

a selection between an autosomal model of Mendelian inheritance, an X-linked model of Mendelian inheritance, and a Y-linked model of Mendelian inheritance; and
a selection of whether to allow de novo mutations.

[0006a] In one embodiment, there is provided a computer implementation system for assessing a causal genetic variant associated with a disease in a proband, comprising:

an input module that, by a processor, receives an input comprising (a) a selection of a Mendelian inheritance model based on a family history of the proband, and (b) the family history; the Mendelian inheritance model selected from the group consisting of autosomal dominant, autosomal recessive, and sex-linked;

a processing module that, by a processor

removes, from a list of genetic variants, variants not compatible with the selected Mendelian inheritance model;

removes from the list (a) variants that are present in unaffected controls below a first specified frequency, (b) variants that are present in unaffected controls above a second specified frequency, and (c) variants that are present in unaffected controls below a specified number of occurrences, each of the specified frequencies and number being determined by the input;

removes from the list variants that are in intergenic regions;

removes from the list deep intronic variants that do not have corresponding records in a database containing (a) known genetic variants and (b) diseases corresponding to the known genetic variants; and

removes from the list synonymous variants that do not have corresponding records in the database;

an identifying module that, by a processor, identifies the causal genetic variant associated with the disease, based on one or more remaining variants in the list; and

an output module that, by a processor, outputs the one or more remaining variants to a display.

[0006b] In one embodiment, there is provided a machine-readable medium comprising machine-readable instructions for causing a processor to execute a method comprising:

(1) receiving an input from a user, the input comprising (a) a selection of a Mendelian inheritance model based on a family history of a proband, and (b) the family history; the Mendelian inheritance model selected from the group consisting of autosomal dominant, autosomal recessive, and sex-linked;

(2) from a list of genetic variants, removing variants not compatible with the selected Mendelian inheritance model;

(3) from the list of variants, removing (a) variants that are present in unaffected controls below a first specified frequency, (b) variants that are present in unaffected controls above a second specified frequency, and (c) variants that are present in unaffected controls below a specified number of occurrences, each of the specified frequencies and number being determined by the input;

(4) from the list of variants, removing variants that are in intergenic regions;

(5) from the list of variants, removing deep intronic variants that do not have corresponding records in a database containing (a) known genetic variants and (b) diseases corresponding to the known genetic variants;

(6) from the list of variants, removing synonymous variants that do not have corresponding records in the database; and

(7) identifying a causal genetic variant associated with a disease of the proband, based on one or more remaining variants in the list.

[0006c] In one embodiment, there is provided a computer-implemented method of assessing a genetic influence for a condition in a proband, the method comprising:

by a processor and from a list of genetic variants, removing variants not compatible with a Mendelian inheritance model determined by an input to the processor and based on a family history of the proband, the Mendelian inheritance model selected from the group consisting of autosomal dominant, autosomal recessive, and sex-linked;

from the list of variants, removing (a) variants that are present in unaffected controls below a first specified frequency, (b) variants that are present in unaffected controls above a second specified frequency, and (c) variants that are present in unaffected controls below a specified number of occurrences, each of the specified frequencies and number being determined by the input;

from the list of variants, removing variants that are in intergenic regions;

from the list of variants, removing deep intronic variants that do not have corresponding records in a database containing (a) known genetic variants and (b) diseases corresponding to the known genetic variants;

from the list of variants, removing synonymous variants that do not have corresponding records in the database; and

by a processor, identifying a causal genetic variant associated with a disease of the proband, based on one or more remaining variants in the list.

[0007] Additional features and advantages of the subject technology will be set forth in the description below, and in part will be apparent from the description, or may be learned by practice of the subject technology. The advantages of the subject technology will be realized and attained by the structure particularly pointed out in the written description and claims hereof as well as the appended drawings.

[0008] It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory and are intended to provide further explanation of the subject technology as claimed.

Brief Description of the Drawings

[0009] The accompanying drawings, which are included to provide further understanding of the subject technology and are incorporated in and constitute a part of this specification, illustrate aspects of the subject technology and together with the description serve to explain the principles of the subject technology.

[0010] FIG. 1 shows an exemplary flow chart illustrating steps for annotating and filtering a vast number of raw variants to produce a short list with rich annotation, according to some embodiments of the present disclosure.

[0011] FIG. 2 shows an exemplary user interface to perform a filter program, according to some embodiments of the present disclosure.

[0012] FIG. 3 shows an exemplary user interface to provide results of a project, according to some embodiments of the present disclosure.

[0013] FIG. 4 shows an exemplary annotation work flow chart, according to some embodiments of the present disclosure.

[0014] FIG. 5 shows an exemplary work flow chart to determine a variant location and DNA level description, according to some embodiments of the present disclosure.

WHAT IS CLAIMED IS:

1. A computer implementation system for assessing a genetic influence for a condition in a proband, comprising:

an input module that, by a processor, receives an input from a user;

an inheritance filtering module that, by a processor, based on the input, and from a list of variants, removes variants not compatible with a Mendelian inheritance model, determined by an input to the processor, and based on a family history of the proband;

a control filtering module that, by a processor, based on the input, and from the list of variants, removes variants that are present in unaffected controls above a specified frequency and specified occurrence that are determined by the input;

an identifying module that, by a processor, identifies a genetic influence for the condition, based on one or more remaining variants in the list; and

an output module that, by a processor, outputs the one or more remaining variants to a display.

2. The computer implementation system of claim 1, further comprising a determining module that, by a processor, determines a treatment plan based on the genetic influence.

3. The computer implementation system of claim 1, wherein the identifying module is configured to identify genetic influence by estimating a probability that one or more of the remaining variants significantly influences at least one of clinical signs or symptoms of the proband.

4. The computer implementation system of claim 1, wherein the input comprises a selection between a recessive model of Mendelian inheritance and a dominant model of Mendelian inheritance.

5. The computer implementation system of claim 1, wherein the input comprises a selection between an autosomal model of Mendelian inheritance, an X-linked model of Mendelian inheritance, and a Y-linked model of Mendelian inheritance.

6. The computer implementation system of claim 1, wherein the input comprises a selection of whether to allow de novo mutations.

7. The computer implementation system of claim 1, wherein the list is an index list, the computer implementation system further comprising a forming module that, by a processor, forms the index list, from a master list of more genetic variants than are in the index list.

8. The computer implementation system of claim 7, wherein the forming module, by a processor:

from the master list, removes common variants that satisfy a user-defined threshold;

from the master list, removes variants in at least one intergenic region; and

from the master list, removes deep intronic variants and synonymous variants that do not have associated records in a selected database.

9. A machine-readable medium comprising machine-readable instructions for causing a processor to execute a method comprising:

(1) receiving an input from a user;

(2) from a list of genetic variants, removing variants not compatible with a Mendelian inheritance model determined by an input to the processor and based on a family history of the proband;

(3) from the list of variants, removing variants that are present in unaffected controls above a specified frequency and specified occurrence that are determined by the input; and

(4) identifying a genetic influence for the condition, based on one or more remaining variants in the list.

10. The machine-readable medium of claim 9, wherein the list is an index list, and wherein the method further comprises forming the index list, from a master list of more genetic variants than are in the index list.

11. The machine-readable medium of claim 10, wherein forming the index list comprises:

within the list of variants, annotating a description of variants based on locations of at least one of (a) mutations in respective genes, or (b) amino acid alterations in affected proteins;

within the list of variants, annotating population frequencies of variants;

within the list of variants, annotating disease information associated with variants;

within the list of variants, annotating with evolutionary conservation indicators at variant positions; and

within the list of variants, annotating at least one prediction of a deleterious effect of at least one variant.

12. The machine-readable medium of claim 10, wherein forming the index list comprises:

from the master list, removing common variants that satisfy a user-defined threshold;

from the master list, removing variants in at least one intergenic region; and

from the master list, removing deep intronic variants and synonymous variants that do not have associated records in a selected database.

13. The machine-readable medium of claim 9, wherein the input comprises:

a selection between a recessive model of Mendelian inheritance and a dominant model of Mendelian inheritance;

a selection between an autosomal model of Mendelian inheritance, an X-linked model of Mendelian inheritance, and a Y-linked model of Mendelian inheritance; and

a selection of whether to allow de novo mutations.

14. A computer-implemented method of assessing a genetic influence for a condition in a proband, comprising:

by a processor and from a list of genetic variants, removing variants not compatible with a Mendelian inheritance model determined by an input to the processor and based on a family history of the proband;

from the list of variants, removing variants that are present in unaffected controls above a specified frequency and specified occurrence that are determined by the input; and

by a processor, identifying a genetic influence for the condition, based on one or more remaining variants in the list.

15. The method of claim 14, further comprising determining a treatment plan based on the genetic influence.

16. The method of claim 14, wherein the identifying a genetic influence comprises estimating a probability that one of more of the remaining variants significantly influences at least one of clinical signs or symptoms of the proband.

17. The method of claim 14, wherein the removing variants not compatible with the Mendelian inheritance model comprises comparing a genotype from the proband to a genotype from at least one family member.

18. The method of claim 14, wherein the removing variants not compatible with the Mendelian inheritance model further comprises removing a heterozygous variant of the proband that exists in at least one unaffected family member as a homozygous variant.

19. The method of claim 14, wherein the Mendelian inheritance model is a dominant model.

20. The method of claim 19, wherein the removing variants not compatible with the Mendelian inheritance model further comprises removing a heterozygous variant of the proband if it exists in at least one unaffected family member or does not exist in at least one affected family member.

21. The method of claim 19, wherein the removing variants that are present in unaffected controls comprises removing a candidate heterozygous variant that presents in at least one unaffected control as either a heterozygous or homozygous variant.

22. The method of claim 14, wherein the Mendelian inheritance model is a recessive model.

23. The method of claim 22, wherein the removing variants not compatible with the Mendelian inheritance model comprises removing a candidate homozygous variant that at least one of (a) presents in at least one unaffected family member or unaffected control as a homozygous variant or (b) does not present in at least one affected family member.

24. The method of claim 22, wherein the removing variants that are present in unaffected controls comprises removing a candidate pair of compound heterozygous variants that present in at least one unaffected control.

25. The method of claim 14, wherein the Mendelian inheritance model is a sex-linked recessive model.

26. The method of claim 25, wherein the removing variants not compatible with the Mendelian inheritance model comprises removing a candidate variant that at least one of (a) presents in at least one unaffected male family member or male unaffected control as a

hemizygous variant, (b) presents in at least one unaffected female family member or female unaffected control as a homozygous variant, or (c) does not present in at least one affected family member.

27. The method of claim 14, wherein the list is an index list, and further comprising forming the index list, from a master list of more genetic variants than are in the index list, by the following steps:

- annotating a description of variants based on locations of at least one of (a) mutations in respective genes, or (b) amino acid alterations in affected proteins;

- annotating population frequencies of variants;

- annotating disease information associated with variants;

- annotating with evolutionary conservation indicators at variant positions; and

- annotating at least one prediction of a deleterious effect of at least one variant.

28. The method of claim 14, wherein the list is an index list, and further comprising forming the index list, from a master list of more genetic variants than are in the index list, by the following steps:

- from the master list, removing common variants that satisfy a user-defined threshold;

- from the master list, removing variants in at least one intergenic region; and

- from the master list, removing deep intronic variants and synonymous variants that do not have associated records in a selected database.

29. The method of claim 28, wherein common variants comprise single nucleotide polymorphisms (SNPs), deletions, insertions, and indels.

CLAIMS

1. A computer implementation system for assessing a causal genetic variant associated with a disease in a proband, comprising:

an input module that, by a processor, receives an input comprising (a) a selection of a Mendelian inheritance model based on a family history of the proband, and (b) the family history; the Mendelian inheritance model selected from the group consisting of autosomal dominant, autosomal recessive, and sex-linked;

a processing module that, by a processor

removes, from a list of genetic variants, variants not compatible with the selected Mendelian inheritance model;

removes from the list (a) variants that are present in unaffected controls below a first specified frequency, (b) variants that are present in unaffected controls above a second specified frequency, and (c) variants that are present in unaffected controls below a specified number of occurrences, each of the specified frequencies and number being determined by the input;

removes from the list variants that are in intergenic regions;

removes from the list deep intronic variants that do not have corresponding records in a database containing (a) known genetic variants and (b) diseases corresponding to the known genetic variants; and

removes from the list synonymous variants that do not have corresponding records in the database;

an identifying module that, by a processor, identifies the causal genetic variant associated with the disease, based on one or more remaining variants in the list; and

an output module that, by a processor, outputs the one or more remaining variants to a display.

2. The computer implementation system of claim 1, further comprising a determining module that, by a processor, determines a treatment plan based on the causal genetic variant.

3. The computer implementation system of any one of the above claims, wherein the identifying module is configured to identify the causal genetic variant by estimating a

probability that one of more of the remaining variants significantly influences at least one of clinical signs or symptoms of the proband.

4. The computer implementation system of any one of the above claims, wherein the input comprises a selection between a recessive model of Mendelian inheritance and a dominant model of Mendelian inheritance.

5. The computer implementation system of any one of the above claims, wherein the input comprises a selection between an autosomal model of Mendelian inheritance, an X-linked model of Mendelian inheritance, and a Y-linked model of Mendelian inheritance.

6. The computer implementation system of any one of the above claims, wherein the input comprises a selection of whether to allow de novo mutations.

7. The computer implementation system of any one of the above claims, wherein the list is an index list, the computer implementation system further comprising a forming module that, by a processor, forms the index list, from a master list of more genetic variants than are in the index list.

8. The computer implementation system of claim 7, wherein the forming module, by a processor:

from the master list, removes common variants that satisfy a user-defined threshold, the common variants comprising single nucleotide polymorphisms (SNPs), deletions, insertions, and indels.

9. A machine-readable medium comprising machine-readable instructions for causing a processor to execute a method comprising:

(1) receiving an input from a user, the input comprising (a) a selection of a Mendelian inheritance model based on a family history of a proband, and (b) the family history; the Mendelian inheritance model selected from the group consisting of autosomal dominant, autosomal recessive, and sex-linked;

(2) from a list of genetic variants, removing variants not compatible with the selected Mendelian inheritance model;

(3) from the list of variants, removing (a) variants that are present in unaffected controls below a first specified frequency, (b) variants that are present in unaffected controls above a second specified frequency, and (c) variants that are

present in unaffected controls below a specified number of occurrences, each of the specified frequencies and number being determined by the input;

(4) from the list of variants, removing variants that are in intergenic regions;

(5) from the list of variants, removing deep intronic variants that do not have corresponding records in a database containing (a) known genetic variants and (b) diseases corresponding to the known genetic variants;

(6) from the list of variants, removing synonymous variants that do not have corresponding records in the database; and

(7) identifying a causal genetic variant associated with a disease of the proband, based on one or more remaining variants in the list.

10. The machine-readable medium of claim 9, wherein the list is an index list, and wherein the method further comprises forming the index list, from a master list of more genetic variants than are in the index list.

11. The machine-readable medium of claim 10, wherein forming the index list comprises:

within the list of variants, annotating a description of variants based on locations of at least one of (a) mutations in respective genes, or (b) amino acid alterations in affected proteins;

within the list of variants, annotating population frequencies of variants;

within the list of variants, annotating disease information associated with variants;

within the list of variants, annotating with evolutionary conservation indicators at variant positions; and

within the list of variants, annotating at least one prediction of a deleterious effect of at least one variant.

12. The machine-readable medium of claim 10, wherein forming the index list comprises:

from the master list, removing common variants that satisfy a user-defined threshold, the common variants comprising single nucleotide polymorphisms (SNPs), deletions, insertions, and indels

.13. The machine-readable medium of any one of claims 9-12, wherein the input comprises:

a selection between a recessive model of Mendelian inheritance and a dominant model of Mendelian inheritance;

a selection between an autosomal model of Mendelian inheritance, an X-linked model of Mendelian inheritance, and a Y-linked model of Mendelian inheritance; and

a selection of whether to allow de novo mutations.

14. A computer-implemented method of assessing a genetic influence for a condition in a proband, the method comprising:

by a processor and from a list of genetic variants, removing variants not compatible with a Mendelian inheritance model determined by an input to the processor and based on a family history of the proband, the Mendelian inheritance model selected from the group consisting of autosomal dominant, autosomal recessive, and sex-linked;

from the list of variants, removing (a) variants that are present in unaffected controls below a first specified frequency, (b) variants that are present in unaffected controls above a second specified frequency, and (c) variants that are present in unaffected controls below a specified number of occurrences, each of the specified frequencies and number being determined by the input;

from the list of variants, removing variants that are in intergenic regions;

from the list of variants, removing deep intronic variants that do not have corresponding records in a database containing (a) known genetic variants and (b) diseases corresponding to the known genetic variants;

from the list of variants, removing synonymous variants that do not have corresponding records in the database; and

by a processor, identifying a causal genetic variant associated with a disease of the proband, based on one or more remaining variants in the list.

15. The method of claim 14, further comprising determining a treatment plan based on the causal genetic variant.

16. The method of any one of claims 14-15, wherein the identifying a causal genetic variant comprises estimating a probability that one of more of the remaining variants significantly influences at least one of clinical signs or symptoms of the proband.

17. The method of any one of claims 14-16, wherein the removing variants not compatible with the Mendelian inheritance model comprises comparing a genotype from the proband to a genotype from at least one family member.

18. The method of any one of claims 14-17, wherein the removing variants not compatible with the Mendelian inheritance model further comprises removing a heterozygous variant of the proband that exists in at least one unaffected family member as a homozygous variant.

19. The method of any one of claims 14-18, wherein the Mendelian inheritance model is a dominant model.

20. The method of claim 19, wherein the removing variants not compatible with the Mendelian inheritance model further comprises removing a heterozygous variant of the proband if it exists in at least one unaffected family member or does not exist in at least one affected family member.

21. The method of claim 19, wherein the removing variants that are present in unaffected controls comprises removing a candidate heterozygous variant that presents in at least one unaffected control as either a heterozygous or homozygous variant.

22. The method of any one of claims 14-21, wherein the Mendelian inheritance model is a recessive model.

23. The method of claim 22, wherein the removing variants not compatible with the Mendelian inheritance model comprises removing a candidate homozygous variant that at least one of (a) presents in at least one unaffected family member or unaffected control as a homozygous variant or (b) does not present in at least one affected family member.

24. The method of claim 22, wherein the removing variants that are present in unaffected controls comprises removing a candidate pair of compound heterozygous variants that present in at least one unaffected control.

25. The method of any one of claims 14-24, wherein the Mendelian inheritance model is a sex-linked recessive model.

26. The method of claim 25, wherein the removing variants not compatible with the Mendelian inheritance model comprises removing a candidate variant that at least one of (a) presents in at least one unaffected male family member or male unaffected control as a hemizygous variant, (b) presents in at least one unaffected female family member or female unaffected control as a homozygous variant, or (c) does not present in at least one affected family member.

27. The method of any one of claims 14-26, wherein the list is an index list, and further comprising forming the index list, from a master list of more genetic variants than are in the index list, by the following steps:

annotating a description of variants based on locations of at least one of (a) mutations in respective genes, or (b) amino acid alterations in affected proteins;

annotating population frequencies of variants;

annotating disease information associated with variants;

annotating with evolutionary conservation indicators at variant positions; and

annotating at least one prediction of a deleterious effect of at least one variant.

28. The method of any one of claims 14-27, wherein the list is an index list, and further comprising forming the index list, from a master list of more genetic variants than are in the index list, by the following steps:

from the master list, removing common variants that satisfy a user-defined threshold, the common variants comprising single nucleotide polymorphisms (SNPs), deletions, insertions, and indels.