Theor. Appl. Genet. 58, 181-191 (1980)

Limits of Artificial Selection Under Unbalanced Mating Systems*

H. Kang

North Central Forest Experiment Station, Forestry Sciences Laboratory, US Department of Agriculture-Forest Service, Rhinelander, Wis. (USA)

G. Namkoong

Southeastern Forest Experiment Station, US Department of Agriculture-Forest Service and Genetics Department, North Carolina State University, Raleigh, N. C. (USA)

Summary. The influence of unbalanced mating systems – factorial mating (FM) and random loss of families after a full diallel crossing (RS) – on the ultimate probability of gene fixation (u(Π)) and the time required to fix or lose a gene (t(Π)) are investigated. The average u(Π) of these systems is smaller than that of random mating, and the range of u(Π) for a given initial parental genotype combination is very large (close to one for most initial genotypic combinations). The average u(II) of different parental genotypic combinations of a given gene frequency are different. These systems accelerate the t(Π).

Key words: Gene fixation – Factorial mating – Diallel crossing

Introduction

Random mating is one of the simplifying assumptions frequently used in theoretical studies of artificial selection models. Under a single-locus monoecious-species model the mating outcome, viz. the progeny genotypic frequency combination before selection, is in Hardy-Weinberg equilibrium. This assumption, however, is hard to realize in many breeding situations, and breeders are interested in finding out the cost of violating the random-mating assumption.

In many models the random mating assumption frequently implies the success of all the zygotes until the time of artificial selection. Therefore, it is convenient to consider the mating period as the time between mating and artificial selection, and the mating outcome as a function of all the factors that influence the population dynamics during that time.

Some of the factors which influence the mating outcome are correlated with the locus of selection interest by means of linkage or by the pleiotropic effects of the genes of the locus. Other factors such as artificial mating systems, random sampling of progenies or progeny families before selection, and dioecious species also influence the final outcome of the mating but are not correlated with the genotypes. Hill and Robertson (1968) showed that dioecious populations are not much different from monoecious populations in the expected fitness when heterozygote genotypes are favored. Kang and Namkoong (1979) showed that balanced artificial mating systems such as pair mating, disconnected diallel, and partial diallel do not influence the probability of gene fixation very much. The above two experiments represent similar situations in the sense that the matings do not alter the gene frequency of the population, and Kang and Namkoong concluded that as long as the mating does not change the gene frequency, the influence of the deviation of the genotypic frequency combination from Hardy-Weinberg equilibrium condition on ultimate probability of gene fixation is negligible.

This paper will discuss two situations where the gene frequencies are altered as the result of matings: (1) a factorial mating (Comstock and Robinson 1948) where the number of male testers is not equal to that of female and (2) random sampling of families after a full diallel crossing.

Models and Assumptions

Genetic Model and Assumptions

A genetic model of two alleles, A and a, at a single locus in a diploid population is used for this study. The standardized gene effect α (distance between the phenotypic values of homozygotes divided by the environmental standard deviation) is small. The three genotypes, AA, Aa,

* This article was written and prepared by U.S. Government employees on official time, and it is therefore in the public domain

0040-5752/80/0058/0181/$ 2.20

and aa, of this locus have genotypic values that are equal to their phenotypic means. The phenotypic values of the individuals are assumed to be identically and normally distributed around the genotypic mean. The sources of the error are the segregation at other loci and the environmental error. Because α is small, the entire population is assumed to be normally distributed. The genotypes of the population do not have to be in Hardy-Weinberg equilibrium. The population is closed and has discrete non-overlapping generations with constant population size (N). The influence of mutation on the population structure is assumed to be negligible.

Selection Model and Assumptions

A generation, which comprises one cycle of selection, is divided into three successive stages: mating, mass selection, and sampling. During the mating stage a total of N parents at time t are crossed randomly or in controlled design. All the progenies obtained from the crossing are allowed to develop until the time of artificial selection. The progeny population is assumed to be large, and natural selection is absent in the process. At the second stage a truncation selection is performed on the progeny population. The phenotypic performances of the members of the population are scored, and individuals with scores less than the truncation point are discarded. Since selection is applied on large populations, a constant truncation point which corresponds to a specific culling fraction is assumed to exist. Finally, N individuals to be used as parents at time t + 1 are sampled from the selected group.

The selected cycle is repeated in the same environment until the population reaches the selection limit. The selection intensity does not change with time.

Computational Procedures

Single Cycle of Selection

The consequences of a single cycle of selection can be described by a transition matrix (Q) with rows representing the genotypic combination of the parents at time t and columns representing that of parents at time t + 1. An element Q_{ij} of Q represents the transition probability of moving from i^{th} state to j^{th} state in one generation, where each state represents a unique combination of the numbers of three different genotypes AA, Aa, and aa. This probability is expressed as:

$$Q_{ij} = P(N_{1,t+1}, N_{2,t+1}, N_{3,t+1} \mid N_{1,t}, N_{2,t}, N_{3,t})$$

where N_1, N_2, and N_3 are arbitrary non-negative integers which satisfies the condition $N_1 + N_2 + N_3 = N$, and the subscripts 1, 2, and 3 refer to AA, Aa and aa individuals, respectively. In order to determine the transition probability Q_{ij} it is necessary to find all the events and their probabilities associated with each step in a complete selection cycle. The transition probability is then found by combining the probabilities of the events in the three steps.

Mating

The probability of obtaining a set of progeny genotypic frequency combinations given parents of $N_{1,t}$ AA, $N_{2,t}$ Aa, and $N_{3,t}$ aa, where t is time in generations, and given a mating scheme will be represented as

$$\theta'_k(q'_{k1}, q'_{k2}, q'_{k3} \mid N_{1,2}, N_{2,t}, N_{3,t}, \text{mating scheme}) \quad (1)$$

where $\{q'_{k1}, q'_{k2}, q'_{k3}\}$ is the k^{th} set of genotypic frequency combination in the progeny population before selection, and the prime implies that the statistics are obtained from the progeny population before selection.

Given a set of parental combinations $\{N_{1,t}, N_{2,t}, N_{3,t}\}$ the derived progeny set will vary depending on the mating system used in the model. The outcome of three different mating systems, random mating, factorial mating, and random sampling of families will be discussed.

Random Mating (RM)

In random mating, the proportion of the progeny genotypes is a function of parental gene frequencies. Assuming Hardy-Weinberg equilibrium in the progeny population, the frequencies of the three genotypes is uniquely determined by the parental gene frequency. When there are $X = 2N_1 + N_2$ A alleles among the parents at generation t, the progeny genotypic frequency combination is:

$$\{q'_1, q'_2, q'_3\} = \left\{ \left(\frac{X}{2N}\right)^2, \left(\frac{X}{N}\right) \left(1 - \frac{X}{2N}\right), \left(1 - \frac{X}{2N}\right)^2 \right\}$$

Factorial Mating (FM)

In factorial mating N parents are divided into two groups and matings are made between, but not within, the groups. For example, N individuals may be divided into n tester males and N – n females in a dioecious species. Given $N_{1,t}$ AA, $N_{2,t}$ Aa, and $N_{3,t}$ aa parents, a random subdivision of N parents into two groups whose sizes are n and m (= N – n) will also divide the three genotypes into two groups; n_1 AA, n_2 Aa, and n_2 aa will be in group 1 and m_1 AA, m_2 Aa, and m_3 aa will be in group 2 where $m_1 = N_{1,t} - n_1$, $m_2 = N_{2,t} - n_2$, and $m_3 = N_{3,t} - n_3$. Because

the subdivision of parents is made at random, n_1, n_2, and n_3 are variable, positive, integers for which $n_1 + n_2 + n_3 = n$. There are $(n + 1)(n + 2)/2$ different possible combinations of n_1, n_2, and n_3. We will use $\{n_{h1}, n_{h2}, n_{h3}\}$ to represent the genotypic combination of a particular set of n testers, where $h = 1, 2, 3, \ldots, (n + 1)(n + 2)/2$.

When the individuals in the two groups are crossed in the factorial manner, the number and the outcome of different crossings will be as follows.

Parental crossing type	Number of crossings	Mating outcome Progeny genotype AA	Aa	aa
AA X AA	$n_{h1}m_{h1}$	1	0	0
AA X Aa	$n_{h1}m_{h2} + n_{h2}m_{h1}$	1/2	1/2	0
AA X aa	$n_{h1}m_{h3} + n_{h3}m_{h1}$	0	1	0
Aa X Aa	$n_{h2}m_{h2}$	1/4	1/2	1/4
Aa X aa	$n_{h2}m_{h3} + h_{h3}m_{h2}$	0	1/2	1/2
aa X aa	$n_{h3}m_{h3}$	0	0	1

Let M_h be a 6 × 1 vector which represents the values in column 2 and C be a 6 × 3 matrix of values in columns 3-5.

Then,

$$\{q'_{h1}, q'_{h2}, q'_{h3}\} = M_h^T C/lM_h$$

where

l = 1 × 6 vector of 1s, and

T = transpose.

The probability of sampling a particular tester genotypic combination $\{n_{h1}, n_{h2}, n_{h3}\}$ from parents of $\{N_{1,t}, N_{2,t}, N_{3,t}\}$ is:

$$\frac{\binom{N_{1,t}}{n_{h1}}\binom{N_{2,t}}{n_{h2}}\binom{N_{3,t}}{n_{h3}}}{\binom{N}{n}}$$

Although there is only one progeny genotypic frequency combination which corresponds to a given tester genotypic combination, matings performed using different tester genotypic combinations could result in the same progeny genotypic frequency combination, say $\{q'_{k1}, q'_{k2}, q'_{q3}\}$.

Therefore, the probability of obtaining $\{q'_{k1}, q'_{k2}, q'_{k3}\}$ is the sum of the probabilities of sampling from the parents the tester genotypic combinations which produce $\{q'_{k1}, q'_{k2}, q'_{k3}\}$, and

$$\theta'_k\ (q'_{k1}, q'_{k2}, q'_{k3} \mid N_{1,t}, N_{2,t}, N_{3,t}, \text{factorial mating}) = \frac{1}{\binom{N}{n}} \sum_{\ell=1}^{L} \binom{N_{1,t}}{n_{\ell 1}}\binom{N_{2,t}}{n_{\ell 2}}\binom{N_{3,t}}{n_{\ell 3}},$$

where L represents the number of the tester genotypic combinations which result in $\{q'_{k1}, q'_{k2}, q'_{k3}\}$, and ℓ indexes such tester genotypic combinations.

Random Sampling of Families (RS)

After a full diallel crossing is made on a monoecious species a few families are randomly sampled, and only those sampled families are used as the basis for artificial selection. The sample families are assumed to be large and equal in size. This scheme is an extreme form of random family size variation in which there are only two types of family sizes; extremely large or none.

Only one crossing combination exists in a full diallel crossing scheme given $\{N_{1,t}, N_{2,t}, N_{3,t}\}$. The number of crosses for each of the six different types are as follows:

Index (g)	Crossing type	Number of crossings
1	AA X AA	$N_{1,t}^2 = M_1$
2	AA X Aa	$2N_{1,t}N_{2,t} = M_2$
3	AA X aa	$2N_{1,t}N_{3,t} = M_3$
4	Aa X Aa	$N_{2,t}^2 = M_4$
5	Aa X aa	$2N_{2,t}N_{3,t} = M_5$
6	aa X aa	$N_{3,t}^2 = M_6$

If we randomly sample y families from the total of N^2 families, y_g families will come from g^{th} crossing type, where $y = \sum_{g=1}^{6} y_g$. Let Y_h be a 6 × 1 vector of y_{hg} obtained as the result of a particular sampling, say h. Then,

$$\{q'_{h1}, q'_{h2}, q'_{h3}\} = Y_h^T C/lY_h$$

The probability of obtaining sample h is:

$$\frac{\binom{M_1}{y_{h1}}\binom{M_2}{y_{h2}}\binom{M_3}{y_{h3}}\binom{M_4}{y_{h4}}\binom{M_5}{y_{h5}}\binom{M_6}{y_{h6}}}{\binom{N^2}{y}}$$

As was the case in factorial mating, the probabilities of obtaining samples which result in the same progeny genotypic combination, say $\{q'_{k1}, q'_{k2}, q'_{k3}\}$, can be added together such that

$$\theta'_k\ (q'_{k1}, q'_{k2}, q'_{k3} \mid N_{1,t}, N_{2,t}, N_{3,t}, \text{Random Sampling of Families}) = \frac{1}{\binom{N^2}{y}} \sum_{\ell=1}^{L} \binom{M_1}{y_{\ell 1}}\binom{M_2}{y_{\ell 2}}\binom{M_3}{y_{\ell 3}}\binom{M_4}{y_{\ell 4}}\binom{M_5}{y_{\ell 5}}\binom{M_6}{y_{\ell 6}}$$

Selection

A truncation selection applied to the progeny population will produce a new set of genotypic frequency combinations', say $\{q_{k1}, q_{k2}, q_{k3}\}$. The selection is made on a conceptually infinite population, and a fixed truncation point, τ, is assumed to exist that satisfies the following equation (Hill 1969).

$$q'_{k1}\,\Phi(-\tau+\alpha/2)+q'_{k2}\,\Phi(-\tau+\alpha.d/2)+q'_{k3}\,\Phi(-\tau-\alpha/2)=F,$$

where $\Phi(\cdot)$ is the distribution function of standardized normal distribution, N(0,1),
d is the degree of dominance, and
F is the fraction selected.

It is convenient to write the above formula as:

$$\frac{q'_{k1}}{F}\,\Phi(-\tau+\alpha/2)+\frac{q'_{k2}}{F}\,\Phi(-\tau+\alpha.d/2)+\frac{q'_{k3}}{F}\,\Phi(-\tau-\alpha/2)=1.$$

The three terms on the left of the equal sign represent q_{k1}, q_{k2}, and q_{k3}, respectively.

Since the mode of selection is deterministic the probability of obtaining $\{q_{k1}, q_{k2}, q_{k3}\}$ from $\{q'_{k1}, q'_{k2}, q'_{k3}\}$ is

$$\theta_k\,(q_{k1}, q_{k2}, q_{k3} \mid N_{1,t}, N_{2,t}, N_{3,t}, \text{Mating System}) = \theta'_k\,(q'_{k1}, q'_{k2}, q'_{k3} \mid N_{1,t}, N_{2,t}, N_{3,t}, \text{Mating System}).$$

Sampling

The probability (p) of randomly sampling $\{N_{1,t+1}, N_{2,t+1}, N_{3,t+1}\}$ from $\{q_{k1}, q_{k2}, q_{k3}\}$ follows a multinomial distribution. Thus,

$$p_k(N_{1,t+1}, N_{2,t+1}, N_{3,t+1} \mid q_{k1}, q_{k2}, q_{k3}) = \binom{N}{N_{1,t+1}\ N_{2,t+1}\ N_{3,t+1}} (q_{k1})^{N_{1,t+1}}\,(q_{k2})^{N_{2,t+1}}\,(q_{k3})^{N_{3,t+1}}$$

Synthesis

Let there be K distinct progeny genotypic frequency combinations after selection $\{q_{k1}, q_{k2}, q_{k3}\}$. Each combination will occur with the probability $\theta_k = \theta'_k$, given $\{N_{1,t}, N_{2,t}, N_{3,t}\}$, and from each combination the parents for the next generation $\{N_{1,t+1}, N_{2,t+1}, N_{3,t+1}\}$ will be sampled with the probability p_k. Therefore, the transition probability Q_{ij} is obtained by summing the K products of the two probabilities:

$$P(N_{1,t+1}, N_{2,t+1}, N_{3,t+1} \mid N_{1,t}, N_{2,t}, N_{3,t}) = \sum_{k=1}^{K} \theta'_k \cdot p_k$$

For random mating,

$$Q_{ij} = \binom{N}{N_{1,t+1}\ N_{2,t+1}\ N_{3,t+1}} (q_{k1})^{N_{1,t+1}}\,(q_{k2})^{N_{2,t+1}}\,(q_{k3})^{N_{3,t+1}}$$

For factorial mating,

$$Q_{ij} = \frac{N_{1,n+1}!\ N_{2,n+1}!\ N_{3,n+1}!}{(N-n)!} \sum_{\ell=1}^{L} \binom{N_{1,t}}{n_{\ell 1}} \binom{N_{2,t}}{n_{\ell 2}} \binom{N_{3,t}}{n_{\ell 3}} \times \sum_{k=1}^{K} (q_{k1})^{N_{1,t+1}}\,(q_{k2})^{N_{2,t+1}}\,(q_{k3})^{N_{3,t+1}}.$$

For random sampling of families,

$$Q_{ij} = \frac{\binom{N}{N_{1,t+1}\ N_{2,t+1}\ N_{3,t+1}}}{\binom{N^2}{y}} \cdot \sum_{\ell=1}^{L} \binom{M_1}{x_{\ell 1}} \binom{M_2}{x_{\ell 2}} \binom{M_3}{x_{\ell 3}} \binom{M_4}{x_{\ell 4}} \binom{M_5}{x_{\ell 5}} \binom{M_6}{x_{\ell 6}} \times \sum_{k=1}^{K} (q_{k1})^{N_{k,t+1}}\,(q_{k2})^{N_{2,t+1}}\,(q_{k3})^{N_{3,t+1}}.$$

Transition Probabilities – Maximum and Minimum

For simplicity, we will use $\{\theta'_k\}$ to represent $\{q'_{k1}, q'_{k2}, q'_{k3}\}$ and $\{\theta_k\}$ to represent $\{q_{k1}, q_{k2}, q_{k3}\}$. Contrary to random mating where only one progeny genotypic frequency combination (Hardy-Weinberg equilibrium) is produced from mating, factorial mating (FM) and random sampling of families (RS) yield more than one possible progeny genotypic frequency combination $\{\theta'_k\}$. Therefore in these mating systems the Q_{ij} is the average probability of obtaining j^{th} parental combination at t + 1 from all the possible $\{\theta'_k\}$ of i^{th} parental combination of t. After selection, the genotypic frequency combination will be changed to produce $\{\theta_k\}$, and it is possible to rank gene frequencies of $\{\theta_k\}$. The $\{\theta_k\}$ with the largest or the smallest gene frequency will be represented as $\{\theta_{max}\}$ or $\{\theta_{min}\}$, respectively. It is interesting to see the consequences of repeatedly choosing $\{\theta_{max}\}$ or $\{\theta_{min}\}$ in a recurrent selection scheme.

The transition probability Q_{ijmax} is:

$$\binom{N}{N_{1,t+1}\ N_{2,t+1}\ N_{3,t+1}} (q_{max\ 1})^{N_{1,t+1}}\,(q_{max\ 2})^{N_{2,t+1}}\,(q_{max\ 3})^{N_{3,t+1}}.$$

where $q_{max\ 1}$ = freq (AA) of the maximum combination after selection, $q_{max\ 2}$ = freq (Aa) of the maximum combination after selection,

and $q_{max\ 3}$ = freq (aa) of the maximum combination after selection. Q_{ijmin} is obtained by replacing the subscript max with min.

Ultimate Probability of Gene Fixation and Time Required to Reach Fixation or Loss

Because the generations are discrete, transition probabilities are independent of time t, and the selection criteria do not change, recurrent selection follows a Markov process. The vector of the ultimate probabilities of gene fixation is obtained by (Carr and Nassar 1970a):

$$\mu(\Pi) = (I-P)^{-1} r$$

where I is a (H–2) × (H–2) identity matrix, P is a (H–2) × (H–2) matrix of transient states, r is a (H–2) × 1 vector of transition probabilities from transient states to the fixation state, Π is the initial parental genotypic combination, and H is the total number of possible genotypic frequency combination given N.

The time required to reach fixation or loss is obtained by (Carr and Nassar 1970b):

$t(\Pi) = (I-P)^{-1} l$ where l is a (H–2) × 1 vector of ls. The notations $\overline{u(\Pi)}$ and $\overline{t(\Pi)}$ will be used to represent the statistics obtained from the transition matrix of average probabilities and $u(\Pi)_{max\ (or\ min)}$ and $t(\Pi)_{max\ (or\ min)}$ will be used for those obtained from maximum or minimum probabilities.

Results

The primary objective of this study is to determine the influence of mating systems on a single cycle of selection and on the selection limit. Therefore, the discussion of results is structured such that the feature which distinguishes the mating outcomes of unbalanced matings from those of balanced matings is identified first. Then, the combined effect of mating, selection, and sampling on a single cycle of selection is discussed. Finally, the influence of mating system on selection limit is presented.

Table 1. Progeny genotypic combinations and related statistics. N = 8, factorial mating, parental gene frequency (q)

No. of testers (No. crossings)	Index (k)	Gene frequency (q'_k)	Genotypic frequency			
			q'_{k1}	q'_{k2}	q'_{k3}	$\theta'_k(\cdot)$
1	1	0.28571	0	0.57143	0.42857	0.25
(7)	2	0.5	0.25	0.5	0.25	0.5
	3	0.71429	0.42857	0.57143	0	0.25
	$\overline{q'}$	0.5				
	v(q')	0.023				
	1	0.33333	0	0.66667	0.33333	0.03571
	2	0.41667	0.14583	0.54167	0.3125	0.28571
2	3	0.5	0.25	0.5	0.25	0.35714
(12)	4	0.58333	0.3125	0.54167	0.14583	0.28571
	5	0.66667	0.33333	0.66667	0	0.03571
	$\overline{q'}$	0.5				
	v(q')					
	1	0.43333	0.11667	0.63333	0.25	0.07143
	2	0.46667	0.2	0.53333	0.26667	0.25
3	3	0.5	0.25	0.5	0.25	0.35714
(15)	4	0.53333	0.26667	0.53333	0.2	0.25
	5	0.56667	0.25	0.63333	0.11667	0.07143
	$\overline{q'}$	0.5				
	v(q')	0.004				
4	1	0.5	0.1875	0.625	0.1875	0.17143
(16)	2	0.5	0.23438	0.53125	0.23438	0.45714
Partial	3	0.5	0.25	0.5	0.25	0.37143
Diallel	$\overline{q'}$	0.5				
	v(q')	0				

$\theta'_k(\cdot)$ = Probability of obtaining k^{th} progeny genotypic combination, $\{q'_{k1}, q'_{k2}, q'_{k3}\}$; $\overline{q'}$ = average progeny gene frequency; v(q') = variance of progeny gene frequency

Mating Outcome

From a set of parents the mating systems, factorial mating (FM) and random sampling of families (RS) can generate many progeny genotypic combinations $\{\theta'_k\}$ with different gene frequencies (Tables 1 and 2). Variance in progeny gene frequency (V(q')) distinguishes these mating systems from balanced matings. The variance results from the unbalanced representation of parents during the mating period. The average progeny gene frequency $\overline{(q')}$ of all possible mating outcomes is the same as the parental gene frequency (q).

Variance and range of progeny frequency decrease as the number of testers (or number of families sampled) increase. When N = 4, the variance of FM is comparable to that of RS with four crossings (Table 2). Because only three crossings are involved in FM, the small difference in progeny gene frequency variation v(q') between the two

Table 2. Progeny genotypic combinations and related statistics[a]. N = 4, FM and RS, q = 0.5, Π = (1, 2, 1)

Mating system	Sample size	Index (k)	Gene frequency (q'_k)	Genotypic frequency			
				q'_{k1}	q'_{k2}	q'_{k3}	$\theta'_k(\cdot)$
		1	0.33333	0	0.66667	0.33333	0.25
		2	0.5	0.25	0.5	0.25	0.5
FM	1 tester	3	0.66667	0.33333	0.66667	0	0.25
		$\overline{q'}$	0.5				
		v(q')	0.014				
		1	0	0	0	1	0.0625
		2	0.25	0	0.5	0.5	0.25
	1 family	3	0.5	0.25	0.5	0.25	0.25
		4		0	0.1	0	0.125
		5	0.75	0.5	0.5	0	0.25
		6	1	1	0	0	0.0625
		$\overline{q'}$	0.5				
		v(q')	0.0625				
		1	0.125	0	0.25	0.75	0.03333
		2	0.25	0	0.5	0.5	0.06667
		3		0.125	0.25	0.625	0.03333
		4	0.375	0.125	0.5	0.375	0.13333
	2 family	5		0	0.75	0.25	0.06667
		6		0.25	0.25	0.5	0.03333
RS		7	0.5	0.5	0	0.5	0.00833
		8		0.25	0.5	0.25	0.18333
		9		0.125	0.75	0.125	0.06667
		10		0	0.1	0	0.00833
		11	0.625	0.5	0.25	0.25	0.03333
		12		0.375	0.5	0.125	0.13333
		13		0.25	0.75	0	0.06667
		14	0.75	0.675	0.25	0.125	0.03333
		15		0.5	0.5	0	0.06667
		16	0.875	0.75	0.25	0	0.03333
		$\overline{q'}$	0.5				
		v(q')	0.029				
		1	0.1875	0	0.375	0.625	0.0022
		37	0.8125	0.625	0.375	0	0.0022
	4 family	$\overline{q'}$	0.5				
		v(q')	0.012				
		1	0.25	0	0.5	0.5	0.00025
		55	0.75	0.5	0.5	0	0.00025
	6 family	$\overline{q'}$	0.007				

[a] Note that there are 37 and 55 different combinations for RS (4 family) and RS (6 family), respectively

Table 3. Probabilities of losing a favorable allele (A) when crossings are randomly sampled after a full diallel[a]

Π			q	Number of samples (families)			
(AA	Aa	aa)		1	2	4	6
0	1	3	0.125	0.5625	0.3	0.06923	0.01049
0	2	2	0.25	0.25	0.05	0.00055	0
0	3	1	0.375	0.0625	0	0	0
0	4	0	0.5	0	0	0	0
1	0	3	0.25	0.5625	0.3	0.06923	0.01049
1	1	2	0.375	0.25	0.05	0.00055	0
1	2	1	0.5	0.0625	0	0	0
1	3	0	0.625	0	0	0	0
2	0	2	0.5	0.25	0.05	0.00055	0
2	1	1	0.625	0.0625	0	0	0
2	2	0	0.75	0	0	0	0
3	0	1	0.75	0.0625	0	0	0
3	1	0	0.875	0	0	0	0

[a] The probabilities of fixing an undesirable allele are the same if A is assumed to be an inferior allele

systems suggests that FM generates smaller variance than RS for a given number of families sampled. A similar conclusion may be made when the ranges of extreme outcomes are compared. When N = 4, the gene frequency range of FM is smaller than that of RS (Table 2). In FM, the variance and range of the outcome seem to depend more on the degree of balanced representation of parents than the number of crossings. For example, when N = 8 (one tester), seven crossings are made but when N = 4 only three crossings are possible. However, $V(q')$ and range of FM is smaller when N = 4 (Tables 1 and 2).

In factorial mating all the parents are represented in the crossings at least once. Although the progeny gene frequency may change due to the unbalanced representation of parents, the systematic mating control in FM prevents the loss of alleles due to crossing. Sampling of families in RS is not restricted and it is possible to have a set of samples that exclude some parents. Therefore, it is possible to lose alleles in this scheme. If a certain fraction of progeny families are lost repeatedly under recurrent selection with RS, the loss probability would influence the population in a manner similar to that of random drift. In this case the sampling of families is an indirect way of limiting the population size, and there are two stages of random drift in a selection cycle of RS.

Contrary to what one might expect, the probability of losing an allele is not directly related to q, but with the number of inferior homozygotes (Table 3). For example, regardless of the q, if the number of aa genotypes is one and one family is sampled, the probability of loss due to mating is 0.0625. However, the average loss probability will depend on q when these probabilities are averaged for different parental genotype combinations of a given q. The loss probability decreases as sample size increases.

The Influence of Mating on a Single Cycle of Selection

Average Statistics

Given the mating system and population size, the average gene frequency of the progeny population increases as the number of testers (or sampled families) increases and the variance of gene frequency and probability of loss (or fixation) decreases (Table 4). It is not surprising to observe the decreases in variance and loss probability as the number of families (or testers) increases. But the change in the average is an unexpected result because the sample size in general does not influence the mean. This change seems to be due to the combined effect of the variable mating outcomes and selection.

When more than one mating outcome is possible, the amount of gene frequency advance due to selection (Δ_q) will vary depending on the specific mating outcome. The average gene frequency advance ($\overline{\Delta_q}$) of the mating system can be calculated from the set of Δ_q which corresponds to all possible mating outcomes. Under an additive gene model the gene frequency advance due to selection is maximum when the initial gene frequency is 0.5. Therefore, Δ_q obtained from initial gene frequencies other than 0.5 will be less than maximum. Regardless of the distributional properties of the progeny gene frequency of mating outcomes, the average gene frequency advance will be smaller than Δ_q obtained from a single genotypic combination with $q' = 0.5$. Because both FM and RS generate more than one gene frequency other than 0.5, $\overline{\Delta_q}$ will be smaller than Δ_q obtained from a random mating when $q' = 0.5$ (or the advance of the conceptual average gene frequency $\Delta_{\overline{q}'} = 0.5$).

When $\overline{q}'$ is not 0.5, the distribution of progeny gene frequencies becomes an important factor. For example, when $\overline{q}' < 0.5$ and the distribution is skewed to the left, $\Delta_{\overline{q}'} \geqq \overline{\Delta_{q'}}$. But when $\overline{q}' < 0.5$ and the distribution is skewed right, $\Delta_{\overline{q}'} \leqq \overline{\Delta_{q'}}$. The distributions of progeny gene frequency of FM and RS, however, tend to be skewed to the right when $\overline{q}'$ is smaller than 0.5 and to the left when $\overline{q}'$ is larger than 0.5. The above distributional properties of FM and RS seem to have caused the values of $\overline{\Delta q'}$ to be smaller than those of $\Delta_{\overline{q}'}$ for all possible parental gene frequencies. The reduction in average gain seems to be most serious at the intermediate gene frequencies (Fig. 1).

Maximum and Minimum Statistics

When a RS system has only one sample, the probability of obtaining a genotype combination with smallest gene frequency is 0.0625 (Table 2). Although the probability is small, it is possible to obtain such minimum combinations,

Table 4. Summary of average transition probabilities when q = 0.5, Π = (1, 2, 1) or (2, 4, 2), α = 0.2, F = 1, and additive gene action

Population size	Mating system	Sample size	$\bar{q}$	V(q)	p(loss)/gen	p(fix)/gen	Range ($\bar{q}_{max}-\bar{q}_{min}$)
4	Random mating	–	0.54374	0.03095	0.00186	0.00759	–
	Factorial mating	1 tester	0.53161	0.03648	0.00238	0.00864	0.33415
	Random sampling	1 family	0.52188	0.07823	0.07379	0.08626	1
		2 families	0153355	0.05298	0.01864	0.03630	0.74895
		4 families	0.53937	0.04038	0.0065	0.01803	0.62434
		6 families	0.54132	0.03618	0.00405	0.01303	0.5066
8	Random mating	–	0.54374	0.01547	346E-8	567E-7	–
	Factorial mating	1 tester	0.53259	0.03466	125E-6	643E-6	0.42897
		2 testers	0.53855	0.01963	109E-7	113E-6	0.33416
		3 testers	0.54022	0.01544	328E-8	488E-7	0.13324
		4 testers	0.54061	0.01441	229E-8	371E-7	0.03563

$\bar{q} = \sum_{j=1}^{H} q_j Q_{ij}$; $V(q) = \sum_{j=1}^{H} (q_j-\bar{q})^2 Q_{ij}$, and $\bar{q}_{max(or\ min)} = \sum_{j=1}^{H} q_j Q_{ij}\ max(or\ min)$

and we may observe the average range of possible outcomes of a complete cycle of selection by comparing maximum and minimum statistics (Table 4). Given a system of mating, the average range of gene frequency ($\bar{q}_{max}-\bar{q}_{min}$) decreases as the number of testers (or families) increases, and FM have smaller ranges than RS. These trends are consistent with those observed in the discussion of the mating outcome, and the average range of gene frequency after a complete cycle of selection is primarily influenced by the mating outcome.

The average ranges of unbalanced mating systems are large when compared to those of balanced mating systems. For example, the range for the partial diallel (factorial with 4 testers) is 0.03563 (N = 8), and that for pair mating is .02184 (N = 4). The variances, which are not shown in Table 4, decrease as the number of testers (or families) decreases. This is because the gene frequencies of maximum or minimum mating outcomes become more extreme as the number of testers (or families) decrease. This trend is more obvious in RS.

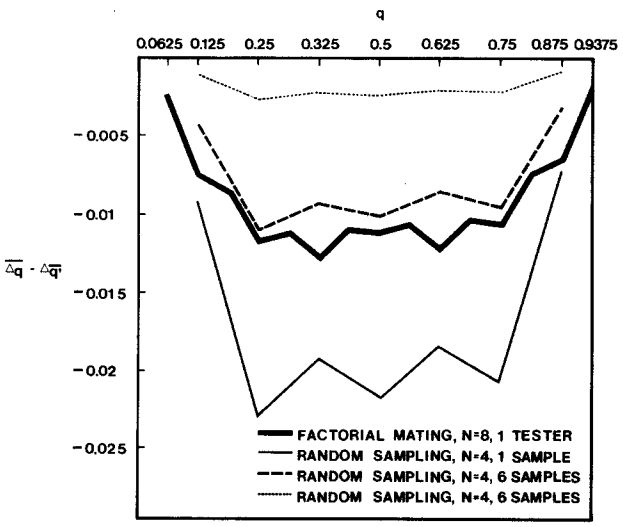

Fig. 1. The difference between the average gene frequency advance (Δ_q) and the advance of average gene frequency ($\Delta_{q'}$) at all possible parental gene frequencies (q). Note the exaggerated ordinate scales

The Influence of Mating on Selection Limits

When N = 4, α = 0.2, and F = 0.1 in unbalanced mating systems, the average ultimate probability of gene fixation $\overline{u(\Pi)}$ generally is depressed and the percent of the depression is greater at lower initial gene frequency (Table 5, Fig. 2). The differences between unbalanced mating systems are relatively large compared to those observed in balanced mating systems. The largest difference in $\overline{u(\Pi)}$ observed in any possible comparisons of balanced mating systems of all possible initial gene frequencies is 0.00026, which is the difference between random mating and the half diallel (or partial diallel) when II = (1,0,3) (Kang and Namkoong 1979). The amount of reduction in $\overline{u(\Pi)}$ under RS decreases as the number of families is increased. As was the case with the variance of progeny gene frequency V(q'), the performance in $\overline{u(\Pi)}$ of FM with 1 tester (N = 4) was similar to that of RS with 4 families. In factorial mating, the percent reduction in $\overline{u(\Pi)}$ under additive gene action tends to be greater than that under complete dominance. The opposite trend is observed in RS (Fig. 2).

The magnitude of the V(q') is determined not only by the mating system and the parental gene frequency but also by the proportions of the parental genotypes. For example, when q = 0.5 and only one genotype (Aa) is available among the parents, the progeny gene frequency has no variance. But the variance will increase as the parents become more homozygous – the more homozygous the founders, the lower the $\overline{u(\Pi)}$ (Fig. 3).

These results, in conjunction with the previous observation that the loss probability of a favorable allele due to

Table 5. $\overline{u(\Pi)}$ and $\overline{t(\Pi)}$ of factorial mating design (FM) and random samplings of families after a random mating (RS). N = 4, α = 0.2, F = 0.1, additive

Π			q	RM	FM-RM 1 tester	RS-RM Sample size			
						1	2	4	6
$\overline{u(\Pi)}$									
0	0	4	0	0	0	0	0	0	0
0	1	3	0.125	0.31575	−0.06044	−0.13990	−0.09905	−0.05642	−0.03501
0	2	2	0.25	0.53761	−0.08043	−0.19701	−0.13479	−0.07423	−0.04533
0	3	1	0.375	0.69357	−0.07704	−0.10033	−0.13210	−0.07062	−0.04253
0	4	0	0.5	0.80324	−0.06122	−0.16646	−0.10761	−0.05625	−0.03353
1	0	3	0.25	0.53761	−0.09814	−0.23632	−0.15495	−0.09347	−0.05763
1	1	2	0.375	0.69357	−0.09187	−0.23591	−0.15832	−0.08541	−0.05161
1	2	1	0.5	0.80324	−0.07361	−0.20032	−0.12948	−0.06755	−0.04020
1	3	0	0.625	0.88042	−0.05024	−0.14334	−0.09041	−0.94619	−0.02724
2	0	2	0.5	0.80324	−0.08600	−0.23485	−0.15210	−0.07917	−0.04701
2	1	1	0.625	0.88042	−0.06057	−0.17515	−0.10912	−0.05501	−0.03222
2	2	0	0.75	0.93475	−0.03443	−0.10369	−0.06340	−0.03150	−0.01834
3	0	1	0.75	0.83475	−0.04302	−0.13346	−0.07927	−0.03834	−0.02204
3		0	0.875	0.97302	−0.01715	−0.05432	−0.03209	−0.01548	−0.00889
4	0	0	1	1		0	0	0	0
$\overline{t(\Pi)}$									
0	0	4	0	0	0	0	0	0	0
0	1	3	0.125	6.2684	−0.8528	−3.6079	−2.4910	−1.3836	−0.8487
0	2	2	0.25	8.5416	−0.8858	−4.6703	−3.1028	−1.6666	−1.0061
0	3	1	0.375	9.1076	−0.6568	−4.4761	−2.8852	−1.5071	−0.8970
0	4	0	0.5	8.6523	−0.4004	−3.7084	−2.3304	−1.1900	−0.7019
1	0	3	0.25	8.5416	−1.1550	−5.6876	−3.9222	−2.1464	−1.3049
1	1	2	0.375	9.1076	−0.7945	−5.2887	−3.4410	−1.8043	−1.0762
1	2	1	0.5	8.6523	−0.4581	−4.3181	−2.7140	−1.3826	−0.8139
1	3	0	0.625	7.5310	−0.2167	−3.1312	−1.9249	−0.9620	−0.5617
2	0	2	0.5	8.6523	−0.5158	−5.1803	−3.1972	−1.6037	−0.9374
2	1	1	0.625	7.5310	−0.2351	−3.7896	−2.2809	−1.1217	−0.6500
2	2	0	0.75	5.8976	−0.0767	−2.3365	−1.3922	−0.6781	−0.3915
3	0	1	0.75	5.8976	−0.0948	−3.0436	−1.7562	−0.8369	−0.4787
3	1	0	0.875	3.7602	−0.0159	−1.3324	−0.7726	−0.3684	−0.2109
4	0	0	1	0	0	0	0	0	0

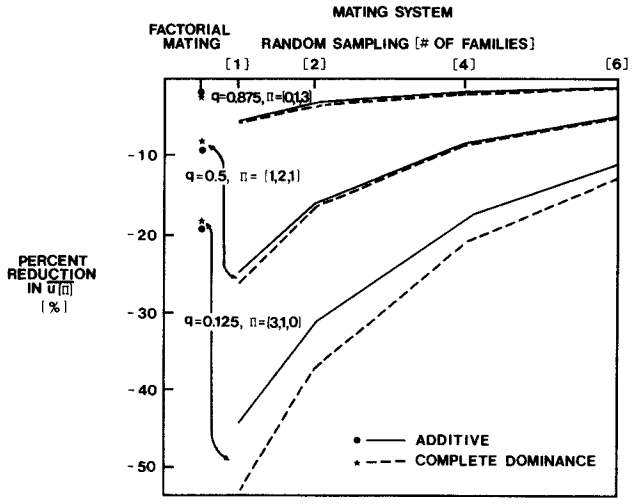

Fig. 2. Relative reduction of $\overline{u(\Pi)}$ in unbalanced mating systems at two levels of degree of dominance (compared to random mating)

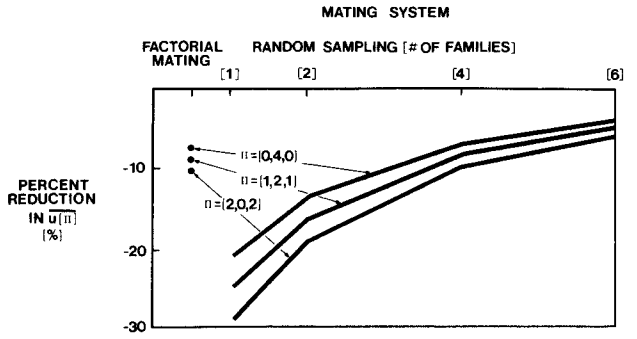

Fig. 3. Reduction of $\overline{u(\Pi)}$ in three different genotypic combinations having the same gene frequency (q = 0.5)

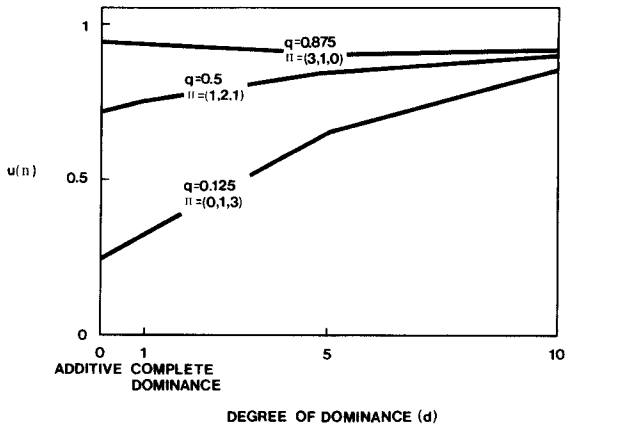

Fig. 4. Ultimate probability of gene fixation $\overline{u(\Pi)}$ under fractorial mating at various degrees of dominance

RS is a function of the number of inferior homozygotes of the parents rather than the parental gene frequency, leads one to appreciate the importance of the genotypic structure of the founders under unbalanced mating systems. Kang and Namkoong (1979) found that under balanced mating systems the difference in u(Π) among different parental genotypic combinations which have the same gene frequency, (Π|q) is negligible ($< 10^{-4}$). Because V(q′) is zero in balanced mating systems, the lack of differences is readily explained, and it is worthwhile emphasizing that the genotypic combination structure is important only when the matings are unbalanced.

The differences in $\overline{u(\Pi)}$ among different (Π|q) under balanced mating systems, however, have never been found to be greater than those found among different q. For example, the $\overline{u(\Pi)}$ of the most heterozygous Π with q = 0.5 is smaller than that of the most homozygous Π with q = 0.625 (N = 4). The lack of overlap in $\overline{u(\Pi)}$ among genotypic combinations of different gene frequencies could be due to the limited number of experiments that have been observed.

In the factorial mating system the differences among various initial gene frequencies become smaller as the degree of dominance increases (Fig. 4). Strong selection for the heterozygote genotype under overdominance gene action seems to reduce the impact of random loss of the favorable allele with small initial gene frequencies. The fixation probability is initially reduced for alleles with a high initial frequency (q = 0.875) before it increases again. Overdominance gene action not only increases the $\overline{u(\Pi)}$ with low initial frequency in an absolute sense, but also reduces the differences between random mating and factorial mating.

The rate of loss (or fixation) under FM becomes faster than that of random mating as the degree of dominance increases (Fig. 5). This is contrary to the intuition that if the difference in $\overline{u(\Pi)}$ between RM and FM becomes smaller because of the reduced loss probability, the time difference should become smaller also. It is interesting to contrast the time required to reach homozygosity $\overline{t(\Pi)}$ with FM with that of pair mating (PM). Although selfings are excluded in both FM and PM, FM is faster than RM and RM is faster than PM (Fig. 5). The exclusion of selfings in PM generates a set of progeny genotypic combinations whose average structure is more heterozygous than the Hardy-Weinberg combination. The retardation is due to the difference in genotypic proportions for a given gene frequency. In FM, on the other hand, the acceleration is primarily due to the variance in progeny gene frequency.

Most of the ranges of $[u(\pi)_{max} - u(\pi)_{min}]$ in unbalanced mating systems are close to one. This is opposite to the results observed in balanced mating systems, where largest observed range was 2.8×10^{-4} (Kang and Namkoong 1979). Factorial mating has a somewhat smaller range than RS.

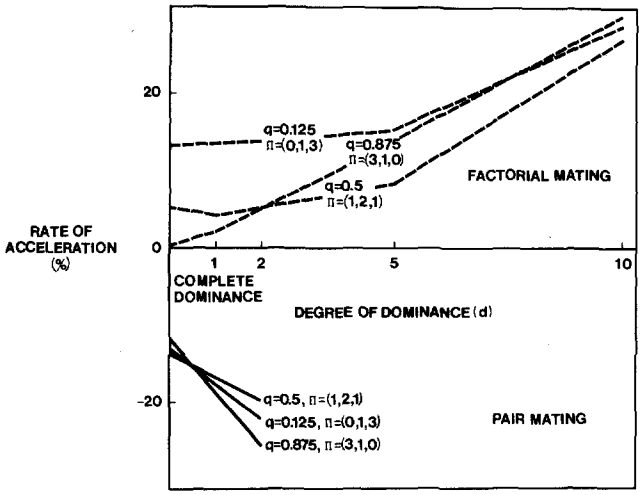

Fig. 5. Rate of acceleration in the average fixation (time) of an allele under factorial and pair mating at various degrees of dominance. The speed under random mating is used as the standard

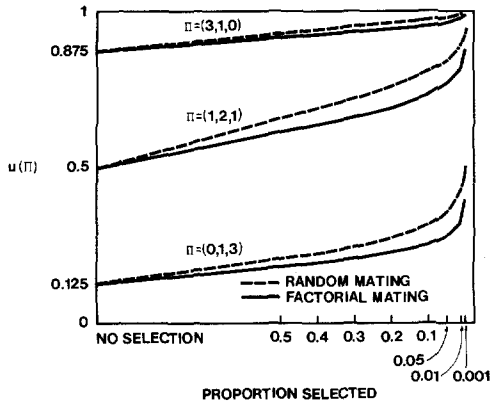

Fig. 6. The comparison of $\overline{u(\Pi)}$ between random mating and factorial mating at various levels of selection proportion: N = 4, $\alpha = 0.2$, N = 4, additive

Discussion

Unbalanced mating systems differ from random mating in several important ways: (1) $\overline{u(\Pi)}$ are smaller, (2) $\overline{u(\Pi)}$ of different parental genotype combinations of a given gene frequency ($\Pi|q$) are different and (3) the ranges of $u(\Pi)$ are large. Variance in progeny gene frequency $V(q')$ is the factor that is primarily responsible for the above outcomes. When combined with selection $V(q')$ will lower $\overline{u(\Pi)}$ as well as cause differences in $\overline{u(\Pi)}$ among different ($\Pi|q$). Without selection $\overline{u(\Pi)}$ is not depressed, but as the selection intensity increases, so does the amount of depression of $\overline{u(\Pi)}$ (Fig. 6). The $\overline{u(\Pi)}$ is, of course, influenced by the population size (N). The influence of population size on $\overline{u(\Pi)}$ was not discussed because we were primarily interested in observing the impact of having different mating systems given N.

In most breeding populations, family sizes are unequal. Variation in family size could occur as the result of genetic correlation of the trait of interest with those that influence family size. Under this type of situation the ultimate probability of gene fixation will vary depending on the sign of the genetic correlation. It is, however, not unusual to find a situation where family size varies at random, and the population size from which the selection is made is small. Therefore, one may expect that the average ultimate probability of gene fixation under this type of situation will be smaller than that obtained from balanced mating.

Acknowledgement

We thank Drs. R. Moll, P. Smouse, and referees for their critical review of the paper.

Literature

Carr, R.N.; Nassar, R.F. (1970a): Effects of selection and drift on the dynamics of finite populations. 1: Ultimate probability of fixation of a favorable allele. Biometrics **26**, 41-49

Carr. R.N.; Nassar, R.F. (1970b): Effects of selection and drift on the dynamics of finite populations. 2: Expected time to fixation or loss of a favorable allele. Biometrics **26**, 221-228

Comstock, R.E.; Robinson, H.F. (1948): The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance. Biometrics **4**, 254-266

Hill, W.G. (1969): On the theory of artificial selection in finite populations. Genet. Res. **12**, 143-163

Hill, W.G.; Robertson, A. (1968): The effect of inbreeding at loci with heterozygote advantage. Genetics **8**, 269-294

Kang, H.; Namkoong, G. (1979): Limits of artificial selection under balanced mating systems. Silvae Genetica **28**, 53-60

Received April 29, 1979
Accepted May 6, 1980
Communicated by R.W. Allard

Dr. H. Kang
North Central Forest Experiment Station
Forestry Sciences Laboratory,
U.S. Department of Agriculture-Forest Service
Rhinelander, Wisc. 5450 (USA)

Dr. G. Namkoong
Southeastern Forest Experiment Station
U.S. Department of Agriculture-Forest Service
and Genetics Department
North Carolina State University
Raleigh, N.C. 27607 (USA)