**2. Materials and methods**

ADAM uses a set of user-defined parameters to specify the genetic model, base population, population structure, selection and mating decisions for simulating breeding schemes. An overview of the software’s functionality and design is provided in Figure 1. As the purpose of this paper is showing multibreed functionalities, we will use breed instead of population through the paper.

### **Genetic model**

There are mainly two models available to generate breeding values for single and multiple traits. One option is the infinitesimal model, mimicking a polygenic makeup. Alternatively, the genomic model by which markers and QTLs are simulated by accounting for linkage disequilibrium (LD) between them. This model requires either a simulated founder population or real genome-wide SNP/sequence data. For the genomic model, it is necessary to either simulate founder haplotypes within ADAM, as detailed in the ADAMplant paper and applied in Jørn’s paper or import real phased haplotypes from genome-wide SNP/sequence data collected from real data or simulations. These haplotypes serve as gametic pool from which the genotypes of the base population are sampled. Specifically, the genotype of each base individual is sampled from the pool of chromosomes in generation of the founder population by simulation or real phased haplotypes. For chromosome ( = …), two chromosomes are randomly sampled without replacement from the th pool of chromosomes and then replaced before the next sampling.

### **Trait simulation**

When an infinitesimal model is used, the true breeding values (TBV) of individuals in the base population are sampled from a normal distribution with a user-specified additive genetic variance for the trait or a multivariate normal distribution for multiple traits. When genomic model is used, QTL effects are sampled to fit the desired genetic variance and covariance between simulated traits. The feature of dominance effect has been added to the genomic model. The procedure to simulate additive effects (), dominance effects (), TBV, dominance deviation (DDV) and total genetic effects for multiple traits are as follows:

1. The additive effect () at locus is sampled from a multivariate normal distribution with mean 0 and user-defined additive genetic variance-covariance matrix where .
2. The dominance effect () at locus is calculated based on a user-defined mean dominance degree and its associated standard deviation. The model described by Wellmann and Bennewitz (2011) is used to account for the contribution of dominance effects to the total genotypic value at each locus. The dominance degree at locus is sampled as follows:

,(3)

(4)

where is the user-defined mean of , is the variance, and is a random number.

The dominance effect at locus is computed as:

.

1. The substitution effect at locus is calculated as , where and are the frequencies of the alleles at locus .
2. The for individual is computed as follows. Denote the genotype of individual at locus as or represents the number of copies of the alternative allele):

The current genetic variance computed as the variance of current breeding values is scaled to satisfy the user-defined additive genetic variance by a square of scaling factor calculated as

,

then the original , and are scaled according to this scaling factor , and the is calculated again after the scaling is done. Then can match the user-defined across the traits.

1. The for individual is computed as:

The TBV and DDV of an animal are calculated depending on the specific time and the breed or herd in which the estimation is conducted. The output includes additive effect, dominance effect and total genetic value, which is the sum of the additive and dominance effects across all loci. These values are assessed at both individual level and population level.