

## Simulation of animal crossbreeding schemes using ADAM software

H. Liu<sup>1,2,\*</sup>, H.M. Nielsen<sup>2</sup>, M. Kargo<sup>2</sup>, J.R. Thomasen<sup>3</sup>, P. Berg<sup>2,4</sup>, M. Slagboom<sup>2</sup>, A.C. Sørensen<sup>5</sup>, V. Milkevych<sup>2</sup>

<sup>1</sup>SEGES, Agro Food Park 15, DK 8200 Aarhus N, Denmark; <sup>2</sup>Center for Quantitative Genetics and Genomics, Aarhus University, Blichers Allé 20, 8830 Tjele, Denmark; <sup>3</sup>VikingGenetics, Ebeltoftvej 16, DK-8960 Randers SØ, Denmark; <sup>4</sup>Norwegian University of Life Sciences, Universitetstunet 3, 1433 Ås, Norway; <sup>5</sup>The Danish Agriculture and Food Council, Agro Food Park 15, DK 8200 Aarhus N, Denmark; \*hliu@seges.dk

### Abstract

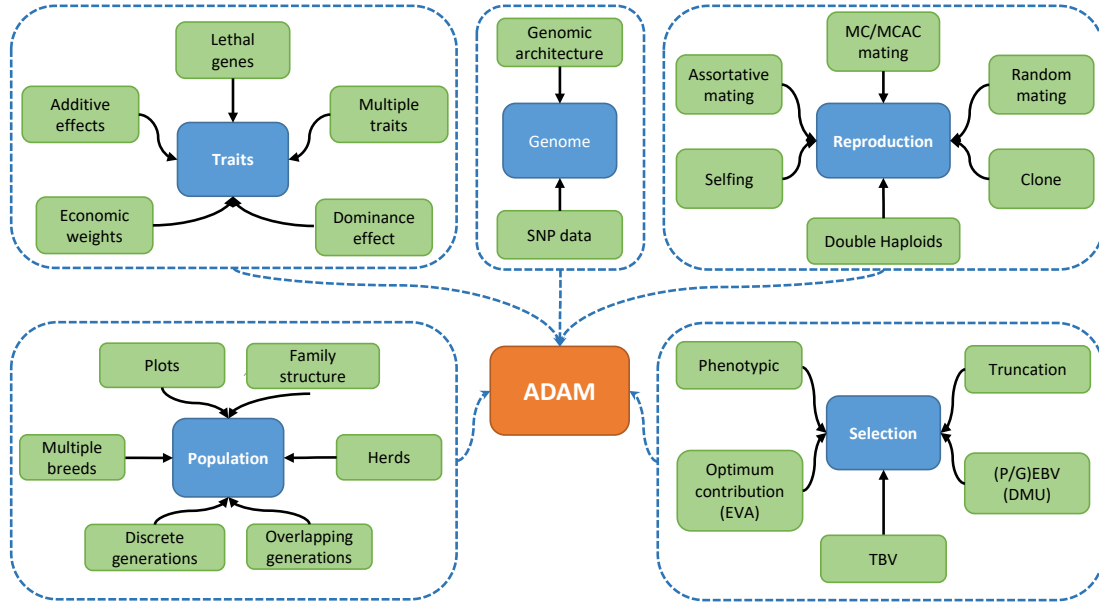
ADAM software is a stochastic simulation tool intended to model a variety of selective breeding schemes for animals and plants. ADAM can be used to generate genetic data to test statistical tools and the implemented algorithms can evaluate breeding schemes by tracing genetic change under different selective breeding scenarios. In practice, crossbreeding has been used extensively in beef cattle, pig, and poultry production systems, and the use of crossbreeding is increasing in dairy cattle breeding system due to changes in the dairy market. The aim of this paper was to present a new crossbreeding feature of the ADAM software. The software was updated to include multiple purebred populations and crossbred populations in animals with different genetic backgrounds considering both additive and dominance effects. Our simulated results showed that, by using the new features in ADAM, breeding organizations can optimize their design of crossbreeding schemes and, thereby, optimize profitability.

### Introduction

The well-known effect of increased performance of crossbreds compared to purebreds has long been used systematically in livestock species such as pig and poultry (Therkildsen et al. 2021; Chen et al. 2021; Han et al. 2020). In dairy cattle, farmers have started to use crossbreeding to utilize the increased performance of crossbreds due to heterosis (Clasen et al. 2017). However, tools for the use of crossbreeding are lacking; including tools to predict genetic merit of crossbreds, tools to manage breeding of several breeds simultaneously, and decision tools for farmers to optimally use crossbreeding in a herd.

One of the major decisions to make when managing crossbreeding is which breeding scheme is optimal. Therefore, efficient computational tools that allow stochastic simulation of a variety of breeding schemes are of high importance. Among the existing software (Faux et al. 2016; Pook et al. 2020) ADAM is a well-established tool intended to model a variety of selective breeding schemes for animals and plants (Pedersen et al. 2009; Liu et al. 2019). The software simulates populations and traces genetic changes in these populations under different scenarios. ADAM accounts for different population structures and genetic models. It can be used to test different phenotyping/genotyping strategies, estimate breeding values, evaluate different selection and mating strategies and generate genetic data to test statistical tools. It is a compiled software written in FORTRAN.

The purpose of this paper is to present the novel features of the ADAM software dedicated to simulation of multiple populations of (possibly) different genetic backgrounds including dominance genetic effects, demonstrated in dairy cattle crossbreeding scheme. The implemented features naturally extend the existing ADAM functionality (Fig. 1) to the case of multiple interacting populations and allow efficient evaluation of crossbreeding effects at genetic and phenotypic levels.



**Figure 1.** General overview of the ADAM software functionality. MC is minimum co-ancestry; MCAC is mating animals by minimizing the covariance between ancestral contributions; TBV is true breeding values.

## Materials & Methods

**Multiple interacting populations.** The model of interacting populations, considered in the context of crossbreed selection, is represented here in terms of evolutionary operators following the approach from Paixão et al. (2015). Given a finite population  $P(t) = (G^n, \varphi, S_{(\iota, \lambda)})$  of size  $n$  at time  $t$  characterized by a finite collection  $G^n = \{g_i\}$  of genotypes, where  $g_i, i = \overline{1, n}$ , genotype-phenotype mapping  $\varphi$ , finite collection  $S_{(\iota, \lambda)} = \{s_j\}$  of selection candidates  $s_j, j = \overline{1, m}$ , where  $\iota$  are animals selected for mating and  $\lambda$  are their offspring. Evolution of population  $P$  (breeding) in general form:

$$P(t+1) = S_{(\iota, \lambda)}(P(t) \cup (V \circ S_{(\iota, \lambda)})P(t)), \quad (1)$$

where  $P(t+1) = (G^l(t+1), \varphi, S_{(\iota, \lambda)})$  is a finite population of size  $l$  at time  $t+1$ ;  $V: G^n \rightarrow G^l$  is a variation operator (mutation, recombination, etc.);  $\cup, \circ$  are union and composition operations respectively.

Let  $a, b, c$  denote two pure breeds ( $a, b$ ) and one crossbred ( $c$ ) populations. Then, the model of interacting populations (crossbreeding):

$$\left\{ \begin{array}{l} P_a(t+1) = S_{(\iota \in a, \lambda \in a)}(P_a(t) \cup (V \circ S_{(\iota \in a, \lambda \in a)})P_a(t)), \\ P_b(t+1) = S_{(\iota \in b, \lambda \in b)}(P_b(t) \cup (V \circ S_{(\iota \in b, \lambda \in b)})P_b(t)), \\ P_c(t+1) = S_{(\iota \in c, \lambda \in c)}(P_c(t) \cup (V \circ S_{(\iota \in c, \lambda \in c)})\{P_a(t), P_b(t)\}), \\ P_a(0) = n_{a0}, \\ P_b(0) = n_{b0}, \\ P_c(0) = 0, \end{array} \right. \quad (2)$$

where  $n_{a0}, n_{b0}, n_{c0}$  are sizes of populations  $a, b$  and  $c$  at the time  $t = 0$  respectively.

**Dominance.** The following sampling was implemented in ADAM to allow a contribution of dominance effects to the total genotypic value (Wellmann and Bennewitz 2011):

$$\delta = \mu + k \cdot \sqrt{\sigma^2}, \quad (3)$$

$$d = \delta \cdot |a|, \quad (4)$$

where,  $\delta$  is a dominance degree,  $\delta \sim N(\mu, \sigma^2)$ ;  $\mu$  is a mean of  $\delta$  with a default value of 0.193;  $\sigma^2$  is a variance of  $\delta$  with a default value of 0.312 (Bennewitz et al. 2017);  $k \sim N(0,1)$  is a random number;  $d$  is a dominance contribution to a genetic value at each locus;  $a$  is an additive contribution to a genetic value at each locus.

**Data and simulation studies.** In order to demonstrate the effect of the novel features implemented in the software, we performed a simple case study simulating a dairy cattle crossbreeding scheme. ADAM can either generate founder haplotypes with user-defined settings for historical populations or read existing haplotypes and genetic architecture into the program. The haplotypes were randomly sampled from the genotype data as used in Thomassen et al., (Thomassen et al. 2020) to resemble the Danish dairy cattle populations. A total of 40,000 SNP markers and 1,000 QTL were randomly distributed in the genome. A trait with heritability of 0.5, additive genetic variance of 1.00 and dominance variance of 0.25 were considered.

**Simulated breeding scheme.** A conventional progeny-testing scheme was performed in three purebred populations (population 1-3) in generation 1-10. In each generation and each population, 200 young bulls were progeny-tested and 20 proven bulls (5- to 6-year-old) were truncation-selected for breeding. In addition, 3,000 cows (1- to 5-year-old) were selected out of 10,000 cows for breeding. All progeny-tested bulls were genotyped and included in the sire reference population. In generation 11 to 13, all males and females were genotyped, and a three-way rotational crossbreeding scheme was performed. In each generation, 10 purebred males were mated to 1,000 crossbred females, and the crossbred animals are considered to be in a crossbred population (population 4). The breeding scheme simulation was replicated 5 times. The average additive and dominance effects in the population were evaluated for purebred animals in generation 0 - 10 in population 1 (representing a purebred population) and for crossbred animals in generation 11 to 13 in population 4.

## Results and Discussion

The simulation results for the selected breeding scheme are depicted in Figure 2, where the trends of average additive and dominance effects across generations are shown for generation 0 - 10 in population 1 population and generation 11 – 13 in population 4. The rate of genetic gain rapidly improved due to crossbreeding in generation 11 - 13 (0.44 in genetic standard deviation unit per generation) compared to generation 1 - 10 (0.24 per generation). The dominance effect decreased during generation 1 - 10 due to selection within purebred populations and increased in generation 11 as the level of heterozygosity increased due to crossbreeding. The dominance effect possibly fluctuated across generations due to a small number of replicates. The introduction of crossbreeding in generation 11 to 13 increased the additive genetic variance from 0.74 to 0.88 (Figure 2).

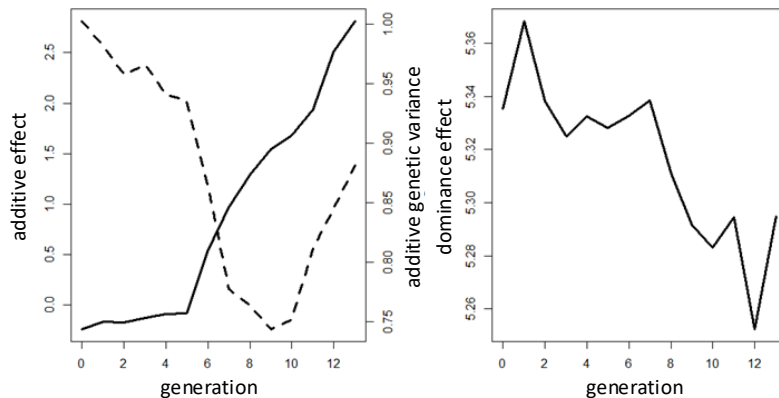


Figure 2. Additive effect (full line), additive genetic variance (broken line) and dominance effect across 13 generations of the simulated breeding schemes.

This work represents the novel advanced features of ADAM software. By using a simple test case we showed that the software can simulate multiple interacting populations within one breeding scheme while accounting for dominance effects. With the new functionalities, ADAM can simulate a broad range of research-based as well as commercial-based breeding and crossbreeding schemes of variable complexity. By using ADAM, breeding organizations have the opportunity to optimize their breeding strategies and use of resources and, thereby, optimize profitability.

## References

- Bennewitz, J., C. Edel, R. Fries, T.H.E. Meuwissen, and R. Wellmann (2017) *Genet. Sel. Evol.* 49 (1):7. <https://doi.org/10.1186/s12711-017-0284-7>
- Chen, C., J. Zhu, H. Ren, Y. Deng, X. Zhang et al. (2021) *Livest. Sci.* 244:104342. <https://doi.org/10.1016/j.livsci.2020.104342>
- Clasen, J.B., E. Norberg, P. Madsen, J. Pedersen, and M. Kargo (2017) *J. Dairy Sci.* 100 (8):6337-6342. <https://doi.org/10.3168/jds.2017-12627>
- Faux, A.-M., G. Gorjanc, R.C. Gaynor, M. Battagin, S.M. Edwards et al. (2016) *Plant Genome* 9 (3). <https://doi.org/10.3835/plantgenome2016.02.0013>
- Han, Z., Q. Li, S. Liu, and L. Kong (2020) *Aquac.* 529:735621. <https://doi.org/10.1016/j.aquaculture.2020.735621>
- Liu, H., B.B. Tessema, J. Jensen, F. Cericola, J.R. Andersen et al. (2019) *Front. Plant Sci.* 9 (1926). <https://doi.org/10.3389/fpls.2018.01926>
- Paixão, T., G. Badkobeh, N. Barton, D. Çörüş, D.-C. Dang et al. (2015) Toward a unifying framework for evolutionary processes. *J. Theor. Biol.* 383:28-43. <https://doi.org/10.1016/j.jtbi.2015.07.011>
- Pedersen, L.D., A.C. Sørensen, M. Henryon, S. Ansari-Mahyari, and P. Berg (2009) *Livest. Sci.* 121 (2):343-344. <https://doi.org/10.1016/j.livsci.2008.06.028>
- Pook, T., M. Schlather, and H. Simianer (2020) *G3: Genes Genomes Genet.* 10 (6):1915-1918. <https://doi.org/10.1534/g3.120.401193>
- Therkildsen, M., S.K. Jensen, and A. G. Kongsted (2021) *Livest. Sci.* 250:104584. <https://doi.org/10.1016/j.livsci.2021.104584>
- Thomassen, J.R., H. Liu, and A.C. Sørensen (2020) *J. Dairy Sci.* 103 (1):597-606. <https://doi.org/10.3168/jds.2019-16974>
- Wellmann, R., and J. Bennewitz (2011) *Genet. Res.* 93 (2):139-154. <https://doi.org/10.1017/S0016672310000649>