Original article

Effect of including major gene information in mass selection: a stochastic simulation in a small population

F Fournet, JM Elsen, ME Barbieri, E Manfredi

Station d'amélioration génétique des animaux, Institut national de la recherche agronomique, 31320 Castanet-Tolosan, France

Summary – A quantitative trait under the control of a major gene plus a finite number of genes with small effects was described using a stochastic model where number, size and linkage between QTL may vary. Selection schemes defined by the selection criteria (individual phenotype, major genotype and combination of both sources of information), the population size and the selection intensities in male and female paths were considered. Different genetic hypotheses were studied concerning the major gene effect, the number of small quantitative loci and the linkage between genes. The ranking of the selection schemes over 30 generations was performed with the following criteria: time taken for the fixation of the favourable A allele at the major locus and differences between the cumulated discounted gains obtained with each scheme. The interactions between the major gene and the flanking QTLs were also studied. The main result was that the inclusion of major gene information in selection schemes was mostly efficient in the medium and long term when the gene was rare and recessive and in the medium term when it was rare and additive, essentially due to a rapid fixation of the favourable A allele and to a limited risk of losing it by genetic drift for a rare recessive gene.

major gene / QTL / selection / Monte-Carlo

Résumé – Inclusion de l'information à un locus majeur en sélection massale : un modèle stochastique dans une petite population. *Un caractère quantitatif sous le contrôle d'un gène majeur et d'un nombre fini de gènes à effets faibles est décrit à l'aide d'un modèle stochastique où le nombre de QTL et leur liaison peuvent varier. Plusieurs schémas de sélection, définis par leurs critères de sélection (performance individuelle, génotype au gène majeur ou combinaison des deux types d'information), la taille de la population et l'intensité de sélection pour les voies mâle et femelle, sont considérés. Différentes hypothèses génétiques sont envisagées, concernant l'effet du gène majeur, le nombre de QTL et la liaison entre locus adjacents. Le classement des schémas de sélection sur 30 générations est effectué à l'aide des critères suivants : le temps nécessaire à la fixation de l'allèle favorable A au locus majeur et les différences entre gains cumulés actualisés obtenus avec chaque schéma. Les interactions entre le gène majeur et les polygènes avoisinants sont également étudiées. Le principal résultat est que l'inclusion de l'information relative au gène majeur dans les schémas de sélection est surtout efficace*

à moyen et long terme quand le gène est rare et récessif, et à moyen terme quand il est rare et additif. Cela est essentiellement dû à une fixation rapide de l'allèle favorable A et à un risque limité de perte du gène par dérive génétique dans le cas d'un gène récessif rare.

gène majeur / QTL / sélection / Monte-Carlo

INTRODUCTION

Many models describing the evolution of genetic variability in response to selection are based on the assumption that a trait is controlled by an infinite number of small independent genes. Nevertheless, the evidence for a small number of QTL with medium to large effects on quantitative traits is increasing in livestock (Mérat and Ricard, 1974; Ollivier, 1980; Piper and Bindon, 1982; Le Roy et al, 1990; Tanksley, 1993). To take more advantage of this genetic variability for animal improvement, specific evaluation methods and selection schemes should be applied (Smith, 1967; Soller, 1978; Smith and Webb, 1981; Smith, 1982; Stam, 1986; Hoeschele, 1990; Kennedy et al, 1990; McLaren et al, 1990; Sehested and Mao, 1992; Gibson, 1994; Ruane and Colleau, 1995; Whittaker et al, 1995; Larzul et al, 1997). These papers pointed out the value of considering the major gene characteristics, ie, the favourable allele initial frequency, and the type of genetic determinism (dominance or additivity, allele effects). They also showed that the evolution of the polygenic distribution depends on the way in which major gene information is taken into account, with the extreme case where maximal extra-response due to a segregating locus (in proportion of fixable locus effect) is obtained when counter-selecting the major gene (Gibson, 1994).

This study attempted to achieve a more precise description of the coevolution, due to selection, of the distribution of a major gene and of the other QTLs controlling the selected trait. In the simulation, the genome of the individuals was described using a stochastic model in which the polygenic inheritance was described by a finite number of linked genes with additive effects. In particular, this model allowed a precise study of the evolution of the genetic variance and the influence of the major gene on its flanking QTLs. The effects of three selection methods, for a trait measurable in the two sexes, were described. To simplify the genetic interpretation of the results, these selection methods were all based on individuals' phenotypes and differed by the way in which the major gene information was included (or not) in the selection criterion. When it was included, the individual genotypes at the major gene and the effects of each possible genotype on the trait were suppposed to be known without error.

METHODS

Description of the model

The algorithm used, introduced by Hospital (1992) and further developed by Fournet et al (1995), was based on a model which describes each individual of the selected population by a finite set of QTLs with a finite number of alleles per locus. The genome was made of identified chromosomes, in the sense that the QTLs

are pooled in sets of equal size (the major gene being located in the middle of the first set), independent from each other but with linkage within-set. Two sizes of genome were simulated in order to study the critical influence of this parameter: 5 chromosomes with 2 QTLs on each (total = 10 QTLs) and 10 chromosomes with 10 QTLs on each (total = 100 QTLs). The recombination rates between any adjacent QTLs, including the major gene (except when the distance between the major locus and the two neighbouring QTLs was varied), were kept identical (0.09) in the two situations. No interference was assumed. In this work, the genes were biallelic and the allele effects were given the values α or $-\alpha$ at any QTL except the major locus. Genotypes in the first generation were given initial frequencies p_i of the favourable allele at each locus, these frequencies being drawn from a $(0,1)$ uniform distribution. Once these first generation genotypes had been simulated, the polygenic genetic variance was calculated from:

$$\sum_{l=1}^{L} 2p_i(1 - p_i)\alpha^2$$

where L is the number of QTLs, and α the gene effect. The environmental variance was adjusted in order to obtain a given within-major genotype heritability $h^2 = \sigma_{a_{pol}}^2/(\sigma_{a_{pol}}^2 + \sigma_e^2)$ in the first generation of selection and added to the polygenic variance $\sigma_{a_{pol}}^2$ to obtain the within-major genotype residual phenotypic variance. This approach implies that initial heritability is constant for all cases studied while the evolution of heritability along generations varies according to each case. Various hypotheses were made on the major gene contribution (initial frequency of the favourable allele A, level of dominance and difference $(G_{AA}-G_{BB})$ between homozygous genotypes). These are detailed below. The values of the three possible genotypes were expressed in residual phenotypic standard deviation units and added to the polygenic effects to give the genetic values of the individuals. Environmental effects were randomly drawn from a Gaussian distribution $N(0, \sigma_e^2)$ and added to genetic values to generate phenotypic values.

The generations did not overlap. As described below, the males and females underwent steps of evaluation and directional selection. The selected individuals were randomly mated, their gametes were formed by the parental chromosomes going through meiosis and recombination, and the offspring genotypes were generated by pairing of the paternal and maternal gametes (see Fournet et al, 1995, for technical details). The new-born individuals then replaced their parents and went through the same steps, with the same cycle being performed until the 30th generation of selection was reached. An entire run of 30 generations of selection was performed 100 times for each case studied. The mean values for total genetic mean and variance (ie, accounting for the major gene and the QTLs), total phenotypic mean and variance, major gene frequency, QTL's genetic mean and variance on the chromosome carrying the major gene and on the non-carrier ones, were calculated per generation over the 100 repetitions and screened out. This oligogenic model, where genetic means and variances are calculated from genotypes at each locus, accounted for the decrease of genetic variance due to changes in gene frequencies and to disequilibrium between loci (Bulmer effect).

Selection methods

Three different selection methods, depending on the way in which the major gene information was included in the selection criteria, were considered. They were chosen to be as simple as possible, in order to avoid any confusing parameter that would make the results difficult to interpret. In particular, 'Animal Model-BLUP' techniques were not considered here; this point will be discussed in the *Discussion and conclusion.*

Phenotypic selection: the male and female candidates were evaluated on the basis of their own performances, with the best ones being selected as breeders. This method was considered as the standard selection method, in which the major gene information is ignored.

Genotypic selection: the candidates were selected first on their major genotypes (*AA* first, then *AB* and possibly *BB*) and then, for the last genotype retained, on their phenotype.

Combined selection: following Larzul et al (1997), the candidates were evaluated on the expected genetic values of their offspring, calculated as the sum of the offspring expected additive polygenic value and their expected value at the major locus, accounting for the known major genotype of the candidate and for the genotype distribution in the population of mates.

These selection methods were, respectively, called S_P, S_G and S_C.

Comparison criteria

Three criteria were chosen to compare the efficiency of the three selection methods.
– The time taken for the fixation of the favourable allele at the major locus described by (1) the generation number when the A allele frequency reached 0.95, and (2) the generation number upon complete fixation.
– The differences between the cumulative discounted gains obtained with combined and phenotypic, combined and genotypic, genotypic and phenotypic selection methods (expressed as a percentage of the cumulated discounted gain for the standard phenotypic selection method). The cumulated discounted gain for a selection method was given by:

$$C(t_f, \theta) = \sum_{t=0}^{t_f} \left(\frac{1}{1+\theta}\right)^t P_t$$

with t the generation number, θ the discounting rate (assumed to be 5% per year), t_f the length of the discounting period and P_t the phenotypic response from generation $t-1$ to t for the given selection method. The choice of this comparison criterion was supported by the fact that, whatever the selection method considered, the favourable alleles would be fixed in the long term, but the dynamics of this fixation and of the phenotypic means would differ from one scheme to another. Discounting is a classical tool used by geneticists (Poutous and Vissac, 1962; Hill, 1974; Cunningham and Ryan, 1975; Smith, 1977; Miller and Pearson, 1979) for taking into account the time when genetic gains are obtained and it has been

applied to numerous selection schemes (Soller et al, 1966; Hinks, 1970; Danell et al, 1976).

Two situations were considered: in both cases 30 years of selection, but represented by either six generations of 5-year olds (cattle) or 30 generations of 1-year olds (rabbit, poultry).
– The differences between total genetic response obtained after 30 years with each selection method were also given.

The statistical significance of all the comparisons presented between methods (except time to fixation) was evaluated with a Student's t-test, using the standard deviations for each parameter over the 100 repetitions of the simulation.

Cases studied

In the first part of the study, a population of 192 individuals (half males, half females) was simulated. Twenty five percent of the males and 50% of the females (24 males and 48 females) were selected as parents for the next generation. A 'within-major genotype' heritability of 0.25 was assumed for the selected trait. The ranking of selection methods was studied for various values of the parameters defining the major gene, and for the two numbers of QTLs, in order to test the sensitivity of this ranking to the characteristics of the genome.

Initial frequency of the favourable allele

The initial frequency of the favourable allele A $fq(A)$ at the major locus was given the values 0.1, 0.5 or 0.9.

Mode of inheritance

Three kinds of genetic determinism at the major gene were tested: additivity of the two alleles A and B, dominance of the A allele, dominance of the B allele.

Effect of the major gene

The difference between the mean values of the two homozygotes $(G_{AA}-G_{BB})$ was assumed to be 1, 2 or 3 within-major genotype phenotypic standard deviations.

In the second part of the study, the evolution of the ranking of the three selection methods with parameters defining the population management (size of the population and selection intensity) was studied, for three cases only, chosen as the most informative after the previous comparison: a recessive favourable allele, with initial frequencies of 0.1 or 0.5, and a rare additive favourable allele at the major locus. For all cases, a medium effect $(2\sigma_P)$ was given to the major gene. Two population sizes N (192 and 480 individuals) and for each size, two proportions p of selected males (25 and 6.25%), were tested, with a small and a large polygenic genome. The proportion of selected females remained equal to 50%.

RESULTS

Characteristics of the genome

100 QTLs and a major gene

The results differed widely depending on the major gene effect, genetic determinism or initial frequency of the favourable allele, as illustrated in figure 1a and b showing the evolution of the phenotypic response in the three selection methods, with an initial A allele frequency of 0.1, respectively, for a gene of large effect (G_{AA}– $G_{BB} = 3\sigma_P$) with A recessive (fig 1a) and for a gene of moderate effect (G_{AA}– $G_{BB} = 1\sigma_P$) with A dominant (fig 1b). The ranking of the selection methods varied with the type of major gene, with S_G being the better method and S_P the worst in the first situation, and the contrary in the second case. As S_C was always intermediate between the other two methods, results presented in the following tables will focuse on S_P and S_G.

For a higher initial frequency, the differences between methods were less striking, and vanished when $fq(A)$ was 0.9. Thus the discussion will focus more on results obtained with an initial frequency of 0.1.

Table I shows the time taken for the fixation of the favourable allele in all the situations studied. For phenotypic selection, fixation speed increased with the initial frequency and the major gene effect. When considering an initial A allele frequency of medium to high, whatever the effect of the gene, a recessive favourable allele was easier to select than an additive one, itself easier to select than a dominant one: if A is dominant, AB animals are eliminated in the genotypic method, they are ranked as AA and kept in the phenotypic method and have a good chance of being chosen in the combined method, while maintaining a high percentage of the B allele in the population as compared with the genotypic selection. The time taken to reach fixation was thus very long when A was dominant, due to the fact that only individual information was used, a well known result in population genetics (Falconer, 1981; Larzul et al, 1997).

When considering a low initial A allele frequency, the fixation of the A allele was easier for an additive allele than a dominant one, itself easier than a recessive one, as expected. The recessive A allele was difficult to select due to the risk of losing it. Indeed, the proportion of AA genotypes was almost zero and AB was rare in the first generations and if the polygenic values of the few heterozygotic animals were low, these animals could be eliminated at the beginning of the selection process in the phenotypic selection method. This phenomenon was observed in the case of a large major gene, where the mean A allele frequency reached a plateau at 0.91: 9% of the runs showed a loss of the favourable major allele. The dynamic aspect of this fixation process can be outlined. In the first generations, the favourable A allele had a random risk of being lost due to its low frequency. If and when the A allele was not lost, its frequency followed the evolution described previously for medium to high frequency. The fixation process of the favourable allele at the major locus then depended on the combination of both phenomena, the risk of being lost in the beginning and the rapid fixation after.
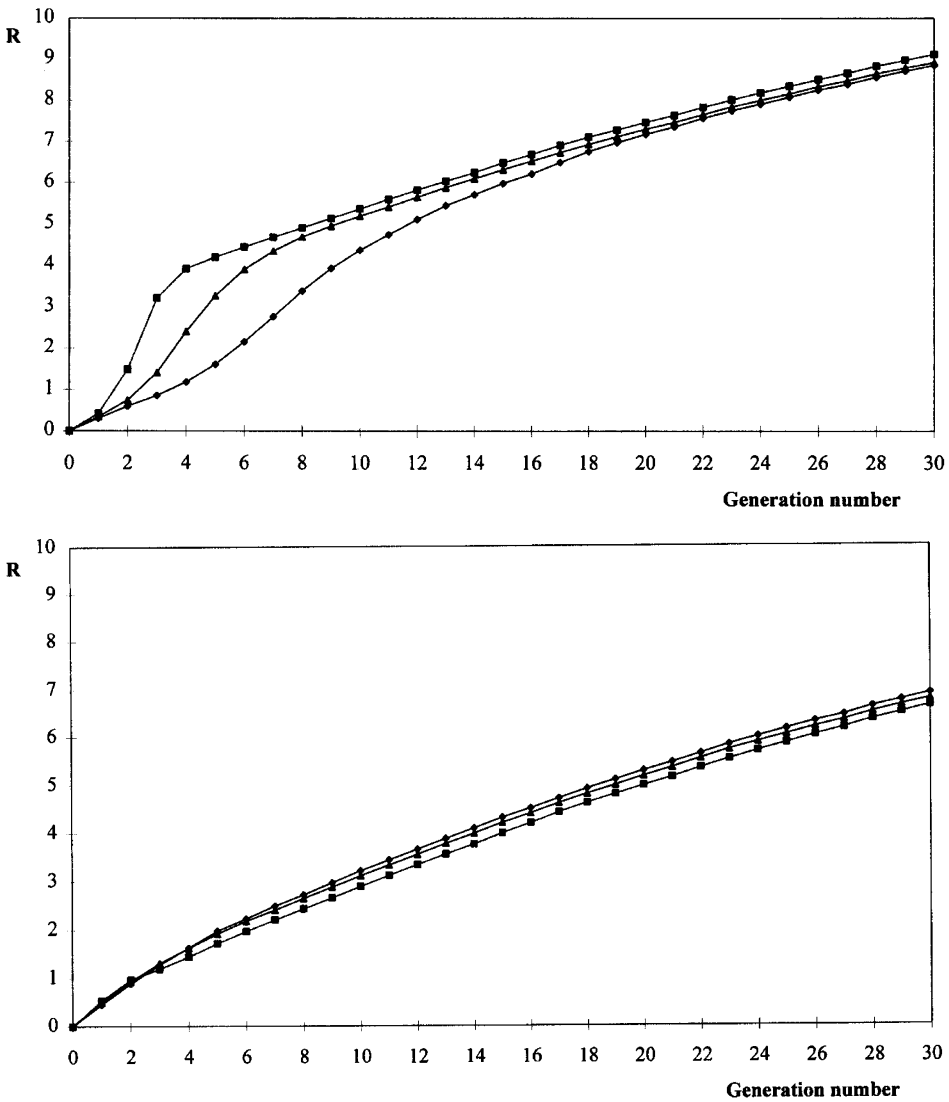
Fig 1. Response to selection (R) in three selection methods: (a) for a rare recessive major gene of large effect $(3\sigma_P)$; (b) for a rare dominant major gene of small effect $(1\sigma_P)$. ♦ S_P, ■ S_G, ♠ S_C.

In genotypic selection, the time taken to reach fixation depended only on the initial frequency of the A allele. The fixation was always very fast (in generation 4 for a low initial frequency and in generation 1 for $fq(A) = 0.9$).

Generally, the ranking of the selection methods concerning time to fixation was S_G, S_C and S_P (fig 2a and b, concerning the same cases as fig 1a and 1b). The other

F Fournet et al

Table I. Generation when the favourable A allele is 95% fixed (completely fixed) for phenotypic (S_P) and genotypic (S_G) selection methods; 100 QTLs, 24 males and 48 females selected among 96 candidates of each sex.

Initial frequency of A	Selection method	A and B additive	A dominant	A recessive
$G_{AA}-G_{BB} = 3\sigma_P$				
0.1	S_P	5 (7)	19 (nf)[a]	0.925 at generation 30[b]
	S_G	4 (4)	4 (4)	4 (4)
0.5	S_P	3 (5)	17 (nf)	2 (3)
	S_G	2 (2)	2 (2)	2 (2)
0.9	S_P	1 (3)	11 (nf)	1 (2)
	S_G	1 (1)	1 (1)	1 (1)
$G_{AA}-G_{BB} = 2\sigma_P$				
0.1	S_P	7 (11)	18 (nf)	0.89 at generation 30
	S_G	4 (4)	4 (4)	4 (4)
0.5	S_P	3 (7)	17 (nf)	3 (4)
	S_G	2 (2)	2 (2)	2 (2)
0.9	S_P	1 (5)	11 (nf)	1 (2)
	S_G	1 (1)	1 (1)	1 (1)
$G_{AA}-G_{BB} = 1\sigma_P$				
0.1	S_P	13 (22)	28 (nf)	0.824 at generation 30
	S_G	4 (4)	4 (4)	4 (4)
0.5	S_P	7 (17)	22 (nf)	5 (9)
	S_G	2 (2)	2 (2)	2 (2)
0.9	S_P	2 (12)	14 (nf)	1 (5)
	S_G	1 (1)	1 (1)	1 (1)

[a] nf = not fixed. [b] $fq(A) = 0.925$.

striking point on these two figures was the genetic lag in polygenic response during the first generations when selecting first on the major gene (S_G): selection pressure was put only on the major gene, and hardly any genetic gain was obtained on the QTLs. This polygenic lag was recovered in the case of a rare recessive favourable allele (fig 2a) but not in the situation of a rare dominant A allele (fig 2b).

The values of the cumulated differences of discounted gains between selection methods are presented in tables II and III.

For the long run (table II), rather small differences were observed between S_G and S_P, except for a favourable A allele rare and recessive where, as explained before, including the major gene information meant avoiding the risk of losing the favourable allele. In this situation, the selection methods ranked from S_G to S_P, with a superiority of the genotypic selection method over standard phenotypic selection significant at the 1% level, less important when decreasing the effect of the major gene. On the contrary, the phenotypic method S_P was found to be better than the other methods for a 'small' major gene with A allele additive or dominant, with much smaller differences. In this case, the largest difference was
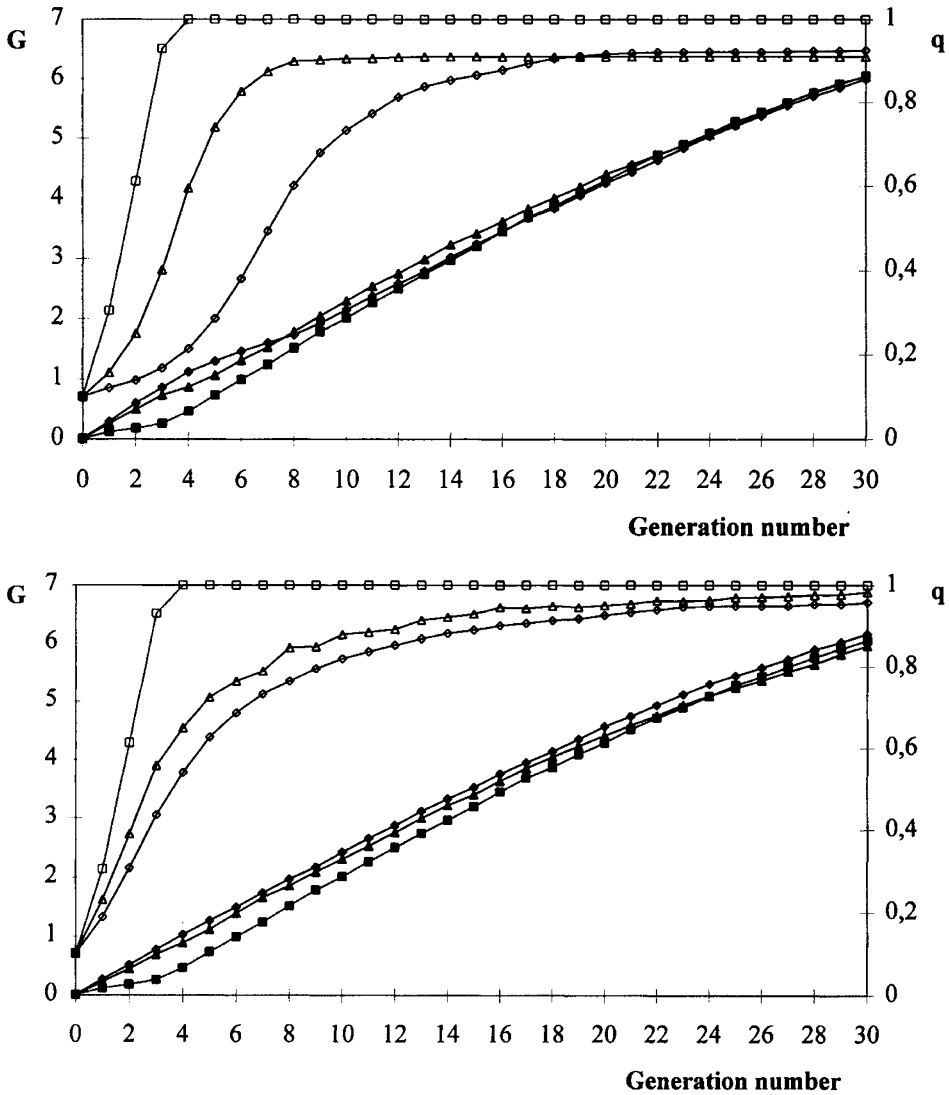
Fig 2. Polygenic genetic mean (G) and A allele frequency (q) in three selection methods: (a) for a recessive major gene of large effect ($3\sigma_P$); (b) for a rare dominant major gene of small effect ($1\sigma_P$). G: ◆ S_P, ■ S_G, ♠ S_C; q: ◊ S_P, □ S_G, △ S_C.

found between phenotypic and genotypic methods, the combined method being intermediate. This suggests that the effort made for the fixation of the favourable allele at the major locus, resulting in a polygenic lag in the genotypic selection method, was too expensive given the weak gain due to the major gene.

Table II. Differences between genotypic and phenotypic selection in cumulative discounted gains, in percentage of the phenotypic method, with a discounting rate of 0.05 and 30 generations of selection; 100 QTLs, 24 males and 48 females selected among 96 candidates of each sex.

Initial frequency of A	A and B additive	A dominant	A recessive
$G_{AA}-G_{BB} = 3\sigma_P$			
0.1	−0.7% NS	−2.1% **	24.0% ***
0.5	−0.9% NS	−1.6% **	0.7% NS
0.9	−0.5% NS	−1.0% NS	−0.1% NS
$G_{AA}-G_{BB} = 2\sigma_P$			
0.1	−1.4% NS	−3.2% **	17.1% ***
0.5	−1.6% NS	−2.6% *	0.7% NS
0.9	−1.5% NS	−1.1% NS	−0.2% NS
$G_{AA}-G_{BB} = 1\sigma_P$			
0.1	−4.7% *	−6.8% ***	8.1% ***
0.5	−2.6% NS	−5.3% **	−1.0% NS
0.9	−0.9% NS	−1.2% NS	−1.3% NS

NS not significant. ** Significant at the 5% level. *** Significant at the 1% level.

The contrasts between S_G and S_P were enhanced when considering the short run (table III) and were generally in favour of the genotypic method, but with a lower significance level: the superiority over the phenotypic method was 168, 105 and 34% for a rare recessive major gene with decreasing effect (3, 2 or 1 σ_P). For a low initial frequency, whatever the effect of the major gene, the value of S_G versus S_P was high for the recessive case, medium for the additive case and low to negative for the dominant case (table III). This was directly linked to the fixation rate of the A allele (see fig 2b). When the initial frequency was 0.9, the superiority of S_G over S_P was medium to low when A was additive, almost zero for A recessive and medium to weakly negative for A dominant. The relative superiority of combined selection over phenotypic selection was less striking than the relative superiority of genotypic selection over phenotypic selection, but the trend was the same.

In conclusion, the results obtained in the medium and long term showed the same tendency, with S_G being valuable not only for the recessive case but also for the additive one in the medium term. The differences in total genetic response between S_G and S_P, expressed in percentage of S_P, are presented in table IV. It can be noted that in some cases, S_G yielded better polygenic genetic responses than

Table III. Differences between genotypic and phenotypic selection in cumulative discounted gains, in percentage of the phenotypic method, with a discounting rate of 0.05 and six generations of selection; 100 QTLs, 24 males and 48 females selected among 96 candidates of each sex.

Initial frequency of A	A and B additive	A dominant	A recessive
$G_{AA}-G_{BB} = 3\sigma_P$			
0.1	9.7%	1.5%	167.7%
	NS	NS	**
0.5	1.7%	3.1%	−0.1%
	NS	NS	NS
0.9	8.2%	−1.1%	0%
	NS	NS	NS
$G_{AA}-G_{BB} = 2\sigma_P$			
0.1	14.1%	0.8%	105.0%
	NS	NS	**
0.5	3.2%	−2.4%	−0.5%
	NS	NS	NS
0.9	0.9%	−2.6%	0.1%
	NS	NS	N
$G_{AA}-G_{BB} = 1\sigma_P$			
0.1	12.4%	−4.2%	33.8%
	NS	NS	NS
0.5	2.1%	−10.9%	−1.7%
	NS	NS	NS
0.9	1.4%	−4.2%	0.8%
	NS	NS	NS

NS not significant. ** Significant at the 5% level. *** Significant at the 1% level.

S_P. As it will be shown later, this can be attributed to smaller loss of polygenic variance in the S_G method. No significant differences were observed, neither in the QTLs nor in the total genetic response. This result was obtained probably because, whatever the selection method used, the favourable allele at the major gene would be fixed in the long term. But the fixation process differs in timing (as shown in fig 2a), leading to a higher phenotypic response in the first generations in S_G, and the comparison of discounted cumulative genetic gains enables this difference to be demonstrated.

Generally, as shown in figures 1a and b, 2a and b, S_G provided, in the first generations, a higher phenotypic response than S_P and S_C, due to a faster fixation of the favourable allele. This was obtained without loss of polygenic variance, these variances remaining very close in the three methods over the 30 years (fig 3a and b), but at the expense of a polygenic lag which was not always recovered (fig 2b). Moreover, polygenic variance might be higher in S_G, due to the lack of selection in the first generations, individuals with poor polygenic values being retained: in table V, the polygenic variance remaining in S_G at generation 5 (this generation being chosen to compare the remaining variances just after the fixation of the A

Table IV. Differences of cumulative genetic responses (for the polygenic part and total) between genotypic and phenotypic selection methods, in percentage of the phenotypic method, for 30 generations of selection; 100 QTLs, 24 males and 48 females selected among 96 candidates of each sex.

Initial frequency of A	Selection method	A and B additive	A dominant	A recessive
$G_{AA}-G_{BB} = 3\sigma_P$				
0.1	Polygenic	4.2%	−0.8%	0.8%
	Total	−1.2%	−1.6%	3.2%
0.5	Polygenic	−0.5%	−1.5%	1.2%
	Total	−0.6%	−1.5%	−0.1%
0.9	Polygenic	0.4%	−3.3%	−0.1%
	Total	−0.4%	−0.4%	−0.1%
$G_{AA}-G_{BB} = 2\sigma_P$				
0.1	Polygenic	1.9%	3.1%	2.9%
	Total	−1.5%	−2.1%	2.3%
0.5	Polygenic	−1.7%	0.6%	0.7%
	Total	−0.8%	−0.7%	0.6%
0.9	Polygenic	−1.7%	−3.1%	1.2%
	Total	−1.0%	−0.1%	0.1%
$G_{AA}-G_{BB} = 1\sigma_P$				
0.1	Polygenic	−3.3%	−1.9%	1.5%
	Total	−4.5%	−3.8%	−0.6%
0.5	Polygenic	0.2%	−2.6%	−0.6%
	Total	−1.1%	−2.0%	−0.8%
0.9	Polygenic	−0.5%	−4.0%	−1.4%
	Total	−0.5%	−0.3%	−0.6%

All differences are not significant.

allele in S_G) was 1.4 and 4.4% higher than the genetic variance obtained with S_P at the same generation, for the cases corresponding, respectively, to figure 2a and b. This might explain the faster evolution of polygenic response observed in S_G after fixation of the major gene (fig 2a and b), although these differences were not significant.

The influence of the major gene on the flanking QTLs was studied more precisely with the deviation between polygenic genetic gain on the chromosome carrying the major gene and on the non-carrying ones, expressed as a percentage of the total polygenic gain. The differences observed (table VI) between the two kinds of chromosomes were globally very weak (no statistically significant differences were found).

For the genotypic selection method, the positive difference of gains obtained on the non-carrier chromosomes relative to the chromosome carrying the major gene depended only on the initial frequency of A, increasing when $fg(A)$ was lower, due to a longer time taken for fixation and a higher polygenic lag on the chromosome carrying the major gene. Indeed, this selection method first concentrated on the fixation of an A allele, which could involve the fixation of unfavourable alleles at
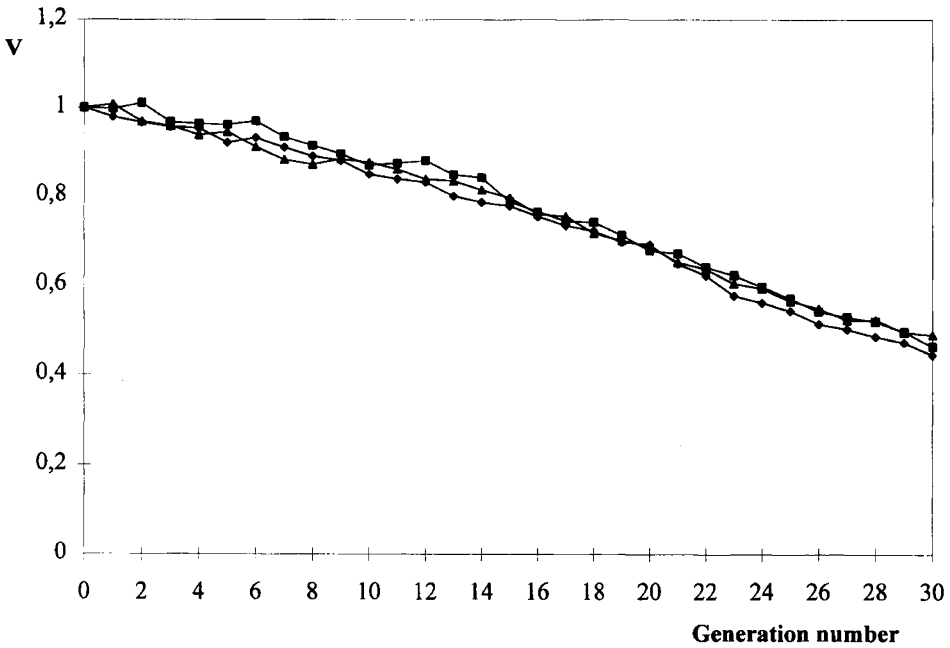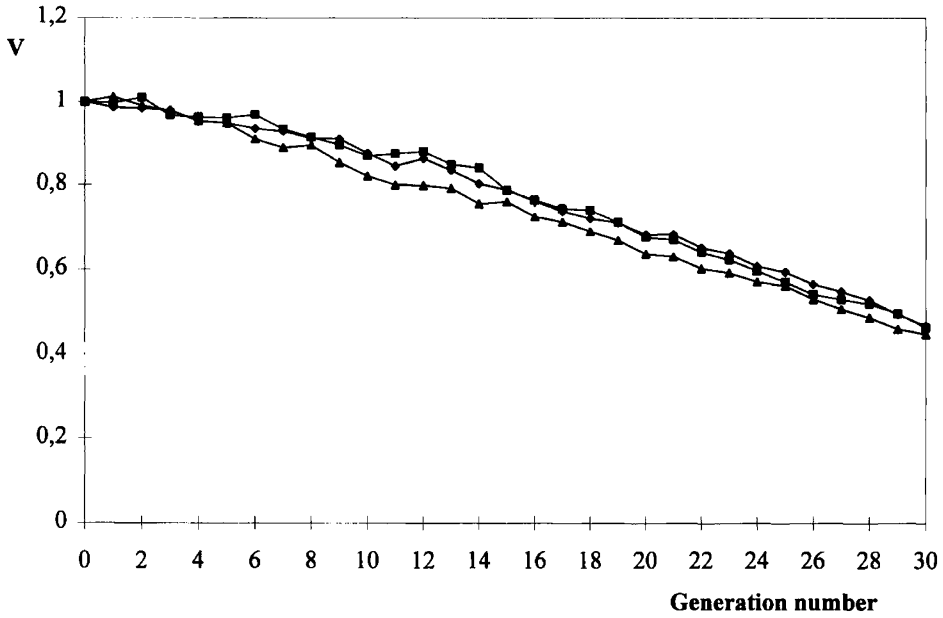
Fig 3. Polygenic variance relative to generation 0 (V) in three selection methods: (a) for a rare recessive major gene of large effect ($3\sigma_P$); (b) for a rare dominant major gene of small effect ($1\sigma_P$). $\blacklozenge$ S_P, $\blacksquare$ S_G, $\spadesuit$ S_C.

Table V. Differences in polygenic variance, at generation 5, between the genotypic selection method S_G and the phenotypic method S_P, expressed in percentage of the phenotypic method; 100 QTLs, 24 males and 48 females selected from 96 candidates of each sex.

Initial frequency of A	A and B additive	A dominant	A recessive
$G_{AA}-G_{BB} = 3\sigma_P$			
0.1	4.4%	1.3%	1.4%
0.5	−1.8%	2.1%	−1.8%
	−1.3%	1.9%	−1.2%
$G_{AA}-G_{BB} = 2\sigma_P$			
0.1	2.9%	3.7%	2.4%
0.5	−1.7%	1.0%	1.1%
0.9	0.1%	1.1%	0.0%
$G_{AA}-G_{BB} = 1\sigma_P$			
0.1	0.1%	4.4%	1.6%
0.5	−2.8%	1.9%	1.2%
0.9	1.8%	2.5%	1.0%

All differences are not significant.

the flanking QTLs. Since the QTLs were not selected during these first generations, the fixation of unfavourable alleles at the QTLs on the chromosome carrying the major gene was not due to the Bulmer's effect but was produced by the genetic linkage between the major gene and the flanking QTLs (hitch-hiking).

For phenotypic and combined selection methods, interactions appeared between the effect and the allelic dominance of the major gene. When the gene effect was large, the maximal difference was observed for A recessive, then A dominant and finally A additive, the opposite being found for a smaller QTL. The difference decreased when the initial frequency increased. This was due to an earlier fixation of the A allele when $fg(A)$ was higher, with the chromosome carrying the major gene and the non-carrier ones behaving identically after the fixation.

Ten QTLs and a major gene

The same study was performed for a small polygenic genome (10 QTLs). In general, the trend was the same for 100 and 10 QTLs, whatever the comparison criteria considered. Therefore, results presented here for 10 QTLs are not as detailed as for the 100 QTLs situation.

The time taken for the fixation of the favourable allele at the major gene was nearly the same between the two genomes. The chance of losing the favourable allele at the major gene when rare and recessive was only slightly higher with 10 QTLs as compared to the 100 QTLs case. Indeed, the risk for AB animals to be eliminated due to a possible low polygenic value, as explained for 100 QTLs, was higher with 10 QTLs as the polygenic part of the genome was smaller. The difference however remains small (1–2%).

Table VI. Observed deviation, at generation 30, between polygenic genetic gain on the chromosome carrying the major gene and non-carrier chromosomes (expressed in percentage of the total genetic gain over 30 generations), for phenotypic (S_P) and genotypic (S_G) selection methods; 100 QTLs, 24 males and 48 females selected from 96 candidates of each sex.

Initial frequency of A	Selection method	A and B additive	A dominant	A recessive
G_{AA}–$G_{BB} = 3\sigma_P$				
0.1	S_P	1%	−5%	−10.6%
	S_G	−7%	−7%	−7%
0.5	S_P	3.7%	1.3%	0.7%
	S_G	−2.9%	−2.9%	−2.9%
0.9	S_P	4.3%	−2.5%	0.7%
	S_G	0.2%	0.2%	0.2%
G_{AA}–$G_{BB} = 2\sigma_P$				
0.1	S_P	−1%	−1.7%	−4.4%
	S_G	−7%	−7%	−7%
0.5	S_P	1.4%	1.6%	−0.1%
	S_G	−2.9%	−2.9%	−2.9%
0.9	S_P	3.7%	−1.4%	0.8%
	S_G	0.2%	0.2%	0.2%
G_{AA}–$G_{BB} = 1\sigma_P$				
0.1	S_P	−4.3%	−1.9%	0.05%
	S_G	−7%	−7%	−7%
0.5	S_P	2.2%	−1.2%	1.5%
	S_G	−2.9%	−2.9%	−2.9%
0.9	S_P	0.2%	−1.7%	2.8%
	S_G	0.2%	0.2%	0.2%

The cumulated differences of discounted gains between the selection methods, for a long term objective, showed a higher superiority of the genotypic method over the phenotypic method with 10 QTLs than for 100 QTLs (26.6, 19.8 and 13.6% for 10 QTLs versus 24.0, 17.1 and 8.1% for 100 QTLs, respectively, for a major gene of 3, 2 and 1 σ_P with a rare recessive A allele). This is consistent with the preceding remark about the time taken for fixation: if the risk for the favourable allele at the major locus to be lost is higher with a small polygenic genome, then a rapid fixation with the genotypic method is more valuable.

When considered in the medium term, the superiority of the genotypic method followed the same trend as with 100 QTLs: the superiority of this method over the phenotypic method was essentially found for a rare recessive allele at major locus, but in the case of 10 QTLs, was extended to the case of a recessive major gene with intermediate initial frequency (C_{GP} is 6.9% for 10 QTLs against −0.5% for 100 QTLs, see table III).

Results concerning the differences in polygenic variance between methods and the behaviour of the QTLs flanking the major locus, with 10 QTLs, were very similar to those with 100 QTLs, and so are not discussed.

Population management

The main conclusion of the first part of this study was that taking account of the major gene information was mostly efficient for a rare recessive major allele, whatever the size of the polygenic genome. Indeed, the genotypic selection method allowed for a rapid fixation of the favourable allele at the major locus, whereas selection based on the phenotype could lead to this allele being lost.

The second part of the study, aimed at testing the effect of the population management parameters, such as population size and selection intensity, was then based on this case of a rare recessive major allele, with a medium gene effect ($2 \sigma_P$). Two other situations were investigated: a recessive favourable allele at a major locus with intermediate initial frequency, because the simulation of a small genome pointed out the value of the genotypic method in this case also, and a rare additive A allele at a major locus, in order to enable a comparison with results in the literature, in which additivity was the most frequent assumption. The following results are presented and discussed for the case of a small polygenic genome (10 QTLs), because the results for 100 QTLs were less pronounced, although very similar.

The results on the time taken for the fixation of the favourable allele showed that the risk of losing a rare recessive favourable allele with the phenotypic selection method was lower when decreasing the selection intensity and increasing the population size. This risk also appeared in the combined selection method for a rare additive favourable allele when the selection pressure was strong and the population size small. The risk of losing the favourable allele at the major gene then depended on a combination between the quantity of favourable alleles in the population and the capacity of the selection method used to pick up the available alleles.

The differences between methods in cumulated discounted gains, when considered for 30 generations, were appreciable only for a rare recessive favourable allele at the major gene. The superiority of the genotypic method over the phenotypic method was the same (around 19%) for any value of N and p except for a large population strongly selected, where this superiority was only 12.6%. In fact, the genetic variability in this case seemed to be large enough and used with sufficient intensity to achieve good progress with the phenotypic selection method, thus reducing the value of the genotypic method. This hypothesis was supported by the fact that the highest phenotypic gain was observed for this combination of N and p.

The same phenomenon was pointed out when considering this comparison criterion for 6 generations, with an extension to the case of a small population strongly selected. The superiority of S_G over S_P, whatever the population size, was found to be lower, although widely positive, when the selection intensity was higher (around 60% for $p = 6.25\%$ against 100% for $p = 25\%$). This was not surprising as the strong selection intensity provided high genetic progress in the first generations, the loss of genetic variability due to the selection (and even more in the small population) being not sensitive enough, at the 6th generation, to reduce this progress.

The differences between the genetic variances in the methods accounting for the major gene information, and in the standard situation, seemed sensitive to population size and selection intensity, although no significant difference was

observed. In the case of a rare recessive favourable allele at the major locus, where the method S_G appeared to be the most appropriate, the polygenic variance at generation 5 was higher when the selection intensity was lower. This phenomenon may be explained by observing the way in which QTLs are selected during the early generations. With a 25% selection rate, all the AA (1%) and AB (18%) animals belonging to the first generation are retained, with polygenic selection occurring only for the BB animals (with a 6/81 selection rate). This lack of selection in carrier genotypes preserves the polygenic variability. On the contrary, with a 6.25% selection rate even the AB are selected with a selection rate of 5.25 over 18 applied to a small AB subpopulation, while, in the phenotypic selection method, the candidate population (the whole population) is much larger, thus leading to a smaller decrease in the polygenic variability. The polygenic variance in method S_G was then, respectively, only 0.9% higher and 2.8% lower than in method S_P for $N = 192$ and $N = 480$ when $p = 6.25\%$, against 11.1 and 13.3% higher when $p = 25\%$.

The last parameter studied was the deviation between polygenic genetic gains on the chromosome carrying the major gene and the non-carrier ones. In the three cases of genetic determinism studied, and whatever the selection method, the observed deviation ranged from the situation where the selected individuals were very few ($N = 192$, $p = 6.25\%$), with a polygenic gain lower on the chromosome carrying the major gene, to the situation where they were the most numerous ($N = 480$, $p = 25\%$), with a polygenic gain higher on the chromosome carrying the major gene (see table VII). The first part of the study had concluded that there was a slight negative influence of the major gene on the flanking QTLs, due to hitch-hiking phenomena during the fixation of the major gene. This effect seems to be more pronounced when selection intensity is higher and population size is smaller. On the contrary, in a large population rather weakly selected, the selection pressure applied to the chromosome carrying the major gene is lower with respect to non-carrier ones. The genetic variability conserved on this chromosome during the generations taken for the fixation of the major gene is higher and leads to a higher gain on this chromosome between the fixation and generation 30. But once again, the observed differences were not significant.

This phenomenon should vary with the recombination rate (r). To test this hypothesis, the change of the observed deviation in polygenic gain between the chromosome carrying the major gene and the non-carrier chromosomes with the recombination rate was studied (see fig 4). The behaviour of the difference was clearly sigmoïd, with positive values (higher polygenic gain on the chromosome carrying the major gene) for intermediate recombination rates (0.01–0.2), and negative values elsewhere. During the first generations before A fixation, very little selection pressure was put on the QTLs linked to the major locus, with all chromosomes carrying the A allele being selected. For very small recombination rates, the hitch-hiking effect explains the loss of favourable QTLs linked to the B allele. For intermediate distances, recombinations occur and all the alleles at these QTLs are kept, without selection. When the recombination rate is higher than 20–30%, selection is possible on these QTLs but, due to residual linkage with the major locus, itself submitted to a strong selection pressure, this selection affects only a subsample of the available alleles at the QTLs, thus reducing the gain. Finally,

F Fournet et al

Table VII. Observed deviation, at generation 30, between polygenic genetic gain on the chromosome carrying the major gene and non-carrier chromosomes in the phenotypic method S_P and genotypic method S_G (expressed in percentage of the total genetic gain over 30 generations); 10 QTLs, $G_{AA}-G_{BB} = 2\sigma_P$, 50% females selected.

Initial frequency of A Mode of inheritance	Selection method	N = 192 p = 25%	N = 192 p = 6.25%	N = 480 p = 25%	N = 480 p = 6.25%
0.1	S_P	−0.6%	−2.7%	5.1%	−0.5%
A recessive	S_G	0.4%	−6.0%	3.4%	0.2%
0.5	S_P	2.3%	−0.8%	4.6%	3.6%
A recessive	S_G	1.6%	−2.0%	4.8%	3.5%
0.1	S_P	0.1%	−1.5%	3.5%	−0.3%
A additive	S_G	0.4%	−6.0%	3.4%	0.2%

for $r = 0.5$, the QTLs and the major gene were independent and no difference in polygenic gain was found between the chromosome carrying the major gene and the non-carrier ones, as expected.
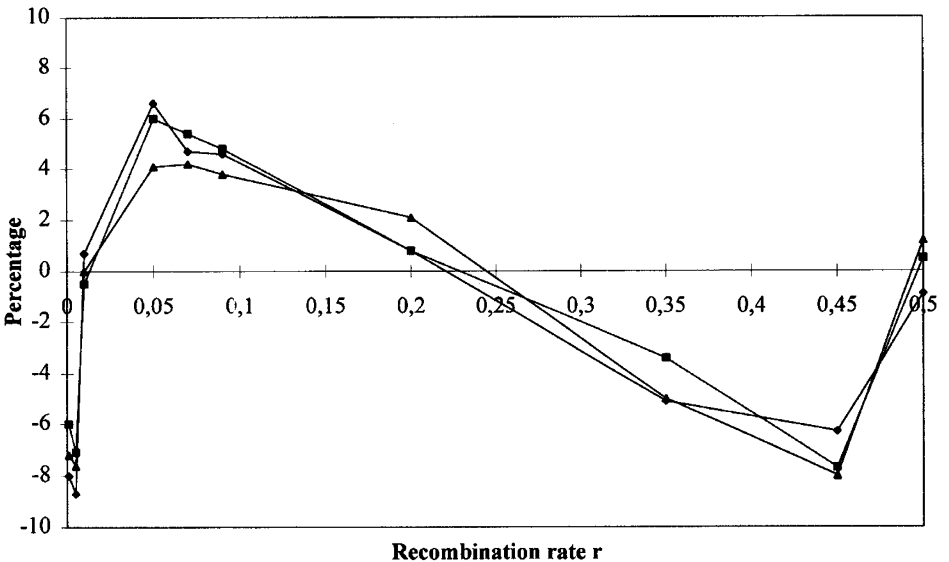

Recombination rate r

Fig 4. Change in the deviation in polygenic gain between the chromosome carrying the major gene and the non-carrier ones (expressed in percentage of the gain on non-carriers) with the recombination rate between the major gene and the 2 QTLs on the same chromosome (total of 10 QTLs), in three selection methods. ♦ S_P, ■ S_G, ♠ S_C.

DISCUSSION AND CONCLUSION

Our objective was to study the value of a new type of modelling in the evaluation of selection methods including major gene information and to describe in which conditions the information on a major gene was useful and how it could bring a higher response. The use of a stochastic model describing polygenic inheritance by a finite number of QTLs, and locating the major gene on a distinct chromosome, allowed an original approach to the subject. To achieve this study, hypotheses were made concerning the selected trait (measured in both sexes), the accuracy of the knowledge on the major genotype (supposed known without error) and the selection method used (one-stage selection based on individual performances). The results presented above are valid under these assumptions, and possible changes of the main conclusions when departing from these hypotheses are then to be discussed.

The first conclusion of this study is that the value of including information on the individual genotypes at a major locus in a one-stage selection scheme, when the selected trait is measured in both sexes, remains very small in 95% of the cases studied. However, and that is the second conclusion to be underlined in this work, the inclusion of major gene information can provide extra-gain in the medium and long term when the favourable allele at the major locus is rare and recessive and in the medium term when it is rare and additive. This is in good agreement with the results of Larzul et al (1997). They also found a higher superiority of the combined method over phenotypic selection for the recessive case with $fq(A)$ of 0.1, with values of 120, 80 and 20% for a decreasing effect of major gene and h^2 0.25. These values are to be compared with 71, 28 and 5% in our model (results not presented), where the size of the population is reduced (thus diminishing the initial available genetic variability) and the selection intensity is lower (the variability is thus less intensively exploited). In the two models, the differences were clearly found to be lower for A additive and even more for A dominant. Taking account of the major gene information essentially leads to a faster fixation of the favourable allele, at the expense of a genetic lag in the selection of flanking QTLs but without loss of polygenic variance. These results partially agree with Gibson (1994) for the additive case in the long term situation, where the phenotypic selection method is always superior to other selection methods.

Our study also shows that, although the number of loci used to simulate polygenic inheritance is below the threshold required to represent the infinitesimal model (1600 QTLs for De Boer and Van Arendonk, 1995, or 1000 QTLs for Fournet and Elsen, 1996), no large differences are found between the results of the two simulated genomes (10 and 100 QTLs).

For a rare recessive favourable allele at the major locus, the value of the genotypic selection method is also to diminish the risk of losing the favourable A allele by genetic drift. The second part of the study provided evidence of this risk in a small population strongly selected, and the rapid fixation of the A allele enabled by the genotypic method prevented this risk.

As a general conclusion, excluding this case of a recessive favourable allele, the value of including major gene information is small. This result was obtained in the case of individual selection on a trait expressed and measurable on the male and female candidates.

(1) The efficiency of the methods studied, as an alternative to a standard selection programme, might be higher when selected traits are not expressed in both sexes (milk yield for dairy bulls) or not measurable in the live animals (carcass quality in pigs), as shown in recent studies concerning marker-assisted selection (Kashi et al, 1990; Meuwissen and Van Arendonk, 1992; Brascamp et al, 1993; Meuwissen and Goddard, 1996; Ruane and Colleau, 1996).

(2) The accuracy of the information on major genotypes might also temperate the conclusions of this study. Here we have assumed the major genotype to be known without error. If the major genotype had been estimated given marker information, the extra-gain due to the inclusion of this information would have been less important, due to the risk of error in the estimation.

(3) The nature of the selection schemes used for the comparison may modify the strength of our conclusions. Gibson (1994) and Larzul et al (1997) show that the effect of including major gene information is lower in the case of progeny test than in individual phenotypic selection. This may be due to a better knowledge of genetic value when including information on relatives. Larzul et al (1997) proposed this explanation when comparing the responses obtained with selection on own performance (scheme I) and with selection on progeny test (scheme II): the extra-gain due to the inclusion of major gene information is higher in scheme I than in scheme II.

Gibson (1994) also compared these two schemes, and the differences, although in disfavour of the inclusion of major gene, were less important when selecting on progeny test than on individual performance. These two examples allow us to conclude that the advantage obtained in our work in favour of the inclusion of major gene information would have been less striking if information on relatives had included in the selection criterion, in the case of a progeny test scheme for example, or when increasing the quantity of information used for the estimation of the breeding values, as in the schemes suggested by Kinghorn et al (1993) and Janss et al (1995).

To complete these conclusions, and in order to evaluate how the cumulated discounted gain obtained in the genotypic selection method, for the most favourable case (ie, a rare and recessive favourable allele at the major locus, considered for six generations of selection), would change when compared to a more precise evaluation method (eg, animal model BLUP), the simulation of a selection method based on true genetic values was performed. So, a new selection method was constructed on the basis of the combined selection method (described in *Selection methods*), but in which the expected genetic values in the offspring were now calculated from the true polygenic genetic value of the parents and from the expected genetic value at the major locus depending on the parents' genotypes. The comparison between the genotypic and the new combined selection methods, when based on the true genetic values in the favourable case of a rare and recessive allele of large effect at the major locus, showed an extra-gain of 48.0% in favour of the genotypic method relative to the new combined method, against 56.4% in the previous comparison with the combined method based on the individual phenotypes. This difference of about 16 percentage points of extra-gain between the two situations constitutes the upper bound of the decrease in extra-gain of the

method S_G obtained when using an animal model. We can then consider that our results would still hold in this situation of genetic evaluation.

Finally, the intermediate ranking of the combined selection may be surprising. This ranking was probably due to the fact that in this method the candidates were not selected on the presence of the major gene but on its effects. A dynamic selection method, ie, the choice of a varying selection strategy, depending on the frequency of the major gene throughout the generations of selection, might be the optimal solution.

REFERENCES

Brascamp EW, Van Arendonk JAM, Groen AF (1993) Economic appraisal of the utilisation of genetic markers in dairy cattle breeding. *J Dairy Sci* 76, 1204–1213

De Boer IJM, Van Arendonk JAM (1992) Prediction of addditive and dominance effects in selected or unselected populations with inbreeding. *Theor Appl Genet* 84, 451–459

Cunningham EP, Ryan J (1975) A note on the effect of the discounting rate and length of the accounting period on the economic value of genetic improvement in cattle populations. *Anim Prod* 21, 77–80

Danell O, Rönningen K, Ström H, Andersson K, Sundgren PE (1976) An extension of the discounted gene flow method with example in pig breeding. *Acta Agriculturae Scandinavica* 26, 203–210

Falconer DS (1981) *An Introduction to Quantitative Genetics.* Longman, London, 2nd ed

Fournet F, Hospital F, Elsen JM (1995) A FORTRAN program to simulate the evolution of genetic variability in a small population. *Comput Applic Biosci* 11, 469–475.

Fournet F, Elsen JM (1996) Numerical comparison of three models for the description of genetic variability in small population undergoing selection. *Genet Sel Evol* (submitted)

Gibson JP (1994) Short-term gain at the expense of long-term response with selection of identified loci. In: *Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, Guelph, Ontario, 7–12 August*, 21, 201-204

Hill WG (1974) Prediction and evaluation of response to selection with overlapping generations. *Anim Prod* 18, 117–139

Hinks CJM (1970) The selection of dairy bulls for artificial insemination. *Anim Prod* 12, 569–576

Hoeshele I (1990) Potential gain from the insertion of major genes into dairy cattle. *J Dairy Sci* 73, 2601–2618

Hospital F (1992) Effets de la liaison génique et des effectifs finis sur la variabilité des caractères quantitatifs sous sélection. Thèse, Université Montpellier II

Janss LLG, Thompson R, Van Arendonk JAM (1995) Application of Gibbs sampling for inference in a mixed major gene-polygenic inheritance in animal populations. *Theor Appl Genet* 91, 1137–1147

Kashi Y, Hallerman E, Soller M (1990) Marker-assisted selection of candidate bulls for progeny testing programmes. *Anim Prod* 51, 63–74

Kennedy BW, Verrinder Gibbins AM, Gibson JP, Smith C (1990) Coalescence of molecular and quantitative genetics for livestock improvement. *J Dairy Sci* 73, 2619–1627

Kinghorn BP, Kennedy BW, Smith C (1993) A method of screening for genes of major effect. *Genetics* 134, 351–360

Larzul C, Manfredi E, Elsen JM (1997) Potential gain from including major gene information in breeding value estimation. *Genet Sel Evol* (accepted)

Le Roy P, Naveau J, Elsen JM, Sellier P (1990) Evidence for a new major gene influencing meat quality in pigs. *Genet Res, Camb* 55, 33–40

McLaren DG, Fernando RL, Lewin HA, Schook LB (1990) Integrated strategies and methodologies for the genetic improvement of animals. *J Dairy Sci* 73, 2647-2656

Mérat P, Ricard FH (1974) Etude d'un gène de nanisme lié au sexe chez la poule: importance de l'état d'engraissement et gain de poids chez l'adulte. *Ann Génét Sél Anim* 6, 211–217

Meuwissen THE, Van Arendonk JAM (1992) Potential improvements in rate of genetic gain from marker-assisted selection in dairy cattle breeding schemes. *J Dairy Sci* 75, 1651–1659

Meuwissen THE, Goddard ME (1996) The use of marker haplotypes in animal breeding schemes. *Genet Sel Evol* 28, 161–176

Miller RH, Pearson RE (1979) Economic aspects of selection. *Anim Breed Abstr* 47, 281–290

Ollivier L (1980) Le déterminisme génétique de l'hypertrophie musculaire chez le Porc. *Ann Génét Sél Anim* 12, 383–394

Piper LR, Bindon BM (1982) The *Booroola Merino* and the performance of medium non-Peppin crosses at Armidale. In: *The Booroola Merino* (LP Piper, BM Bindon, RD Nethery, ed), CSIRO, Melbourne

Poutous M, Vissac B (1962) Recherche théorique des conditions de rentabilité maximum de l'épreuve de descendance des taureaux d'insémination artificielle. *Ann Zootech* 11, 233–000

Ruane J, Colleau JJ (1995) Marker assisted selection for genetic improvement of animal populations when a single QTL is marked. *Genet Res* 66, 71-83

Ruane J, Colleau JJ (1996) Marker-assisted selection for a sex-limited character in a nucleus breeding population. *J Dairy Sci* (submitted)

Sehested E, Mao IL (1992) Responses to selection on genotypic and phenotypic values in the presence of genes with major effects. *Theor Appl Genet* 85, 403–406

Smith C (1967) Improvement of metric traits through specific genetic loci. *Anim Prod* 9, 349–358

Smith C (1977) Estimation of the value of genetic improvement in the presence of inflation and risk. In: *28th Annual Meeting of the EAAP, Brussels, 22–25 August*

Smith C (1982) Estimates of trends in the Halothane gene in pigs stocks with selection. *Z Tierzüchtg Züchtgsbiol* 99, 232–240.

Smith C, Webb AJ (1981) Effects of major genes on animal breeding strategies. *Z Tierzüchtg Züchtgsbiol* 98, 161–169

Soller M (1978) The use of loci associated with quantitative effects in dairy cattle improvement. *Anim Prod* 27, 133–139

Soller M, Bar-Anan R, Pasternak H (1966) Selection of dairy cattle for growth rate and milk production. *Anim Prod* 8, 109–119

Stam P (1986) The use of marker loci in selection for quantitative characters. In: *Exploiting New Technologies in Animal Breeding. Genetic Developments*. Oxford University Press, Oxford, 170–182

Tanksley SD (1993). Mapping polygenes. *Annu Rev Genet* 27, 205–233

Whittaker JC, Curnow RN, Haley CS, Thompson R (1995) Using marker maps in marker assisted selection. *Genet Res* 66, 255–265