CLAIMS

1. A computer-implemented method comprising:

identifying nucleotide reads corresponding to a target genomic sample;

accessing a target-variant-reference panel including a customized data file comprising (i) marker-variant indicators for marker variants from an initial reference panel and (ii) target-variant indicators within a data field of at least one target-variant position added to the initial reference panel for phased alleles of genomic samples of different haplotypes, the target-variant indicators indicating a presence or absence of a target variant at a genomic coordinate within the genomic samples; and

imputing a genotype call for the target variant within the target genomic sample based on a comparison of the target-variant-reference panel and the nucleotide reads corresponding to the target genomic sample.

2. The computer-implemented method of claim 1, wherein the target-variant indicators indicate a presence or absence of the target variant in the at least one target-variant position for the phased alleles of the genomic samples.

3. The computer-implemented method of claim 1 or 2, further comprising:

determining phased alleles of the target genomic sample based on the comparison of the target-variant-reference panel and the nucleotide reads corresponding to the target genomic sample; and

imputing the genotype call by imputing a phased genotype call for the target variant within the target genomic sample based on the phased alleles of the target genomic sample.

4. The computer-implemented method of claim 1, wherein computing the genotype call for the target variant comprises generating a prediction of whether the target genomic sample comprises the target variant.

5. The computer-implemented method of claim 4, wherein generating the prediction comprises predicting whether the target genomic sample comprises a pathogenic variant at an allele present on a maternal haplotype or a paternal haplotype.

6. The computer-implemented method of any one of claims 1-5, wherein imputing the genotype call comprises:

identifying, within the nucleotide reads corresponding to the target genomic sample, one or more single-nucleotide polymorphisms (SNPs) as one or more marker variants within the target-variant-reference panel for the target variant; and

determining the genotype call further based on the one or more SNPs within the nucleotide reads.

7. The computer-implemented method of any one of claims 1-6, wherein imputing the genotype call for the target variant comprises imputing the genotype call for a repeat expansion.

8. The computer-implemented method of any one of claims 1-7, wherein imputing the genotype call comprises imputing the genotype call utilizing a genotype imputation model.

9. The computer-implemented method of any one of claims 1-8, wherein imputing the genotype call for the target variant comprises imputing the genotype call for a Replication Factor C Subunit 1 (RFC1) gene, a Cytochrome P450 Family 2 Subfamily D Member 6 (CYP2D6) gene, Cytochrome P450 Family 2 Subfamily B Member 6 (CYP2B6) gene, Cytochrome P450 Family 21 Subfamily A Member 2 (CYP21A2) gene, Survival Motor Neuron 1 (SMN1) gene, Survival Motor Neuron 2 (SMN2) gene, Glucosylceramides Beta (GBA) gene, Blood Group Rh(CE) (RHCE) gene, Lipoprotein(A) (LPA) gene, a Fragile X Mental Retardation 1 (FMR1) gene, a Hexosaminidase Subunit Alpha (HEXA) gene, Hemoglobin Subunit Alpha 1

(HBA1) gene, Hemoglobin Subunit Alpha 2 (HBA2) gene, or a Hemoglobin Subunit Beta (HBB) gene.

10. The computer-implemented method of any one of claims 1-9, wherein the target-variant-reference panel comprises marker-variant indicators for marker variants within a target genomic region for the target variant and does not comprise additional marker-variant indicators for additional marker variants outside of the target genomic region.

11. The computer-implemented method of any one of claims 1-10, wherein the at least one target-variant position comprises a target-variant position for target-variant indicators of a biallelic target variant or a multi-allelic target variant.

12. The computer-implemented method of any one of claims 1-8 or 10-11, wherein the target variant comprises a repeat expansion, a deletion, an insertion, a duplication, an inversion, a translocation, or a copy number variation (CNV) transmitted within a population.

13. The computer-implemented method of any one of claims 1-12, wherein the target variant satisfies one or more of a threshold carrier frequency, a threshold linkage disequilibrium (LD) with respect to particular marker variants, or a threshold mutation rate.

14. A system comprising at least one processor and a non-transitory computer readable medium comprising instructions that, when executed by the at least one processor, cause the system to perform a method according to any one of claims 1-13.

15. A non-transitory computer-readable medium storing instructions that, when executed by at least one processor, cause a computing device to perform a method according to any one of claims 1-13.