Chapter 9

Use of Simulated Annealing Algorithms for Optimizing Selection Schemes in Farm Animal Populations

Edo D'Agaro

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/50665

1. Introduction

The design of optimal mating schemes is a mean to improve farm animal performances. During the last decades, breeding strategies and techniques addressing both genetic improvement and inbreeding control have been well documented and applied in several countries [1]. The detrimental effects of inbreeding have been reported in farm animals and, in the recent years, many selection and mating strategies were proposed to restrict inbreeding in selection programmes [2]. Recent advances in animal breeding theory have clearly shown the importance of mating design optimization by means of new analytical models as the optimum contribution selection method [3] and simulated annealing algorithms [4]. Stochastic simulation programs are generally used to create farm animal populations under artificial selection, and, by this way, genetic and inbreeding effects are easily modelled and studied for several generations. In this study, a stochastic simulation (Monte Carlo method) was used to evaluate and optimize different mating schemes of farm animals under a restricted inbreeding rate.

2. Breeding strategies in farm animal populations

2.1. Selection based on phenotype

The selection of individuals based on the phenotype has been a breeding strategy widely practiced over time as it allowed to obtain significant benefits in the economic sector. Most of these selection schemes have been applied by isolating a specific phenotype, ignoring the genetic structure of selected traits [5]. However, in practice, selection programs based only on phenotypes have shown effects both positive and negative. In fact, although some

significant improvements were observed in animal production and reproduction performances, at the same time, several undesirable characteristics were also selected.

2.2. Marker assisted selection

In recent years, the importance of molecular genetics to understand the genetic nature of quantitative characters has worldwide been recognized, identifying specific regions of genes or chromosomes that affect production and reproduction traits [5]. For example, there are some traits, such as the resistance to a specific pathogen, that can be studied only by a selection method based on the genotype or studying the correlation between the phenotype and genotype. Different types of molecular markers can be used to identify specific gene variants and, a marker assisted selection scheme or MAS (Marker Assisted Selection) implemented in a population. The MAS is a direct selection technique which is based on the association between a trait and several molecular markers. This technique allows to select at a very early stage of development, since it is not necessary to wait for the phenotypic expression of the trait. To date, for many species, a very large number of molecular markers or sequences of DNA are available. An important goal, realized for many species of commercial importance and under way for others, is to set up a map of the genome, identifying several molecular markers and then use them in association mapping, linkage analysis or QTL (Quantitative Trait Loci) studies. The technique of QTLs mapping, based on quantitative genetic laws and molecular methods, allows to associate, for a quantitative trait one or more genetic markers. In this way, for example, it is possible to find markers associated with resistance to a certain pathology or high growth rates. However, in order to obtain a successful QTL analysis, it is necessary to use a large number of polymorphic molecular markers linked to measurable and heritable traits. So, the ideal situation is to perform the analysis in a population showing a high degree of polymorphism and high variability in the genes that control the expressed phenotypes. All individuals of the segregating population are identified for both the molecular markers and quantitative traits.

2.3. Genomic maps

Genomic maps are used in order to get more information concerning the genome of individuals of a given species, describing the order in which genetic loci or markers are displaced and the distance between them on each chromosome. There are two ways to map the genome, using physical or genetic (linkage) methods. Maps are a useful tool for the isolation and cloning of genes of interest.

2.4. Physical maps

A physical map is set up to show the position of specific genes. A physical map consists on a set of markers or physically identifiable regions of DNA and is constructed without using the recombination analysis between genes. The main role is to measure the order and distance between two markers. Physical maps can have different resolutions. For example,

the location of a marker on a specific chromosome is given by the hybridization technique of somatic cells. By this way, it is possible to produce a chromosome map in which each chromosome is characterized by a particular banding, observable after staining under a microscope. Another type of physical map has a medium-high resolution, allowing to make eukaryotic metaphase fluorescent chromosomes or specific DNA sequences, and DNA specific fluorochrome-labeled (FISH, fluorescence in situ hybridization). The third type of physical maps has a high resolution of thousands of STS (Sequence-Tagged Site), that defines unique portions of the genome. The STS are short segments of DNA, long approximately 60-1000 bp, which represent points of reference in the genome.

2.5. Maps of association or linkage

Association maps maps show the distances between various genes, their position and other features. Distances between genes are determined by the frequency by which two markers, located on the same chromosome, are inherited together. Alleles which are very closed, they have a higher probability of being transmitted together than those found on distant loci. A unit of genetic map or cM (centimorgan) represents the distance between two genes (1% of recombination). A linkage map, then, defines the distance between markers and their positions on the genome, determining the frequency by which two markers are associated. Maps that use genes as markers show generally a low density and therefore are not always informative. In the construction of a linkage map, DNA sequences should be preferably used.

2.6. Genomic selection

More recently, the availability of high-throughput sequencing techniques has allowed to discover in several livestock species, thousands of Single Nucleotide Polymorphisms (SNPs) spread across the whole genome. Currently, beadchips for genotyping bovines at more than 750,000 marker loci are commercially available. Such a map density is enough to find Linkage Disequilibrium (LD) between markers and QTLs and, by this way looking for associations between traits and markers without specific knowledge of the population structure. These new techniques give rise to new perspectives for the genetic evaluation of farm animals with a so called genome-wide approach. On one hand, this new advance allows to explore the genome looking for QTLs and associations between SNPs and phenotypes. On the other hand, it allows to use directly the marker information to estimate the genomic breeding value (GEBV). In the former case, we talk about the genome-wide association (GWA) studies, while in the latter, the term genomic selection (GS) is generally adopted. Briefly, the GS rely on the segmentation of the genome using a dense marker map in several thousands of bits, each contributing to the explanation of part of the genetic variance of a quantitative trait. The effect of each segment is estimated in a reference population (animals with known phenotypes and genotypes). SNPs effects are then used to predict the breeding values of another set of genotyped animals (prediction population) without phenotypes. Meuwissen et al. in 2001 [6] proposed to use dense marker information

to predict the breeding values of animals. Afterwards, several of models and approaches – mainly on simulated dataset – have been proposed to solve the main statistic issue of practical implementation of GS: the great asymmetry of data matrix i.e., the number of effects (single markers or haplotypes) which is much greater than the number of phenotypic records available. In brief, potential advantages of using high density markers in genomic evaluation are the following: i) each QTL is expected to be in LD with at least one marker; ii) all the genetic variance is taken into account in the estimation of breeding values; iii) the animals can be genotyped early in life, and this may guarantee a reduction of generation interval; iv) a better estimation of mendelian sampling (deviation of the individual from the average family effect) term may give rise to a lower inbreeding rate. Genotypes of a particular marker provide a direct information on variability at the locus and frequently at the closely linked loci. When the marker map is not very dense, we may get a biased measure of the variation of the non-genotyped part of the genome. There is an active development of new molecular genetic technology allowing for high-throughput genotyping. Dense marker maps cover the entire genome giving a detailed picture of the genetic variability. This technology has also facilitated the detection of important regions or loci with adaptive effects [7]. Loci could be studied further over breeds and individuals using a technique called re-sequencing [8]. As the technology in molecular genetics advances, it is very likely that sequencing of the whole genome of individuals will soon replace the marker typing. This would result in increments in the accuracy of the estimation of genomic variation and, correspondingly, in the power of strategies devoted to the management of the genetic diversity (and also in selection efficiency). Over generations, alleles at different loci are recombined. If population size has stayed large over a long period, there has been time to produce recombinations even over a very narrow genome area. On the other hand, in a very small population, variants tend to be transmitted over longer genome stretches. Such blocks would therefore indicate a small population size (bottleneck) in the recent history of the population [7]. Furthermore, considering different populations within species, allele frequency differences are used to quantify relationships (through the calculation of different genetic distances) among all groups [9-10].

3. Calculation of inbreeding and additive relationship coefficients in farm animals and small populations

3.1. Calculation of the inbreeding coefficient

Alleles at one locus can be classified into two categories: alleles identical in structure (IS) and identical by descent (ID). The inbreeding coefficient (F) of an animal is defined as the probability that both alleles at a locus are identical by descent (copies of the same allele present in a common ancestor) [11]. The presence of a common ancestor is a key element. Figure 1 shows a diagram of half sibs mating.

F_u coefficient can be computed as the probability F of the animal U to get two copies of an allele from a common ancestor. In the example, animals W and V can have each four

different genotypes and inherit the same allele (A_3 or A_4) from the common grandfather (Z) but not from the grandmothers (X and Y). Animal U receives one allele from each parent and so we can get 16 different genotypes (A_1A_3, A_1A_4, A_1A_5, A_1A_6, A_2A_3, A_2A_4, A_2A_5, A_2A_6, A_3A_3, A_3A_4, A_3A_5, A_3A_6, A_3A_4, A_4A_4, A_4A_5, A_4A_6). In the example of half sibs, the condition in which both alleles at a locus are identical by descent is satisfied only for the genotypes A_3A_3 and A_4A_4. The probability to obtain these genotypes is equal to two out of 16 possible combinations (2/16 or 1/8).

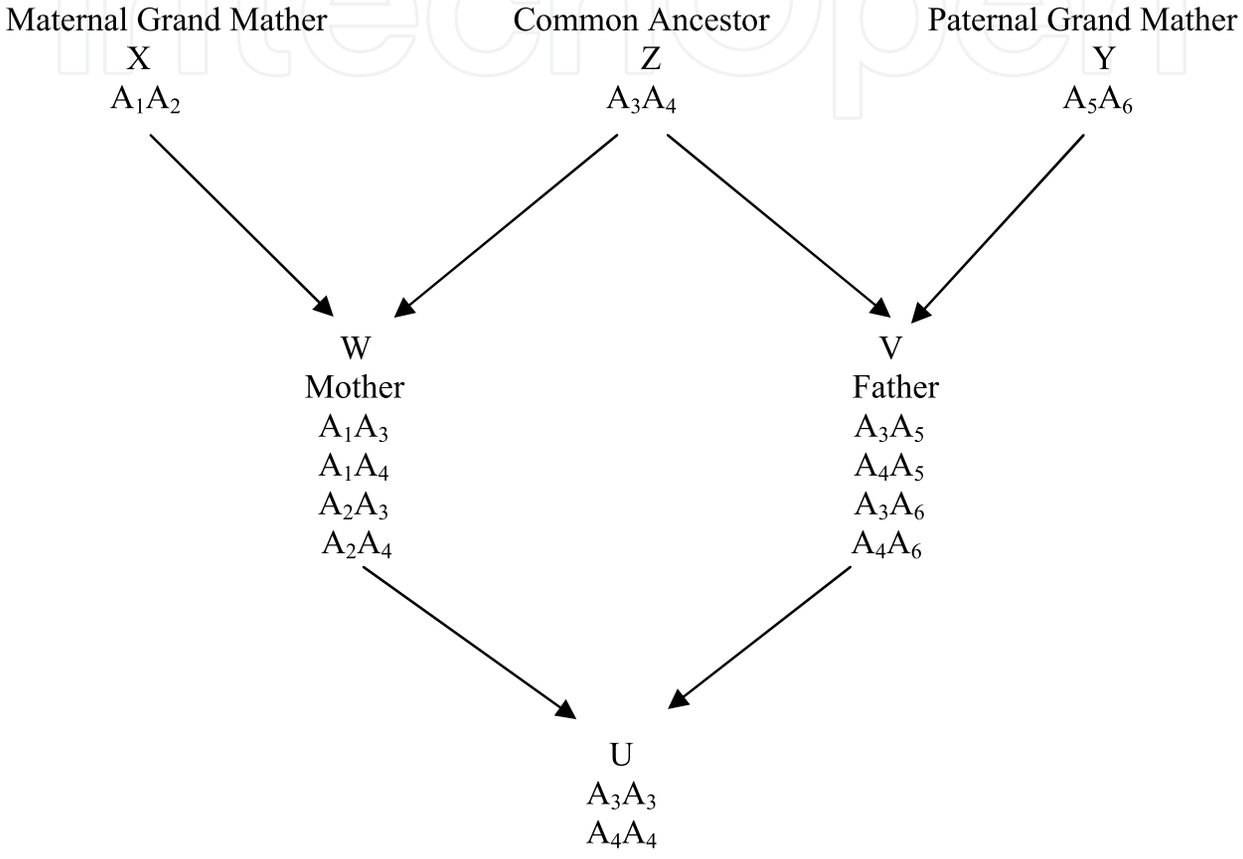


Figure 1. Example of half sibs mating

F coefficient can be computed as follows:

1. for one offspring, the probability to inherit one allele from his father is equal to 1/2;
2. the result of gamete segregation is independent to other segregations that occur at the same or in previous generations. Fu coefficients can be computed as: for allele A_3, the probability that the animal U inherits the allele A_3 from Z and W is equal to: 1/2 x 1/2 = 1/4. Similarly, the probability that animal U inherits the same allele via the Z * V * U path is equal to 1/4. The probability that U inherits the A_3 allele from both parents is equal to the product of the two probabilities:

$$P(A_3A_3) = (1/2 \times 1/2) \times (1/2 \times 1/2) = (1/2)^4 = 1/16$$

and the probability that U inherits the allele A_4 is equal to:

$$P(A_4A_4) = (1/2 \times 1/2)\ (1/2 \times 1/2) = (1/2)^4 = 1/16$$

As the two genotypes A_4A_4 and A_3A_3 are mutually exclusive:

$$P\ (U = ID) = (1/2)^4 + (1/2)^4 = (1/2)^3 = 1/8$$

Fu coefficient is equal to 1/8. Note that this result is equal to 1/2 powered 3.

In general, the coefficient of inbreeding of an animal U is computed using the following formula:

$$F_U = \sum \left(1 + F_Z\right)\left(\frac{1}{2}\right)^n \quad (1)$$

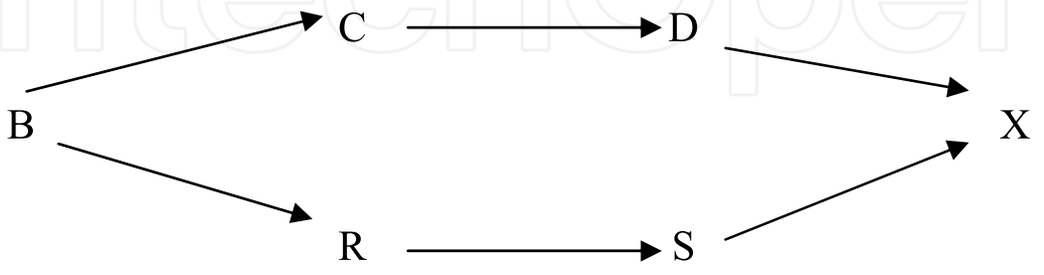
where:
n = number of individuals in the path connecting U and Z. Where Z is descendent of U.
1/2: probability that one allele is transmitted to the next generation
F_Z: inbreeding coefficient of Z (common ancestor).

Values of F range between 0 and 1. In the reference population (assuming no homozygous animals), F coefficient is equal to 0.
In the example, there are three ancestors (W, Z and V) which are considered in the transmission of alleles but X and Y are ignored since they don't influence the inbreeding coefficient. If the common ancestor Z is inbred, the Fz is calculated and multiplied by 1/8. So, the inbreeding coefficient is calculated as:

$$\begin{aligned} Fu &= 1/8 + 1/8Fz \\ &= \left(1 + Fz\right)\ 1/8 \\ &= \left(1 + Fz\right)\ \left(1/2\right)^n \end{aligned}$$

If the inbreeding coefficient of the common ancestor is not specified, it is assumed that is 0.

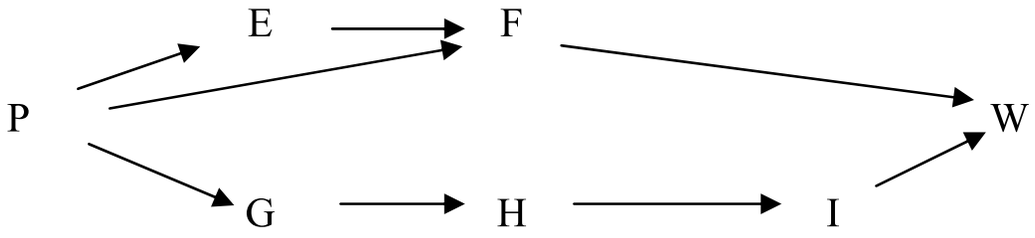
Example 3.1 – One common ancestor (one path)



The common ancestor is B: SR<u>B</u>CD; n = 5

$$Fx = (1/2)^5 = 1/32 = 0{,}03125$$

Example 3.2 - One common ancestor (two paths)



The two paths are:

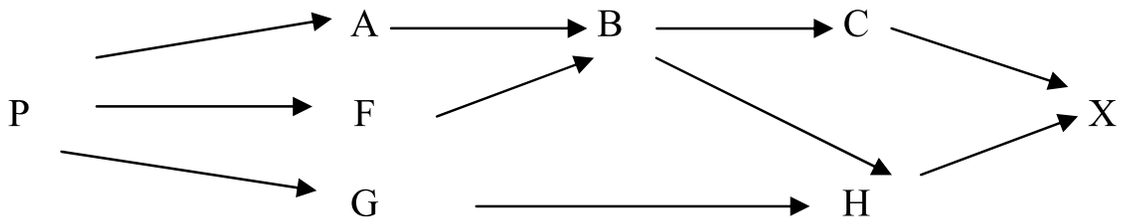
$$IHG\underline{P}EF \qquad (1+ F_P)(1/2)^6$$

$$IHG\underline{P}F \qquad (1+ F_P)(1/2)^5$$

If P is not inbred:

$$P_W = (1/2)^6 + (1/2)^5 = 3/64 = 0{,}0469$$

Example 3.3 – Two common ancestors (one animal is inbred):



P and B are the common ancestors of C. H, B are inbred but they don't contribute to the inbreeding coefficient of X. The inbreeding coefficient of B is equal to:

$$F_B = (1/2)^3 = 0.125$$

There are three paths:

$$H\underline{B}C \qquad (1+F_B)(1/2)^3$$

$$HG\underline{P}ABC \qquad (1 + F_P)\,(1/2)^6$$

$$HG\underline{P}FBC \qquad (1 +F_P)\,(1/2)^6$$

$$F_X = (1 +1/8)(1/2)^3 + (1/2)^6 + (1/2)^6 = 0.172$$

All paths can be easily identified using the following rules:

- one animal appears only once in a path;
- the path has a direct trend;
- all individuals, with the exception of the common ancestors, are ignored in the calculation of the inbreeding coefficient.

Inbreeding occurs in the progeny of related parents increasing the degree of genetic homozygosity, at the expense of heterozygous genes. The increase of inbreeding rate in the population induces two genetic events:

a. a progressive fixation of alleles;
b. a gradual reduction of dominance effects;
c. an increment of inbreeding depression effects due to the higher frequency of recessive genes.

Inbreeding coefficients refer also to the inbreeding level averaged across all individuals living in a population.

3.2. Effective population size

In small populations, the effective number of reproducing animals or effective population size (Ne) determines the expected increment of inbreeding per generation (rate of inbreeding) [11]:

$$\Delta F = \frac{1}{2N_e} \tag{2}$$

Note that equation 2 is appropriate only if the population is in Hardy-Weimberg equilibrium (panmictic population).

3.3. Unequal numbers of males and females

Sometimes, in small populations, the number of males (Mm) and females (Nf) is not 1:1. In this case, sexes are not contributing equally and Ne is calculated as:

$$\frac{1}{N_e} = \frac{1}{2}\left(\frac{1}{2Nm} + \frac{1}{2Nf}\right) = \frac{1}{4Nm} + \frac{1}{4Nf} \tag{3}$$

The equation 3 can be also written as:

$$N_e = \frac{4N_m N_f}{N_m + N_f}$$

The effective population size is primarily determined by the less numerous sex. The increase of F in one generation is computed as:

$$\Delta F = \frac{1}{2N_e} = \frac{1}{8N_m} + \frac{1}{8N_f} \tag{4}$$

Example 3.3.1: 2 males and 50 females

$$1/Ne = 1/8 + 1/200 = 0.13$$

$$Ne = 1/0.13 = 7.7$$

$$\Delta F = 1/2(7.7) = 0.0650$$

In terms of increment of inbreeding per generation, this population of 52 individuals is equivalent to a population of 8 animals: 4 males and 4 females.

Example 3.3.2: the number of males is 2 and the number of females is assumed to be infinite

$$1/Ne = 1/4Nm = 1/4(2) = 1/8$$

$$\Delta F = 1/16 = 0.0625$$

The result obtained in the example 3.3.2 is similar to the value calculated in the previous example.

3.4. Non-random distribution of the family size

The family size is the number of offspring of each family that become parents in the next generation. Under an ideal situation, the size of the population remains constant in successive generations and each of the parents has to be replaced by another animal. In this case, the average number of offspring per parent is equal to 1 with an average size of the family of 2. Ne is function of the variance of the family size:

$$N_e = \frac{4N}{2 + V_k} \tag{5}$$

where:
N = total number of animals in the population
V_K: variance of the family size
Note that if $V_K = 2$ than Ne = N

Example 3.4: Vk = 6 for both sexes. Population size: 25 males and 25 females

$$Ne = \frac{4(50)}{2+6} = \frac{200}{8} = 25$$

In terms of inbreeding rate, this population is equivalent to a population made of 12 males and 12 females.

If each male mate with more than one female, then the number of offspring and the variance of family size will be different within sexes. In this case, equation 5 becomes:

$$N_e = \frac{8N}{4 + V_{km} + V_{kf}}$$

V_{km} and V_{kf} are variances of family size for males and females.

3.5. Variable number of breeding animals over generations

If the number of parents is not constant over generations, the effective population size can be calculated by the harmonic mean as follows [11]:

$$\frac{1}{N_e} = \frac{1}{t}\left(\frac{1}{N_1} + \frac{1}{N_2} + ... + \frac{1}{N_t}\right) \quad (6)$$

where:
t = number of generation.
N_1 = number of reproducing animals at the first generation

Example 3.5: Numbers of parents over four generations: 10, 10,50 and 10 animals

$$\frac{1}{Ne} = \frac{1}{4}\left(\frac{1}{10} + \frac{1}{10} + \frac{1}{50} + \frac{1}{10}\right)$$

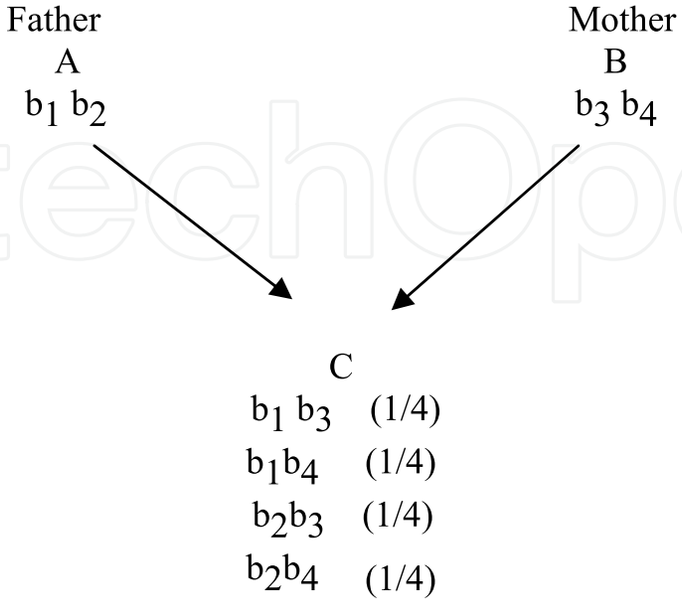
$$1/4(0.32) = 0.08$$

$$\overline{N_e} \cong 13$$

After four generations, the expected inbreeding coefficient will be the same as for a population of 13 animals in each generation. Note that, the increase in the number of breeding animals up to 50 in the third generation will not modify the value of inbreeding rate.

3.6. The kinship coefficient and the additive relationship

The kinship coefficient f_{IJ} between two individuals (A and C) is measured by the probability of taking a given allele at a locus of an animal that is identical by descent to another allele on the same locus in a second animal [12]:



The probability that two alleles taken at random from A and C are identical is equal to:

Possible combinations	Probability (P)	
b_1b_1	1/2 x 1/4	$\Rightarrow$ 1/8
b_1b_2	1/2 x 1/4	
b_1b_3	1/2 x 1/4	
b_1b_4	1/2 x 1/4	$\Rightarrow$ 1/4
b_2b_1	1/2 x 1/4	
b_2b_2	1/2 x 1/4	$\Rightarrow$ 1/8
b_2b_3	1/2 x 1/4	
b_2b_4	1/2 x 1/4	

The probability to get two identical alleles (b_1b_1 or b_2b_2) from A and C (father and son) is equal to ¼ (P = 1/8 + 1/8=1/4). Because each locus contains two alleles, the process must be repeated two times. The coefficient of kinship or additive relationship is defined as:

$$2f_{AC} = a_{AC} = = 2 \times 1/4 = 1/2 \quad (7)$$

where:
f_{AC} = kinship coefficient
a_{AC} = additive relationship

The coefficient of inbreeding of an animal C is equal to the coefficient of kinship of his parents (A and B) [11]:

$$F_C = f_{AB} \quad (8)$$

The inbreeding coefficient is equal to half of the additive relationship coefficient of his parents:

$$a_{AB} = 2F_C$$
$$F_C = \frac{1}{2} a_{AB} \quad (9)$$

3.7. Calculation of the additive relationship and relationship coefficients using the tabular method

The simplest method to determine the additive relationship coefficient between individuals and inbreeding coefficient is the tabular method [13]. This method is called tabular method because, the result takes the form of a table. The tabular method allows to construct a matrix relationship in the following way:

1. the number of columns is equal to the number of animals in the population. In the following population, there are 6 individuals and 6 columns. Animals are sorted by birth date starting from the left;
2. parents are indicated above each individual (-: missing record);
3. the value of 1 (on the diagonal) indicates the relationship of each individual with himself ($a_{xx} = 1$);

4. the calculation of the additive relationship starts on the first animal on the first row (A) and continues with the other individuals in the same line;
5. the additive relationship of each animal is computed as 1/2 of the sum of the additive relationship coefficients of his/her parents at the left on the same row;
6. the inbreeding coefficient is calculated adding to 1 values on the diagonal, 1/2 of the additive relationship of the animal's parents (e.g. see in table 1 the DD cell: 1+1/8).

	- -	**A -**	**A -**	**B C**	**D -**	**D C**
	A	**B**	**C**	**D**	**X**	**Y**
A	1	1/2	1/2	1/2	1/4	1/2
B	1/2	1	1/4	5/8	5/16	7/8
C	1/2	¼	1	5/8	5/16	13/16
D	1/2	5/8	5/8	1+1/8	1/2	13/16
X	1/4	5/16	5/16	1/2	1	13/32
Y	1/2	7/8	13/16	13/16	13/32	1

Table 1. Example of calculation of the additive relationship and inbreeding coefficient using the tabular method

Parameters obtained from the pedigree analysis provide a useful information for predicting the genetic consequences of a given management scheme or for designing future resources of a conservation programme, where biodiversity has to be maintained. Use of molecular information (combined with pedigree data or alone) may be the most useful for dealing with adaptive variation and to unveil the old history of populations (i.e. before pedigree recording started) [7].

4. Restricted inbreeding strategies in farm animals

There are a number of approaches described in the literature to assess the acceptable rate of inbreeding or conversely the minimum effective population size to maintain a relatively 'safe' population. Regarding the short-term prevention of inbreeding depression problems, there is a consensus among animal breeding researchers that ΔF of 0.5 to 1% is the acceptable rate. Therefore, an effective size of 50-100 could be sufficient to keep a population in a healthy state. Meuwissen and Woolliams in 1994 [14] also considered balancing the depression due to inbreeding, which decreases fitness, against the genetic variation available for natural selection, which improves fitness. Depending on the fitness parameters assumed, the critical effective size varied between 50-100 individuals. When taking into account other criteria (i.e. long-term potential to evolve and accumulation of mutations), figures should be higher, with the value depending on the assumptions about the mutational model (i.e. the mutational rate and the mean effect of spontaneous mutations). Some organisations (e.g. FAO [15]) often use the effective population size to define the level of endangerment. Breeding programmes for mainstream breeds are focused on achieving significant gains in the trait of interest but the programmes should also deal with the problems associated with the loss of diversity. One way to cope with the situation is an

efficient monitoring process to detect undesirable changes in fitness traits that are sensitive to inbreeding depression. However, a more reasonable strategy is to incorporate restrictions on the level of expected kinship (or inbreeding) in the animals with the objective of maximising their gain. The maintenance of variation is related to the effective population size or rate of inbreeding. From the definition of Ne itself and the factors maximising Ne (or minimising the genetic drift), some basic recommendations can be extracted. First, we should obviously keep the highest possible number of parents and try to have the same number of sires and dams. Then, we should try to equalise the number of offspring (contributions) to be obtained from every potential parent. The idea behind this is to give the same opportunity to every parent of effectively transmitting their alleles. And finally, we should prolong the generation interval as genetic drift occurs always when parents' alleles are sampled in creating offspring. Note that, this last recommendation and that of using many parents decreases the annual rate of response in a selection scheme. In practice, in many livestock species it is impossible to reach the 1:1 sex ratio. To cope with this situation some hierarchical (several dams mated to each sire) and regular systems have been developed [16,17]. The idea is always to equalise the contributions from each individual to the next generation. Basically, these strategies consist of a more or less optimised form of within-family selection. Hierarchical methods have the advantage of being simple and easy to implement for non specialised personnel and of providing predictions on the evolution of inbreeding over the years. The disadvantage of them is that they are very sensitive to deviations from the assumed conditions (i.e. related founders, mating failures, number of females not being an exact multiple of the number of males, fluctuating population size) as shown by Fernandez et al. [18] and, therefore, they are not applicable in most real situations. When no pedigree is available there are two options. To begin with, we could use molecular information to complete or replace the genealogical information. In its simpler form, it is very common to carry out a paternity analysis, useful for determining the probabilities for the sire candidates (and sometimes also for the dams) in free range animals, and consequently filling the gaps in the pedigree. In more complex situations, we could determine the general relationships in a group of animals through a set of available kinship estimators [19,20] or a IBD (Identical By Descent) matrix is constructed. Fernàndez et al. (18) studied the accuracy of molecular kinship in maintaining the genetic diversity in a conservation programme when replacing or complementing the genealogy with molecular genetic information. The study relied on the use of microsatellites and conclusions should be re-evaluated in the context of dense SNP maps. The genomic information could also be utilised for comparing the genetic value of individual animals for quantitative traits. The pedigree-based relationships can be augmented or even replaced by marker-based information. This is probably easier to envisage by considering a new genomic selection method [6], where the genetic value of an animal is determined by summing the effects of tens or hundreds of thousands markers over the whole genome. Marker effects are estimated from a sufficiently large reference population. Management of variation is very important in genomic selection because, as a very efficient method, it is expected to lead to a long-term depletion of variation with a higher risk compared than conventional methods [21].

5. Use of the optimum contribution selection (OCS)

The kinship between individuals is directly related to the genetic diversity of the population (measured as the expected heterozygosity) and also related to the expected inbreeding in the next generation. The kinship between individuals also reflects the proportion of common genes and, thus, the redundancy of the alleles in the individuals. From this, it follows that a good methodology should consist of finding the combination of contributions from available parents to minimise the expected average kinship in the next generation. This is achieved by applying the OCS [22]. Long-term selection schemes also benefit from it by restricting the average kinship to a desired level in the in the objective function (with a negative sign), directed at maximising the gain [3]. There are interesting similarities behind the two terms. In finding the best candidates for selection, the comparison of genetic values also use the information from relatives. The well-known additive relationship matrix, used in such an evaluation using the BLUP methodology, equals twice the kinship matrix. In conclusion, with OCS one can either minimise the rate of inbreeding (ΔF), or constrain it into a predefined value and maximise genetic gain simultaneously. Recently, software has been developed for choosing the sires and dams and allocating the contributions for them both in conservation and selection programmes. GENCONT [23] is able to perform OCS selection for a given rate of inbreeding. EVA [24] produces a similar kind of outcome but puts cost weights against the kinship instead of restricting the rate of inbreeding. Once the parents and the optimal number of offspring from each of them have been decided, we should determine the mating scheme. It should be noted that the optimisation of contributions is the main task in the management, leaving little margin for any improvement in the mating design. With a one generation horizon, the genetic level and average kinship do not depend on the way the parents were mated. Inbreeding is greatly influenced, because the inbreeding of the descendants is, by definition, the kinship between the mating pairs. If we are worried about inbreeding, it is sensible to implement strategies that prevent matings between close relatives [22]. In a general non-regular population, this methodology is called the minimum kinship mating and consists of finding the combinations of couples that yield the minimum average kinship between each pair of individuals to be mated. As pointed out by some authors [25], the prevention of mating between relatives is not the best method in the long term but the method they proposed implies a large increase in inbreeding in the short term, which would not be acceptable in most conservation programmes. Other strategies like compensatory mating [8] have been proposed. This methodology works by mating the most related females with the least related males, and vice versa, trying to balance the genetic contributions from under- and over-represented lineages. However, performances are not really very different from that of the minimum kinship mating, so the former may be recommended. Henryon et al. [26] proposed to reduce the covariance between ancestral contributions (MCAC mating), showing that lower levels of inbreeding can be reached when performing truncation selection. When physiologically feasible, some authors [27] have proved that performing a factorial mating design (i.e. mating each parent to several mates) would reduce the levels of inbreeding achieved due to the reduced correlation between the contributions of mates.

Moreover, factorial mating increases the flexibility in breeding schemes for achieving the optimum genetic contributions. Sometimes, for practical reasons (e.g. a female is not able to mate with more than one male), and results from the OCS methodology cannot be fitted into a realistic mating design. In that situation, we would like to determine, at the same time, not only how many offspring an animal should have, but also with which animal it should be mated. The simultaneous optimisation of selection and mating is called 'mate selection' and, instead of deciding just on the number of offspring to be had from each candidate, it also looks into the number of offspring produced from every possible couple. It is easy to include some restrictions on the number of matings per particular animal or the maximum number of full-sibs to generate among the progeny.

6. A stochastic simulation program (Matlab), based on a simulated annealing algorithm, for optimizing farm animal breeding schemes under restricted inbreeding

In species with large families, the management of the pedigree to minimize ΔF can be combined with appropriate selection techniques within families. The high reproductive potential, in some commercial species (pig, chicken, fish), allows high genetic gains by applying high selection intensities. This means that, a very small number of individuals are used to generate successive generations and hence the rate of inbreeding can be high [27]. The detrimental effects of inbreeding are well documented in several commercial species. In recent years, many selection and mating strategies have been proposed to restrict inbreeding in selection programmes [27]. In this study, a stochastic simulation model was used to simulate and optimize mating schemes of farm animals using different genetic parameters and under restricted inbreeding. The structure of the simulated breeding scheme was that of a closed nucleus. An animal population under artificial selection was modelled by stochastic (Monte Carlo) simulation using the Matlab software. Selection was applied for a single trait measured on both sexes and based on estimated breeding values (EBVs) using the ASREML2 statistical package. Generations were discrete (equal number of sires and dams were selected at each generation). The trait under selection was assumed to be determined by an infinite number of unlinked additive loci, each with an infinitesimal effect. The trait was considered to be standardized, so the initial phenotypic variance is unity. Phenotypes of unrelated base population animals (generation 0) were generated as the sum of a normally distributed environmental and genetic effects. Phenotypic values of the offspring born every generation were generated as:

$$P_i = \mu + \left(\sigma_A \mathrm{RND}(0,1)_S + \sigma_A \mathrm{RND}(0,1)_D\right)/2 + \sigma_E \mathrm{RND}(0,1) + \\ + \left(0.5\left(1-\left(F_s+F_d\right)/2\right)\right)^{1/2} \sigma_A * \mathrm{RND}(0,1) \quad (10)$$

where:

$\sigma_A = \sigma_{A(o)}/(1+kh^2)$

$k=(0.5)(k_m+k_f)$
$k_y= i_y(i_y-x_y)$ y=male or female
i= selection intensity
RND = random number

Phenotypic values (P*i)* were calculated as $P_i = \mu + \sigma_{Gi} + \sigma_{Ei}$ where σ_{Gi} is the genetic effect and σ_{Ei} is the environmental effect, which were sampled from *N* (0,1) making the base phenotypic variation (σ^2_P) equal to 1.0. The base generation additive genetic variance, σ^2_A was 0.1, or 0.25, or 0.50 corresponding to a heritability, h^2, of 0.1 or 0.30 or 0.50, respectively. Later generations were obtained by simulating progeny genotypes from $\sigma_{Gi} = 0.5\ \sigma_{Gs} + 0.5\ \sigma_{Gd} + mi$, where *s* and *d* denote sire and dam of progeny *i*, respectively, and *mi* = mendelian sampling component, which was sampled from $(0.5(1-(F_s+F_d)/2\ ))^{1/2}\ \sigma_A$ *RND(0,1), where $(F_s+F_d)/2$ is the average of the inbreeding coefficients of the sire and the dam. ΔF was restricted to 0.010 per generation, which is an indication of the maximum acceptable rate of inbreeding. Selection was directional upwards and by truncation. Total number of offspring born per generation and numbers of selected males and females were constant (10 or 20 offspring per mating) over generations and varied according to the mating schemes. The simulated breeding schemes are described in Table 2.

Number of selection candidates per generation	90 or 180 or 360 or 720
Number of generations	10
Number of replicated simulations	100
Mating schemes	factorial 3 x 3; 6 x 6; nested 6 (males) x 18 (females)
Heritability coefficient	0.1 or 0.3 or 0.5

Table 2. Parameters of the closed nucleus scheme

The OCS and simulated annealing was used to select animals. The OCS theory maximises the genetic gain while constraining the rate of inbreeding or the relationships among selection candidates. These methods choose the selected parents and assign genetic contributions to the next generation for each selected candidate. This method maximises the genetic level of the next generation of animals:

$$G_{t+1} = c_t' EBV_{t\,1} \tag{11}$$

c_t is a vector of genetic contributions of selected candidates to the generation t_{+1};

EBVt is a vector of best linear unbiased prediction (BLUP) estimates of candidates in generation t.

The objective function, $ct'EBV_t$, is maximized for c*t* under two restrictions: the first is on the rate of inbreeding and the second is on the contribution per sex. The desired rate of inbreeding, ΔF is obtained by constraining the average kinship of the selection candidates to:

$$C_{t+1} = 1 - (1 - \Delta F)t \tag{12}$$

The actual contributions of the individuals are then obtained in such a way that they fulfil the constraint:

$$C_{t+1} \geq ct' A tct/2 \quad (13)$$

where:
A_t is a ($n \times n$) relationship matrix among the selection candidates. Note that the level of the constraint C_{t+1}, can be calculated for every generation before the breeding scheme commences.

The contribution of each sex is constrained to ½ :

$$Q'ct = 1/2 \quad (14)$$

where:
Q is a ($n \times 2$) incidence matrix of the sex of the selection candidates (the first column yields ones for males and zeros for females, and the second column yields ones for females and zeros for males). The contributions of the male and those of the female candidates will sum to ½.

In order to obtain the optimal ct that maximize $G_{t+1,}$ Lagrangian multipliers were used. An additional restriction was to select only one full sibs per family.

Using the lagrangian method for restricted optimization, the optimum solution is obtained as follows:

$$c = A^{-1} (EBV - Q \lambda) /2\lambda_0) \quad (15)$$

where:
λ and λ_0) are lagrangian multipliers.

The minimum kinship mating (reduce the average relationship of sires and dams and therefore also the inbreeding of their progeny is minimized) is obtained by applying the simulated annealing algorithm according to Press et al.[4]. The output from the selection method is a vector with genetic contributions for each selection candidate, c*t*. The ultimate goal, in this mating tool, is to reduce the average inbreeding coefficient in the following generation. Input parameters included all possible relationships between pairs of selected dams and selected sires. The scheme with the lowest average inbreeding coefficient in the next generation is considered as the optimal one. The essential steps of the simulated annealing algorithm can be summarized as follows: 1) sires and dams are mated at random according to their frequencies in vector **c**, and than the resulting average inbreeding coefficient is stored as reference value *00*; 2) change of mating partners and comparison of the new resulting average inbreeding value *01* with *00*; 3) if the value *01 is* < *00* than *it* is replaced with *01* and so for all possible matings. By using simulated annealing, inbreeding is avoided as much as possible. The rate of inbreeding (ΔF) and the genetic gain (increase in animal performance through a genetic programme, ΔG) for the three mating designs are reported in Table 3.

	Number of offspring per family					
	10			20		
Heritability	0.1	0.3	0.5	0.1	0.3	0.5
	Full factorial (3males x 3 females)					
ΔF	3.96	5.00	4.55	4.50	4.70	4.65
ΔG	0.29	0.60	0.75	0.37	0.60	0.97
	Full factorial (6 males x 6 females)					
ΔF	2.25	2.20	2.10	2.05	1.80	1.85
ΔG	0.38	0.64	0.76	0.47	0.69	0.94
	Nested (6 males x 18 females)					
ΔF	2.60	2.25	2.36	2.06	2.90	2.05
ΔG	0.43	0.72	0.86	0.57	0.66	0.89

Table 3. Rate of inbreeding (ΔF)(x100) and genetic gain (ΔG))(σ_P) for different mating schemes and genetic parameters.

The full factorial design gives the best results in terms of ΔF and ΔG (1.85 and 0.94 or 1.80 and 0.69) using a higher number of sires and dams (6 x 6), family size per mating ,a family size per mating of 720 offspring and heritability coefficients of 0.3 or 0.5. According to Sorensen *et al*. [2], the superiority of the factorial mating compared to hierarchical scheme can be explained in terms of the different genetic structure of populations obtained showing, in the factorial design, small full-sibs families, more paternal half-sibs and a group of maternal halfsibs. At a lower heritability (0.1) the nested design become competitive with the full factorial mating (6 x 6). This selection approach have been already evaluated in practice for several domestic species such as dairy cattle [7], the Hanoverian horses for show jumpers [19], fish and pigs.

7. Conclusions

Additive relationships among individuals are generally used for weighting records of relatives in the genetic evaluation of farm animals and to calculate inbreeding coefficients. The tabular method, used for computing the additive relationships and inbreeding coefficients of farm animals is the most efficient and widely used method. According to the present simulation study, the best mating scheme of farm animals under a restricted inbreeding rate is a full factorial mating (6 males and 6 females) with full-sibs families of 20 animals. Furthermore, the present work has clearly shown that, the most suitable approach for long-term selection activities under inbreeding restrictions, is to use together the optimum genetic contribution and simulated annealing methods.

Author details

Edo D'Agaro
Faculty of Veterinary Medicine, University of Udine, Udine, Italy

8. References

[1] Woolliams J.A. Genetic contributions and inbreeding. In: Oldenbroek, K. (ed.) Utilisation and conservation of farm animal genetic resources. Wageningen: Academic Publishers; 2007 p 147-165.

[2] Sorensen A.C., Berg P. and Woolliams J.A. The advantage of factorial mating under selection is uncovered by deterministically predicted rates of inbreeding. Genetics Selection Evolution 2005;37 57-81.

[3] Meuwissen, T.H.E.,. Maximizing the response of selection with a predefined rate of inbreeding. Journal of Animal Science 1997;75 934-940.

[4] Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T. Minimization or maximization of function. In: Numerical recipes. The act of scientific computing. Cambridge: University of Cambridge Press;1989 pp. 274-334.

[5] Dekkers J.C.M., Hospital F. The use of molecular genetics in the improvement of agricultural populations. Nature Rev. Genet. 2002;3 22-32.

[6] Meuwissen T.H.E., Hayes B.J., Goddard, M.E.. Prediction of total genetic value using genome-wide dense marker maps. Genetics 2001; 157 1819-1829.

[7] Toro M.A., Màki-Tanila, A. Genomics reveals domestication history and facilitates breed development. In: Oldenbroek, K. (ed.) Utilisation and conservation of farm animal genetic resources. Wageningen: Wageningen Academic Publishers; 2007 p 75-102.

[8] Sellner E.M., Kim J.W., McClure M.C., Taylor K.H., Schnabel R.D., Taylor J.F. Board-invited review: Applications of genomic information in livestock. Journal of Animal Science 2007; 85 3148-3158.

[9] Caballero A., Toro, M.A. Characterization and conservation of genetic diversity in subdivided populatìons. Phil. Trans. R. Soc. B 2005;360 1367-1378. Caballero, A., Santiago, E. and Toro, M.A., 1996. System of mating to reduce inbreeding in selected populations Animal Science 2005;62 431-442.

[10] Toro M.A., Fernàndez J., Caballero A. Molecular characterization of breeds and its use in conservation. Livestock Science 2009; 120 174-195.

[11] Falconer DS Introduction to quantitative genetics. New York: Longman Inc.; 1981

[12] Malecòt G. Les mathèmatiques de l'ereditè. Paris: Masson and Cie ; 1948

[13] Van Vleck L.D., Pollack E.J., Oltenacu E.A.B. Genetics for the animal sciences. New York: Freeman and Company; 1987

[14] Meuwissen T.H.E., Woolliams, J.A.. Effective sizes of livestock populations to prevent a decline in fitness. Theoretical and Applied Genetics 1994; 89 1019-1026.

[15] FAO Secondary Guidelines on the Management of Small Populatìons at Risk. UNEP, Rome;1998 p 201-204.

[16] Wang J. More efficient breeding systems for controlling inbreeding and effective size in animal populations. Heredity 1997; 79 591-599.

[17] Sànchez-Rodrfguez L., Bijma P., Woolliams J.A. Reducing inbreeding rates by managing genetic contributions across generations. Genetics 2003; 164 1589-1595.

[18] Fernàndez J., Toro M.A., Caballero, A.. Fixed contributions designs versus minimization of global coancestry to contrai inbreeding in small populations. Genetics 2003;165 885-894.

[19] Fernàndez J. and Toro M.A.. A new method to estimate relatedness from molecular markers. Molecular Ecology 2006; 15 1657-1667.

[20] Oliehoek P.A., Windig J.J., Van Arendonk J.A.,Bijma, P., Estimating relatedness between individuals in generai populations with a focus on their use in conservation programs. Genetics 2006;173 483-96.

[21] Hayes, B.J., Bowman, P.J., Chamberlain, A.J. and Goddard, M.E. Invited review: Genomic selection in dairy cattle: progress and challenges. Journal of Dairy Science 2009;92 433-43.

[22] Meuwissen T.H.E. Operation of conservation schemes. In: Oldenbroek, K (ed.) Utilisation and conservation of farm animal genetic resources. Wageningen: Wageningen Academic Publishers; 2007 p 167-193.

[23] Meuwissen T.H.E.. GENCONT: An operational tool for controlling inbreeding in selection and conservation schemes. Proceedings of 7th World Congress on Genetics Applied to Livestock Production, Montpellier,19-23 August 2002; 33 769-770.

[24] Berg P., Nielsen I., and Sorensen M.K. Computing realized and predicting optimal genetic contributions by ÈVA. Proceedings of 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte; 2006 p 120-121.

[25] Crow J.F. and Kimura, M., 1972 An introduction to population genetics theory. Harper and Row; New York. Daetwyler, H.D., Villanueva, B., Btjma, P., Woolliams, J.A.. Inbreeding in Genome-Wide Selection. Journal of Animal Breeding and Genetics 2007; 124 369-376.

[26] Henryon M., Sorensen A.C. and Berg P. Mating animals by minimising the covariance between ancestral contributions generates more genetic gain without increasing rate of inbreeding in breeding schemes with optìmum-contribution selection. Animal Production 2009;3 1339-1346.

[27] Sonesson A.K., Gjerde B., Meuwissen T.H.E. Truncation selection of BLUP-EBV and phenotypic values in fish breeding schemes. Aquaculture 2005; 243:61-68.