

# A theoretical basis for measuring the efficiency of selection in plant breeding

KATSUEI YONEZAWA\*†, KENTARO YANO‡, TAKURO ISHII§, & TETSURO NOMURA†

† Department of Biotechnology, Kyoto Sangyo University, Kyoto 603-8555, Japan, ‡ Faculty of Agriculture, Kyoto University, Kyoto 606-8520, Japan and § Chugoku National Agricultural Experiment Station, Fukuyama 721-8514, Japan

A new index of selection efficiency, i.e. the ratio of the probability of achieving the desired genetic gain in a target population to the cost expended for that gain, is introduced to define the optimum selection procedures that give the most opportunities for producing new commercial varieties. Monte Carlo simulations of mass selection showed that the index gives solutions to some important optimization problems which cannot be solved based on the traditional indices, the expected genetic gain or its ratio to the cost. The optimum cycle and intensity of selection and optimum population size for a target population could be defined by the new index, but not by the traditional indices.

**Keywords:** optimization of selection, resource allocation, selection efficiency.

## Introduction

Selection theories for animal and plant breeding have developed with the progress of quantitative genetics theory since the 1950s, and numerous ideas and sophisticated methods for various aspects of artificial selection have been proposed from theoretical standpoints. Few of these methods, however, have been applied in practical breeding projects. Breeders are often reluctant to employ new refined methods, because they feel that the new methods are not rewarding enough to replace the traditional or conventional ones. Disagreement in the attitude towards new refined methods exists, not only between breeders and researchers, but also among researchers themselves (Quinton *et al.*, 1992; Wei, 1995). In our opinion, one of the main reasons for this disagreement is that the theoretical basis for evaluating the efficiency of selection has not adequately been integrated in the selection theory developed hitherto.

Selection methods until now have been discussed in terms of the expected genetic gain, which, for a single cycle of selection, is presented usually by an equation,  $E(R) = ih^2\sigma$ , where  $E(R)$  = expected genetic gain,  $\sigma$  = phenotypic standard deviation for the character concerned,  $i$  = standardized selection differential, and  $h^2$  = heritability which is formulated in different ways according to the selection procedures adopted

(e.g. Falconer, 1981). Efficiency of selection is not determined by the result of a single cycle of selection alone. The equation has been extended to the cases of mid- and long-term selection with a finite population size (Robertson, 1960, 1970; Wei *et al.*, 1996). In the previous discussion, a selection method that gives the largest expected genetic gain has been regarded as the most efficient.

This criterion would be valid in a situation where, as is the case in animal breeding, selection occurs to obtain the largest possible genetic gain from a target population. Selection in animal breeding usually occurs on a current herd or flock population, and any additional, if not the largest possible, improvement could produce economic profit. In this selection, no particular degree of genetic gain is imposed as a critical minimum.

Selection in plant breeding does not operate in this way. In the case of crop plants, for instance, selection occurs on target populations that have been newly produced by hybridization between breeding stock materials, with the aim of obtaining lines or varieties which are superior to the best current commercial variety (check variety). The populations and selected lines are discarded and all efforts of breeders are ruined when no lines surpassing the check variety are obtained. In this situation, what is essential to breeders is not the expected (average) genetic gain by selection, but whether or not they can obtain lines superior to the check variety. As will be discussed in this paper, optimum selection procedures are different according to which of

\* Correspondence. E-mail: yonezaw@cc.kyoto-su.ac.jp

the two criteria, the expected genetic gain or the likelihood of obtaining the desired lines, is used.

Another problem in using  $E(R)$  as a criterion of the selection efficiency is that the variable related to cost is not explicitly incorporated in the equation of  $E(R)$ , and consequently, optimum resource investment for a target population cannot be defined by  $E(R)$ . This problem is important to plant breeders who try many target populations for either a single or different breeding purposes under limited resources and facilities. The expected genetic gain per unit cost,  $E(R)/C$ , where  $C$  is the cost expended for the selection, has been used by Namkoong (1970), who discussed optimum allocation of selection intensity in two-stage selection. As will be shown later in this paper, however,  $E(R)/C$  cannot be used to define the optimum selection procedures with optimum resource investment per target population.

Selection procedures in selfing crops have often been discussed based on the criterion of achieving a sufficiently high probability of obtaining the desired genotypes (Akemine, 1958; Iyama, 1979; Jansen, 1992). Selection procedures thus defined are not always the optimum either. To satisfy this criterion with polygenic characters, an impracticably large population size and/or many generations will be needed for a target population, which severely limits the number or range of the target populations tried. A disadvantage resulting from too heavy an investment per target population must be duly considered in order to define the optimum selection procedures.

To cope with the above-mentioned difficulties of the traditional approaches, a new criterion of selection efficiency is introduced in this paper. Differences of the present from traditional criteria are examined based on Monte Carlo simulations of mass selection.

### Index to measure the efficiency of selection

A selection method can be regarded as being more efficient than others if it provides more opportunities for success at a given cost (investment). In this criterion, the efficiency of a selection method can be measured by the number of successful trials achievable at a given cost. A successful trial in this context means a trial where the desired result (genetic gain) was achieved in a target population.

The expected number of successful trials,  $T$ , is given by:

$$T = \frac{C'}{C} dS = C'd \left( \frac{S}{C} \right), \quad (1)$$

where:  $C'$  = the total cost;  $C$  = the cost per target population;  $d$  = the probability that a target population tried by the breeder is really desirable, i.e. having the potential to give the desired response to selection (the value of  $d$  would largely depend on the experience of breeders and availability of breeding materials and information); and  $S$  = the probability that the desired genetic gain is achieved in a desirable target population.

The term  $C'/C$  in eqn (1) is equal to the number of target populations tried under a given total cost  $C'$ . Both  $C'$  and  $d$  are independent of the selection method applied, and can be taken as constants in the present discussion. According to eqn (1), a selection method that gives a larger value of the term  $S/C$  can be regarded as more efficient.  $S$  and  $C$  are interrelated and formulated in different ways according to the selection method employed. Cost  $C$  may be measured in an arbitrary unit because the relative value, not the absolute value, of  $C$  is important. In the following discussion, the probability  $S$  is simply referred to as the chance of success. This terminology was once used by Nicholas (1980), who discussed the population size necessary to reduce the influence of random drift. In the terminology of Nicholas (1980), success indicates the event that the selection response of a target population falls within a short distance from the expected response, whereas success in the present paper is the event that a target population shows a selection response superior to a particular critical minimum (level of check variety).

In formulating eqn (1),  $S$  was assumed to be the same for all desirable target populations. In reality,  $S$  should have different values for different target populations. In the case where the chance of success is  $S_i$  for a fraction  $d_i$  of all target populations to be tried, the variables  $d$  and  $S$  in eqn (1) should be substituted by the total fraction  $d = \sum_i d_i$  and the average chance,  $\bar{S} = \sum_i d_i S_i / d$ , respectively. For simplicity it is assumed below that all desirable target populations have the same chance of success.

The  $T$  of eqn (1) is not the only criterion of selection efficiency. Breeders may, as is often the case in the breeding of crop plants, try two or more target populations one by one until they finally achieve the desired genetic gain in a target population. The selection efficiency in this situation may best be measured by the total cost that is expended until the success is achieved; a selection method that requires the least total cost may be taken to be the most efficient.

The total cost that is expected to be spent until the success,  $C_a$ , is formulated as:

$$\begin{aligned}
 C_a &= \sum_{i=1}^{\infty} (Ci) \{ (1-d) + d(1-S) \}^{i-1} (dS) \\
 &= CdS \sum_{i=1}^{\infty} (1-dS)^{i-1} \\
 &= \frac{1}{d} / \left( \frac{S}{C} \right), \tag{2}
 \end{aligned}$$

where *i* stands for the number of target populations tried until the success, other variables being defined as in eqn (1). Equation (2) shows that *C<sub>a</sub>* becomes smaller as the term *S/C* becomes larger, indicating that the efficiency of selection in this situation is also determined by the term *S/C*.

The efficiency of selection should be evaluated by other criteria when breeders try more than one target population for one breeding objective, not in turn but in parallel in the same period of years. The efficiency in this case depends on the probability that the desired result is achieved in at least one of these target populations. A selection method that maximizes this probability with a given total cost may be regarded as the most efficient.

With *m* populations being chosen to be tried, the above probability, *P*, is equal to one minus the probability that the selection fails in all of the *m* populations, which is formulated as:

$$P = 1 - (1 - dS)^m \approx 1 - \exp(-mdS).$$

With *C* being spent for each of the *m* target populations under a total *C'*, *m* equals *C'/C*, and then, the above equation becomes:

$$P = 1 - \exp \left\{ -C'd \left( \frac{S}{C} \right) \right\}, \tag{3}$$

which again indicates that a selection method that maximizes *S/C* is the best.

Here, we show that the chance of success per unit cost, i.e. *S/C* determines the efficiency of selection in any of the three typical situations discussed above. *S/C* may be used as a basal index to define the optimum selection procedures with optimum cost investment per target population.

The probability *d* may be substantially increased if some preselection procedures, such as collecting useful breeding stocks and/or information, are applied at certain costs before deciding the target populations (cross combinations) to be tried. When considering cost for these procedures, *C<sub>d</sub>*, a ratio:

$$\frac{dS}{C_d + C} = \frac{d}{1 + C_d/C} \cdot \frac{S}{C} \tag{4}$$

instead of *S/C* should be used as the basal index. The term *C<sub>d</sub>/C* would be negligibly small in most breeding projects. By eqn (4), the preselection procedures as costly as *C<sub>d</sub> = C* are not rewarding unless *d* is increased more than twice. Detailed discussion about the preselection part is beyond the scope of this paper. Only the selection part is discussed below based on *S/C*.

### Application for solution of some optimization problems in mass selection

Mass selection is one of the conventional selection systems that has been widely used in plant breeding. Optimization of the procedures for mass selection has been one of the core issues in the research of selection methodology. In this section, the difference of *S/C* from the traditional indices *E(R)* and *E(R)/C* is discussed with respect to some optimization issues in mass selection.

Suppose that mass selection is used for breeding of a monoecious, annual and allogamous plant, with *N* individuals being tested with a fraction *α* of them being selected per target population per year (selection cycle). Then, the cost *C* could be represented by:

$$C = At + BNt,$$

where: *A* = economic loss per year because of delay in achieving the desired genetic gain for a new variety, which may be quantified by the annual economic advantage or income that is expected to be earned by release of the desired new variety; *B* = cost per individual per year, which is expended for caring and scoring the individuals (*BN* being the cost per year per target population); and *t* = cycles (years) of selection.

Cost *A* was introduced to measure the expenditure of time (years). With more years being spent to achieve a new line, more rival candidate lines will be supplied from rival breeding companies or institutions, and then the risk of the newly obtained line losing out to the competition will become larger. Cost *A* could be taken as the economic burden of this risk. Cost *A* should be much larger than *BN* and then *C* could be approximated by *At*. In this case, the efficiency of selection is determined by the index *S/t*, because *A* can be taken as a constant.

In the above definition of *C*, only the failure of the expected benefit was considered as cost, whereas the cost *BNt* that is actually expended for selection was totally neglected. It depends on various factors, including the policy of breeders, as to which of the costs *At* and *BNt* should be considered more important. In some circumstances, only the actually expended cost *BNt*, will be important. In this case, cost *C* equals *BNt*, and the efficiency of selection is determined by index *S/(Nt)*

when the cost  $B$  is fixed, or by  $S/(Bt)$  when the population size  $N$  is fixed. Cost in most actual selection projects should comprise both  $At$  and  $BNt$ , but only two extreme cases, i.e.  $C = At$  and  $C = BNt$  (with  $B$  being fixed) are discussed in what follows.

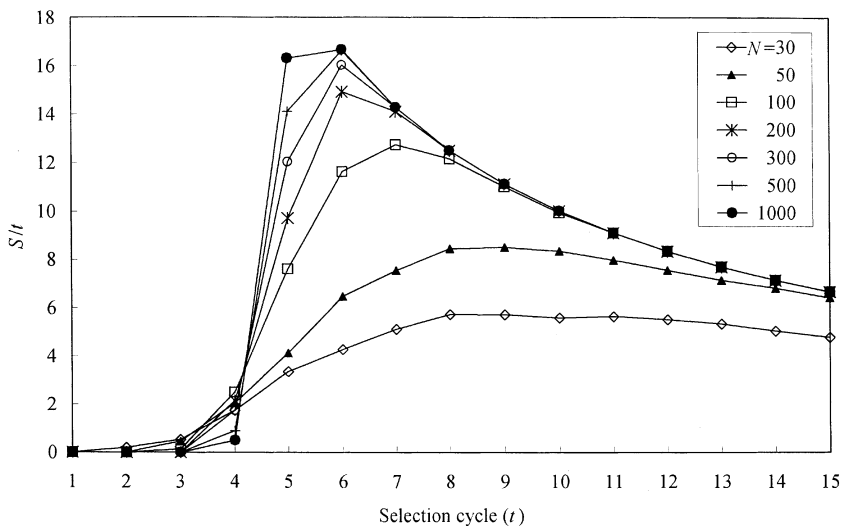
The chance of success  $S$  depends on various procedural and genetic variables, i.e. population size ( $N$ ), fraction selected in each cycle of selection ( $\alpha$ ), and selection cycles ( $t$ ) as procedural variables, and heritability ( $h^2$ ), number of loci concerned ( $L$ ), genetic effects and linkage relations of the genes as genetic variables.  $S$  may be formulated in an analytically tractable equation under the assumptions of the infinitesimal model (Hill, 1977), but would be impossible otherwise. Here, we obtained  $S$  by the relative frequency of successful runs (among 500 Monte Carlo simulation runs) that showed a genetic gain exceeding a critical minimum ( $r$ ). The expected genetic gain  $E(R)$  was calculated by the average gain of the 500 runs.

The target population of size  $N$  ( $N = 10 - 1000$ ) was initiated as an  $F_2$  population produced by crossing two inbred lines. In each run, genetic gain after  $t$  cycles ( $t = 1 - 15$ ) of mass selection was evaluated by the average genotypic value of 1000 individuals, generated by intermating between the upper  $\alpha N$  individuals ( $\alpha = 0.01 - 0.20$ ) selected in the  $t$ th cycle of mass selection. Populations of size  $N$  for further cycles of selection were generated separately from the above 1000 individuals.  $S/C$ ,  $E(R)$  and  $E(R)/C$  were calculated for arbitrarily chosen conditions  $h^2 = 0.3$  (defined for the initial population),  $r = 2$  (in units of phenotypic standard deviation of the initial population), and  $L = 20$ . Independent inheritance and additive actions of the genes, with genotypic values  $-1$  and  $1$  for the two homozygotes and  $0$  for the heterozygote at each of the  $L$  loci, were assumed.

Results of calculations of  $S/t$  with different population sizes are illustrated in Fig. 1. In this figure,  $S/t$  is maximized at six to eight cycles of selection according to decreasing population size. The selection efficiency becomes large with increasing population sizes, but is little improved with population sizes larger than 500. It is noted that, at selection cycles fewer than five, a small population may give a higher efficiency than a large population. This is because, with a small population size, random drift caused a larger fluctuation among populations (runs), so that a higher fraction of populations showed genetic advances surpassing the critical minimum  $r$ .

Figure 2 describes how selection intensity influences the selection efficiency  $S/t$ . The selection intensity in a range  $0.05 < \alpha < 0.10$  is optimum, giving the highest selection efficiency ( $S/t$ ) with the optimum selection cycles of six and seven. It is noted that, coinciding with the previous conclusion that lower selection intensities are desirable for more cycles of selection (Robertson, 1960; Hospital & Chevalet, 1993), not very high selection intensities ( $\alpha = 0.05 - 0.10$ ) are appropriate to achieve the maximum efficiency with the optimum cycles of selection, whereas a very high selection intensity ( $\alpha = 0.01$ ) is desirable when the selection is closed in two or three cycles. This trend is more prominent with a smaller population size (data not presented).

Calculations of  $S/(Nt)$  are presented in Fig. 3. Comparison of Fig. 3 with Fig. 1 shows that the rank of the different population sizes regarding the contribution to the selection efficiency is largely changed; the efficiency  $S/(Nt)$  is now maximized with a population size as small as 30, being decreased with increasing population size over 30. The optimum cycle of selection to give the largest efficiency is seven with  $N = 30$ . The influence of selection intensity on the efficiency  $S/$



**Fig. 1** Selection efficiency  $S/t$  under different selection cycles and population sizes ( $\alpha = 0.1$ ).

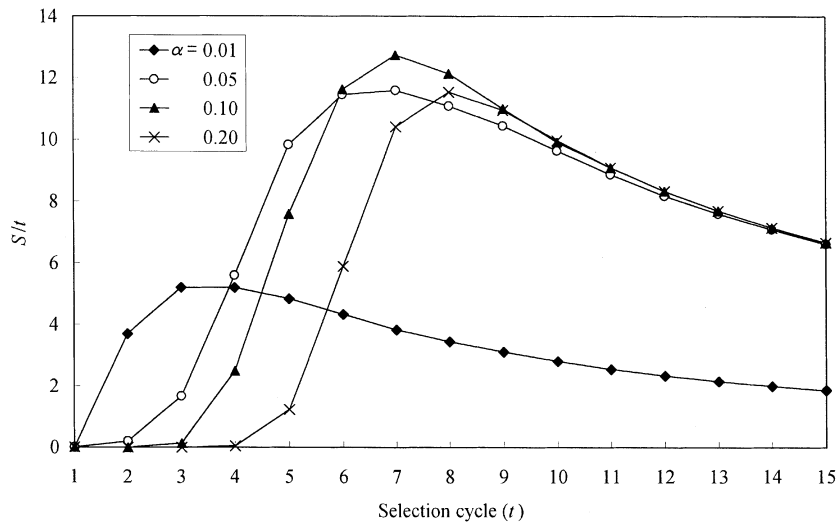


Fig. 2 Selection efficiency  $S/t$  under different selection cycles and intensities ( $N = 100$ ).

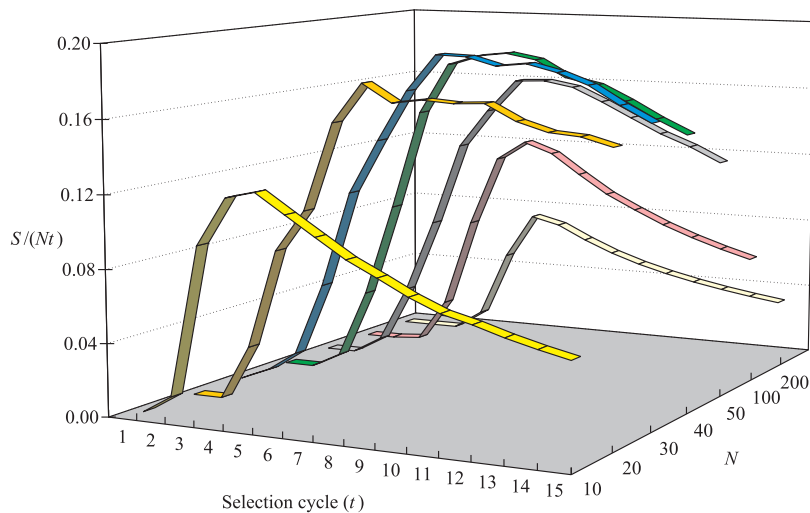


Fig. 3 Selection efficiency  $S/Nt$  under different selection cycles and population sizes ( $\alpha = 0.1$ ).

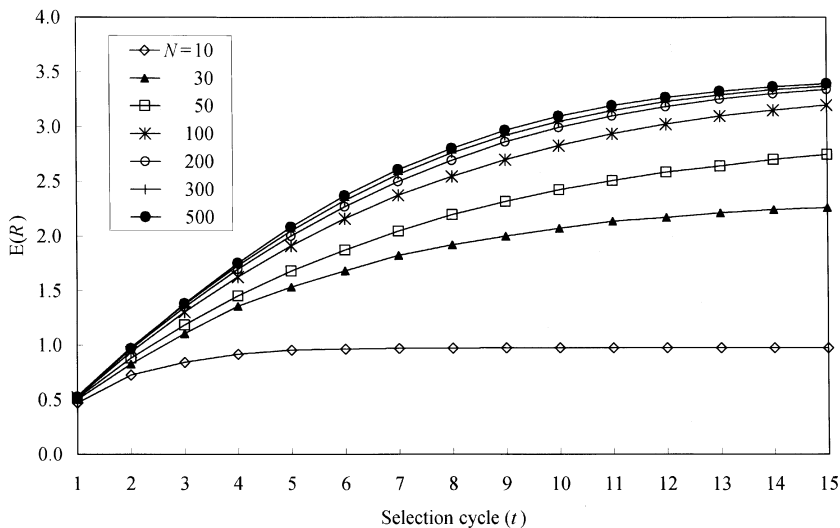


Fig. 4 Expected genetic gain  $E(R)$  under different selection cycles and population sizes ( $\alpha = 0.1$ ).

( $Nt$ ) is the same as mentioned above for  $S/t$  (data not presented).

Calculations of  $E(R)$  are presented in Figs 4 and 5, which were obtained with the same conditions as Figs 1 and 2, respectively. In these figures,  $E(R)$  increases as expected consistently with increasing cycles of selection, and the optimum cycle of selection cannot objectively be pointed out. As for the population size (Fig. 4), a larger population size gives a higher value of  $E(R)$  in any selection cycle, and the difference in  $E(R)$  resulting from different population sizes inflates with increasing selection cycle. However, because random drifts in plus and minus directions cancel each other out when calculating  $E(R)$ , different population sizes cause much smaller differences in  $E(R)$  than those observed in Figs 1 and 3 with the optimum selection cycles.  $E(R)$  reaches a plateau when  $N$  becomes as large as 200. Because the optimum selection cycle cannot be defined by  $E(R)$ , neither can the optimum selection intensity be (Fig. 5). However, a similar trend to that observed in Fig. 2 is observed in Fig. 5; a very high intensity is desirable with a few selection cycles, but not for more cycles.

Calculations of  $E(R)/t$  and  $E(R)/(Nt)$  are presented in Figs 6 and 7, which were obtained with the same conditions as in Figs 1 and 3, respectively. In all cases calculated, both  $E(R)/t$  and  $E(R)/(Nt)$  decline consistently with increasing  $t$ . If the index  $E(R)/C$  really measures the efficiency of selection, the results of these figures should indicate that selection with a single cycle is the most efficient. This obviously does not make sense. A new variety could seldom be obtained if selection is closed with only one cycle. The optimum selection cycle cannot be defined by  $E(R)/C$  either.

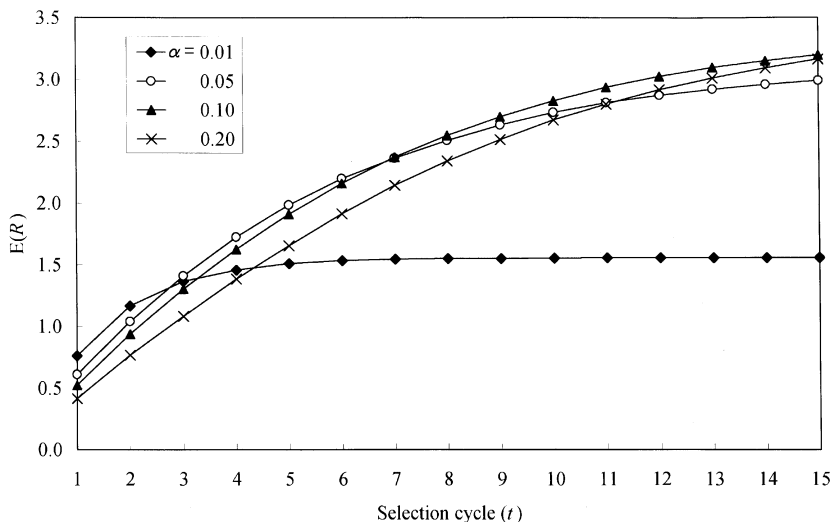
As for the effect of population size  $N$ ,  $E(R)/t$  (Fig. 6) gives the same trend as  $E(R)$  (Fig. 4), in the sense that at

each cycle of selection as  $E(R)/t$  increases with increasing  $N$ , reaching a plateau as  $N$  approaches about 200. The contribution of  $N$  is reversed when evaluated by  $E(R)/(Nt)$ ; in Fig. 7, smaller population sizes give higher values of  $E(R)/(Nt)$  in all selection cycles. As observed in Fig. 4, however, the genetic gain should be quite limited with a population as small as 10. On the other hand, only a limited number of populations could be tried with a large population size, which restricts the opportunity for trying really desirable populations. Therefore, there should be an optimum population size. However, the optimum population size recognized by  $S/(Nt)$  (Fig. 3) cannot be defined by  $E(R)/(Nt)$ .

In conclusion, the optimum cycle (Figs 1 and 3) and intensity (Fig. 2) of selection and optimum size of population (Fig. 3) could be defined by index  $S/C$ , but not by  $E(R)$  and  $E(R)/C$ .

## Discussion

We have shown here that optimum selection procedures with optimum resource allocation per target population can be defined by the newly introduced index  $S/C$ . These selection procedures may not give the largest possible genetic gain in individual target populations, but they give the most opportunities of achieving a sufficiently high genetic gain required for recognition of a new commercial variety. The traditional indices  $E(R)$  and  $E(R)/C$  may define selection procedures to give the largest expected genetic gain in a target population, but not those to give the most opportunities. In the context of mass selection considered in this paper, the optimum cycle and intensity of selection and optimum population size could be defined by the index  $S/C$ , but not by  $E(R)$  and  $E(R)/C$ .



**Fig. 5** Expected genetic gain  $E(R)$  under different selection cycles and intensities ( $N = 100$ ).

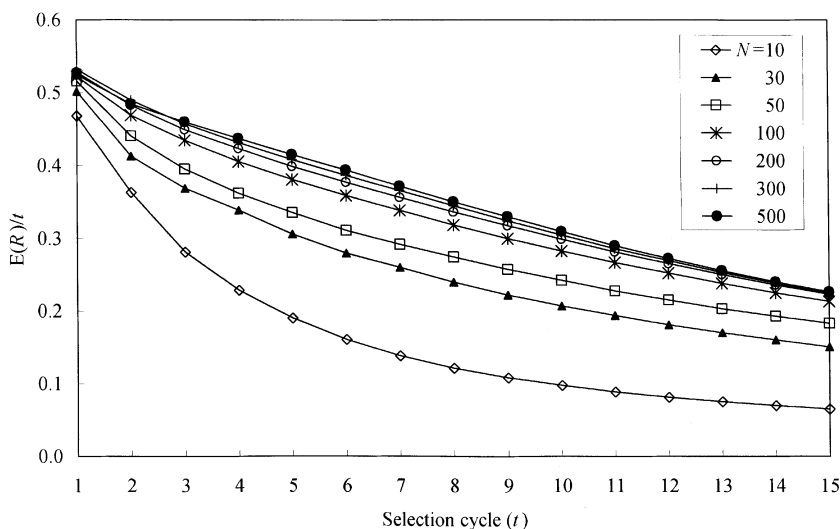


Fig. 6 Ratio of the expected genetic gain to cost  $E(R)/t$  under different selection cycles and population sizes ( $\alpha = 0.1$ ).

The advantage of using  $S/C$  is not confined to that discussed above.  $S/C$  can be used to settle some other optimization issues which cannot be approached by  $E(R)$  and  $E(R)/C$ . It is of great practical importance to decide which strategy should be employed: trying many target populations with a relatively low investment for each population, or a single or a few target populations with a high investment for each. This problem has been discussed previously by Yonezawa & Yamagata (1978) and Weber (1979) for a quite specific case of minimizing the risk of missing desirable genotypes in  $F_2$  populations of selfing crop. The conclusion was that the number of  $F_2$  populations (cross combinations) rather than the size of population was important. Selection was not considered in this discussion. A later discussion by Yonezawa

& Yamagata (1982), taking into account the selection in  $F_2$  and  $F_3$  generations, led to a modified conclusion that several hundred plants for an  $F_2$  population and a few to several hundred plants for an  $F_3$  population are desirable. This problem needs to be considered based on a more general model of selection and index of efficiency. As will be discussed elsewhere, this problem can effectively be dealt with by  $S/C$ .

It is another problem of practical importance to decide which is more efficient: testing many individuals with a simple low input assessment procedure, or few individuals with a refined high input assessment. This problem, which was previously discussed by Yonezawa (1983) for a single cycle of selection in breeding of a selfing crop plant, also needs to be investigated with a

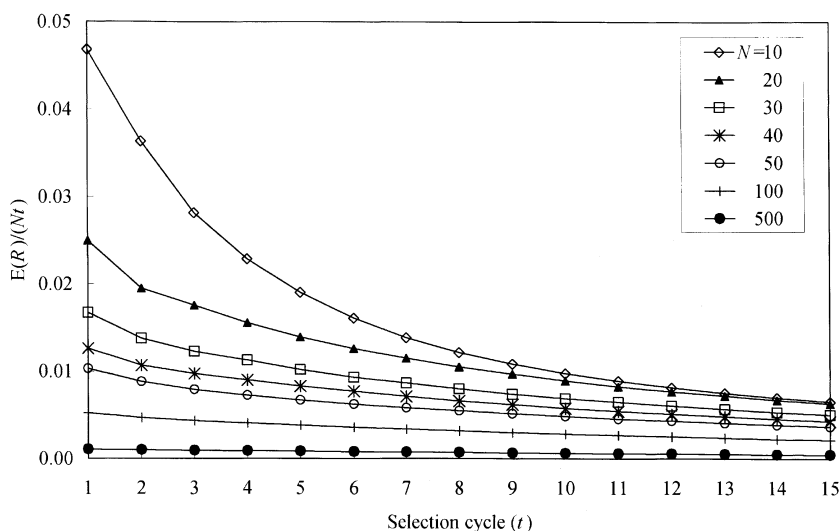


Fig. 7 Ratio of the expected genetic gain to cost  $E(R)/(Nt)$  under different selection cycles and population sizes ( $\alpha = 0.1$ ).

more general theoretical model. It will be shown elsewhere that  $S/C$  provides a much better solution to this question.

In the present simulations, the result of selection in individual runs was evaluated by the genetic gain in population means. In selfing plants, however, it is not the improvement of population means but the acquisition of a particular desirable genotype that is the aim of selection. In this case, the result of selection for a target population may be better evaluated by the presence in or absence from the population of the desired genotype. Index  $S/C$  can also be used in this case if  $S$  is redefined as the probability that the desired genotype is included in a target population (run). This probability is calculated by the relative frequency of the simulation runs in which the desired genotype is found.

The concept of the chance of success  $S$  can be used to optimize not only the selection procedures for genetically segregating populations, but also the yield-screening trials of candidate lines or varieties. The yield trial system to date has been discussed in terms of the expected mean yield of selected lines over that of all candidate lines tested (Finney, 1958; Young, 1972; Bos, 1983). We consider that the probability of high-yielding lines being selected is a more suitable criterion, because the breeder is interested in the acquisition of high-yielding lines, not the mean yield of selected lines. An approach using this probability adds some points that are not derived by the expected mean yield. This issue will be discussed in detail elsewhere.

The criterion discussed here is applicable also to animal breeding if selection is made to achieve a particular sufficiently large, although not the largest possible, genetic gain from a target population. A problem in using  $S$  as the index of selection efficiency is that  $S$  is, in general, more difficult than  $E(R)$  to formulate in analytically tractable equation. However, what is most important is not the mathematical simplicity of an index but how well it works in solving practical issues. With increasing accessibility to high-speed computing systems, mathematical simplicity is becoming less important.

## References

- AKEMINE, H. 1958. Practical aspects of bulk method of breeding. In: Sakai, K., Takahashi, R. and Akemine, H. (eds) *Studies on the Bulk Method of Plant Breeding*, pp. 265–281. Yokendo, Tokyo (in Japanese).
- BOS, I. 1983. The optimum number of replications when testing lines or families on a fixed number of plots. *Euphytica*, **32**, 311–318.
- FALCONER, D. S. 1981. *Introduction to Quantitative Genetics*, 2nd edn. Longman, Essex.
- FINNEY, D. J. 1958. Plant selection for yield improvement. *Euphytica*, **7**, 83–106.
- HILL, W. G. 1977. Variation in response to selection. In: Pollak, E., Kempthorne, O. and Bailey, T. B. (eds) *Proceedings of the International Conference on Quantitative Genetics*, pp. 343–365. Iowa State University Press, Ames, IA.
- HOSPITAL, F. and CHEVALET, C. 1993. Effects of population size and linkage on optimal selection intensity. *Theor. Appl. Genet.*, **86**, 775–780.
- IYAMA, S. 1979. Breeding system of self-fertilizing crops by generation acceleration. *Rec. Adv. Breeding*, **20**, 116–120 (published in Japanese by Japanese Society of Breeding).
- JANSEN, R. C. 1992. On the selection for specific genes in double haploids. *Heredity*, **69**, 92–95.
- NAMKOONG, G. 1970. Optimum allocation of selection intensity in two stages of truncation selection. *Biometrics*, **86**, 465–476.
- NICHOLAS, F. W. 1980. Size of population required for artificial selection. *Genet. Res.*, **35**, 85–105.
- QUINTON, M., SMITH, C. and GODDARD, M. E. 1992. Comparison of selection methods at the same level of inbreeding. *J. Anim. Sci.*, **70**, 1060–1067.
- ROBERTSON, A. 1960. A theory of limits in artificial selection. *Proc. R. Soc. B*, **153**, 234–249.
- ROBERTSON, A. 1970. A theory of limits in artificial selection with many linked loci. In: Kojima, K. (ed.) *Mathematical Topics in Population Genetics*, pp. 246–288. Springer, Berlin.
- WEBER, W. E. 1979. Number and size of cross progenies from a constant total number of plants manageable in a breeding program. *Euphytica*, **28**, 453–456.
- WEI, M., CABALLERO, A. AND HILL, W. G. 1996. Selection response in finite populations. *Genetics*, **144**, 1961–1974.
- WEI, R. P. 1995. Optimal restricted phenotypic selection. *Theor. Appl. Genet.*, **91**, 389–394.
- YONEZAWA, K. 1983. Practical implication of improving the precision of genotype assessment in selection – A theory. *Euphytica*, **32**, 543–555.
- YONEZAWA, K. AND YAMAGATA, H. 1978. On the number and size of cross combinations in a breeding programme of self-fertilizing crops. *Euphytica*, **27**, 113–116.
- YONEZAWA, K. AND YAMAGATA, H. 1982. Selection strategy in breeding of self-fertilizing crops. IV. The sizes of  $F_2$  and  $F_3$  populations in consideration of selection. *Jap. J. Breed.*, **32**, 35–44.
- YOUNG, J. C. 1972. An investigation of procedures for multiple-stage selection for a variate subject to errors of measurement. *Biometrika*, **59**, 323–334.