Original article

Pedigree and marker information requirements to monitor genetic variability

Roswitha BAUMUNG*, Johann SÖLKNER

Department of Livestock Sciences, University of Agricultural Sciences,
Gregor Mendel-Str. 33, 1180 Vienna, Austria

Abstract – There are several measures available to describe the genetic variability of populations. The average inbreeding coefficient of a population based on pedigree information is a frequently chosen option. Due to the developments in molecular genetics it is also possible to calculate inbreeding coefficients based on genetic marker information. A simulation study was carried out involving ten sires and 50 dams. The animals were mated over a period of 20 discrete generations. The population size was kept constant. Different situations with regard to the level of polymorphism and initial allele frequencies and mating scheme (random mating, avoidance of full sib mating, avoidance of full sib and half sib mating) were considered. Pedigree inbreeding coefficients of the last generation using full pedigree or 10, 5 and 2 generations of the pedigree were calculated. Marker inbreeding coefficients based on different sets of microsatellite loci were also investigated. Under random mating, pedigree-inbreeding coefficients are clearly more closely related to true autozygosity (*i.e.*, the actual proportion of loci with alleles identical by descent) than marker-inbreeding coefficients. If mating is not random, the demands on the quality and quantity of pedigree records increase. Greater attention must be paid to the correct parentage of the animals.

autozygosity / inbreeding / microsatellite / quality of pedigree

1. INTRODUCTION

In the initial stages of conservation, populations may not be concerned with genetic progress, but simply with conserving genetic variation. This means that the rate of inbreeding should be minimised. Various suggestions have been made to achieve this. In principle, two questions have to be answered: which animals to select and how to mate them? Caballero and Toro [3,4] discuss that the optimal choice of breeding individuals requires minimisation of the average coancestry among the reproductive individuals weighted by their contribution to the next generation. The same authors point out that the choice of the mating system is less simple because it depends on the

* Correspondence and reprints
E-mail: baumung@boku.ac.at

time scale of interest. In many practical breeding programmes the interest of conservation is more in the short rather than in the long term. In this case, avoidance of inbred matings seems to be appropriate [4]. If special breeding strategies are a condition for the financial support of endangered breeds, a critical judgement of the mating system becomes more important. This can be done with measures based on molecular genetic information [16]. Another way of measuring the mating system within a population is the comparison of the expected inbreeding coefficient under random mating (mean kinship in generation t) and the observed mean inbreeding coefficient in generation t + 1 based on pedigree information. This simple comparison allows a statement whether the average level of inbreeding is higher or lower than that expected under random mating conditions.

One weakness especially of the latter option is that the inbreeding coefficient depends very much on the quality of pedigree information. Developments in molecular genetics make it possible to calculate several measures based on genetic marker information. The aim of this study was to compare measures based on pedigree or genetic marker information with regards to the monitoring of endangered populations. Minimum requirements for the quantity and quality of the underlying source of information necessary to detect autozygous individuals (*i.e.*, individuals with a high proportion of alleles identical by state) were investigated. The correlation between such measures and true autozygosity serves as an indicator of the quality of the underlying source of information.

2. SIMULATION STUDY

A simulation study was carried out for a population with ten sires and 50 dams. For each animal a genome was modelled consisting of 20 pairs of chromosomes with 50 loci each. In most situations, a total length of 30 Morgans for the whole genome was assumed. Two further genome lengths with 100 Morgans and 10 Morgans were investigated as well. The recombination rate between neighbouring loci was 0.03, 0.001 and 0.01, respectively. All loci were assumed to be neutral with regards to selection. No mutation events were modelled.

Three simple mating schemes were examined: Random mating (scheme I), mating of full sibs was avoided (scheme II) and a third scheme in which mating of half sibs and full sibs was not permitted (scheme III). In all schemes, each female was permitted to produce a maximum of two offspring (full sibs). Ten males and 50 females were generated as potential parents for the next generation. Animals were observed over a period of 20 discrete generations.

The level of true autozygosity (proportion of loci with alleles identical by state) and homozygosity (proportion of loci with alleles alike in state) at the loci

of the whole genome was investigated for each animal of a reference population. The reference population was defined as the last simulated generation. True autozygosity was used as a reference for the measures described below.

The correlation between true autozygosity and several measures based on pedigree or marker information was calculated within each repetition for all animals of the reference population. The correlations presented in this paper are means and corresponding standard deviations of 100 repetitions for each situation.

2.1. Measures based on pedigree information

When pedigree inbreeding coefficients are computed in the sense of Malécot [13] or Wright [21], it is necessary to define the base population to which the present inbreeding is referred. In this case a "real" base population with unrelated individuals is present. In such a situation, the average inbreeding coefficient of the reference population can be taken as a measure for the true autozygosity. Under practical circumstances the "real" base population is never known, very often gaps and sometimes false parentage occur in pedigree records. These pedigree weaknesses influence the value of our measures for autozygosity. Three situations with regards to the quality of pedigrees were considered.

2.1.1. Length of pedigrees

Using a method described by VanRaden [19], we calculated inbreeding coefficients taking only 2, 5, 10 or the complete 20 generations into account. Studies on the genetic variability of several cattle breeds in Austria [15] and France [2] showed that a maximum number of 10 to 18 generations of animals in a defined reference population could be traced back. In the case of two highly endangered cattle breeds in Austria [1], this number was clearly lower (6 and 9).

2.1.2. Completeness of pedigrees

The maximum number of traceable generations gives no reliable information about gaps in the pedigree. A good way of describing the quality of a pedigree is the average complete generation equivalent (*i.e.*, number of generations in a comparable complete pedigree) [2]. This measure was found to be very high (15.22) in Lipizzan horses [22], but clearly lower (1.73–6.18) in many cattle breeds [1,2,15]. We reduced the simulated pedigrees to mirror the quality of pedigrees in a rare Austrian cattle breed. The maximum number of traceable generations was set to 6 and known ancestors in the reduced pedigree were randomly exchanged against unknown ancestors. This resulted in average complete generation equivalents of 2 to 3. These strongly reduced pedigrees were used to calculate inbreeding coefficients.

2.1.3. Correctness of pedigrees

Errors in pedigrees are known to occur due to mis-mothering, misidentification and incorrect recording procedures. Several studies [5,7,14] show that the misidentification rate in cattle pedigrees varies between about 3 and over 20%. Even in Lipizzan horse pedigrees where a great importance is attached to correct pedigree recording, a small number of pedigree errors has been revealed by mtDNA analysis [10]. To take this into account, 1, 5, 10 and 20% of the sires were exchanged randomly against wrong animals in each generation. Inbreeding coefficients were calculated according to these incorrect pedigrees.

2.2. Measures based on genetic marker information

Molecular technologies provide direct information on genotypes at polymorphic loci. Therefore it is possible to analyse the system of mating of a population as a deviation from the heterozygosity expected under Hardy-Weinberg equilibrium using the following formula [8,16]:

$$f = 1 - Ho/He \tag{1}$$

where He is the expected heterozygosity calculated from allele frequencies in a defined base population with random mating, and Ho is the observed heterozygosity in a reference population. We used a similar formula to derive individual inbreeding coefficients based on marker information:

$$f_{\text{gen}} = \frac{1}{n} \sum_{L=1}^{n} \left(1 - \frac{H_{oL}}{H_{eL}} \right) \tag{2}$$

where H_{eL} is the expected heterozygosity for marker locus L (with $L = 1, 2, 3, \ldots, n$) derived from the allele frequencies at locus L in the base population and H_{oL} the observed heterozygosity at locus L. Assuming that all alleles at homozygous loci in the simulated true base population are alike in state but not identical by descent, the increase of homozygosity (*i.e.* the proportion of homozygous loci) can be used to estimate the true autozygosity of a reference population. In addition to these marker-based inbreeding coefficients, the level of homozygosity was calculated for each animal. In reality, the number of analysed marker loci is restricted, and allele frequencies in the "true" base population are usually not known. Also, genetic markers show different polymorphism and allele frequencies. Several scenarios described below were considered.

2.2.1. Number of genetic markers

A marker inbreeding coefficient was calculated for different sets of 20, 50, 100 and 200 marker loci equally spaced over the whole genome to cover information from each chromosome. Genetic markers were assumed to be fully informative.

Table I. Assumed situations with regards to the number of alleles per locus and the initial allele frequencies in the base population.

Situation	Number Alleles/locus	Allele frequencies in the base population						
A	7	0.40	0.20	0.20	0.05	0.05	0.05	0.05
B	2	0.50	0.50	–	–	–	–	–
C	2	0.30	0.70	–	–	–	–	–
D	2	0.10	0.90	–	–	–	–	–

2.2.2. Number of alleles per marker locus and allele frequencies

Various types of genetic markers are currently used. We considered one situation with seven different marker alleles per locus in the base population to mimic a microsatellite marker. In addition, marker loci with two alleles with different frequencies in the base population were simulated to evaluate the effect of SNP markers. The marker loci represent just a small part of the total genome which was modelled in the same way as the marker loci (Tab. I). This results in an ideal situation because marker loci mirror the rest of the genome. All loci were assumed to be in Hardy-Weinberg equilibrium.

2.2.3. Definition of the base population

Inbreeding coefficients must be related to a base population or they are meaningless [9]. The expected heterozygosity for each marker locus was calculated from allele frequencies in the true base population and for base populations 2, 5 and 10 generations back from the reference population.

3. RESULTS AND DISCUSSION

3.1. Level of autozygosity and average pedigree inbreeding coefficients

Table II gives an overview of the results for pedigree inbreeding coefficients under random mating. The average inbreeding coefficient of animals in the reference population is a good measure for true autozygosity when pedigrees can be traced back to the true base population. This is still the case with complete pedigrees with 20% false parentage in each generation. With pedigrees reduced in length, true autozygosity is severely underestimated. Generally the standard deviations of replicates were similar for true average autozygosity and average pedigree inbreeding coefficients.

As expected, the level of true autozygosity was lower after 20 generations with avoidance of mating with close relatives (Tab. II). The potential to infer the average level of true autozygosity from pedigree inbreeding coefficients was not influenced by the mating scheme.

Table II. Arithmetic means $\bar{x}$ for investigated pedigree-inbreeding and marker-inbreeding coefficients of animals in the reference population and correlation with true autozygosity $\bar{r}$ for situation A, a total genome length of 30 Morgans and different mating schemes; means calculated from 100 repetitions, standard deviation in italics.

(continued on the next page)

	Random mating		No full sib mating		No full sib + no half sib mating	
	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$
True autozygosity	**0.254**		**0.249**		**0.232**	
	0.015		*0.017*		*0.014*	
True homozygosity	**0.441**	**0.949**	**0.436**	**0.944**	**0.425**	**0.900**
	0.011	*0.021*	*0.015*	*0.030*	*0.009*	*0.027*
Pedigree inbreeding coefficients based on different number of generations						
20 generations	**0.253**	**0.791**	**0.251**	**0.763**	**0.231**	**0.583**
	0.013	*0.080*	*0.015*	*0.114*	*0.011*	*0.083*
10 generations	**0.129**	**0.791**	**0.129**	**0.762**	**0.111**	**0.583**
	0.010	*0.080*	*0.010*	*0.115*	*0.008*	*0.084*
5 generations	**0.058**	**0.788**	**0.058**	**0.756**	**0.044**	**0.580**
	0.009	*0.083*	*0.009*	*0.116*	*0.005*	*0.083*
2 generations	**0.015**	**0.670**	**0.013**	**0.609**	**0.000**	–
	0.007	*0.134*	*0.006*	*0.161*	*0.000*	
Pedigree inbreeding coefficients based on reduced pedigree information (average complete generation equivalent 2.79)						
2.79 generations	**0.018**	**0.613**	**0.016**	**0.547**	**0.006**	**0.242**
	0.008	*0.163*	*0.007*	*0.174*	*0.003*	*0.181*
Pedigree inbreeding coefficients based on complete pedigrees with different percentage of false paternity per generation						
1%	**0.253**	**0.763**	**0.251**	**0.728**	**0.232**	**0.539**
	0.013	*0.104*	*0.014*	*0.119*	*0.011*	*0.127*
5%	**0.252**	**0.651**	**0.248**	**0.605**	**0.233**	**0.437**
	0.013	*0.169*	*0.014*	*0.199*	*0.011*	*0.123*
10%	**0.251**	**0.580**	**0.247**	**0.537**	**0.234**	**0.270**
	0.013	*0.194*	*0.014*	*0.167*	*0.012*	*0.170*
20%	**0.247**	**0.402**	**0.245**	**0.397**	**0.236**	**0.104**
	0.012	*0.178*	*0.013*	*0.183*	*0.019*	*0.154*

Table II. Continued.

	Random mating		No full sib mating		No full sib + no half sib mating	
	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$
Marker inbreeding coefficient based on different number of marker loci						
200 loci	**0.252**	**0.766**	**0.247**	**0.738**	**0.232**	**0.643**
	0.020	*0.069*	*0.023*	*0.095*	*0.016*	*0.089*
100 loci	**0.254**	**0.613**	**0.243**	**0.593**	**0.231**	**0.460**
	0.025	*0.104*	*0.027*	*0.109*	*0.021*	*0.119*
50 loci	**0.251**	**0.469**	**0.241**	**0.470**	**0.222**	**0.321**
	0.033	*0.117*	*0.032*	*0.097*	*0.030*	*0.127*
20 loci	**0.257**	**0.282**	**0.249**	**0.272**	**0.220**	**0.176**
	0.066	*0.145*	*0.043*	*0.144*	*0.052*	*0.141*

3.2. Relationship between true autozygosity and pedigree inbreeding coefficients

Even with 2 generation pedigrees, the correlation between autozygosity and pedigree inbreeding coefficients was rather high (0.670) in situations with random mating (Tab. II). Therefore it seems to be possible to identify the most autozygous animals assuming parents and grandparents are known. Taking more than five generations of a correct pedigree into account leads only to a marginal increase of the correlation of pedigree inbreeding coefficients and autozygosity. Under random mating, inbreeding coefficients based on pedigrees with very low quality (reduced pedigrees) were still highly related to true autozygosity (0.613). In the case of 10% and 20% incorrect paternity in a complete pedigree, this correlation dropped to 0.58 and 0.40, respectively. The occurrence of false parentage of 20% or more seems to be a more severe problem for the identification of the most autozygous animals than does incompleteness and shortness of pedigrees. In cases of incomplete or short pedigrees the autozygosity of single animals or all animals within a generation is underestimated to the same extent, respectively. If there is false parentage, two types of errors might occur using pedigree inbreeding coefficients: underestimation and overestimation of the true autozygosity of individuals. The relationship between true autozygosity and pedigree inbreeding coefficients must become less close if correct parents are exchanged randomly against false ones so that the correlation drops. A noticeable higher standard deviation for the mean correlation over the 100 repetitions was observed in cases of pedigrees with false parentage compared to short and incomplete pedigrees.

In conservation breeding programmes, mating of close relatives is purposely avoided. Common ancestors could not be detected in short (two generations) and strongly reduced pedigrees when full- and half sib mating was strictly avoided (Tab. II). With false parentage it is almost impossible to make a statement about which individuals are highly autozygous.

3.3. Level of autozygosity and marker-based inbreeding coefficients

The results of inbreeding coefficients based on genetic marker information with an underlying total length of the genome of 30 Morgans are shown in Tables III and IV. Table III comprises the results for microsatellite markers (situation A). As in the case of pedigree-inbreeding coefficients, the true base population must be known to get a good estimate of the true level of autozygosity in a population. Otherwise true autozygosity is underestimated and the average marker-based inbreeding coefficient simply estimates the increase in homozygosity with regards to the defined base population (2, 5 or 10 generations back). In reality, allele frequencies several generations back are usually not known. In such a situation, no meaningful results for the level of autozygosity are obtained, if the expected heterozygosity is calculated on the allele frequencies of the current (reference) population (Tabs. III and IV). To get meaningful results on the evolution of genetic variability, genotyping of animals in different generations is necessary.

If only animals from the reference population were genotyped, the observed heterozygosity was usually slightly higher than the expected one resulting in negative values for the average level of marker inbreeding in the reference population (equation (1)). Templeton and Read [16] state that such negative values can be expected in finite populations with separate sexes because of random differences in allele frequency between sexes.

The average level of marker based inbreeding in the reference population was not influenced by the number of marker loci (Tab. III) but the standard deviation between replicates decreased with increasing number of marker loci. This can be explained by the increase of sample size (*i.e.* number of loci): the higher the number of marker loci the more reliable is the marker-based inbreeding coefficient.

Clearly, a high number of marker loci and knowledge about allele frequencies in the base population are necessary to get a reliable estimator for the level of autozygosity. In contrast, a quite low number of marker loci (20) was sufficient to measure the level of true homozygosity (Tab. III).

3.4. Relationship between true autozygosity and marker inbreeding coefficients

The correlations between marker inbreeding coefficients and true autozygosity show clearly that a rather high number of polymorphic loci must be analysed

Table III. Arithmetic means $\bar{x}$ for investigated marker-inbreeding coefficients with a different number of genetic markers of animals in the reference population and correlation with true autozygosity $\bar{r}$ for situation A; a total genome length of 30 Morgans and random mating, means calculated from 100 repetitions, standard deviations in italics.

				Number of marker loci				
	20		50		100		200	
	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$
Marker homozygosity								
	0.443	**0.282**	**0.438**	**0.469**	**0.440**	**0.613**	**0.439**	**0.766**
	0.049	*0.145*	*0.025*	*0.117*	*0.019*	*0.104*	*0.015*	*0.069*
Marker inbreeding coefficients with expected heterozygosity calculated from allele frequencies for different numbers of generations back from the reference population								
20 generations	**0.257**	**0.282**	**0.251**	**0.469**	**0.254**	**0.613**	**0.252**	**0.766**
	0.066	*0.145*	*0.033*	*0.117*	*0.025*	*0.104*	*0.020*	*0.069*
10 generations	**0.126**	**0.269**	**0.117**	**0.448**	**0.121**	**0.586**	**0.117**	**0.730**
	0.080	*0.140*	*0.042*	*0.111*	*0.030*	*0.102*	*0.021*	*0.079*
5 generations	**0.043**	**0.247**	**0.037**	**0.424**	**0.038**	**0.552**	**0.033**	**0.690**
	0.067	*0.147*	*0.040*	*0.118*	*0.028*	*0.121*	*0.023*	*0.109*
2 generations	**0.009**	**0.238**	**0.007**	**0.410**	**0.007**	**0.526**	**0.005**	**0.664**
	0.052	*0.149*	*0.033*	*0.134*	*0.023*	*0.127*	*0.020*	*0.116*
0 generations	**−0.023**	**0.229**	**−0.026**	**0.391**	**−0.024**	**0.512**	**−0.024**	**0.661**
	0.028	*0.161*	*0.018*	*0.135*	*0.014*	*0.128*	*0.013*	*0.097*

in order to detect autozygous individuals. The correlations between homozygosity calculated with different numbers of marker loci and true autozygosity are also shown. These correlations are identical to those derived using marker inbreeding coefficients, where the expected heterozygosity was calculated from allele frequencies in the real base population. It must be pointed out that even with 100 marker loci, this correlation is only as high as with very poor pedigree information (reduced pedigree information, r = 0.613) under random mating. The standard deviation for the mean correlation of the 100 replicates with marker based inbreeding coefficients is quite high compared to the correlations with pedigree inbreeding coefficients based on short pedigrees. Only with 200 marker loci is the standard deviation for the correlation as low or even lower than for pedigree inbreeding coefficients based on five or more generation-pedigrees. Several studies dealing with marker-based kinship measures [6, 12] showed that a high number of polymorphic markers is necessary to obtain reliable estimates for the relatedness of individuals. Eding and Meuwissen [6] concluded that by studying the scenarios presently used in the studies of genetic

Table IV. Arithmetic means $\bar{x}$ for investigated marker-inbreeding coefficients based on 200 genetic marker loci of animals in the reference population and correlation with true autozygosity $\bar{r}$ for total genome length of 30 Morgans, random mating and different situations with regards to the number of alleles per marker and initial allele frequencies (see Tab. I); means calculated from 100 repetitions, standard deviation in italics.

Measure	Situation A		Situation B		Situation C		Situation D	
	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$
True	**0.254**		**0.253**		**0.254**		**0.254**	
autozygosity	*0.015*		*0.016*		0.015		*0.016*	
True	**0.441**	**0.949**	**0.626**	**0.863**	**0.688**	**0.837**	**0.866**	**0.667**
homozygosity	*0.011*	*0.021*	*0.009*	*0.045*	*0.009*	*0.062*	*0.007*	*0.109*
Marker inbreeding coefficients with expected heterozygosity calculated from allele frequencies for different numbers of generations back from the reference population								
20 generations	**0.252**	**0.766**	**0.253**	**0.589**	**0.259**	**0.538**	**0.262**	**0.384**
	0.020	*0.069*	*0.027*	*0.107*	*0.037*	*0.107*	*0.065*	*0.131*
10 generations	**0.117**	**0.730**	**0.119**	**0.549**	**0.129**	**0.453**	**0.134**	**0.259**
	0.021	*0.079*	*0.032*	*0.112*	*0.048*	*0.114*	*0.113*	*0.146*
5 generations	**0.033**	**0.690**	**0.032**	**0.500**	**0.035**	**0.403**	**0.035**	**0.257**
	0.023	*0.109*	*0.035*	*0.125*	*0.044*	*0.111*	*0.096*	*0.129*
2 generations	**0.005**	**0.664**	**0.000**	**0.483**	**0.004**	**0.403**	**0.017**	**0.262**
	0.020	*0.116*	*0.032*	*0.113*	*0.035*	*0.110*	*0.066*	*0.141*
0 generations	**−0.024**	**0.661**	**−0.025**	**0.463**	**−0.022**	**0.406**	**−0.026**	**0.240**
	0.013	*0.097*	*0.013*	*0.117*	*0.012*	*0.109*	*0.016*	*0.129*
Marker homozygosity								
	0.439	**0.766**	**0.626**	**0.589**	**0.689**	**0.538**	**0.867**	**0.384**
	0.015	*0.069*	*0.013*	*0.107*	*0.016*	*0.107*	*0.012*	*0.131*

diversity with 10–15 loci, it is impossible to distinguish even full sibs from half sibs. Our study shows that with such a low number of marker loci, the detection of highly inbred (autozygous) animals is not possible.

The level of polymorphism and the allele frequencies also influence the usefulness of a genetic marker. Table IV shows that higher polymorphic markers and even allele frequencies are more useful in identifying autozygous animals, which corresponds with results from Toro *et al.* [17].

The information content of a set of genetic marker loci is also influenced by the length of the whole genome. In our simulation study, we considered three simple cases in which the length of the genome depended only on the given recombination rate between neighbouring loci. Table V shows the results

Table V. Arithmetic means $\bar{x}$ for pedigree inbreeding coefficients and marker inbreeding coefficients and their correlation with true inbreeding $\bar{r}$ under random mating and different lengths of the total genome in Morgans, means calculated from 100 repetitions, standard deviations in italics.

Measure	Length of total genome					
	10 Morgans		30 Morgans		100 Morgans	
	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$	$\bar{x}$	$\bar{r}$
True	**0.255**		**0.254**		**0.253**	
autozygosity	*0.018*		*0.015*		*0.013*	
True	**0.441**	**0.968**	**0.441**	**0.949**	**0.439**	**0.935**
homozygosity	*0.014*	*0.011*	*0.011*	*0.021*	*0.011*	*0.028*
Pedigree inbreeding coefficients based on a different numbers of generations or reduced pedigree (average complete generation equivalent 2.79)						
20 generations	**0.252**	**0.628**	**0.253**	**0.791**	**0.252**	**0.896**
	0.013	*0104*	*0.013*	*0.080*	*0.013*	*0.047*
2 generations	**0.014**	**0.503**	**0.015**	**0.670**	**0.014**	**0.693**
	0.007	*0170*	*0.007*	*0.134*	*0.007*	*0.132*
2.79 generations	**0.017**	**0.457**	**0.018**	**0.613**	**0.018**	**0.693**
	0.007	*0177*	*0.008*	*0.163*	*0.007*	*0.160*
Marker inbreeding coefficients with expected heterozygosity based on allele frequencies 20 generations back, and for different numbers of typed loci						
200 loci	**0.257**	**0.838**	**0.252**	**0.766**	**0.251**	**0.681**
	0.024	*0.058*	*0.020*	*0.069*	*0.022*	*0.097*
100 loci	**0.257**	**0.714**	**0.254**	**0.613**	**0.253**	**0.539**
	0.030	*0.080*	*0.025*	*0.104*	*0.027*	*0.105*
50 loci	**0.257**	**0.561**	**0.251**	**0.469**	**0.257**	**0.416**
	0.037	*0.106*	*0.033*	*0.117*	*0.032*	*0.126*
20 loci	**0.267**	**0.319**	**0.257**	**0.282**	**0.258**	**0.264**
	0.055	*0.140*	*0.066*	*0.145*	*0.049*	*0.136*

for a fixed number of loci and chromosomes but different lengths of the total genome (10, 30 and 100 Morgans, respectively). For the same number of marker loci the correlation between the marker inbreeding coefficient and true autozygosity is closer for smaller genomes due to the smaller recombination rate.

The correlations between true autozygosity and marker inbreeding coefficients drop slightly in the case of non-random mating schemes. The decrease of these correlations is lower than the observed decrease for the pedigree inbreeding coefficients (Tab. II).

3.5. Pedigree inbreeding coefficients *versus* marker inbreeding coefficients

In all situations under random mating with less than 100 marker loci, the correlation between autozygosity and marker inbreeding coefficients (Tab. III) was lower than with pedigree inbreeding coefficients (except for pedigrees with 20% false paternity, see Tab. II). Only ideal genetic marker loci with regards to codominance and possible mutations were considered. Mutations and genotyping errors would lead to even worse results for marker based inbreeding coefficients. Therefore, marker inbreeding coefficients do not appear to be a favourable method for identifying autozygous animals when reliable pedigree information is available. An additional point is that in most cases we assume the allele frequencies in a defined base population to be known, which is quite unrealistic. Wang [20] points out that allele frequencies have to be estimated from samples and are therefore subject to sampling errors. For highly polymorphic markers and realistic sample sizes, the sampling errors for allele frequencies cannot be ignored. In our study we assumed that the total population was genotyped. Therefore, in reality even less favourable results for marker inbreeding coefficients must be expected. Toro *et al.* [18] made an extensive comparison of several estimators of coancestry based on molecular markers. They also calculated the correlation between such estimators and the genealogical coancestry. They observed that the use of the true allelic frequencies in the base population increases the correlation of the investigated estimators with the genealogical coancestry. These results are in agreement with our results (Tab. IV). Toro *et al.* [18] also calculate coancestry estimators using allelic frequencies in the actual population. They found discrepancies between the estimations of coancestry obtained from molecular information and the values of the genealogical coancestry, which are comparable to our results in Tables III and IV, where the expected heterozygosity was calculated from allele frequencies in the reference population. The lack of information on the true allelic frequencies in the base population cannot be compensated for with a larger number of investigated loci. This has clearly been shown for coancestry estimators [18] and the marker inbreeding coefficients (Tab. III) calculated here.

Under non-random mating conditions the picture changes slightly. A complete pedigree with less than 5 percent of false paternity or a correct and complete 5-generation pedigree leads to a closer correlation of pedigree inbreeding coefficients with true autozygosity than marker pedigree coefficients based on 100 or less microsatellite marker loci.

Of course, under the avoidance of full and half sib mating, no conclusions can be drawn from pedigree inbreeding coefficients (based on a 2-generation pedigree) about the most inbred animals.

Marker inbreeding coefficients based on 200 evenly spaced microsatellite marker loci would lead to almost the same (0.738) or an even closer (0.643) correlation than the corresponding pedigree inbreeding coefficients based on perfectly complete pedigree records (0.763 and 0.583) under avoidance of full sib mating or avoidance of full and half sib matings, respectively (Tab. II).

4. CONCLUSIONS

Several papers that deal with the coancestry or kinship between individuals within a population were published only recently [6, 12, 18, 20]. The average kinship between the parents selected from generation t equals the average inbreeding coefficient of their offspring in generation t + 1 (assuming random mating of the parents). This suggests that we can control inbreeding by controlling the average kinship of the parents [11]. For that reason many papers published focus on kinship. Nevertheless many conservation breeding programmes, including a number of them in Austria, are based on the maximum avoidance of inbreeding, *i.e.* mating pairs are chosen such that the inbreeding coefficients of the offspring are as small as possible. Animals with high inbreeding coefficients are excluded from financial support and further utilization as breeding animals. This rule serves to avoid the deliberate mating of close relatives. To control the compliance with such rules it is important to have reliable information on average and individual inbreeding coefficients. If special breeding strategies are a condition for the financial support of endangered breeds, a critical judgement of the previous mating systems becomes more important.

The simulations show that even pedigrees of low quality allow the identification of the most autozygous animals in a random mating population. Measures based on codominant marker loci lead to comparable results only when more than 100 (better 200) microsatellite loci are typed.

With regards to conservation breeding programmes, it seems wise to use pedigree information whenever available especially considering the costs of genotyping for 100 or more microsatellite marker loci. The situation might change if typing of a massive number of genotypes per individual at a reasonable cost is a feasible task. Genotyping of animals for the purpose of monitoring the genetic variability in small populations needs to be a continuous process. Molecular markers must be investigated over several generations. Under a non-random mating scheme, the demands on quality and quantity of pedigree records increase. Deeper pedigrees are necessary to draw conclusions from pedigree inbreeding coefficients on autozygosity. To control past breeding strategies in practical conservation breeding programmes, complete five generation pedigee data should be obtained. Greater attention must be paid to the correct parentage of animals. Parentage verification with genetic markers is a valuable tool.

ACKNOWLEDGEMENTS

REFERENCES

[1] Baumung R., Sölkner J., Analysis of pedigrees of Tux-Zillertal, Carinthian Blond and Original Pinzgau cattle populations in Austria, J. Anim. Breed. Genet. 119 (2002) 175–181.

[2] Boichard D., Maignel L., Verrier É., The value of using probabilities of gene origin to measure genetic variability in a population, Genet. Sel. Evol. 29 (1997) 5–23.

[3] Caballero A., Toro M.A., Interrelations between effective population size and other pedigree tools for the management of conserved populations, Genet. Res. 75 (2000) 331–343.

[4] Caballero A., Toro M.A., Analysis of genetic diversity for the management of conserved subdivided populations, Conserv. Genet. 3 (2002) 289–299.

[5] Christensen L.G., Madsen P., Petersen J., The influence of incorrect sire-identification on the estimates of genetic parameters and breeding values, in: Proceedings of the 2nd World Congress on Genetics Applied to Livestock Production, Madrid, 4–8 October, 1982, Vol. 7, pp. 200–208.

[6] Eding H., Meuwissen T., Marker-based estimates of between and within population kinships for the conservation of genetic diversity, J. Anim. Breed. Genet. 118 (2001) 141–159.

[7] Geldermann H., Pieper U., Weber W.E., Effects of misidentification on the estimation of breeding value and heritability in cattle, J. Anim. Sci. 63 (1986) 1759–1768.

[8] Jacquard A., Inbreeding: One word, several meanings, Theor. Pop. Biol. 7 (1975) 338–363.

[9] Falconer D.S., Mackay T.F.C., Introduction to quantitative genetics, 4th edn., Longmann, London, 1996.

[10] Kavar T., Brem G., Habe F., Sölkner J., Dovč P., History of Lipizzan horse maternal lines as revealed by mtDNA analysis, Genet. Sel. Evol. 34 (2002) 635–638.

[11] Lacey R.C., Ballou J.D., Princée F., Starfield A., Thompson E.A., Pedigree analysis for population management, in: Ballou J.D., Gilpin M., Foose T.J. (Eds.), Population management for survival and recovery, Columbia university Press, New York, 1995, pp. 57–75.

[12] Lynch M., Ritland K., Estimation of pairwise relatedness with molecular markers, Genetics 152 (1999) 1753–1766.

[13] Malécot G., Les mathématiques de l'hérédité, Masson, Paris, 1948.

[14] Ron M., Blanc Y., Band M., Ezra E., Weller J.I., Misidentification rate in the Israeli dairy cattle population and its implications for genetic improvement, J. Dairy Sci. 79 (1996) 676–681.

[15] Sölkner J., Filipcic L., Hampshire N., Genetic variability of populations and similarity of subpopulations in Austrian cattle breeds determined by analysis of pedigree, Anim. Sci. 67 (1998) 249–256.

[16] Tempelton A.R., Read B., Inbreeding: One word, several meanings, much confusion, in: Loeschke V., Tomiuk J., Jain S.K. (Eds.), Conservation Genetics, Birkhäuser Verlag Basel, Switzerland, 1994, pp. 91–105.

[17] Toro M.A., Silió L., Rodrigánez J., Rodriguez C., Fernández J., Optimal use of genetic markers in conservation programmes, Genet. Sel. Evol. 31 (1999) 255–261.

[18] Toro M.A., Barragán C., Óvilo C., Rodriganez J., Rodriguez C., Silió L., Estimation of coancestry in Iberian pigs using molecular markers, Conserv. Genet. 3 (2002) 309–320.

[19] VanRaden P.M., Accounting for inbreeding and crossbreeding in genetic evaluation for large populations, J. Dairy Sci. 75 (1992) 3136–3144.

[20] Wang J., An Estimator for Paorwise Relatedness Using Molecular Markers, Genetics 160 (2002) 1203–1215.

[21] Wright S., Coefficients of inbreeding and relationship, Am. Nat. 56 (1922) 330–338.

[22] Zechner P., Sölkner J., Druml T., Baumung R., Achmann R., Bodo I., Marti E., Habe F., Brem G., Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information, Livest. Prod. Sci. 77 (2002) 137–146.