

CLAIMS

What is claimed is:

1. A method for classifying a variant of interest (200), the method comprising:
executing, on one or more processors of a computer system, stored instructions for classifying the variant of interest (200), wherein the stored instructions, when executed cause the one or more processors to perform acts comprising:
receiving as an input a pathogenicity score (206) for a variant of interest (200) of a gene;
comparing the pathogenicity score (206) for the variant of interest (200) to a gene-specific pathogenic threshold (212); and
classifying (210) the variant of interest (200) as pathogenic in response to the pathogenicity score (206) exceeding the gene-specific pathogenic threshold (212).
2. The method of claim 1, wherein the pathogenicity score (206) is based on an extent of purifying selection in humans and non-human primates for the variant of interest (200).
3. The method of claim 1, further comprising:
determining a correlation between mean pathogenicity scores for pathogenic variants and percentiles of pathogenicity scores;
wherein the correlation between the mean pathogenicity scores for pathogenic variants and the percentiles of pathogenicity scores is used to define the gene-specific pathogenic threshold (212).
4. The method of claim 1, wherein the pathogenicity score (206) is generated based on an amino acid sequence processed by a neural network (102) trained to generate the pathogenicity score (206) from amino acid sequences.
5. The method of claim 4, wherein a central amino acid of the amino acid sequence corresponds to the variant of interest (200).
6. The method of claim 4, wherein the neural network (102) is trained using both human sequences and non-human sequences.

7. The method of claim 1, wherein the gene-specific pathogenic threshold (212) is in the range defined by and including the 51th percentile to the 99th percentile.
8. The method of claim 1, wherein the gene-specific pathogenic threshold (212) is in the range defined by and including the 75th percentile to the 99th percentile.
9. The method of claim 1, wherein the stored instructions, when executed cause the one or more processors to perform further acts comprising:
 - comparing the pathogenicity score (206) for the variant of interest (200) to a gene-specific benign threshold (218);
 - in response to the pathogenicity score (206) being less than the gene-specific benign threshold (218), classifying the variant of interest (200) as benign.
10. The method of claim 9, further comprising:
 - determining a correlation between mean pathogenicity scores for benign variants and percentiles of pathogenicity scores;
 - wherein the correlation between the mean pathogenicity scores for benign variants and the percentiles of pathogenicity scores is used to define the gene-specific benign threshold (218).
11. The method of claim 9, wherein the gene-specific benign threshold (218) is in the range defined by and including the 1st percentile to the 49th percentile.
12. The method of claim 9, wherein the gene-specific benign threshold (218) is in the range defined by and including the 1st percentile to the 25th percentile.
13. A non-transitory computer-readable medium storing processor-executable instructions that, when executed by one or more processors, cause the one or more processors to perform steps comprising:
 - processing a variant of interest (200) for a gene using a pathogenicity scoring neural network (102) to generate a pathogenicity score (206) for the variant of interest (200);

comparing the pathogenicity score (206) for the variant of interest (200) to one or more thresholds specific to the gene; and

classifying the variant of interest (206) as pathogenic, benign, or neither pathogenic nor benign based on the comparison of the pathogenicity score (206) to the one or more thresholds specific to the gene.

14. The non-transitory computer-readable medium of claim 13, wherein the one or more thresholds comprise a gene-specific pathogenic threshold (212) and wherein the variant of interest (200) is classified as pathogenic in response to the pathogenicity score (206) exceeding the gene-specific pathogenic threshold (212).

15. The non-transitory computer-readable medium of claim 13, wherein the one or more thresholds comprise a gene-specific benign threshold (218) and wherein the variant of interest (200) is classified as benign in response to the pathogenicity score (206) being below the gene-specific benign threshold (218).

16. The non-transitory computer-readable medium of claim 13, wherein the one or more thresholds specific to the gene comprise a gene-specific pathogenic threshold (212) and a gene-specific benign threshold (218) and wherein the variant of interest (200) is classified as neither pathogenic nor benign in response to the pathogenicity score (206) being between the gene-specific benign threshold (218) and the gene-specific pathogenic threshold (212).

17. The non-transitory computer-readable medium of claim 13, wherein comparing the pathogenicity score (206) for the variant of interest (200) to one or more thresholds specific to the gene comprises comparing the pathogenicity score (206) for the variant of interest (200) to a first threshold prior to a second threshold, wherein comparison to the second threshold is contingent upon the outcome of the comparison to the first threshold.

18. The non-transitory computer-readable medium of claim 13, wherein comparing the pathogenicity score (206) for the variant of interest (200) to one or more thresholds specific to the gene comprises comparing the pathogenicity score (206) for the variant of interest (200) to a first threshold and to a second threshold in parallel.

19. A variant classification system, comprising:
one or more memory structures encoding processor-executable instructions;
one or more processors in communication with the one or more memory structures,
wherein the one or more processors execute the instructions to perform operations comprising:
receiving as an input a pathogenicity score (206) for a variant of interest (200)
of a gene;
comparing the pathogenicity score (206) for the variant of interest (200) to one
or more thresholds specific to the gene; and
classifying the variant of interest (200) as pathogenic, benign, or neither
pathogenic nor benign based on the comparison of the pathogenicity score (206) to the
one or more thresholds specific to the gene.
20. The variant classification system of claim 19, wherein the one or more thresholds
comprise a gene-specific pathogenic threshold (212) and wherein the variant of interest (200)
is classified as pathogenic in response to the pathogenicity score (206) exceeding the gene-
specific pathogenic threshold (212).
21. The variant classification system of claim 19, wherein the one or more thresholds
comprise a gene-specific benign threshold (218) and wherein the variant of interest (200) is
classified as benign in response to the pathogenicity score (206) being below the gene-specific
benign threshold (218).