM resistant and susceptible lines along with a few segregating F_1 progenies.

[00182] Prediction of disease response was made based on PCR product analysis whether or not the F_1 plant contains a heterozygous genotype similar to PM resistant parent or a homozygous genotype similar to susceptible parent. The predicted phenotypes (PM response) were confirmed with the observed response

of F_1 plants in an artificially inoculated PM trial. The results are as below (Table 3). All the predicted resistant F_1 plants had the resistant genotype while the all the susceptible F_1 plants had the susceptible genotype. Table 1: Results of the PM resistant QTL linked SNP markers using agarose gel-based T-ARMS marker assays along with their predicted and observed responses in a PM inoculation trial.

5

Table 3: Confirmation of Predicted phenotype

Plant Id	**Marker genotype for 91K.chr2Ap83851294[T/C] [Amplicon size and SNP genotype]**	**Marker genotype for 91K.chr2Ap83929791[C/T] [Amplicon size and SNP genotype]**	**Predicted phenotype based on genotype of PM QTL SNP markers**	**Observed response of plants (phenotype) in a control PM inoculation trial**
Resistant Parent plant 1	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
Resistant Parent plant 2	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
Susceptible Parent plant 1	302 and 203bp; CC	316, and 169bp; TT	S	S
Susceptible Parent plant 2	302 and 203bp; CC	316, and 169bp; TT	S	S
F_1-1902	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
F_1-1913	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
F_1-1923	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
F_1-1936	302 and 203bp; CC	316, and 169bp; TT	S	S
F_1-1951	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
F_1-1958	302 and 203bp; CC	316, and 169bp; TT	S	S
F_1-1962	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
F_1-1972	302 and 203bp; CC	316, and 169bp; TT	S	S
F_1-1981	302 and 203bp; CC	316, and 169bp; TT	S	S

F_1-1990	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
F_1-1991	302, 203 and 150bp; TC	316, 200 and 169bp; CT	R	R
F_1-2033	302 and 203bp; CC	316, and 169bp; TT	S	S
F_1-2083	302 and 203bp; CC	316, and 169bp; TT	S	S
F_1-2091	302 and 203bp; CC	316, and 169bp; TT	S	S
R: PM Resistant (average PM score 0-1, S: PM Susceptible (mean PM score 3-4) based on the 0-4 disease scoring scale.				

Fluorescent-based allele-specific SNP genotyping assays

[00183] Competitive allele specific PCR for SNP genotyping assay based on dual FRET (Fluorescent Resonance Energy Transfer) including KASP (Kompetitive Allele Specific PCR (LGC, Biosearch), and/or PACE (PCR Allele Competitive Extension, (3CR Bioscience, UK) assays were developed as per the manufacturer guidelines. Briefly, for each genotyping assay, two allele-specific primers, one primer for each allele of the SNP and one common genome specific primer was designed. The allele specific primers were designed with the melting temperature between 60-67 °C (64 °C optimum) with its 3' end ending at the SNP. The common reverse primers were designed with the melting temperature between 60–67 °C (66 °C optimum) and min and max length of 18-30 bp. The two allele-specific primers contain universal tail sequences on their 5' ends correspond with a universal FRET cassette: one labelled with FAM™ dye and the other with HEX™ dye. The reaction master mix contains all the components required for PCR and universal FRET reporting cassettes, which binds to the corresponding tail sequences and emits fluorescence (https://www.biosearchtech.com/how-does-kasp-work). If the genotype at a given SNP is homozygous, only one of the two possible fluorescent signals will be generated, and a mixed fluorescent signal will be generated If the genotype is heterozygous. PCR reactions were optimized with both high-quality genomic DNA using commercial DNA extraction kits as well as crud DNA extracts using alkaline lysis solutions. Genotyping reactions were carried out in a volume of 10 µL (96 well plate) or 5 µL volume (384 well plate). Genotyping reactions were carried out using enzyme activation (94°C for 15 min), 10 cycles of template denaturation at 94°C for 20 secs and annealing and extension at 65°C to 57°C for 60 secs (with 0.8 °C drop per cycle), followed by 30 cycles of denaturation at 94°C for 20 secs and annealing and extension at 57°C for 60 secs. Finally, read the plate at 30 °C for 1 minute. Additional one or two recycling was performed with three cycles of 94 °C for 20 secs and 57 °C for 60 secs. Fluorescence detection of the reactions was performed using a QuantStudio™ 6 Flex real-time PCR system (Thermo Fisher) and the data analysed

using the genotyping tool in the QuantStudio™ Design and Analysis Software Version 2.6. This generates allelic discrimination plot, with VIC or HEX as X-axis and FAM as Y-axis. PACE/KASP primer sequences linked with the QTL for PM resistance are listed in Table 4.

Table 4: PACE/KASP primer sequences

PACE/KASP primer Id	Primer Sequences
91Kv2Chr2Ap83738925-A1	5'-GAAGGTGACCAAGTTCATGCTTGAAAATGAGGTTGCCCAGAGT- 3'
91Kv2Chr2Ap83738925-A2	5'-GAAGGTCGGAGTCAACGGATTTGAAAATGAGGTTGCCCAGAGC- 3'
91Kv2Chr2Ap83738925-Com	5'-CTCTCTTGGCATGCAGCACA- 3'
91Kv2Chr2Ap83851939-A1	5'-GAAGGTGACCAAGTTCATGCTCGAGTCTTTTCACACCTTTGCTC- 3'
91Kv2Chr2Ap83851939-A2	5'-GAAGGTCGGAGTCAACGGATTCGAGTCTTTTCACACCTTTGCTA- 3'
91Kv2Chr2Ap83851939-Com	5'-TTGTCTGAGCCACTCAAGGCA- 3'
91kv2Chr2Ap84141949-A1	5'-GAAGGTGACCAAGTTCATGCTCCATTCGTGTCAACATTCCAAACA- 3'
91kv2Chr2Ap84141949-A2	5'-GAAGGTCGGAGTCAACGGATTCCATTCGTGTCAACATTCCAAACC- 3'
91kv2Chr2Ap84141949-Com	5'-GAAAGCCTGGTTGAACATGAGTCT- 3'
91kv2Chr2Ap84085089-A1	5'-GAAGGTGACCAAGTTCATGCTACAGGAAAAGTATTGGAATGCAATGT- 3'
91kv2Chr2Ap84085089-A2	5'-GAAGGTCGGAGTCAACGGATTACAGGAAAAGTATTGGAATGCAATGA- 3'
91kv2Chr2Ap84085089-Com	5'-GCTTTAGTATTAAGTTTGTTGAAGGGACT- 3'
91kv2Chr2Ap84026493-A1	5'-GAAGGTGACCAAGTTCATGCTGCAAGCAACCACCAAAACTGAT- 3'
91kv2Chr2Ap84026493-A2	5'-GAAGGTCGGAGTCAACGGATTGCAAGCAACCACCAAAACTGAA- 3'
91kv2ChrAp84026493-Com1	5'-GCTTGCGCTGCAATTCGGTAT- 3'
91kv2Chr2Ap84275859-A1	5'-GAAGGTGACCAAGTTCATGCTCCTGTTAAATCCATGCTGCATAGAG- 3'
91kv2Chr2Ap84275859-A2	5'-GAAGGTCGGAGTCAACGGATTCCTGTTAAATCCATGCTGCATAGAA- 3'
91kv2Chr2Ap84275859-Com	5'-GTCAACTGTGGTAACCCCCA- 3'

[00184] KASP/PACE SNP markers were validated on randomly selected 72 plants including parental lines and F_1 progenies from multiple populations and standard checks. Response of these lines against the PM infection was scored in an artificially inoculated PM trial. The 25 F_1 plants with the PM resistant phenotype also had the genotype like the resistant plant (source of PM resistance QTL), while the 37 F_1 plants with the PM susceptible phenotype had the marker genotype like the susceptible parents (Table 5) (FIG. 7).

Table 5: Validation of PACE/KASP based SNP marker assays associated with the QTL for resistance to cannabis powdery mildew.

plant Id	Predicted phenotype based on genotype of KASP/PACE marker. Resistant (R) and Susceptible (S)	Observed response of F_1 plants on 28 days post inoculation. Plants scoring 0-1 are considered resistant (R) and 2-4 as susceptible (S)

Resistant parent	R	R
Susceptible parent-1	R	R
Susceptible parent-2	S	S
Susceptible parent 3	S	S
F1-164	R	R
F1-165	S	S
F1-166	S	S
F1-167	R	R
F1-168	R	R
F1-169	S	S
F1-170	S	S
F1-171	S	S
F1-172	R	R
F1-175	S	S
F1-176	S	S
F1-177	R	R
F1-178	S	S
F1-179	S	S
F1-180	R	R
F1-181	R	R
F1-182	S	S
F1-183	S	S
F1-184	S	S
F1-185	S	S
F1-186	S	S
F1-187	S	S
F1-188	S	S
F1-189	R	R
F1-190	R	R
F1-191	R	R
F1-192	S	S
F1-194	R	R
F1-195	S	S
F1-196	S	S
F1-197	R	R
F1-198	R	R
F1-200	R	R
F1-201	S	S
F1-202	S	S
F1-203	R	R

F1-204	R	R
F1-205	S	S
F1-206	S	S
F1-207	S	S
F1-208	R	R
F1-211	S	S
F1-212	S	S
F1-218	S	S
F1-219	S	S
F1-220	R	R
F1-221	S	S
F1-222	S	S
F1-224	R	R
F1-225	R	R
F1-226	S	S
F1-227	R	R
F1-228	S	S
F1-229	R	R
F1-230	R	R
F1-231	S	S
F1-232	R	R
F1-233	S	S
F1-234	R	R
F1-235	S	S
F1-244	S	S
F1-247	S	S
Susceptible check 1	S	S
Susceptible check 2	S	S

Example 9 - Validation of powdery mildew linked SNP markers and Molecular market development for breeding

[00185] To validate the powdery mildew linked SNP markers and determine the effect of PM QTL in different genetic background, we developed three different F_1 populations by crossing the PM resistant cultivar Res-1 (donor of PM resistant QTL) with the three different susceptible cannabis cultivars (Sus-1, Sus-2, and Sus-3). The F_1 populations (212 plants) along with the resistant and susceptible checks were grown under standard growth conditions and artificially inculcated with PM spores as described in Example 2 - "Phenotyping the F_1 population". PM inoculated plants were scored after early visible signs of powdery

mildew growth were detected (approximately two weeks after inoculation) up to four weeks post-inoculation, using a 0-4 disease scoring scale. SNP markers linked with the PM resistant QTL on chromosome 2 were used to select for the donor QTL region from 91K (PM resistant plant). For that leaf samples from the young F_1 seedlings were collected and genotyped using PM linked markers. Based on the analysis of multiple PM linked marker, F_1 plants with heterozygous allele like the resistant parent were predicted as PM resistant, while the plants with homozygous allele (like the susceptible parent) were predicted as PM susceptible. Out of 212 F_1 plants, 105 plants were predicted as PM resistant and 107 as PM susceptible plants. The predicted PM phenotype was compared with the observed response of plants scored using 0-4 disease scoring scale. All three population were analysed separately. The mean score of the predicted resistant F_1 plants per population were ranged from 0.3 to 0.7 per population, while the mean score of predicted susceptible F_1 plant per population ranged from 2.9 to 3.8 (Table 6). This high correlation between marker based predicted plant response (phenotype) with the observed plant response (phenotype) in an artificially inoculated PM trial revealed utility of the identified QTL region, sequences, and SNP markers for marker-assisted selection for powdery mildew resistance in cannabis.

Table 6: Results of PM marker validation on three different F_1 populations.

F_1 population	Number of F_1 plants predicted as PM resistant based on the genotype of PM QTL markers	Average observed PM score of predicted resistant F_1 plants on 28 days post inoculation using 0-4 disease scoring scale.	Number of F_1 plants predicted as susceptible based on the genotype of PM QTL markers	Average observed PM score of predicted susceptible F_1 plants on 28 days post inoculation using 0-4 disease scoring scale.
Res-1 X Sus-1	30	0.3 (R)	27	3 (R)
Res-1 X Sus-2	36	0.4 (R)	36	2.9 (S)
Res-1 X Sus-3	39	0.7 (R)	44	3.8 (S)
Observed PM scores of parental material and checks used in the PM screening experiment on 28 days post inoculation.				
checks and parents		Number of plants	Average observed PM score based on 0-4 scale and	
Res-1 (PM resistant parent)		10	0 (R)	
Sus-1 (PM susceptible parent)		10	3.2 (S)	
Sus-2 (PM Susceptible parent)		10	3.6 (S)	
PM Susceptible check 1)		10	3.1 (S)	
PM Susceptible check 2)		10	4 (S)	
Plants scoring 0-1 disease score are considered resistant (R) and 2-4 as susceptible (S).				

INCORPORATION BY REFERENCE

[00186] All references cited in this specification, and their references, are incorporated by reference herein in their entirety where appropriate for teachings of additional or alternative details, features, and/or technical background.

EQUIVALENTS

[00187] While the disclosure has been particularly shown and described with reference to particular embodiments, it will be appreciated that variations of the above-disclosed and other features and functions, or alternatives thereof, may be desirably combined into many other different systems or applications. Also, that various presently unforeseen or unanticipated alternatives, modifications, variations or improvements therein may be subsequently made by those skilled in the art which are also intended to be encompassed by the following embodiments.

BIBLIOGRAPHY

Bengtsson, Therése, Inger Åhman, Outi Manninen, Lars Reitan, Therese Christerson, Jens Due Jensen, Lene Krusell, Ahmed Jahoor, and Jihad Orabi. "A novel QTL for powdery mildew resistance in nordic spring barley (Hordeum vulgare L. ssp. vulgare) revealed by genome-wide association study." Frontiers in plant science 8 (2017): 1954.

Di Marzo, V., & Piscitelli, F. (2015). The Endocannabinoid System and its Modulation by Phytocannabinoids. In *Neurotherapeutics*. https://doi.org/10.1007/s13311-015-0374-6

Doerge, Rebecca W., and Gary A. Churchill. "Permutation tests for multiple loci affecting a quantitative character." Genetics 142, no. 1 (1996): 285-294.

Edwards, J.H. (1919). Haldane and the mutation rate. Journal of Genetics 8:299-309.

ElSohly, Mahmoud A., Mohamed M. Radwan, Waseem Gul, Suman Chandra, and Ahmed Galal. "Phytochemistry of Cannabis sativa L." Phytocannabinoids (2017): 1-36.

Farnsworth, Norman R. "Pharmacognosy and Chemistry of" Cannabis Sativa"." Journal of the American Pharmaceutical Association 9, no. 8 (1969): 410.

Fournier, G., Richez-Dumanois, C., Duvezin, J., Mathieu, J. P., & Paris, M. (1987). Identification of a new chemotype in Cannabis sativa: Cannabigerol-dominant plants, biogenetic and agronomic prospects. *Planta Medica*. https://doi.org/10.1055/s-2006-962705

Ge, Xintian, Weiwei Deng, Zheng Zhou Lee, Francisco J. Lopez-Ruiz, Patrick Schweizer, and Simon R. Ellwood. "Tempered mlo broad-spectrum resistance to barley powdery mildew in an Ethiopian landrace." Scientific reports 6, no. 1 (2016): 1-10.

Golicz, A. A., P. E. Bayer, and D. Edwards. "Skim-based genotyping by sequencing In Plant Genotyping." (2015): 257-270.

Henning, J. A., D. H. Gent, M. S. Townsend, J. L. Woods, S. T. Hill, and D. Hendrix. "QTL analysis of resistance to powdery mildew in hop (Humulus lupulus L.)." Euphytica 213, no. 4 (2017): 98.

Hyma, Katie E., Paola Barba, Minghui Wang, Jason P. Londo, Charlotte B. Acharya, Sharon E. Mitchell, Qi Sun, Bruce Reisch, and Lance Cadle-Davidson. "Heterozygous mapping strategy (HetMappS) for high resolution genotyping-by-sequencing markers: a case study in grapevine." PloS one 10, no. 8 (2015): e0134880.

Jansen, Ritsert C. "Interval mapping of multiple quantitative trait loci." Genetics 135, no. 1 (1993): 205-211.

Kao, Chen-Hung. "On the differences between maximum likelihood and regression interval mapping in the analysis of quantitative trait loci." Genetics 156, no. 2 (2000): 855-865.

Li, Yahui, Xiaohan Shi, Jinghuang Hu, Peipei Wu, Dan Qiu, Yunfeng Qu, Jingzhong Xie et al. "Identification of a recessive gene PmQ conferring resistance to powdery mildew in wheat landrace Qingxinmai using BSR-Seq analysis." Plant disease 104, no. 3 (2020): 743-751.

McPartland, J. M., & Guy, G. W. (2017). Models of Cannabis Taxonomy, Cultural Bias, and Conflicts between Scientific and Vernacular Names. *Botanical Review*. https://doi.org/10.1007/s12229-017-9187-0

McPartland, J. M., & Small, E. (2020). A classification of endangered high-THC Cannabis (Cannabis sativa subsp. indica) domesticates and their wild relatives. *PhytoKeys*. https://doi.org/10.3897/PHYTOKEYS.144.46700

Nekrasov, Vladimir, Congmao Wang, Joe Win, Christa Lanz, Detlef Weigel, and Sophien Kamoun. "Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion." Scientific reports 7, no. 1 (2017): 1-6.

Pépin, N., Z. K. Punja, and D. L. Joly. "Occurrence of powdery mildew caused by Golovinomyces cichoraccarum sensu lato on Cannabis sativa in Canada." Plant Disease 102, no. 12 (2018): 2644-2644.

Punja, Zamir K., Danielle Collyer, Cameron Scott, Samantha Lung, Janesse Holmes, and Darren Sutton. "Pathogens and molds affecting production and quality of Cannabis sativa L." Frontiers in plant science 10 (2019): 1120.

Ren, Yan, Weixiu Hou, Caixia Lan, Bhoja R. Basnet, Ravi P. Singh, Wei Zhu, Xiyong Cheng, Dangqun Cui, and Feng Chen. "QTL analysis and nested association mapping for adult plant resistance to powdery mildew in two bread wheat populations." Frontiers in plant science 8 (2017): 1212.

Schultes, R. E., Schultes, R. E., Klein, W. M., Plowman, T., & Lockwood, T. E. (1974). Cannabis: an

example of taxonomic neglect. *Botanical Museum Leaflets, Harvard University.*

Scott, Cameron, and Zamir K. Punja. "Evaluation of disease management approaches for powdery mildew on Cannabis sativa L.(marijuana) plants." Canadian Journal of Plant Pathology (2020): 1-19.

Seifi, Alireza, Dongli Gao, Zheng Zheng, Stefano Pavan, Luigi Faino, Richard GF Visser, Anne-Marie A. Wolters, and Yuling Bai. "Genetics and molecular mechanisms of resistance to powdery mildews in tomato (Solanum lycopersicum) and its wild relatives." European journal of plant pathology 138, no. 3 (2014): 641-665.

Small, E., & Beckstead, H. D. (1973). Cannabinoid phenotypes in Cannabis sativa. *Nature*. https://doi.org/10.1038/245147a0

Small, Ernest. "Interfertility and chromosomal uniformity in Cannabis." Canadian Journal of Botany 50, no. 9 (1972): 1947-1949.

Small, Ernest, and Arthur Cronquist. "A practical and natural taxonomy for Cannabis." Taxon (1976): 405-435.

Sun, Suli, Zhongyi Wang, Haining Fu, Canxing Duan, Xiaoming Wang, and Zhendong Zhu. "Resistance to powdery mildew in the pea cultivar Xucai 1 is conferred by the gene er1." The Crop Journal 3, no. 6 (2015): 489-499.

Tabone, Tania, Diane E. Mather, and Matthew J. Hayden. "Temperature switch PCR (TSP): Robust assay design for reliable amplification and genotyping of SNPs." BMC genomics 10, no. 1 (2009): 1-14.

Thompson, George Richard, J. M. Tuscano, M. Dennis, A. Singapuri, S. Libertini, R. Gaudino, A. Torres, A. M. P. Delisle, J. D. Gillece, and J. M. Schupp. "A microbiome assessment of medical marijuana." Clin Microbiol Infect 23, no. 4 (2017): 269-270.

Xu, Xuewen, Ting Yu, Ruixue Xu, Yang Shi, Xiaojian Lin, Qiang Xu, Xiaohua Qi, Yiqun Weng, and Xuehao Chen. "Fine mapping of a dominantly inherited powdery mildew resistance major-effect QTL, Pm1. 1, in cucumber identifies a 41.1 kb region containing two tandemly arrayed cysteine-rich receptor-like protein kinase genes." Theoretical and applied genetics 129, no. 3 (2016): 507-516.

Zhang, Luyan, Lei Meng, Wencheng Wu, and Jiankang Wang. "GACD: Integrated software for genetic analysis in clonal F 1 and double cross populations." Journal of Heredity 106, no. 6 (2015): 741-744.

Zhao, Fukai, Yinghui Li, Baoju Yang, Hongbo Yuan, Cong Jin, Lixun Zhou, Hongcui Pei et al. "Powdery mildew disease resistance and marker-assisted screening at the Pm60 locus in wild diploid wheat Triticum urartu." The Crop Journal 8, no. 2 (2020): 252-259.

Braun, U., and Cook, R. T. A. 2012. Taxonomic Manual of the Erysiphales (Powdery Mildews). CBS Biodiversity Series 11. CBS-KNAW Fungal Biodiversity Centre, Utrecht, The Netherlands. Google

Scholar

Farinas, C., and F. Peduto Hand. "First Report of Golovinomyces spadiceus Causing Powdery Mildew on Industrial Hemp (Cannabis sativa) in Ohio." Plant Disease 104, no. 10 (2020): 2727-2727.

McPartland, J. M. "Cannabis pests." Journal of the International Hemp Association 3, no. 2 (1996): 49-52.

Wiseman, Michele S., Taylor Bates, Andrea Garfinkel, Cynthia M. Ocamb, and David H. Gent. "First report of Powdery Mildew Caused by Golovinomyces ambrosiae on Cannabis sativa in Oregon." Plant Disease ja (2021).

Medrano, Ruan Felipe Vieira, and Camila Andréa de Oliveira. "Guidelines for the tetra-primer ARMS–PCR technique development." Molecular biotechnology 56, no. 7 (2014): 599-608.

CLAIMS

What is claimed is:

1. A method for creating a population of Cannabis plants with enhanced powdery mildew resistance, the method comprising:

a) providing a first population of Cannabis plants;

b) detecting the presence of a genetic marker that is genetically linked to a powdery mildew resistance locus on a defined linkage group by 20 cM or less in the first population;

c) selecting one or more Cannabis plants containing said marker from the first population of Cannabis plants; and

d) producing a population of offspring from at least one of said selected Cannabis plants.

2. The method of claim 1, wherein the genetic marker detected is genetically linked to the powdery mildew resistance locus by less than about 15 cM.

3. The method of claim 1, wherein the genetic marker detected is genetically linked to the powdery mildew locus by less than about 10 cM.

4. The method of claim 1, wherein the genetic marker detected is genetically linked to the powdery mildew locus by less than about 5 cM.

5. The method of any one of claims 1 to 4, wherein the genetic marker detected is located within a chromosome interval comprising and flanked by a first marker and a second marker.

6. The method of claim 5, wherein the chromosome interval is located on Chromosome 2.

7. The method of claim 6, wherein the chromosome interval is located on Chromosome 2 at 83660977-84353662 bp.

8. The method of any one of claims 1 to 4, wherein the genetic marker detected is located within a chromosome interval comprising and flanked by a first nucleic acid sequence and a second nucleic acid sequence, optionally wherein the chromosome interval is located on Chromosome 2 at 83660977-84353662 bp.

9. The method of any one of claims 1 to 8, wherein the genetic marker is selected from the markers set forth in FIGs. 4A-4E.

10. The method of any one of claims 1 to 8, wherein the genetic marker has a sequence selected from the sequences set forth in FIGs. 5A-5M and sequences having at least, greater than or about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity thereto.

11. A method of creating a population of Cannabis plants comprising at least one allele associated with enhanced powdery mildew resistance comprising at least one sequence selected from a group of resistance marker nucleic acid sequences, the method comprising the steps of:

a) genotyping a first population of Cannabis plants, said population containing at least one allele associated with enhanced powdery mildew resistance, the at least one allele associated with enhanced powdery mildew resistance comprising at least one sequence selected from the group consisting of resistance marker nucleic acid sequences;

b) selecting from said first population one or more identified Cannabis plants containing said at least one allele associated with enhanced powdery mildew resistance comprising at least one sequence selected from the group consisting of resistance nucleic acid sequences; and

c) producing from said selected Cannabis plants a second population, thereby creating a population of Cannabis plants comprising at least one allele associated with enhanced powdery mildew resistance comprising at least one sequence selected from the group consisting of resistance nucleic acid sequences.

12. The method of claim 11, wherein the at least one allele comprises at least one allele set forth in FIGs. 4A-4E.

13. The method of claim 11 or 12, wherein the at least one sequence is selected from the sequences set forth in FIGs. 5A-5M and sequences having at least, greater than or about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity thereto.

14. A method for creating a population of Cannabis plants with enhanced powdery mildew resistance, the method comprising:

a) providing a first population of Cannabis plants;

b) concurrently detecting the presence of at least one genetic marker that is genetically linked to at least one resistance loci by about 20 cM or less in the first population;

c) selecting one or more Cannabis plants containing said at least one marker from the first population of Cannabis plants; and

d) producing a population of offspring from at least one of said selected Cannabis plants.

15. The method of claim 14, wherein the at least one genetic marker detected is genetically linked to at least one resistance loci by less than about 15 cM.

16. The method of claim 143, wherein the at least one genetic marker detected is genetically linked to at least one resistance loci by less than about 10 cM.

17. The method of claim 14, wherein the at least one genetic marker detected is genetically linked to at least one resistance loci by less than about 5 cM.

18. The method of any one of claims 14 to 17, wherein the at least one resistance loci is located on Chromosome 2.

19. The method of claim 18, wherein the at least one resistance loci is located within the chromosome interval on Chromosome 2 at 83660977-84353662 bp.

20. The method of any one of claims 14 to 19, wherein the at least one genetic marker is selected from the markers set forth in FIGs. 4A-4E.

21. The method of any one of claims 14 to 19, wherein the at least one genetic marker has a sequence selected from the sequences set forth in FIGs. 5A-5M and sequences having at least, greater than or about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity thereto.

22. The method of any one of claims 1 to 21, wherein the powdery mildew resistance in the population is enhanced at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% compared to a second population of plants that is susceptible to powdery mildew under the same growing conditions.

23. A population of Cannabis plants having enhanced powdery mildew resistance compared to a second population of wild-type plants under the same growing conditions.

24. A population of Cannabis plants having enhanced powdery mildew resistance compared to a second population of plants that is susceptible to powdery mildew under the same growing conditions.

25. A population of Cannabis plants created according to the method defined in any one of claims 1 to 22.

26. The population of Cannabis plants of any one of claims 23 to 25, wherein at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% of the plants in the population have enhanced powdery mildew resistance compared to wild-type plants.

27. The population of Cannabis plants of any one of claims 23 to 26, comprising at least one powdery mildew resistance chromosome interval or at least one resistance loci, optionally wherein the at least one powdery mildew resistance chromosome interval is located on Chromosome 2 at 83660977-84353662 bp, or the at least one resistance loci is located within the chromosome interval on Chromosome 2 at 83660977-84353662 bp.

28. The population of Cannabis plants of any one of claims 23 to 27, comprising at least one powdery mildew resistance genetic marker.

29. The population of Cannabis plants of claim 28, wherein the at least one powdery mildew resistance genetic marker is selected from the markers set forth in FIGs. 4A-4E.

30. The population of Cannabis plants of claim 28, wherein the at least one genetic marker has a sequence selected from the sequences set forth in FIGs. 5A-5M and sequences having at least, greater than or about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity thereto.

31. The population of Cannabis plants of any one of claims 23 to 30, comprising at least one genetic marker that is genetically linked to at least one resistance loci by about 20 cM or less.

32. The population of Cannabis plants of any one of claims 23 to 30, comprising at least one genetic marker that is genetically linked to at least one resistance loci by about 15 cM or less.

33. The population of Cannabis plants of any one of claims 23 to 30, comprising at least one genetic marker that is genetically linked to at least one resistance loci by about 10 cM or less.

34. The population of Cannabis plants of any one of claims 23 to 30, comprising at least one genetic marker that is genetically linked to at least one resistance loci by about 5 cM or less.

35. The population of Cannabis plants of any one of claims 23 to 34, wherein the Cannabis plants have a low THC and high CBD chemotype.

36. The population of Cannabis plants of any one of claims 23 to 35, wherein the Cannabis plants have a tetrahydrocannabinol (THC) content of between about 0.05% and about 0.25% by weight.

37. The population of Cannabis plants of any one of claims 23 to 36, wherein the Cannabis plants have a cannabidiol (CBD) content of between about 0.01% and about 10% by weight.

38. A Cannabis plant, plant part, tissue or cell thereof, wherein the Cannabis plant, plant part, tissue or cell thereof has enhanced powdery mildew resistance compared to a wild-type Cannabis plant, plant part, tissue or cell.

39. The Cannabis plant, plant part, tissue or cell of claim 38, wherein the Cannabis plant, plant part, tissue or cell comprises at least one powdery mildew resistance chromosome interval or at

least one resistance loci, optionally wherein the at least one powdery mildew resistance chromosome interval is located on Chromosome 2 at 83660977-84353662 bp, or the at least one resistance loci is located within the chromosome interval on Chromosome 2 at 83660977-84353662 bp.

40. The Cannabis plant, plant part, tissue or cell of claim 38 or 39, wherein the Cannabis plant, plant part, tissue or cell comprises at least one at least one powdery mildew resistance genetic marker.

41. The Cannabis plant, plant part, tissue or cell of any one of claims 38 to 40, comprising at least one genetic marker that is genetically linked to at least one resistance loci by about 20 cM or less, by about 15 cM or less, or by about 10 cM or less, or by about 5 cM or less.

42. The Cannabis plant, plant part, tissue or cell of claim 41, wherein the at least one genetic marker is selected from the markers set forth in FIGs. 4A-4E.

43. The Cannabis plant, plant part, tissue or cell of claim 41, wherein the at least one genetic marker has a sequence selected from the sequences set forth in FIGs. 5A-5M and sequences having at least, greater than or about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity thereto.

44. The Cannabis plant, plant part, tissue or cell thereof of any one of claims 38 to 43, wherein the Cannabis plant, plant part, tissue or cell thereof is a hybrid Cannabis plant, plant part, tissue or cell thereof.

45. The Cannabis plant, plant part, tissue or cell thereof of claim 44, wherein the hybrid plant, plant part, tissue or cell thereof is an asexual clone.

46. An isolated nucleic acid molecule comprising a nucleotide sequence having at least, greater than or about 75% sequence identity to the sequence of a powdery mildew resistance genetic marker.

47. The isolated nucleic acid molecule of claim 46, wherein the powdery mildew resistance genetic marker is selected from the markers set forth in FIGs. 4A-4E.

48. The isolated nucleic acid molecule of claim 46, wherein the powdery mildew resistance genetic marker has a sequence selected from the sequences set forth in FIGs. 5A-5M and sequences having at least, greater than or about 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity thereto.

49. The isolated nucleic acid molecule of any one of claims 46 to 48, wherein the nucleotide sequence has at least, greater than or about 80%, 85%, 90%, 95%, 96%, 97%, 98% or 99% sequence identity to the sequence of the powdery mildew resistance genetic marker.

50. An isolated polypeptide encoded by the isolated nucleic acid molecule as defined in any one of claims 46 to 49.

51. A cDNA molecule that codes for the isolated polypeptide of claim 50.

52. An antibody that specifically binds the isolated polypeptide of claim 50.

53. An organism, tissue or cell comprising the isolated nucleic acid molecule of any one of claims 46 to 49, the isolated polypeptide of claim 50, and/or the cDNA molecule of claim 51.

54. The organism, tissue of cell of claim 53, wherein the organism, tissue of cell is a plant, a plant tissue, or a plant cell.

55. The organism, tissue of cell of claim 53 or 54, wherein the organism, tissue of cell is a Cannabis plant, a Cannabis tissue, or a Cannabis cell.

56. The method of claim 7, wherein the genetic markers for enhanced powdery mildew resistance are selected from the genetic markers identified in Table 3.

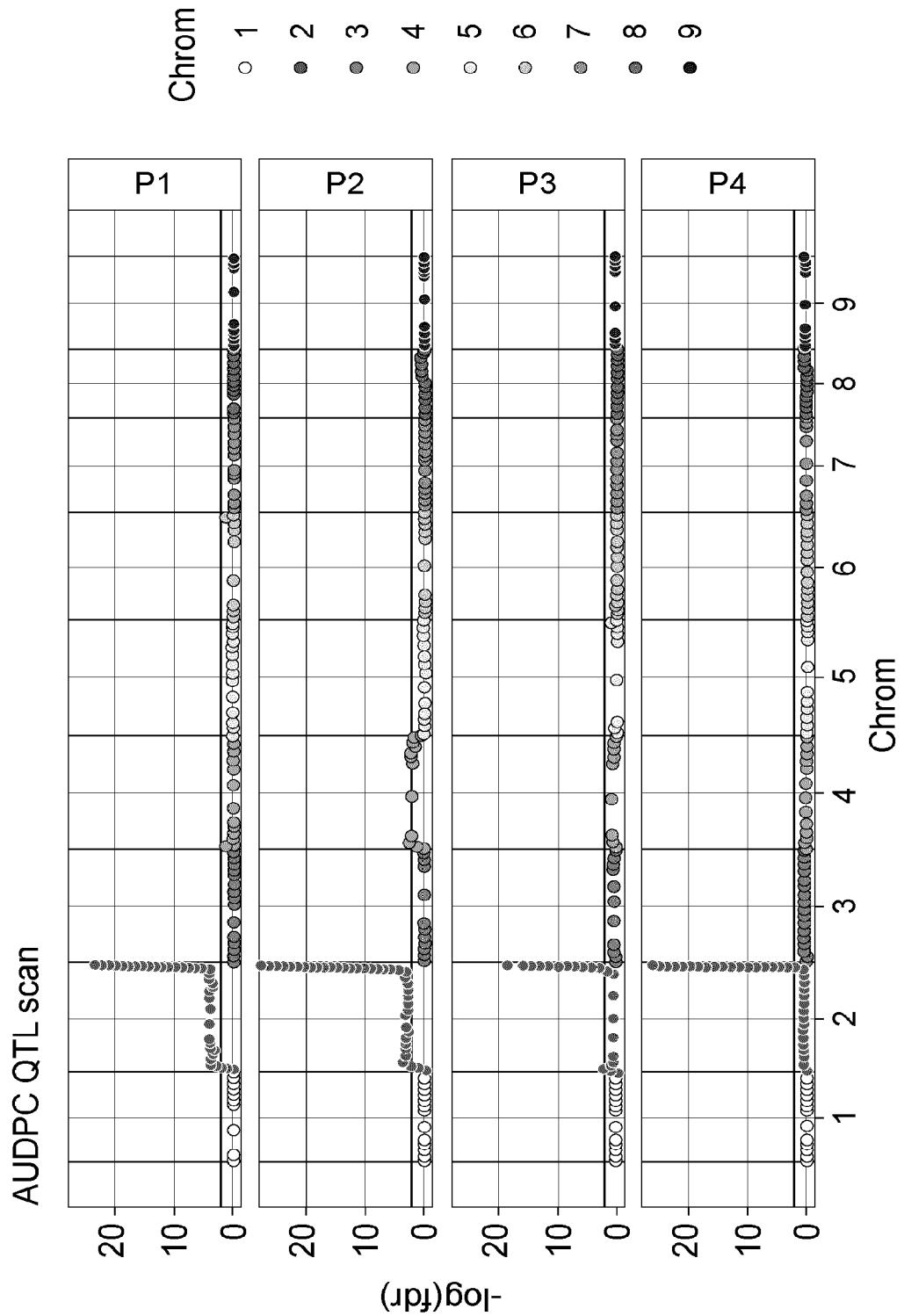
AUDPC QTL scan
Chrom
1
2
3
4
5
6
7
8
9
P1
P2
P3
P4
-log(fdr)
0
10
20
Chrom
1
2
3
4
5
6
7
8
9

FIG. 1

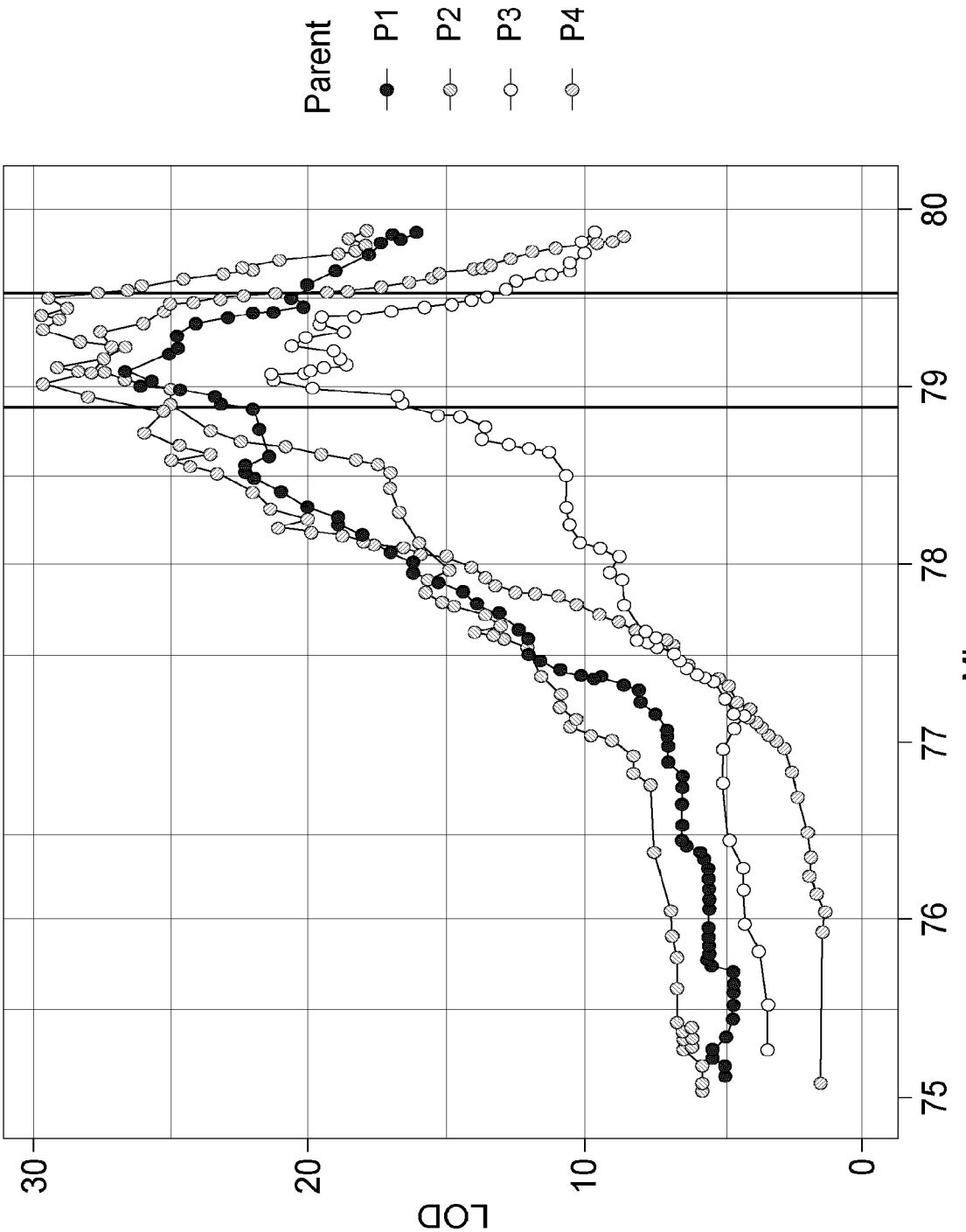

FIG. 2

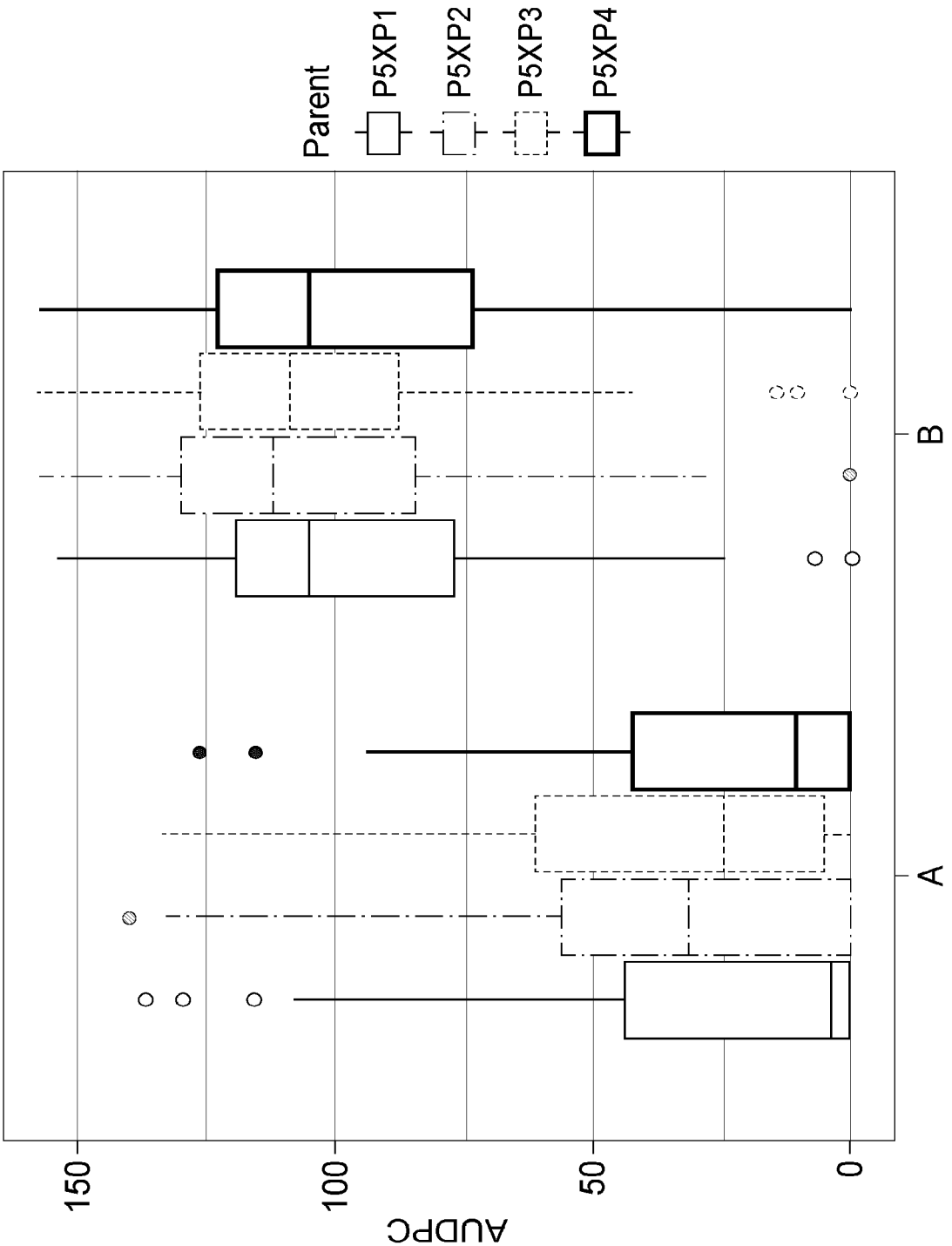

FIG. 3

Marker	A_pos	B_pos	A_allele	B_allele
1	83661138	78892192	C	A
2	83670312	78902120	A	T
3	83677748	78909556	A	T
4	83678729	78910507	G	T
5	83680451	78912282	C	G
6	83688402	78920586	C	G
7	83688797	78921175	G	C
8	83710646	78947661	T	G
9	83731248	78977665	G	T
10	83731879	78978296	A	T
11	83732009	78978426	C	G
12	83732597	78979014	C	T
13	83742301	78989694	A	T
14	83742417	78989810	A	G
15	83742707	78990100	C	A
16	83742816	78990209	T	C
17	83746443	78994471	C	G
18	83746515	78994543	G	A
19	83746792	78994820	T	C
20	83748273	78996281	T	C
21	83748951	78996959	T	G
22	83749509	78997517	T	C
23	83751001	78999074	G	T
24	83752130	79000203	G	A
25	83755671	79003640	C	T
26	83757594	79005546	T	A
27	83758836	79006788	C	T
28	83774612	79017405	T	A
29	83776153	79018946	A	G
30	83777214	79020010	T	G
31	83777886	79020687	A	C
32	83778537	79021338	G	A
33	83778851	79021652	T	A
34	83779449	79022247	T	C
35	83781145	79023944	G	A
36	83781200	79023999	C	T
37	83786804	79030483	C	T
38	83788620	79032697	T	C
39	83790907	79035044	T	C
40	83800911	79059529	C	A
41	83801699	79060317	A	C
42	83802355	79060986	T	C
43	83802547	79061178	C	T
44	83802826	79061457	T	C
45	83804625	79063241	A	C
46	83810334	79068533	A	T
47	83810542	79068741	G	A
48	83812042	79070212	T	G
49	83812462	79070633	G	A
50	83812902	79071056	A	T
51	83813059	79071213	T	A
52	83813801	79071886	C	G
53	83822038	79078975	T	G
54	83822447	79079383	T	C
55	83827350	79084284	C	T
56	83827890	79084824	G	T
57	83828180	79085114	C	T
58	83846539	79100549	C	T
59	83851294	79105413	T	C
60	83851939	79106058	C	A
61	83854472	79108866	A	G
62	83857828	79113382	T	G
63	83859465	79115014	T	C
64	83860078	79115627	C	A
65	83860155	79115704	T	C
66	83860288	79115837	C	T
67	83870181	79125227	G	A
68	83870373	79125419	G	A
69	83872374	79127420	C	T
70	83872603	79127649	C	A
71	83876864	79131862	C	T
72	83883440	79139342	A	T

To FIG. 4B

FIG. 4A

From FIG. 4A

73	83886385	79142300	C	T
74	83886777	79142695	A	G
75	83891056	79149746	T	A
76	83891581	79150271	T	C
77	83892886	79151621	T	G
78	83896020	79154720	T	C
79	83898118	79156913	C	T
80	83899274	79158067	T	A
81	83910159	79160817	T	C
82	83913362	79164021	C	G
83	83913977	79164636	T	C
84	83920656	79169549	T	C
85	83922130	79171003	C	G
86	83922204	79171077	T	C
87	83923617	79172493	T	G
88	83924935	79173783	T	A
89	83925647	79174495	A	G
90	83926127	79174975	T	C
91	83926430	79175278	C	T
92	83928911	79177759	A	G
93	83929304	79178152	G	A
94	83929791	79178639	C	T
95	83937766	79186662	C	T
96	83944003	79193148	A	G
97	83944454	79193599	A	G
98	83944962	79194107	C	T
99	83945639	79194784	A	G
100	83951447	79205344	A	C
101	83951501	79205398	A	G
102	83951922	79205819	T	C
103	83954597	79208497	T	C
104	83954651	79208551	C	T
105	83957405	79211303	G	A
106	83965250	79217619	A	C
107	83967543	79219858	G	A
108	83971930	79224280	T	G
109	83974490	79226993	C	G
110	83977660	79232177	C	A
111	83983288	79235740	G	C
112	83984445	79236737	A	T
113	83984694	79236986	A	T
114	83987265	79239553	T	A
115	83987484	79239772	A	G
116	83991563	79242411	C	A
117	83991669	79242517	A	G
118	83992602	79243453	A	G
119	83993173	79244024	G	T
120	83996141	79246966	G	T
121	83996748	79247573	A	C
122	83997051	79247876	T	A
123	83998160	79248985	T	C
124	83998359	79249184	G	A
125	84009119	79255624	A	T
126	84010480	79257003	T	A
127	84015214	79261737	T	G
128	84024725	79271249	C	T
129	84025902	79272576	A	C
130	84026493	79273161	T	A
131	84030579	79277228	T	C
132	84082978	79319521	A	T
133	84084484	79321022	T	A
134	84084912	79321450	C	A
135	84085089	79321627	A	T
136	84086727	79323258	T	C
137	84086840	79323371	T	C
138	84086904	79323435	A	G
139	84098065	79332257	A	G
140	84099435	79333648	C	T
141	84099503	79333716	T	A
142	84101266	79335284	G	A
143	84108363	79347387	C	T
144	84109102	79348129	C	T

To FIG. 4C

FIG. 4B

From FIG. 4B

145	84109586	79348613	G	A
146	84109667	79348694	C	T
147	84110163	79349192	C	T
148	84110533	79349562	T	C
149	84111161	79350187	T	C
150	84111521	79350546	C	T
151	84111790	79350815	A	T
152	84115956	79354935	G	A
153	84117120	79356096	T	C
154	84117231	79356207	T	C
155	84126498	79365502	A	G
156	84127451	79366455	C	T
157	84130728	79369735	T	A
158	84133211	79372212	C	T
159	84136245	79375239	T	A
160	84137267	79376287	T	A
161	84137471	79376491	C	G
162	84141244	79380211	G	A
163	84141949	79380925	A	C
164	84144767	79383945	A	T
165	84145357	79384535	C	A
166	84146470	79385648	T	G
167	84146584	79385762	T	A
168	84157362	79396327	A	G
169	84158537	79397527	T	G
170	84162803	79548782	T	A
171	84163350	79549353	C	T
172	84165585	79570517	C	T
173	84166530	79571483	A	T
174	84172025	79576992	C	T
175	84173494	79578497	T	A
176	84196098	79599790	A	T
177	84198707	79602387	C	T
178	84198854	79602534	G	A
179	84200808	79604474	T	C
180	84202123	79605785	G	T
181	84202739	79606402	A	G
182	84203199	79606862	T	C
183	84203715	79607378	T	A
184	84207617	79612611	T	C
185	84208220	79613217	T	G
186	84208860	79613857	G	A
187	84209525	79614520	C	T
188	84216277	79621273	T	C
189	84216884	79621880	T	G
190	84217212	79622219	C	T
191	84218881	79623910	T	C
192	84219063	79624092	C	T
193	84220392	79625400	A	G
194	84221039	79626046	T	A
195	84228838	79633113	C	A
196	84229877	79634160	T	G
197	84239862	79640657	T	C
198	84240090	79640885	C	T
199	84240369	79641164	C	T
200	84240578	79641373	C	T
201	84246889	79647365	A	G
202	84247654	79648136	G	C
203	84247867	79648349	A	G
204	84248307	79648789	C	T
205	84248964	79649446	T	C
206	84249414	79649896	A	G
207	84250079	79650561	A	T
208	84268906	79658900	A	C
209	84275859	79665933	G	A
210	84284257	79674308	A	C
211	84284689	79674740	G	A
212	84285205	79675249	T	C
213	84287775	79677983	C	G
214	84287851	79678059	G	T
215	84296530	79683154	G	T
216	84300337	79686961	A	G

To FIG. 4D

FIG. 4C

From FIG. 4C

217	84301454	79688081	C	A
218	84325518	79713361	G	A
219	83660977	78892031	T	C
220	83688285	78920469	G	C
221	83722470	78968934	C	T
222	83732730	78979147	T	A
223	83737546	78984800	T	C
224	83738491	78985745	C	T
225	83738925	78986200	T	C
226	83752526	79000593	C	T
227	83758653	79006605	G	A
228	83772365	79015173	G	A
229	83773669	79016463	A	C
230	83775690	79018483	C	T
231	83778028	79020829	G	C
232	83782315	79025216	C	T
233	83881431	79137356	G	A
234	83881561	79137486	T	A
235	83886072	79141986	G	A
236	83912253	79162911	T	C
237	83913523	79164182	A	G
238	83913677	79164336	T	G
239	83923122	79171998	A	G
240	83923211	79172087	A	T
241	83928527	79177375	A	G
242	83930070	79178918	C	T
243	83930863	79179711	T	C
244	83952313	79206210	A	G
245	83957074	79210972	G	A
246	83957237	79211135	A	G
247	83964889	79217261	C	A
248	83965085	79217454	C	T
249	83967741	79220056	A	G
250	83968215	79220531	A	T
251	83974800	79227303	C	A
252	83977507	79232024	T	C
253	83978145	79232646	A	T
254	83978987	79233486	T	G
255	83984934	79237226	A	C
256	83987023	79239311	C	T
257	83991361	79242209	T	G
258	84002990	79251640	A	G
259	84029139	79275788	G	A
260	84034792	79281270	C	T
261	84043263	79297879	A	G
262	84044008	79298624	C	T
263	84044364	79298977	T	C
264	84052299	79307175	A	T
265	84083078	79319621	C	T
266	84096494	79330692	A	G
267	84101667	79335685	C	T
268	84102485	79336494	C	T
269	84114328	79353306	A	T
270	84115189	79354168	C	T
271	84117849	79356824	G	C
272	84120052	79359046	C	T
273	84121378	79360372	C	T
274	84127188	79366192	G	A
275	84128993	79368001	C	T
276	84138203	79377223	C	T
277	84149712	79388918	T	A
278	84150685	79389891	A	G
279	84159070	79398089	G	C
280	84170529	79575496	A	C
281	84188800	79592480	A	G
282	84189391	79593068	C	T
283	84189451	79593128	T	G
284	84195913	79599605	T	A
285	84196692	79600384	T	C
286	84198130	79601811	A	G
287	84209737	79614732	T	G
288	84219469	79624503	A	G

To FIG. 4E

FIG. 4D

From FIG. 4D

289	84220293	79625301	A	G
290	84237473	79638294	T	C
291	84238521	79639327	A	T
292	84246349	79646823	C	T
293	84246575	79647049	T	C
294	84248093	79648575	G	A
295	84252075	79654201	A	G
296	84268573	79658567	T	C
297	84272997	79663086	A	G
298	84274500	79664573	T	G
299	84275357	79665431	C	A
300	84282401	79672474	T	C
301	84284044	79674095	G	C
302	84286664	79676864	G	A
303	84288029	79678237	G	A
304	84300583	79687207	A	T
305	84301370	79687997	G	C
306	84325076	79712919	T	C
307	84353662	79717276	A	C
308	83732124	78978541	C	G
309	83738265	78985519	G	A
310	83793745	79038394	C	T
311	83828271	79085205	T	G
312	83842003	79095943	G	T
313	83846813	79100773	C	G
314	83867834	79122893	T	A
315	83923016	79171892	T	C
316	83931139	79179987	A	G
317	83975300	79227800	A	G
318	83985099	79237391	A	G
319	84087004	79323535	T	C
320	84101104	79335122	G	T
321	84163272	79549275	A	C
322	84198204	79601885	T	C
323	84208440	79613437	G	T
324	84217665	79622671	C	T
325	84276972	79667046	T	C
326	84316233	79701992	C	T
327	83788948	79033025	G	A
328	84276356	79666430	C	A

FIG. 4E

Marker	Sequence with both SNP allele [A_allele/B_allele]
1	TTTAATTAGTACGATAATTAATTAAGTTATTAAATTTAAAAGACATTATT[C/A]ATTAATATATTATTTATAGATGTATAGATATAGCTAGAGTCAAGTTTGAA
2	TTTTAGGTGAGAACGTCACTTTTATTTATATAGAAGTATAGAGATAATTT[A/T]TATTCGGTATATATACACGACAATATACATTATGAGCTAATTAAAATTTC
3	GTATTAAATTTCAAACATTTAATCAAAGCATATTTAGTACTTACTTACAC[A/T]TCACTCCATCTCTTCAAACACAAAATAATCAAAAGGAAGTTGGAGCCTAAA
4	ATTCAATCCAAAAATAACTATAAATACTCATTGAAATATCTTCCTGAGTCA[G/T]CTTTTCCTTAAAGGGAAAAATACCCTTTTTCCAAAATTGGAAAGTTCATT
5	AGTGTTTATATCAACTCCCATTCTGTAAATCAATCCATGAATTTCCAGCC[C/G]TGTTTGAACTGAACCCGTCAATCCACACACGGGGAGTAATCCCGTGATGG
6	ACCCAATCACACAGTATCATTTCATTTTCTTCTTCTACCTTAGCTTCAAA[C/G]TTCCTCCTACAACAAATCAATTCCAATAACACTATACCATAGCTATACAC
7	TACTATAACCATTAATGCTTATAACTCCTTTGAAAACAGTACCAAAAGCA[G/C]CACTTCCTAGTTGCTCACTGAACTTATTGGTTGCTTTCTCTAGCTCCTCA
8	TCTACACCAACAATGTCAAAGATGGTGTCTTCCAAAAAATTTGAGATCGA[T/G]CTTTTTTAATGGGAAGAATAATTTCAATTTATGACAGAGCATGGTGAAAGA
9	TGGAGTGCAAAACTGGTTTCTGGGACATACCCGATGGCTTTAATGCGCTC[G/T]ATCAAGCTTGCAAGCAATTCATATATTTGTTGGGATCGAGGATGTCGCCT
10	CATGCATGCCATAACCAGCCATCAATGATGTCCATGAAACAAAATTCTTG[A/T]GTTTCATCTTGTTAAACACAACTTGAGCATCACCAATGTCTCCAGATTTG
11	AAAAGGTGTCGTGGAATCATGTTCATTGCGAATTACATAGGCATGAATTT[C/G]TTTACCGAATCTTAGTGCCCCTAAATGAGCACATGCCATCAAAGCACAAG
12	AAATAAGCCAATAGCATCTTCAAATCTACCAACATGAGAATACGCTGTAA[C/T]CATAGCATTCCACGATACCACATCCCTAACCTTCATTTGCCCAAAAACTT
13	CTTCTTCTCCTTCTCTTACAACATTGTCCAACTTTTTCACTGCAACAAAA[A/T]TACCATCACAAAGAACCCCTTTGAATACAACGGCGAAAGCGCCTCTTCCT
14	CCATTTGTGGCTTCTTCCAACTCTTCATAGGTGAAACTTTGTAAATTCAT[A/G]CCTTCCTTGAATTGGGTTTGTTTAGAAACTTTTGCTTTTCTATACAAGTG
15	CTTCCAACACTCACCATTTCTAAAAATAGCCACAGCACAAAAACAACATCTC[C/A]TAAACAAGCCTTCCTGCACCAATTCTCATCCACCAACCGAAAATACTCAT
16	AGAGGCCAGTCTGTGTTTTCCATGGTGTTGAAGTAGAAATTTTTCTTCATC[T/C]TTAGAATCTTTATCACAACTTTGAGCCTCAAAGTTTTGTTGGCAACCTTT
17	TAGAGACTAATGTTGCCTTTTAGCCTTTCTCCTCCTTAATTAGCTTAAAT[C/G]ATAATCAAAGATTGCATTTTCAATGCAAATATATATTTATTGAAACTT
18	AATGCAAATATATATTATTTATTGAAACTTGTAGTTATTAAGGAGAGAGATGAT[G/A]TGATGTGATTGAGAATGGTACACTATAGTATACACTCACTCCGGCTGTGC
19	AGACAACCCACCACCAATGGCGACCTCTCTTGGCGTTTAAACTTCGAAAGT[T/C]TTTAAACTTTCCAACACCCACAAAGAGAAGAAAACCCTCACCCCCTGGCCT
20	ATTCTTTGGTTTACTGCTTTCTCAATGCAAACTCCAAATCCGTATCAATA[T/C]CCTATTGGAAAAAGGCGTATGCAGCCGTTGGGTATCCTTGTGTTTGCGTC
21	CGTGCCTTTGTCCATTTAGATTACGAGTTCTCACACAAACCCGAGCATGC[T/G]CAGTCTCAGTCTTAGTGTTAAGCTGTAAAACCAGTTGTTATTATCTTGTT
22	GCCAAACGAGATAACTTAGTTGGTTTAGTATTTGCAATGATCTGGAACAT[T/C]GGTTCGAGTATGATCAGTCTTCAAAATTTTGGTCAGTCCAAAATCAAAAA
23	ATTCGAGTCCTTTGGTTCCAACTCGGTTTAGGACCCCGAAATAGAAAACT[G/T]GCTAAGGTTCCATTGCTCATGTACTCATACACCAGAAGCCTGTGCTTTCT
24	AATTTCCTTTTGCATCAAAGTTGATTTGATAATCATGACCACCAATATTA[G/A]AGCCTATTTTATGGTATGTAACAGGAGTTTCAAAATTGAACAAAGATGTG
25	GCTTTGAGTTTTGAGTGAGTGATATACTTACTACTTTACAATCTACAAAA[C/T]AGAAAAACTCACTATATTCCAAAGGCATATAGATACATCTTATATTCTAT

To FIG. 5B

FIG. 5A

From FIG. 5A

26	CAGTTCAAGAAATCAATGCTATCAATATAAAAGTTTTTGTAATCTTCTAA[T/A]TCATCATTGCAACTTTGAGCTTCAAAAATTGGTATGCACCCTTTCATCTT
27	AACTTGACACACAGTTTAATCTAATTTCTTCTTCATTTTT[C/T]TATTGGAGGGGGGGTTGGGTTAAAAATTGACTTGGTCTATAATAATT
28	GCTTTGAGGGTCCAAAAAGAAAGCTTGCTAAGGATCCATTGCTCATGTAC[T/A]CATAAACAAGAAGCCGGTGTTGTCCCTCGTTGCAGAACCCAATTAGCTTT
29	TCTCTCAAAACTAATGAGCCAAGTTTGGTTAGTTCAACAACTGAGCCTTG[A/G]TGAACAAGATTGTCTCCATTGGCTGACCAAACAATGGTTTTTTTCGGGTAT
30	CAAATCATAATTTTGAACCAAACCAATCTATTCTTGAGTGGGATTTATTTG[T/G]TTTCCTTTGATCATTTTAGTTTGGAATTTCAAAAAACCTTAATGGTTTGG
31	TTCGGACAAGCTTTCTTTCAAAGCGTTGGAAGCTCATGTGGTATTCAGTT[A/C]CCACTCTCTATTTCCCTGATTTAAATTTAACTGCCAATTATGTGGACAAT
32	TGTACAAGGCAATCACAAATCTCACTCTAAATTGTGATTGGCATACGAAC[G/A]ACTCGTCCTCAGTAGAATATGTTATTTCAAACCTTCCTCTTCTTGAGAAA
33	ATTCACGTGCCAGAAAACTATGACGGAAATTGGTTTGTTAGTTTGATGAA[T/A]TTTTTTTCTGAATCTCAATTGTCCATGGAACACAGTAAGCTTGCATGTTGG
34	GTAAGTAACATCAACCACTCTCTACTCTTTTATGGATCATGGGAAATTAA[T/C]AGAACTTGCGAAAGAAAAGAAGGAAATTAATGCGTGTTTGTTTGTCTGGT
35	ATTTGAAAAGAGTTAGGTTAATTTATAATCTGTCATTCACTTCAAGAGAA[G/A]GATAGCAATTAGACATTAACATTGGTAACTTATCAACAAGATTCGTCTCC
36	GCAATTAGACATTAACATTGGTAACTTATCAACAAGATTCGTCTCCCTAT[C/T]TTCTTATCTTGATTAATCTTCTTTTGTTATTTTTTAATTTTCTGAATTGT
37	CCTGTAATTAAAGGGTCAGAGCGTTATAATTGTCAAATTCTTTGTATGAA[C/T]CGAAAAATATTGATGATACGCCCAAAGATGTGGCTTGCAAATGATGGGAA
38	GCTTCTGGCTTTTTGCTTCCTCTATCTGGTGGAGCAGATAGCTCATCTGT[T/C]GCTGCTATAGTTGGCTGTATGTGCCAGCTTGTTGTCAAAGGTTTGTTTTC
39	AAATTCATTCCTTGATAAATAAGCTAACCAACACTAGCTTCATAAGATAG[T/C]ATATTGATATTTTTGTCTCTTAATTTTCGTCTATATGGAAGAATAATGTA
40	GAAGATGGAATCGTTTCGAATTTCCAATGAGGAATACTTAATAACCTTAG[C/A]TGGATTAGATTCCCAATTTCTTTAGGCAAATTCAAAGTAGATGTGGCATA
41	AATGATGGCTAATGGCAGACCAGAGCACTTTCTAAGCATCTCTACCGCTA[A/C]TTCTTTCTTCCTTTCTTCATCATCATTTGAGTCTGAATTTATATATCCAG
42	AAAGCTCCTTGGCTAATTCATCATCGCTTATCATTTTTGATTTCTTCTCTTT[T/C]TTTCTTTGGTGGGAGAAGTGAAAGCGAAGTAAATTGCTTCGAAGACATTG
43	TAGTTTTGCCCAAACCACCCATCCCACATACAGAGATCACCTTATGCTTA[C/T]GAGAATTCTCTTTTTCAGTCAAAAGGGCAACCAACTCCTCAATATCTTTA
44	ATCCAACTTTGTGGACATTAATTCCCTCTCTGAAGATGTTAATAGATCTTT[T/C]TGAGCACACCTTCATTCTTCTTCAAAGTCACTTTGAGGACATAAGTCTCA
45	ACAAATTTTGATTCATTCAATTACACAACACATCAACAGAAGAACACAAAAA[A/C]CAAAAGATAGGAAAAAGAGTCACTAATGTAATTATTGAACAAACACACAC
46	TTAAGTCAGGCACAAGTTTCACATGAGATATTGAAGGAGCCATGGATTAA[A/T]ATTCAAACCACCATACTAAAAATAAATGGAAAAAATATTACATTTTTCAT
47	TTTATTTTCCCCACATATATATAATGTGTTTGTTTAATATCATAAATCTTT[G/A]ATTCTTTACATATTTTAATCATCTATCTTTTTATTTGATATTTTATTTAAT
48	TAAACAAACACAATTGATTTTACAAGCACAATTACATAATCAAAAGCAAT[T/G]CACAACACACACATTAAACATGAACATAAACCAGAAAGTTAGAGAGACAT
49	GATTTCATAATTAAGAACCCAAGCAATAGTCGCCTAGAATGTGTTTCTAG[G/A]ATTAGATGAAGCTTGTAAATCGGGTGTTCTTGCTTTACTTTCTGTTTTAA
50	TAATAACTACATATATATATTGAAGATATTGTTTTTGATAATTGAAAATT[A/T]TAATTACACAATAATATATAAGAGAAAAAAAAAAACAAATAAAGTAAATTG

To FIG. 5C

FIG. 5B

From FIG. 5B

51	ATTACTCATGTCAAATTATGGTTCTTTGACACCATTTAGGATATGTAATT[T/A]AGTTTAAATCCAAAACAAATAATATAATTACAAAAGATAGCTAATTGATG
52	AATTCGTAATTTCTTGAGACTCTCCAACTGTAGAAGATCACTCAGTAGAA[C/G]ATCCGTGGTACAAATAGATACCAATGTTTGTAAATTTCTAGACTTAGTTG
53	AACACTACGGTATTTTGAGTCAACGTCATGTTGTTCAGCTTCACCTA[T/G]GCAGCGAATTACATTTGCTTTCAGTTGCTCCCAATCATATAAGGTGTGTT
54	AAAGCTAAACATATATCTGTTTAACACACTAGGAACATATACAAAGCCAA[T/C]GAATTTGACATTATAAATATAAATGGTAGCAAAGATTAAAAAAAATAAAG
55	TCGTTGGAGCTTAAAAAGATGGAGTTAGAAATCGAGTCAAACATAGTAAA[C/T]GAAGCTCGGACTATGTTGACTAGTTCAATTGAGAACTCAGTTGAGGATGA
56	CTTTCTGATGAAGTTCAAATGCTCCATCAAGAGGTTTCGTCTTCAAGATC[G/T]TCCCTTCAGGAACTATCTACAAGAAAATCTAGCATCCAACAGGATATAGC
57	CGAGAAAGAAATTAATGACACTGTTGAGAAACTGGAGGAGACAGAGGTGG[C/T]CATTTCCTCCAAGGAAAAAGAGTTAGCCATGGCCAGATTTGAGAGGCTGG
58	AATGAGACAGGTAATAATGTGTGCTTATTTAACGAAAATTGTTGTTTTAA[C/T]GGTTTTTTTTTATTTAAATATTTAAATTAATAATTAGGTAATATTGTAATGT
59	AACACTTGCATAGCTTCACCTAATTCGCCATATTTACCGAATGTATCGAT[T/C]ATACTATTGTAAATTTGTCTGTCTAATGACATTCCCATATCCTTAATCTC
60	CTTTCAGTATTGAATTTCACTTCAAGACCGAGTCTTTTCACACCTTTGCT[C/A]TGCCTTGAGTGGCTCAGACAACTTGAAACGAGACTAGTTCTTCTGGGGAT
61	AATGAGCAAAACTACATCGCACCAGCACATGGAGAGCAGTTTTAGGGGCA[A/G]CAGAGCAAAAGGGACATAAAGTATCAACTTGAACATGTTTAGAAGACAAA
62	ACCCTTGGCCAAAATCTTATGAAGCCCGTGGATTTCAATAATGTCATTTT[T/G]AGTTAAAGCTTATCTTTTCTTAAAAATACCATTTTTCATTCTAAGCTTTTT
63	CTTAGCCACTTTCTTTGAATTTCTACTTTTGTTGATTGTTCATGGCCATC[T/C]CTTCTAACATCTCATAAGCGTCATTAGTACTCTTTTTCATGAATGCCTCA
64	GATCCCTGTCATTGTGGGTAGCACATACTCATGCAAACTTCTATTTTGAC[C/A]ATTAGCCTGATCTACTACACACCCCATTGTTATCATTTTGAGCCATATCTTT
65	GTTATCATTTTGAGCCATATCTTTTTCTAACTCAACATCTCTAATAGCAG[T/C]TCTTTCCAACATCTTCTTTCCTTTATTCTGTCGACAAGCTTCTCAATTTC
66	GACGTTGCATAAACTGCTTGATCCCGAGAAAAGAGAAAGATTAACACACA[C/T]AGAGTTAGTAAAAAGTTGCAAAGAATGGCCCTGGCTTGGATGTGCTCAAC
67	TATTAAGCCAAACCAAGTTCGAATCCAACCTCTCCCAATATCGAAAAAAAC[G/A]AAAAAAAAAAGGAAAGAAAATAGTTGATTTTTTTTGAACAATTTACATGTTT
68	AGATTGTTGGATTCAAGGATTCTTGCCCGATTCCGTTCAATTTTACTAAC[G/A]TGGTGATTGTCCATGTATTAAATTATATCTCCAAATTTTAATAAGTACCA
69	ACATCATCTCATTTTCAAATCATACAAAACAATTCAAAAAACAAAAAAAAT[C/T]CCAACAATGGCTTCTCACTTTTTGAAGGCTGTCTTTGCATTTTGTACATT
70	GTTATTTATTTATTGAAAATGGTCATGTTCAATTGTGATTTTTGTGCAGAT[C/A]AAGTTACAAAATGTGACAAGGTTGTGTAAAACAAAGAGTATTTTTGACTGT
71	TCTAACTCTTATTTTAATAAAGATTCTAAACAGAGATAGTACTGATGAATT[C/T]TGGAGAAAAAATTACAGTTTTACAAATATTGGGTACTATGTCGTTAAAAAA
72	GAGCCATTTTTTGCTATAATAATATTTTATGAGATTTTTTGGTTACTGGGAT[A/T]ACTAGAGTAGCATAACATGAGTAGTGATATGAGATTTTCTACTTTTCTAG
73	TCAAAGATGCTGACATTGACAGCGGGTGGTCTTATACAAGACTGAACCCC[C/T]GTCATAATAGGTAGCATATAATCCCCCAGACTTCTGGCAGGTAGATTCTT
74	ATAAATAAATCACAAAAAATACTAAATGTCAGATTATCATTACAGTTAAA[A/G]ATACTGTAATTGACTGAAAAAGAGCTTGTTTCAAATTGTAATGGTAATGG
75	AACTTTTTTTTTTTTTGTTTTTGTTTATGTTTTTGTGTTTTTTTTTTTCAA[T/A]TACATTGTCTGGCGTACATACATTAAATTTACACAATATGGTGAATTCAA

To FIG. 5D

FIG. 5C

From FIG. 5C

76	AGCTTTGGAGTGGTGTTACTTGAGATTGTTTGTTGTAGAAGAAATGTGGA[T/C]ATGGAGATTGGTGAGGAAGAGAAAGAGATCTTGACAGATTGGGCTTATGA
77	ATTTTATCACTCTTAGTCCATACTCTCAAAGTCTGCAGGCATCTTAAATT[T/G]CCAAAAGAAGATGGAATCTTTTCAATCTGCCAATAAGGAATACTTAATAA
78	TTGGTCATAAATAGCTTGTTAAGCTTTTTCTAACTAACTCGTCATGTTTG[T/C]CAGAATCACTACTTTGATTCTTGCTCTGCTACATTTTTCTTTCTCCTTTG
79	ATCATAATCAGAACTTCCATTAACATAGATGGCAATTCTTCTAACATTTG[C/T]TTCTGTTGGTTCTTCTCCTTTATTCCTCCAATCAATAAAATGCATAAAGT
80	GGCGTGTATTACATTGCTGAGAAATTGAGGCCCAAGCATGACAATCAAAG[T/A]GTTCCCTCACATCAGGATGCTGATAGACCTTTCTTGCAAGAGTAGTCTTC
81	GTTTGATAGATACTTTCTCTTATATTTTCTCAATATTTCATACTCTAGAC[T/C]GATTGCATAACAAAGTTGTTTTCAATTATTTATGGATAATACAAAAATAA
82	CGAATAAGACACACATAGACAAACTCAGTTGACTTGACCCAAAGTCTTCC[C/G]AAGAAAAATATAATAAATATTAATTTGACTCTGGGCTTCTTTCAATCACCC
83	GCATACAACACTTAGATCATCCACTGTGACATAAAAGGAATAAAAGGGTA[T/C]ATAGTTCTGGTGTGGTACTGCTTGGAAATTGTTTGTTGTCAAAGAAACTT
84	TTGAGGAATATAAACGAGAAGATCCAATAAAAGAGTTGACATCTTTCATGA[T/C]AATGATTATGAAGTAACTTTGCATTAAAATATTCAGTTTTTTAATAATATA
85	GATAAAGGAAATCAATAGACAAGGCAGATGAAGCTTCTTCAAGCTATGTC[C/G]AAAAGCAAAGACAGTTGAGGCAAGCTTATTCTTATGTTGAAGACAATCAT
86	GCTTATTCTTATGTTGAAGACAATCATGTTGTTGGATTCGATAAAGATAT[T/C]CAAGAGTTGGTTGCCCTTTTAACTGAAAAAGAGAACCCTCGTAAGCATAA
87	CCTTCTCCATTCATGATCTCATGAGAGAATTGTGCGTGTCTAAGGCTCAA[T/G]AAGAAAACTTTCTACATTCAATTGATTGTAGGAATCAATGGGAAGAACCA
88	ATACTAAATATTATCGAAATATATATTGAAAAACTCACTAGAGTGCAATA[T/A]TATCACAATAAACAACTCTTATACCAGATAATAAGATGGTGAGGGACATG
89	AACTCATACACCAACAAATGATTCTTCCCATCACAATATCCAAGTAG[A/G]CGAACGAGATTCTTGTGATGTGTTTGGCCAATGGATTTCAATTCAGCCTT
90	CCTTTTCGAATTTTGATGAAAGCAATAGATGGACGGTTGTGATCAACTCTT[T/C]CCATTTGATAAGGGAAACCTCTTCTTCCAACAAGTTCCTCCTCCTAAGAAC
91	GGATTGATTAGAGAGTACCCTTTTAAGCACTCACACTTTGGTCTCTTCTC[C/T]TTATTGAGTCTGCAAATGGTATTGTATCCACAAACTCCAATTCCTTCCCA
92	TCACAATGTATGATCTGAGTGTTGCACTCCTCGTGTAAATATAAAAGCCC[A/G]TTTGCAATTCCTAAAGCAAGTTCACATCTTTGAGTCCAACTAGGCTTTAT
93	TCCTTGTAGGTAAAGCAACGCAAATTTGACTCGGGAACTTCTTGTGGAGC[G/A]AAAGACCTTTTATTCCTCTTCTTGTAAATGATGAAAAACCCCATAGACAT
94	CTTTGCAGCCTTGAATGAAATCTGGTTTACAACTTCCATAACTGTCATTT[C/T]GGTTGAGTAGAGAGTACCCTTCTAAGCATTCACACTTCGGTCTCTTATCT
95	GTAACTCCTTTATTTGCAATCCCACATGGCACTATGTGCTTAAAATAGTT[C/T]AAATCTGGATCAATATTAAATGCCAACCCATGAGAGGTGACTGCATACGA
96	ACCTCAACTAAAGAATCAATAAGCCACTAAAGAATTACTAAGCCACTAAC[A/G]TTGAAATATATTAGACTTCCAAACTAAATGAAGAGTAAATAAACAAATCA
97	ATTGTAAAAGCAACTTCTTTTCAAACCCAGGAAAAGAAACTGGGAAAAAA[A/G]GCCAAGAATTTTCTTATAAATATCGTGGGGAGAAACACAAACAAGTCAAG
98	AGTGGAACCAACAACTCCTCTTCTGGGATGTTAACATCAGCTTTGATGTT[C/T]GATAGTAACCGAAAGAAATTTGTGTAATCGACTTTGTCAACAGCCATATT
99	AAGGGCACACTAGGATAACCCAACAAGAAAATTTCAAGATACAAACAAGT[A/G]AAATGGCGAAAAAGTCAGCTAACTTATGTTTCAACTTTTGTCGTAATACT
100	CCATTGTATATTAGTTTGACATGAATGGTTTTTGTTTAGCATAGATCTGTTG[A/C]CTTTTAATAGTTCTGTTCATGCCACTTGTTTTTCAGAGTTTAGAATGTTAGT

To FIG. 5E

FIG. 5D

From FIG. 5D

101	TAATAGTTCTGTTCATGCCACTTGTTTTCAGAGTTTAGAATGTTAGTATA[A/G]GTTTTAAGTTCACAGTTTAGTGTGGAGATATTAGAAGGTATTCAAGTTCT
102	AAGCTCAATCAGATGCCATCACTTGGGTAGTTGTTTGACTTTTCATTTTC[T/C]TATTTATGATGTTTTCTTGACAATAGGTATGAATGTTTAGTACTTTGTCT
103	TTTTTCTTTTCTTTTCTTCTTTTTTGCTCTGGAAACAGATGAGAGAGATCA[T/C]CATTTTTCTCAAGGACAAACAATGGATCACGAACGGGGATGGAGGATGGG
104	TTTTCTCAAGGACAAACAATGGATCACGAACGGGGATGGGTGA[C/T]GGGCGAGCAATGGAAGGATCCAACGGCGATAAAATGGCGAGGGAAGGAAG
105	ATAGAAAAGGGGTGAACAAGAACAACAAGGCTATCATTATCTTGGTG[G/A]TCTTTGTTGATCTCTCTTTGGGGTTAAAAACCTTTGCTTGGAGTTTTAC
106	CTTTCAACACTAAGCTCCTTAGCTCCTGAGCCGCGGCATATTCCGGCGAA[A/C]TTTTCCGGTAGAAGACGACGGCGTTGGTAAATATAAGAAGAAGGTCTCTA
107	AACACGACTTAGATCTTAACCAAAAAACCAAAACTCCCGACTCAGATCTC[G/A]AACAAGAACAAAGAGATTTACACGATCGAAACGTCGCGTCGCCGAACCTC
108	TTTAAGCCTCCCCACCAACCCCAACATCCCCCTCACTATTATCTCTCCCTC[T/G]TTTATTGGTAGTTACCTCCGTCTTGGCCGCGGCCACCACCGTCGCCTTCT
109	CGAAGCCGTCGATGGAAAACAGAACCGAACCGGTGAGATCGAACGGTCCC[C/G]AGAAATTTAACCAACGGCTCAGATTTAAGACCCGAGACTCGATTTTTGGT
110	TTTGTTTCAGTTTTTTGTTCAAAGTTTGAATCTTTATTATCTGTTTGGTTG[C/A]TGAGAAAAATAAGGGGTTTTTATTTTCTTGGTGGGTTTTCAGTTCAAAAGC
111	CACACTGCAAAAAGATTTTAAATCACATTTTTTTGTACGATATCGTGCGATG[G/C]AGATAATTTATATAAACAATTTAATCAACAATCCTAGAAGAAAAGAGCCTA
112	GTTTCCAACCTCATCTGTTTCGAAAAAGCATTTTGTGTAATTCAGTAAGT[A/T]ATTCGCTAAAGGAAGAGAAAAAAGTTTCCGAGTTAAAGTGTTTGGAAGGT
113	TTTGTATAGTTCTTCAAACACAACAATATCAGGTTCATTTGTTATAGCAT[A/T]ATATTAGACAAATTATAAAGAATTAAGTCACTAATAAGCGATGTTAAATA
114	AATTTAAGAGAAGAAGAAGTAAAGGAACGAGAAGAAAACGATGAAAATGTG[T/A]GAAGAACTTGAAAGCTTAGACCATGCAGCTCCTTTGAATGTGGCTTTTCG
115	GAGATAAAGTTTGAATTTTTACTTACAAATTGAGGAGAGGAAGTAGTAAT[A/G]AGGAGATGAAGAAAAATGTTAAATTGATTTGGGTTTCTTCTCTTTCT
116	ATCAAATGTGTCAGTTCAAAAGTTGCAAATAAATAACAGTGAGAAATGGA[C/A]CGATTTGGAAGAAAAAAAAAAATAAGAAAATCTATGTAATTTGTATCATCT
117	ATCTGCCAAAAACAAGAAAATTGAAACTATGAATCTGTTAGATATCATGA[A/G]GTCCTTTTATTTGAAGAGTTTATAATCTTAATCCATATAAATAAATAGGG
118	GGTCGGGATGAAAAGCAGGATTCTATAATAGGCCACTACGCTTTTTACTA[A/G]AAAATAAGTTTCTGCATTTCATTCAACTGCATTAAAATTTAGATATATTC
119	GGTTTTTGAGATTTAATCTTCCTCTCTCATTCACTCCATTTCATCAGTAT[G/T]TTATACCAATTCAAAGTTCTATCTTCACTCTAAAATCAAGCCAATTTTCA
120	CTAAAATAATCCAAAGTTCATACATGGTTATTCTCACACACTCACATAAC[G/T]AAGATAGCATTTCAAATATAATAGATTTGTATTCTCTGTAGATAGATGTG
121	CGTAATAATCCCTTCAACATCTTAAAAATATCTGATTTATCGTAATAGTA[A/C]ATACTTTAAACTCTTTCGAAAACTCGATTTTGATCTTCTGATTCTTCAAC
122	AAACTACTCACATAATTCCCAAGAGATAGAGGCCATACAAAGTGAGCT[T/A]CCGAAGCTCTTAATCCCATCTGGGGATTGTTCATTAAAAAATTCAAGTTG
123	TTGAGGACCTTGAACCTCATACAACTCCTCTTCACGAACTTTCTCGAGAT[T/C]CCACTTCTTAGCAGCTGCAACAAATACTTGCGCGACACGTGGAATAGGGT
124	CATTTACCAATATCTTCAAATAGACCAAAATAGTGTTTGAGAAAAGTGA[G/A]CCAAGATTTAATGAGAAGTAAAAACAACAAAAGTATGAAGCTTTGTTGTT
125	TTGTGGTCTAAATTCAAATTAATAACAAAATAAAATAGATGAAATTATTG[A/T]ATATATATATATATATGCTAAATAATGTTGACAGCTAATTTCTACCCT

To FIG. 5F

FIG. 5E

From FIG. 5E

126	CAAAATCTTAAAATAAGTTGTTTGAAGCATGCTATATATACACGTGAAAT[T/A]AGTTATTTTAAGTTTAATTAAAGTACCTTAAATTATTAAAAAAATTAAAA
127	AATATCTTCAATCCCGCTTTCGGTGATAGTTTCCACAATATTGTTATTAT[T/G]TGTGGTATTATGGCAACTATTGGTGTCATTGTTGGTCATATTCAAATTAG
128	CCGGCTAGCCAAGGTGGCAATCGGATTCCTGTAAAAGGGTACAAGATTCC[C/T]GCCGCAATTGGCATCCCGAGGATGTTGTAGCCGAGGGCCCAAACATAATT
129	TGTTCACTGTTCGCCTGCATAGGGAAAAAGACATCTTACAAAAGAAACTC[A/C]AAAAATTCGGGCTGCAGAGATTTTCTTGTAGTAATTCAAATCTCTAACCA
130	GTTGGTGTTGCTAGACCGAGAGCACACGGGCAAGCAACCACCAAAACTGA[T/A]ATACCGAATTGCAGCGCAAGCTCAAATTCATCCATGTCTTTTGGTATCCA
131	TAATTCACAGAACAAAAAAGAAAGTGAGAAAAATACAATGAAAGTTAAGT[T/C]TTTGAAACAAACCCATGTCAAATTCAGCAGAGAAAGAACAAACCCAGTTG
132	CTTCTTCCATTACCAACCAATGATGTGCACATCTTTCTTAACCATTCTTC[A/T]TTCTTTTCTTCCTCATCTTCCTATATATTATGATAATACAATATATGT
133	ATGCATTGTATCTTCTAATCCTCCATTTTCATTCTCCAATCCACTTAAAA[T/A]TTCTATAACATCCTTTTTCCTGCAATAACAAAGAAATTCACTTAAAAAAA
134	TAAGAAACTGCAAACCGCCAATTGAGATAAGGCGTTTAGCATTGTTGGTT[C/A]GCTCCTCTTCGCCAAAATTAGTAAAAAGTTGCTCTAAGATCAAAAGGGTT
135	TTTTCCAGCTTTCTATTACTCGCTTTAGTATTAAGTTTGTTGAAGGGACT[A/T]CATTGCATTCCAATACTTTTCCTGTTGTTGGGCATGTTTTGTTTCCTTGA
136	TCGAGAAAATCAAGAATAGACAAACAACTCCCGCGAATCATTACTCGAAA[T/C]TCCTTTTCTTTAAGAAACCTTTTGATGTATCCACTTAAGATTGACACCAT
137	AACTTCATCAAGAGCAACCTTCTCATAACTCAACTTCTTACTTTGAACTT[T/C]TTCAAACACTTTTTTTCCCTTAACTTCATCCGAAAATATGTCATTTTCTT
138	TTTCCCTTAACTTCATCCGAAAATATGTCATTTTCTTCATCTTCATCTTC[A/G]TCTTGAACCTCAACAATCTCATTATTTGAATGTGAAGTGATTGATTCTCT
139	CGAGAGGCCAAAGAGTGTCTTTACGTTGAACCAGATGATGAATCTTCGAA[A/G]AAAGGTGTTCTTTGGTACAAGAACAAACTTTGATCAAATCGGCAGTGGAA
140	ATAACTAAAGGTGGCACCGCCTATAGAAGGTGTTGGACACCACTAATACC[C/T]AATAACAATGCTCTTATTATTACTTCCTAACGTACCATAATGTTTGCATT
141	TATTACTTCCTAACGTACCATAATGTTTGCATTACCGAAAGAAAAAGAAC[T/A]GGGACATTATCGTCCAAAGAAAGTTTTTATTGATTTTGAACTGTCTCCTC
142	AAAATAGCACATTTAATATACTAGACAAGTATGGTGCTTGTTTGTTGATT[G/A]AATTTAAAGGTTTTTGAAAAGGTAGAAGATGCCATTAAAATCTAGATTGAG
143	CCTATTCGGATTTTCAGCTACGGTAGTTAAGCGGCCAGTTTAGTGTGATG[C/T]TTTAGCCGAAGGGTAAAACAGTATTTTCATAAACCGAAAAACGGACCTCT
144	TTGTCTTTTTTCAGCTTCTGAACTTTGGTTTTTTGTTGTTGAAATGCTTA[C/T]AGGGAAGATTTTGATTGGAGTTTGAGCAATAATGACTTCTAGAGGGAGGT
145	TTTATATTTTTTGGTCATGTACACACAACAAATCATCTTGGCAAATGGTTTA[G/A]TATGGAAATGCCAGTTAATGTATCATGAGTTTAGGCTTAGTTTGTAAATG
146	TTAGGCTTAGTTTGTAAATGCATCTTCTTTCGCTCTGTGCCTGCTTAAGG[C/T]CTTTAAGTATCTCCTTTCAGCGTACTTCTATATTGTCTTCCTTGTTTCTA
147	GCTTTGCAAAATGCTGCAGATAAAGCTGACCGGTGGGTTATATTTTATGT[C/T]TGGGACGATCACTTTGGTTTTTGCATTAGTTGACAATCAATCTTTTGCTG
148	AAGACTCAGATTCTAGTTCCATATAATATTCTTCCACTAGATCCTGATAG[T/C]GCAAATCAGGCAATTATGAGATATCCCGAGGTATGTTTGTAACATTCAGG
149	TAATATTTTCTGTTGTTAGGAGTCTTTTACCAGCCATATTCAATCTATTG[T/C]TTCCTAATTATGAGTATGTAGAACTTTTAGACTACTGACATTGCTGACAA
150	CTCTTCAAGAACTACAAAAAAATGGTGCAAGTACTTGGGTCGTAAAAGTAG[C/T]CTTTGGTGAGTGCTTTTTGTGCTGAGCTATTTGTCATTAGATTTTGAATG

To FIG. 5G

FIG. 5F

From FIG. 5F

151	CAAGAAGTTCAGCAGCGCAAACTTCTATACATGGGACTGTATCTTTTAAT[A/T]TGGGGGGGAGGCTGCAAATTTAAGATTTATGCCCGAATGCCTTTGTTATAT
152	TCTTTTTTTGATGCATATGGGAATTGAATGAACACCTTCCCCTAACTCGC[G/A]CATACCCCTCCTCTCATACCATTTGAGTTTACCTTTGTTAGTTCCAATTT
153	AGAGTGTGCGAGGGCCATCCTATATTGAGATCATTCGCATTATTTACTAA[T/C]CTCCTTGATTAACTTTCCAGTGTTGTTGAGCATATATTCAATGTGGTTGA
154	GCAGATGGTGATCTACTTACTGAATTCAAGATGAGTTCCCTCCCTACACT[T/C]TATGGGCACTTTGTTAAGCTTATTGAATTTATGGTAAGTTTTTGAATGTCC
155	TTGTGGCGGAAGAATATCGTGATTACAAGATTTTAGCATTTTAAAACACT[A/G]AAACATACATCAACCAAAAATAACTCATAGGAAGCAATTTTTTAGTCTCT
156	TTGGTTTGTTTTCTTATGCAGCTGAGACGTGTGATGGAAACTGATGCAGC[C/T]ATGACGGACAACTTAATTGCTTATAATATTATTCCACTTGATGCTCCTAG
157	TTGTACAACATTTACATGGGCACTGAAAGTTCATTTTAATTTCTTGGTGA[T/A]GTGATGATGGAAAAATTTCTTATCTGTTGCTGAGTTCCACAAAACAGATG
158	TTTGTTTTTTTATATGTTTGACCATAACCTTACACCTTCAAAATTTCTGT[C/T]GTCATCTCCAGATTAGATCCTTGGAATTAGTTCACAAACTTTTTGAAGAG
159	ATTATACCTCAACTGTTGTGGATTTTTAATTCGAGGGTTTCAATATATTTT[T/A]TCTTTTGTAACACCAATCTCTCCCTTAATGATGTTTCTTGTTTCCTCTGA
160	GTTAGAAATGAAGCCCTTCGTGTTGCTTTTATTGATGAGGTTGAAACACT[T/A]AAGGATGATGGTAAAGTACAGAGGGAATATTTCTCAAAGCTCGTTAAGGG
161	TGCTTTTCTTTCCGTTCCCAAGTTCTCCCTTTCTGGTAGATTACTTAGTG[C/G]TATATGGTATATAACATATAGGGATTTATAGGCTGAAAAAGTCTAAATTA
162	ACTATTCTTAGTCTAAGGTTCTTTATTTTCCAGTTTGGTGTTGTCTACAA[G/A]CTAGATTTGCAAGGAAACAATACATCGCTTTCGGTATTACCTCTATTCAT
163	ATCCCGATTGCATTCTTTTCTTGGTTCCCATTCGTGTCAACATTCCAAAC[A/C]AGACTCATGTTCAACCAGGCTTTCAGCCGAGGTTTGGAGATCTCTCTCAT
164	CAAAAATAAATAAGATTTTAATTTAAATTCAAAAAAGAAAAAAAAA[A/T]TAAATCCTCTAAAATCTTCTTTTGAAATCCATAACTATTCGGCTTCCACG
165	TCCTCTTCGCCATGGCTATAATAAGGTTTGGTCTTCTCAACAACAACAAC[C/A]ACAAAACCAGCAAAATCAACAGAAGCAAGTTTCAGAATTCATCGA
166	GGAGAATGGTACTTGTTCGGAGAAGAAGAGCATTGGTGAAGAGGATTAATT[T/G]TAAGTCTGCATCTTCTGAACAGATAACCCAATTGGATTTCGGAATATTTT
167	TGGACATAAAATGATGATGATGCCTTTTGTTATGTAAATGAGGATAT[T/A]ACATTTTCGTAAGAACTGGCTTTGTTCTCCAAATTCTACCTTAGCAATAA
168	ACCACCAAGAGTTTGGTGGAGAGCTACAACAATAATCAAACTCTTTTCTT[A/G]GTTGATTTTGTCAAAGCTATGATTAAGATGGGTAATATTAGTCCTCTTAC
169	CTTTTGTTTTTTGTTTTTTAAGAGAAACGACACAAACATTAGAAAACGACAC[T/G]TAAAAACAGAGGTGGGTTACAAAATGATTATTGAGAAACTGAACTGTGCT
170	TGTCTTCCTTCTTCTGGGTTCTTTTTCTTGTTTGAGTTCACACCCTTGTT[T/A]ATATTATGCATAATTTTACTCTGTAAGGCTAAGAAACCAGAGTAGAAAAG
171	GACGATCATCGCTTGATCACATACCCATTGCTTACTATTATCCATGGTGA[C/T]GATGATGATAGTAAGAAGGATCATTTAAGAACTGCCACAGCGGATCT
172	ATGTTCACCCTAGCCGGTGGTGGAGTCGAGCCCTTCCTCTGCGCAATTTT[C/T]GCAGGAGGTAACTTCACACTAGGCTTGGGCCTGCTTTGTTTCTGGGATTT
173	CAAGCTAATTTCCAAAAAGATAACGAAATGGCTTTGATTAAGAACTTTTC[A/T]GCACAATTTAATACCTTTCCACATCCCTCGTAATGACAGACATATTGTCT
174	CCCATTCTCTTATATCAGTCGACGTCGTTTTATTACTAATCTCCATAGCT[C/T]TTCTCCCACTTAAAATCTCCAAACCAAAACGCCAAAACTAAACACGTCG
175	GAAAATGTCTCTCAAAAAATATATATAAATAAGAGAAAAATGGGTGGTGTG[T/A]TTTTTTGGTCAAATCTATGTTTAATTATTTAATTTCATGCACAGCATC

To FIG. 5H

FIG. 5G

From FIG. 5G

176	CTGAGTTGCTCGCTGTCAAAACAAACTGAAACACGCATTTTGTTTTACCA[A/T]TGCAATGAGAATACATCCACTTCAATTTAAACAACTGGGAACAGAAAGGC
177	TCTAATTTTCTCGTCAACCAAACAGAAAAAGTAATAAGAAAAGACAATAA[C/T]TGAGAGACCTGAAGTAGAAAACAATAACCCGATGAAGACAACAAGCTTCT
178	CCGAAGCTAATCCAAACACTTTAATTGCTTCCATTTTTGATCAAAACTGA[G/A]AGACTTCGAAAGCAACAGAAAAAAAGATAGAGAAGAAACAAAAGCAGTGAA
179	TCCACATCAAGTCTTACTTGATTAAGCTGTTTTCTAGCTTCCTCGCATTTT[T/C]AGTTCATTTTCTTTTTCAAGTTGCTTTTTGTATTCCTCCAACTGCTCCAC
180	TCAATCTATGTACATGAAAAGACTAAGTATTTCATTTATTCCTTCAGTGA[G/T]ACATTTGATAAGTCTTAGGCTTATCAATACTTACCCTGCATTTTACATTT
181	AATCCAAGATGACCCAACCCAGATCATCATATACTACTAATTGTTCAAAA[A/G]TATTCTCCCTTCAACAAAATATTAGTTCACAAAGAGAGTCAAGTGAAGAA
182	TGAAAGCCCATTACATCTGTTGGAAATCAATCCAACCCCAATCTCAACTC[T/C]CTGTACACATAAACACAGCAAAGCACAAAACAGCGCAGCAAATCCAAGAA
183	TAAGAATGAAAATAGGAATATAAATAAAAATGGAATTAAATTTAAAATGC[T/A]TATATTAATTTGTTTATCATCTTGCATTAGAATGGTCTTTTCTTCCATTT
184	CCCGAGTTTGTTGTGACTCGGCCGAGTTCGAACTCAGTTCAGCATTGATT[T/C]TATATCTTATCATTGAAGTTTCAGTCCCAAACAATAATGGTTGTTGTGGC
185	TCCGTTGGATTTTTACTATAATTTTGTAGTTTTTGTATAGATGGTGATAGT[T/G]AGAAATGAGATTTGGGAATTTTGTGATAGTGAAGAAGGGGATAGTTGAAA
186	AATTTAATACTTTCATTATTGGTTACACTGATTAGATAGAGTGGTAGGCT[G/A]TATTGGATCTGTTCTACTAGCTCTGCGGTTGTTGGTGCTTAAGTTTGATG
187	AGGCTACATACAAATTATACCTACCATGACATGTTGAATTTACAGTATTT[C/T]AGTGCTGATGATCATTTTATTTGCAGTGCTGTACTTGCACATATTATTCT
188	TTTGTTAGTTTTGCATCAATGAGAAACATCCCTGCATATGAAAATGCTGT[T/C]AAGGAGTCCTTCGAACGATGCTTGGATCTTTACTTGTGTCCTAGAGTACG
189	AATCCTGTGCGGGATCTTTCTATACTGGCTGCTTCTACGTAAGGCCTTAT[T/G]CTACTGTTAGGTTTTTGCATCTTTAAGTAAATTAACATTTGAGTATGATTA
190	GACACTGTGTCAATTGCAAGTTGGCATCAAGATGACAAACACGGTGGAAT[C/T]AGATTGAAGCATTTTAAGGTGTGTTAATAATTGCCCTTAGTTTTTTCTTT
191	AGGTGAAACTTCTTTTCTTTAATATAATTTATTCATCAAACAAAAAAACA[T/C]TATGGATCTTGCTTAAATGTGTATCGAATTCCACATATAATGTTATTTAC
192	TGTCACATATTTAGATTATTACAAAAAATGTGATTCCCACACGTATATAAAT[C/T]TGTATGGAATTTAAGTGCATCTCTAGATGTATTAGGATACCAAATTGAAG
193	ACAGCCTTTACAACTGTACATTTATGAAAAAACATTAGAGATACAACAAA[A/G]AGATAAACTAGTTGGAAGGAATAAAAACACAATTAGGAATAAGGATGGAA
194	AGATAACAAAGAATCAAATATTCCTAATTACCCAAGAAAGAAAACTATAA[T/A]TATCTCTAGCTAAAAAAAAATAACAAAATTCTTACCCTAAGAAATATGGTC
195	GGATGCTTTAGCATAGCATACAGAAACACATACGATGCCATTTATTATTTT[C/A]CTCTCTTAAAACAAACTTATATTCCATTTACATCATGTCACTTAAGTGCG
196	TATTACTATATATCTACCCAATGTCGTATTTACAGGGAAAAAAGCTTGAC[T/G]GTGGGCTGATGAATAATTTTACATGGAAACTTTAGATTTGTATTTGAAGT
197	TAACTTACAAAAAGGAAAACCTTTGAAAAGACATGCTTTCCCAAGATTCA[T/C]GGGTTAGTTCATCTTGGAGGGTTGATTCACTAAACGCGTTTGTGTTTCTA
198	GCAGATATTGGGCTCATCAGTTTTCCAAAAAATTTCCAGGTATGTGGTTT[C/T]AGTATCATTCAAGCAACATAGCCTAACTGGTTTTAATGCCATGGTAAAGT
199	TAAAGGTAATACATGTGATTTGGTTGTACTGCTTATGGTAATTCCAAATT[C/T]CTTTCCCCTTTCCAAAAATCCCACATTGCTTCTGGTGAGATTTGATCATA
200	TTTAGTACTAGGGTACATGATATTGTGCTTCTATTGCTTGCACACTCAAC[C/T]GCAAATGCTCTTTTTTCTTTTCCTTCTTAGAGGAGTGTACCACTGTAAACT

To FIG. 5I

FIG. 5H

From FIG. 5H

201	ACCGATGCAATAAGAGCAAAGATAGCTCAAGAGGAGTACAAAGAACATCA[A/G]AATGCTAGGCAAGAAGCATTGCAGAGAAAGAAAGCCGAGAGGAGAAAACT
202	TACCGGCTGAGCTAAGACGGCAATGAATTTTTCAATTAAAATTTTACTAT[G/C]TAACACAGTGCCTCATCGAGCCATTGGGCCTAGTTTTAATCGGCCCATTT
203	TGGCTGTCCAGACAAAGCTTCCACACAGCTTCATTCCCACTCGAACCTTC[A/G]TCAACGCCAGAGTCAAATGGGTACGAGACCCATACCTAGATTTCGCAGTA
204	GAGAAGCTAAAATGGGATATGGGTCTTCCTCACAATTTCATATCTTCGCT[C/T]GTAATCGATTTTCCCGATTACTTTCAGGTTTGCGAAATTGAAAACCCTAA
205	GACTCAACAAGACACTGAATTTGAAGTGAGTTTAATCATGGTGGTTGTGC[T/C]TGAATAATGTGCATGCATATTTGCTTTAATGCTCATATTTGGTCTTTGCA
206	TTTAAAATTGGATTATACAGAACTTCCCTTTATTGAAATATAGCTTGGCC[A/G]CATTTACTATCAACATTTTTCCACTTTAGTTTTTTGTCATTTTCTGTATT
207	TGACTTTGGGAGCTGTTTCATTGCAGCATCATTTAGATAATCAATCCTTA[A/T]GGTTATAGCACAGAGAAGATGTTTTCTTTGAAACGGGTGCAATCATACCA
208	AATTTTTTTGACAATAATTGTTGGTCTATAGGGGAGAGAGGACTCTGTTA[A/C]TGACTCTCCCAATAAAGGACTCATATTATAAATAGAAGCATTATAGTCCC
209	TGCTGTTGATGCTGAAGTTGCTGTTGCCTGTTAAATCCATGCTGCATAGA[G/A]TCTCTCTCTTGATTGGGGGTTACCACAGTTGACAAAGGCATTTGATTAAC
210	GCTTCAGCAAAGCTATTGGTATTAACAGCATACAATGCCAATCCACCAAA[A/C]CCAATAAGGAACATAGTTGCAAATAGAAGAGCAAGCAGCTTGGCATAAGG
211	GTTCATATCTATACAATTACAAATCTAAACGAAATTAAACAACAGATCCA[G/A]ATGAAATTGATATATGATTACCTTAAGTTGCTCAACTCTTCCATGGAGAT
212	AGTAAACTAGCATCGGATGTATTGGCTATATATCAGATTGGACTGATATTT[T/C]ACAGCAGAAAAGATATCGTAAACTTACCGTAATATATTTCTAAGTAAAT
213	TCCGAATCCAGGTGATCCTACTACACCAGGGCTTAGACTGAGTCGACTCA[C/G]TAGAAGAAGAACGATTAGGAACGAGTTTCACCTTCTCACTCTTCTTACAC
214	TTTCACCTTCTCACTCTTCTTACACGACGGAGCAGAGATCCGATGAGGAG[G/T]TCGCGAATCGGAGGAGGAAGGGAGTTTGTGAGCCACGCGCTTCGTTGTTT
215	TAACATATAATCATACTACTAATTACTATATTAATTACCTTAATTTCGGG[G/T]TTAATCGGAGTTCTAGAAGAGACGGCGGCCGGAGCAAAGATGCAGGAGTG
216	CAGTGATACATTACATTGTTGAACCCTTCCTTAAGAAAAAACTGGTAAAGT[A/G]ACAGTATCAACCACAGTTCCCAAAACATTGGTCTCAACATCATTGTAGAT
217	GTAGAGAAACTTACAAAAATATTAGATTTTGGGCAAATATTTACAATTTT[C/A]CTACCATATGGCAAAAAGAGAAATTTGCTAAACTAATTATTTGAATTACT
218	TATATTATTAGAAGTTTGACCCAAATAATGTGATCATTTTAGTAATTGTA[G/A]GTCATTAAATATTAAAGTGTCATAAACTAAATATGTAAATACTTTAAAAG
219	CAGAAACCGATAAAACGAATTAAAAGAGCAAGAAGAAGAAACCCTAAA[T/C]TGAGTTGAAGAGAGTTTTTGTAGTGATTAAAAATATGCCATTGTTTGATAT
220	TCTTGAATACACCACAATGCAACCATTACAAACCTCTCTACTCTGTTCAT[G/C]TCACCCAAGGCTTCTTCATCATTCACCAAAACCATATCCATTTTCCCATC
221	ACTGCAAAGAAACTGGTATTGGAAGGCACAAATCTGCACCAAGGACACCA[C/T]AGCAGAATGGCTTGGCTGAGAGGATGAATAGAACTCTTATTGAAAGAGTT
222	TACATGTCCACAATGGCATTACCCACATAGACATCTTCATTAAAACCATT[T/A]CGGACAGCAAAACCATGAACTTCCTTACCCCACATTGATAAACCCATAGA
223	CCTAAGGTTTCGGGAAAGGTTGGGATGTATGTCTGTGGAGTTACTGCCTA[T/C]GATCTTAGCCATCTCGGCCATGCCCGAGCTGCCATTAATTTTGATGTCCT
224	TTCTTCCTAAGATCTTATTTATGGGATTGTATTTTAATATAATATTTTTT[C/T]CATCGTACTTTGCAGATCATCAAGAATGACTATGCATATGCAGTTGATGG
225	GGTGGAATTGACTTGATCTTTCCACATCATGAAAATGAGGTTGCCCAGAGT[T/C]TGTGCTGCATGCCAAGAGAGCAATGTAAGTTACTGGATGCATAATGGACA

To FIG. 5J

FIG. 5I

From FIG. 5I

226	TGGTAATAAGCACCAAGCCAGTGTGGATATTGAGCTCAAGGGTGGATCCT[C/T]TTGGCACAGTGGTTGCATTTGCATGCCAAACTATGGTTACATCAGGACCG
227	GTCCAACACAGCAGCATCAAAAGAAATAGTAAGCTTAGAAATGTTTGGTT[G/A]AGATGTGTACTACTCATCATCATTGTGTCTTAATAATATAACTAATAATA
228	AGAGCGAGGCCACGAATTGGTGGACTATATCCCAGGAATCCGTTCAACTC[G/A]GCTATCAGATATCTCAGAGAATTTTTTTTGAGAAAGAAGATAGTAAAACAA
229	TGGTTCATTAGTCTTGTTAATCTCAGGGAGAGTTTAATTTAATGAACCAAT[A/C]AATAATGATTCAAGTGGATTAAGGCCTTATTTGCTAGGGAAGTTAAGGTG
230	AAAACCATTTGAATGATTAGAAGATGACATTGGTATCAACTCAT[C/T]AAAGATGGTTGTTCCATTCTGTTCTTGCAAGTAAATAGAGCCAGATTGGT
231	TTCGGTTATCACTAGGTTGGAGATTAGAGAGTTTTGTTATGATTATGAAA[G/C]TTCTTTGTGTGGCCCCTTGGATAAATTGTTAGCTTTTGCAGTTGAGAATA
232	ATCAAAATCAAGAGTTGCTCTATAATAGTAATTTGCAATAGAAGCACTCT[C/T]TGTACTAAACATATCCATTTCATTATTCTCTCTCAGTATATACAAACTAC
233	TTCAGTGATGAAAATCTTTTTACTTCTTCTACTACCATTCTTTTGGCGTAA[G/A]AAGAAGAATCCTACACAAACTGTGGCTGCCAATAAAATATTAACAAAAAC
234	TGATCAACTCATCTTGGTTTTTCCTCCTCTTCTACATAAGGAATCCCA[T/A]CTAAGGTGATATTGTCTTTACGAACTTTGATAAACGCCCTCGAGTTAAGA
235	ACTGATATAAGTTGTTAACTTCCCCTCTCAATTTAGCTGCTTTTGTTGGA[G/A]GAAAAATTTAGAGAGGAATTTCATCGGCAAGTCACCACATGTGTTGATAG
236	TTTGTCCCACCAACAATGTGTCTGTTGGTTCATCGAAACTCTCCCACAAA[T/C]TGGTTTTGTTGTTTGCTAGGACAAGATTACCTGTGTCTAACACAGCC
237	CAGATAAAGAATTCCAAGTTGAAGTGAATGTTATAGGCCAAACTCACCAC[A/G]AGAATCTTGTTGGGTTGAATGGATATTGTGATGAGAATAATCATAGGTTG
238	GTCTTAAACCAAGTTGAAATCAAAGGACTCTAATTGCTGTTGGCATTGCA[T/G]ACAACACTCAGATCAAAGCCAATTTAATACTAAGACTATGATTAGAGGAA
239	TTTTTGTGAATGTTCAATGGATAAATTCAGACTCAAATGTTGATGAAAAA[A/G]GCAAGGAAAAACTAGCACTAGAGATGCTTAGAAAGTGCTCTGGTCTGCCA
240	TCTGGTCTGCCATTAGCTATCATTGTGCTCGCTGGGCTTCTTTCTAAGAA[A/T]CACACCTTATATGATTGGGAGCTGATGAAAGAAAATGTAATTCGCTGCAT
241	CACCACATGGAAATCATCACATATATCTCTAACCTCTTCTTATCACCCAA[A/G]ACATCATCTTCATAATCCACACAAGAACATCCAATGACCCTTCTCTGAAGCA
242	AGTAATCCCTTGTTGACACTACTCTTCCTTGTGTTAAATTGAACCTTTCG[C/T]CATTTCCTCTCGAAATATACAAGTAGCCCGTTTCGTTGAAGATTATCTGA
243	CTTTACATTCAACTCAGTAACTACTTATGATTTTGGTCTCTAAATTTGTG[T/C]TCAACTAATCAAAGACCAAACTCAAGTACAGTGGAGTAACCTGGTGATAA
244	CCGTTACCTAGTCCTACTGTCTCAGCATCTCGTAAAAAACAGAAACCATC[A/G]CAATCTGTAGCTTCGATGTCTATTGGTGCTCCCTCGCCAGCGCTGCCTCC
245	AAAACCCATCAACCCACCAAATTAAATTCAAATAAAAATCCAAACTTTAA[G/A]ATAAACAAGATCACTCACTCTATCAAATCACAGATCCCCAAATTGAAAGA
246	CGAACCCAGTTACAAACAGCAACAACATCTTTGATTTCAGAGAATCCGGT[A/G]AGAGAAGCTCTTCCGGTGGATTTTCCGGCACCTCTAGAATCAAAGGTGAC
247	AGCTCATTAACTCCAAACTCATGCTGATCATTTGTACTTTTCCTCCCAAT[C/A]ACTCTTTCTTCACTCCTCTTTTTCTTAATACCCTTTTCATCAATCGAAAC
248	TTTCTTCAGAAGACTTATTACTCCTTTTACCACAAGTTACTATAATAGAT[C/T]GAGAAGACTTGTTTGATTTGGGAGCTGATGAATCTGTTTTCTCTACAACT
249	AACCTGTCTTTACAGTTGTCTGGACTCCGATACACAGAGGCGGTTCGGTT[A/G]GAGATTTCCATGGCGATGGAGTCCCAGTTCTTGGTGCCGTACCGGTTAAC
250	TCACGTACTTTTTTTGGGTTGTTGGGCCAGCCCGATTATGATTTTTATTTTATATTTATTATTGGGTAAATATATTTTTCTATTATTGAGGA

To FIG. 5K

FIG. 5J

From FIG. 5J

251	TAACGACTTTTCCGGGTTCGGATTCGGGTCGTGGAGGATCCGAATCATTT[C/A]TCCTCTCTGGTTCGGGTTCCGTTTCGGATTTCACGGCGCCGTTTCGTCGA
252	ATATATGTTTCTATGTATATTTGGTTATAAAGTATTGATTTTTTCTGATTG[T/C]TAATGAAGAAAAAAGAGATGAGTTGGCAAAACCACAGTCAAAGAAATTCA
253	GCAACAATGAGGAACCCTAGAACCTATACAAAAAAGATCTAGGCCAACAGA[A/T]ACTTCTCCAACCAGAGAAGGTAAAAAGTTTTCATTTTTTAACTTTTAATGAA
254	ATAACTGTTGACAAAAATAGTCGCCAACTCGTTTGTAAGTCTACTCTACT[T/G]TTTATCAGATTTTAATAGTCATTAGTTGTTGTGTTTATCCACCTTAATTG
255	TAATAAAAATTCGAGGTGTAAAATATGATAGTGGTCATAATTCGAGAGGT[A/C]AAATAGCTAATTATTTAACGAAAGTCCTTCGAGAAAAATATTTAACATTT
256	ATCATCAGCAACACCAGCAATAAAAGCTCTCTTACCTAAACAAAACAAAA[C/T]CCACTTATCATATTAGCAAAAATGTCACATTATTAAATTTCATTGAGACC
257	CACAAACGGGTCTTGACAGCTGGAGGAGAGGCGCCATCTGGGTAAGATGG[T/G]GGAACTACAGGGTTCCTATTGGGTGGAGGGAGGGCTGGGCTGCCACCAGC
258	ACCTAGCAAAAGATATGCATATTTCCACCCTCCTGATGATCTTGTTGTTA[A/G]GTTCTTCTTCTCATAATTTTCAGATTCATCATCATCTCCAAAACTTGCAT
259	TCATCATAAGTCCTTTCGATTTCTAAAAACAAGAATTATTTACTCTCTTAT[G/A]GAAAAGAACAATGCACACAAACTTCATAATAGCTTACATTAATCAACTCTGG
260	CATAGCATATAAGGCAAGAAGAAGTGGATATCCATAACTTTCAATTTTAT[C/T]TCAAAATATTATCAAATTATCTAAATCACTTCAAAAATGACACTCTAATG
261	ATCACTGATTGTGGATTACTGCTTTGATAATTTTGTGTGATTTGATCACC[A/G]CGTCAAAAAGAAAAAAAAGATGAATAAACACTTGTCTTGCCAGAAAAGAGC
262	TTAGGTTTTGCATATTCTAATGTATAGTAAATATTATGAAACATCATTGT[C/T]ATTAAATTTTAGTGGTGTTTAATTGAGATGATGCCCATGGCTAACATAAA
263	AATAATATATCACGTGATACTGGACACTATTCGTAGTCTTTAACAAATCGT[T/C]TAAACACTTGTAACACGTCACCTATAATATAATTAGAGACCACATGTGCA
264	AGTTAATTTATGTATTTTGGTTGTTTTTTGTTATTTTTTAAGTTGGG[A/T]TTTTTTTGTTATTTGATGTTTTTTATGAAATTAAATGTCGAATTAATGTAGTA
265	TGGCACATGTTAGTTCAGTGACGGACTCAAAAATTTTGATTTAATAGAAG[C/T]GAAAATTGAATAAATATGTGTATCAATAAGGACATAATATTAGATTGTCC
266	TATAATATTTCTTCACATAGCTTTTTTCACTCTCTCTTGTCACACCACAAA[A/G]AATAATATTACATGCTAACACTTCTTACATCTTCTTAAAGTTCTCTTCAA
267	GAGGGAGCTATCGACACTTCGAGAATGGCACGATTGATCCTTTCGGTTTA[C/T]ATTTGTATGGAATTGTGCTCAAAGAGGAAGGCAATGAGAATCTTGCTCGT
268	CCGAATGATTCTCGGTTGTGGATCGCCATGGCCCGGTGTTATGAAACCGA[C/T]CAGCTTCGAATGCTTGATGAAGCAATTAAGTGTTACAAAAGGGCAGTTAA
269	GTTACTTATGCTTATACTTGGGAGAACCCTCCTGGATTTGCTCAAACAATA[A/T]AAAAATTGGTTTGGAAACAATTCAAGTTCGCCTTCCTTGTTTATTTTGGC
270	ATAATTATATACCTTGCTTGTTTTCATCTGATTAATGACATGCTGATGAG[C/T]GATTATACATGTGTGTCTTGGATACAGATAAAACCTCTAGTAGGACCGAC
271	GAATCTGCCATGGATGTGCCATCCAATTTGGAAGCAAGAAGAAGTGTC[G/C]TTTTTCTCAAATTCATTGTTTATGGATATGCCTGATGCCCCTAAAGTTCG
272	GAATTGCAGTACAGATCAGATGTGTATCATGTAATTCTTTGATGTTTTTGG[C/T]GTGTTCTGTTAAGTTTGCTAAAACAGTGTGTTATCTTATTGTATGAAAAT
273	GTCGATGATTGGACTGATTGGAATAAGTGGATAAGCAACAGAGGAGGTAT[C/T]GGTGTTCCGCCTGAAAAGAGTTGGGAATCATGGTGGGAGGAGGAGCAAGA
274	CTTCAGCAGTTGACTGAGGAAATTCCTGATGAGGTTTGTTGTATGTCTTTT[G/A]TTATCAGTTGTATTTTTGAACATCCGAAATTTTTGTGAAGCATTTGTGTTG
275	GCCATGGGGACCTTTGTCTCATGGACAGGAGGGGGCTGTTAAATGAGTATA[C/T]GCTTAAAGAGCATAGTGAAATAAAGTTGTGATATATCTTATGGGGGTGCT

To FIG. 5L

FIG. 5K

From FIG. 5K

276	GAGGAAGCTTTGAAAATGAGAAACCTGCTTGAAGAATTCCACCGTGATCA[C/T]GGCATTCGTCCTGCCACAATTCTTGGTGTTAGGGAGCATGTCTTTACTGG
277	ATATTTGGCACAAGAACTTGTTAAATAATATGAAATGGGACGTTCATTCA[T/A]ATCTTCTTGATTCTTATTCAAGGACTAGTAGAGAGTCTAACTAGTAGCAA
278	GAAAAAAGCTTACTTAGCAGCAAAATCACCTCTAAAACTCCCTTATCTGT[A/G]GATATAGCAGAGAGCAACTTTTGCCTGATTTCTGTCTTCTTTTGTAACAC
279	TGCCAAGTGATGTTGTGGAGAAAATCATGTTGTTGGTGCCTGCAGATTCT[G/C]TAGTAGAGTGTAAATCTGTGAACAAGTTCTGGTATTCTTGTATTTCTACT
280	ATTTTCTTATGTTGACTTGGTCAATGGAGACTTTATTCGAGTGTTGTCTA[A/C]TATATATATACATACATAAATTTTTTTTATATATATAAATTGTAACA
281	GGCGAAACCATCGAGTGGCTGCTCCAGCAAGCCGAACCTGCTATCATTGC[A/G]GCCACTGGTACCGGAACGATTCCGGCTAATTTTTCGAGTTTGAATATTTC
282	AGACCTTCTAATATTTTACCGGCTACTGCCATGTGGGCAGTAGCGCCAGC[C/T]CAAGCTAGTGGAGCTGGCAGCACCTTTTGGATGCTGCCAGTCTCGGCTGG
283	GGAGCTGGCAGCACCTTTTGGATGCTGCCAGTCTCGGCTGGAGCTGGAGT[T/G]TCTCCAGTTCCAGTAGCTGCTGCTGCTGTTACGGTCAGTCAGCAGCA
284	AGTCTCATCATCAGTAATGTCAGTTGGAATTAAAACATTTGAAAATTTAG[T/A]ATGATTTATAAGATGCAAAAAAGATGATAGCTTTAGAGCTTACGATAATG
285	AAAAAGAATGCCCAGTTAGATTCAAAAAAGCTTAGCCGTTTGTCCAGTTC[T/C]ACATTATGAAGATTCCATTTGTACCTTACAATTTCAAAGGAGAAATTAGT
286	TTAGATGGAAAAGAAAAGTATAATGTTATTTCGTTGAATAGGACTACACT[A/G]CTAGGCAAAAATATTGAATTTGAATATAAATCTTCAACTGCATACTTACAT
287	AGTTTGGGATCTTGCCATGGAGCCAGGTAACTATAGATTTTCATGCTGCA[T/G]TATGATGCAAATTTTCAGTCATATATATTGTTGTTTCATTAGATGCTCCA
288	AGAACAAATATAAAATGAGCCTAATTCCTTTAGAGAAGCTGAACTGCAGC[A/G]ATTGCCGGAATTTTTTCCCTCAATTTCTGTGTAACAGCAACTCTTACAGT
289	TTCCCACCATCTGCAATGAGATATGGTCGAACTTCATCCAATACAGTTTC[A/G]ACGTTCTCTTCAGTTAGTGGTAATTCCACTGCTGAATCTGGAGTAGCAAC
290	ATCATTGAAGCAAGAAGTTGTCAAAACAATTCATCATTTTAAAGCTTCTG[T/C]GATCAAATTGATCATTTTATTATTCTTTTATTTATATCAATAGAAATATA
291	TTTTTGTGTATCTTTCTAGTTATTAAATGCATTTCATTGTTTTTCTCAAA[A/T]GTAAATCGATTCTAAATGTTGGGTTATTTCTATAGGTCGGTGCTCTTGAA
292	TTAGCATTTCTCTAACATTTTTTCCTTTGATAAGGAGAGGAAATTAGGAAAT[C/T]AACCATCTAACTAGGATTATTAACTTTACTTTAGCAAATATTATGTTATGT
293	GGCGAGAAAGCTTTCGAGAAATATGGAAAGGGCTTCATCTCAATGCATTTT[T/C]TCAGATCAACACTTAGGCACAACCAAGAAGATGCTGCTGTTCAAGTTTGC
294	ATCCAAATTCTTCAACAAATACCCATCTTTCTTCTCCATCTTCCAACCAA[G/A]TCCAGGTCTTCCTCCACGTGTCAAACTCACACAAGACAAGCTCTCTCAATCC
295	CAACCTCAGGGTAGTGCAACATGTCCAAGGCCGCAACAACAACCTTTACA[A/G]CAGACTCAGCAATCTCCTAATGTGCAAGGAAGCGCCTCCCAATCTCCTCT
296	CTAAAAACTATTCTCAATAGTAGTATGATACCGAGAAGGGTGCTAGGCAC[T/C]AATAATACCCTTTAGCATTTCTCAATAATAATACTAAAAAAATATTAAAA
297	GACATATGAGGTTCCCTTTCGAGTACAGCAATTACATCCTTGATTGTTAT[A/G]TTACTACGGGGCACCTTTGTTTGAGGCGAGACGCCTTGGCTTCTTCCGAA
298	CGGCTATCCTCCTTGGGTACAGAAAAGAGCTGTTCTTCCTCCTCTATAA[T/G]TGTGAAACATCATAAGATCAAATCATTGCATAATTAACTAGTAGACTAC
299	GGAAGAAAGTGAATGTGGGGGTGCCATGCTTGTGGATGATGAAGGAACAA[C/A]CCCAGCAGAAGTTGGAAAACCCATTCTTGATGGCTGCTTCTCTAAAGACT
300	TGCAATGTGGAATTATGTAAAGATTTTAAGGCTCGTTTGGTACGATATAT[T/C]TGGTCGTATTATATTTTTATATAATACTATGTTAAATTTTAATTATAAGC

To FIG. 5M

FIG. 5L

From FIG. 5L

301	AGAATGAGTATATGATTCCTTTCAATGACTTCACTCTTCCCTTTCCTTAA[G/C]GAATCCACCTTCTCCGATATGGCATCTGAAACAAGCCCGAGCATCATCGC
302	AAATTCAGGGGGCATGATTTAGTACATGTCAAAGTTCGGAGAGAAAAATT[G/A]CTAATTAGCCTAAATATAACTCTTCACCTCATAACACATTTTGACTCACT
303	AGATCGGTTCTGGCGGCGGAGAAGGGTTGAGTTGTAGTGGGCCTCCGGTT[G/A]GGCCCATCACCGGCAATGGTTTTGGACCTTTTCGGAACCGGTGGGCTGTC
304	AGGGTGTGATAAAACTTGTTCACAAACCTAAATTGTACTAGGGAATCAGC[A/T]GGTGCCCACAACATAATTTTCTCCACAAGTTCACTTGGCAAATCATAAAA
305	CTTGTGACGGGCCTCGTGGATGGGCCTATCATTTGGGCTGAAAATGAGGT[G/C]CACATGTAGAGGCCCAACCTCGTGTTAATGATGGTAGAGAAACTTACAAA
306	CTTTAGAGTACATATACGACATGTCGTTTTCTTGGGTAAACATTTTACGC[T/C]TTTTAGCCCTTTAGAGTATATATACGACAAGTCGTTTTCATGGATAAATG
307	AATTTTTCAATTTTGTTTTAGCTAAACCACACAAAGCTTTGTCAGTATTA[A/C]TATTTTAGTTTAATGTTAAAAACAACTGATGATAACAAATTGTTTTTTCA
308	TTAGGCCTTTGGTAGTTATGTTGACCAAGCATTTGGTGAAATAGAAGTAA[C/G]GCATCATTGGCTTCCCCATGTTGAGCATACCCACCAATCATGACTGTCCA
309	AATGAGATTGGGGAAGATCCACTAAGTTTAAGTAATCGGTTCTGCCAAGA[G/A]TACCTTTCTGACATGGATGCTCTTCAGTGTCTCCTTCCTACCCATCAGCC
310	AAGGTATATATAGCCTGATTGATTGAAGATGAGTTGAAAGCCACTGTTTA[C/T]AGTGTTGCTAGACCAATAAGGGAAATTTTGAGTGTCAAGTGGGAAAGTTC
311	GAGAGGCTGGTTTTAACGGCGGGATATGCTGAAGCAGAAAGAGCAGTTGC[T/G]TTGGAATTTGGAGATGCTGAAGAGGCTAAGCTTCTCCTTGCTGAGAAAGA
312	TAATAATGTATTTATATATATTGTTTATTTTTCCTAATTTTAAGATAGAA[G/T]ACACACAATGTTACAAATATTATACATTTAAGCTATAAAATACAAATTAT
313	GAGTTGTTAGTGATTTTTATCATAGTTATTAAATTATACATATATGAAAT[C/G]TAACATTGTTTATATACATGTATCATGATCAGGCAATCTATATGTTTTGA
314	GACTTTGGATTAGCAAAGCTATTGTTGATTAACCAGAGTCATACTAACAC[T/A]GCCATTAGAGGGACAAAGGGATATGTGGCACCAGAATGGTTTAGCAACAT
315	CTGACAGGTCTCAATATAATTATAATTTATGCATATGTTTGACATGATGA[T/C]CATCCAGTAGTATGTAAGTCTCTATAATTAATGTTTCATTACTTAATAAT
316	TAGATAATTGTTAAATCACATATTGGAGAAAATTCTAGTAGCATATTTGA[A/G]TAAGACACTCAAGTTCTCTGATTGGATTTAGTTATTATTATGATTCATCA
317	TGAAGTGAAACGTCGTTTGAGTTCATCAAACCTGTGTTTACAGGTATCTG[A/G]AGTCCATGGGGATTCAGTTCGGTTGGAGAGTTCCATGGCGATGGAGTCCC
318	AATCACAAAATACAAAGCACAAAAATTGTCTTATGAAAAAAGAGTGATAA[A/G]GTACCAGCCGAGATCGTGTTGACTCTAACTCGGTGTTTTCTACCCGCTTC
319	TTGCACCAAATCTTCCAACAAGATTGTCTCTTGCCCTAACACCGTCGATT[T/C]TTGGTGATTCAACGCGCTTATTATAACTATACCGAGCAACATCTGATCTT
320	TATGGTAGGGCAATATCTGGTTTTCCCACATTCTTTCTATCTTTCCTTTG[G/T]CTAGAAAATCTCTTATCTAATTTGTCTTAACACACAATATTTCAAGCATC
321	TTTGCAGCTAAGCACCTCCTCAATACCAAAAACCAATCTTCTATGTCCTT[A/C]CTTTGTTTCAAAGACCCTTCCCCCCATGACGATCATCGCTTGATCACATA
322	TATAAATCTTCAACTGCATACTTACATAAAAGAGTTTAATGAGTCCAAAC[T/C]GTAAGAAGGAATAAAATTTCAAAAGACCACCAAAAGAGCAGAGTTCTTGA
323	CTAAATTAATTCTCTCATTAGTGATCTACTATGGCAACTTATGTCTAACT[G/T]AATTCTCTCATTAGTGATTAAATAGAGTGGTAACTTGTATTGGATCTGTT
324	AGAATTCCATTCAAGTTACACGGGCTTCCAGTCACTTCAGTTTTCCACCC[C/T]ACTCGTTCTATTTTCTTCATTGCAACAAAAAAGAATGTTCGTGTTTATGA
325	ATGATTCACTACTTGCTCCTGCACAGATGAATTCATTGATTGTGGCTGTG[T/C]ATTGTTCTGGTCTTGTTGAATTTGAAAGTTTGCACTTTCATCTAAAGTAG
326	TGGTTCTTGATGGAGAGTTCATTTCCATGGATTTGAGGATCACATTCACA[C/T]TCATTTCCAAATCCAAATAAGCTCATTTTTTGTTGCCTTCCAAAAAGACC
327	ATTATACAGACGGACAGTTTCCTACTGACAGTAGAGAATTTGCAAAACGC[G/A]TATTTTACACAGTTTATATGGGAACAGAAAACAGGTGGACACGCTCAAGA
328	ACCTCTAATGAGACGGACAAATGACTCTTTGGTGATTTCTTCCCTCTGAA[C/A]ATACAAAAACCAATAAGAAATAGTAAAAGAAAACTAATAATACAAATGAG

FIG. 5M

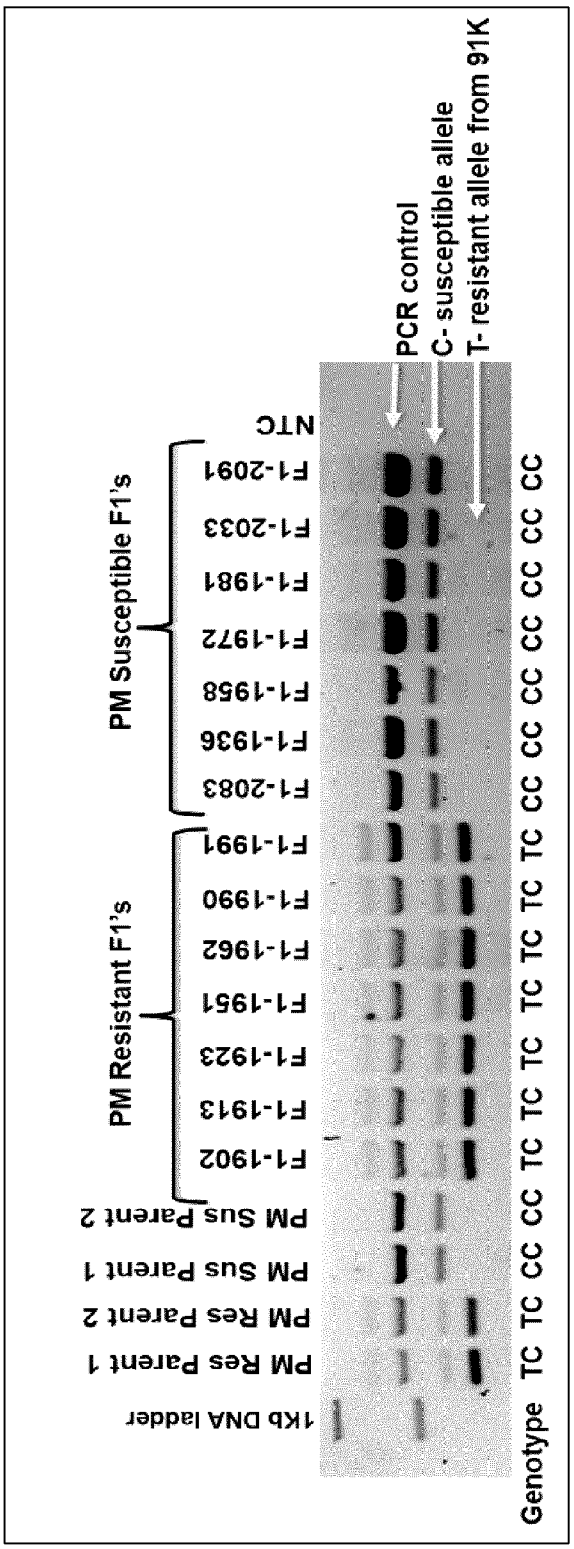
PCR control
C- susceptible allele
T- resistant allele from 91K
NTC
F1-2091
F1-2033
F1-1981
F1-1972
F1-1958
F1-1936
F1-2083
PM Susceptible F1's
F1-1991
F1-1990
F1-1962
F1-1951
F1-1923
F1-1913
F1-1902
PM Resistant F1's
PM Sus Parent 2
PM Sus Parent 1
PM Res Parent 2
PM Res Parent 1
1Kb DNA ladder
Genotype
TC TC CC CC TC TC TC TC TC TC TC CC CC CC CC CC CC

FIG. 6

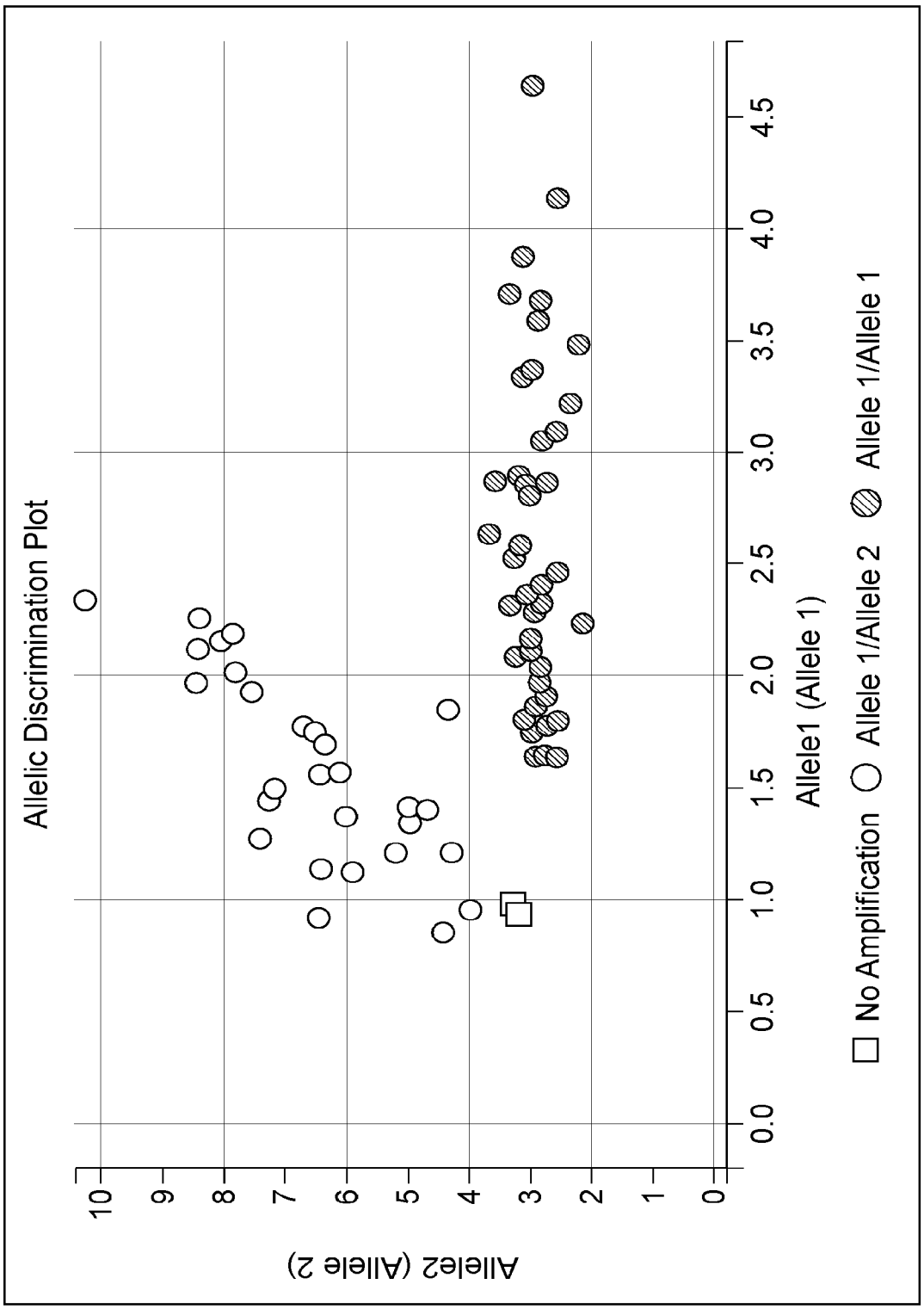
Allelic Discrimination Plot
Allele2 (Allele 2)
Allele1 (Allele 1)
0
1
2
3
4
5
6
7
8
9
10
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
No Amplification
Allele 1/Allele 2
Allele 1/Allele 1

FIG. 7

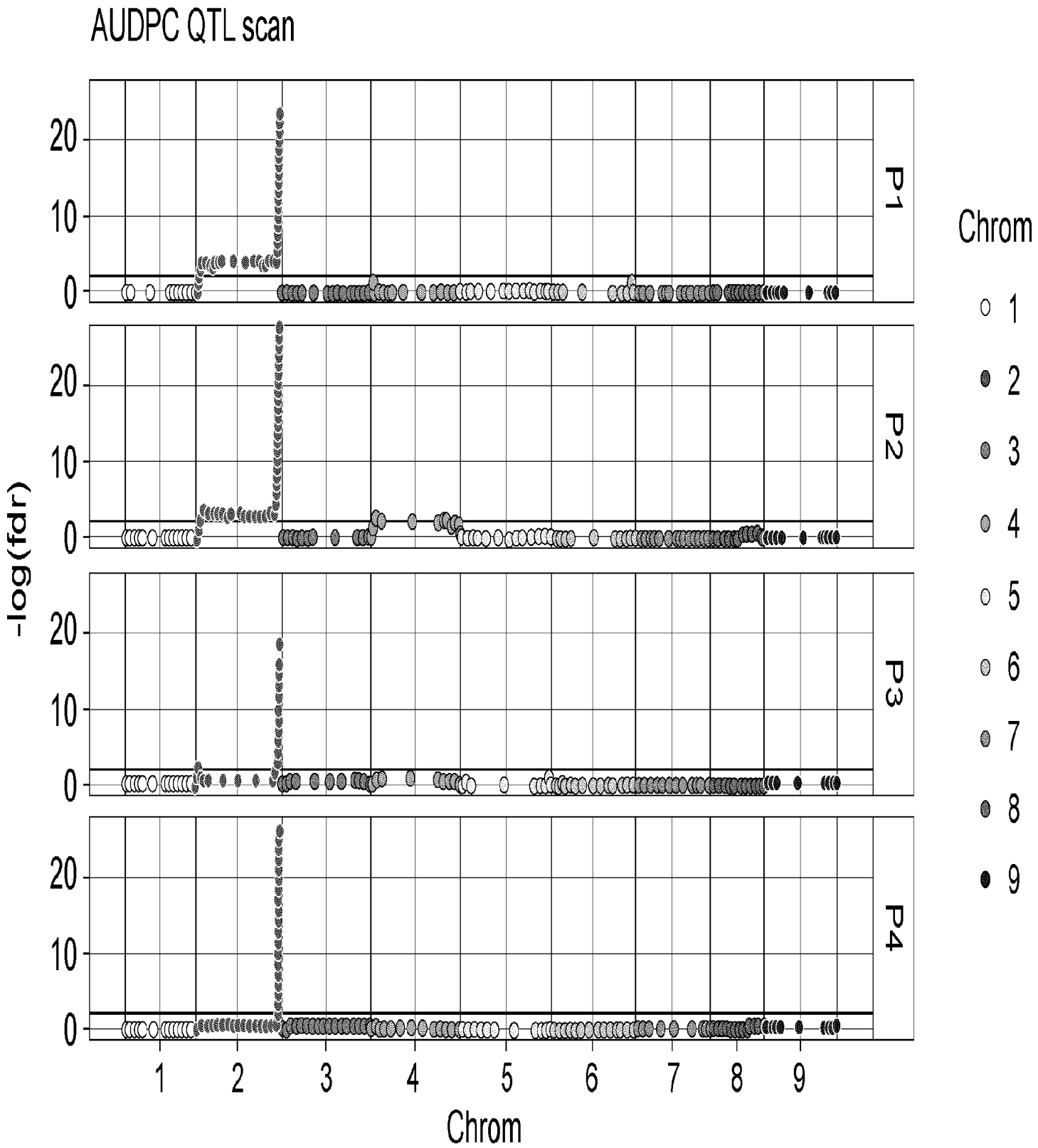
AUDPC QTL scan
P1
P2
P3
P4
20
10
0
-log(fdr)
1
2
3
4
5
6
7
8
9
Chrom
Chrom
1
2
3
4
5
6
7
8
9

FIG. 1