ALLELE SURVIVAL IN STORAGE OF GAMETES AND EMBRYOS FOR CONSERVATION

G.C.Gandini, R.Leonarduzzi, A.Bagnato
Istituto di Zootecnica, Facoltà di Medicina Veterinaria, Via Celoria 10, 20133 Milano, Italy.

SUMMARY

The paper presents the analysis of the probability distribution of founder alleles in the semen storage of a small cattle population. The probability distribution is estimated by simulating mendelian segregation of distinct (not identical by descent) founder alleles throughtout the pedigree at an autosomal locus (gene dropping technique). The marginal probability distributions of 1) the number of distinct founder alleles survived, 2) the contribution by founder of distinct alleles, 3) the number of copies for distinct allele present in the population are investigated. Given gene frequencies at a BoLA locus in the founders, for each allele survival probability and probability to be present with a single copy in the semen storage is computed. The analysis can be used to select the animals to be stored.

INTRODUCTION

Storage of frozen semen and embryos has been proposed for conservation of genetic variability of endangered livestock populations (e.g. Maijala et al., 1984). Type of cells (gametes, zygotes) and sample size have to be considered to minimize genetic drift in sampling and during first usage of stocks (Smith, 1984). Because the aim is to store a representative sample of the population variability, it is important to evaluate the genetic variation of the stocked animals. Genetic variation can be estimated by polymorphic DNA analysis. Whenever genealogies are available, estimates can also be obtained from pedigree data. Pedigree analysis assumes founders unrelated (relationships being unknown), thus we do not estimate the true genetic variation but we measure changes during the registered history, from founder to current population. Considering local endangered cattle breeds, pedigree analysis often shows a large number of founders, several generations of progressive decrease of population size, and large variances in reproductive success resulting in high rates of genetic drift (unpublished data). In addition, during sampling for storage we produce further bottlenecks. Founders, being numerous, can be considered representative of the population variation before the beginning of the population size decrease process. Heterozigosity, inbreeding, kinship and proportionate founder contribution can be used to estimate the genetic variation of the stocked pedigree population. When a population is exposed to genetic drift the rate of loss of alleles can be more rapid than the loss of heterozigosity (Kimura and Crow, 1964). Therefore it is interesting to study the distribution of founder alleles at an autosomal locus. This allows estimates of the genetic variation among the stocked animals and to compute survival probabilities of desirable alleles. In this paper we analyze the distribution of the number of distinct founder alleles and of their copies in the frozen semen stock of the Pisana cattle breed by simulating the mendelian segregation of founder alleles throughout the pedigree. This simulation tecnique has been recently implemented for zoo population studies and is often called the gene dropping method (MacCluer et al., 1986). The Pisana is a dual purpose breed farmed in Tuscany. Population size from 10.000 heads in 1940 decreased to less than 300 animals today. A conservation programme, based on genetic management of the population and technical and economic support to breeders, started in the late eighties.

MATERIAL AND METHODS

With the gene dropping simulation technique two hypothetical distinct alleles are assigned to each founder. Then the genotype of descendants are constructed working down the pedigree by simulating the mendelian segregation. The entire procedure is repeated several times (here 10,000 iterations) and the information from the genotypes of the current population summarized over iterations. In this paper the current population is the group of stocked bulls (frozen semen storage) or alternatively the living population. In the simulation we do not consider selection and linkage, assuming all loci be indipendent. By gene dropping simulation we obtain an estimate of the joint probability distribution (determined by pedigree structure) of the number of copies per distinct founder allele present in the current population. This is a NFA-dimensional probability distribution, where NFA is the number of founder alleles, equal to twice the number of founders. Considering the constraint of the number of alleles in the current population (NCA) equal to twice the number of current animals, in the distribution there are

(NFA+NCA-1)!/(NCA!(NFA-1)!) terms. The high number of terms does not allow usefull direct examination. Geyer and Thompson (1988), looking at the joint probability distribution of the number of distinct alleles by founder, have shown complex correlation patterns in allele surival. However we can examine some marginal distributions: 1) the number of distinct founder alleles survived among current individuals, 2) the contribution to the current population, by each founder, of distinct founder alleles (from 0 to 2 alleles), 3) the number of copies, for distinct founder allele, present in the current population.

The probability distribution of the number of distinct founder alleles survived in the current population can be used to compare the genetic variation of different stocks and the living population. Founder alleles are all considered distinct (i.e. not identical by descent), however they might be biochemically identical (i.e. same type of allele). Assuming the founders a random sample from the population, an allele with frequency p in the population will be present among founders with frequency $p \pm (p(1-p)/\mathrm{NFA})^{0.5}$. If the number of founders is sufficiently large, we can compute the probability that this allele is present in the current population as $\sum_{w}(1-(1-p)^{w})P(w)$, where w is the number of distinct founder alleles survived and P(w) is its probability distribution.

Given NCA and w, the expected number of copies E(n) by distinct founder allele is NCA/w; however, because mendelian segregation and differential reproductive success, we expect a certain variance. We estimate, for any w, the marginal distribution of the number of copies for each distinct founder allele (P(n_w)). We can now compute, for an allele with frequency p among founders, the probability to be present with a single copy in the current population. This probability is given by $\sum_{w}(p(1-p)^{w-1}(w!/1!(w-1)!))(\Pr(1_w))(P(w))$.

Hypothetical parents are assigned to those individuals with only one known parent. Individuals in the pedigree will therefore have either both parents or none (founders). Moreover, in order to save CPU time, the uninformative parts of the pedigree are removed and only those individuals connected with the current population are kept. The herdbook of the Pisana includes 776 animals. The pedigree of the living population (livings) and the pedigree of the group of stocked bulls (stocked) are analysed separately (Table 1).

TABLE 1. Pedigree files after editing

file	total n. of animals in pedigree	n. of current (1) animals	n. of founders	n. of generations	generation completeness (2) 1°	2°	3°	4°	5°	6°
herdbook	776	-	135	-				-		
livings	292	174	51	6	1	.9	.6	.4	.2	.3
stocked	78	23	27	5	1	.7	.5	.3	.2	

(1) genotypes considered for the analysis of the probability distribution of founder alleles.
(2) generation completeness is given as number of parents observed/number of parents expected.

RESULTS AND DISCUSSION

Founders represented in the livings and in the stocked are respectively .38 and .20 of the 135 founders in the herdbook (Table 1). We first consider the probability distribution of the number of distinct founder alleles surviving in the livings and in the stocked. The maximum range of these distributions depends obviously from the number of founders and the number of current individuals. The mean number of distinct alleles per locus is 47.61 (s.d. 2.71, range 38-57) and 17.03 (s.d. 1.87, range 11-24) respectively for the livings and the stocked, corresponding to .47 and .31 of the alleles present among founders. The probability distributions, considering classes with frequency ≥ 005, are given in Table 2. In the livings, .68 of the loci are expected to have 46 to 49 distinct alleles. In the stocked, .70 of the loci shows 15 to 18 distinct alleles, .025 11 to 13 alleles.

Table 2 Livings: probability distribution of the number of distinct founder allelels

41	42	43	44	45	46	47	48	49	50	51	52	53	54
.009	.017	.034	.062	.092	.117	.146	.151	.126	.105	.066	.039	.020	.008

Stocked: probability distribution of the number of distinct founder allelels

13	14	15	16	17	18	19	20	21	22
.020	.060	.121	.194	.210	.177	.117	.064	.024	.008

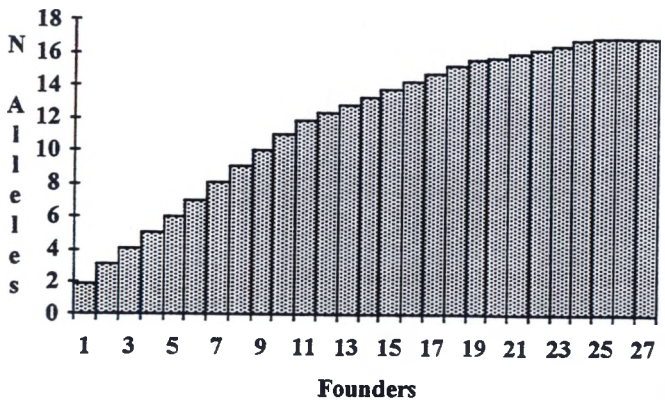

Figure 1: cumulative distribution of the number of distinct alleles by founder.

In Figure 1 the contribution of distinct alleles by each founder to the stocked, averaged over loci, is given. Each founder can obviously contribute a maximum of 2 distinct alleles. The cumulative distribution in the Figure shows as two founders contributes more than one allele (1.9 and 1.2), eight founders contributes one allele each and the remaining 17 founders less than one allele each. Thirteen of the 27 founders (.48) contributes 13 alleles of the 17.03 present at the mean locus (.70).

The expected number of copies by each distinct founder allele in the stocked is 2.7. However, under genetic drift we expect some variance. The number of copies can range from 1 to twice the number of the stocked. In Table 3 the probability distribution of the number of copies by distinct founder allele is given. Missing values may be related to insufficient iterations.

TABLE 3. Probability distribution of the number of copies by distinct founder allele in the stocked, given the number of distinct alleles survived, and at the mean locus (.000 equal < .001).

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
11	.136	.182	.197	.106	.136	.061	.075	.015	.015	.045	.015			.015								
12	.213	.213	.167	.122	.072	.062	.053	.023	.021	.009	.007	.007	.005	.005	.007	.002	.002	.009				
13	.275	.208	.143	.109	.090	.053	.038	.019	.013	.011	.009	.008	.007	.003	.004	.005	.003	.015	.000	.001		
14	.322	.191	.147	.108	.076	.048	.030	.017	.015	.012	.009	.006	.005	.005	.003	.002	.000	.000	.000	.000	.000	.000
15	.366	.192	.139	.096	.069	.043	.026	.017	.013	.010	.008	.005	.004	.004	.003	.001	.001	.001	.000	.000	.000	.000
16	.406	.185	.134	.093	.061	.037	.022	.016	.012	.009	.007	.005	.004	.003	.002	.001	.001	.000	.000	.000		
17	.439	.184	.128	.085	.056	.034	.020	.014	.001	.008	.006	.005	.003	.003	.001	.001	.000	.000	.000	.000		
18	.473	.175	.124	.083	.051	.030	.018	.013	.009	.007	.006	.004	.003	.003	.002	.001	.001	.000	.000	.000		
19	.506	.170	.119	.078	.045	.023	.017	.013	.008	.006	.006	.003	.003	.002	.001	.000	.000	.000	.000		.000	
20	.535	.165	.114	.073	.040	.020	.014	.012	.009	.005	.005	.003	.002	.001	.000	.000	.000	.000				
21	.554	.162	.117	.066	.035	.020	.014	.009	.007	.005	.003	.003	.002	.001	.000							
22	.578	.159	.112	.063	.027	.021	.014	.010	.007	.002	.002	.004	.001		.001							
23	.601	.167	.112	.044	.018	.018	.007	.015	.003	.008	.004	.004										
24	.667	.097	.097	.042	.028	.056			.014													
Mean	.443	.180	.128	.086	.056	.033	.020	.015	.011	.008	.006	.005	.003	.002	.002	.001	.001	.001	.000	.000	.000	.000

The maximum observed number of copies of a founder allele is 22 . However this high number is expected at the mean locus only with a probability of .00001. A number of copies from 1 to 3 is expected with probability of .75.

We can finally consider the frequencies of 13 alleles at the locus BoLa-DRB3 exon 2 (van Eijk et al., 1992) and compute, as an example, their probabilities of occurence and of being present with a single copy in the stocked. The BoLa frequencies of Table 4 are from a group of Fresian bulls (Damiani et al., in prep.) but we assume being the same in the Pisana founder population. Table 4 gives, for the different alleles, frequency among founders, probability of survival and probability of being present with a single copy in the stocked animals.

TABLE 4. Probability of survival and to be present with a single copy in the stocked for 13 Bola alleles.

BoLA-DRB3 allele	22	16	24	23	7	11	3.9	26	27	2.12
frequency among founders	.213	.175	.162	.125	.075	.062	.050	.037	.025	.013
Pr (survival)	.981	.960	.948	.894	.732	.661	.580	.472	.349	.199
Pr (one copy)	.034	.059	.070	.107	.158	.164	.163	.150	.124	.079

Survival probabilities are high for all BoLA alleles, included those with a low frequency in the founder population. This result was expected because the relative low complexity of the pedigree file of the stocked. Registration in the Pisana started recently and for several individuals (Table 1) the pedigree has only one or two generations known. However it should be noted that alleles 2 and 12 that have the lowest survival probability (.199), if not extinct, are expected with an high probability (.4) to be present in the stocked with only one copy.

CONCLUSIONS

The relationship matrix among founders and current animals gives the founders genetic contributions to each animal. These contributions are often averaged over the current animals as proportionate founder contributions. It should be noted that these contributions do not add since they are not independent. Consider, as an example, two full sibs: the probability for an allele from one grandparent to be present in both sibs is not simply given by the product of the probabilities to be present in each sib because it is conditional to the occurence in the parents. Therefore survival probabilities can not be computed from the relationship matrix. If the registered population has gone through tight bottlenecks and high differential reproductive success, survival probabilities of founder alleles can be low and therefore should be considered in evaluating the genetic variation of the stock. By gene dropping simulation we can also compute survival probabilities for each founder. This information can be used to select the animals to be stored in order to maximize the genetic variation in stock (Gandini et al. in prep.).

REFERENCES

GEYER, C.G. and THOMPSON, E.A. (1988) Zoo Biology, 7 : 313-327.

KIMURA, M. and CROW,J.F. (1964) Genetics, 49 : 725-738.

MAIJALA, K., CHEREKAEV, A.V., DEVILLARD, J.M., REKLEWSKI, Z., ROGNONI, G., SIMON, D.L., STEANE, D.E. (1984) Livestock Production Science, 11 : 3-22.

MACCLUER, J.W., VANDEBERG, J.L., READ, B., RYDER, O.A. (1986) Zoo Biology, 5 . 147-160.

SMITH, C. (1984) Livestock Production Science, 11 : 37-48.

VAN EIJK, M.J.T., STEWART-HAYNES, J.A., LEWIN, H.A. (1992) Animal Genetics, 23 : 483-496.