

# The use of frozen semen to minimize inbreeding in small populations

ANNA K. SONESSON<sup>1\*</sup>, MIKE E. GODDARD<sup>2</sup> AND THEO H. E. MEUWISSEN<sup>1†</sup>

<sup>1</sup>ID-Lelystad, Institute for Animal Science and Health, PO Box 65, 8200 AB Lelystad, The Netherlands

<sup>2</sup>Institute of Land and Food Resources, University of Melbourne, Parkville, Victoria 3052 and Victorian Institute of Animal Science, Attwood, 3049 Victoria, Australia

(Received 19 September 2001 and in revised form 16 April 2002)

## Summary

In this study, we compared the average coancestry and inbreeding levels for two genetic conservation schemes in which frozen semen from a gene bank is used to reduce the inbreeding in a live population. For a simple scheme in which only semen of generation-0 ( $G_0$ ) sires is used, the level of inbreeding asymptotes to  $1/(2N)$ , where  $N$  is the number of newborn sires in the base generation and rate of inbreeding goes to zero. However, when only sires of  $G_0$  are selected, all genes will eventually descend from the founder sires and all genes from the founder dams are lost. We propose an alternative scheme in which  $N$  sires from generation 1 ( $G_1$ ), as well as the  $N$  sires from  $G_0$ , have semen conserved, and the semen of  $G_0$  and  $G_1$  sires is used for dams of odd and even generation numbers, respectively. With this scheme, the level of inbreeding asymptotes to  $1/(3N)$  and the genes of founder dams are also conserved, because 50% of the genes of sires of  $G_1$  are derived from the founder dams. A computer simulation study shows that this is the optimum design to minimize inbreeding, even if semen from later generations is available.

## 1. Introduction

A common goal of genetic conservation schemes is to maintain a living population with minimum genetic drift and inbreeding. One way to reach this goal is to use old, less-related males. This slows down the turnover of generations and thus reduces genetic drift and inbreeding. Smith (1977) took this idea to the extreme and proposed the use of frozen semen of only the least-related sires from the founder generation (generation 0;  $G_0$ ), in order to keep genetic drift to a minimum. However, if only sires from  $G_0$  are selected, all genes will eventually descend from the founder sires: the sires of  $G_0$  contribute 50% of the genes in generation 1, 75% of the genes in generation 2, 87.5% of the genes in generation 3, etc. Therefore, there will eventually be no genes of the females of  $G_0$  (if frozen oocytes and embryos cannot be used). However, if sires from generation 1 ( $G_1$ ) as well as  $G_0$  were used, genes of founder dams would be conserved in the

progeny because males from  $G_1$  have 50% of their genes from founder dams. Here, we propose to mate sires of  $G_0$  to dams from odd generation numbers and sires of  $G_1$  to dams from even generation numbers.

The aim of this paper is to compare the inbreeding and average coancestry ( $\bar{C}_t$ ) of a scheme in which only sires of  $G_0$  are used (Smith, 1977) with those of a scheme in which sires of both  $G_0$  and  $G_1$  are used. The average coancestry is a good measure of genetic diversity because it accounts for both founder effects and the effects of genetic drift with respect to contributions and variance of family size, and because it has a direct relationship with the expected heterozygosity (Lacy, 1995; Caballero & Toro, 2000).

## 2. Methods

### (i) Derivation of average coancestry

The coancestry of two individuals is defined as the probability that two gametes taken at random (one from each individual) carry alleles that are identical by descent (Malécot, 1948). The average coancestry of animals in a population,  $\bar{C}_t$ , can be calculated as the

\* Corresponding author. Tel: +31 (0)320 238238. Fax: +31 (0)320 238050. e-mail: a.k.sonesson@id.wag-ur.nl

† Present address: Agricultural University of Norway, PO Box 5025, 1432 Ås, Norway.

average of all elements in a matrix of coancestry between all individuals. We can split up the total average coancestry  $\bar{C}_t$  in three terms taken from the coancestry matrix

$$\bar{C}_t = 0.25\bar{C}_s + 0.25\bar{C}_d + 0.50\bar{C}_{sd}, \tag{1}$$

where  $\bar{C}_s$  and  $\bar{C}_d$  are the average coancestries of sires and dams, respectively, and  $\bar{C}_{sd}$  is the average coancestry between sires and dams.

(a) *Repeated use of frozen semen from sires of generation 0*

For this scheme,  $N$  sires from  $G_0$  are used every generation. Every sire is randomly mated to one of the  $N$  dams to produce one female offspring per mating pair.

The coancestry of  $G_0$  sires, inclusive of the coancestry of each sire with itself,  $C_s = 0.5 \times I_N$  is a matrix in which the diagonal element, 0.5, represents the coancestry of a sire with itself (0.5) and the off-diagonal terms represents the coancestry between different  $G_0$  sires, which is 0 here. It follows that  $\bar{C}_s = N \times 1/(2N^2) = 1/(2N)$ .

The average coancestry between sires and dams,  $\bar{C}_{sd}$ , is the average of the  $N \times N$  matrix  $C_{sd} = J\bar{F}$ , where  $J$  is an  $N \times N$  matrix of ones and the average coancestry of sires and dams is, by definition, the average inbreeding of the population ( $\bar{F}$ ) when there is random mating. It follows that  $\bar{C}_{sd} = \bar{F}$ .

$\bar{F}$  can be derived from  $c_{x,y} = 0.5(c_{x, sy} + c_{x, dy})$ , where  $c_{x,y}$  denotes the coancestry of the sire  $x$  and dam  $y$ , and the sire and dam of  $y$  are denoted by  $sy$  and  $dy$ . At equilibrium, the expected value of  $c_{x, sy}$  is  $\bar{C}_s$  and the expected value of both  $c_{x, dy}$  and  $c_{x, sy}$  is  $\bar{F}$ , because  $x$  is a sire and  $y$  and  $dy$  are dams, and  $\bar{F}$  has stabilized over generations. It follows that  $\bar{F} = 0.5[1/(0.5 \times N) + \bar{F}]$  and thus that  $\bar{F} = 1/(2N)$ .

The average coancestry of different dams, inclusive of the coancestry of each dam with itself,  $\bar{C}_d$ , is the average of the  $N \times N$  matrix

$$C_d = \begin{bmatrix} 1/2 + 1/2\bar{F} & \cdots & C_d^* \\ \cdots & \cdots & \cdots \\ C_d^* & \cdots & 1/2 + 1/2\bar{F} \end{bmatrix},$$

where  $C_d^*$  is the coancestry of two different dams.  $C_d^*$  can be derived from  $c_{x,y} = 0.25(c_{sx, sy} + c_{sx, dy} + c_{sy, dx} + c_{dx, dy})$ , where  $c_{x,y}$  is the coancestry between dams  $x$  and  $y$ . Because family sizes are equal, in order to minimize inbreeding, every sire has one daughter (i.e. different daughters have different founder sires). It follows that  $c_{sx, sy} = 0$ . At equilibrium, the expected value of both  $c_{sx, dy}$  and  $c_{sy, dx}$  is  $\bar{F}$ , and the expected value of both  $c_{dx, dy}$  and  $c_{x,y}$  is  $C_d^*$ . Hence,  $C_d^* = 0.25(0 + 1/N + C_d^*)$  and thus  $C_d^* = 1/(3N)$ .

Hence the average of the elements of the  $C_d$  matrix is:

$$\begin{aligned} \bar{C}_d &= \{N[0.5 + 0.5/(2N)] + N[N - 1]/[3N]\} \div N^2 \\ &= [(5/6) - 1/(12N)] \div N. \end{aligned}$$

Finally, from Eqn 1, the equilibrium average coancestry is

$$\bar{C}_t = 7/(12N) - 1/(48N^2).$$

(b) *Alternating the use of sires from  $G_0$  and  $G_1$*

For this scheme,  $N$  sires of  $G_0$  and  $N$  sires of  $G_1$  are used. At odd generation numbers,  $G_0$  sires are used and, at even generation numbers,  $G_1$  sires are used. Every sire is randomly mated to one of the  $N$  dams to produce one male and one female offspring per mating pair. As for the previous scheme, there is random mating of sires and dams every generation.

The average coancestry of sires from different sires of the same generation, inclusive of the coancestry of each sire with itself, is  $\bar{C}_s = 1/(2N)$ ; that is, the same as for the previous scheme. The  $\bar{C}_s$  is the same for the two schemes because  $G_1$  sires are unrelated and non-inbred and have thus the same  $\bar{C}_s$  as  $G_0$  sires.

The average coancestry between sires and dams ( $\bar{C}_{sd}$ ) is, as in the previous scheme, equal to  $\bar{F}$ . The value of  $\bar{F}$  can be derived from  $\bar{F} = c_{x(0), y} = 0.5[c_{x(0), sy(1)} + c_{x(0), dy}]$ , where  $c_{x(0), y}$  is the coancestry between a  $G_0$  sire ( $x(0)$ ) and one dam ( $y$ ). At equilibrium, the expected value of  $c_{x(0), sy(1)}$  is  $1/(4N)$ , where the coancestry between a father of  $G_0$  and a son of  $G_1$  is 0.25. The value of  $c_{x(0), dy}$  is given by  $c_{x(0), dy} = 0.5[c_{x(0), sdy(0)} + c_{x(0), ddy}]$ , where  $sdy(0)$  refers to the  $G_0$  sire of  $dy$  and  $ddy$  refers to the maternal grand-dam of  $y$ . The expected value of  $c_{x(0), sdy(0)}$  is  $\bar{C}_s$ , because both sires are from  $G_0$ . The expected value of  $c_{x(0), ddy}$  is  $\bar{F}$ , because  $x(0)$  is a sire and  $ddy$  is a dam. It follows that  $\bar{C}_{sd} = \bar{F} = 0.5\{1/[4N] + 0.5[1/(2N) + \bar{F}]\}$  and thus  $\bar{F} = 1/(3N)$ .

The average coancestry of different dams, inclusive of the coancestry of each dam with itself,  $\bar{C}_d$ , is the average of the  $N \times N$  matrix

$$C_d = \begin{bmatrix} 1/2 + 1/2\bar{F} & \cdots & C_d^* \\ \cdots & \cdots & \cdots \\ C_d^* & \cdots & 1/2 + 1/2\bar{F} \end{bmatrix},$$

where  $C_d^*$  is the coancestry between two different dams. The value of  $C_d^*$  can be derived from  $c_{x,y} = 0.25(c_{sx, sy} + c_{sx, dy} + c_{sy, dx} + c_{dx, dy})$ , where  $c_{x,y}$  denotes the coancestry between females  $x$  and  $y$ . At equilibrium, the expected value of  $c_{sx, sy}$  is 0, because females  $x$  and  $y$  will be from different sires  $sx$  and  $sy$ , and every sire gets one female offspring in order to equalize family sizes. The expected value of both  $c_{sx, dy}$

and  $c_{sy,dx}$  is  $\bar{F}$ , and the expected value of both  $c_{dx,dy}$  and  $c_{x,y}$  is  $C_a^*$ . It follows that  $C_a^* = 0.25(0 + 2\bar{F} + C_a^*)$  and thus  $C_a^* = (2/3)\bar{F}$ . Using  $\bar{F}$  as calculated above

$$\bar{C}_d = [N(0.5 + 0.5\bar{F}) + N(N-1) \times (2/3)\bar{F}] \div N^2 = [13 - (1/N)] \div 18N.$$

From Eqn 1, we get the equilibrium coancestry for the scheme in which both  $G_0$  and  $G_1$  sires are used

$$\bar{C}_t = 17/(36N) - 1/(72N^2).$$

$\bar{C}_t$ ,  $\bar{C}_s$ ,  $\bar{C}_{sd}$ ,  $\bar{C}_d$  and  $\bar{F}$  were calculated for schemes with 6, 10 or 20 animals (male plus female) born per generation (i.e.  $N = 3, 5$  or  $10$ ).

(ii) *Simulation study*

We also conducted a simulation study, in which the average coancestry of the selected parents, and thus the inbreeding of the offspring, was minimized each generation (Sonesson & Meuwissen, 2001) but that considered a discrete generation structure. Female candidates for selection were taken from the current generation,  $t$ , whereas male candidates were either  $G_0$  sires only or sires from all generations ( $G_0, G_1 \dots G_t$ ). The second scheme tests whether younger sires than  $G_0$  and  $G_1$  sires will also be used when average coancestry is minimized.

**3. Results and discussion**

We attempted here to minimize inbreeding in genetic conservation schemes that combine the conservation of live animals with the use of a semen bank. A scheme proposed by Smith (1977) in which sires of  $G_0$  were used on all later generations of dams resulted in steady-state average coancestry of 0.192, 0.116 and 0.058 for schemes with 6, 10 and 20 animals per generation, respectively (Table 1). The total average coancestry was ~ 23% higher for the scheme in which

only sires of  $G_0$  were used than for the schemes in which sires of  $G_0$  and  $G_1$  were used. Similarly, the average inbreeding was 50% higher for the scheme where only sires of  $G_0$  were used than for the schemes in which sires of  $G_0$  and  $G_1$  were used.

Selection of only  $G_0$  sires or alternating use of  $G_0$  and  $G_1$  sires are only two of many possible designs. For instance, we could make equal use of  $G_0, G_1$  and  $G_2$  sires. To study the optimum design, we compared the predicted results with a selection algorithm that minimized the average coancestry each generation. It is assumed that this also minimizes the average coancestry in the longer term. Either  $G_0$  or all sires were selection candidates. The average coancestry and inbreeding levels from the simulated schemes were nearly identical to the two schemes of Table 1 (results not shown). The simulated scheme with all sires as selection candidates resulted in a scheme in which  $G_0$  and  $G_1$  sires were selected alternately (i.e. the scheme that we propose here). The simulation results indicate that our proposed scheme is optimal, because the minimization of coancestry every generation is expected to result in the minimum possible coancestry. Males of later generations are not used in the optimum design because they are related to the current females through intermediate generation dams as well as through the founder sires and dams.

The proposed method is developed for schemes in which the population is maintained for several generations when the number of sires is less than or equal to the number of dams and when the number of sires is limited. When the purpose of the use of gene-bank material is to restore a large live population, the same principle holds: more genes are conserved when semen of sires from both  $G_0$  and  $G_1$  is stored. If the number of  $G_0$  sires is larger than the number that can be stored, it is always better first to select all the sires of  $G_0$ , because they contain more genes from the founder generation than sires from  $G_1$ . Also, if founder

Table 1. *Inbreeding ( $\bar{F}$ ), coancestry among sires ( $\bar{C}_s$ ), among dams ( $\bar{C}_d$ ), between sires and dams ( $\bar{C}_{sd}$ ) and in total ( $\bar{C}_t$ ) at equilibrium for two schemes in which only sires from generation zero ( $G_0$ ) can be selected or where sires from generations zero and one ( $G_0$  and  $G_1$ ) are used alternately*

| Scheme                                   | $\bar{F}$ | $\bar{C}_s$ | $\bar{C}_d$ | $\bar{C}_{sd}$ | $\bar{C}_t$ |
|--|-----------|-------------|-------------|----------------|-------------|
| <i>6 newborn animals per generation</i>  |           |             |             |                |             |
| $G_0$                                    | 0.167     | 0.167       | 0.269       | 0.167          | 0.192       |
| $G_0$ and $G_1$                          | 0.111     | 0.167       | 0.235       | 0.111          | 0.156       |
| <i>10 newborn animals per generation</i> |           |             |             |                |             |
| $G_0$                                    | 0.100     | 0.100       | 0.163       | 0.100          | 0.116       |
| $G_0$ and $G_1$                          | 0.067     | 0.100       | 0.142       | 0.067          | 0.094       |
| <i>20 newborn animals per generation</i> |           |             |             |                |             |
| $G_0$                                    | 0.050     | 0.050       | 0.083       | 0.050          | 0.058       |
| $G_0$ and $G_1$                          | 0.033     | 0.050       | 0.072       | 0.033          | 0.047       |

female gametes or embryos can be frozen, this will reduce average coancestry more than the scheme proposed here, where only semen could be frozen.

A drawback of the conservation schemes presented is that they will allow no further evolution (genetic adaptation and accumulation of mutations) in the conserved population. Further research is needed to optimize combined *in situ* and *ex situ* conservation schemes that do allow for evolution of the conserved population. For practical schemes, the selection strategies presented here are probably oversimplified, but the general principle that storage of semen from two generations of sires conserves more genetic diversity than that of only one generation of sires is expected to hold true in more complicated situations.

## References

- Caballero, A. & Toro, M. A. (2000). Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genetical Research* **75**, 331–343.
- Lacy, R. (1995). Clarification of genetic terms and their use in the management of captive populations. *Zoo Biology* **14**, 565–578.
- Malécot, G. (1948). *Les mathématiques de l'hérédité*. Paris: Masson.
- Smith, C. (1977). Use of frozen semen and embryos to measure genetic trends in farm livestock. *Zeitschrift für Tierzucht und Züchtungsbiologie* **94**, 119–227.
- Sonesson, A. K. & Meuwissen, T. H. E. (2001). Minimisation of rate of inbreeding for small populations with overlapping generations. *Genetical Research* **77**, 285–292.