



HAL
open science

Estimation by simulation of the efficiency of the French marker-assisted selection program in dairy cattle (Open Access publication)

François Guillaume, Sébastien Fritz, Didier Boichard, Tom Druet

► To cite this version:

François Guillaume, Sébastien Fritz, Didier Boichard, Tom Druet. Estimation by simulation of the efficiency of the French marker-assisted selection program in dairy cattle (Open Access publication). *Genetics Selection Evolution*, 2008, 40 (1), pp.91-102. hal-00894611

HAL Id: hal-00894611

<https://hal.science/hal-00894611>

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.

Estimation by simulation of the efficiency of the French marker-assisted selection program in dairy cattle (*Open Access publication*)

François GUILLAUME^{1,2*}, Sébastien FRITZ³, Didier BOICHARD¹,
Tom DRUET¹

¹ INRA, UR337 Station de génétique quantitative et appliquée, 78350 Jouy-en-Josas, France

² Institut de l'élevage, 149 rue de Bercy, 75595 Paris Cedex 12, France

³ Union nationale des coopératives d'élevage et d'insémination animale, 149 rue de Bercy,
75595 Paris Cedex 12, France

(Received 19 January 2007; accepted 3 September 2007)

Abstract – The efficiency of the French marker-assisted selection (MAS) was estimated by a simulation study. The data files of two different time periods were used: April 2004 and 2006. The simulation method used the structure of the existing French MAS: same pedigree, same marker genotypes and same animals with records. The program simulated breeding values and new records based on this existing structure and knowledge on the QTL used in MAS (variance and frequency). Reliabilities of genetic values of young animals (less than one year old) obtained with and without marker information were compared to assess the efficiency of MAS for evaluation of milk, fat and protein yields and fat and protein contents. Mean gains of reliability ranged from 0.015 to 0.094 and from 0.038 to 0.114 in 2004 and 2006, respectively. The larger number of animals genotyped and the use of a new set of genetic markers can explain the improvement of MAS reliability from 2004 to 2006. This improvement was also observed by analysis of information content for young candidates. The gain of MAS reliability with respect to classical selection was larger for sons of sires with genotyped progeny daughters with records. Finally, it was shown that when superiority of MAS over classical selection was estimated with daughter yield deviations obtained after progeny test instead of true breeding values, the gain was underestimated.

marker-assisted selection / simulation / efficiency / dairy cattle

* Corresponding author: francois.guillaume@jouy.inra.fr

1. INTRODUCTION

Marker-assisted selection (MAS) is expected to be particularly valuable for dairy cattle breeding [2,6]. Indeed, several conditions in which MAS improves the efficiency of classical selection are met: most traits of interest are sex-limited, generation interval is long and progeny-test is a long and costly step. Furthermore, MAS can increase the reliability of breeding values [7]. This would be particularly beneficial for bull dams, which are often selected on pedigree information only [2] or for functional traits, with a low heritability, that are gaining emphasis in breeding goals. Therefore, since the end of 2000, a MAS program has been implemented in France. Breeding companies joined this program in order to improve their selection efficiency. However, since MAS programs are recent and relatively rare, little is known about their efficiency. Indeed, the progeny testing step is relatively long and a comparison of breeding values predicted by MAS before and after progeny testing can be done only more than four years after first MAS predictions. In addition, the number of progeny tested bulls remains limited to estimate MAS efficiency and to draw conclusions. Finally, the true breeding values are unknown and this adds some sampling error. Simulation studies offer the possibility to increase the number of animals and to repeat the analysis, to know the true breeding values and to have direct answers. Different simulation studies [6, 8, 12] have already proven the efficiency of MAS for predicting breeding values. However, simulation studies are often based on simple hypotheses. Thanks to the information accumulated in the French MAS program since 2000, it is now possible to make more realistic assumptions regarding the population structure, the marker informativity, the number of genotyped animals, the number of animals with records and the precision of these records, *etc.* Variances of the QTL used are also better known because they have been estimated recently on a large data sample [3]. The objective of this study was to estimate by simulation the efficiency of the French MAS evaluation for two different time periods.

2. MATERIAL AND METHODS

2.1. French MAS data

Data sets used for French MAS evaluation of April 2004 and 2006 were used in this study. Two different time periods were studied to observe the evolution of the efficiency of MAS. Indeed, the efficiency of MAS should be improved in 2006 because more families were genotyped, dams of young animals were more often genotyped and some new microsatellite markers were used.

Table I. Population structure of the French MAS program in April 2004 and 2006.

	April 2004	April 2006
Animals in pedigree	34 318	55 336
Animals with records	23 137	38 859
Genotyped animals	16 629	27 551
Male candidates	1180	1689
Sires ¹	72	79
Dams ¹	793	1130
Genotyped dams ¹	486	887
Maternal grand-sires ¹	70	114
Sires ¹ with more than 20 genotyped progeny daughters	11	12
Progeny tested sire families with 30 sons or more	47	64

¹ Of male candidates.

Three files were used at each evaluation: the pedigree file, the markers file containing the probabilities of transmission for each QTL and the data file.

The pedigree used in the French MAS includes different types of animals. First, candidates are young males or females aged from 1 month to 1 year of age. These animals can be chosen to be parents in the next generation. Males can be selected for progeny testing while females can be used as bull dams. The purpose of MAS is to improve the prediction of breeding values of these candidates, which are therefore genotyped. It is also advised to genotype dams of candidates in order to follow QTL transmission as accurately as possible. Families of progeny tested bulls or groups of progeny daughters were genotyped in order to estimate QTL effects of old bulls or younger bulls (sire of candidates), respectively. Thanks to the genotyped animals, the genotypes of some other animals (*e.g.* sires) were reconstructed. In addition, the pedigree file contained parents over two generations of all these animals. Table I indicates the number of candidates (with their sires and dams), genotyped dams, number of genotyped progeny tested bulls or progeny daughter families.

Animals were genotyped for 43 and 45 microsatellite markers before and after first of January 2005, respectively. These markers are used to follow the transmission of 14 QTL regions [1]. Seven of these QTL affecting milk production or composition traits were used in this study. Two to five microsatellite markers are available for each QTL. These were used to estimate probability of identity-by-descent (pid) matrices using a method similar to that of Wang *et al.* [15] extended to the use of multiple markers as in Pong-Wong *et al.* [10].

Finally, phenotypic records were twice the daughter yield deviations (DYD) for males and yield deviations (YD) for females computed for milk, fat and protein yields and fat and protein percentages, pooled from the first three lactations jointly as in VanRaden and Wiggans [14]. These records were obtained from the official genetic evaluation of April 2004 [11]. Respective weights were estimated as in VanRaden and Wiggans [14] with a correction for the number of cows in each herd. DYD of sires were obtained by using only records of daughters not included in the pedigree file. These phenotypic records were replaced by simulation.

2.2. Simulation

The pedigree file and the file containing pid were exactly the same as in the real MAS program. The structure of the performance file was also kept: the same animals had records and the weights of the records were conserved. Only the records were simulated with the following method. The genetic effect of animal i is computed as

$$g_i = u_i + \sum_{j=1}^{n_qtl} (v_{ij1} + v_{ij2})$$

where u_i is the polygenic effect of individual i (excluding QTL effects), v_{ij1} and v_{ij2} are allelic effects at QTL j for the paternal and maternal alleles, respectively, and n_qtl is the number of QTL.

For animals without parents, the polygenic effect was sampled from $N(0, \sigma_u^2)$ while for animals with parents, the polygenic effect was equal to the sum of the mean polygenic effects of the parents and the Mendelian sampling drawn from a normal distribution with the variance adjusted for number of known parents. The polygenic variance (σ_u^2) was defined according to the heritability of the traits and the proportion of genetic variance explained by QTL (Tab. II).

For each QTL j , a biallelic gene with substitution effect α_j was simulated. The estimated percentage of heterozygous sires in the population was used to approximate the allelic frequency in the population. The substitution effect was derived from the simulated QTL variance and the allelic frequencies. The variances used for each QTL for each trait are presented in Table II. These were obtained from Druet *et al.* [3] and from our knowledge of these QTL. For all founder animals, QTL alleles were sampled thanks to the allelic frequencies. Then, the alleles were transmitted to the entire population using the estimated pid. By definition, the pid gives the probability for an offspring to receive the

Table II. Proportion of genetic variance used to simulate QTL effects for dairy traits and polygenic effect (in %).

Trait	Number of the chromosome on which the QTL is located							Polygenic effect	Heritability of the traits
	3	6	7	14	19	20	26		
Milk yield	0	0	5	15	0	10	10	60	0.30
Fat yield	0	5	5	15	5	0	20	50	0.30
Protein yield	0	5	5	10	5	0	10	65	0.30
Fat content	0	5	5	40	0	10	0	40	0.50
Protein content	10	15	0	10	0	15	0	50	0.50

paternal or the maternal allele from its parent. Therefore, these probabilities were used to simulate which QTL allele an offspring had received from its parent. For instance, if the pid was equal to 0.5, the progeny had equal chances to receive the paternal or the maternal allele of its parent while if the paternal pid was equal to 1 then the progeny received the paternal allele of the corresponding parent.

To simulate records, a residual value was sampled from $N(0, \sigma_e^2)$ where the residual variance is adjusted by the weight from actual phenotypes in the MAS data set. The simulated records were the sum of the genetic and residual values. Additionally, for male candidates, records were simulated with a weight corresponding to the first EBV obtained after progeny testing.

Simulations were repeated 100 times for each trait and both time periods.

2.3. MAS evaluation

The model used in this study was a single trait and multi-QTL model as proposed by Fernando and Grossman [4]:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \sum_{i=1}^{n_{\text{qtl}}} \mathbf{Z}_{v_i} \mathbf{v}_i + \mathbf{e} \quad (1)$$

where \mathbf{y} is a vector containing records, $\boldsymbol{\beta}$ is a vector of fixed effects (the mean), \mathbf{u} is a vector of random polygenic effects, \mathbf{v}_i is a vector of random gametic effects for QTL i and \mathbf{e} is a vector of random residual terms. \mathbf{X} , \mathbf{Z} and \mathbf{Z}_{v_i} are known design matrices that relate records to fixed, random polygenic and gametic effects, respectively.

Four to five QTL were used for each production trait and the variance components (see Tab. II) were assessed based on a previous study [3].

Table III. Mean information content measured as $|1-2p|$ weighted by QTL variance for each trait for 2004 and 2006 candidates and their parents.

Traits	Mean information content weighted by QTL variance					
	Candidates		Sires		Dams	
	2004	2006	2004	2006	2004	2006
Milk	0.61	0.72	0.36	0.47	0.44	0.54
Fat yield	0.58	0.72	0.33	0.48	0.41	0.54
Protein yield	0.60	0.71	0.34	0.48	0.43	0.55
Fat content	0.70	0.76	0.40	0.49	0.50	0.56
Protein content	0.72	0.75	0.41	0.53	0.52	0.60

3. RESULTS

3.1. Simulated data

The results were obtained for two different sets of candidates (Tab. I). They included males born during the previous AI season, *i.e.* from October to September. The first set was constituted of candidates of year 2004 whereas the second set of candidates of year 2006. Informativity was estimated as $|1 - 2p|$ where p was the probability transmission of a given paternal or maternal QTL allele [2]. When the transmitted allele is known, p is equal to 0 or 1 and $|1 - 2p|$ is one while when there is no information on which allele was transmitted, p is equal to 0.5 and $|1 - 2p|$ is zero. So this information content indicates how well the QTL transmission is followed in the population. For each trait, mean information content was computed by weighting the information content of each QTL by the proportion of genetic variance explained by this QTL. This weighted mean information content is presented in Table III for candidates of years 2004 and 2006 and for their sires and dams. For all the traits, information content increased in 2006 with respect to 2004: for candidates, mean information content gains ranged from +0.03 up to +0.14 while for sires they ranged from +0.09 up to +0.15. The gains were comprised between +0.06 and +0.13 for dams.

3.2. Estimation model

Marker-assisted selection was compared to classical selection (model with only a polygenic effect). Accuracies of breeding values (squared correlation R^2 between estimated and true genetic effects) were estimated and are presented in Table IV. For all traits, MASEBV were more reliable than classical EBV.

Table IV. Reliabilities (R^2) of classical polygenic EBV (POL) and MAS EBV (MAS) for male candidates from 2004 and 2006.

Trait	April 2004			April 2006		
	POL	MAS	Difference	POL	MAS	Difference
Milk yield	0.294	0.327	+0.033	0.313	0.361	+0.048
Fat yield	0.281	0.296	+0.015	0.310	0.373	+0.063
Protein yield	0.254	0.273	+0.019	0.303	0.341	+0.038
Fat content	0.313	0.407	+0.094	0.342	0.453	+0.111
Protein content	0.214	0.301	+0.087	0.342	0.418	+0.076

Table V. Reliabilities of classical polygenic EBV (POL) or marker-assisted EBV (MAS) of candidates of 2004, depending on the status of their sires.

Trait	Sires of candidates without genotyped progeny daughters			Sires of candidates with genotyped progeny daughters		
	POL	MAS	Difference	POL	MAS	Difference
Milk yield	0.266	0.302	+0.036	0.291	0.353	+0.062
Fat yield	0.255	0.263	+0.008	0.277	0.312	+0.035
Protein yield	0.243	0.265	+0.022	0.267	0.307	+0.040
Fat content	0.269	0.384	+0.115	0.304	0.476	+0.172
Protein content	0.200	0.301	+0.101	0.210	0.372	+0.162

In 2004, the gain of reliability ranged from 0.015 for fat yield up to 0.094 for fat content. Gain was relatively limited for yield traits (0.033, 0.015 and 0.019 for milk, fat, and protein yields, respectively) and larger for content traits (0.094 and 0.087 for fat and protein contents, respectively). In 2006, the difference between MAS EBV and classical EBV was larger, especially for yield traits (0.048, 0.063 and 0.038 for milk, fat and protein yields, respectively). Among all 100 replications for 2004, MAS was less efficient than classical selection for eleven and nine replications for fat and protein yields, respectively. In 2006, MAS resulted in lower reliabilities for a single replication for milk yield. For these few negative results, the difference between evaluation methods was close to zero.

In 2004, MAS and classical EBV were also compared with respect to the amount of information available to estimate gametic effects of the sires (Tab. V). Two classes of sires were defined: sires with or without genotyped progeny daughters (at least 20). The improvement of accuracy due to MAS is larger for all traits when a group of progeny daughters is also genotyped. The difference between MAS selection and classical selection when sires of candidates have no genotyped progeny represent only 59, 23, 55, 67 and 63%

Table VI. Correlations between classical polygenic EBV (POL) or marker-assisted EBV (MAS) and simulated DYD of candidates of 2004.

Trait	POL	SAM	Difference			
			Mean	St. Dev.	Minimum	Maximum
Milk yield	0.476	0.502	+0.026	0.013	-0.002	0.074
Fat yield	0.466	0.477	+0.011	0.018	-0.041	0.051
Protein yield	0.441	0.457	+0.016	0.014	-0.021	0.056
Fat content	0.526	0.599	+0.073	0.022	0.022	0.135
Protein content	0.522	0.572	+0.052	0.020	0.013	0.097

of the difference obtained when sires have genotyped progeny for milk, fat and protein yields and fat and protein contents, respectively.

Finally, the comparisons between MAS and classical EBV with simulated DYD (with an accuracy corresponding to first EBV after progeny testing) are shown in Table VI. As expected, MAS EBV are better predictors but the difference between MAS and classical selection varies across replications. The mean correlation gain is equal to 0.026, 0.011, 0.016, 0.073 and 0.052 for milk, fat and protein yield and fat and protein content, respectively. These gains are lower than when comparison is done with true genetic values (comparison on an accuracy scale). The minimum and maximal gains ranged from -0.002 to 0.074, -0.041 to 0.051, -0.021 to 0.056, 0.022 to 0.135 and from 0.013 to 0.097 for milk, fat and protein yields and fat and protein contents, respectively. For some samples, MAS appeared to perform worse than the classical model for fat or protein yield.

4. DISCUSSION

Files involved in the French MAS are increasing on a regular basis as a consequence of continuous addition of new genotyped animals (see Tab. I). Therefore, the MAS evaluation is more demanding in computational terms but the information on QTL is increasing with time. More families are genotyped and QTL transmission is better observed. Both these information improve the estimation of QTL effects and therefore the efficiency of MAS. The increment of genotyped animals is not only due to the continuous application of the MAS program but also to strategic choices decided to improve the French MAS program. For instance, breeding companies genotype dams of candidates more frequently than at the start of the MAS program. At the beginning, neither the dams of sire nor the progeny daughter families were genotyped. During the MAS program, breeding companies were advised to genotype these animals.

The impact of all these decisions is visible in Table III where increasing information can be noted. Some technical changes were also implemented to improve the efficiency of MAS. Some microsatellite markers are no longer used while some more informative markers were integrated in the program. All these elements improved the efficiency of MAS to follow QTL transmission in the population (see Tab. III). The changes in precision of the pid between 2004 and 2006 are important and are consequences of efforts made by breeding companies. Efficiency of MAS can still be improved by the use of denser markers. For instance, if informativity is increased by replacing the microsatellite markers by ten SNP close to the QTL (within 1 cM), the gain of reliability of MAS with respect to classical selection is increased from 43% up to 79% (data not shown). As shown, the gain of efficiency achieved by improving the accuracy of the pid is important but to obtain even larger gains, other MAS strategies must be applied (such as the use of linkage disequilibrium).

Some previous studies showed the advantage of MAS in predicting breeding values [12, 13]. The present study focused on accuracy gain rather than genetic progress gain achievable by MAS; in fact the latter criterion is greatly dependent on the selection strategy whereas accuracy of prediction reflects the methodology efficiency more. In the present study, many conditions were those really applied in the French breeding schemes (pedigree, markers, genotyped animals, *etc.*). Under these conditions, MAS improved the reliability of breeding values but the gain remained limited.

Accuracy improvement appeared larger for content traits than for yield traits. This can be explained by several facts. For content traits, QTL explained in general a larger part of the genetic variation. Part of genetic variance explained by the QTL has a major impact on the efficiency of MAS. Indeed, the gain of reliability achieved by MAS ranked similarly to the part of variance explained by the QTL. However, other parameters influence the efficiency of MAS. For instance, QTL variance is equal for fat yield and protein content but MAS performed better with protein content. Mean information content was higher for content traits. The influence of mean information content can also be seen when comparing the results for yield traits in 2004 and 2006. Efficiency of MAS improved clearly at constant QTL variance thanks to better mean information content. In addition, MAS is more beneficial, at constant part of genetic variance explained by the QTL, when there are fewer QTL (but with larger effects). Indeed, the polygenic model is more appropriate for a situation with many QTL (closer to the infinitesimal model) than with a few QTL. Therefore, the superiority of MAS will be reduced with many small QTL. Finally, QTL effects are estimated more accurately when QTL have larger effects

and when there is less environmental noise. However, for low heritability traits, gains of reliability of MAS are expected to be larger because there is much room for improvement since classical selection performs poorly. In the present study, efficiency of MAS was studied only for heritabilities above 0.30 and no conclusions can be drawn for low heritability traits.

The number of QTL and proportion of total genetic variance explained by them are greater than parameters usually assumed by previous simulation studies [8, 12]. This should enhance MAS efficiency, by reducing the risk that parents are homozygous at all the QTL.

On the contrary to various simulation studies [8, 13], population structure is fairly unbalanced. As shown in Table I, a few sires and maternal grandsires contribute heavily to the population. It is essential to evaluate their gametic effects as accurately as possible. Therefore, it is very important to genotype many animals such as dams and progeny daughters' families. Indeed, the results showed that when sires of candidates have genotyped progeny daughters, MAS was more efficient. This approach has some similarities with the Bottom-up scheme proposed by Mackinnon and Georges [7] and which was shown to increase MAS efficiency. Sires of candidates with genotyped progeny daughters were just a few (11 out of 72 in 2004 and 12 out of 79 in 2006) but generally contributed to a large proportion of candidates (20% in 2004 and 25% in 2006). Since the start of the French MAS, efforts have been made to increase the information available. In 2006, gains of accuracy obtained with MAS were better than in 2004. All the accumulated information improves the French MAS programs.

The study also showed that if the efficiency of MAS is assessed with field data, on DYD for instance, the estimated gain is reduced. Indeed, MAS EBV are better predictors of true genetic values. DYD still contain some errors and MAS EBV do not predict these error terms well.

Although many parameters were estimated on real data, the simulation performed in this study might depart from the underlying biological reality. Therefore, the results presented might over- or under-estimate MAS efficiency. Variance of the QTL was estimated on a large sample independent from the sample used for QTL detection. Still, the variances used might be incorrect. Therefore, the efficiency of MAS was also tested by using under- or over-estimated (by 25%) QTL variances and the differences were marginal: MAS was achieving the same gains. Allelic frequencies or effects might be wrong or the QTL could be multi-allelic. The evaluation model should be robust to these changes and the accuracy of the estimation of QTL effects should not vary much. For instance, the evaluation model does not assume a fixed number of alleles but

rather an infinite number of alleles and could easily handle a multi-allelic QTL. Although reliabilities obtained by MAS might be only slightly affected by different allelic frequencies or multi-allelic QTL, the polygenic model might be more sensitive to the changes and therefore the difference between MAS and the polygenic model might be over- or under-estimated. However, it is difficult to predict if the polygenic model would be penalised under different hypotheses. When more parents are heterozygous (due to multi-allelic QTL or changes in frequencies) and transmission of genetic values departs from the rules used with the polygenic model (half of the breeding value is transmitted), the polygenic model should achieve lower reliabilities.

This is also true when QTL effects are larger because they have a larger influence in the ranking of the animals. On the contrary, the polygenic model performs better with many QTL because this situation is closer to the infinitesimal model.

In this study, QTL were assumed additive but if some QTL had non-additive effects (dominance, epistasis), the impact on the results would be larger since the model would be less robust to it.

Finally, MAS will certainly evolve in the future towards more efficient models using denser marker maps (*e.g.*, Meuwissen *et al.*, [9]) and exploiting linkage disequilibrium. For instance, Hayes *et al.* [5] presented advantages of LD-MAS. With dense maps, small haplotypes around the QTL will be in linkage disequilibrium with the QTL. Therefore, gametic effects will be estimated across families and no longer within each sire family. As a consequence, the effects will be estimated more accurately and less genotyped animals will be required to estimate these effects. In addition, follow-up of transmission of gametic effects will be more precise because information content will improve.

5. CONCLUSIONS

In the French MAS program, accuracy of breeding values of young candidates was shown to be improved thanks to the use of molecular information. The obtained gains of accuracy (in comparison with classical selection) were relatively limited and strongly dependent on the accumulated information in the program. By genotyping more animals (such as dams or progeny daughters of sires of candidates) or using better markers, the efficiency of this program was clearly improved.

Thanks to the development of new genotyping technologies, still improved results are expected with the use of denser marker maps and of linkage disequilibrium.

REFERENCES

- [1] Boichard D., Fritz S., Rossignol M.N., Boscher M.Y., Malafosse A., Colleau J.J., Implementation of marker-assisted selection in French dairy cattle, in: Proceedings of the 7th World Congress on Genetics Applied to Livestock Production, 18–23 August 2002, Montpellier, Communication no. 22-03.
- [2] Boichard D., Grohs C., Bourgeois F., Cerqueira F., Faugeras R., Neau A., Rupp R., Amigues Y., Boscher M.Y., Levéziel H., Detection of genes influencing economic traits in three French dairy cattle breeds, *Genet. Sel. Evol.* 35 (2003) 77–101.
- [3] Druet T., Fritz S., Boichard D., Colleau J.J., Estimation of genetic parameters for quantitative trait loci for dairy traits in the French Holstein population, *J. Dairy Sci.* 89 (2006) 4070–4076.
- [4] Fernando R.L., Grossman M., Marker assisted selection using best linear unbiased prediction, *Genet. Sel. Evol.* 21 (1989) 467–477.
- [5] Hayes B.J., Chamberlain A.J., Goddard M.E., Use of markers in linkage disequilibrium with QTL in breeding programs, in: Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, 13–18 August 2006, Belo Horizonte, Brazil, Communication no. 30-06.
- [6] Kashi Y., Hallerman E., Soller M., Marker-assisted selection of candidate bulls for progeny testing programmes, *Anim. Prod.* 51 (1990) 63–74.
- [7] Mackinnon M.J., Georges M.A.J., Marker-assisted preselection of young dairy sires prior to progeny testing, *Livest. Prod. Sci.* 54 (1998) 229–250.
- [8] Meuwissen T.H.E., Goddard M.E., The use of marker haplotypes in animal breeding schemes, *Genet. Sel. Evol.* 28 (1996) 161–176.
- [9] Meuwissen T.H.E., Hayes B.J., Goddard M.E., Prediction of total genetic value using genome-wide dense marker maps, *Genetics* 157 (2001) 1819–1829.
- [10] Pong-Wong R., George A.W., Woolliams J.A., Haley C.S., A simple and rapid method for calculating identity-by-descent matrices using multiple markers, *Genet. Sel. Evol.* 33 (2001) 453–471.
- [11] Robert-Granié C., Bonaiti B., Boichard D., Barbat A., Accounting for variance heterogeneity in French dairy cattle genetic evaluation, *Livest. Prod. Sci.* 60 (1999) 343–357.
- [12] Schrooten C., Bovenhuis H., van Arendonk J.A.M., Bijma P., Genetic progress in multiple stage dairy cattle breeding schemes using genetic markers, *J. Dairy Sci.* 88 (2005) 1569–1581.
- [13] Spelman R., Bovenhuis H., Genetic response from marker-assisted selection in an outbred population for differing marker bracket sizes and with two identified quantitative trait loci, *Genetics* 148 (1998) 1389–1396.
- [14] VanRaden P.M., Wiggans G.R., Derivation, calculation, and use of national Animal Model Information, *J. Dairy Sci.* 74 (1991) 2737–2746.
- [15] Wang T., Fernando R.L., van der Beek S., Grossman M., van Arendonk J.A.M., Covariance between relatives for a marked quantitative trait locus, *Genet. Sel. Evol.* 27 (1995) 251–274.