



**HAL**  
open science

## Effect of linkage on the control of inbreeding in selection programmes

Blanca Fernández, Enrique Santiago, Miguel Toro, Armando Caballero

► **To cite this version:**

Blanca Fernández, Enrique Santiago, Miguel Toro, Armando Caballero. Effect of linkage on the control of inbreeding in selection programmes. *Genetics Selection Evolution*, 2000, 32 (3), pp.249-264. 10.1051/gse:2000117 . hal-00894334

**HAL Id: hal-00894334**

**<https://hal.science/hal-00894334>**

Submitted on 11 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Effect of linkage on the control of inbreeding in selection programmes

Blanca FERNÁNDEZ<sup>a</sup>, Enrique SANTIAGO<sup>b</sup>,  
Miguel A. TORO<sup>c</sup>, Armando CABALLERO<sup>a,\*</sup>

<sup>a</sup> Departamento de Bioquímica, Genética e Inmunología, Facultad de Ciencias, Universidad de Vigo, 36200 Vigo, Spain

<sup>b</sup> Departamento de Biología Funcional, Facultad de Biología, Universidad de Oviedo, 33071 Oviedo, Spain

<sup>c</sup> Departamento de Mejora, Genética y Biotecnología, INIA, Carretera de la Coruña Km. 7, 28040 Madrid, Spain

(Received 29 June 1999; accepted 14 February 2000)

**Abstract** – Selection and mating methods for controlling inbreeding in selection programmes are based on relationships obtained from pedigrees. The efficiency of these methods has always been tested by studies using genetic models of independent loci. However, under linkage the rate of inbreeding obtained from pedigrees can be different from the probability of identity by descent of genes. We simulated a quantitative trait under artificial selection controlled by a large number of genes spread on genome regions of different sizes. A method to control inbreeding based on minimising the average coancestry of selected individuals with a restriction in the loss of selection response, and a mating procedure to control inbreeding were applied. These methods, that use coancestry relationships, were not effective in controlling inbreeding when the genome sizes were smaller than five morgans or so. However, for larger genome sizes the methods were sufficiently efficient. For very tight linkage, methods that utilise molecular information from markers should be used. We finally discuss the effects of the selection of individual major genes on the neutral variability of adjacent genome regions.

**inbreeding / effective population size / genetic drift / heterozygosity / BLUP**

**Résumé** – L'effet de la liaison des gènes sur le contrôle de la consanguinité dans les programmes de sélection. Les méthodes de sélection et les systèmes d'accouplement permettant de freiner l'augmentation de la consanguinité dans les schémas de sélection sont fondés sur les relations génétiques obtenues à partir des généalogies. Jusqu'à maintenant, leur efficacité avait toujours été testée à partir de modèles génétiques sans linkage. Cependant, si les gènes sont liés le coefficient de consanguinité obtenu à partir de la généalogie et la probabilité d'identité des gènes neutres peuvent être différents. Aussi, nous avons simulé un caractère quantitatif déterminé par un grand nombre de gènes répartis également dans le génome. Le

\* Correspondence and reprints  
E-mail: armando@uvigo.es

degré de liaison entre les gènes était imposé par le choix de la taille du génome. Pour contrôler l'augmentation de la consanguinité, on a utilisé une méthode fondée sur la parenté moyenne des candidats sélectionnés et comportant une restriction sur la perte de gain génétique associée, ainsi qu'un système d'accouplement particulier. Ces méthodes qui emploient les coefficients de parenté ne sont pas efficaces si le génome fait moins de cinq morgans environ, mais elles le deviennent pour des tailles supérieures. En cas de liaison très forte entre les gènes, les méthodes qui emploient l'information apportée par les marqueurs moléculaires neutres devraient être utilisées. Enfin, les effets de la sélection de gènes majeurs sur la variation des marqueurs neutres dans les régions adjacentes ont été discutés.

**consanguinité / effectif génétique / dérive génétique / hétérozygotie / BLUP**

## 1. INTRODUCTION

Inbreeding leads to the exhaustion of genetic variation and depression in reproductive traits, reducing the capability of populations for its commercial exploitation. This has been a matter of concern, particularly in recent years, because of the increased rates of inbreeding in small nucleuses of selection. A great deal of research effort has been invested in developing selection and mating procedures with the objective of controlling inbreeding in selection programmes. All selection methods proposed are based, in some way or another, on the same principle. The idea is to reduce the average coancestry among selected individuals, the probability of identity by descent of two genes taken from the population. This can be done indirectly by reducing the weight given to family information in the selection decisions (see Grundy *et al.* [10] for the latest development of this idea). Alternatively, a more direct method is the maximisation of a selection objective combining breeding values and coancestries of individuals (*e.g.* [5,9,11]). Mating schemes proposed to control inbreeding include a reduction in the number of full-sib progeny through factorial designs, mating among the least related animals, and compensatory mating [3,20].

All the above methods have been investigated using genetic models of independent loci. This can be a valid assumption when loci affecting a given selected quantitative trait are spread over a large genome, but when the size of the genome is not large, or when the quantitative trait loci are restricted to few chromosomes or genome regions of low recombination, linkage can be important. Selection under linkage produces a particularly large increase in the build up of linkage disequilibrium, with a consequent reduction in genetic response and increase in inbreeding [13]. Under selection and linkage, the rate of inbreeding obtained from pedigrees can be very different from the probability of homozygosity of genes within individuals. This is because the two copies of a neutral gene in an individual may not have the same probability of being transmitted to the next generation, since they are embedded in different chromosomes carrying different combinations of selected genes [16].

The objective of this work was to find out the genome size below which linkage reduces the effectiveness of the methods for controlling inbreeding

in selection programmes. We simulated a quantitative trait under artificial selection controlled by a large number of genes spread on genome regions of different sizes. A method to control inbreeding based on minimising the average coancestry of selected individuals with a restriction in the loss of selection response was applied. We also studied the efficiency of a mating procedure to control inbreeding, the so-called compensatory mating. This was implemented by sorting selected individuals according to the average coancestry with all other selected individuals, and mating males with the highest average value with females with the lowest [3]. Both the selection and mating methods used coancestry relationships obtained from pedigrees and, therefore, assumed no linkage. However, we found that they were effective in controlling inbreeding unless the genome sizes were smaller than five morgans or so.

## 2. MODEL ASSUMPTIONS AND SIMULATIONS

We simulated a population with  $M$  male and  $F$  female parents, where each male was mated to  $F/M$  females. From each cross,  $n$  offspring of each sex was obtained, so the proportion selected was  $M/nF$  for males and  $1/n$  for females. Ten generations of selection with discrete generations were carried out, averaging results over 1 000 (phenotypic selection) or 600 (BLUP selection) replicates. In the initial generation, the quantitative trait was normally distributed with a phenotypic variance of one, and additive variance  $\sigma_A^2$ . Environmental deviations common to members of the families and those for individuals were normally distributed with variance  $\sigma_C^2$  and  $\sigma_E^2 = 1 - \sigma_A^2 - \sigma_C^2$ , respectively.

The quantitative trait was controlled by 200 biallelic loci with equal effects, initial average frequency 0.5, and additive within and between-locus gene action. Other numbers of selected loci, variable effects and variable gene frequencies were also run, but the main results obtained were similar, and are not shown. Two hundred neutral multiallelic loci were also simulated between selected loci. Individuals at generation zero carried different alleles at these neutral loci in order to calculate probabilities of identity by descent per locus.

Theoretical studies have shown that, under selection with the infinitesimal model, the critical parameter regarding genetic drift [16] and linkage disequilibrium [14] is the total genome length, irrespective of the number of chromosomes. Thus, in the simulations, neutral and selected loci were equally spaced in a single chromosome  $L$  morgans long. Cross-overs occurred without interference, so that the number of cross-overs followed a Poisson distribution. Genome sizes of  $L = 0$  (no recombination), 0.5, 1, 5, 10, 20, and  $L = \infty$  (free recombination) were investigated. The population was assumed to be initially at linkage equilibrium.

Every generation, the coancestries among individuals were calculated from pedigrees ( $f_{\text{ped}}$ ) and from neutral molecular loci ( $f_{\text{mol}}$ ), building up the corresponding coancestry matrices of size  $(M + F) \times (M + F)$ . The molecular coancestry between two individuals is the probability of identity in state of two genes taken at random from each individual. It was calculated by comparing the alleles in a locus of one individual with those of the other individual, counting the proportion of allelic coincidences among the four comparisons,

and averaging this proportion over all loci. Average coancestries were calculated giving a weight  $F/M$  to those combinations where a male was involved, and a weight  $(F/M)^2$  to those where two males were involved. While pedigrees give the expected coancestries between individuals, the neutral loci give the actual probability of identity in state averaged for all 200 loci. The degree of coincidence between molecular and pedigree coancestries was measured by the correlation ( $r$ ) between the corresponding values for the whole coancestry matrix. The amount of variation in the relationships was measured by the variances,  $\sigma_{f_{ped}}^2$  and  $\sigma_{f_{mol}}^2$ , respectively, of pairwise coancestries among individuals (including individuals with themselves).

### 2.1. Selection methods

Truncation selection was made so that  $M$  males and  $F$  females were chosen as parents, but the selection criteria varied as follows:

- MS and BLUP: Selection based on the phenotypic values (mass selection), or on the estimated breeding values from BLUP evaluations,
- MCSped: Minimum coancestry selection using pedigrees. Selection of individuals (phenotypic selection or BLUP) with minimum average coancestry, being calculated from pedigree relationships,
- MCSmol: Minimum coancestry selection using molecular information. Selection of individuals (phenotypic selection or BLUP) with the minimum average molecular coancestry obtained from the probabilities of identity in state of neutral molecular markers. The number of markers used was one or four per morgan (for  $L = \infty$ , 20 or 100 markers were used) equally spaced along the chromosome, with 8 alleles each, initially segregating at equal frequencies. Cases were also run where all neutral loci (200) with all the alleles (initially different) were used as markers in the selection decisions.

Methods MCSped and MCSmol were applied with the restriction that the selection differentials were above a certain percentage of that obtained with mass or BLUP selection. Both methods were implemented by simulated annealing algorithms (*e.g.*, [5]). This was carried out as follows. For every generation, the individuals with the highest phenotypic (BLUP) score were first selected to obtain the expected selection differential and average coancestry among selected individuals for MS (BLUP). Selection differentials were calculated with an  $F/M$  weight given to males. Then, a random combination of  $M$  males and  $F$  females was chosen from all the available individuals with the only restriction that the selection differential was within a given percentage (typically 2.5%) of that obtained previously under MS (BLUP). For each iteration, a selected individual was randomly taken and substituted by another one of the same sex taken randomly from among the non-selected individuals. The selection differential was calculated again after the substitution. If the restriction was not held, the substitution was reversed and a new individual was chosen. If it was held, the average coancestry was calculated with the new individual incorporated. If the average coancestry was smaller than the previous one, the substitution was maintained. Otherwise the change was still maintained with a given probability of  $\Omega = \exp(-\Delta/T)$ , where  $\Delta$  is the difference between

the average coancestry of the new combination and that of the previous one, and  $T$  is the so-called temperature. The initial temperature was one, and it was reduced by 0.1 every 200 iterations. With this approximate method, it was possible to reach the optimum solution or a close one when the number of iterations was sufficiently large. The process was finished when no change was made after 200 iterations, or when 20 000 iterations had been done.

## 2.2. Systems of mating

The usual system of mating followed is random mating, but a non-random mating system used to control inbreeding was also assessed. This was a modification [3] of the system of mating proposed by Santiago and Caballero [15], and called compensatory mating. In essence, selected individuals were sorted according to the average coancestry with the other selected individuals, the males being weighted by the ratio of sexes as explained before. Males with the highest average were mated to females with the lowest, in sequence. With this system of mating, transmission lines of families with low and high selective success are mixed up, and genetic drift due to selection is compensated by genetic drift due to random sampling [15]. The result is that, under free recombination, the cumulative effect of selection on drift due to associations between neutral genes and selective values is practically removed in one generation time. Coancestries could again be obtained from pedigrees (CMped: compensatory mating using pedigree relationships) or from molecular markers (CMmol: compensatory mating using molecular relationships).

## 3. RESULTS

Table I shows cumulative responses to generations 5 ( $R_5$ ) and 10 ( $R_{10}$ ), averaged over replicates. Rates of average coancestry [ $\Delta f = (f_t - f_{t-1})/(1 - f_{t-1})$ ], and variances and correlation of pedigree and molecular coancestries are averages of replicates and generations 6 and 10 (asymptotic values for these parameters were generally obtained by generation 6).

MS columns indicate the results in absolute values for mass selection. MCSped columns indicate the proportional reductions in response and in the rate of coancestry with the MCSped method relative to mass selection. The proportional loss in response when the method of control of inbreeding was applied was similar for the different genome sizes (about 1–4%), and somewhat higher than the maximum restriction in selection differential (2.5%). This was due to the fact that the accuracy of selection was also reduced when using the method for controlling inbreeding, because proportionately more weight was given to the deviations within families. From the observational components of variance between sires, between dams within sires and within dams obtained in the simulations, it is possible to predict these reduced selection responses [4]. The loss in response was proportionally less in later generations ( $R_{10}$  vs.  $R_5$ ) because of the preservation of genetic variation with the method for controlling inbreeding.

**Table I.** Response to selection and rates of coancestry for different genome sizes.

	$L = 0$		$L = 0.5$		$L = 1$		$L = 5$		$L = \infty$	
	MS	MCSped	MS	MCSped	MS	MCSped	MS	MCSped	MS	MCSped
$R_5$	1.986	-3%	2.154	-2%	2.233	-4%	2.358	-4%	2.389	-4%
$R_{10}$	2.293	-1%	3.234	-2%	3.553	-3%	4.105	-3%	4.330	-3%
$\Delta f_{ped}$	0.0229	-21%	0.0278	-23%	0.0283	-23%	0.0296	-23%	0.0298	-23%
$\Delta f_{mol}$	0.1427	-9%	0.0805	-14%	0.0683	-20%	0.0424	-22%	0.0297	-24%
$\sigma_{f_{ped}}^2$	0.01		0.01		0.01		0.01		0.01	
$\sigma_{f_{mol}}^2$	0.058		0.026		0.019		0.012		0.01	
$r$	0.20		0.49		0.62		0.89		0.99	

$L$ : genome size in morgans. MS: mass selection. MCSped: minimum coancestry selection based on pedigrees.  $R_i$ : response to selection at generation  $t$ .  $\Delta f$ : rates of increase in coancestry between generations 6 and 10 obtained from pedigrees (ped) and molecular (mol) coancestries.  $\sigma_f^2$ : variance of pedigree (ped) and molecular (mol) coancestries.  $r$ : correlation between pedigree and molecular coancestries. Population with  $M = 8$  males,  $F = 16$  females selected and  $n = 4$  offspring per sex and mating. Restriction in selection differential of 2.5%. Initial heritability,  $h^2 = 0.4$ . No common environmental variance ( $c^2 = 0$ ). S.E. of simulations range 0.3-0.7% for selection response. S.E. for rates of coancestry range 2-4% for  $L = 0$ , and 0.2-2% for other  $L$ .

The rate of coancestry obtained with pedigree ( $\Delta f_{\text{ped}}$ ) did not change with the genome length, but this is a correct measure of the coancestry only with free recombination ( $L = \infty$ ). On the contrary, the rate of increase in true coancestry ( $\Delta f_{\text{mol}}$ ) changed substantially with the level of linkage, and the variance of these coancestries ( $\sigma_{f_{\text{mol}}}^2$ ) increased for tighter linkage. Accordingly, the variance of pedigree coancestries between pairs of individuals ( $\sigma_{f_{\text{ped}}}^2$ ) did not vary for different genome sizes, indicating the failure in the estimation. With no recombination at all ( $L = 0$ ), the reduction in  $\Delta f_{\text{mol}}$  with the method of control was 9%, and this was increased for larger values of  $L$ , so that for  $L = 5$  the reduction was similar to that for free recombination. This is also illustrated by the correlation ( $r$ ) between pedigree and molecular coancestries. This correlation took small values for tight linkage, but about 0.9 for  $L = 5$ . Results for  $L = 10$  and 20 were also very similar to those for  $L = \infty$ , and are not shown in any table. Thus, from Table I we can already conclude that the method using pedigree coancestries was of little efficiency in controlling the rate of inbreeding for tight linkage, but with genome sizes of the order of five morgans or so, the method was sufficiently effective.

Table II shows results similar to those from Table I with different sets of parameters, such as different population size, selection intensity, restriction in the selection differential, initial heritability, and common environmental variation. We only show the proportional reductions with MCSped relative to MS.

None of these changes qualitatively affected the above reached conclusions. The larger the population size and selection intensity the larger the reduction in the rate of inbreeding with the method of control, but this effect was soon irrelevant. With smaller heritabilities, the efficiency of the method was slightly higher in terms of reduction of the rate of inbreeding (*cf.* Tab. I), but the loss in response was also larger because of the loss in selection accuracy. The opposite was observed for large heritability ( $h^2 = 1$ ), for which the loss in response was very small, and practicably non-existent by generation 5. In this case the loss in response was only that caused by the 2.5% drop in the selection differential imposed by the method, and this loss of response was compensated after a few generations by the higher preservation of genetic variation with the method for controlling inbreeding.

Because with tight linkage the methods of control of inbreeding are not effective, we simulated a modification of the method where molecular coancestries were used instead of pedigree information. We used all 200 neutral loci to compute true coancestries, but selection decisions were based on coancestries calculated from probabilities of identity in state using one or four of these neutral loci per morgan (using them as genetic markers), with 8 alleles each. In other runs, we used all 200 neutral loci as genetic markers. Table III shows the results from these simulations.

We can observe that the apparent reductions from pedigree information ( $\Delta f_{\text{ped}}$ ) were very different from the real ones ( $\Delta f_{\text{mol}}$ ). These latter were very substantial, particularly with complete linkage. Those obtained by using one marker per morgan (first panel) were already important, and with 4 markers (second panel) the results were similar to those where all the neutral loci were used (third panel). However, the reductions in  $\Delta f_{\text{mol}}$  were coupled to



**Table II.** Proportional reduction in response and in rate of coancestry with the MCSped method relative to mass selection.

	$L = 0$	$L = 0.5$	$L = 1$	$L = 5$	$L = \infty$
	$M = 16$	$F = 32$ $n = 4$	Res = 2.5%	$h^2 = 0.4$	$c^2 = 0$
$R_5$	-4%	-4%	-3%	-4%	-4%
$R_{10}$	-1%	-2%	-3%	-3%	-4%
$\Delta f_{\text{mol}}$	-14%	-14%	-19%	-25%	-24%
	$M = 16$	$F = 32$ $n = 8$	Res = 2.5%	$h^2 = 0.4$	$c^2 = 0$
$R_5$	-5%	-5%	-5%	-4%	-5%
$R_{10}$	-1%	-3%	-3%	-3%	-4%
$\Delta f_{\text{mol}}$	-15%	-14%	-25%	-29%	-29%
	$M = 16$	$F = 32$ $n = 8$	Res = 5%	$h^2 = 0.4$	$c^2 = 0$
$R_5$	-6%	-7%	-8%	-7%	-8%
$R_{10}$	-1%	-5%	-5%	-7%	-7%
$\Delta f_{\text{mol}}$	-17%	-24%	-32%	-40%	-40%
	$M = 8$	$F = 16$ $n = 4$	Res = 2.5%	$h^2 = 0.2$	$c^2 = 0$
$R_5$	-9%	-9%	-9%	-9%	-10%
$R_{10}$	-4%	-7%	-7%	-9%	-9%
$\Delta f_{\text{mol}}$	-22%	-28%	-30%	-34%	-33%
	$M = 8$	$F = 16$ $n = 4$	Res = 2.5%	$h^2 = 0.2$	$c^2 = 0$
$R_5$	-7%	-7%	-6%	-5%	-6%
$R_{10}$	+2%	-3%	-4%	-5%	-5%
$\Delta f_{\text{mol}}$	-20%	-25%	-27%	-32%	-33%
	$M = 8$	$F = 16$ $n = 4$	Res = 2.5%	$h^2 = 1.0$	$c^2 = 0$
$R_5$	*	-2%	+0%	-1%	-1%
$R_{10}$	*	-0%	+1%	-0%	-0%
$\Delta f_{\text{mol}}$	*	-4%	-9%	-18%	-22%

Res: restriction in the selection differential. Definition of parameters as in Table I.

\* Case not given because genetic variance was completely lost in the first few generations.

an additional loss in response in all cases (*cf.* Tab. I). This loss was due to the fact that with molecular information, members of a given family are likely to have different coancestries with other individuals, unlike what happens with pedigree information. Thus, with MCSped, the individuals with the highest phenotype in each family were always chosen, while with MCSmol this was not always the case. The additional loss of response was larger with the smaller the heritability, and did not occur for  $h^2 = 1$  (data not shown). This additional loss in response was also responsible for the larger reduction in  $\Delta f_{\text{mol}}$  observed for MCSmol than for MCSped (Tab. I) for  $L = \infty$ . With a very large number of segregating markers, such a difference would not be expected, because the

**Table III.** Reduction in response and in rate of coancestry with the MCSmol method relative to mass selection (Tab. I).

	$L = 0$	$L = 0.5$	$L = 1$	$L = 5$	$L = \infty$
Markers	1	1	1	5	20
$R_5$	-9%	-5%	-5%	-5%	-4%
$R_{10}$	-5%	-2%	-4%	-4%	-4%
$\Delta f_{\text{ped}}$	-3%	-11%	-12%	-22%	-27%
$\Delta f_{\text{mol}}$	-58%	-30%	-29%	-27%	-25%
Markers	4	4	4	20	100
$R_5$	-9%	-6%	-5%	-6%	-5%
$R_{10}$	-6%	-3%	-3%	-5%	-5%
$\Delta f_{\text{ped}}$	-3%	-8%	-11%	-21%	-28%
$\Delta f_{\text{mol}}$	-60%	-43%	-39%	-36%	-33%
MCSmol (all neutral genes used as markers)					
$R_5$	-8%	-8%	-7%	-7%	-5%
$R_{10}$	-4%	-4%	-4%	-7%	-5%
$\Delta f_{\text{ped}}$	-3%	-12%	-14%	-21%	-28%
$\Delta f_{\text{mol}}$	-57%	-45%	-46%	-41%	-31%

$M = 8$ ,  $F = 16$ ,  $n = 4$ ,  $h^2 = 0.4$ ,  $\text{Res} = 2.5\%$ .

See Tables I and II for definition of parameters.

use of markers would be equivalent to the use of pedigree information when  $L = \infty$ . However, the use of a restricted number of markers generates a random component within families. The molecular coancestries between members of the same family and any other individual may be different whereas the true coancestries are identical. This reduces the accuracy for the estimation of breeding values when  $h^2 < 1$ .

Table IV shows simulations similar to those presented above, but referred to selection on BLUP evaluations. The first panel presents the results of the MCSped method. The conclusions are the same as before. The method for controlling inbreeding was effective unless the genome length was less than 5 morgans. The loss of response was rather small in all cases, as previously observed for phenotypic selection with large heritability (see Tab. II).

The second and third panels show relative results when molecular coancestries were used (MCSmol). Substantial reductions in the rate of inbreeding were obtained for small genome lengths, although some additional losses in response also occurred, for the reasons explained above. For  $L = \infty$ , the losses in response and in rates of coancestry were the same under MCSped and MCSmol, as observed for phenotypic selection with  $h^2 = 1$  (not shown).

We also investigated a different method for controlling inbreeding with BLUP selection, the use of an artificially raised heritability in the BLUP evaluations [8]. Although sub-optimal, this method has the advantage of being very simple and quite effective. While the true heritability was 0.2, the evaluations

Table IV. Response to selection and rate of coancestry under BLUP selection, and relative reductions using the MCSped and MCSmol methods.

	$L = 0$		$L = 0.5$		$L = 1$		$L = 5$		$L = \infty$	
	BLUP	MCSped	BLUP	MCSped	BLUP	MCSped	BLUP	MCSped	BLUP	MCSped
$R_5$	1.262	-3%	1.344	-3%	1.377	-1%	1.413	-0%	1.431	-1%
$R_{10}$	1.503	-1%	2.000	-0%	2.126	+1%	2.398	+1%	2.538	+1%
$\Delta f_{ped}$	0.037	-29%	0.043	-25%	0.061	-23%	0.064	-24%	0.064	-20%
$\Delta f_{mol}$	0.153	-14%	0.091	-17%	0.102	-17%	0.077	-20%	0.063	-20%
$r$	0.21		0.46		0.55		0.83		0.98	
MCSmol										
Markers	4		4		4		20		100	
$R_5$	-8%		-7%		-4%		-2%		-1%	
$R_{10}$	-5%		-4%		-1%		-1%		+1%	
$\Delta f_{ped}$	-2%		-6%		-9%		-15%		-20%	
$\Delta f_{mol}$	-58%		-38%		-31%		-26%		-22%	
All neutral genes used as markers										
$R_5$	-9%		-7%		-4%		-1%		+1%	
$R_{10}$	-4%		-4%		-3%		-0%		+1%	
$\Delta f_{ped}$	-1%		-9%		-14%		-14%		-19%	
$\Delta f_{mol}$	-53%		-43%		-36%		-27%		-21%	

$M = 8$ ,  $F = 16$ ,  $n = 4$ ,  $h^2 = 0.4$ ,  $Res = 2.5\%$ .

See Tables I and II for definition of parameters.

**Table V.** Reduction in response and in rate of coancestry with the selection MCS method and mating CM method using pedigree (ped) or molecular information (mol).

	$L = 0$	$L = 0.5$	$L = 1$	$L = 5$	$L = \infty$
MCSped and CMped					
$R_5$	-6%	-7%	-7%	-8%	-7%
$R_{10}$	-1%	-4%	-4%	-6%	-7%
$\Delta f_{\text{mol}}$	-12%	-26%	-27%	-37%	-37%
MCSmol and CMmol					
Markers	4	4	4	20	100
$R_5$	-10%	-7%	-7%	-7%	-6%
$R_{10}$	-7%	-4%	-5%	-6%	-5%
$\Delta f_{\text{mol}}$	-65%	-51%	-46%	-43%	-38%
MCSmol and CMmol (all neutral genes used as markers)					
$R_5$	-14%	-11%	-11%	-9%	-7%
$R_{10}$	-8%	-6%	-8%	-8%	-6%
$\Delta f_{\text{mol}}$	-67%	-60%	-60%	-51%	-39%

$M = 8$ ,  $F = 16$ ,  $n = 4$ ,  $h^2 = 0.4$ ,  $\text{Res} = 2.5\%$ . See Tables I and II for definition of parameters.

were made using a value of 0.4. The results obtained were similar to those above for MCSped. The reductions in the rate of coancestry were 14, 18, 21, 22 and 28% for  $L = 0, 0.5, 1, 5$  and  $\infty$ , respectively, with losses in response at generation 5 of 4, 1, 0, 3 and 1%.

In all simulations above, mating among selected individuals was random. Table V shows the results of simulations of phenotypic selection in which compensatory mating (CM) was carried out. In the first panel of the table, the method of selection minimising coancestries and compensatory mating were based both on pedigree coancestries (*cf.* Tab. I), while in the second and third panels, both methods were based on molecular information (*cf.* Tab. III). Compensatory mating has two general effects. First, because of its component of negative assortative mating, it slightly reduces selection response [15]. Second, it increases the effective rate of recombination between transmission lines with little and large reproductive success. For  $L = 0$  there was no possibility of increasing the recombination rate, so the reduction in the rate of coancestry was only due to the loss in response. For  $L \geq 0.5$ , however, there was a more substantial reduction in the rate of coancestry, mainly due to the increase of the effective recombination rate.

#### 4. DISCUSSION

In recent years, methods for controlling inbreeding in selection programmes have been sought, both for unselected conserved populations (*e.g.*, [1,2]) and for selected populations (see references above). Rates of inbreeding can be largely

increased in artificially selected populations, particularly in species with small genome sizes [16], and this may have a large impact on the medium and long-term response [13]. Inbreeding reduces genetic variability, the probability of fixation of favourable genes, and the vigour and reproductive capacity of the animals. All methods proposed to control inbreeding in selection programmes have been tested simulating genetic models with unlinked genes. However, under linkage rates of inbreeding calculated from pedigree relationships can be very different from those calculated from the probability of identity in state of neutral markers. We used a method based on the general principle of minimising average coancestries among selected individuals with a restriction in the loss of gain. This principle is followed in some way or another in the majority of the proposed methods to control inbreeding.

The framework investigated was that of a fixed number of parents in each generation. More flexible schemes that also give an optimum number of selected sires and dams [5,9,11] may be somewhat more vulnerable to errors in coancestries, but it is unlikely that they would have substantially changed the results obtained in this paper. This was supported by the high correlations observed between pedigree and molecular coancestries for  $L = 5$  or higher values (see  $r$  in Tabs. I and IV).

The genetic model studied assumed a large number of genes of small effect. This is effectively an infinitesimal model with linkage [14], which would produce the largest effects of linkage on the rate of inbreeding [16]. For this model, the number of genes, the magnitude of effects and the initial gene frequencies are mostly irrelevant, provided the number of genes is large enough. The key factor is the genetic density of a given chromosome region, *i.e.* the amount of genetic variation for the selected trait spread in that region. The genome simulated consisted of a single chromosome,  $L$  morgans long. In terms of the effects of linkage under selection, this is the only parameter of importance so, for example, a chromosome  $L$  morgans long is approximately equivalent to  $L$  chromosomes of one morgan. This is because the effect of linkage is manifested only in neutral genes very close to the selected gene in the chromosome [16].

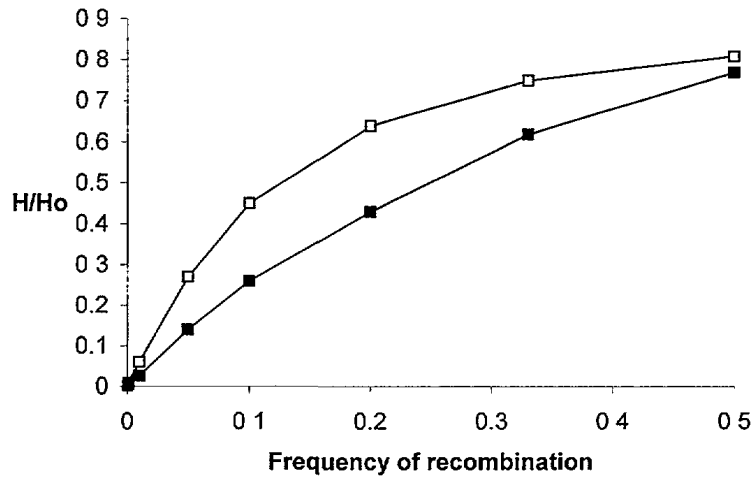
Our results indicate that the use of pedigree information is useful unless the quantitative trait loci for the selected trait are spread in a genome size of the order of five morgans or less. Most domestic animals have genomes with 20 or 30 chromosomes of one morgan or so [12]. If we assume that genes affecting a given quantitative trait subject to selection are spread over genome regions accounting for five morgans or more then we can conclude that linkage does not pose an important problem when the control of inbreeding is made by procedures using pedigree information. However, for small genome lengths or regions of low recombination, methods of control of inbreeding using pedigree data would be inefficient, and molecular markers should be used for a proper control. In conservation programmes without selection, the use of molecular markers can be useful for controlling rates of inbreeding even in species with large genomes, especially when combined with pedigree information [17]. With selection, however, and due to the restrictions imposed by the phenotypic scores, the effectiveness of using molecular information is much more restricted.

In order to check the above aspect, we carried out simulations with the MCSped and MCSmol methods without artificial selection for the case with

$M = 8$ ,  $F = 16$ ,  $n = 4$ , and variable genome sizes. With the MCSped method, the rate of coancestry was 0.0120, 44% lower than the rate with no control of inbreeding (0.0215). This reduction was larger than that obtained with the classical method of Gowe *et al.* [7], *i.e.* 35%, and was similar to that obtained with the modification proposed by Wang [19]. With MCSmol the rates of coancestry were reduced by 99, 49, 36, 42, and 42% using 1, 1, 1, 5 and 20 markers, respectively, for  $L = 0, 0.5, 1, 5$  and  $\infty$ , respectively. With 4, 4, 4, 20 and 100 markers the corresponding reductions were 93, 72, 61, 65 and 68%. For all genes used as markers, the reductions were 99, 94, 90, 78, and 59%. The corresponding reductions with artificial selection and a restriction in the selection differential of 2.5% are given in Table III. Note that, in the absence of selection, the series did not decrease monotonically with  $L$  when the number of markers was variable. For example, there was more efficiency with 5 markers and  $L = 5$  than with one marker and  $L = 1$ . This effect was not observed under selection (see Tabs. III and IV).

We investigated the effects of linkage considering a whole genome where genes with equal effects are distributed uniformly along the whole chromosome. Linkage effects can also occur in populations with large genome sizes referred to chromosome regions with low recombination. The use of marker-assisted selection in order to select genes of large effect associated to markers is another concern in terms of a reduction of genetic variability. As initially shown by Gibson [6], selection on a major gene has the effect of decreasing the long-term response from other genes of minor effect (polygenes) on the genome. This is mostly due to a reduction in the selection intensity applied on the polygenes, but also due to a lesser extent, to an increase in the rate of inbreeding. Recently, successful methods have been proposed for overcoming this difficulty, either based on controlling inbreeding or adjusting the selection pressure on the major gene (see Villanueva *et al.* [18] and references therein). These methods again refer to situations of free recombination. Under tight linkage, the increase of the rate of inbreeding can be largely magnified for the neutral genes (or polygenes) close to the major gene (see Fig. 1).

Because the effective recombination rate increases with the number of generations until fixation, one effective way of avoiding the loss of variability in regions close to a major gene must be to decrease the selection intensity on this one. This allows recombination to be increased between the major gene and the adjacent genes before fixation. We ran some simulations to check this hypothesis. Figure 1 represents the effect of doubling the number of generations to fixation of a major gene on the final heterozygosity of neutral genes at different distances (recombination frequencies) from it. The major gene was initially considered to be in single copy in a population of 30 individuals, and it was fixed in either 5 or 10 generations. This was achieved by increasing its frequency in a regular way until fixation. Although the increase of the frequency in consecutive generations was nearly even in the simulations, a wide range of differences in the rate of increase of frequency had a very small influence on the final heterozygosity (data not shown). The most important parameter was the total number of generations to fixation. If the recombination rate between the major gene and the neutral gene was not large, the effect of increasing the number of generations by a given factor has an effect equivalent to increasing



**Figure 1.** Proportion of the remaining heterozygosity of neutral loci after the fixation of a major gene initially present as a single copy in a population of 30 individuals. The results are shown for different values of the frequency of recombination between the major gene and the neutral loci. Fixation of the major gene was carried out after either 5 generations (filled boxes) or 10 generations (open boxes) of selection.

the recombination rate by the same factor. Two generations of recombination reduce the association between the two genes by  $(1 - r)^2$ , which is nearly equal to  $1 - 2r$ , that is the effect of one single generation with two times the original recombination rate. This effect can be observed in Figure 1 as an equivalent heterozygosity in the two lines for neutral loci with a recombination rate with the major gene differing by a factor of two. Thus, a reduction in selection intensity may have led the major gene to fixation without so much loss of variability in its surroundings. It should be noted, however, that if the major gene is selected *via* a linked marker, a decrease in the selection intensity might also allow for recombination between the major gene and the marker.

The effect on the whole chromosome depends on its genetic size and the location of the major gene. For a chromosome size of one morgan, the effect of doubling the number of generations to fixation from 5 to 10 increased the average chromosome heterozygosity by 29% if the major gene was in the middle of the chromosome. The increase was only 15% if it was near the tip of the chromosome. Comparatively, the effect of selection of a major gene on the variability of neutral genes on other chromosomes was small unless fixation was extremely quick.

#### ACKNOWLEDGEMENTS

We are grateful to J. Fernández, W.G. Hill, J. Wang and two anonymous referees for useful comments on the manuscript. This work was supported by grants PB96-0343 (A.C.) and PB95-0909-C0202 (E.S.) from Ministerio de Educación y Cultura, by grant 64102C856 (B.F. and A.C.) from Universidad de Vigo, and by INIA Project SC98-083 (M.A.T.)

## REFERENCES

- [1] Ballou J.D., Lacy R.C. Identifying genetically important individuals for management of genetic diversity in captive populations, in: Ballou J.D., Gilpin M., Foote T.J. (Eds.), *Population Management for Survival and Recovery*, Columbia University Press, New-York, USA, 1995, pp. 76–111
- [2] Caballero A., Toro M.A., Interrelations between effective population size and other pedigree tools for the management of conserved populations, *Genet. Res.*, (2000) in press.
- [3] Caballero A., Santiago E., Toro M.A., Systems of mating to reduce inbreeding in selected populations, *Anim. Sci.* 62 (1996) 431–442.
- [4] Falconer D.S., Mackay, T.F.C., *An Introduction to Quantitative Genetics*, 4th edn., Longman, Harlow, 1996.
- [5] Fernández J., Toro M.A., The use of mathematical programming to control inbreeding in selection schemes, *J. Anim. Breed. Genet.* 116 (1999) 447–466.
- [6] Gibson J.P., Short term gain at the expense of long term response with selection on identified loci, in: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production*, 7–12 August, University of Guelph, Guelph, Canada, 1994, Vol 21, pp. 201–204.
- [7] Gowe R.S., Robertson A., Latter B.D.H., Environment and poultry breeding problems. 5. The design of poultry control strains, *Poult. Sci.* 38 (1959) 462–471.
- [8] Grundy B., Caballero A., Santiago E., Hill W.G., A note on using biased parameter values and nonrandom mating to reduce rates of inbreeding in selection programmes, *Anim. Prod.* 59 (1994) 465–468.
- [9] Grundy B., Villanueva B., Woolliams J.A., Dynamic selection procedures for constrained inbreeding and their consequences for pedigree development, *Genet. Res.* 72 (1998) 159–168.
- [10] Grundy B., Luo Z.W., Villanueva B., Woolliams J.A., The use of mendelian indices to reduce the rate of inbreeding in selection programmes, *J. Anim. Breed. Genet.* 115 (1998) 39–51.
- [11] Meuwissen T.H.E., Maximizing the response of selection with a predefined rate of inbreeding, *J. Anim. Sci.* 75 (1997) 934–940.
- [12] Nicholas F.W., *Veterinary Genetics*, Oxford University Press, 1987.
- [13] Robertson A., A theory of limits in artificial selection with many linked loci, in: Kojima K. (Ed.), *Mathematical Topics in Population Genetics*, Springer-Verlag, Berlin, 1970, pp. 246–288.
- [14] Santiago E., Linkage and the maintenance of variation by mutation-selection balance in finite populations: an infinitesimal model, *Genet. Res.* 71 (1998) 161–170.
- [15] Santiago E., Caballero A., Effective size of populations under selection, *Genetics* 139 (1995) 1013–1030.
- [16] Santiago E., Caballero A., Effective size and polymorphism of linked neutral loci in populations under selection, *Genetics* 149 (1998) 2105–2117
- [17] Toro M.A., Silió L., Rodríguez J., Rodríguez C., Fernández J., Optimal use of genetic markers in conservation programmes, *Genet. Sel. Evol.* 31 (1999) 255–261.
- [18] Villanueva B., Pong-Wong R., Grundy B., Woolliams J.A., Potential benefit from using an identified major gene in BLUP evaluation with truncation and optimal selection, *Genet. Sel. Evol.* 31 (1999) 115–133.



- [19] Wang J., More efficient breeding systems for controlling inbreeding and effective size in animal populations, *Heredity* 79 (1997) 591–599.
- [20] Woolliams J.A., Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle, *Anim. Prod.* 49 (1989) 1–14.