



## A dynamic deterministic model to evaluate breeding strategies under mixed inheritance

Eduardo Manfredi, Maria Barbieri, Florence Fournet, Jean Michel Elsen

### ► To cite this version:

Eduardo Manfredi, Maria Barbieri, Florence Fournet, Jean Michel Elsen. A dynamic deterministic model to evaluate breeding strategies under mixed inheritance. *Genetics Selection Evolution*, 1998, 30 (2), pp.127-148. hal-00894200

**HAL Id: hal-00894200**

**<https://hal.science/hal-00894200>**

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.

## A dynamic deterministic model to evaluate breeding strategies under mixed inheritance

Eduardo Manfredi\*, Maria Barbieri, Florence Fournet,  
Jean Michel Elsen

Station d'amélioration génétique des animaux,  
Institut national de la recherche agronomique,  
Centre de Toulouse, BP. 27,  
31326 Castanet-Tolosan cedex, France

(Received 9 June 1997; accepted 10 March 1998)

**Abstract** – A dynamic deterministic model is proposed to study the combined use of an identified major gene and performance information for selection of traits expressed in one sex. The model considers simultaneously combined adult selection via within genotype thresholds, mating structures according to major genotypes and preselection of young males. The application given indicates that an optimum combination between performances and genotypic information yields better results, in terms of polygenic means, genotype frequencies and cumulated discounted genetic progress, than classical selection ignoring the genotype information. The greatest advantage of combined selection occurs for rare recessive alleles of large effect on phenotypes (up to +49 % for polygenic gains; +26 % for total genetic gain). Optimum within genotype proportions of selected individuals and mating structures vary with generations thus highlighting the value of a dynamic approach.  
© Inra/Elsevier, Paris

**dynamic selection / major gene / genetic marker / model**

**Résumé** – Un modèle déterministe et dynamique pour comparer des stratégies de sélection en hérédité mixte. Un modèle déterministe et dynamique est proposé pour étudier l'utilisation conjointe des performances et des génotypes à un locus majeur pour la sélection des caractères exprimés dans un sexe. Le modèle prévoit la sélection des adultes au-delà de seuils de performances intra-génotype, des accouplements en fonction des génotypes et la présélection de jeunes mâles sur leur génotype. L'application présentée indique que la combinaison optimale des performances et des génotypes permet d'obtenir des meilleurs résultats, en terme de moyennes polygéniques, de fréquences génotypiques et du progrès génétique actualisé et cumulé, que la sélection classique ignorant l'information

---

\* Correspondence and reprints  
E-mail: manfredi@toulouse.inra.fr

génomique. Les avantages de la sélection combinée sont plus importants quand l'allèle favorable est rare et récessif, les différences avec la sélection classique pouvant atteindre +49 % en gains polygéniques et +26 % en gain génétique total. Les taux optimaux de sélection intra-génotype et les structures d'accouplement optimales varient au cours des générations, confirmant l'intérêt de l'approche dynamique. © Inra/Elsevier, Paris

**sélection dynamique / gène majeur / marqueur génétique / modélisation**

## 1. INTRODUCTION

For many years, selection and matings among animals have been based on classical genetic evaluations where performances are adjusted under a polygenic model. The rapid evolution of molecular genetics allows genotyping at known major loci at a reasonable cost for males and females at any age. However, the advantages of adding genotypic information at the major locus in order to improve the gains obtained by classical selection may vary widely according to the time horizon, the genetic determinism of the trait (relative importance of the major gene and the polygenic effects, allele frequencies at the major locus, additive and dominance effects at the major locus), the age and sex where trait expression occurs, the type of selection practised (mass or family selection), and the strategy combining the performances and the genotypic information at the major locus.

The problem has been recursively addressed in the literature through stochastic or deterministic simulations based on genetic models including a polygenic background plus marked QTL or known major gene effects. Precise comparison of results is difficult because genetic models, simulated selection methods, methods for prediction of genetic gains, criteria for comparing selection schemes and situations studied vary widely.

Several studies have reported disadvantages or modest gains when combining genotype or marker information with performances in indexes for single-threshold adult selection [5, 15]: in the short term, classical selection yielded lower responses than combined selection using performance and major genotype information because combined selection resulted in a rapid fixation of favourable alleles at the major locus; however, classical selection performed better than combined selection in the long term since selection intensity applied to the polygenic background was reduced by combined selection. Advantages of combined selection have been reported for situations such as multiple trait objectives [2], especially when traits are negatively correlated [14], or when favourable alleles are recessive [9].

The use of genotype or marker information in multi-stage selection appears to be more profitable than combining genotype and performance information for adult selection [8], especially when traits are expressed in only one sex.

Most of these literature results are obtained by fixing, a priori, rules to combine genotype and performance information. Here, we propose a procedure for mixed inheritance (one major gene plus polygenes) aiming to find optimum dynamic rules through a deterministic simulation model for infinite size populations without overlapping generations. The model allows simultaneous consideration of adult combined selection through multiple within genotype thresholds, genotypic preselection of animals and mating structures according to major genotypes.

## 2. THE SELECTED POPULATION

We concentrate on the case of selection of traits expressed only in females. Females are selected on own performances and males are selected at two stages: genealogical selection through planned matings and progeny test selection. The model for phenotypes is:

$$P_{ij} = m_i + a_{ij} + e_{ij}$$

where  $P_{ij}$  is the phenotype,  $m_i$  is the fixed effect of the  $i$ th genotype at the major locus,  $a_{ij}$  additive polygenic value of the  $j$ th individual bearing the  $i$ th genotype,  $a_{ij} \sim N(\mu_i, \sigma^2)$  and  $e_{ij}$  is the random residual,  $e_{ij} \sim N(0, \sigma_e^2)$ , such that:

$$h^2 = \sigma^2 / (\sigma^2 + \sigma_e^2)$$

The population where a major gene segregates is divided into five classes of animals, each one subdivided according to genotypes at the major locus:

males born: 'M'

males in progeny testing ('males in test'): 'Y'

males selected after progeny testing ('tested males'): 'S'

unselected females: 'F'

females selected as dams of males ('dams of males'): 'D'

Accordingly, five transmission paths are defined: dam to son, tested male to son, female to daughter, tested male to daughter and males in test to daughter (*figure 1*).

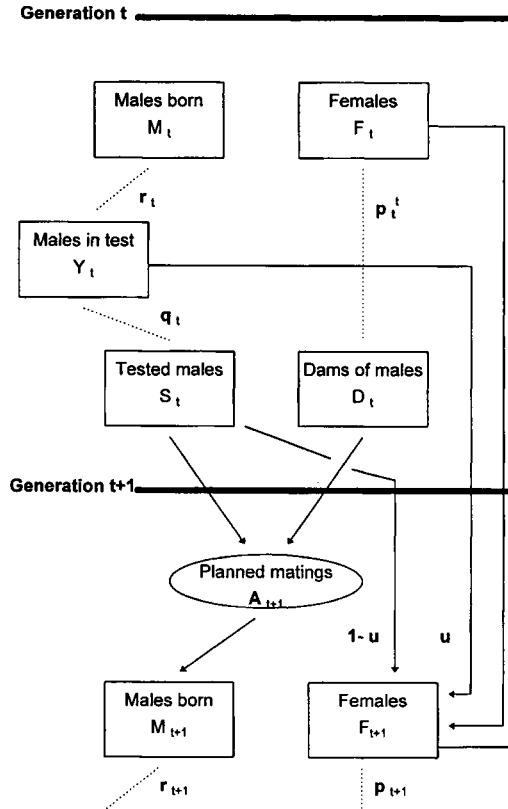
The model allows for two types of selection:

1) combined genotypic and polygenic selection of dams of males and tested males. Here, combined selection implies the use of an index including a fixed genotypic effect and a random polygenic effect, but also consideration of different proportions of individuals selected within major genotypes. These proportions are the ratios 'parents kept after selection/candidates for selection' defined separately for each major genotype. This implies that the classical single threshold selection is replaced by multiple thresholds, one threshold per major genotype, in the proposed model. The within genotype proportions selected may change at each generation and they are represented by the vectors  $\mathbf{q}_t$  (males) and  $\mathbf{p}_t$  (females) in *figure 1*; the order of the vectors equals the number of genotypes and  $t$  indicates the generation number. These vectors are variables whose values are obtained via maximization of an objective function defined below.

2) genotypic selection, before progeny testing, of males born. The proportions selected, i.e. the ratios 'males kept for progeny testing/males born', are defined for each major genotype and for each generation  $t$ . In *figure 1*, they are represented by the vectors  $\mathbf{r}_t$  of order equal to the number of genotypes. This step is an across family genotypic selection.

Selection of dams of daughters is not considered.

Also, the model allows for consideration of proportions of males born from matings between dams of males and tested sires according to the major genotypes of these parents. These proportions are defined for each generation  $t$  and represented by the  $\mathbf{A}_t$  matrices of order 'number of maternal genotypes  $\times$  number of paternal genotypes' in *figure 1*. The elements of the  $\mathbf{A}_t$  matrices are variables whose elements



**Figure 1.** The selected population. For each generation five classes of animals (in the rectangles), five genetic transmission paths (arrows), planned matings (matrix  $A$ ), three selection steps (discontinuous traits) quantified by the vectors  $r$ ,  $p$  and  $q$ , and the rate of use of males in test  $u$ .

are found by optimization, subject to constraints, of an objective function defined below.

The approach is dynamic since, for a given user-defined objective function, for instance the cumulated polygenic gains or the cumulated global genetic gains in a given animal class, the model locates the optimum within genotype selected proportions and the optimum mating structures at each generation of a user-defined time horizon.

### 3. MATHEMATICAL MODEL

The variables and parameters of the model are described in *table I*. Model equations are listed in *table II*. These equations, in scalar notation, represent the selection process modelled.

1) Selection of dams of males by combining genotypic and performance information. In equation (1.1) the optimum proportions  $p_{gt}$  of females selected within genotype  $g$  at generation  $t$  are used to compute within genotype selection thresholds and their corresponding selection differentials. A constant correlation between true and estimated polygenic value ( $\rho_F$ ) is applied to female selection for all genotypes and all generations. In equation (1.2) the genotype frequencies of dams of males are functions of the proportion of females selected within genotypes  $p_{gt}$  and the genotype frequencies of females  $f_{Fgt}$ . Equation (1.3) sets a necessary constraint tying the overall proportion of females selected  $P$  to the within genotype proportions selected. Equation (1.4) sets bounds for the solutions of optimum proportions selected.

**Table I.** Variables and parameters of the model.

Symbol	Description
Subscripts	
$t(t = 1 \text{ to } T)$	generation number
$i(M, Y, S, F, D)$	animal class
$g(AA, AB, BB)$	genotypes
Variables characterizing the state of the population at generation $t$	
$\mu_{igt}$	polygenic mean of animal class $i$ and genotype $g$
$f_{igt}$	frequency of genotype $g$ in animal class $i$
Variables describing the selection process, to be optimized	
$p_{gt}$	proportion of females selected as dams of males, within genotype $g$ at generation $t$
$r_{gt}$	proportion of males born kept for progeny testing, within genotype $g$ at generation $t$
$q_{gt}$	proportion of males in test selected after progeny testing, within genotype $g$ at generation $t$
$\alpha_{hkt}$	proportion of males born from matings between tested sires of genotype $h$ and dams of males of genotype $k$ at generation $t$
Parameters (demographic, genetic, economic)	
$P, Q, R$	total (across genotypes) selection rates of females, males born and tested males, respectively
$u$	proportion of daughters of males in progeny testing
$\tau_{ghk}$	probability of having a son of genotype $g$ given the genotype $h$ of the tested sire and the genotype $k$ of the dam
$\rho_i (i = F \text{ or } Y)$	within genotype correlation between true and estimated polygenic values
$\sigma$	polygenic standard deviation
$i[\cdot]$	selection intensity (as a function of proportions of selected individuals)
$m_g$	effect on performances of genotype $g$
$d$	discounting coefficient

**Table II.** Model equations.

Selection of dams of males			
$\mu_{Dgt} = \mu_{Fgt} + i[p_{gt}]\rho_F\sigma$			(1.1)
$f_{Dgt} = p_{gt}f_{Fgt}/P$			(1.2)
$\sum_g p_{gt}f_{Fgt} = P$	(1.3)	$p_{gt} \in [0, 1]$	(1.4)
Selection of progeny tested males (S)			
$\mu_{Sgt} = \mu_{Ygt} + i[q_{gt}]\rho_Y\sigma$			(2.1)
$f_{Sgt} = q_{gt}f_{Ygt}/Q$			(2.2)
$\sum_g q_{gt}f_{Ygt} = Q$	(2.3)	$q_{gt} \in [0, 1]$	(2.4)
Production of males (M)			
$\mu_{Mgt} = \frac{1}{2} \left\{ \sum_h \sum_k \tau_{ghk} \alpha_{hkt} [\mu_{Sh(t-1)} + \mu_{Dk(t-1)}] \right\} / f_{Mgt}$			(3.1)
$f_{Mgt} = \sum_h \sum_k \tau_{ghk} \alpha_{hkt}$			(3.2)
$\sum_h \alpha_{hkt} = f_{Dk(t-1)}$	(3.3)	$\sum_k \alpha_{hkt} = f_{Sh(t-1)}$	(3.4)
		$\alpha_{hkt} \in [0, 1]$	(3.5)
Production of females			
$\mu_{Fgt} = \frac{1}{2} \sum_h \sum_k \tau_{ghk} \{ u f_{Yh(t-1)} f_{Fh(t-1)} [\mu_{Yh(t-1)} + \mu_{Fk(t-1)}] + (1-u) f_{Fk(t-1)} f_{Sh(t-1)} [\mu_{Sh(t-1)} + \mu_{Fk(t-1)}] \} / f_{Fgt}$			(4.1)
$f_{Fgt} = \sum_h \sum_k \tau_{ghk} \{ u f_{Yh(t-1)} + (1-u) f_{Sh(t-1)} \} f_{Fk(t-1)}$			(4.2)
Genotypic selection of males before progeny testing			
$\mu_{Ygt} = \mu_{Mgt}$			(5.1)
$f_{Ygt} = r_{gt}f_{Mgt}/R$			(5.2)
$\sum_g r_{gt}f_{Mgt} = R$	(5.3)	$r_{gt} \in [0, 1]$	(5.4)

2) Selection of tested males. Equations are analogous to female selection equations. In equation (2.1) within genotype directional selection on an index is considered, as in female selection.

3) Production of young males. The model allows planned matings between dams of males and tested males according to their genotypes at the major locus. The plan is automatically given by the optimum solutions of  $\alpha_{hk}^t$  (elements of the  $\mathbf{A}^t$  matrices of figure 1) corresponding to the optimum proportions of males born at generation  $t$  from parents of genotypes  $h$  and  $k$ . Thus, in equation (3.1), the polygenic means  $\mu_{Mgt}$  of males

of genotype  $g$  born at generation  $t$  are functions of the parental polygenic means weighted by the proportion of males born  $\alpha_{hkt}$  and the probability  $\tau_{ghk}$  of obtaining a son of genotype  $g$  from matings between a paternal genotype  $h$  and a maternal genotype  $k$ . Equations (3.3) and (3.4) are necessary constraints tying the proportions of males born to the parental genotypes (e.g. equation (3.3) states that the sum, across paternal genotypes, of the proportions of sons of dams of genotype  $k$  must be equal to the genotype frequency of dams of genotype  $k$ ).

4) Production of females. Female replacements are always produced by random mating. Thus, polygenic means and genotype frequencies in equations (4.1) and (4.2) are functions of the genotype frequencies and the polygenic means of males in test, tested males and females of the previous generation.

5) Genotypic selection of males before progeny testing. Equation (5.1), where the within genotype polygenic means of males born and males put in test are identical, means that the males born to be tested are chosen solely according to their major genotype. The proportions selected within genotypes  $\tau_{gt}$  are obtained by optimization.

As described in the Introduction, literature results indicate that combined selection, when compared to classical selection, leads to a rapid fixation of a favourable allele at the major locus but it may penalize selection intensity on the polygenic background. The proposed model is designed to verify if this assertion is general or if it is only valid for the combined selection rules defined a priori in previous studies, and to find general trends for selection and mating rules when combined selection is used during a given number of generations. By defining as decision variables all selection (polygenic for adults; genotypic for young males) and mating decisions and an objective function including total genetic gains (major genes + polygenes), the model finds a compromise between rapid gains at the major locus and selection intensity applied to the polygenic background. Note also that selection decisions are not conditioned a priori by mating decisions: constraints (3.3) and (3.4) concerning matings allow for parents of all major genotypes and all possible matings among them. Other constraints could be useful to accelerate fixation rates at the major locus but they would add a priori rules to the model.

#### 4. OPTIMIZATION

The objective function chosen here was the cumulated discounted genetic gain of the female class:

$$\sum_{t=0}^T \left( \frac{1}{1+d} \right)^t \left\{ \sum_g f_{Fgt} [m_g + \mu_{Fgt}] \right\}$$

where the ratio  $\frac{1}{1+d}$  is raised to the power  $t$  thus giving a relatively high weight to gains obtained in the short term.

Note that the model equations in *table II* are general enough to allow the definition of other objective functions.

The selection process was optimized by maximizing the objective function subject to linear and nonlinear constraints. For each generation, variables were not only the decision variables, i.e. the proportions of selected individuals  $p_{gt}$ ,  $r_{gt}$  and  $q_{gt}$  and the proportions of males born  $\alpha_{hkt}$ , but also the genotype frequencies  $f_{igt}$  for the five defined classes of animals. As a consequence, bounds were defined by expressions (1.4), (2.4), (3.5) and (5.4), expressions (3.2), (3.3) and (3.4) were linear constraints, and expressions (1.2), (1.3), (2.2), (2.3), (4.2), (5.2) and (5.3) were nonlinear constraints.

This optimization approach was oriented towards programming simplicity: the frequencies could have been computed from the decision variables but they were considered as



variables in order to avoid complex algebraic expressions. Alternatively, the use of recursive formulae for representing the genotype frequencies of all classes of animals as functions of the starting genotype frequencies, the proportions of selected individuals and the mating structures would diminish the number of variables to solve while increasing the computation time of the objective function and complicating the setting of the constraints.

The subroutine E04UCF of the NAG library (Numerical Algorithms Group Ltd.) was used to find the optimum solutions. The subroutine uses a sequential quadratic programming approach. Personal programming was limited to providing the objective function computation, the bounds for variables and the linear and nonlinear constraints and some of their derivatives. Gradients were estimated by finite differences by the NAG routine.

## 5. THE REFERENCE MODEL

The results of the optimization were compared to a 'classical' selection scheme where genotypes are ignored at all selection stages. While keeping the basic structure of five classes of animals and the transmission paths among them, single threshold selection was modelled at each generation for dams of males and tested males and matings among them were at random. Proportions of selected individuals were obtained by solving, at each generation:

$$P = \sum_g f_{Fgt} \left\{ 1 - \Phi \left( \frac{K_F - m_g - \mu_{Fgt}}{\rho_F \sigma} \right) \right\} \quad \text{for female selection} \quad (6.1)$$

$$Q = \sum_g f_{Ygt} \left\{ 1 - \Phi \left( \frac{K_Y - m_g - \mu_{Ygt}}{\rho_Y \sigma} \right) \right\} \quad \text{for male selection} \quad (6.2)$$

where  $\Phi$  represents the normal cumulative distribution function, integrating the normal density function between  $-\infty$  and the selection threshold, and  $K_{Ft}$  and  $K_{Yt}$  represent the female and the male thresholds computed at each generation.

As before, the objective function, the genetic gain and the polygenic gain were computed for this 'classical' strategy.

## 6. APPLICATION

Three main cases were simulated according to the interaction between alleles: recessive, dominant and additive.

For each case, four situations were simulated for a major locus with two alleles (A, favourable, and B) by combining a high ( $P(A) = 0.8$ ) or low ( $P(A) = 0.2$ ) frequency of the favourable allele and a large or small effect of the major genotype on performances. For the additive case, large and small genotype effects were  $[4 \ 2 \ 0]$  and  $[1 \ 0.5 \ 0]$  times the polygenic standard deviation for the genotypes  $[AA \ AB \ BB]$ , respectively. Corresponding values for the recessive case were  $[4 \ 0 \ 0]$  and  $[1 \ 0 \ 0]$  and, for the dominant case  $[4 \ 4 \ 0]$  and  $[1 \ 1 \ 0]$ . For each situation, the three selection strategies compared were 'classical' selection, optimized selection without genotypic preselection of males born ('optimal 1') and optimized selection including a preselection of males born based on their genotypes ('optimal 2'). The time horizon was fixed at six generations of selection.

For the 36 parameter combinations examined, results included the objective function, the polygenic gain and the total genetic gain (polygenic + genotypic) as well as the

polygenic means and the genotype frequencies of the five animal classes at each generation, the within genotype selection proportions of 'tested males', 'males in test' and 'dams of males' at each generation and the mating structure among tested males and dams of males at each generation.

Constants common to the 36 runs were taken from a dairy goat scheme studied by Barbieri [1]:

polygenic standard deviation  $\sigma = 1$ ;

within genotype correlation between true and estimated breeding values of dams of males ( $\rho_F = 0.7$ ) and tested males ( $\rho_Y = 0.9$ ) corresponding to an intermediate heritability (polygenic) of 0.30. These correlations imply the use of individual and ancestors' performances for female indexes and ancestors' and progeny performances for male indexes.

The total (across major genotypes) proportion of tested males selected ( $Q$ ) was 0.30 and the proportion of daughters sired by males in test ( $u$ ) was 0.30.

For the classical and the optimal 1 strategies, there was no selection of males born ( $R = 1.0$ ) and  $P$ , the total proportion of selected females, was 0.10. In the optimal 2 strategy, 30 % of males born were eliminated at birth by genotypic selection ( $R = 0.7$  and, accordingly, the proportion of selected females was increased to 0.10/0.7 ( $P = 0.14$ ). Thus, in optimal 2 the same number of males enter progeny testing as in the optimal 1 and classical strategies. The proportion of selected females took into account culling for conformation and other complementary traits. The discount rate per generation ( $d$ ) was 0.10, with a generation interval of 4 years. Six generations of selection were simulated.

Barbieri [1] showed that the model is extremely sensitive to initial genotype frequencies and major gene effects but less sensitive to the discount rate. Relatively small changes in total proportions selected ( $P = 0.05$  or  $P = 0.10$ ) and time horizons (from 6 to 8 generations) did not alter the observed general behaviour of optimized solutions.

## 7. RESULTS

The additive case is presented first, with a detailed description on the evolution of genetic means, frequencies and mating structures along generations. An overview is given for the recessive (*table VI*) and the dominant (*table VII*) cases.

### 7.1. Additive case – gains

In *table III*, the optimized strategies, optimal 1 and 2, were always better than classical selection but differences were negligible when the initial frequency of the favourable allele was high. For low initial frequencies and small genotype effects, 'optimal 2' outperformed classical selection by 5 % in terms of cumulated discounted gains and by 6 % in terms of genetic gain. This superiority of the optimal 2 scheme over classical selection was due to a more rapid fixation of the favourable allele A in the female population ( $p(A) = 0.82$  in the optimized scheme at generation 6 versus  $p(A) = 0.62$  in classical selection), without losses in polygenic gains. The optimized strategies were more useful when the favourable allele is rare and has a large effect on the phenotype: both optimized schemes outperformed the classical one in terms of cumulated discounted gains, genetic gain and polygenic gain. Note that 'optimal 2', the scheme which has an additional stage of selection and has a higher initial proportion of females selected ( $P = 0.14$ ), had an advantage of 21 % in polygenic gains over the classical scheme while keeping a faster rate of fixation of the favourable allele A.

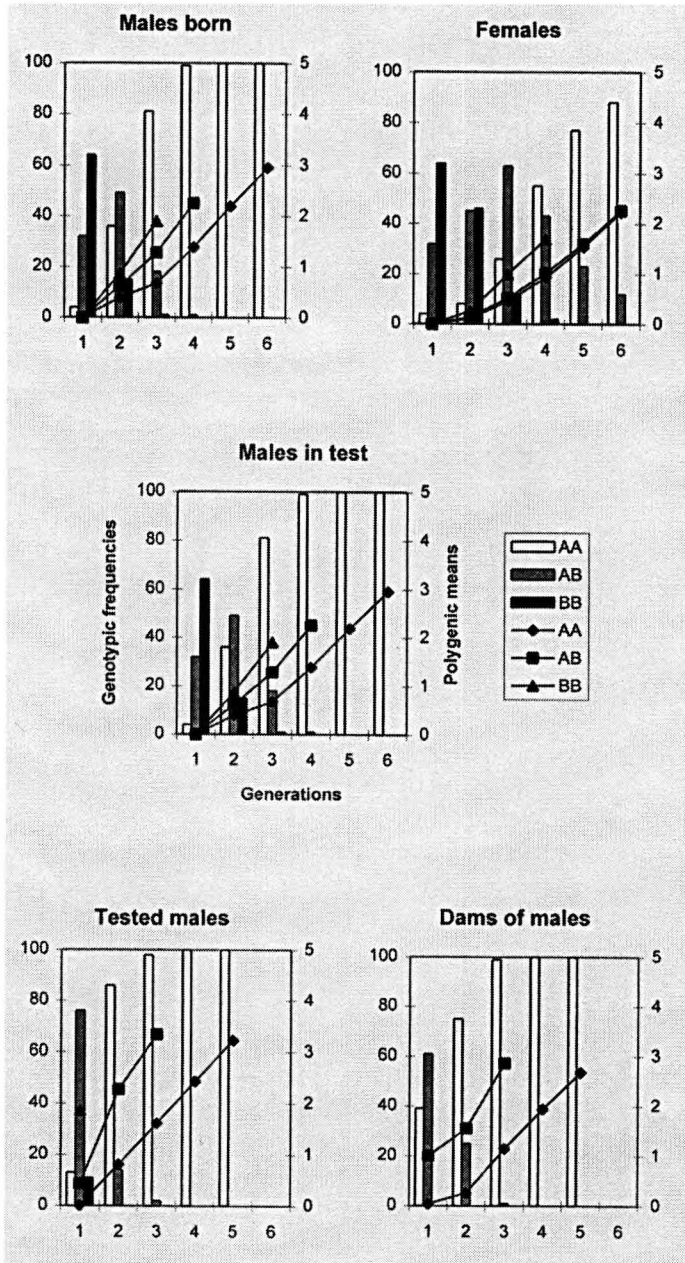
The evolution of the polygenic means and genotype frequencies for all animal classes are presented in *figures 2* (classical), *3* (optimal 1) and *4* (optimal 2). For the female

**Table III.** Additive case. Objective function, polygenic gain and genetic gain for four situations and three selection strategies.

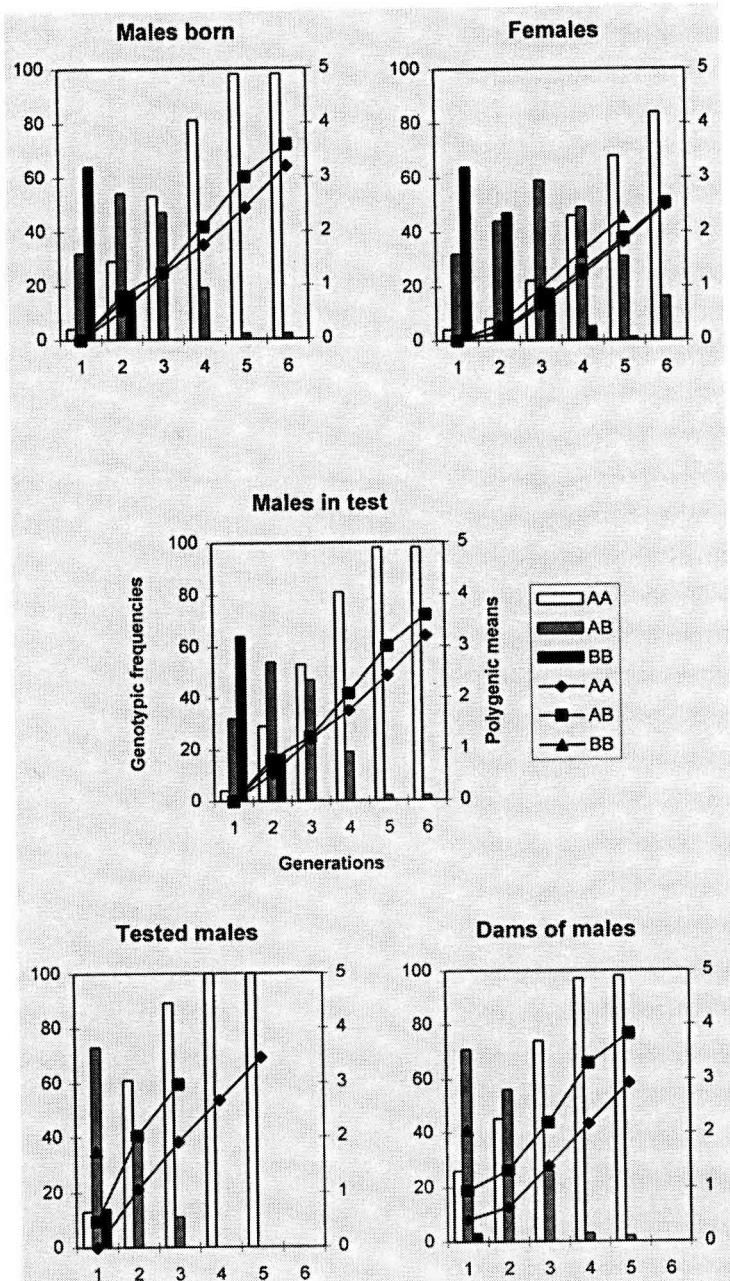
Strategy <sup>a</sup>	Final genotype frequencies for females (%)			Objective function <sup>b</sup>	Polygenic gain <sup>b</sup>	Genetic gain <sup>b</sup>
	AA	AB	BB			
High initial frequency of the favourable allele; small genotype effects <sup>c</sup>						
Classical	93	7	0	13.4	3.29	3.45
Optimal 1	90	10	0	100	101	100
Optimal 2	98	2	0	101	100	100
High initial frequency of the favourable allele; large genotype effects <sup>c</sup>						
Classical	98	2	0	27.9	3.17	3.94
Optimal 1	97	3	0	100	101	100
Optimal 2	98	2	0	102	103	102
Low initial frequency of the favourable allele; small genotype effects <sup>c</sup>						
Classical	37	49	14	10.4	3.12	3.54
Optimal 1	36	50	14	100	101	101
Optimal 2	65	33	2	105	100	106
Low initial frequency of the favourable allele; large genotype effects <sup>c</sup>						
Classical	88	12	0	18.6	2.24	5.20
Optimal 1	84	16	0	101	112	103
Optimal 2	89	11	0	108	121	109

<sup>a</sup> Optimal 1: without male genotypic preselection; optimal 2: with male genotypic preselection. <sup>b</sup> Values for classical selection expressed as genetic standard deviation units. Other values are expressed as percentage of classical selection results. <sup>c</sup> Low and high initial frequencies of the favourable allele are, respectively, 0.2 and 0.8. Small and large genotype effects are, respectively, [ 1 0.5 0 ] and [ 4 2 0 ] genetic standard deviation units.

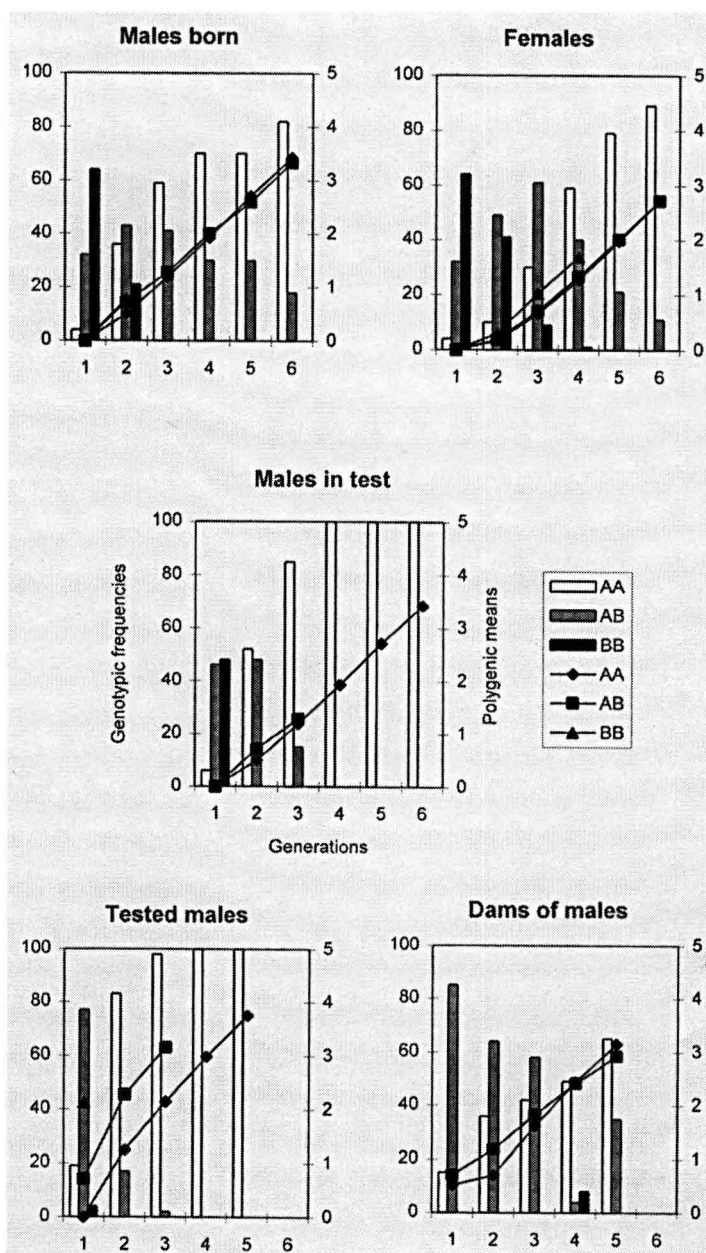
class, optimal 2 performed better than both classical and optimal 1 in terms of rate of fixation of the A allele and polygenic mean of the AA genotypes (at generation 6, polygenic means were 2.23, 2.49 and 2.71  $\sigma$  for AA females under the classical, optimal 1 and optimal 2 schemes, respectively; corresponding values for the frequencies of AA females were 0.88, 0.84 and 0.89). The superiority of optimal 2 in female characteristics reflects a better efficiency in sire selection: fixation of the A allele in the males in test class occurred at the 4th generation in optimal 2 versus the 5th generation for classical and no fixation for optimal 1. For tested males, fixation occurred at generation 4 for the three schemes compared. However, polygenic means of tested males at generation 5 were 3.23, 3.46 and 3.75  $\sigma$  for classical, optimal 1 and optimal 2, respectively. Female selection showed a different behaviour: the A allele was fixed very rapidly in the classical scheme (at generation 4) and it was not fixed in the optimized schemes at the



**Figure 2.** Classical strategy for the additive case when the favourable allele is rare and has a large effect on performance. Polygenic means (plots) and genotype frequencies (histograms) for five classes of animals.



**Figure 3.** Optimal 1 strategy for the additive case when the favourable allele is rare and has a large effect on performance. Polygenic means (plots) and genotype frequencies (histograms) for five classes of animals.



**Figure 4.** Optimal 2 strategy for the additive case when the favourable allele is rare and has a large effect on performance. Polygenic means (plots) and genotype frequencies (histograms) for five classes of animals.

horizon of six generations, especially for optimal 2 (see next section for explanation). This was compensated by higher polygenic means in the optimized schemes: 2.68, 2.93 and 3.15 s in generation 5 for the classical, optimal 1 and optimal 2 schemes. Recall that optimal 2 started with an increased proportion of females selected.

The optimized schemes balanced short and long term gains. Optimal 2 outperformed classical in terms of total genetic gains for the female class at each generation: percentage superiorities were +14, +9, +10, +9 and +9 % for generations two to six, respectively. For all generations, polygenic and major gene contributions to the total genetic means in optimal 2 were higher than those in classical. Corresponding values for optimal 1 were 0, -2, +1, +2 and +3 %, respectively. Optimal 1 had higher polygenic means than classical for all generations but lower gains at the major gene in the first two generations since selection pressure on the major gene in optimal 1 was more evenly spread over generations than in classical.

## 7.2. Additive case – proportions of selected individuals

*Table IV* compares the proportions of dams and sires selected under the three selection strategies. In the classical scheme, the proportions selected within the favourable genotype were very high (for the first generation, 100 % of AA tested males were kept and 98 % of AA females were kept as dams of males) thus leading to a rapid fixation of the favourable genotype (*figure 2*). The behaviour of this single threshold selection is not surprising in this situation where the A allele has a very large favourable effect and it is rare. However, as discussed before, the strategy led to losses in polygenic gains.

Polygenic selection within AA and AB genotypes in optimal 1 was more intense than corresponding polygenic selection in classical, especially in selection of dams of males where 65 and 55 % of AA females were kept as dams of males in the 1st and 2nd generation, respectively, while corresponding figures for the classical scheme were 98 and 89 %. Inclusion of an early genotypic selection of males (optimal 2) allowed even higher selection intensities within favourable genotypes than optimal 1 (e.g. in generation 1, 54 % of AA females were kept as dams of males in optimal 2 versus 65 % kept in optimal 1). The early selection of males allowed higher selection intensities within the favourable genotypes while completing genotypic selection with an additional selection step. In the male side, polygenic selection of favourable genotypes was more intense in optimal 2 than in optimal 1 but the A allele was fixed more rapidly as already shown in *figures 3* and *4*. Concerning the female side, optimal 2 yielded a lower frequency of AA dams of males than optimal 1 but higher frequencies for the favourable genotype of the female class (*table III*; *figures 3* and *4*). The surprising low frequencies for the AA dams of males produced by optimal 2 are due to our choice, for simplicity, of keeping constant across generations the global selection rate  $P$  ( $P = 0.14$  in optimal 2). In fact, what is important in order to maximize the objective function is to have, as soon as possible, at least 70 % of males born having the AA genotype which are translated into 100 % of AA males entering the progeny test. This is achieved when passing from generations 3 to 4 and repeated until the 6th generation (*figure 4*). In these situations, it is useless to keep 14 % of AA females as dams of males and accordingly, the optimum solution proposes a better selection of the AA females (12 % AA females were kept in generations 4 and 5) while completing the 14 % forced by the constant  $P$  with females of other genotypes (e.g. in generation 4, 1 % of the females were BB and they were all selected as dams of males thus representing 9 % of this class. They were all mated to AA sires, all their sons AB were culled before the progeny testing and they did not contribute to the value of the objective function).

**Table IV.** Proportions of selected parents of males (%) under three breeding strategies. Additive case for a low initial frequency of a favourable allele with large effect on performance.

	Tested sires			Dams of males		
	AA	AB	BB	AA	AB	BB
Generation 1						
Classical	100	72	5	98	19	0
Optimal 1	100	68	7	65	22	0
Optimal 2	100	51	2	54	38	0
Generation 2						
Classical	72	8	0	89	6	0
Optimal 1	63	22	0	55	13	0
Optimal 2	48	11	0	52	19	0
Generation 3						
Classical	36	3	4	39	1	0
Optimal 1	51	7	0	33	4	0
Optimal 2	35	5	4	20	13	0
Generation 4						
Classical	30	0	0	18	0	0
Optimal 1	37	0	0	21	1	0
Optimal 2	30	0	0	12	15	100
Generation 5						
Classical	30	0	0	13	0	0
Optimal 1	30	0	0	14	1	0
Optimal 2	30	0	0	12	24	0

### 7.3. Additive case – mating patterns

Table V illustrates the optimized mating patterns for the optimal 1 and optimal 2 strategies. The mating patterns affected the objective function before fixation of the A allele in the female or male sides, i.e. during three or four generations according to the situation studied. In some cases, parents of the favourable genotype were mated to parents of the unfavourable genotype (e.g. in table V, for generation 2 of optimal 1, 56 % of the dams of males were AB and all of them were mated to AA tested males; 39 % of the tested males were AB and all of them were mated to AA dams of males). This 'complementary' or 'heterogametic' pattern was found (as a solution maximizing the objective function) in generation 2 (table V) and later generations of optimal 1 and optimal 2. It was also found for situations where the starting frequency of the favourable allele was high ( $P(A) = 0.8$ ) for large or small genotype effects on performance.



Matings among parents of favourable genotype or 'homogametic matings' were found for the first generation of optimal 2 (in *table V*, 15 % of dams of males were AA and they were all mated to AA tested males; 4 % of tested sires were BB and all of them were mated to AB dams, the 'worst' female genotype available). In some cases, e.g. generation 1 for optimal 1, the mating pattern cannot be strictly described as homogametic or heterogametic.

**Table V.** Parental genotype frequencies (%; bold type) and proportions of males born from planned matings under two breeding strategies. Additive case for a low initial frequency of a favourable allele with large effect on performance.

Optimal 1				Optimal 2			
Paternal genotype	Maternal genotype			Paternal genotype	Maternal genotype		
Generation 1							
	AA	AB	BB		AA	AB	BB
	26	71	3		15	85	0
AA 14	10	1	3	AA 19	15	4	0
AB 72	2	70	0	AB 77	0	77	0
BB 14	14	0	0	BB 4	0	4	0
Generation 2							
	AA	AB	BB		AA	AB	BB
	44	56	0		36	64	0
AA 61	5	56	0	AA 82	18	64	0
AB 39	39	0	0	AB 18	18	0	0
BB 0	0	0	0	BB 0	0	0	0
Generation 3							
	AA	AB	BB		AA	AB	BB
	74	26	0		42	58	0
AA 89	63	26	0	AA 98	40	58	0
AB 11	11	0	0	AB 2	2	0	0
BB 0	0	0	0	BB 0	0	0	0
Generation 4							
	AA	AB	BB		AA	AB	BB
	97	3	0		49	43	9
AA 100	97	3	0	AA 100	49	43	9
AB 0	0	0	0	AB 0	0	0	0
BB 0	0	0	0	BB 0	0	0	0
Generation 5							
	AA	AB	BB		AA	AB	BB
	98	2	0		65	35	0
AA 100	98	2	0	AA 100	65	35	0
AB 0	0	0	0	AB 0	0	0	0
BB 0	0	0	0	BB 0	0	0	0

The heterogametic pattern was frequently observed and it contributed to the objective function in two ways. First, the probability of obtaining homozygous sons with the unfavourable genotype BB was reduced (this mating pattern does not maximize the probability of obtaining AA sons but it does maximize the probability of heterozygous sons). Second, the relatively high polygenic values of BB or AB parents were combined with the relatively low polygenic values of the less intensely selected AA parents (*figures 3 and 4*). This mating structure between individuals having the best polygenic values within genotype is not equivalent to negative assortative mating in a polygenic context but it is a negative assortative mating at the major locus. Gomez-Raya and Gibson [6], working in a context where a major gene affected traits not included in the adult selection criteria, reported that negative assortative mating combined with preselection of young individuals outperformed random mating in terms of selection response against the unfavourable allele.

The homogametic pattern was observed when the initial frequency of the favourable allele was low and the genotype effect on performance was large, especially when the selection scheme included a genotypic pre-selection of males. This pattern produced more AA sons and more BB sons than expected under random mating but it affected the objective function positively since the sons BB were culled before progeny testing.

#### 7.4. Recessive case

Figures for the objective function, the genetic gain and the polygenic gain in the animal class 'females' as well as the genotype frequencies of females after six generations of selection are presented in *table VI*. The advantages of the optimal strategies were negligible when the initial frequency of the favourable allele was high. Optimal 2 always yielded the best results but it never outperformed the classical strategy by more than 3 %. These high initial frequency situations will not be discussed further for the recessive case.

When the initial frequency of the favourable allele was low and the major gene effect was small, the superiorities of optimal 2 were 11 % for the objective function and 14 % for the total genetic gain with negligible losses in polygenic gain. Corresponding values for optimal 1 were 5 and 6 %, without advantages in polygenic gains. Genotype frequencies of the AA female class at generation 6 were 73 % for the optimal 2 strategy versus 39 % for optimal 1 and only 19 % for classical selection. The low frequency of the favourable allele in the classical strategy was expected under classical selection theory. Under our model assuming constant variance across genotypes it was observed that, after the first generation of selection, the homozygous BB individuals were preferred to AB individuals which had the same genotype effects but smaller polygenic means.

When the major gene effect was large, advantages of the optimal 2 strategy were 34 % for the objective function, 49 % for the polygenic gain and 26 % for total genetic value. Corresponding figures for optimal 1 were 21, 33 and 17 %. Again, optimal 2 was the most rapid strategy for fixing the favourable allele with 90 % of females having the AA genotype versus 86 % in optimal 1 and 80 % in the classical strategy. Here, the classical strategy was rather efficient for fixing the favourable allele but very inefficient in polygenic gains ( $1.80\sigma$ ; *table VI*) since selected animals were almost exclusively AA. Then, the advantages of the optimized strategies came mainly from the polygenic gains through an adequate compromise between polygenic and genotypic selections.

When the favourable allele was rare and the genotype effects were either large or small, the within genotype thresholds in the optimized schemes took profit from the favourable alleles hidden in the heterozygous parents. The optimized strategies used both the AA and AB subpopulations as candidates for selection thus allowing increased selection pressure within the AA genotypes and avoiding losses in polygenic means while the BB animals were seldom kept as parents thus leading to a relatively fast increase in the frequencies of the favourable allele A.

**Table VI.** Recessive case. Objective function, polygenic gain and genetic gain for four situations and three selection strategies.

Strategy <sup>a</sup>	Final genotype frequencies for females (%)			Objective function <sup>b</sup>	Polygenic gain <sup>b</sup>	Genetic gain <sup>b</sup>
	AA	AB	BB			
High initial frequency of the favourable allele; small genotype effects <sup>c</sup>						
Classical	97	3	0	12.9	3.22	3.55
Optimal 1	96	4	0	100	101	101
Optimal 2	98	2	0	103	102	101
High initial frequency of the favourable allele; large genotype effects <sup>c</sup>						
Classical	98	2	0	26.4	3.17	4.54
Optimal 1	98	2	0	100	100	100
Optimal 2	98	2	0	103	103	101
Low initial frequency of the favourable allele; small genotype effects <sup>c</sup>						
Classical	19	50	31	8.9	3.13	3.28
Optimal 1	39	49	12	105	100	106
Optimal 2	73	27	0	111	98	114
Low initial frequency of the favourable allele; large genotype effects <sup>c</sup>						
Classical	80	20	0	12.4	1.80	4.85
Optimal 1	86	14	0	121	133	117
Optimal 2	90	10	0	134	149	126

<sup>a</sup> Optimal 1: without male genotypic preselection; optimal 2: with male genotypic preselection. <sup>b</sup> Values for classical selection expressed as genetic standard deviation units. Other values are expressed as percentage of classical selection results. <sup>c</sup> Low and high initial frequencies of the favourable allele are, respectively 0.2 and 0.8. Small and large genotype effects are, respectively, [ 1 0 0 ] and [ 4 0 0 ] genetic standard deviation units.

Details on the proportions of selected individuals and mating structures of the optimized schemes are not given since the general trends were similar to those for the additive case. For the classical scheme, selected proportions of BB individuals were always higher than selected proportions of AB individuals. The proportions of AA females selected and AA tested males selected in the classical scheme were higher than those in optimal 1 and optimal 2. Within optimized strategies, optimal 2 selected fewer AA individuals than optimal 1 but fixation was more rapid in optimal 2 because this strategy could eliminate young males of unfavourable genotypes. As in the additive case, the optimized strategies produced homogametic matings in early generations and heterogametic matings afterwards.

### 7.5. Dominant case

Table VII gives an overview concerning the dominant case. As in the previous recessive and additive cases, a high initial frequency of the favourable allele did not allow discrimination of selection strategies over a six generation time horizon. When the major gene effect was small, optimal 2 offered the advantage of fixing the favourable allele rapidly but the sum of frequencies of the favourable female genotypes AA and AB approached 90 % for the three strategies at generation 6.

**Table VII.** Dominant case. Objective function, polygenic gain and genetic gain for four situations and three selection strategies.

Strategy <sup>a</sup>	Final genotype frequencies for females (%)			Objective function <sup>b</sup>	Polygenic gain <sup>b</sup>	Genetic gain <sup>b</sup>
	AA	AB	BB			
High initial frequency of the favourable allele; small genotype effects <sup>c</sup>						
Classical	74	24	2	14.0	3.34	3.36
Optimal 1	75	23	2	100	101	101
Optimal 2	96	4	0	100	99	100
High initial frequency of the favourable allele; large genotype effects <sup>c</sup>						
Classical	98	2	0	29.6	3.23	3.42
Optimal 1	86	13	1	101	103	101
Optimal 2	96	4	0	101	101	100
Low initial frequency of the favourable allele; small genotype effects <sup>c</sup>						
Classical	40	47	13	11.31	3.01	3.52
Optimal 1	35	50	15	103	105	104
Optimal 2	55	40	5	106	105	107
Low initial frequency of the favourable allele; large genotype effects <sup>c</sup>						
Classical	51	41	8	22.02	2.72	4.98
Optimal 1	70	28	2	104	100	104
Optimal 2	83	17	0	109	105	109

<sup>a</sup> Optimal 1: without male genotypic preselection; optimal 2: with male genotypic preselection. <sup>b</sup> Values for classical selection expressed as genetic standard deviation units. Other values are expressed as percentage of classical selection results. <sup>c</sup> Low and high initial frequencies of the favourable allele are, respectively, 0.2 and 0.8. Small and large genotype effects are, respectively, [ 1 0 0 ] and [ 4 0 0 ] genetic standard deviation units.

The advantages of the optimized strategies were slightly better for a low initial frequency of the favourable allele, especially when the major gene had a large genotype effect. The superiorities of optimal 2 were +9, +5 and +9 % for the objective function, the polygenic gain and the total genetic gain, respectively. Corresponding values for optimal 1

were +4 %, 0 and +4 %. Here, classical selection kept breeding animals from both the AA and AB subpopulations, thus allowing moderate proportions selected and a good polygenic response. In this situation, the main advantage of both optimal 1 and optimal 2 was the ability to eliminate more rapidly the unfavourable allele hidden in the AB individuals.

## 8. DISCUSSION

The main point to highlight is that using more information, performances plus genotypes, yields better results than ignoring information in terms of polygenic and genotypic gains when both types of information are optimally combined by allowing dynamic rules for within genotype selection and nonrandom mating. For some situations studied here (e.g. additive or recessive alleles with a large favourable effect on genotypes) advantages of the optimized schemes came from both the polygenic gain and the fixation rate of the favourable allele. For the additive case, the advantages of the optimal 2 scheme in total genetic gain of 6 and 9 % for rare genes of small and large effect, respectively, are slightly higher than literature results assuming major genes identified without error [5, 9, 15].

These previous studies, based on single generation optimization and a partial use of genotype information (i.e. to just combine genotypes and performances in an index without altering the single threshold selection framework nor considering the mating structures), reported losses in polygenic response over time for mass or family selection. As shown here, this disadvantage can be avoided when applying optimum dynamic within genotype selection rules allowing nonrandom mating. The relative contributions of, on the one hand, dynamic rules for within genotype selection and, on the other hand, dynamic nonrandom mating, were not quantified in the present application and this topic merits further research.

We confirmed, as in Kashi et al. [8], the benefits of a preselection of young males based on their major genotype. In this situation, adult selection is not the only source of gains at the major locus since the selection of young males creates additional selection pressure on the major gene. This allows a relatively high polygenic selection within parents of favourable genotypes while keeping a relatively rapid fixation of the favourable allele at the major locus. For the additive case, preselection of young males yielded an extra polygenic gain of 20 % over selection ignoring the genotypic information, for the intermediate value of polygenic heritability chosen here and a large effect of the major gene on performances. Adverse effects of this across family genotypic selection on the inbreeding rate and its consequences for genetic response merit further research under a dynamic model. In actual applications where the genotype information is available, e.g. Manfredi et al. [10] for dairy goat selection, breeders can maximize the probability of success at later selection stages by selecting within family the sons of favourable genotypes.

As in Larzul et al. [9], important advantages of using genotypic information should be expected when a favourable allele is rare and recessive. The superiorities in total genetic gain reported here are similar (optimal 1) or higher (optimal 2) than that reported by Larzul et al. [9] using single generation optimization and single threshold selection. But, their strategy leads to polygenic losses which diminished the long term benefits of using genotype information.

Dynamic selection is particularly useful to the design of breeding schemes when information on performance is combined with information on a major gene and selection is complex (traits expressed in only one sex; combined progeny, mass and ancestor selection). We have shown that all variables studied (proportions selected and mating structures) may change over time. Other time horizons as well as multiallelic loci can be described with the proposed model as in Barbieri [1] who studied goat selection strategies including the complex polymorphism of the alpha-s1 casein. Also, in the present application total

proportions selected have been kept constant over generations but they could be considered as variables with minor changes in the model.

The model does not take into account the changes in polygenic variances over time and across genotypes. Also, selection is for only one trait and overlapping generations are not considered. Inclusion of the changes in genetic variances in the proposed model is possible for infinite size populations but major changes in the mathematical model would be needed in order to take into account the effect of finite population sizes on the polygenic variances. Interactions between the mating structures, the within genotype proportions selected and the corresponding selection intensities could affect polygenic variances in both the optimized and the classical schemes making it difficult to speculate on how the benefits of the optimized over classical schemes reported here could be changed by modelling polygenic variances.

Including overlapping generations adds no theoretical difficulty to the model proposed. Animals within genotype would then be subdivided into age classes, as proposed by Elsen and Mocquot [4] and Hill [7]. More variables, polygenic means and frequencies for age-major genotype-animal classes, should be added to the optimization problem while using the ageing expressions proposed by Elsen [3].

The combined use of genotypic and performance selection may be enhanced in a multiple trait selection context. Multiple trait selection could also be considered in our model by changing the objective function and adding the computation of polygenic means for each trait. This affects mainly the computation of the objective function but the structure of the model will remain as presented here.

MAS results, obtained under quite different genetic models and schemes but comparable conditions of intermediate heritability and short term selection of traits expressed in females only, reported advantages of using molecular information between 9 and 30 % [8, 11–13] but polygenic losses occurred [13]. Consideration of QTL markers is another potential extension of the proposed model. In this case, the genotype frequencies of progeny would be functions of the parental frequencies and the recombination rates at the population level.

The advantage of this deterministic approach is the possibility of evaluating many possible strategies with a general model including simultaneous adult selection, preselection of young animals and mating structures, to summarize the main trends and to study by stochastic simulation the most interesting situations.

## ACKNOWLEDGMENT

We thank three anonymous referees for their useful revisions of our manuscript.

## REFERENCES

- [1] Barbieri M.E., Polymorphisme de la caséine  $\alpha$ s1. Effets des génotypes sur des performances zootechniques et utilisation en sélection caprine, Ph.D. thesis, Institut national agronomique, Paris-Grignon, 1995.
- [2] de Koning G.J., Weller J.I., Efficiency of direct selection on quantitative trait loci for a two-trait breeding objective, *Theor. Appl. Genet.* 88 (1994) 669–677.
- [3] Elsen J.M., Mocquot J.C., Recherches pour une rationalisation technique des schémas de sélection des bovins et des ovins, *Bull. Tech. Département. Génét. Anim., Inra*, Paris, 1974.
- [4] Elsen J.M., Prediction of annual genetic gain and improvement lag between populations, *Genet. Sel. Evol.* 25 (1993) 75–82.

- [5] Gibson J.P., Short term gain at the expense of long-term response with selection of identified loci, Proceedings of the 5th World Congress Genet. Appl. Livest. Prod., 7–12 August, Guelph, vol. 21, University of Guelph, Guelph, Ontario, Canada, 1994, pp. 201–204.
- [6] Gomez-Raya L., Gibson J.P., Within-family selection at an otherwise unselected locus in dairy cattle, *Genome* 36 (1993) 433–439.
- [7] Hill W.G., Prediction and evaluation of response to selection with overlapping generations, *Anim. Prod.* 18 (1974) 117–140.
- [8] Kashi Y., Hallerman E., Soller M., Marker-assisted selection of candidate bulls for progeny testing programmes, *Anim. Prod.* 51 (1990) 63–74.
- [9] Larzul C., Manfredi E., Elsen J.M., Potential gain from including major gene information in breeding value estimation, *Genet. Sel. Evol.* 29 (1997) 161–184.
- [10] Manfredi E., Ricordeau G., Mahé M.F., Leroux C., Elsen J.M., Martin P., Grosclaude F., Importance of the  $\alpha s1$ -casein polymorphism in dairy goat genetics, 46th EAAP Meeting, session IV, 48, 4–7 September, Prague, Wageningen Pers, Wageningen, 1995.
- [11] Meuwissen T.H.E., Van Arendonk J.A.M., Potential improvements in rate of genetic gain from marker-assisted-selection in dairy cattle breeding schemes, *J. Dairy Sci.* 75 (1992) 1651–1659.
- [12] Meuwissen T.H.E., Goddard M.E., The use of marker haplotypes in animal breeding schemes, *Genet. Sel. Evol.* 28 (1996) 161–176.
- [13] Ruane J., Colleau J.J., Marker-assisted selection for a sex-limited character in a nucleus breeding population, *J. Dairy Sci.* 79 (1996) 1666–1678.
- [14] Verrier E., Colleau J.J., Foulley J.L., Marker assisted selection for increasing the sum of two antagonistic traits under mixed inheritance, 46th EAAP Meeting, session III, 37, 4–7 September, Prague, 1995.
- [15] Woolliams J.A., Pong-Wong R., Short versus long-term responses in breeding schemes, 46th EAAP Meeting, session III, 35, 4–7 September, Prague, 1995.