Throughout the document, I assume that we have established a base population (G0), either through real genotypes or simulations. Otherwise, we can simply simulate founder/historical populations (generations say -1000 to 0) to reach a base population relatively easily. The details of the genome is not so important, so I just excluded them from here. Just have a look, and we can discuss further if you want.

There is a great interest now on crossbreeding schemes, and I’ve simulated 4 different data sets for 4 projects/papers including mine (soon to be published in GSE), using my own Julia implementation. I would definitely use XSim, if it was available/straightforward and simple to run without user tricks. Yet, not many softwares are available for simulating crossbreeding schemes. Since you have already shown, in your earlier XSim paper, most of the details of your program, perhaps focusing on crossbreeding is a good strategy. Up to you, but I also added a simple more general simulation scenario below. Coincidently, the software’s name has a direct link to “cross breeding” with its “X” at the beginning ☺

I will also develop a software package for the models I implemented in Julia for an EU project, and will try to organize a really short (perhaps one day long) workshop about how to analyze crossbred data together with multi-breed data, around next August for the final project meeting. The workshop will include my software, and a single-step implementation in another commercial software for large scale analyses. We have lots of partners from all over the Europe there, and perhaps we can integrate your simulation software to the program, so that it gets complete from data to large scale industry analysis. This may also give your excellent work more visibility within EU institutions. Of course, depends on your decision of whether we pay special attention to simulation of crossbred data here ☺

**1) A simple livestock example**

* n generations
* sire:dam ratio
* male:female ratio
* n\_offspring per dam

**Example scenario:** 10 generations of individuals with 20 sires per generation, and 10 dams per sire, 2 offspring per dam with 1:1 male female ratio (1 male 1 female). This leads to 400 individuals at each generation, of which 200 is males and 200 is females. The 20 sires can be either selected randomly (with or without replacement, to allow some sires to be heavily used) or based on some phenotype.

**2,3) A rotational cross-breeding and a multi-breed example**

* n breeds and rotation order
* breed sizes (n male + n female), also for crossbred population
* n generations
* sire:dam ratio (per breed. For crossbreds, sire should come from pure breed)
* male:female ratio (per breed, and for crossbreds)
* n\_offspring per dam (per breed, and for crossbreds)
* same set of sires across parallel pure breed simulations and crossbred populations???

**Example scenario:** 10 generations of individuals. Parallel purebred populations are simulated as well as a crossbred population. 1 small 2 large pure breeds, and a crossbred (X) population, for example. Breed 1 has 50 males 500 females, Breeds 2-3 have 100 males 2000 females at G0.

*Breed 1 (small breed):* 50 sires per generation, and 10 dams per sire, 2 offspring per dam with 1:1 male female ratio (1 male 1 female). This leads to 1,000 individuals at each generation, of which 500 is males and 500 is females. The 50 sires can be either selected randomly (with or without replacement, to allow some sires to be heavily used) among the males, or based on some phenotype.

*Breeds 2,3 (large breeds):* 100 sires per generation, and 20 dams per sire, 2 offspring per dam with 1:1 male female ratio (1 male 1 female). This leads to 4,000 individuals at each generation, of which 2,000 is males and 2,000 is females. The 100 sires can be either selected randomly (with or without replacement, to allow some sires to be heavily used) among the males, or based on some phenotype. The critical point here may be the selection of sires. The same set of sires can be used when selecting sires for pure and crossbreds, or they can be selected separately at random.

*Breed X:* Sires come from pure breed and dams come from crossbred population at each rotation cycle (except from round 1, obviously), so n sires cannot be greater than the number of sires at the smallest breed (if selection is done among sires used in pure breed without replacement). Or the sires cannot be smaller than the number of males available, regardless of they are used as sires in pure breed or not.

*Rotation Scheme*

Suppose that males from Breed A and females from Breed B used. Population size set for crossbreds controls the number of offspring. Let’s say we set population size at 2,000 for the crossbred population, and sire:dam ratio controls the sires from pure population and dams from the crossbreds (though, Breed B in round 1). Suppose we set sire:dam at 1:10, and n offspring at 2 with 1:1 male female ratio (1 male 1 female).

* (G1) Bred 2 males and Breed 3 females at first round: 100 sires from Breed 2, and 10 dams per sire from Breed 3 (1,000 dams), 2 offspring per dam with 1:1 male female ratio (1 male 1 female). This leads to 2,000 individuals, of which 1,000 is males and 1,000 is females.
* (G2) Breed 1 sires and females of G1 crosses are used at second round: Still can use same 50 sires as in parallel pure breed, or can select among the available 500 males. For the **former**, the software should decide based on the male:female ratio for crossbred population: 50 sires, and 20 dams per sire (1,000 crossbred dams from G1 crosses). With our setting of 2 offspring per dam with 1:1 male female ratio (1 male 1 female), this leads to 2,000 individuals, of which 1,000 is males and 1,000 is females. For the **latter**, better to set the number of sires at the size of available sires (50 in this case), to avoid any problems.
* (G3) Breed 3 sires and females of G2 crosses are used at second round: Still can use same 100 sires as in parallel pure breed, or can select among the available 1,000 males. For the **former**, the software should decide based on the male:female ratio for crossbred population: 100 sires, and 10 dams per sire (1,000 crossbred dams from G2 crosses). With our setting of 2 offspring per dam with 1:1 male female ratio (1 male 1 female), this leads to 2,000 individuals, of which 1,000 is males and 1,000 is females. For the **latter**, better to set the number of sires at the size of available sires (100 in this case), to avoid any problems.
* Continues with Breed 2 sires and so on…

**4) An admixed population example**

* n main breeds
* n outside breed (mainly a breed that is introduced at some point into the population, but maybe not so regularly)
* n generations
* generations to introduce outside breed
* prop sires to be used from the outside breed(s)
* dam:sire ratio (per breed, and for crossbreds)
* male:female ratio (for crossbreds)
* n\_offspring per dam (for crossbreds)
* same set of sires across parallel pure breed simulations and crossbred populations???

**Example scenario:** 10 generations of individuals. Parallel purebred populations are simulated as well as a crossbred population. 1 small 2 large pure breeds, and a crossbred (X) population, for example. Breed 1 has 50 males 500 females, Breeds 2-3 have 100 males 2000 females at G0.

Very similar to above, except that from F1 we simulate F2…..FN. So the populations become admixed. At some generation(s), we introduce a third breed, either permenantly or temporarily. If for instance, we set prop sires to be used from outside breed at 100% at G2, than it turns to be a rotational crossbreeding as above.