

# Effective Size and $F$ -Statistics of Subdivided Populations. I. Monoecious Species with Partial Selfing

Jinliang Wang

College of Animal Science, Zhejiang Agricultural University, Hangzhou 310029, People's Republic of China

Manuscript received January 22, 1996  
Accepted for publication April 28, 1997

## ABSTRACT

Assuming discrete generations and autosomal inheritance involving genes that do not affect viability or reproductive ability, we have derived recurrence equations for the inbreeding coefficient and coancestry between individuals within and among subpopulations for a subdivided monoecious population with arbitrary distributions of male and female gametes per family, variable pollen and seed migration rates, and partial selfing. From the equations, formulas for effective size and expressions for  $F$ -statistics are obtained. For the special case of a single unsubdivided population, our equations reduce to the simple expressions derived by previous authors. It is shown that population structure (subdivision and migration) is important in determining the inbreeding coefficient and effective size. Failure to recognize internal structures of populations may lead to considerable bias in predicting effective size. Inbreeding coefficient, coancestry between individuals within and among subpopulations accrue at different and variable rates over initial generations before they converge to the same asymptotic rate of increase. For a given population, the smaller the pollen and seed migration rates, the more generations are required to attain the asymptotic rate and the larger the asymptotic effective size. The equations presented herein can be used for the study of evolutionary biology and conservation genetics.

**T**HE effective size of a population is a key parameter in population and quantitative genetics. The concept was introduced by WRIGHT (1931) as the size of an ideal population whose genetic composition is influenced by random processes in the same way as the real population. It was distinguished subsequently by CROW and coworkers (CROW and KIMURA 1970, pp. 345–364) as inbreeding effective size, which predicts changes in heterozygosity, and variance effective size, which predicts changes in gene frequency drift variance; however, these two effective sizes are identical in a population that is not completely subdivided and of constant size (KIMURA and CROW 1963; CROW and KIMURA 1971; JOHNSON 1977; HILL 1979).

Most of the previous work on effective size considers a single unsubdivided population. There is evidence that many organisms are arrayed into complexes of breeding units or subpopulations (SELANDER 1970; CHESSER 1983) due to intrinsic factors, such as behavioral segregation, or extrinsic factors, such as geographic distance and habitat fragmentation. It is found that differentiation among subpopulations is very common in plants (BARRETT and HUSBAND 1990). Thus it is more appropriate to take population structure into account in deriving expressions for effective size. CHESSER *et al.* (1993) considered the effective size and  $F$ -statistics of a dioecious population with subdivision

and migration; SUGG and CHESSER (1994) extended the model to include multiple paternity. In this article, we broaden the scope of previous inquiries by considering the effective size of a monoecious diploid population of finite size subdivided into a number of subpopulations. We incorporate both pollen and seed migrations in WRIGHT's (1969) island model, partial selfing and an arbitrary distribution of family size in our analysis. By removing the restrictive assumptions made by previous authors, we obtain more general equations for effective size and  $F$ -statistics that preserve concordance with the classical models when the same assumptions are added.

## ASSUMPTIONS AND PARAMETER DEFINITIONS

We consider a monoecious population of plants subdivided into  $s$  subpopulations, each consisting of  $N$  individuals. In each generation, each subpopulation receives a proportion ( $d_p$ ) of migrant pollen drawn at random from the whole population at the time when plants are flowering. Since plants may reproduce by self-fertilization to varying degrees (SCHEMSKE and LANDE 1985), we incorporate partial selfing in our model. After pollen migration, each subpopulation reproduces by selfing with probability  $\beta$  and by random outbreeding with probability  $1 - \beta$ . After pollination each subpopulation receives a proportion ( $d_s$ ) of seeds taken randomly from the whole population. Thus both pollen and seed migrate in the island model (WRIGHT 1943), and female gametes do not migrate. Throughout the

Address for correspondence: Jinliang Wang, Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Rd., Edinburgh EH9 3JT, United Kingdom.  
E-mail: jinliang.wang@ed.ac.uk

article we shall deal just with the case of stable census size and population structure for simplicity. Other assumptions are discrete generations, autosomal inheritance involving genes that do not affect viability or reproductive ability so that natural selection is not operating to eliminate them.

Because of the symmetry of the island model, only the following probabilities of identity by descent are necessary to describe the second moments of gene frequencies for a subdivided monoecious population:

$F$ : the inbreeding coefficient defined as the probability that two genes at a given locus of an individual drawn at random from the population are identical by descent.

$\theta(\theta^*)$ : coancestry of individuals chosen at random from the same subpopulation after (before) seed migration. Coancestry is defined as the probability that two genes at a given locus, one taken at random from each of two randomly selected individuals, are identical by descent.

$\alpha(\alpha^*)$ : coancestry of individuals drawn at random from different subpopulations after (before) seed migration.

WRIGHT's (1969, pp. 294–295)  $F$ -statistics are rather useful for describing gene flow and breeding structure of subdivided populations. This is made obvious by their extensive use in the published literature and by the attention to proper interpretation and estimation of the coefficients (NEI 1977; WEIR and COCKERHAM 1984; WEIR 1990). The  $F$ -statistics are defined as

$$F_{IS} = \frac{F - \theta}{1 - \theta}; \quad F_{ST} = \frac{\theta - \alpha}{1 - \alpha}; \quad F_{IT} = \frac{F - \alpha}{1 - \alpha} \quad (1)$$

(COCKERHAM 1969, 1973), where subscripts  $I$ ,  $S$ , and  $T$  represent individuals, subpopulations, and the total population, respectively. The  $F$ -statistics can be thought of as inbreeding coefficient and also as correlations between gametes. In a subdivided population,  $F_{IS}$  is concerned with inbreeding in individuals relative to the subpopulation to which they belong,  $F_{ST}$  is concerned with inbreeding in subpopulations relative to the total population of which they are a part, and  $F_{IT}$  is concerned with inbreeding in individuals relative to the total population. Equally,  $F_{IS}$  is the correlation of uniting gametes relative to gametes drawn at random from within a subpopulation,  $F_{ST}$  is the correlation of gametes drawn at random within subpopulations relative to gametes drawn at random from the entire population, and  $F_{IT}$  is the correlation of uniting gametes relative to gametes taken at random from the entire population.  $F_{IS}$  is a function of the nonrandom mating in the subpopulation, it being negative, zero, and positive with avoidance of close inbreeding, random mating, and close inbreeding, respectively.  $F_{ST}$  is a measure of the increase in inbreeding due to the finite size of the population. It is always greater than or equal to zero because of the Wahlund effect. Finally,  $F_{IT}$  is the most inclusive measure of inbreeding in that it takes into account both

the effects of nonrandom mating within subpopulations ( $F_{IS}$ ) and the effects of population subdivision ( $F_{ST}$ ). The relation among the three coefficients can be expressed as

$$(1 - F_{IT}) = (1 - F_{ST})(1 - F_{IS}) \quad (2)$$

(CROW and KIMURA 1970, pp. 106).

#### RECURRENCE EQUATIONS FOR PROBABILITIES OF IDENTITY BY DESCENT

First we consider the recurrence equation for the inbreeding coefficient,  $F$ . Each zygote is formed by sampling one male gamete and one female gamete from the parents. The female gamete is assumed to be taken from within a subpopulation. The male gamete, however, is chosen at random from the same subpopulation as the female gamete with probability

$$q_{mf} = 1 - d_p(1 - 1/s) \quad (3)$$

and from a different subpopulation with probability  $1 - q_{mf}$ . For the first case, the male gamete comes from the same parent as that of the female gamete with probability  $\beta$ , from a different individual with probability  $q_{mf} - \beta$ . Thus it is clear that  $\beta \leq 1 - d_p(1 - 1/s)$  or  $d_p \leq s(1 - \beta)/(s - 1)$ . The inbreeding coefficients of a zygote in generation  $t$ , from the two gametes that descend from the same individual, separate individuals in the same subpopulation and in different subpopulations:  $1/2(1 + F_{t-1})$ ,  $\theta_{t-1}$  and  $\alpha_{t-1}$ , respectively. Thus the average inbreeding coefficient in generation  $t$  is

$$F_t = 1/2\beta(1 + F_{t-1}) + (q_{mf} - \beta)\theta_{t-1} + (1 - q_{mf})\alpha_{t-1}. \quad (4)$$

Next we consider the recurrence equations for  $\theta$ . Now the two genes under consideration are from different individuals in the same subpopulation. Three cases can be distinguished.

**Both genes come from male gametes with probability  $1/4$ :** Since pollen migrates with migration rate  $d_p$ , we can distinguish three situations in deriving the probability (denoted by  $q_{mm}$ ) that two male gametes taken at random and without replacement are from the same subpopulation.

*Both male gametes are nonmigrants:* This occurs with probability

$$(1 - d_p) \left( 1 - \frac{Nd_p}{N-1} \right). \quad (5a)$$

In such a situation, the gametes certainly come from the same subpopulation. Thus the probability that they are from the same subpopulation is also given by (5a).

*One is a migrant and the other a nonmigrant:* This occurs with probability  $2Nd_p(1 - d_p)/(N - 1)$ . Since the migrant comes equally from the  $s$  subpopulations, the

probability that the two gametes are from the same subpopulation is

$$\frac{2Nd_p(1 - d_p)}{s(N - 1)} \tag{5b}$$

*Both male gametes are migrants:* This occurs with probability  $d_p(Nd_p - 1)/(N - 1)$ . The probability that two migrants taken at random and without replacement come from the same subpopulation is  $(Nd_p - 1)/(sNd_p - 1)$ . Thus in this situation the probability that the two gametes are from the same subpopulation is

$$\frac{d_p(Nd_p - 1)^2}{(N - 1)(sNd_p - 1)} \tag{5c}$$

Combining (5a), (5b), and (5c) we get

$$q_{mm} = (1 - d_p) \left( 1 - \frac{Nd_p}{N - 1} \right) + \frac{2Nd_p(1 - d_p)}{s(N - 1)} + \frac{d_p(Nd_p - 1)^2}{(N - 1)(sNd_p - 1)} \tag{6}$$

For the special cases of no migration ( $d_p = 0$ ) and complete dispersion ( $d_p = 1$ ) of pollen, (6) reduces to  $q_{mm} = 1$  and  $q_{mm} = (N - 1)/(sN - 1)$ , respectively, as expected. Expression 6 is generally quite complicated. If the subpopulation size  $N$  is relatively large, (6) reduces to

$$q_{mm} = 1 - (1 - 1/s)d_p(2 - d_p) \tag{7}$$

approximately. Note that if the gametes are sampled with replacement we also obtain (7). MARUYAMA and TACHIDA (1992) used the latter sampling method without explanation. We think, however, sampling without replacement is more practical in reality. However, for simplicity we still use (7) instead of (6) for relatively large values of  $N$  in this article.

It has been shown that the probability that a random pair of male gametes from the same subpopulation are descended from the same parent is  $p_{mm} = \sigma_m^2/(N - 1)$  (WANG 1996), where  $\sigma_m^2$  is the variance of the number of male gametes per parent. Thus, the two genes are identical by descent with probability

$$\frac{1}{4} \left\{ q_{mm} \left[ p_{mm} \left( \frac{1 + F_{t-1}}{2} \right) + (1 - p_{mm})\theta_{t-1} \right] + (1 - q_{mm})\alpha_{t-1} \right\} \tag{8}$$

**One gene is from a male gamete and the other is from a female gamete with probability  $1/2$ :** Since only pollen migrates, the probability that the two genes (or gametes) come from the same subpopulation is  $q_{mf} = 1 - d_p(1 - 1/s)$ . For partial selfing populations, the probability that two gametes of separate sexes taken at random that do not unite to produce one zygote and come from the same parent is  $p_{mf} = (\sigma_{mf} + 1 - \beta)/$

$(N - 1)$ , where  $\sigma_{mf}$  is the covariance of the numbers of male and female gametes contributed per parent (WANG 1996). Thus the two genes are identical by descent with probability

$$\frac{1}{2} \left\{ q_{mf} \left[ p_{mf} \left( \frac{1 + F_{t-1}}{2} \right) + (1 - p_{mf})\theta_{t-1} \right] + (1 - q_{mf})\alpha_{t-1} \right\} \tag{9}$$

**Both genes come from female gametes with probability  $1/4$ :** Since female gametes do not migrate, the two genes are from the same subpopulation with probability one. Thus the probability of identity is

$$\frac{1}{4} \left\{ p_{ff} \left( \frac{1 + F_{t-1}}{2} \right) + (1 - p_{ff})\theta_{t-1} \right\}, \tag{10}$$

where  $p_{ff} = \sigma_f^2/(N - 1)$  and  $\sigma_f^2$  is the variance of the number of female gametes per parent. In (8-10),  $\sigma_v^2$  and  $\sigma_{mf}$  are theoretical variance and covariance; the corresponding observed values are with Gaussian corrections  $S_v^2 = \sigma_v^2 N/(N - 1)$  and  $S_{mf} = \sigma_{mf} N/(N - 1)$ , respectively, where  $v = m$  or  $f$ .

Combining Expressions 8-10, the coancestry between two random individuals before seed migration is

$$\theta_t^* = 1/8 (q_{mm}p_{mm} + 2q_{mf}p_{mf} + p_{ff}) (1 - 2\theta_{t-1} + F_{t-1}) + 1/4 (q_{mm} + 2q_{mf} + 1) (\theta_{t-1} - \alpha_{t-1}) + \alpha_{t-1} \tag{11}$$

Similarly, we can derive the coancestry between two individuals, taken at random from separate subpopulations before seed migration. The equation is

$$\alpha_t^* = 1/8 (q'_{mm}p_{mm} + 2q'_{mf}p_{mf}) (1 - 2\theta_{t-1} + F_{t-1}) + 1/4 (q'_{mm} + 2q'_{mf}) (\theta_{t-1} - \alpha_{t-1}) + \alpha_{t-1}, \tag{12}$$

where  $q'_{mv}$  is the probability that a male gamete and a gamete of sex  $v$  ( $v$  refers to sex,  $v = m$  or  $f$ ), taken at random and without replacement from separate subpopulations, are from the same subpopulation before seed migration. Using a procedure similar to the derivation of (6), we can obtain

$$q'_{mf} = d_p/s, \tag{13}$$

$$q'_{mm} = \frac{2d_p(1 - d_p)}{s} + \frac{d_p^2(Nd_p - 1)}{sNd_p - 1} \tag{14}$$

For large subpopulation size ( $N$ ), (14) reduces to

$$q'_{mm} = d_p(2 - d_p)/s \tag{15}$$

as an approximation.

Now we consider seed migration. It changes the coancestry between individuals,  $\theta_t^*$  and  $\alpha_t^*$ , but does not affect the inbreeding coefficient,  $F_t$ . After seed migration, the probabilities that two individuals, taken at random and without replacement from within subpopulations (denoted as  $q_{ss}$ ) and from different subpopula-

tions (denoted as  $q'_{ss}$ ), come from the same subpopulation before seed migration can also be expressed by (6) and (14), respectively, or by (7) and (15) approximately, only to replace  $d_p$  with  $d_s$ . Thus the coancestry between two individuals from the same subpopulation after seed migration is obtained as

$$\theta_t = q_{ss}\theta_t^* + (1 - q_{ss})\alpha_t^* \tag{16}$$

Substituting (11) and (12) into (16) yields

$$\begin{aligned} \theta_t = & \frac{1}{8}\{q_{ss}[(q_{mm} - q'_{mm})p_{mm} + 2(q_{mf} - q'_{mf})p_{mf} \\ & + p_{ff}] + q'_{mm}p_{mm} + 2q'_{mf}p_{mf}\}(1 - 2\theta_{t-1} + F_{t-1}) \\ & + \frac{1}{4}[q_{ss}(1 + q_{mm} - q'_{mm} + 2q_{mf} - 2q'_{mf}) \\ & + q'_{mm} + 2q'_{mf}](\theta_{t-1} - \alpha_{t-1}) + \alpha_{t-1}. \end{aligned} \tag{17}$$

Similarly we can derive the coancestry between two individuals from different subpopulations after seed migration,

$$\begin{aligned} \alpha_t = & \frac{1}{8}\{q'_{ss}[(q_{mm} - q'_{mm})p_{mm} + 2(q_{mf} - q'_{mf})p_{mf} + p_{ff}] \\ & + q'_{mm}p_{mm} + 2q'_{mf}p_{mf}\}(1 - 2\theta_{t-1} + F_{t-1}) \\ & + \frac{1}{4}[q'_{ss}(1 + q_{mm} - q'_{mm} + 2q_{mf} - 2q'_{mf}) \\ & + q'_{mm} + 2q'_{mf}](\theta_{t-1} - \alpha_{t-1}) + \alpha_{t-1}. \end{aligned} \tag{18}$$

Expressions 4, 17, and 18 are the complete recurrence equations for  $F$ ,  $\theta$ , and  $\alpha$ . These can be presented in a matrix form,

$$\mathbf{S}_t = \mathbf{TS}_{t-1} + \mathbf{C}. \tag{19}$$

In (19),  $\mathbf{S}_t$  is a column vector of the three probabilities of identity by descent at generation  $t$ ,

$$\mathbf{S}_t = \begin{pmatrix} F_t \\ \theta_t \\ \alpha_t \end{pmatrix}. \tag{20}$$

The transition matrix  $\mathbf{T}$  defining the probabilistic changes of the vector of variables is

$$\mathbf{T} = \begin{pmatrix} \frac{1}{2}\beta & q_{mf} - \beta & 1 - q_{mf} \\ aq_{ss} + b & (c - 2a)q_{ss} + e - 2b & 1 - cq_{ss} - e \\ aq'_{ss} + b & (c - 2a)q'_{ss} + e - 2b & 1 - cq'_{ss} - e \end{pmatrix}, \tag{21}$$

where we put

$$\begin{aligned} a &= \frac{1}{8}[(q_{mm} - q'_{mm})p_{mm} + 2(q_{mf} - q'_{mf})p_{mf} + p_{ff}], \\ b &= \frac{1}{8}(q'_{mm}p_{mm} + 2q'_{mf}p_{mf}), \\ c &= \frac{1}{4}(1 + q_{mm} - q'_{mm} + 2q_{mf} - 2q'_{mf}), \\ e &= \frac{1}{4}(q'_{mm} + 2q'_{mf}). \end{aligned} \tag{22}$$

The constant column vector  $\mathbf{C}$  is just the first column of matrix  $\mathbf{T}$ ,

$$\mathbf{C} = \begin{pmatrix} \frac{1}{2}\beta \\ aq_{ss} + b \\ aq'_{ss} + b \end{pmatrix}. \tag{23}$$

Since the initial values of all variables,  $\mathbf{S}_0$ , are zero, we can use (19) to predict the inbreeding coefficient and coancestry and thus  $F$ -statistics and effective size in any generation.

The effect of pollen and seed migration rates on inbreeding coefficient in a subdivided population is shown in Figure 1. The graphs are generated using 20 individuals per subpopulation, 20 subpopulations with parameters  $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = 0.1$ , and  $\beta = 0.1$  (independent Poisson distributions of the number of progeny per parent from selfing and outcrossing). Results applicable to all subdivided populations are as follows.

(1) The inbreeding coefficients in the first few generations are decreased by pollen or seed migration. However, the lines converge slightly and eventually cross; larger values of migration rate give lower inbreeding coefficients initially than smaller values, but in later generations the order is reversed. The convergence rate is dependent on values of  $d_s$ ,  $d_p$ ,  $N$ ,  $s$ ,  $\beta$ ,  $S_m^2$ ,  $S_f^2$ , and  $S_{mf}$ . Using the recurrence equations for inbreeding coefficient and coancestry, the generation at which a reversal takes place for different values of migration rate can be calculated. In Figure 1A, for example, the line for  $d_p = 0.1$  will cross lines for  $d_p = 0.2, 0.3$ , and  $0.4$  in generations 900, 864, and 837, respectively, while in Figure 1B the line for  $d_s = 0.1$  will cross lines for  $d_s = 0.2, 0.3$ , and  $0.4$  in generations 864, 855, and 850, respectively. (2) Seed migration has a larger effect on inbreeding coefficient than pollen migration. At a given migration rate, seed migration gives lower inbreeding in initial generations and higher inbreeding in later generations than pollen migration. For example in Figure 1,  $d_s = 0.4$  and  $d_p = 0$  results in lower inbreeding than  $d_p = 0.4$  and  $d_s = 0$  before generation 431 and thereafter the order is reversed. This is because, at the same rate of migration, seed migration disperses more genes than pollen migration.

### EFFECTIVE SIZE

For a finite population size, the values of  $F$ ,  $\theta$ , and  $\alpha$  will accrue with generations. Although the instantaneous rates of increase for the three variables,

$$\begin{aligned} \Delta F_t &= \frac{F_t - F_{t-1}}{1 - F_{t-1}}; & \Delta \theta_t &= \frac{\theta_t - \theta_{t-1}}{1 - \theta_{t-1}}; \\ \Delta \alpha_t &= \frac{\alpha_t - \alpha_{t-1}}{1 - \alpha_{t-1}} \end{aligned} \tag{24}$$

may be quite different in initial generations, they will eventually converge to the same value ( $\Delta F$ ) if the sub-

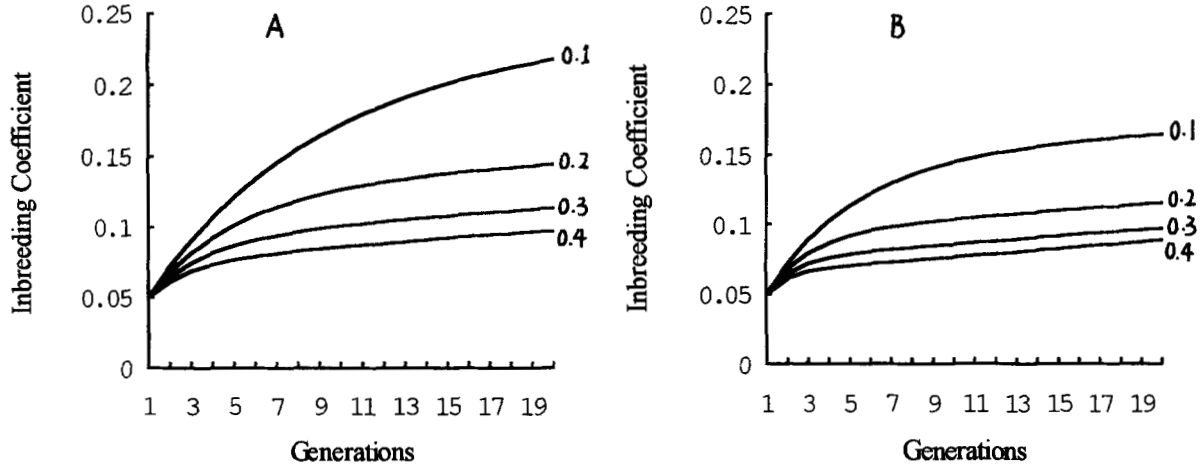


FIGURE 1.—Effect of migration rate ( $d_p$  or  $d_s = 0.1, 0.2, 0.3,$  and  $0.4$ ) on inbreeding coefficient over the first 20 generations for a population subdivided into 20 subpopulations, each consisting of 20 individuals, with parameters  $S_m^2 = S_f^2 = 1, S_{mf} = 0.1,$  and  $\beta = 0.1$ . (A) Effect of  $d_p$  with constant value of  $d_s$  (0). (B) Effects of  $d_s$  with a constant value of  $d_p$  (0).

populations are not completely isolated ( $d_p + d_s > 0$ ).  $\Delta F_t, \Delta \theta_t,$  and  $\Delta \alpha_t$  correspond to instantaneous inbreeding ( $N_{el,t}$ ), coancestry ( $N_{\theta,t}$ ), and variance ( $N_{eV,t}$ ) effective sizes, respectively, with relations

$$\Delta F_t = \frac{1}{2N_{el,t}}; \quad \Delta \theta_t = \frac{1}{2N_{\theta,t}}; \quad \Delta \alpha_t = \frac{1}{2N_{eV,t}}. \quad (25)$$

Thus  $N_{el,t}, N_{\theta,t},$  and  $N_{eV,t}$  will also converge to the same asymptotic value ( $N_e$ ). The number of generations required for convergence is dependent mainly upon the total rate of migration with low rates requiring more generations until convergence is achieved. Other factors, such as subpopulation number ( $s$ ) and selfing proportion ( $\beta$ ), also have some effect on the convergence rate. Figure 2 shows the changes in instantaneous inbreeding, coancestry, and variance effective sizes over successive generations for two populations with Poisson distribution of family size. It can be seen that although the three effective sizes may be quite different and change dramatically over the first few generations, they converge gradually to the same value in later generations. Generally,  $N_{eV,t}$  decreases,  $N_{\theta,t}$  increases, and  $N_{el,t}$  changes erratically (depending on parameters such as  $\beta, d_s,$  and  $d_p$ ) over the initial disequilibrium generations.

The asymptotic effective size,  $N_e$ , can be obtained by using the transition matrix **T**. A matrix like **T**, which has only positive elements, has a single dominant positive characteristic root  $\lambda$ , which equals to  $1 - \Delta F$ . Thus  $N_e$  can be derived as

$$N_e = \frac{1}{2(1 - \lambda)}. \quad (26)$$

The characteristic equation of matrix **T** can be obtained from (21), which reads

$$\begin{aligned} \lambda^3 - [1 - 2b - 2aq_{ss} + \frac{1}{2}\beta + c(q_{ss} - q'_{ss})]\lambda^2 \\ - [(2a - c - 2ae + 2bc - \frac{1}{2}\beta c + aq_{mf}) \\ \times (q_{ss} - q'_{ss}) - \frac{1}{2}\beta + aq'_{ss} + b]\lambda \\ - (\frac{1}{2}\beta c - aq_{mf} + ae - bc)(q_{ss} - q'_{ss}) = 0 \quad (27) \end{aligned}$$

The equation has three roots; the largest one lying between zero and one is the dominant characteristic root,  $\lambda$ , which is required. Using Newton's tangential method of approximation and omitting second and higher orders of  $1/N$ , the solution for  $\lambda$  can be found analytically from (27) as

$$\begin{aligned} \lambda = 1 - x + \frac{x}{(1 - \frac{1}{2}\beta)[1 - c(q_{ss} - q'_{ss})]} \\ \times \{a(1 - q_{mf})(q_{ss} - q'_{ss}) + b + aq_{ss} \\ - x[\frac{1}{2}\beta + c(1 - \beta)(q_{ss} - q'_{ss})]\}, \quad (28) \end{aligned}$$

where

$$x = \frac{b + aq'_{ss} + (ae - bc)(q_{ss} - q'_{ss})}{(1 - \frac{1}{2}\beta)[1 - c(q_{ss} - q'_{ss})]}.$$

Substituting (28) into (26), we therefore get the asymptotic effective size

$$\begin{aligned} N_e = \frac{\frac{1}{4}(2 - \beta)[1 - c(q_{ss} - q'_{ss})]}{b + aq'_{ss} + (ae - bc)(q_{ss} - q'_{ss})} \\ + \frac{b + aq_{ss} + a(1 - q_{mf})(q_{ss} - q'_{ss})}{2[b + aq'_{ss} + (ae - bc)(q_{ss} - q'_{ss})]} \\ + \frac{1}{2[1 - c(q_{ss} - q'_{ss})]}. \quad (29) \end{aligned}$$

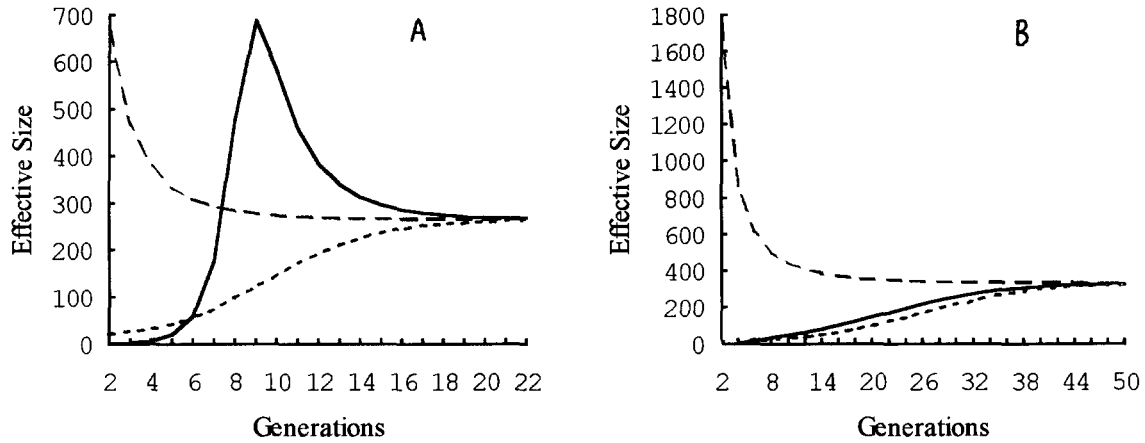


FIGURE 2.—Changes in instantaneous inbreeding ( $N_{e,t}$ , —), coancestry ( $N_{\theta,t}$ , ----) and variance ( $N_{eV,t}$ , - - -) effective sizes over successive generations for two populations with Poisson distribution of family size. For population A, the parameters are  $s = 20$ ,  $N = 20$ ,  $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = \beta = 0.8$ ,  $d_p = 0.3$ , and  $d_s = 0$ . For population B, the parameters are as that of population A except that  $d_p = 0.1$ .

The equation is complicated and it is not easy to see the relations between  $N_e$  and the parameters of the population. When both seed and pollen migration rates are small enough so that second and higher order terms of  $d_p$  and  $d_s$  can be ignored,  $N$  is large so that (7) and (15) can be used for an approximation, (29) is simplified considerably to

$$N_e = \frac{2s(2 - \beta)N}{S_k^2 + 2 - 2\beta} + \frac{1 + s(1 - 2d_s)}{2(d_p + 2d_s)} + \frac{sd_p(S_f^2 - S_m^2)}{2(d_p + 2d_s)(S_k^2 + 2 - 2\beta)}, \quad (30)$$

where  $S_k^2 = S_m^2 + 2S_{mf} + S_f^2$  is the variance in the number of total gametes per parent.

The predictions from (29) or (30) (for small values of  $d_p$  and  $d_s$ ) fit in very well with the numerical results by recurrence equations. The effects of  $d_p$ ,  $d_s$ ,  $N$ , and  $s$  on  $N_e$  are shown in Figure 3 for a population with fixed total size ( $sN = 8000$ ) and equal migration rates of seed and pollen ( $d_s = d_p$ ). The other parameters are  $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = \beta = 0.2$ . It can be seen that decreasing migration rate always increases  $N_e$ , this is more evident when the number of subpopulations is large. Some common conclusions are as follows. (1) For a given population (given values of all parameters except migration rates), the smaller the migration rates, the larger their effects on the increase of effective size. (2) For large size ( $N$ ) and small number ( $s$ ) of subpopulations, the effective size is determined mainly by parameters  $N$ ,  $s$ ,  $\beta$ , and  $S_k^2$ ; migration parameters  $d_p$  and  $d_s$  have relatively small effects on  $N_e$ . If  $s$  is large and  $N$  is small, however, small migration rates have relatively large effects on  $N_e$ . (3) The absolute magnitude of the effects of migration rates on  $N_e$  is irrespective of the census size of the subpopulation ( $N$ ), while it increases rapidly with the increment in the number of subpopulations ( $s$ ).

The effect of selfing proportion on effective size is influenced by the variance of family size; self-fertiliza-

tion increases effective size when the variance of the number of gametes contributed per parent is small, and *vice versa*. For equal gamete contribution per family ( $S_m^2 = S_f^2 = S_{mf} = 0$ ) and Poisson distribution of the number of gametes per family ( $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = \beta$ ), the effects of selfing proportion on effective size are shown in Figure 4 for a population with structure and breeding parameters  $s = 20$ ,  $N = 20$ ,  $d_p = 0.1$ , and  $d_s = 0.05$ . The reason that close inbreeding (here selfing) has a differential effect depending on the variance of family size has been explained by ROBERTSON (1964) and CABALLERO (1994).

The effects of pollen and seed migrations are asymmetrical in their influence on effective size. For a given migration rate, pollen migration always has a smaller effect on effective size than seed migration. For a population with a Poisson distribution of family size ( $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = \beta = 0.2$ ) with parameters  $s = N = 20$ , the effects of pollen and seed migration rates on effective size are shown in Figure 5. Clearly, the value of  $N_e$  for a given subdivided population is smallest when either  $d_p$  or  $d_s$  is large and increases dramatically as both  $d_p$  and  $d_s$  approach zero.

For a single unsubdivided population with partial selfing, CABALLERO and HILL (1992) and POLLAK and SABRAN (1992) derived the effective size

$$N_e = \frac{2(2 - \beta)N}{S_k^2 + 2 - 2\beta} \quad (31)$$

by different methods. For the special case in our model,  $s = 1$ ,  $d_p = d_s = 1$  (all pollen and seeds disperse back into the population from which they come), and the general expression (29) reduces to  $N_e = 2(2 - \beta) / (p_{mm} + 2p_{mf} + p_{ff})$  approximately, which again reduces to (31) if  $p_{mm}$ ,  $p_{mf}$ , and  $p_{ff}$  are substituted, as expected.

When the population is subdivided permanently without migration among subpopulations ( $d_p = d_s = 0$ ), then the effective size can not be calculated by (29)

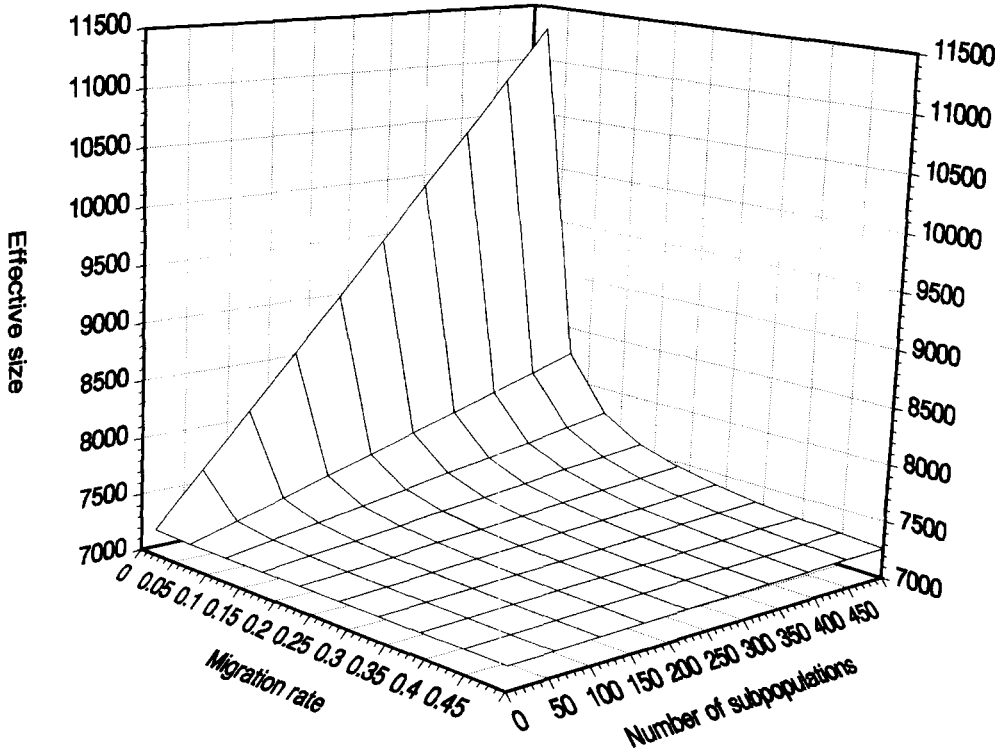


FIGURE 3.—Three-dimensional diagrams depicting the relationship between migration rate ( $d_s = d_p$ ) and the value of  $s$  or  $N$  on the effective size. The population has a fixed total size ( $sN = 8000$ ), with other parameters being  $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = \beta = 0.2$ .

or (30). In such a case, inbreeding and genetic drift will follow different fates, and  $N_{eI,t}$  and  $N_{e\theta,t}$  will converge to the same value, the asymptotic inbreeding effective size as given by (31). The equation can be derived from (4) and (17) reduced for  $d_p = d_s = 0$  following the same procedure as the derivation of (29). The variance effective size  $N_{eV,t}$ , however, does not converge to the same value as  $N_{eI,t}$  and  $N_{e\theta,t}$ . It reaches another asymptotic value, which is much larger than inbreeding effective size. Thus the variance effective size is maximized and the inbreeding effective size minimized with complete subdivision (ROBERTSON 1964).

F-STATISTICS

The instantaneous rate of increase in coancestry between individuals within subpopulations in generation  $t$  can be presented as

$$\Delta\theta_t = \frac{\theta_t - \theta_{t-1}}{1 - \theta_{t-1}} = \frac{1}{2N_{e\theta,t}} \quad (32)$$

Substituting (17) and (22) into (32) yields

$$\frac{1}{2N_{e\theta,t}} = (aq_{ss} + b) \left( 1 + \frac{F_{I,t-1} - \theta_{t-1}}{1 - \theta_{t-1}} \right) + (cq_{ss} + e - 1) \left( \frac{\theta_{t-1} - \alpha_{t-1}}{1 - \theta_{t-1}} \right) \quad (33)$$

If we denote the instantaneous  $F$ statistics in generation  $t - 1$  as

$$F_{IS,t-1} = \frac{F_{t-1} - \theta_{t-1}}{1 - \theta_{t-1}}; \quad F_{ST,t-1} = \frac{\theta_{t-1} - \alpha_{t-1}}{1 - \alpha_{t-1}}; \quad F_{IT,t-1} = \frac{F_{t-1} - \alpha_{t-1}}{1 - \alpha_{t-1}} \quad (34)$$

and substitute these relations into (33), we obtain

$$\frac{1}{2N_{e\theta,t}} = (aq_{ss} + b) (1 + F_{IS,t-1}) + (cq_{ss} + e - 1) \left( \frac{F_{ST,t-1}}{1 - F_{ST,t-1}} \right) \quad (35)$$

Similarly, we can derive the instantaneous variance and inbreeding effective sizes in generation  $t$  from (18) and (4) as

$$\frac{1}{2N_{eV,t}} = (aq'_{ss} + b) (1 - F_{ST,t-1}) (1 + F_{IS,t-1}) + (cq'_{ss} + e) F_{ST,t-1} \quad (36)$$

$$\frac{1}{2N_{eI,t}} = \frac{\beta(1 + F_{IS,t-1})}{2(1 - F_{IS,t-1})} + \frac{q_{mf}F_{ST,t-1} - F_{IT,t-1}}{1 - F_{IT,t-1}} \quad (37)$$

If there is migration among subpopulations ( $d_p + d_s > 0$ ), the instantaneous effective sizes ( $N_{eI,t}$ ,  $N_{eV,t}$ , and  $N_{e\theta,t}$ ) will change over generations initially before they converge to the same asymptotic effective size ( $N_e$ ). Instantaneous  $F$ statistics as given in (34) also change in the first few generations and will converge to their respective asymptotic values in later generations. Using (4), (17), (18), and (34), the instantaneous  $F$ statistic

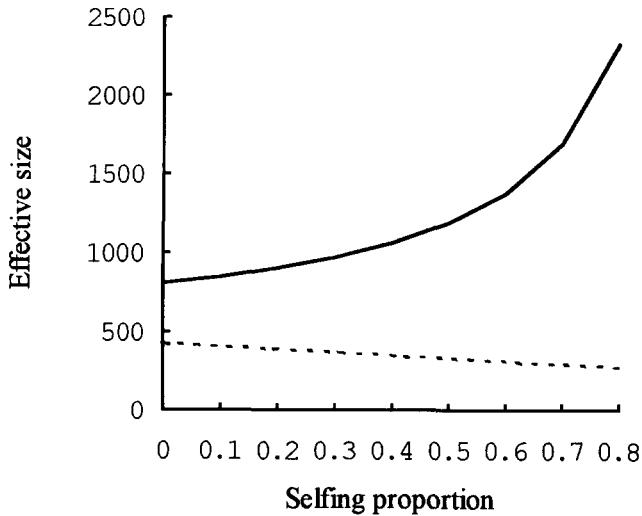


FIGURE 4.—Effect of partial selfing on the effective size of a population with parameters  $s = 20$ ,  $N = 20$ ,  $d_p = 0.1$ , and  $d_s = 0.05$ . —, equal gamete contribution per family ( $S_m^2 = S_f^2 = S_{mf} = 0$ ); ----, Poisson distribution of the number of gametes per family ( $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = \beta$ ).

tics can be calculated, which are shown in Figure 6 for a population with parameters  $s = 20$ ,  $N = 20$ ,  $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = \beta = 0.5$ ,  $d_p = 0.1$ , and  $d_s = 0$ . It can be seen that  $F_{IS}$  reaches its asymptotic value in fewer generations than  $F_{ST}$  and  $F_{IT}$ . This is true for all populations with different sets of parameters.

After a sufficient number of generations when the population attained the steady state, the three effective sizes arrive at the same asymptotic value ( $N_e$ ) and the instantaneous  $F$ -statistics come to their respective asymptotic values (denoted as  $F_{IS}$ ,  $F_{IT}$ , and  $F_{ST}$ ). In such a case, (35–37) reduce to

$$\frac{1}{2N_e} = (aq_{ss} + b)(1 + F_{IS}) + (cq_{ss} + e - 1) \left( \frac{F_{ST}}{1 - F_{ST}} \right), \quad (38)$$

$$\frac{1}{2N_e} = (aq'_{ss} + b)(1 - F_{ST})(1 + F_{IS}) + (cq'_{ss} + e)F_{ST}, \quad (39)$$

$$\frac{1}{2N_e} = \frac{\beta(1 + F_{IS})}{2(1 - F_{IS})} + \frac{q_{mf}F_{ST} - F_{IT}}{1 - F_{IT}}. \quad (40)$$

Solving the equations for  $F_{IS}$  and  $F_{ST}$ , we therefore obtain

$$F_{ST} = \frac{a(q_{ss} - q'_{ss})}{2N_e[b + aq'_{ss} + (ae - bc)(q_{ss} - q'_{ss})]}, \quad (41)$$

$$F_{IS} = \frac{\beta - 2(1 - q_{mf})F_{ST}/(1 - F_{ST})}{2 - \beta}, \quad (42)$$

and  $F_{IT}$  can be calculated by (2).

When both seed and pollen migration rates are small so that second and higher order terms of  $d_p$  and  $d_s$  can be dropped as an approximation, the general expression for  $F_{ST}$  can be reduced to

$$F_{ST} = \frac{s}{2(d_p + 2d_s)N_e} = \frac{1}{\frac{s+1}{s} + \frac{4(2-\beta)(d_p+2d_s)N}{S_k^2+2-2\beta}}. \quad (43)$$

Expression 41 or 43 can be compared to the classical results derived by WRIGHT (1931, 1943) and later extended by many authors (MARUYAMA 1970; LATTER 1973; LI 1976; PROUT 1981). In WRIGHT's island model of population structure it is assumed that a population consists of an infinite number of subpopulations each of the same size  $N$ , and in each subpopulation the proportion  $m$  of the total gene pool is derived from immigrants that may be considered a random sample of the entire population. Then the approximate equation for  $F_{ST}$  is derived as

$$F_{ST} = \frac{1}{1 + 4Nm} \quad (44)$$

approximately when  $m$  is small. PROUT (1981) extended this model by considering both pollen and seed migrations and obtained (in our notation)

$$F_{ST} = \frac{1}{1 + 2N(d_p + 2d_s)} \quad (45)$$

approximately for small values of  $d_p$  and  $d_s$ . Note that, in PROUT's (1981) pollen-seed migration model, the proportion of migrant genes is  $m = \frac{1}{2}d_p + d_s$ , inserting the relation into (45) we get (44).

For the special case shown above,  $s$  is large and gametes are selected at random from the population (in independent Poisson distributions); thus  $S_m^2 = S_f^2 = 1$ ,  $S_{mf} = \beta$  when the proportion of self-fertilization is  $\beta$ . Seed and pollen migration rates are small enough so that second and higher order terms in  $d_s$  and  $d_p$  can be dropped. With these assumptions our expression for effective size reduces to

$$N_e = \frac{1}{2}(2 - \beta)sN + \frac{1 + s(1 - 2d_s)}{2(d_p + 2d_s)} \quad (46)$$

approximately from (29) or (30), and (41) or (43) reduces to

$$F_{ST} = \frac{1}{1 + (2 - \beta)(d_p + 2d_s)N}. \quad (47)$$

$F_{ST}$  is a monotone increasing function of selfing rate  $\beta$ . Expression (47) is also derived by MARUYAMA and TACHIDA (1992) using an infinite allele model. If self-fertilization occurs at a random proportion ( $1/N$ ), (47) again reduces to (45), as expected. It is clear that



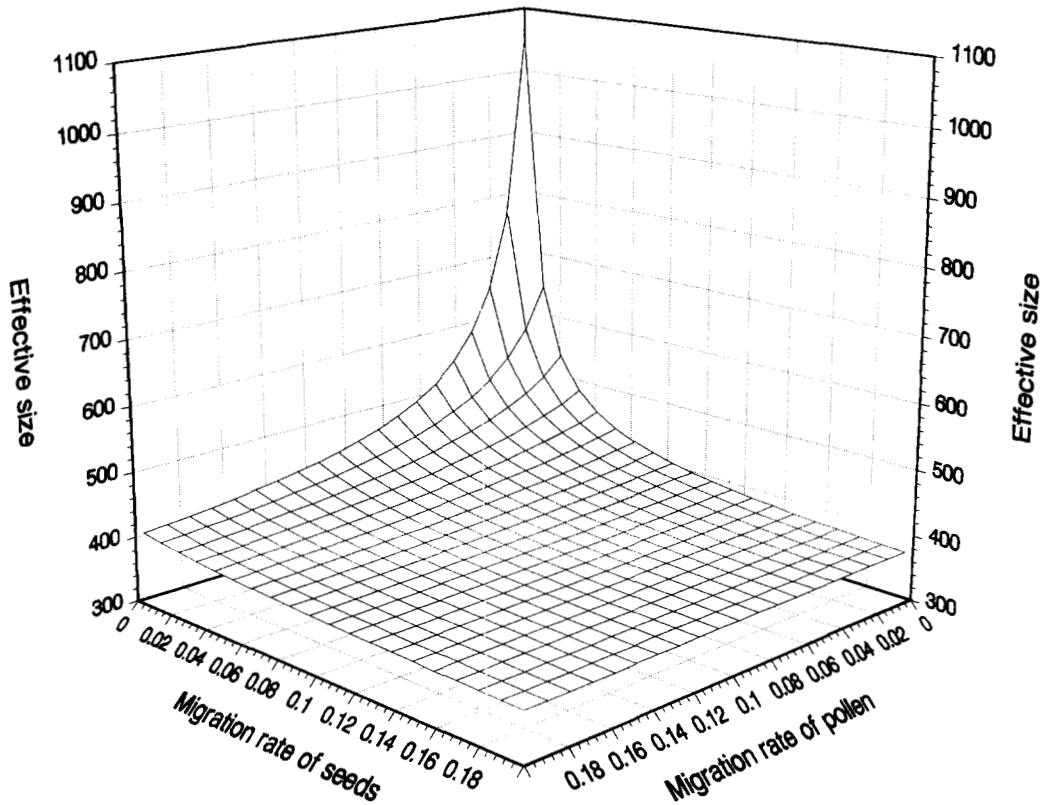


FIGURE 5.—Three-dimensional diagrams depicting the influence of pollen ( $d_p$ ) and seed ( $d_s$ ) migration rates on the effective size. The graph is generated for a population with parameters  $S = 20$ ,  $N = 20$ ,  $S_m^2 = S_f^2 = 1$ , and  $S_{mf} = \beta = 0.2$ .

(41) and (43) derived in this article are more general expressions considering a finite number of subpopulations, partial selfing, variance of family size, and both pollen and seed migrations.

For a single unsubdivided population ( $s = 1$ ,  $d_p = d_s = 1$ ), we have  $F_{ST} = 0$  from (41), and (42) reduces to  $F_{IS} = \beta / (2 - \beta)$ , a classical result that was first derived by HALDANE (1924).

If the population is completely subdivided with no migration among the subpopulations, then we have the instantaneous  $F$ -statistics as  $F_{IS,t} = (F_t - \theta_t) / (1 - \theta_t)$ ,  $F_{ST,t} = \theta_t$ , and  $F_{IT,t} = F_t$  from (34) because the value of  $\alpha$  is zero in any generation.  $F_{IS,t}$  can be negative when self-fertilization is avoided;  $F_{ST,t}$  and  $F_{IT,t}$ , however, are always positive and incremental over generations. If self-fertilization occurs at a random proportion ( $\beta = 1/N$ ), the value of  $F_{IS,t}$  is slightly smaller than zero and  $F_{ST,t} = F_{IT,t+1}$  in any generation. Irrespective of the value of  $\beta$ ,  $F_{IS,t}$  will reach an asymptotic value, while both  $F_{ST,t}$  and  $F_{IT,t}$  will increase steadily over generations toward unity with the same rate of change.

DISCUSSION

A common assumption made by most previous authors in deriving expressions for effective size is a single isolated population. Thus these equations for  $N_e$  are not readily applicable to natural populations or practical crop breeding populations, which are generally somewhat subdivided. The expressions presented

herein provide the basis for predicting the effective size of subdivided populations with arbitrary selfing proportions, variance of family size, pollen and seed migration rates, and different extent of subdivision. It is shown that our expressions reduce to the classical results for the special case of a single unsubdivided population.

Subdivision and gene dispersion are important in determining both inbreeding coefficient and effective size. Failure to take population structure into account can lead to underestimation or overestimation of the true effective size, especially when both pollen and seed migration rates are small and the number of subpopulations is large. For complete seed migration, the population behaves as a single unsubdivided one, and there is little difference between the values of effective size predicted by traditional and the newly derived equations. If seeds disperse randomly and pollen do not migrate ( $d_s = 1$ ,  $d_p = 0$ ), (29) reduces to

$$N_e = \frac{2(2 - \beta)}{\left(\frac{N - 1}{sN - 1}\right)(p_{mm} + 2p_{mf} + p_{ff})} \tag{48}$$

approximately. Noting that  $p_{mm}(N - 1) / (sN - 1) = \sigma_m^2 / (sN - 1)$  is the probability that a random pair of male gametes from the total population are descended from the same parent (and similar for the other two probabilities), (48) is also the prediction of effective size of a single unsubdivided population of size  $sN$ . In the opposite situation ( $d_s = 0$  and  $d_p = 1$ ), (29) is simplified to

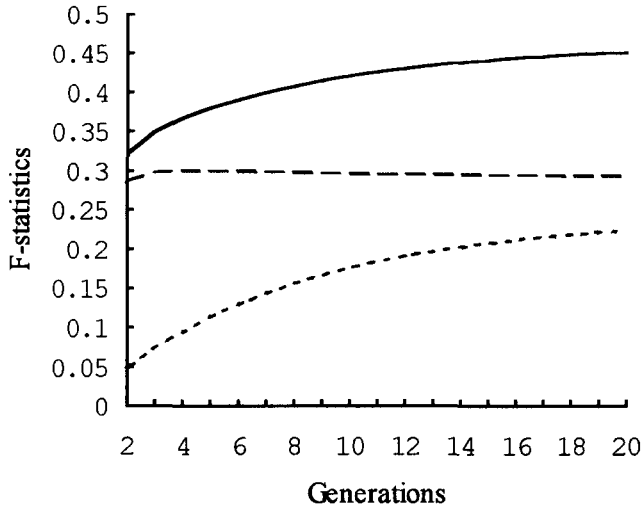


FIGURE 6.—Instantaneous  $F$ -statistics over generations for a population with  $s = 20$ ,  $N = 20$ ,  $d_p = 0.1$ ,  $d_s = 0$ ,  $S_m^2 = S_f^2 = 1$ , and  $S_{mf} = \beta = 0.5$ . —, - - -, and ..... represent  $F_{IT}$ ,  $F_{IS}$ , and  $F_{ST}$  respectively.

$$N_e = \frac{6(2 - \beta)}{\left(\frac{N - 1}{sN - 1}\right)(3p_{mm} + p_{ff}) + \left(\frac{2}{s}\right)(3p_{mf} + p_{ff})} \quad (49)$$

approximately, which always gives values of  $N_e$  slightly smaller than that by (48), especially when  $N$  is small. This implies that even with random dispersion of pollen there is differentiation among the subpopulations. Inserting  $d_s = 0$  and  $d_p = 1$  into (41), we have  $F_{ST} = p_{ff}/[3(2 - \beta)]$  approximately, while  $d_s = 1$  and  $d_p = 0$  results in a value of  $F_{ST}$  of zero. However, when  $N$  is large, complete pollen migration has similar effect to that of complete seed migration. In this case, both (48) and (49) reduce to  $N_e = 2(2 - \beta)s/(p_{mm} + 2p_{mf} + p_{ff})$  approximately and  $F_{ST} \approx 0$ .

For a subdivided population, we should note that the asymptotic effective size is attained after a number of generations. The smaller the migration rates, the more generations are required. The initial values of variance, coancestral and inbreeding effective sizes may be quite different and change at different rates over the initial generations. The reason is that gene correlations are accruing at different rates within individuals, within subpopulations, and among subpopulations. Therefore, the most feasible way to depict accurately the dynamics of all relevant gene correlations is to iterate numerically Equation 19 using a simple computer program. In this manner, the inbreeding coefficients, coancestry,  $F$ -statistics, and effective size can be precisely tracked for each generation. For general planning of programs for genetic conservation or study of long-standing populations in the wild, the expressions for asymptotic effective size can be used as an adequate guide.

It is well-known that predominantly selfing plant or animal species have lower genetic variability within colony and greater values of fixation index ( $F_{ST}$  or  $C_{ST}$ ) compared with outcrossing species (BROWN and RICHARDSON 1988; JARNE 1995). This is usually explained theoretically by Equation 47, assuming that male and female gametes per parent are independently Poisson distributed (MARUYAMA and TACHIDA 1992; JARNE 1995). The present study shows that self-fertilization has differential effects on  $F_{ST}$ . It increases  $F_{ST}$  when  $S_k^2$  is large ( $S_k^2 > 2$ ) and decreases  $F_{ST}$  when  $S_k^2$  is small ( $S_k^2 < 2$ ). If  $S_k^2 = 2$ , the proportion of selfing has little influence on  $F_{ST}$ , as can be seen from (43). For the special case of equal contributions of male and female gametes per parent (which might seldom occur in natural populations but can exist in controlled experimental populations),  $S_k^2 = 0$  and (43) reduces to

$$F_{ST} = \frac{1}{\frac{(s + 1)/s + 2[1 + 1/(1 - \beta)](d_p + 2d_s)N}{}} \quad (50)$$

This is a monotone decreasing function of selfing proportion. The reason that selfing-fertilization decreases colony differentiation ( $F_{ST}$ ) when  $S_k^2$  is small is due to its incremental effect on effective size.

Although it is commonly found that  $F_{ST}$  for selfing populations is larger than that for outcrossing populations in many species, the magnitude of the difference varies drastically for different species observed. JARNE (1995) concluded that  $F_{ST}$  value of selfers was about twice that of outcrossers for many species. The result can be accounted for satisfactorily by Expression 47, which, when the number of effective migrant genes ( $d_p + 2d_s$ )  $N$  is not too small, gives an  $F_{ST}$  value for selfers ( $\beta = 1$ ) close to twice that for outcrossers ( $\beta = 0$ ). HAMRICK and GODT (1990) found that  $F_{ST}$  in a selfing species can be five times that of an outcrossing species. The result may be partly due to the difference between the family size variances of selfing and outcrossing populations. It is possible that selfers have a much larger variance of family size than outcrossers and thus, from Equation 43, a much larger value of  $F_{ST}$ . If, for example, the variances of family size for outcrossing and selfing populations are 2 and 10, respectively, then  $F_{ST}$  of selfing populations can be five times that of outcrossing populations from (43). Although no estimates of  $S_k^2$  for selfers and outcrossers are available from the literature, we suspect that selfers have a larger variance of the number of total gametes contributed per parent than outcrossers. Consider, for example, a population with large fertility variation among individuals from environmental causes. When an individual is in a especially unfavorable niche, it will produce few gametes of both sexes. The successful gametes (that unite to form progeny) contributed by the individual are even fewer for a selfer than that for an outcrosser, because the outcrosser may fertilize (or be fertilized by) another indi-

vidual which occupies a favorable niche and thus has a greater output of gametes. When an individual is in a favorable niche, it will produce more male and female gametes. This also means more successful gametes if the individual is a selfer. If the individual is an outcrosser, however, more gametes do not necessarily result in a larger number of offspring since the latter is also determined by other individuals. Therefore, we surmise that selfers generally have larger values of  $S_k^2$  than outcrossers.

The author is grateful to two anonymous reviewers for their helpful comments on an earlier draft of this manuscript.

#### LITERATURE CITED

- BARRETT, S. C. H., and B. C. HUSBAND, 1990 The genetics of plant migration and colonization, pp. 254–277 in *Plant Population Genetics, Breeding and Genetic Resources*, edited by A. H. D. BROWN, M. T. CLEGG, A. L. KAHLER and B. S. WEIR. Sinauer Associates, Sunderland, MA.
- BROWN, K. M., and T. D. RICHARDSON, 1988 Genetic polymorphism in gastropods: a comparison of methods and habitat scales. *Am. Malacol. Bull.* **6**: 9–17.
- CABALLERO, A., 1994 Developments in the prediction of effective population size. *Heredity* **73**: 657–679.
- CABALLERO, A., and W. G. HILL, 1992 Effective size of nonrandom mating populations. *Genetics* **130**: 909–916.
- CHESSER, R. K., 1983 Genetic variability within and among populations of the black-tailed prairie dog. *Evolution* **37**: 320–331.
- CHESSER, R. K., O. E. RHODES, JR., D. W. SUGG and A. F. SCHNABEL, 1993 Effective sizes for subdivided populations. *Genetics* **135**: 1221–1232.
- COCKERHAM, C. C., 1969 Variance of gene frequencies. *Evolution* **23**: 72–84.
- COCKERHAM, C. C., 1973 Analysis of gene frequencies. *Genetics* **74**: 679–700.
- CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory*. Harper & Row, New York.
- CROW, J. F., and M. KIMURA, 1971 The effective number of a population with overlapping generations: a correction and further discussion. *Am. J. Hum. Genet.* **24**: 1–10.
- HALDANE, J. B. S., 1924 A mathematical theory of natural and artificial selection. Part I. *Proc. Camb. Phil. Soc.* **23**: 19–41.
- HAMRICK, J. L., and J. W. GODT, 1989 Allozymes diversity in plant species, pp. 43–63 in *Plant Population Genetics, Breeding and Genetic Resources*, edited by A. H. D. BROWN, M. T. CLEGG, A. L. KAHLER and B. S. WEIR. Sinauer Associates, Sunderland, MA.
- HILL, W. G., 1979 A note on effective population size with overlapping generations. *Genetics* **92**: 317–322.
- JARNE, P., 1995 Mating system, bottlenecks and genetic polymorphism in hermaphroditic animals. *Genet. Res.* **65**: 193–207.
- JOHNSON, D. L., 1977 Inbreeding in populations with overlapping generations. *Genetics* **87**: 581–591.
- KIMURA, M., and J. F. CROW, 1963 The measurement of effective population number. *Evolution* **17**: 279–288.
- LATTER, B. D. H., 1973 The island model of population differentiation: a general solution. *Genetics* **73**: 147–157.
- LI, W. H., 1976 Effect of migration on genetic distance. *Am. Nat.* **110**: 841–847.
- MARUYAMA, T., 1970 Effective number of alleles in a subdivided population. *Theor. Pop. Biol.* **23**: 216–240.
- MARUYAMA, K., and H. TACHIDA, 1992 Genetic variability and geographical structure in partially selfing populations. *Jpn. J. Genet.* **67**: 39–51.
- NEI, M., 1977 *F*-statistics and analysis of gene diversity in subdivided populations. *Ann. Hum. Genet.* **41**: 225–233.
- POLLAK, E., and M. SABRAN, 1992 On the theory of partially inbreeding finite populations. III. Fixation probabilities under partial selfing when heterozygotes are intermediate in viability. *Genetics* **131**: 979–985.
- PROUT, T., 1981 A note on the island model with sex dependent migration. *Theor. Appl. Genet.* **59**: 327–332.
- ROBERTSON, A., 1964 The effect of non-random mating within inbred lines on the rate of inbreeding. *Genet. Res.* **5**: 164–167.
- SCHEMSKE, D. E., and R. LANDE, 1985 The evolution of self-fertilization and inbreeding depression in plants. 2. Empirical observations. *Evolution* **39**: 41–52.
- SELANDER, R. K., 1970 Behavior and genetic variation in natural populations. *Am. Zool.* **10**: 53–66.
- SUGG, D. W., and R. K. CHESSER, 1994 Effective population size with multiple paternity. *Genetics* **137**: 1147–1155.
- WANG, J., 1996 Inbreeding and variance effective sizes for nonrandom mating populations. *Evolution* **50**: 1786–1794.
- WEIR, B. S., 1990 *Genetic Data Analysis*. Sinauer Associates, Sunderland, MA.
- WEIR, B. S., and C. C. COCKERHAM, 1984 Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370.
- WRIGHT, S., 1931 Evolution in Mendelian populations. *Genetics* **16**: 97–159.
- WRIGHT, S., 1943 Isolation by distance. *Genetics* **28**: 114–138.
- WRIGHT, S., 1969 *Evolution and the Genetics of Population*, Vol. 2. *The Theory of Gene Frequencies*. University of Chicago Press, Chicago.

Communicating editor: B. S. WEIR