CLAIMS

1. A method, comprising:

a) ingesting a plurality of batches of sequence reads derived from a set of samples of genetic material extracted from a plurality of dogs, wherein each sample of genetic material comprises a raw genotype, and wherein each batch is associated with a respective dog among the plurality of dogs;

b) generating, for each batch, a plurality of phased haplotypes based on the sequence reads associated with the batch;

c) determining, for each batch and based on linkage disequilibrium in canids and haplotype diversity, a window size of 5 centimorgan (cM) for a plurality of windows of which at least one of the plurality of phased haplotypes is to be partitioned;

d) executing, for each batch, one or more machine learning models comprising a positional Burrows-Wheeler transform to partition the at least one phased haplotype into the plurality of windows, each of the plurality of windows being sized as 5 cM;

e) comparing, for each batch, the plurality of windows against a reference panel comprising a plurality of reference haplotypes associated with a plurality of reference populations of reference dogs;

f) generating, for each batch, one or more local assignments for one or more genetic populations of reference dogs based on the comparison;

g) determining, for each batch, a degree for smoothing the one or more local assignments for the one or more genetic populations of reference dogs based on whether the genetic material contains single-origin chromosomes or admixed chromosomes;

h) executing, for each batch, a hidden Markov model (HMM) to smooth the one or more local assignments for the one or more genetic populations of reference dogs based on the determined degree;

i) determining, for each batch and based on the one or more smoothed local assignments for the one or more genetic populations of reference dogs, one or more source populations associated with the dog associated with the batch;

j) partitioning, for each batch, the one or more smoothed local assignments for the one or more genetic populations of reference dogs into one or more of a maternally-inherited group or a paternally-inherited group;

k) identifying, for each batch and based on the one or more smoothed local assignments for the one or more genetic populations of reference dogs and the one or more source populations, a genetic ancestry of the dog associated with the batch and a plurality of genetic traits associated with the dog associated with the batch, the plurality of genetic traits

comprising at least (1) a predicted adult weight of the dog and (2) a risk prediction or a predisposition to a genetic skin disease; and

l) administering each of the plurality of dogs a respective treatment diet determined based on (1) a current weight and the predicted adult weight of the dog; and (2) the risk prediction or predisposition to the genetic skin disease, the treatment diet comprising energy density less than 4100 kcal/kg, an amount of fat less than 20% w/w, an amount of carbohydrates less than 25% w/w, an amount of protein more than 20% w/w, and an amount of dietary fiber more than 5% w/w, linoleic acid in an amount from about 7 g/Mcal to about 9 g/Mcal, and zinc in an amount from about 40 mg/Mcal to about 60 mg/Mcal.

2. The method of claim 1, wherein determining the one or more source populations comprises:

aggregating the one or more smoothed local assignments for the one or more genetic populations over both maternal and paternal chromosomes;

calculating proportions associated with the one or more source populations based on the aggregations; and

determining the one or more source populations based on the calculated proportions.

3. The method of claim 1, wherein the partitioning is based on one or more clustering algorithms.

4. The method of claim 1, wherein identifying the plurality of genetic traits is further based on one or more of genome-wide statistics, genomic principal component analysis (PCA) projections, DNA methylation profiles, or polygenic risk scores.

5. The method of claim 1, wherein the plurality of genetic traits further comprise one or more of:

a risk prediction or a predisposition to a genetic disease;

a nutrition recommendation;

a behavior and temperament class prediction;

a longevity estimation;

an all-causes mortality prediction in years;

a predicted pharmacological response; or

a recovery time range in hours for injectable anesthetics.

6. The method of claim 1, further comprising:

updating the one or more machine learning algorithms based on one or more new reference samples added to the reference panel.

7. The method of claim 6, wherein the updating comprises:

applying a cross-validation across all samples in the reference panel;

identifying, based on results associated with the cross-validation by a detection algorithm, one or more outliers; and

removing the identified one or more outliers from the reference panel.

8. The method of claim 6, wherein the updating is repeatedly iterated until a predetermined accuracy level of the one or more machine learning algorithms is reached.

9. The method of claim 6,

wherein the updating further comprises generating one or more labels for one or more unlabeled samples in the reference panel; and

wherein the updating is based on the generated one or more labels.

10. The method of claim 1, further comprising:

generating, based on the raw genotype, a consensus genotype, wherein generating the plurality of phased haplotypes is further based on the consensus genotype, and wherein generating each of the plurality of phased haplotypes comprises phasing the raw genotype and the consensus genotype into maternal and paternal chromosomes.

11. The method of claim 1, wherein executing the HMM to smooth the one or more local assignments for the one or more genetic populations of reference dogs comprises:

removing one or more errors associated with the one or more local assignments for the one or more genetic populations of reference dogs.

12. One or more computer-readable non-transitory storage media embodying software that is operable when executed to perform the method of any one of claims 1-11.

13. A system comprising: one or more processors; and a non-transitory memory coupled to the processors comprising instructions executable by the processors, the processors operable when executing the instructions to execute the method of any one of claims 1-11.