Efficiency of Marker Assisted Selection(MAS) over The Phenotypic Selection for Economic Traits in Economic Animals

Gwang-Joo Jeon

Department of Genomic Engineering, Genomic Informatics Center, Hankyong National University, Ansung City, Kyeonggi-Do, Korea

경제동물의 주요 경제형질에 대한 표지인자를 이용한 선발(MAS)의 효율성

전 광 주

한경대학교 생명공학과, 유전정보연구소

ABSTRACT

The efficiency of marker assisted selection(MAS) over conventional selection index based sorely on phenotypic records was studied by deterministic simulation model. Parameter combination of heritability and amount of genetic variation due to the markers included in the index was employed. For the index with own phenotypic information vs. the index with own phenotypic plus marker information, the relative efficiency of MAS over the selection with phenotypic records was about 38% high when heritability was low(0.05). However, when heritability was high(50%), the relative efficiency of MAS was vary low and almost negligible. For more practical situation of selection index which included information on own, sire and dam, MAS was less effective than when selection criteria was only on own performance.

(**Key words** : Marker assisted selection, Selection response)

Ⅰ. INTRODUCTION

Animal quantitative geneticists have long been working on the improvement of economically important traits. The main tools that quantitative geneticists have conventionally used are only the phenotypic information relying on the infinitely large numbers of genes affecting the traits although they never deal with "so-called black box" of DNAs that actually act on the performances. The dramatic improvement of DNA technologies have enabled animal breeders to better understand the systems of major loci affecting those economic traits. The special

"본 논문은 한경대학교 2000년도 학술연구비 지원에 의하여 수행된 연구 결과입니다."

Corresponding author : G. J. Jeon, Dept. of Genomic Engineering, Hankyong National Univ., 67 Seok-Jung Dong Ansung city, Kyonggi-Do 456-749

attention is carefully given that the polymorphic markers are not likely the gene loci or the quantitative trait loci(QTL) but most likely be the genetic markers linked to the QTL. The commonly used as genetic markers are RFLPs, minisatellites, microsatellites, and etc. The determination of QTL mapping is required prior to MAS and is achieved by various structure of populations such as half-sib mating design, back-cross design, F_2 design, granddaughter design and etc. The main advantage of utilizing the marker information is that more information is added on the traditional selection method sorely based on phenotypic records. However, as pointed out by Spelman and Garrick(1997), MAS still needs phenotypic records information by two major reasons; 1) the majority of economic traits are quantitative and affected by many genes with small effects(infinitesimal model theory) but a few number of genes with large effects. Therefore, traditional methods with a large number of animals would be needed (such as progeny testing schemes in dairy cattles) to identify those small gene effects. 2) Single(major) genes of large effects are deleterious pleiotropic effects or tightly linked to other genes with detrimental effects, which requires that the monitoring of the offspring's phenotype has to be continued(Lander and Thompson, 1990). Kashi et al.(1990) and Mackinnon and Georges(1998) studied the possible preselection of young bulls before entering the progeny testing schemes, which most benefited from the short generation intervals to accelerate the rate of genetic improvement. From the simulation study of Mackinnon and Georges (1998), MAS yielded 8% to 23% more genetic gains than the conventional progeny testing schemes. However, an antagonistic result was also found by Spelman and Garrick(1997) that MAS showed even less than 2.5% superior to the traditional progeny testing scheme in a commercial dairy population, in which the response estimated was the very upper bound of the MAS scheme assuming the complete linkage between QTL and genetic markers. More importantly, the success of MAS would be dependent on the precise identification of QTLs linked to the markers. Otherwise, the efficiency of MAS wouldn't be preferable to the conventional selection schemes. The efficiency of MAS is also determined by the amount of variation of QTL attributed to the total genetic variation of the traits. The objective of this study was to find the possible superiority of MAS over the conventional selection based on phenotypic records using simulated data by deterministic modelling. The objective of this study was to evaluate the possible selection responses from MAS compared with the conventional selection schemes that are sorely based on phenotypic records.

1. Utilization of Genetic Markers in Animal Breeding

The genetic markers useful in animal breeding strategies are involving four steps(Meuwissen and Van Arendork, 1992): 1) Identifying genetic markers, 2) Establishment of a linkage map of the markers, 3) Detection of associations between markers and QTL, and 4) Use of marker-QTL associations in the breeding program. However, step 3) and 4) are essential for MAS breeding scheme. Recently, marker and QTL associations may be found easily by multiple regression of performance data on the number of marker alleles present for all markers

(Lander and Thompson, 1990). For detection of QTL and marker association, linkage disequilibrium is necessary between markers and QTL. However, over generations, linkage disequilibria gets smaller. This strategy is useful when populations are hybridized recently, i.e., mixture of two different breeds or species, or large number of markers available, which increases the linkage disequilibria by increasing the probability of closeness between markers and QTL. For nucleus herd breeding schemes such as a MOET scheme, many full-sib families are available. Though the linkage disequilibria within family is rather loose but can be effectively used within family selection. The simple detection of possible existence of QTL could be made by a simple ANOVA method using various polymorphic markers as:

$$y_{ij} = \mu + \Sigma M_i + e_{ij} \quad [1]$$

where

y_{ij} = the phenotypic record of ith marker genotype on the jth animal;

μ = overall mean;

M_i = the effect of ith marker genotype; and,

e_{ij} = random residual associated with y_{ij}.

However, the F-test for the marker genotype for the model [1] merely gives the association of phenotypic record and marker genotype, but doesn't guarantee that the marker genotype itself is the gene(QTL) affecting the trait or the marker being linked to the QTL. The theoretical F-test in ANOVA for the null hypothesis is that the recombination θ being 0 or the marker and QTL being the same. However, single marker analysis is not useful because of the confounding effects of the linkage distance of QTL to the marker and the size of QTL. Various algorithms such as maximum likelihood method, EM algorithm, Gibbs algorithm, interval mapping using regression method, and etc have been developed. The history of utiliza- tion of genetic markers goes back to 1950's with protein and blood polymorphism as genetic markers. It's still hard to prove the poly- morphism has an effect. In general, the probability that a random polymorphism has an effect on a trait will be small. QTL can be detected only if the marker and QTL are in linkage disequilibrium. The linkage disequilibrium could be generated from several factors such as mutation, random genetic drift, selection and migration. Fortunately, in most of animal and plant species, linkage disequilibrium exists. The recent development of interval mapping using regression method is easily understood (Haley and Knott, 1992)

2. The application of marker information into breeding schemes(MAS)

Many papers have been published on MAS (Lander and Thompson, 1990; Hospital et al., 1997; Spelman and Garrick, 1997; Laurence et al., 1998; Mackinnon and Geoges, 1998; Spelman and Garrick, 1998; Wang and Hill, 2000; Wang, 2001; Moreau et al., 2000). MAS can be applied to hybridization programs and within family(or breed) selection. MAS will be most beneficial when trait-based selection is not effective. With information on markers linked to QTLs, QTL genotypes can be identified and individuals with desired genotypes can be selected. If genetic marker information is combined with phenotypic and pedigree information, the selection accuracy will be much higher. In dairy population for example, most benefits could be achieved by preselecting young

bulls before the progeny test, which will reduce the generation intervals resulting in faster genetic responses. As defined earlier, detection and estimation of QTL location and size will be used for selection of animals, which is termed as marker assisted selection(MAS). A bit confusion might be noticed that a traditional selection is based on the phenotypic records, which is direct selection from physical records only and thus, traditional selection is indirectly selecting animals by estimating genetic merits from phenotypic records, so called "indirect selection". However, MAS is a selection scheme based on DNA information, ie., QTL, but also QTL is estimated from phenotypic records, too. Therefore, MAS is also indirect selection. The efficiency of MAS depends on several factors. Many papers have been published on the efficiency of MAS over conventional selection schemes. The simplest case of MAS is when selecting for linkage disequilibrium state, where the marker is associated with the trait of interest and the linkage disequilibrium exits between the marker and QTL. for example, marker M is linked to Q(QTL), then if linkage disequilibrium exist, the MQ gametes may be more frequent than Mq, which will increase a favourable gene of Q rather than an unfavorable gene of q. For the utilization of linkage disequilibrium, the recombination ratio of θ between a marker and a QTL must be smaller than 1/4Ne, for Ne being effective population size(Lander and Thompson, 1990). And also, two different genetic groups are suggested as two parental groups, then in their progeny generation, quite a large linkage disequilibrium may occur. Nevertheless the linkage disequilibrium will be halved by each generation as given in the formula:

$$D_t = D_0 (1-\theta)^{t-1}$$

where θ = recombination rate. As θ gets less than 1/t, then more than 30% of D_0 will remain. From the paper of Lander and Thompson(1990), the number of markers to cover the genome is given :

$$2tL + C, \quad \text{or } L/\theta + C, \text{ or } \min[2tL, 8N_eL] \qquad [2]$$

where

t = at time t;

L = chromosome length in cM;

θ = recombination rate;

C = number of chromosomes in haploid form; and,

N_e = effective population size.

3. Selection Indexes utilizing the marker information

The phenotypic measurement, P of an individual can be expressed as a deviation from population mean and is comprised of two parts:

$$P = P_f + P_w$$

Where P_f = Deviation of its family mean from the population mean; and, P_w = Deviation of the individual from the family mean(within family deviation). From the above equation, three simple selection strategies are utilized in selection of domestic animals; 1) Individual selection(or mass selection), 2) Family selection, 3) Within family selection. The main differences are the selection criteria and breeding structure. The weighting factor is depending on the selection criteria. For family selection, zero weighting factor is given to P_w and for within-family selection, zero weighting factor is given to P_f.

Ⅱ. MATERIALS AND METHODS

The simple form of selection index for the efficiency of MAS over a conventional selection index using phenotypic information can be examined to see the general trend, rather than the exact one. The selection index including marker information can be set up as(Lander and Thomson, 1990) :

$$I = b_P X_P + b_m m \quad [3]$$

where

X_P = individual phenotypic record;

m = marker score; and,

b_p and b_m = index weights.

Then, the weights bs are solved by :

b = $P^{-1}Ga$, for a being economic weights, i.e., [1 0].

The index can be extended to any general situation and to solve for b;

$$b = P^{-1}Ga$$

where a = [1 0]', for marker being zero weight

$$\mathbf{P} = \begin{vmatrix} V(X_p) & Cov(m, X_P) \\ Cov(m, X_P) & V(m) \end{vmatrix}$$

$$\mathbf{G} = \begin{vmatrix} V(A) & Cov(A,m) \\ Cov(A,m) & V(m) \end{vmatrix}$$

where V(m) = R × V(A), Cov(A,m)=R × V(A), R=% of genetic variance due to the marker, and V(A)=additive genetic variance.

Then, solving for bs' ;

$$b_P = 1/2(\sigma_A^2 - \sigma_M^2)/(\sigma_P^2 - \sigma_M^2)$$

where : $1/2(\sigma_A^2 - \sigma_M^2) = 1/2(h^2\sigma_P^2 - R\sigma_A^2)$

$= 1/2(h^2\sigma_P^2 - Rh^2\sigma_P^2)$

$= 1/2h^2\sigma_P^2(1-R)$; and,

$(\sigma_P^2 - \sigma_M^2) = (\sigma_P^2 - Rh^2\sigma_P^2)$

$= \sigma_P^2(1-Rh^2)$

Thus, $b_P = 1/2(\sigma_A^2 - \sigma_M^2)/(\sigma_P^2 - \sigma_M^2)$

$= 1/2h^2\sigma_P^2(1-R)/\sigma_P^2(1-Rh^2)$

$= 1/2h^2(1-R)/(1-Rh^2)$

$$b_M = 1/2(\sigma_P^2 - \sigma_A^2)/(\sigma_P^2 - \sigma_M^2)$$

where $1/2(\sigma_P^2 - \sigma_A^2) = 1/2(\sigma_P^2 - h^2\sigma_P^2)$

$= 1/2\sigma_P^2(1-h^2)$; and,

$(\sigma_P^2 - \sigma_M^2) = \sigma_P^2(1-Rh^2)$

Thus, $b_M = 1/2(\sigma_P^2 - \sigma_A^2)/(\sigma_P^2 - \sigma_M^2)$

$= 1/2\sigma_P^2(1-h^2)/\sigma_P^2(1-Rh^2)$

$= 1/2(1-h^2)/(1-Rh^2)$

Finally, the estimated genetic response from MAS is given after some mathematical manipulation as;

$$\Delta G_{MAS} = 2iCov(G, I)/\sqrt{(Var(I)}$$

The selection response from Index using phenotypic records only(ΔG_{SI}) is given as;

$$I = b_P X_P \quad [4]$$

$$\Delta G_{SI} = ih^2\sigma_p$$

Relative efficiency(RE) of MAS over conventional breeding scheme is then expressed as:

$$RE = \Delta G_{MAS} \ / \Delta G_{SI}$$

After some numerical manipulation,

$$\Delta G_{MAS} / \Delta G_{SI} = \sqrt{\frac{p}{h^2} + \frac{(1-p)^2}{1-h^2 p}}$$

where

p = proportion of genetic variation attributed to the marker loci; and,

h^2 = heritability of the trait of interest.

The relative economic weight for marker information is given zero because no intrinsic economic relations are valid. The practical index selection criteria with use of marker information could be expressed as:

$$I = b_1x_1 + b_2x_2 + b_3x_3 + b_mm \qquad [5]$$

where

b_1, b_2, b_3, b_m = index weights;

x_1 = own phenotypic record;

x_2 = sire's phenotypic record;

x_3 = dam's phenotypic record; and,

m = own marker score.

To compute the relative efficiency of marker assisted selection over the traditional phenotypic selection, the traditional selection index without marker information was set as:

$$I = b_1x_1 + b_2x_2 + b_3x_3 \qquad [6]$$

In above equation, only marker information was not included as compared with equation [5]. For the computation of selection responses, the selection intensities of sire and dam were assumed as top 5% and top 10%, respectively.

Ⅲ. RESULTS AND DISCUSSION

The relative efficiency of MAS over the conventional selection schemes was listed in Table 1 and 2. For a simple case of model calculations(Table 1) regarding the efficiency of MAS with own marker information over the conventional selection index with own phenotypic record only, the estimated responses from the index using the marker information is more efficient when heritability is lower and amount of genetic variation accounted for by the marker included in the index was smaller. The eficiency of MAS was negligible when heritability was close to 0.5. The more practical situation was simulated and the results were presented in Table 2. Similar results were found as were in Table 1. Trait with higher heritability reduced the efficiency of MAS over conventional selection criteria. And also, as the amount

Table 1. Approximated relative efficiency of the index with marker information included(Index [3]) over the index with a phenotypic record only (Index [4])

h^2	p[1]	RE[2]
0.05	0.05	1.380
	0.10	1.677
0.10	0.05	1.186
	0.10	1.348
0.20	0.05	1.077
	0.10	1.151
0.50	0.05	1.012
	0.10	1.026

[1]p = % of genetic variance accounted for by marker loci.

[2]RE = relative efficiency of the index with marker information and the index with a phenotypic record only.

Table 2. Relative efficiency of the index with marker information included(Index [5]) over the index with phenotypic records only(index [6])

h^2	P[1]	$\triangle G_{MAS}$[2]	$\triangle G_{AI}$[3]	RE[4]
0.05	0.05	0.109	0.086	1.267
	0.10	0.128		1.349
	0.20	0.116		1.489
0.1	0.05	0.191	0.170	1.124
	0.10	0.210		1.235
	0.20	0.245		1.441
0.2	0.05	0.346	0.330	1.048
	0.10	0.362		1.097
	0.20	0.394		1.194
0.3	0.05	0.494	0.481	1.027
	0.10	0.506		1.052
	0.20	0.532		1.106
0.5	0.05	0.771	0.765	1.008
	0.10	0.776		1.014
	0.20	0.792		1.035

p[1] = % of genetic variance accounted for by marker loci.
$\triangle G_{MAS}$[2] = genetic gain from selection with marker information.
$\triangle G_{AI}$[3] = genetic gain from selection with phenotypic information only.
RE[4] = relative efficiency of the selection with marker information and the selection with phenotypic records only.

of genetic variation due to marker increased and heritability increased, the expected genetic response was small within MAS schemes(Table 2). For the future breeding schemes, MAS will be highly efficient in the traits with lower heritability such as reproductive traits.

Ⅳ. 요 약

유전표지인자를 이용한 선발(marker assisted selection; MAS)과 전통적인 표현형기록에 근거한 반응의 비교분석을 위하여 고정효과모델(deterministic model)을 이용하여 시뮬레이션을 하였다. 자신의 단일 기록을 이용한 경우와 자신의 기록과 자신의 표지인자 정보를 이용할 경우 유전력이 높을수록 MAS의 효율성이 38% 정도 높게 나타났다. 그러나 유전력이 높은 경우(50%) MAS의 효율성은 약 1%정도로서 효율성이 대단히 낮은 것으로 나타났다. 자신의 기록과 부모의 표현형 정보에 표지인자 정보를 추가할 경우 MAS의 효율성은 27% 정도였으며 마찬가지로 유전력이 높은 경우에는 효율성이 0에 가깝게 나타났다. MAS의 효율성은 유전력이 낮을수록 그리고 이용한 표지인자의 유전적 변이가 클수록 효율성이 높아지는 것으로 나타났다.

Ⅴ. ACKNOWLEDGEMENT

This study was supported by Hankyong National University Academic Research Fund of 2000.

VI. REFERENCES

1. Haley, C. S. and Knott, S. A. 1992. A simple regression method for mapping quantitative trait loci in line crosses sing flanking markers. Heredity 69:315-324.
2. Hospital, F., Moreau, L., Lacoudre, F., Charcosset, A. and Gallais, A, 1997. More on the effiency of marker-assisted selection. Theor. Appl. Genet. 95:1181.
3. Kashi, Y., Hellerman, E. and Soller, M. 1990. Marker assisted selection of candidate bulls for progeny testing programmes. Anim. Prod. 51:63.
4. Lander, R. and Thompson, R. 1990. Efficiency of marker assisted selection in the improvement of quantitative traits. genetics. 14:743-756.
5. Laurence, M., Lemarie, L. S., Charcosset, A. and Gallais, A. 2000. Economic Efficiency of One Cycle of Marker-Assisted Selection. Crop Sci. 40: 329.
6. Laurence, M., Charcosset, A., Hospital, F. and Gallais, A. 1998. Marker-Assisted Selection Efficiency in Populations of Finite Size. Genetics. 148:1353.
7. Mackinnon, M. J. and Georges, M. A. J. 1998. Maker assisted preselection of young dairy sires prior to progeny testing. Livest. Prod. Sci. 54: 229-250.
8. Meuwisen, T. H. E. and Van Arendork, J. A. M. 1992. Potential improvements in rate of genetic gain from marker assisted selection in dairy cattle breeding schemes. J. Dairy Sci. 75:1651-1659.
9. Spelman, R. J. and Garrick, D. J. 1998. Genetic and Economic Responses for Within-Family Marker-Assisted Selection in Dairy Cattle Breeding Schemes. J. Dairy Sci. 81:2942.
10. Spelman R. and Garrick, D. 1997. Utilisation of marker assisted selection in a commercial dairy cow population. Live. Prod. Sci. 47:139.
11. Wang, J. and Hill, W. G. 2000. Marker-Assisted Selection to Increase Effective Population Size by Reducing Mendelian Segregation Variance. Genetics. 154:475.
12. Wang, J. 2001. Optimal Marker-Assisted Selection to Increase the Effective Size of Small Populations. Genetics. 157:867.

(접수일자 : 2002. 9. 11 / 채택일자 : 2002. 11. 5)