(1995). — LOACH, K.: Hormone applications and adventitious root formation in cuttings. A critical review. Acta Horticulturae **227**: 126–133 (1988). — LORENZI, R. and CECCARELLI, N.: Vivaismo e propagazione per talea in relazione al ciclo vegetativo: *Picea abies pungens* kosteriana glauca e *Cupressus sempervirens*. Rivista di Ortoflorofrutticoltura Italiana **65**(1): 11–20 (1981). — RADDI, P.: Variabilità della resistenza al cancro nell'ambito del cipresso comune (*Cupressus sempervirens*) e di altre specie. In: GRASSO, V. and RADDI, P. (Eds.): Il cipresso: Malattie e Difesa. pp. 185–193 (1979). — SINISCALCO, C. and PAVOLETTONI, L.: Cutting propagation of *Cupressus sempervirens* clones resistant to *Seiridium cardinale*. Proceedings XXIII International Horticultural Congress, 1. Oral, Firenze (Italy) August 27 to September 1, 1990, p. 176 (1990). — SPANOS, K. A., PIRRIE, A. and WOODWARD, S.: Micropropagation of *Cupressus sempervirens* L. and *Chamaecyparis lowsoniana* (A. MURR.) PAR.. Silvae Genetica **46**: 291–295. (1997). — SPANOS, K. A., PIRRIE, A. and WOODWARD, S.: The effect of fertilizer and shading treatments on rooting efficiency in cuttings of the *Cupressaceae*. Silvae Genetica **48**: 248–254 (1999). — STANKOVA, T. and PANETSOS, K.: Vegetative propagation of of *Cupressus sempervirens* L. of Cretan origin by softwood stem cuttings. Silvae Genetica **46**: 137–144 (1997). — STUBBS, H. L., BLAZICH, F. A., RANNEY, T. G. and WARREN, S. L.: Propagation of 'Carolina Sappire' smooth Arizona cypress by stem cuttings: effects of growth stage, type of cutting, and IBA treatment. Journal of Environmental Horticulturae **15**(2): 61–64 (1997). — TEISSIER DU CROS, E., FERRANDÈS, P., HALLARD, F., DUCATILLON, C. and ANDRÉOLI, C.: Cypress genetic improvement in France: aims and results. In: PANCONESI, A. (Ed.): Il Cipresso. National Research Council of Italy, Firenze. pp. 121–127 (1991). — XENOPOULOS, S.: Screening for resistance to cypress canker (*Seiridium cardinale*) in three Greek provenances of *Cupressus sempervirens*. European Journal of Forest Pathology **20**: 140–147 (1990). — XENOPOULOS, S.: Pathogenic variability of various isolates of *Seiridium cardinale*, *S. cupressi* and *S. unicorne* inoculated on selected *Cupressus* clones and seedlings. European Journal of Forest Pathology **21**: 129–135 (1991).

Maximizing Gain at Restricted Group Coancestry in Selection from Populations with a Hierarchical Structure

By R.-P. WEI[1]) and F. C. YEH

Department of Renewable Resources, University of Alberta, Edmonton, Alberta, Canada T6G 2H1

(Received 13th September 2000)

Abstract

A general model was derived to find a set of optimal family contributions within a single cycle of selection from populations with a strictly hierarchical structure. The model maximized genetic gain at restricted selection proportion and group coancestry, or minimized group coancestry at restricted selection proportion and genetic gain. Populations generated from single-pair/open-pollinated and nested mating designs, as special cases of hierarchical populations, were considered in order to exemplify optimal selection through numerical analyses and simulations. Numerical analyses were made with the assumption that family numbers were finite, while family sizes were infinitely large. Monte Carlo simulations generated breeding populations of finite family number and size. The contribution of a full-sib family was a function of within-family variation, the breeding values of the different types of families involved, and the constraints considered in optimization. Results concerning the optimal solutions were discussed in terms of selection intensity, group coancestry, heritability and gain.

Key words: Breeding population, optimal selection, family contributions, gain, group coancestry, effective size.

Introduction

Selection to improve mass performance involves two basic considerations. The first is how best to increase the expected genetic gain in the breeding population; various selection methods for this purpose have been proposed (e.g. FALCONER and MACKAY, 1996). Genetic gain increases when information on relatives is used to identify individuals with the highest breeding values (LUSH, 1947; OSBORNE, 1957; HENDERSON, 1984; FALCONER and MACKAY, 1996). It is well established that BLUP is the best selection method in a single cycle of selection.

The second consideration is the increase in the level of group coancestry (COCKERHAM, 1967; hereafter, referred to as coancestry for short unless otherwise noted) in the breeding population. This increase can hinder the realization of the expected gain in production populations and of the long-term breeding goals due to increased probability of inbreeding (depression) and the reduction of genetic variation for further selection. This issue is of particular significance for outcrossing species like forest trees. It is inevitable that selection increases coancestry or reduces genetic variability (BULMER, 1971; BURROWS, 1984; SANTIAGO and CABALLERO, 1995; WEI, 1995). In fact, the maximum gain by using BLUP is obtained at the expense of available genetic variances for later generations of breeding. There are selection methods that result in low or minimum coancestry. For instance, within-family selection leads to minimum coancestry (e.g. WEI, 1995; WEI and LINDGREN, 1995). In addition to gain and coancestry, selection intensity is also often considered as an important factor in selection.

There are many studies on the effects of selection on genetic gain or coancestry alone (e.g. LUSH, 1947; OSBORNE, 1957; ROBERTSON, 1970; BULMER, 1971; JAMES, 1972; BURROWS, 1984; SANTIAGO and CABALLERO, 1995). Most of the practical applications of selection emphasized gain but gave little or no attention to the resultant increase in coancestry. Only recently have some studies compared selection alternatives, and developed selection methodology, to take account of both gain and coancestry (TORO and PEREZ-ENCISO, 1990; QUINTON *et al.*, 1992; WRAY and GODDARD, 1994; WEI, 1995; WEI and LINDGREN, 1995; BRISBANE and GIBSON, 1995; VILLANUEVA and WOOLLIAMS, 1997; MEUWISSEN, 1997; ROSVALL and ANDERSSON, 1999). Coancestry consideration differs from situation to situation (e.g. among species). A method that allows breeders effectively and flexibly

---

[1]) Corresponding author and the current address: Sino-Forest Corporation, 1602-3, Emperor Group Centre, 288 Hennessy Road, Wanchai, Hong Hong. Tel. 852-2893-9880; Fax: 852-2892-2661; Email: runpengwei@sinoforest.com

to compromise expected gain with coancestry seems more allround, powerful and appealing than those methods that either maximize gain or minimize coancestry alone.

As regards coancestry or inbreeding for one or multiple generations of breeding goals, different selection procedures have been proposed to improve gain. These include: selection combining restrictions on family contributions (e.g. Toro and Perez-Encico, 1990; Brishane and Gibson, 1995; Wei, 1995); an index selection with variable weights given to between-family and within-family information (e.g. Toro and Perez-Encico, 1990; Wei, 1995; Villanueva and Wooliams, 1997); selection with constraint on inbreeding rather than on selection intensity (Quinton *et al.*, 1992; Wei, 1995; Villanueva and Woolliams, 1997); selection on a criterion obtained by the linear composition of breeding value and the average relationship of the individual with other selections or estimate of inbreeding rate between the neighboring generations (e.g. Wray and Goddard, 1994; Brisbane and Gibson, 1995; Lindgren and Mullin, 1997). All these alternatives could flexibly compromise gain and coancestry, and it seems more logical to combine both breeding value and coancestry into a single selection criterion (Wray and Goddard, 1994; Brisbane and Gibson, 1995). However, there is no theoretical proof that maximum gain is attained at a certain level of coancestry or at a certain generation amongst multiple generations.

Meuwissen (1997) published a dynamic selection procedure that maximizes the genetic merit of the selected parents while restricting their average relationship. Grundy *et al.* (1998) further improved this procedure by modifying the numerator relationship matrix. The algorithm combines one-generation's selection, unequal mating and offspring numbers into a single solution *a priori,* which may be difficult in many breeding programs, particularly for tree species. A conventional method is that parents (selections) equally contribute to the offspring population from which selections are made as the parents of the next generation. Allowing for selection intensity, gain and effective family number, Lindgren *et al.* (1993) and Wei and Lindgren (1995) derived the optimal selection method appropriate for unrelated families of uniform but infinitely large sizes. This method ensures that maximum gain is achieved at a pre-set selection intensity and effective family number, or maximum effective family number is achieved at a pre-set selection of intensity and gain. This study attempts to develop optimal selection applicable to finite breeding populations of unrelated families and related families with a hierarchical structure. Family sizes are balanced. The constrained factors are either selection intensity and coancestry, or selection intensity and genetic gain. A single cycle of selection is considered.

Basic Assumptions and Model Development

A breeding population of N individuals has a hierarchical structure of families. There are totally k types of families. A family type (sib type) is identified by the coefficient (r_j) of the genetic relatedness between the members belonging to the family. Let $r_1, r_2, ...r_k$ rank from low to high, indicating that family members become closely related as r_j increases. An individual can be a member of different types of families at the same time. For instance, it will be a member of a full-sib family in which it has the same father and mother as other members, and also a member of a half-sib family in which all members have one common parent. The phenotype, $x_{1i2i...(k+1)i}$, of an individual can be written as the sum of the population mean (μ), the independent variables ($a_{1i}, a_{1i2i}, ..., a_{1i2i...ki}$) corresponding to the respective classes of families and its deviation ($a_{1i2i...(k+1)i}$) from the family mean of the kth type of family (full-sib):

$$x_{1i2i...(k+1)i} = \mu + a_{1i} + a_{1i2i} + \ldots + a_{1i2i...ji} + \ldots + a_{1i2i...ki} + a_{1i2i...(k+1)i}, \tag{1}$$

where the first number $j(=1, 2,...,k)$ of the subscript ji $(=1i, 2i,...,ki)$ denotes the family type to which the individual belongs, and the two together denote the rank of the family within that type of family (j=1, 2,...,k), or the rank of an individual within a family of the kth type (j=(k+1)). The combinations of different ji explain that the specific family or individual may be lower in the hierarchical structure than are other family types with lower r_j values. For instance, the subscript, $1i2i...ji$, means that family $1i2i...ji$ is also the member of family $1i$, $2i$, ..., or $(j\text{-}1)i$ in the respective types.

The value of a family variable ($a_{1i2i...ji}$) is its family mean (phenotype) minus the average of all families in the same family type, except for the first type of families (a_{1i}), that are obtained by subtracting the population mean (μ) from the family means. For example, in a population produced by mating each of n_m male parents to a different set of n_f female parents, $a_{1i}=F_h\text{-}\mu$, $a_{1i2i}=F_f\text{-}F_h$ and $a_{1i2i(3)i}=x_{1i2i(3)i}\text{-}F_f$, where F_h and F_f are the respective means of half-sib and full-sib families. Therefore, all independent variables, $a_{1i2i...ji}$ (j=1,2,...,k, (k+1)), have zero mean. Rank families or individuals within the kth type of families with their values decreasing when the subscripts are high, e.g. the best family or individual with subscript ...j1 (j=1,...k) or ...(k+1)1. The meanings of these subscripts will be the same in other situations unless otherwise stated.

The total phenotypic variance, σ^2, is the sum of family variances σ^2_j, $j \leq k$) of different family types, and within family variance ($\sigma^2_{(k+1)}$) of family type k:

$$\sigma^2 = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_j^2 + \ldots + \sigma_k^2 + \sigma_{(k+1)}^2. \tag{2}$$

Total variance and the family components are observed parameters of a population. An individual performance is decided by genetic and environmental factors as measured by the respective variance components. Small and additive effects are exclusively assumed for all genes. The relative effect of the genetic element is described by individual heritability, h^2, the ratio of genetic variance to the total variance. The family variance of any type of family can be further decomposed into the respective genetic and environmental components, and the genetic effects can thus be expressed by the respective family heritabilities, i.e. h_1^2, h_2^2, ..., h_k^2 for family types with r_j from low to high. Heritability within the kth family (the (k+1)th variable) is denoted by $h^2_{(k+1)}$.

Taking all parts of the phenotypic value into the account, the best estimate of an individual's breeding value or genetic value is given by the multiple regression equation:

$$g_{1i2i...(k+1)i} = a_{1i}h_1^2 + a_{1i2i}h_2^2 + \ldots + a_{1i2i...ji}h_j^2 + \ldots + a_{1i2i...ki}h_k^2 + a_{1i2i...(k+1)i}h_{(k+1)}^2. \tag{3}$$

Family heritabilities (h_1^2, h_2^2, ..., h_k^2) and within family heritability ($h^2_{(k+1)}$) are, therefore, defined as the partial regression coefficients. A number (n) of superior individuals will be selected. Let $n_{1i2i...ji}$ denote the number of selections contributed by family $1i2i...ji$, then we have:

$$\sum\sum\cdots\sum n_{1i2i...ki} = \sum\sum\cdots\sum n_{1i2i...(k-1)i} = \ldots = \sum\sum n_{1i2i} = \sum n_{1i} = n, \tag{4}$$

where different Σ's imply different number of summands. Genetic gain (ΔG) under selection is estimated by averaging the breeding values of selections, which can be written as the sum of the components contributed by different family types and within full-sib family deviations:

$$\Delta G = E\left(g_{1i2i\ldots(k+1)i}\right) \tag{5}$$

$$= \frac{1}{n}\Big(\sum n_{1i}a_{1i}h_1^2 + \sum\sum n_{1i2i}a_{1i2i}h_2^2 + \ldots + \sum\sum\cdots\sum n_{1i2i\ldots ji}a_{1i2i\ldots ji}h_j^2 + \ldots +$$

$$\sum\sum\cdots\sum n_{1i2i\ldots ki}a_{1i2i\ldots ki}h_k^2 + \sum\sum\cdots\sum n_{1i2i\ldots ki}\bar{a}_{1i2i\ldots ki}h_{(k+1)}^2\Big),$$

where $\bar{a}_{1i2i\ldots ki}=\Sigma a_{1i2i\ldots(k+1)i}/n_{1i2i\ldots ki}$, the mean of the $a_{1i2i\ldots(k+1)i}$ variable upward truncated, or the mean deviation from $a_{1i2i\ldots ki}$ of $n_{1i2i\ldots ki}$ selections.

Assume that all individuals in the population are non-inbred. For each individual within a group of n selections, there are n possible pair combinations of relationships including self-pairing. The relationships among selections are: 0.5 for an individual self-pairing, 0.5 r_j with any of its ($n_{1i\ldots ji}$-1)-($n_{1i\ldots(j+1)i}$-1) sibs of the jth-type family with $j<k$, 0.5 r_k with its $n_{1i2i\ldots ki}$-1 sibs in the kth-type family, and zero with n-n_{1i} other unrelated individuals. The mean contribution to the group coancestry (COCKERHAM, 1967) by the 1*i*2*i*...(*k*+1)*i*th selection is therefore

$$c_{1i2i\ldots(k+1)i} = \frac{0.5}{n}\Big[1 + r_1\left\{\left(n_{1i}-1\right)-\left(n_{1i2i}-1\right)\right\} + r_2\left\{\left(n_{1i2i}-1\right)-\left(n_{1i2i3i}\right)\right\} + \ldots + \tag{6}$$

$$r_{(k-1)}\left\{\left(n_{1i2i\ldots(k-1)i}-1\right)-\left(n_{1i2i\ldots ki}-1\right)\right\} + r_k\left(n_{1i2i\ldots ki}-1\right)\Big]$$

$$= \frac{0.5}{n}\Big\{1 + r_1\left(n_{1i}-1\right) + \left(r_2-r_1\right)\left(n_{1i2i}-1\right) + \ldots + \left(r_j - r_{(j-1)}\right)\left(n_{1i2i\ldots ji}-1\right) + \ldots +$$

$$\left(r_k - r_{(k-1)}\right)\left(n_{1i2i\ldots ki}-1\right)\Big\},$$

depending on family contributions, $n_{1i2i\ldots ji.}$ Averaging over all n selections, the group coancestry of the whole selected population is

$$C = \frac{0.5}{n^2}\Big\{n + r_1\sum n_{1i}\left(n_{1i}-1\right) + \left(r_2-r_1\right)\sum\sum n_{1i2i}\left(n_{1i2i}-1\right) + \ldots + \tag{7}$$

$$\left(r_j - r_{(j-1)}\right)\sum\sum\cdots\sum n_{1i2i\ldots ji}\left(n_{1i2i\ldots ji}-1\right) + \ldots + \left(r_k - r_{(k-1)}\right)\sum\sum\cdots\sum n_{1i2i\ldots ki}\left(n_{1i2i\ldots ki}-1\right)\Big\}.$$

If self-pairing is excluded, the average coancestry of the selected group is

$$C^* = \frac{1}{n}\left\{0.5 + \left(n-1\right)C\right\}. \tag{8}$$

Effective population size of the selected population is therefore defined as

$$N_e = \frac{0.5}{C}, \tag{9}$$

which summarizes the status of the breeding population after selection in comparison with an ideal population (BURROWS, 1984; FALCONER and MACKAY, 1996; NOITON and ALSPACH, 1996; LINDGREN and MULLIN, 1998).

Equations 5 and 7 hold for any kind of selection. They also show that a selection solution or decision by any means can be expressed in terms of family contributions, $n_{1i2i\ldots ki}$, which in turn decide the contributions of other types of families (eq. 4). Therefore, the problem of optimization is to find a set of family contributions ($n_{1i2i\ldots ki}$) that maximize genetic gain (eq. 5) under the constraints

$$n=n_c,\ C=C_c, \text{ and } 0\le n_{1i2i\ldots ki}\le s, \tag{10}$$

where s is the size of family 1*i*2*i*...*ki*. In operation, n_c and C_c are specified in advance as constant numbers. Trying equivalently to minimize -ΔG and rewriting the constraint $0\le n_{1i2i\ldots ki}\le s$ as -$n_{1i2i\ldots ki}\le 0$ and $n_{1i2i\ldots ki}$-s≤0, we defines the LAGRANGE function as follows

$$L = -\Delta G + \lambda_1\left(n-n_c\right) + \lambda_2\left(C-C_c\right) + \sum\sum\cdots\sum \gamma_{1i2i\ldots ki}\left(-n_{1i2i\ldots ki}\right) + \tag{11}$$

$$\sum\sum\cdots\sum \eta_{1i2i\ldots ki}\left(n_{1i2i\ldots ki}-s\right),$$

where λ_1, λ_2, $\gamma_{1i2i\ldots ki}$ and $\eta_{1i2i\ldots ki}$ are the LAGRANGE multipliers. Let $n_{1i2i\ldots ki}$ be the optimal solution. The KUHN-TUCKER theorem (LUENBERGER, 1989) asserts that there are real constants λ_1 and λ_2, $\gamma_{1i2i\ldots ki}\ge 0$ and $\gamma_{1i2i\ldots ki}\ge 0$ such that

$$\frac{\partial L}{\partial n_{1i2i\ldots ki}} = 0, \quad \gamma_{1i2i\ldots ki}\left(-n_{1i2i\ldots ki}\right) = 0, \quad \eta_{1i2i\ldots ki}\left(n_{1i2i\ldots ki}-s\right) = 0.$$

Considering the situations with $0<n_{1i2i...ki}<s$ where $\gamma_{1i2i...ki}=0$ and $\eta_{1i2i...ki}=0$, respectively, we get

$$\frac{\partial L}{\partial n_{1i2i...ki}} = \frac{\partial}{\partial n_{1i2i...ki}}\{-\Delta G + \lambda_1(n-n_c)+\lambda_2(C-C_c)\} \tag{12}$$

$$= -\frac{1}{n}\{a_{1i}h_1^2 + a_{1i2i}h_2^2 + \ldots + a_{1i2i...ki}h_k^2 + t(n_{1i2i...ki})h_{(k+1)}^2\} + \lambda_1 +$$

$$\frac{\lambda_2}{n^2}\{r_1 n_{1i} + (r_2 - r_1)n_{1i2i} + \ldots + (r_k - r_{(k-1)})n_{1i2i...ki} - 0.5r_k\}$$

$$=0$$

where $t(n_{1i2i...ki})$ derived from $\partial(n_{1i2i...ki}\bar{a}1_{i2i...ki})/\partial(n_{1i2i...ki})$ is the truncation point above which $n_{1i2i...ki}$ top-ranking individuals are selected. Selection based on eq. 12 is referred to as optimal selection. When $\gamma_{1i2i...ki}\geq 0$ and $\eta_{1i2i...ki}=0$ then $n_{1i2i...ki}=0$; when $\eta_{1i2i...ki}\geq 0$ and $\gamma_{1i2i...ki}=0$, $n_{1i2i...ki}=s$. Clearly 0 and s are the both limits of $n_{1i2i...ki}$. It should be noted that using C^* instead of C, and minimizing C or C^* with constraint on gain (ΔG) instead of C or C^* in eq. 10 gives the same selection model as eq. 12.

Family contributions, $n_{1i2i...ki}$, which could be obtained from eq. 12, are functions of the independent variables or breeding values ($a_{1i2i...ji}h_j^2$) corresponding to the respective types of families. They differ with families for a given family type. The model in eq. 12 exhibits complexity in giving optimal solutions. We will instantiate the model by considering unrelated and related families, which are the special cases of hierarchical populations.

Examples

Single-Pair / Open-Pollinated and Nested Mating Design

Generally, eq. 12 may apply to any population with a strictly hierarchical structure involving different types (totally k) of sib relationships or families. To illustrate and exemplify the results above, we consider populations produced by using open-pollinated, single-pair and nested mating. Open-pollinated families can be seen as the special case of nested mating in which each female parent has many randomly sampled male partners but only female parents are identified. All parents used to generate these populations are unrelated. For populations of open-pollinated or single-pair families, there are m_1 unrelated families, each of s members. This is the simplest situation in breeding. Populations produced by nested mating have m_1 male parents, and m_2 female parents as the partners of each male parent. Therefore, there are m_1 unrelated half-sib families and m_1m_2 full-sib families. Parameters needed for making optimal selection from these two kinds of populations are formulated in *table 1*.

Table 1. – Formulations for calculations of important parameters for populations under single-pair/open-pollinated and hierarchical mating design, respectively (*cf.* OSBORNE, 1957; FALCONER and MACKAY, 1996).

Finite m_1, (m_2) and s	Infinite s
Single-pair/open-pollinated mating	
$k=1$	$k=1$
$r_1 = 0.5$ or 0.25	$r_1 = 0.5$ or 0.25
$x_{1i(2)i} = a_{1i} + a_{1i(2)i}$	$x_{1i(2)i} = a_{1i} + a_{1i(2)i}$
$\sigma^2 = \sigma_1^2 + \sigma_{(2)}^2$	$\sigma^2 = \sigma_1^2 + \sigma_{(2)}^2$
$\sigma_1^2 = \sigma^2\{1 + r_1h^2(s-1)\}/s$	$\sigma_1^2 = \sigma^2 r_1 h^2$
$\sigma_{(2)}^2 = \sigma^2(1-r_1h^2)(s-1)/s$	$\sigma_{(2)}^2 = \sigma^2(1-r_1h^2)$
$h_1^2 = h^2\{1+r_1(s-1)\}/\{1+r_1h^2(s-1)\}$	$h_1^2 = 1$
$h_{(2)}^2 = h^2(1-r_1)/(1-r_1h^2)$	$h_{(2)}^2 = h^2(1-r_1)/(1-r_1h^2)$
Hierarchical nested mating	
$k=2$	$k=2$
$r_1 = 0.25, r_2 = 0.5$	$r_1 = 0.25, r_2 = 0.5$
$x_{1i2i(3)i} = a_{1i} + a_{1i2i} + a_{1i2i(3)i}$	$x_{1i2i(3)i} = a_{1i} + a_{1i2i} + a_{1i2i(3)i}$
$\sigma^2 = \sigma_1^2 + \sigma_2^2 + \sigma_{(3)}^2$	$\sigma^2 = \sigma_1^2 + \sigma_2^2 + \sigma_{(3)}^2$
$\sigma_1^2 = \sigma^2[1 + h^2\{r_2(m_2s-1) - r_1s(m_2-1)\}]/(m_2s)$	$\sigma_1^2 = \sigma^2h^2\{r_2m_2 - r_1(m_2-1)\}/m_2$
$\sigma_2^2 = \sigma^2(m_2-1)\{1+h^2(r_1s-r_2)\}/(m_2s)$	$\sigma_2^2 = \sigma^2r_1h^2(m_2-1)/m_2$
$\sigma_{(3)}^2 = \sigma^2(1-r_2h^2)(s-1)/s$	$\sigma_{(3)}^2 = \sigma^2(1-r_2h^2)$
$h_1^2 = h^2(1+r_1m_2s+r_1s-r_2)/\{1+h^2(r_1m_2s+r_1s-r_2)\}$	$h_1^2 = 1$
$h_2^2 = h^2(1+r_1s-r_2)/\{1+h^2(r_1s-r_2)\}$	$h_2^2 = 1$
$h_{(3)}^2 = h^2(1-r_2)/(1-r_2h^2)$	$h_{(3)}^2 = h^2(1-r_2)/(1-r_2h^2)$

Selection When Family Sizes Are Infinite

Infinite or even large family size (s) is not realistic in practice. However, the results of selection based on this assumption would provide general profiles of the relationships amongst concerned factors, thereby providing a better understanding of a selection model. In addition, the consequences of selection are good approximations of those results taken from populations with large enough family size.

Optimal selection applied to populations of full-sib or half-sib family (unrelated) that are infinite in number and size has been well-formulated and studied in WEI and LINDGREN (1995). In this study, we focus on selection from populations with a nested structure (k=2) and assume infinite family size (s) but finite family numbers (for both m_1 and m_2). Instead of n_{1i} or n_{1i2i}, the proportion (p_{1i} or p_{1i2i}) of the selected individuals from a family is used, that is, $p_{1i}=n_{1i}/m_2s$ or $p_{1i2i}=n_{1i2i}/s$. The overall selected proportion of the initial population is $P=n/N=\Sigma p_{1i}/m_1$ $=\Sigma\Sigma p_{1i2i}/(m_1m_2)$. Accordingly, the proportion (w_{1i} or w_{1i2i}) of the selected individuals from a family to the selected population $w_{1i}=n_{1i}/n=p_{1i}/(m_1P)$ or $w_{1i2i}=n_{1i2i}/n=p_{1i2i}/(m_1m_2P)$ so that $\Sigma w_{1i}=\Sigma\Sigma w_{1i2i}=1$. Genetic gain for eq. 5 with k=2 can be rewritten in the form,

$$\Delta G=\sum\sum w_{1i2i}\left(a_{1i}h_1^2+a_{1i2i}h_2^2+\bar{a}_{1i2i}h_{(3)}^2\right)=\frac{1}{m_1m_2P}\sum\sum p_{1i2i}\left(a_{1i}h_1^2+a_{1i2i}h_2^2+\bar{a}_{1i2i}h_{(3)}^2\right). \tag{13}$$

Then, eqs. 7 and 8 with k=2 converge at

$$C=0.5\left\{r_1\sum w_{1i}^2+(r_2-r_1)\sum\sum w_{1i2i}^2\right\}=\frac{0.5}{(m_1m_1P)^2}\left\{r_1\sum\left(\sum p_{1i2i}\right)^2+(r_2-r_1)\sum\sum p_{1i2i}^2\right\}. \tag{14}$$

The selection model expressed in eq. 12 is rewritten as

$$a_{1i}h_1^2+a_{1i2i}h_2^2+t(p_{1i2i})h_{(3)}^2-m_1m_2P\lambda_1-\frac{\lambda_2}{(m_1m_2P)}\left\{r_1m_2p_{1i}+(r_2-r_1)p_{1i2i}\right\}=0, \tag{15}$$

which is always true since $0<p_{1i2i}<1$. We know that $p_{1i2i}\rightarrow 1$ as $t(p_{1i2i})\rightarrow -\infty$, $p_{1i2i}\rightarrow 0$ as $t(p_{1i2i})\rightarrow\infty$, and $p_{1i2i}=0.5$ as $t(p_{1i2i})=0$. It can easily be proven that when both λ_1 and λ_2 have zero values $P=\Sigma\Sigma p_{1i2i}/(m_1m_2)=0.5$, which could, in fact, be extended to general situations, that is, $P=\Sigma\Sigma\ldots\Sigma p_{1i2i\ldots ki}/(\Pi m_j)=0.5$.

There is no analytical solution for both λ_1 and λ_2 at constrained selection proportion (P) and coancestry (C). For a given pair of λ_1 and λ_2, there exists an optimal solution. Therefore, a „trial and error" method could be employed to search for p_{1i} and p_{1i2i} when λ_1 and λ_2 are given, and in turn to search for λ_1 and λ_2 when P and C or N_e are specified. In this study we used the Golden Ratio Method (RÅDE and WESTERGREN, 1990). Both p_{1i} and p_{1i2i} have values between 0 and 1. Given λ_1 and λ_2, a preset value of p_{1i}, denoted by $\tilde{p}_{1i}$, leads to a sole set of p_{1i2i} that in turn results in a new value of p_{1i}, denoted by $\hat{p}_{1i}$. Try different values of $\tilde{p}_{1i}$ from 0 to 1 until $\tilde{p}_{1i}=\hat{p}_{1i}$ then $\tilde{p}_{1i}$ and the corresponding set of p_{1i2i} are the final solutions, or optimal family contributions.

Selection from Finite Populations

When there is only one family type (k=1), eq. 12 reduces to

$$a_1h_1^2+t(n_{1i})h_{(2)}^2-n\lambda_1-\frac{\lambda_2}{n}\left(r_1n_{1i}-0.5r_1\right)=0, \tag{16}$$

which is similar to the expression given by LINDGREN *et al.* (1993) and WEI and LINDGREN (1995). As $t(n_{1i})$ is the truncation point corresponding to n_{1i}, family contributions (n_{1i}) are solely decided by family performances. Using nested mating to produce populations, there are two types (k=2) of families, half-sib and full-sib, and eq. 12 becomes

$$a_{1i}h_1^2+a_{1i2i}h_2^2+t(n_{1i2i})h_{(3)}^2-n\lambda_1-\frac{\lambda_2}{n}\left\{r_1n_{1i}+(r_2-r_1)n_{1i2i}-0.5r_2\right\}=0. \tag{17}$$

The truncation points, $t(n_{1i2i})$, correspond to full-sib family contributions, n_{1i2i}. It is clear that full-sib family contributions (n_{1i2i}) are proportional to half-sib family breeding value ($a_{1i}h_1^2$) and full-sib family breeding value ($a_{1i2i}h_2^2$) if λ_1 and λ_2 are constant.

There are s+1 possibilities, that is, 0, 1, 2,...,s, for the contribution (n_{1i} or n_{1i2i}) of a full-sib family with s members, each corresponding to a specific truncation point, $t(n_{1i})$ or $t(n_{1i2i})$, in the present situations. Strictly, an optimum solution of selection should have an n_{1i} and/or n_{1i2i}, together with an exact truncation point which satisfies eqs. 16 or 17. In practice, it is difficult because of the discontinuity of n_{1i} or n_{1i2i}, and the possible random error associated with $a_{1i(2)i}$ or $a_{1i2i(3)i}$, the within-family deviations. To simplify optimal selection, it is reasonable to approximate the truncation point by

$$t(n_{1i})=\begin{cases}a_{1i(2)q},\ q=n_{1i}, & \text{if } n_{1i}>0\\ 2a_{1i(2)1}-a_{1i(2)2}, & \text{if } n_{1i}=0\end{cases}, \tag{18}$$

when unrelated families are considered, or

$$t(n_{1i2i})=\begin{cases}a_{1i2i(3)q},\ q=n_{1i2i}, & \text{if } n_{1i2i}>0\\ 2a_{1i2i(3)1}-a_{1i2i(3)2}, & \text{if } n_{1i2i}=0\end{cases}, \tag{19}$$

when related families are considered. The optimal solution at pre-set λ_1 and λ_2 is thus approximated by a pair of n_{1i} and $t(n_{1i})$ or n_{1i2i} and $t(n_{1i2i})$ that minimizes

$$\left|\left\{a_{1i}h_1^2+t(n_{1i})h_{(2)}^2\right\}-n\lambda_1-\frac{\lambda_2}{n}\left(r_1n_{1i}-0.5r_1\right)\right|$$

or

$$\left|\left\{a_{1i}h_1^2+a_{1i2i}h_2^2+t(n_{1i2i})h_{(3)}^2\right\}-n\lambda_1-\frac{\lambda_2}{n}\left\{r_1n_{1i}+(r_2-r_1)n_{1i2i}-0.5r_2\right\}\right|.$$

As there is a relationship between the truncation point and contribution of a finite family, this algorithm, which is similar to MEUWISSEN (1997), should be superior to the one that directly rounds the on-integer values of optimal solutions.

There are explicit solutions of n_{1i} when there is only one type of families, or k=1. However, the solutions of n_{1i2i} are implicit when k=2 because of the unknown n_{1i} in eq. 17. A search for n_{1i} is necessary before solving n_{1i2i} with this procedure.

The n_{1i}, the contribution of half-sib family 1i, is the sum of the n_{1i2i} over full-sib families within the 1ith half-sib family, and is distributed over integers 0,1,2,...,$min(m_2s,n)$. Given a n_{1i}, denoted by $\tilde{n}_{1i}$, we can obtain a set of n_{1i2i} which sum to $\hat{n}_{1i}$. For each $\tilde{n}_{1i}$, there is a corresponding $\hat{n}_{1i}$. At pre-set λ_1 and λ_2, try all, i.e. 1+*min*(m_2s,n), possibilities of n_{1i}. The optimal solution can be approximated by the $\tilde{n}_{1i}$ that is equal to or has the minimum distance with the corresponding $\hat{n}_{1i}$. A similar "trial and error" method as above described for infinite family size can be used to find the values for λ_1 and λ_2 when P and C are specified (note that, for a finite population, a specified C value may not be realized, rather, a value closest to it would be obtained). The same procedure can be extended to other types of breeding populations having a hierarchical structure. However, it becomes difficult when more family types are involved.

Results

Normal distribution is further assumed for the average performances of all types of families and the within-family deviations of full-sib families. Population mean (μ) is set to zero. Numerical and simulation results are demonstrated for unrelated families (k=1) and related families generated by a nested-mating (k=2). Numerical results for unrelated families of infinite large sizes are detailed in WEI and LINDGREN (1995). Thus, optimal selection is only applied to populations of infinitely large family size (s) produced by using nested mating design. Variables for family means are mimicked by order statistics, and within-family deviations have continuous values. Optimal family contributions are illustrated in *figure 1* for a mating involving 40 male parents (m_1) and 40 female partners (m_2) for each of them. Expected gains at given selected proportion and coancestry are given in *table 2* for a mating with m_1=5 and m_2=4.

Monte Carlo simulation is performed for optimal selection from populations of finite family sizes. ROSVALL and ANDERSSON (1999) suggested group merit selection to maximize the selected population's merit that takes gain and relatedness into account (WRAY and GODDARD, 1994; BRISBANE and GIBSON, 1995). For comparison, an algorithm for group merit selection developed by LINDGREN and MULLIN (1997), which will be referred to as L & M algorithm, is included in the simulation. Family means and within-family deviations are generated at given population structure (unrelated families or related families generated using nested mating), h^2, m_1, m_2, and s. Only full sibs (r_1=0.5) are considered for unrelated families. In a simulated population, selection is conducted for the same selected number (n) and for a variety of selection decisions (varying λ_1 and λ_2 values and gain and coancestry level). At a pre-set λ_1, a value can be found for λ_2, which corresponds to n. The value of λ_2 will vary over repeated populations. Simulations are repeated 50 times. Averaging over simulations, gain and group coancestry and their respective variation coefficients are calculated for unrelated families (*Table 3*) and for related families (*Table 4*).

Unlike classical selection methods that lead to a single solution, optimal selection in eq. 12 provides a series of solutions (*Figure 1; Tables 2, 3* and *4*). Given selection proportion, optimal solutions are confined to a domain of coancestry with both limits corresponding to combined-index and within-family selection, respectively (WEI and LINDGREN, 1995). Combined-index selection yields the high limit, while within-family selection combines the low limit. The high limit varies with populations of the same family number and size because of sampling errors, while the low limit is constant (results not shown, WEI and LINDGREN, 1995). When family size (s) is infinitely large, the number of optimal solutions is infinite. In practice, the number is limited because the combinations of selections are countable, and it is always smaller in number than these possible combinations.

Optimal family contribution following selection is a non-linear function of the mean performances or breeding values of half-sib and full-sib family (*Figure 1*). Generally, increase in family breeding values leads to increase in family contributions. The magnitude of the change relies on family types and the values of λ_1 and λ_2 that decide the constraints, P and C. Given P, low λ_2 drives family contributions to be more differential, indicating an increase in C and also in ΔG. Low λ_1 at cer-

Table 2. – Expected gain (ΔG) for optimal selection at given selection proportion and coancestry from a population produced using hierarchical mating (m_1=5, m_2=4 and s64).

		Heritability (h^2)					
P	$N_e(C)$	0.01	0.10	0.25	0.50	0.75	1.0
0.001	5(0.10)	0.0969	0.4070	0.7937	1.4109	2.0811	2.8739
	10(0.05)	0.0713	0.3355	0.6944	1.2951	1.9676	2.7782
0.010	6(0.0874)	0.0874	0.3503	0.6653	1.1566	1.6832	2.3015
	10(0.05)	0.0674	0.2962	0.5925	1.0764	1.6105	2.2489
0.100	9(0.0556)	0.0667	0.2534	0.4663	0.7904	1.1333	1.5336
	13(0.0385)	0.0456	0.1977	0.3932	0.7118	1.0631	1.4833

Table 3. – Estimated gain ()*G*) and realized average coancestry (*C*) and their respective variation coefficients (*CV*) following optimal and group merit selection from simulated populations of 20 full-sib unrelated families for two family sizes (*s*), one selected number (n=40), two heritabilities (h^2) and four specified effective population sizes (N'_e). Effective population size (N_e) realized in selection is given by 0.5/C.

			Optimal selection			Group merit selection		
s	h^2	N'_e	)G/CV	C/CV	N_e	)G/CV	C/CV	N_e
5	0.1	25	0.077/0.026	0.0200/0.004	25.03	0.077/0.024	0.0200/0.003	25.02
		24	0.084/0.023	0.0205/0.007	24.36	0.084/0.024	0.0206/0.006	24.30
		22	0.103/0.016	0.0227/0.009	22.06	0.102/0.017	0.0227/0.009	22.05
		20	0.114/0.016	0.0248/0.009	20.18	0.114/0.017	0.0248/0.009	20.17
	0.25	25	0.186/0.023	0.0200/0.005	25.04	0.187/0.020	0.0200/0.003	25.02
		24	0.203/0.030	0.0205/0.007	24.36	0.203/0.022	0.0205/0.007	24.35
		22	0.238/0.019	0.0223/0.012	22.38	0.239/0.018	0.0224/0.008	22.32
		20	0.265/0.020	0.0247/0.012	20.21	0.266/0.020	0.0248/0.012	20.19
50	0.1	23	0.157/0.021	0.0214/0.007	23.36	0.157/0.022	0.0214/0.008	23.35
		20	0.203/0.018	0.0247/0.012	20.22	0.201/0.020	0.0248/0.010	20.20
		17	0.242/0.029	0.0290/0.010	17.22	0.240/0.027	0.0291/0.010	17.20
		14	0.276/0.053	0.0352/0.011	14.19	0.274/0.052	0.0352/0.012	14.22
	0.25	23	0.338/0.020	0.0215/0.007	23.29	0.337/0.022	0.0215/0.007	23.30
		20	0.416/0.017	0.0248/0.010	20.19	0.413/0.019	0.0247/0.013	20.22
		17	0.484/0.023	0.0290/0.011	17.24	0.481/0.022	0.0291/0.011	17.19
		14	0.545/0.042	0.0352/0.012	14.22	0.542/0.040	0.0352/0.012	14.20

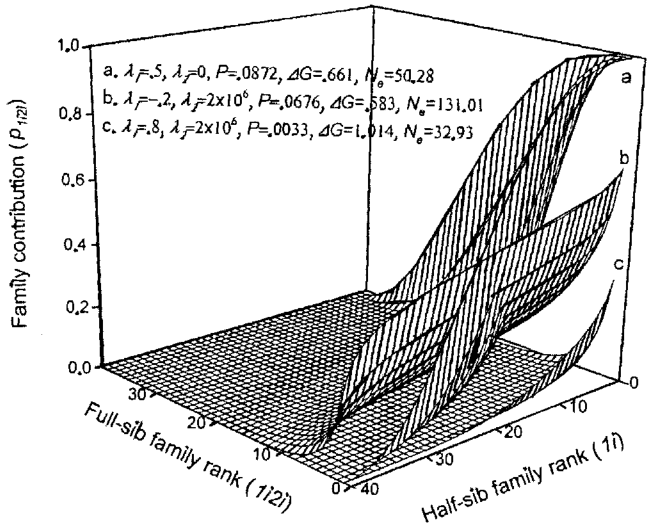

Figure 1. – Illustration of family contributions (p_{1i2i}) against family performances for three different selection schemes (denoted by a, b and c, respectively) from a population generated by mating each of 40 male parents to 40 female parents (nested mating). Individual heritability (h^2) of the character concerned is 0.25.

tain λ_2 largely increases P and N_e but decreases ΔG. The distribution of family contributions over families can be very different between half-sib and full-sib families. To increase gain without changing P much, selection could focus on families with higher ranking in both half-sib and full-sib family performances (e.g. a in *Figure 1*). The most effective way to increase N_e is to keep family contributions uniform over half-sib families (e.g. b in *Figure 1*). The distribution over full-sib families would not matter much.

Tables 2 demonstrates the relationship between heritability (h^2), selection intensity (P), group coancestry (C) and genetic gain (ΔG). As expected, ΔG increases as h^2 increases, selection becomes intensive with low P, and C increases or N_e decreases (also see *Tables 3* and *4*). Genetic gain is high if h^2 is high. Given P and C, ΔG seems linearly proportional to h^2. It is also shown from *table 2* that an intensive selection may yield both higher ΔG and N_e (lower C) than a weak selection, especially when h^2 is relatively high. A set of λ_1 and λ_2 decides an optimal solution with certain P, C and ΔG. Given P and λ_1, λ_2, C and ΔG vary over replicated populations (*Tables 3* and *4*), implying the influences of population random processes. The coefficients (*CV*) of variation trend to increase as ΔG and C increase. At the low limit of C, *CV* for C is 0 or the distribution of selections among families is constant over populations, regardless of h^2 and ΔG.

Optimal selection procedure developed in this study and L & M algorithm (LINDGREN and MULLIN, 1997) yield the same or almost the same ΔG given P and N'_e (*Tables 3* and *4*). For most of situations (repeated simulations) and solutions, both methods are identical (results not shown). The trivial devia-

Table 4. – Estimated gain ()*G*) and realized average coancestry (*C*) and their respective variation coefficients (*CV*) following optimal and group merit selection from simulated populations derived from a nested mating (m_1=5, m_2=4) for two family sizes (*s*), one selected number (*n*=40), two heritabilities (h^2) and four specified effective population sizes (N'_e). Effective population size (N_e) realized in selection is given by 0.5/C.

			Optimal selection			Group merit selection		
s	h^2	N'_e	)*G*/*CV*	*C*/*CV*	N_e	)*G*/*CV*	*C*/*CV*	N_e
5	0.1	13	0.070/0.030	0.0383/0.002	13.04	0.071/0.029	0.0384/0.002	13.02
		12	0.099/0.025	0.0412/0.008	12.13	0.100/0.020	0.0413/0.007	12.10
		11	0.116/0.019	0.0448/0.009	11.17	0.116/0.019	0.0448/0.009	11.16
		10	0.129/0.015	0.0493/0.012	10.14	0.129/0.016	0.0493/0.010	10.14
	0.25	13	0.173/0.033	0.0384/0.002	13.03	0.174/0.030	0.0384/0.002	13.02
		12	0.233/0.022	0.0412/0.006	12.14	0.234/0.018	0.0413/0.007	12.11
		11	0.267/0.018	0.0448/0.011	11.15	0.267/0.013	0.0449/0.008	11.14
		10	0.288/0.015	0.0493/0.013	10.13	0.288/0.016	0.0492/0.013	10.16
50	0.1	13	0.147/0.035	0.0384/0.003	13.03	0.148/0.029	0.0384/0.001	13.01
		12	0.199/0.023	0.0413/0.006	12.12	0.200/0.019	0.0413/0.004	12.09
		11	0.233/0.024	0.0449/0.009	11.14	0.234/0.019	0.0450/0.006	11.12
		10	0.261/0.020	0.0494/0.012	10.12	0.261/0.019	0.0495/0.010	10.11
	0.25	13	0.350/0.028	0.0384/0.002	13.03	0.351/0.025	0.0384/0.001	13.02
		12	0.440/0.018	0.0413/0.005	12.11	0.440/0.019	0.0413/0.005	12.10
		11	0.495/0.020	0.0447/0.009	11.19	0.496/0.019	0.0448/0.009	11.15
		10	0.544/0.017	0.0496/0.009	10.08	0.544/0.018	0.0496/0.008	10.08

tions seem random and negligible. The results, in addition to examining selection based on eq. 12, indirectly verify that L & M algorithm yields optimal solutions or excellent approximations to optimal solutions.

Tables 3 and *4* also provide a comparison of single-pair and nested mating in term of ΔG and C or N_e. For both of mating designs, full-sib family number is the same. The selected number (n_s) is chosen so that the next generation of breeding populations with the same full-sib family number can be generated for the respective mating designs. Thus, the comparison is made at the same level of heritability, full-sib family number and size, and total population size. There appears to be no advantage to having a more complicated mating design in the present situation. The coancestry range with single-pair is narrower and lower than it is with the nested-mating design. No significant difference in gain range is shown between the two mating designs. Indeed, nested mating is much worse in terms of gain, given family size and heritability, and at a similar level of coancestry (*Tables 3* and *4*).

Discussion

A selection can always be expressed as a set of family contributions. Selection model developed in eq. 12 exhibits an implicit and complicated relationship of family contributions ($n_{1i2i...ki}$) with the breeding values of all types of families involved, the corresponding truncation points, Lagrange multipliers, and the contributions of families at lower levels ($j<k$). For given constraints, say, on P and C, multifold iterative techniques have to be employed in numerical computation for searching for λ_1 and λ_2, truncation points and, then, family contributions, especially when k increases. The model in eq. 12 involves only two LAGRANGE multipliers, λ_1 and λ_2, regardless of k values. If more constraints are superimposed, such as on the contributions of certain family types, then more multipliers will be involved.

Optimal selection results in a range of solutions that correspond to sets of selection intensity (P), gain (ΔG) and average coancestry (C). A solution guarantees that no other sets of $n_{1i2i...ki}$ exist which could results in higher ΔG at given P and C or lower C at given P and ΔG. Classical selection methods provide a single choice with specified values of ΔG and C when P is set, but either ΔG, C or both of the two may be far from desired (BULMER, 1971; BURROWS, 1984; FALCONER and MACKAY, 1996; WEI and LINDGREN, 1995). Optimal selection is, therefore, flexible. If the interaction of ΔG and C is known, then a single solution that further optimizes breeding objectives with specified values of P, ΔG and C could be found (LINDGREN *et al.*, 1993; WEI and LINDGREN, 1995).

WEI and LINDGREN (1995) concluded that combined-index/BLUP and within-family selection are the two special cases of optimal selection. Within-family selection combines a low limit of C, and combined-index selection yields the global maximum of ΔG that is obtained at λ_2=0 or when constraint on C in eq. 10 is released. By using eq. 12, family selection is also an extreme case of giving a high limit of C. Combined-index selection always produces C in between both limits. They further concluded, however, that optimal solutions between family selection and combined-index selection should not be treated as optimum, since superior selection solutions that yield both higher ΔG and lower C always exit. Obviously, in lieu it is more natural to view selection based on combined-index or BLUP (Henderson, 1984) as the high limit of C in optimal selection, thus, $\lambda_2 \geq 0$ (Wei and Lindgren, 1995). This conclusion should hold for any type of breeding populations.

Family contributions corresponding to different types of families are simply related by eq. 4, and the optimal contributions ($n_{1i2i...ki}$) of the kth type of families decide the contributions of other types of families. Given constraints on P and C, $n_{1i2i...ki}$ exclusively depend on within full-sib family deviations and on the breeding values of the different types of families. They are also functions of the constraints imposed. At high C or λ_2 close to 0 and if P is not high (e.g. $P<0.01$), then $n_{1i2i...ji}$ are high for few families with high breeding values of all types of families involved but are low for most of the families with relatively low breeding values (*Figure 1*). The difference declines when C decreases, and vanishes at low limit of C. In other words, weak restriction on C (low λ_2) means intense use of between-family variation, while enhanced restriction on C (high λ_2) forces selection to be close to within-family selection. The distribution of the $n_{1i2i...ji}$ can be very different over different types of families. For k dimensions of families, reducing difference in n_{1i} over the first type of families is most important for controlling coancestry.

Genetic gain following optimal selection depends on the choice of P and C. It is higher with lower P and higher C. Investigations have contributed to the relationship between P and ΔG for long-term individual selection without considering C or N_e (e.g. Robertson, 1970; James, 1972). As P is a measure of census population size, N_e is probably more relevant to ΔG obtained in long-term breeding programs (Quinton *et al.*, 1992; Wray and Goddard, 1994; Brisbane and Gibson, 1995; Santiago and Caballero, 1995). Taking both census and effective population size into account would probably give a better understanding of selection mechanisms in breeding programs. Optimal selection maximizes ΔG at given P and N_e for a single generation of breeding, and is likely to perform best in the multi-generation evolution of breeding as well, if the same constraints (e.g. N, mating method and ratio, P and N_e) are considered. Yet, more studies are needed in order to truly unveil the relationship of P, ΔG and N_e over a time horizon.

Genetic gain is a function of heritability. In general, gain is approximately linear to heritability for all optimal solutions (Wei and Lindgren, 1995; *Table 2*). A high limit of C is lower with higher h^2 due to the decreasing use of sibs information. This suggests that a small gain is obtained with high coancestry at low h^2, whereas a substantial gain is obtained with low coancestry at high h^2 (Wei and Lindgren, 1995; *Tables 3* and *4*).

Tables 3 and *4* imply that mating design for parents has an effect on the consequences of optimal selection. The objective of a complicated design like nested mating is to enrich knowledge on the genetic composition of a population in question, to improve gain by using more genetic information in selection and to use non-additive genetic variation. Divergence appears when considering C or N_e. Given the number (e.g. $m_1 m_2$) of full-sib families, a simple mating design can be more efficient than a complicated one in compromising ΔG and C. The reason for this has two aspects. First, C following selection mainly depends on the number (m_1) and contributions (n_{1i}) of families at the lowest value of sib relatedness, that is, r_1. C increases rapidly when m_1 decreases and n_{1i} diverge (*Figure 1*). A complicated design produces fewer m_1 families than a simple one, 5 half-sib versus 20 full-sib families in *tables 3* and *4*. Second, the weighted selection differential corresponding to the family component decreases with small parent numbers (both m_1 and m_2 in *Tables 3* and *4*), a decrease which contributes the main component of ΔG at C not close to the low limit. Selection is supposed to be more efficient in ΔG when more sib information is included. However, the increase may not supplement the loss caused by reduction in the selection differential. If the restriction on the number of full-sib families (e.g. $m_1 m_2$) is released, and other restrictions (e.g. N and P or their combinations) are invoked in lieu, then the conclusions would certainly be different. More treatments are needed to cover the effects of variations in mating design, and family numbers and sizes.

L & M algorithm screens the individuals with the highest merit one by one (either added or subtracted) by imposing a penalty due to increase in relatedness/coancestry on the selection criterion (breeding value) (Wray and Goddard, 1994; Brisbane and Gibson, 1995; Lindgren and Mullin, 1997; Rosvall and Andersson, 1999). It is not theoretically or experimentally proved that the group merit or genetic gain by this algorithm is maximized given selection intensity and coancestry. Simulation results in this study show that L & M algorithm is always optimal or very close to optimal. There are other existing search methods for the integer solutions of group merit selection (e.g. Wray and Goddard, 1994; Brisbane and Gibson, 1995; Andersson *et al.*, 1999). However, no mathematical evidence exists that these methods lead to highest group merit at given selection proportion and coancestry.

Monoecious populations, balanced family size (s) and mating rate (m_2) are considered in selection model development and in numerical calculations or simulations. These constraints can easily be released with minor or without modifications of the selection model by eq. 12. For doecious species like animal species, optimal selection can separately be performed to the male and female populations as long as selection number (n), mating rate and respective C are set. Optimal selection could also be used to attain optimal n and mating ratio at constrained C in a way that is similar to the treatments by Villanueva and Woolliams (1997) and Meuwissen (1997). If family size is unbalanced, an adjusted size over all families is needed to obtain estimates of genetic parameters involved in selection. For family $1i2i...ki$ with size $s_{1i2i...ki}$, the contribution should be confined to $[0, s_{1i2i...ki}]$ instead of $[0, s]$ in selection. In theory, the selection model by eq. 12 can be applied to any breeding population with a hierarchical structure ($k \leq 2$). In practice, the most possible situations are k=1, 2 and 3.

Acknowledgements

We thank an anonymous reviewer's useful comments and suggestions. Dr. R. C. Yang reviewed an early version of the manuscript. This study was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada, A2282, to F.C.Y.

References

Andersson, E. W., Sánchez, L. and Andersson, B.: Group coancestry controlled selection in a *Pinus silvestris* L. breeding program. Theor. Appl. Genet. **99**: 73–80 (1999).— Brisbane, J. R. and Gibson, J. P.: Balancing selection response and rate of inbreeding by including genetic relationship in selection decisions. Theor. Appl. Genet. **91**: 421–431 (1995). — Bulmer, M. G.: The effect of selection on genetic variability. Am. Nat. **105**: 201–211 (1971). — Burrows, P. P.: Inbreeding under selection from related families. Biometrics **40**: 895–906 (1984). — Cockerham, C. C.: Group inbreeding and coancestry. Genetics **56**: 89–104 (1967). — Falconer, D. S. and Mackay, T. F. C.: Introduction to Quantitative Genetics. 4th ed. Longman Scientific & Technical, Harlow/New York (1996). — Grundy, B., Billanueva, B. and Woolliams, J. A.: Dynamic selection procedure for constrained inbreeding and their consequences for pedigree development. Genet. Res. **72**: 159–168 (1998). — HENDERSON, C. R.: Applications of Linear Models in Animal Breeding. University of Guelph, Guelph, ON, Canada (1984). — JAMES, J. W.: Optimum selection intensity in breeding programmes. Anim. Prod. **24**: 287–305 (1972). — LINDGREN, D. and MULLIN, T. J.: Balancing gain and relatedness in selection. Silvae Genet. **46**: 124–129 (1997). — LINDGREN, D. and MULLIN, T. J.: Relatedness and status number in seed orchard crops. Can. J. For. Res. **28**: 276–283 (1998). — LINDGREN, D., WEI, R.-P., and BONDESSON, L. F.: Optimal selection from families. Heredity **70**: 619–621 (1993). — LUENBERGER, D. G.: Linear and Nonlinear Programming. 2nd ed. Addison-Wesley Publishing Company, USA (1989). — LUSH, J. L.: Family merit and individual merit as bases for selection. Am. Nat. **81**: 241–261 and 262–379 (1947). — MEUWISSEN, T. H. E.: Maximizing the response of selection with a predefined rate of inbreeding. J. Anim. Sci. **75**: 934–940 (1997). — NOITON, D. A. M. and ALSPACH, P. A.: Founding clones, inbreeding, coancestry, and status number of modern apple cultivars. J. Am. Soc. Hort. Sci. **121**: 773–782 (1996). — OSBORNE, R.: The use of sire and dam family averages in increasing the efficiency of selective breeding under a hierarchical mating system. Heredity **11**: 93–116 (1957). — QUINTON, M., SMITH, C. and GODDARD, M. E.: Comparison of selection methods at the same level of inbreeding. J. Anim. Sci. **70**: 1060–1067 (1992). — RÅDE, L., and WESTERGREN, B.: Mathematics Handbook. 2nd ed. Studentlitteratur, Lund, Sweden (1990). — ROBERTSON, A.: Some optimum problems in individual selection. Theor. Pop. Biol. **1**: 120–127 (1970). — ROSVALL, O. and ANDERSSON, E.W.: Group-merit selection compared to conventional restricted selection for trade-offs between genetic gain and diversity. For. Genet. **6**: 1–14 (1999). — SANTIAGO, E. and CABALLERO, A.: Effective size of populations under selection. Genetics **139**: 1013–1030 (1995). — TORO, M. and PEREZ-ENCISO, M.: Optimization of selection response under restricted inbreeding. Genet. Sel. Evol. **22**: 93–107 (1990). — VILLANUEVA, B. and WOOLLIAMS, J. A.: Optimization of breeding programmes under index selection and constrained inbreeding. Genet. Res. **69**: 145–148 (1997). — WEI, R.-P.: Predicting genetic diversity and optimizing selection in breeding programmes. PhD Thesis, Swedish University of Agricultural Sciences, Umeå, Sweden. ISBN 91-576-4900-1 (1995). — WEI, R.-P. and LINDGREN, D.: Optimal family contributions and a linear approximation. Theor. Pop. Biol. **48**: 318–332 (1995). — WRAY, N. R. and GODDARD, M. E.: Increasing long-term response to selection. Genet. Sel. Evol. **26**: 431–451 (1994).

Inheritance of Two Chlorophyll Mutants in *Eucalyptus globulus*

By B. PATTERSON[1]), C. M. WOLBANG, R. E. VAILLANCOURT and B. M. POTTS

Cooperative Research Centre for Sustainable Production Forestry and School of Plant Science, University of Tasmania,
G.P.O. Box 252-55, Hobart, Tasmania 7001, Australia

(Received 10th November 2000)

Summary

Mutations in two independent genes causing chlorophyll deficiency (*chl1* and *chl2*) were identified in two *Eucalyptus globulus* trees from remnant native forest. In each tree the proportion of albino progeny, following selfing, was found to be consistent with the segregation of a single gene. One of these genes (*chl1*) was found to be linked to the isozyme locus *Gpi-2*. The frequencies of the mutants were monitored in open-pollinated seed progeny and, from this, outcrossing rates were calculated. The rate of outcrossing in the *chl1* mutant tree was also determined using a multi-locus, maximum likelihood estimation based on three isozyme loci, this agreed closely with the *chl1* single locus estimate. This result highlights a role for rare, easily scored, mutants in seed orchards, where they can be utilized to monitor selfing rates.

Key words: outcrossing, selfing, albino, eucalypt, seed orchard.

Introduction

High levels of deleterious recessive mutations may accumulate, particularly in outbreeding species (BYERS and WALLER, 1999). Such alleles are exposed upon inbreeding, and abnormal phenotypes have been reported for many forest tree taxa (e.g. FRANKLIN, 1970; ELDRIDGE, 1970). The expression of these recessive mutations is believed to be the main cause of inbreeding depression in forest tree species (LEDIG, 1986), including eucalypts (GRIFFIN, MORAN and FRIPP, 1987; JAMES and KENNINGTON, 1993; HARDNER and POTTS, 1995). There are some reports of abnormal seedling phenotypes in eucalypts which are believed to be due to single recessive genes (ELDRIDGE, 1970; HODGSON, 1976; POTTS, 1990; POTTS and JORDAN, 1994). However, the genetic control of these abnormal phenotypes has been studied in very few cases as a result of long generation time and poor seed set following selfing. Such traits are of interest in forestry because easily scored mutants, controlled by single genes expressed at the seedling stage, allow a simple and cost-effective way to monitor outcrossing rates in open-pollinated seed (e.g. ELDRIDGE, 1970; HODGSON, 1976; POTTS, 1990).

During germination of *E. globulus* open-pollinated seed, two unrelated native trees were each found to produce some progeny that lacked chlorophyll (C. M. HARDNER, unpublished data). Both trees produced albinos that were discernable at the cotyledon stage and, initially, albino and non-albino phenotypes were equally vigorous. The albinos, however, were never observed to produce leaves and died a few weeks after germination. In one case (513) the albinos were yellow, and in the other (309) they were more pink, presumably the result of differing levels of anthocyanin. The aim of this study was to determine the genetic control of these albino phenotypes in *E. globulus* and to test their use in monitoring outcrossing rates.

Materials and Methods

The two trees described above (tree numbers 309 and 513) are located in southern Tasmania and are separated by a

---

[1]) Corresponding author (briony.patterson@utas.edu.au)