

Review article

Developments in predicting the effective size of subdivided populations

JINLIANG WANG*† & ARMANDO CABALLERO‡

†*Institute of Cell, Animal and Population Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK, and ‡Departamento de Bioquímica, Genética e Inmunología, Facultad de Ciencias, Universidad de Vigo, 36200 Vigo, Spain*

The effective population size is the parameter that summarizes the magnitude of genetic drift and increase in inbreeding occurring in a population. In this paper, developments in the prediction equations for the effective size of populations subdivided under various models are reviewed, and extensions are made in several cases. Derivations are shown for some simple models, and the relationships among these equations and with those for a single unsubdivided population are discussed. The effect of population subdivision on neutral genetic variation and its implications are explained.

Keywords: conservation, effective size, genetic drift, heterozygosity, inbreeding, migration.

Contents

- 1 Introduction
- 2 Population structure and genetic differentiation
- 3 Effective size
 - 3.1 Equal contribution from subpopulations
 - 3.1.1 *Island model*
 - 3.1.2 *Stepping stone model*
 - 3.1.3 *The neighbourhood model*
 - 3.1.4 *The general model with an equal contribution from subpopulations*
 - 3.2 Variable contributions from subpopulations
 - 3.2.1 *Effective size for monoecious species*
 - 3.2.2 *Constant and equal subpopulation size*
 - 3.2.3 *Extinction and recolonization*
 - 3.3 The pattern of inheritance
 - 3.3.1 *Dioecious species*
 - 3.3.2 *Sex-linked loci*
 - 3.3.3 *Haploid species*
 - 3.3.4 *Non-nuclear gene inheritance*
- 4 The effects and implications of population subdivision
 - 4.1 The effects
 - 4.2 Implications
- 5 Some other considerations of subdivision
 - 5.1 Complete subdivision
 - 5.2 More than two levels of subdivision
 - 5.3 Realization of the asymptotic effective size

References

1 Introduction

Effective population size is a key parameter in population and quantitative genetics. It has important applications in evolution theory, domestic animal and plant breeding and conservation biology. As a measure of the strength of the stochastic process in a finite population, effective size determines the rate of decay of neutral genetic variation. It also affects the behaviour of genes under selection and other systematic forces, and thus influences the variance of selection response (Hill, 1985), the selection limits (Robertson, 1960) and the short- and long-term survival of populations under conservation (Lande & Barrowclough, 1987; Lynch *et al.*, 1995).

The concept of effective size was introduced by Wright (1931) and developed subsequently by Crow and colleagues (see Crow & Kimura, 1970, pp. 345–364). They distinguished between a variance effective size, which predicts the variance of change in gene frequency resulting from one generation of genetic sampling, and an inbreeding effective size, which predicts the rate of decrease in heterozygosity. In simpler cases, the two effective sizes are the same but, for more complex situations, it is necessary to

*Correspondence. E-mail: jinliang.wang@ed.ac.uk

make a distinction (Kimura & Crow, 1963a). Since this pioneer work, there has been much progress on developing equations to predict this important parameter. A long list of factors, such as sex ratio, mating system, variance in reproductive success, fluctuation of population size over generations, selection, overlapping generations and pattern of inheritance (e.g. haploid, polyploid, sex-linked) have been identified, and their effects on effective size quantified. These developments were reviewed by Caballero (1994). However, an important factor, the subdivision of populations, was only included briefly in his review.

The importance of population subdivision is reflected in both its substantial impact on effective size (as shown below) and its ubiquity among organisms. Many species naturally form subpopulations in the form of herds, flocks, schools, colonies or other types of aggregations because of intrinsic factors, such as behavioural segregation. In addition, natural habitats are typically patchy, with favourable areas interspersed among unfavourable areas. Even in the ideal case of a uniformly favourable habitat, the population is still subdivided to some extent if the habitat is large, because the distance of individual migration is usually much smaller than the entire range of the habitat; individuals are more likely to reproduce locally, leading to the so-called 'isolation by distance' (Wright, 1943). Now, more and more habitats are fragmented by human activities, and many organisms live in more or less isolated islands.

Recognizing the prevalence of population subdivision and its effects on genetic variation, several authors have developed prediction equations for effective size of subdivided populations (also called metapopulations) in recent years. In this paper, we will review these equations and their inter-relations and implications for conservation biology, evolutionary theory and breeding applications. Derivations and extensions to previous developments are made for some cases. Particular attention is paid to keeping the symbols, parameters and equations for subdivided populations consistent with those reviewed previously (Caballero, 1994) for a single unsubdivided population

2 Population structure and genetic differentiation

The effective size of a metapopulation is the size of a Wright–Fisher idealized population that would give rise to the variance of change in gene frequency or the rate of inbreeding observed in the actually subdivided population under consideration. An

idealized population is a monoecious population with constant size over discrete generations, random mating including selfing in random amounts, an equal probability of contributing gametes to the next generation from different parents and with respect to autosomal loci without mutation and selection (Fisher, 1930; Wright, 1931; see Caballero, 1994). The main violation of the above ideal situation in a subdivided population is nonrandom mating; mating is more likely to occur within subpopulations rather than among subpopulations. This kind of nonrandom mating will inevitably lead to genetic differentiation (the difference in gene frequency) among subpopulations, which is reflected as the relative difference between the average observed heterozygosity over subpopulations and the theoretical heterozygosity assuming no subdivision (completely random mating). Genetic differentiation may result from natural selection favouring different genotypes in different subpopulations and may also result from random genetic drift acting independently within local demes.

Migration, the gene flow among subpopulations, tends to impede genetic differentiation. As a homogenizing factor, it offsets the differentiating forces so that the population may reach an equilibrium. With one generation of complete migration (random mating in the entire population), the differentiation among subpopulations disappears completely. With no migration, the subpopulations differentiate continuously until alleles become either lost or fixed in different subpopulations. We are generally interested in incomplete subdivision in which genetic differentiation among subpopulations is counterbalanced by a certain degree of gene flow.

As another homogenizing force, mutation has little effect on genetic differentiation compared with migration, because it occurs at a rate of the order of 10^{-5} or less. Throughout the paper, we concentrate on selectively neutral variation and consider genetic drift as the sole differentiating force and migration among subpopulations as the sole homogenizing force.

The interaction between genetic drift and migration, which determines the equilibrium differentiation, the steady-state rate of decay of genetic variation and effective size, is largely influenced by the hierarchy of population structure. A hierarchical structure means that individuals can be grouped into progressively inclusive (nested) levels such as sublines, lines and the whole population. The actual geographical structure of a natural population may be quite complicated. For mathematical tractability, some simplified (idealized) population structure

models, such as the island model, stepping stone model and neighbourhood model, have been proposed. Results from these models are, however, still pertinent to the real world and have implications for evolutionary theory, plant and animal breeding and conservation biology.

Given the geographical structure and demographic properties of a population, its genetic differentiation will eventually reach an equilibrium after a sufficiently long time. Wright's F -statistics (Wright, 1951) are very useful in describing the genetic structure (or the partition of genetic variation) of a metapopulation irrespective of its geographical structure, the degree of isolation among subdivisions and without getting into details such as gene and genotypic frequencies. As will be shown below, the F -statistics can also be used to derive the effective size of metapopulations, making the expressions more concise and meaningful. F -statistics can be predicted using information on the geographical structure and demographic properties of the population and can also be estimated using gene and genotypic frequency data on allozyme and DNA polymorphisms (see Hartl & Clark, 1997).

3 Effective size

The variance and inbreeding effective sizes are not always the same in a metapopulation. However, if the total metapopulation is not completely subdivided, and its size and structure are constant over generations, then, after a sufficient number of generations, the rate of inbreeding and the variance of change in gene frequency will gradually reach their respective asymptotic values, which correspond to a unique effective size called eigenvalue effective size (Ewens, 1982) or asymptotic effective size (Chesser *et al.*, 1993; Wang, 1997a,b). The long-term effect of drift and inbreeding in a metapopulation on neutral or weakly selected alleles can be predicted by the asymptotic effective size. In this part, we are concerned only with the asymptotic effective size, which will be called effective size (N_e) for simplicity hereafter.

3.1 Equal contribution from subpopulations

Most prediction equations for the effective size of a subdivided population assume, explicitly or implicitly, that all subpopulations contribute equally to the next generation and have an equal and constant size over generations. These assumptions simplify the prediction considerably, although they may not be true in natural populations.

3.1.1 Island model Wright (1943) considered a population subdivided into n subpopulations, each being an idealized population of size N except for receiving a proportion m of immigrants taken randomly from the entire population per generation. This is the so-called island model. He derived, using the approach of the variance in gene frequency, an expression for effective size

$$N_e = nN \left[1 + \frac{(n-1)^2}{4Nmn^2} \right], \quad (1)$$

which was also derived by Nei & Takahata (1993) using Slatkin's (1991) formula for the coalescence time of two randomly chosen homologous genes. Equation (1) indicates that subdivision always results in an increase in N_e with this model. When Nm is large, $N_e \approx nN$, as expected for a panmictic population; but N_e can be much greater than the metapopulation size (nN) when Nm is small. If $n \gg 1$, we have

$$N_e = nN + n/(4m) \quad (2)$$

approximately. When migration is low ($4Nm \ll 1$), each subpopulation acts as a single individual, and $N_e \approx n/(4m)$ is mainly determined by the number of subpopulations and migration rate and can be much larger than the census size nN . However, a strong assumption made in deriving eqns (1) and (2) is that different subpopulations contribute equally to the next generation. As will be shown below, violation of this assumption may lead to the opposite.

Natural populations are seldom so ideal; there may be nonrandom mating, non-Poisson variance in reproductive success and other complications within subpopulations as reviewed by Caballero (1994). Wang (1997a) considered a monoecious metapopulation with partial selfing at rate β , an arbitrary distribution of family size (with variance S_k^2) and both pollen and seed migration with rates d_p and d_s , respectively. The resulting equation for N_e is complicated, but a good approximation can be obtained from eqn (2) by replacing N with the effective size of a subpopulation (N_{es}), i.e.

$$N_e = nN_{es} + n/(4m), \quad (3)$$

where the total migration rate is $m = \frac{1}{2}d_p + d_s$. The effective size of a subpopulation is

$$N_{es} = \frac{4N}{2(1 - F_{IS}) + S_k^2(1 + F_{IS})}, \quad (4)$$

where

$$F_{IS} = \frac{\beta}{2 - \beta} \quad (5)$$

(see Caballero, 1994). Equation (3) shows that any factor that acts to increase (or decrease) the effective size of subpopulations will also increase (or decrease) the effective size of the metapopulation. However, this is important only when migration is high ($4N_{eS}m \gg 1$).

For a subdivided dioecious population with different numbers of males (N_m) and females (N_f) within subpopulations and different migration rates of males (d_m) and females (d_f), predictions of N_e have been made by Chesser and colleagues for equal family size and Poisson distribution of family size (Chesser *et al.*, 1993) and by Wang (1997b) for an arbitrary distribution of family size. To a good approximation, Wang's eqn (28) also reduces to eqn (3) above with $m = \frac{1}{2}(d_m + d_f)$ and $N_{eS} = 16 / (P_{m,mm} + 2P_{m,mf} + P_{m,ff} + P_{f,mm} + 2P_{f,mf} + P_{f,ff})$ (Nagylaki, 1995), where $P_{u,vw}$ is the probability that two individuals of sexes v and w taken at random from the same subpopulation before migration come from the same parent of sex u ($u, v, w = m$ for males, f for females). $P_{u,vw}$ is given by

$$P_{u,vv} = \frac{1}{N_v - 1} \left[\left(\frac{N_u}{N_v} \right) \sigma_{uv}^2 + \frac{N_v}{N_u} - 1 \right]$$

and

$$P_{u,mf} = \left(\frac{N_u}{N_m N_f} \right) \sigma_{u,mf} + \frac{1}{N_u},$$

where σ_{uv}^2 denotes the variance of the number of offspring of sex v from a parent of sex u , and $\sigma_{u,mf}$ represents the covariance of the numbers of male and female offspring from a parent of sex u .

Equation (3) is only an approximate prediction. From exact solutions of recurrence equations for the probability of identity by descent (PIBD) or more accurate equations for N_e , it has been shown that male migration is more important than female migration when $N_m < N_f$ (Wang, 1997b). This is because a male immigrant is expected to contribute a higher proportion of genes to the next generation than a female migrant when $N_m < N_f$. However, the difference between the effects of male and female migration is small.

3.1.2 Stepping stone model In natural populations, individuals are often distributed more or less discontinuously to form colonies (subpopulations), and individuals or gametes are mainly exchanged

between adjacent or nearby colonies. Kimura (1953) first proposed a model, called the 'stepping stone model', to analyse such a structure. Based on the spatial arrangements of the subpopulations, the model is classified into one, two, three or higher dimensions. Maruyama (1970) considered a monoeious population subdivided into an even number of n colonies, each of size N and arranged in a circular stepping stone model. Migration between adjacent colonies occurs at a rate of $\frac{1}{2}m$ in each generation ($m \neq 0$). By two different methods, Maruyama (1970) derived that

$$N_e \approx \frac{n^2}{2m\pi^2}, \text{ if } 2Nm < n/\pi^2, \quad (6)$$

and

$$N_e \approx nN, \text{ if } 2Nm > n/\pi^2. \quad (7)$$

These equations show that, when the absolute number of migrant gametes from (or to) each subpopulation ($2Nm$) is small, N_e is approximately proportional to n^2 , i.e. the square of the number of subpopulations, as pointed out by Kimura & Crow (1963b). On the contrary, when $2Nm$ is large, there is little differentiation among subpopulations, and the metapopulation behaves as a large panmictic population. The equality $2Nm = n/\pi^2$ represents a turning point at which one of the equations becomes valid and the other one breaks down. At the turning point, however, both approximate equations give the same and the least level of accuracy, underpredicting N_e by a factor of 2 (Maruyama, 1970).

Comparing eqns (6) and (7) with eqn (2), we see that the stepping stone model is, in fact, similar to the island model, in that a turning point determined by the effective number of migrant gametes ($2N_{eS}m$) and the number or squared number of subpopulations divides the population into two classes, effectively panmictic and subdivided populations. Extrapolating the expression of effective size for the stepping stone model from that for the island model, we obtain

$$N_e = nN + \frac{n^2}{2m\pi^2}. \quad (8)$$

Although eqn (8) is not derived analytically, it provides better predictions than eqns (6) and (7) across the whole range of all parameters (N , n and m), the greatest improvement occurring at the turning point, as expected. This can be verified by the numerical examples of Maruyama (1970; tables 1 and 3).

Prediction equations for N_e for other stepping stone models are not available, but the relation between N_e and n , N and m should, in principle, be similar to that shown above. When subpopulations are not ideal, N_{es} should be used instead of N in eqn (8).

3.1.3 The neighbourhood model The island and stepping stone models assume discrete subpopulations, i.e. there is an area between colonies that is not inhabited. A natural population may have a continuous distribution of its individuals across the habitat, but it is not necessarily a random mating unit, because the distance of individual migration is usually much smaller than the entire distribution range of the population. Wright (1943, 1946) considered a model in which a population is distributed uniformly over a large territory, but parents of any given individual are drawn at random from a small surrounding region. The individuals in this region in a continuum constitute a random mating unit, called a 'neighbourhood' (Wright, 1946). The neighbourhood size is determined by the migration distance (or rather the variance in migration distance) and population density (number of individuals per unit area). For a diploid, monoecious population distributed uniformly along a circular (one-dimensional) habitat, Maruyama (1971) derived that

$$N_e \approx \frac{N_T L}{\pi^2 D \sigma^2}, \text{ if } D \sigma^2 < L / \pi^2, \quad (9)$$

and

$$N_e \approx N_T, \text{ if } D \sigma^2 > L / \pi^2, \quad (10)$$

where σ^2 is the variance of dispersion distance, D is the population density, L is the length of habitat and $N_T = LD$ is the metapopulation size. When the population is distributed on a linear habitat with two ends, eqns (9) and (10) and the conditional inequalities hold if L is replaced by $2L$ (Maruyama, 1971). Therefore, when $D \sigma^2$ is small, the effective size of a linear population is twice that of a circular population. When $D \sigma^2$ is large, however, both populations behave as panmictic units and have the same effective size, N_T . The turning point that determines roughly whether the population behaves like a panmictic one or not depends on both $D \sigma^2$ (a local property) and L (a property of the metapopulation or habitat). This property is similar to that of the circular stepping stone model (see eqns 6 and 7). With the same arguments given for the stepping stone model, we immediately see that the equations

before and after the turning point could be combined to give more general predictions,

$$N_e = N_T \left(1 + \frac{L}{\pi^2 D \sigma^2} \right), \quad (11)$$

for a circular structured population, or the same expression replacing L with $2L$ for a linear structured population.

For a population distributed uniformly and continuously on a habitat of a torus-like surface (two-dimensional) with size $L \times L$ and density D , numerical examples indicated that

$$N_e \approx \frac{N_T}{D \sigma^2}, \text{ if } D \sigma^2 < 1 \quad (12)$$

and

$$N_e \approx N_T, \text{ if } D \sigma^2 > 1 \quad (13)$$

(Maruyama, 1972). Comparing eqn (9) with eqn (12), we note that, for a given size N_T and a given small $D \sigma^2$, one-dimensional structures result in a much larger N_e than two-dimensional structures. When the habitat occupied by a population is rectangular, eqns (12) and (13) and the conditional inequalities are valid if $D \sigma^2$ is replaced by $\frac{1}{2} D \sigma^2$ (Maruyama, 1972). Therefore, for small values of $D \sigma^2$, a rectangular structure results in a larger N_e than the torus-like structure. By analogy with eqns (8) or (11), a combination of eqns (12) and (13) gives more general and accurate predictions.

In contrast to one-dimensional populations (eqns 9 and 10), the turning point for two-dimensional populations depends only on $D \sigma^2$, independently of the habitat size. In this respect, two-dimensional, continuously distributed populations are similar to populations in the island model (see eqn 2).

3.1.4 The general model with an equal contribution from subpopulations In the above, different geographical structures of a metapopulation lead to different prediction equations for effective size. The parameters in these equations, such as migration rate and variance in dispersion distance, are difficult to estimate in practice (Slatkin, 1985). Moreover, if the number of dimensions in the stepping stone model and neighbourhood model exceeds two, it is also difficult to derive an analytical expression for effective size. Furthermore, the structure of natural populations may be much more complicated than that described by the three models. Therefore, a general expression for effective size that is indepen-

dent of the geographical structure of the population will be very useful.

F -statistics describe the genetic architecture of a subdivided population in any geographical structure, and they can be estimated using protein or DNA polymorphism data. If effective size can be expressed in terms of F -statistics, there will be no need to consider the geographical structure of the population and to estimate m and σ^2 . Here, we derive such an equation for effective size, according to the approach of the variance in gene frequency (although a similar derivation can be made by the inbreeding approach).

Assume that a population is subdivided into n subpopulations with the same census size N and effective size N_{eS} . Let q_i and q'_i be the frequencies of a given allele at an arbitrary neutral locus in the i th subpopulation in generations t and $t+1$, respectively. The change in gene frequency caused by genetic drift in the i th subpopulation is $\Delta q_i = q'_i - q_i$. The variance of change in gene frequency can be derived using a procedure shown in Caballero (1994) as $V_{\Delta q_i} = q_i(1 - q_i)/(2N_{eS})$. The mean change in gene frequency for the metapopulation is

$$\Delta q = \frac{1}{n} \sum_{i=1}^n \Delta q_i$$

Ignoring the second-order terms introduced by correlations among the Δq_i because of the finite number of subpopulations, we obtain

$$V_{\Delta q} = \frac{1}{n^2} \sum_{i=1}^n [V_{\Delta q_i}] = \frac{\sum_{i=1}^n q_i(1 - q_i)}{2n^2 N_{eS}}. \quad (14)$$

In the subdivided population, the average heterozygosity is

$$\frac{2}{n} \sum_{i=1}^n q_i(1 - q_i) = 2(1 - F_{ST})\bar{q}(1 - \bar{q})$$

(Hartl & Clark, 1997), where

$$\bar{q} = \frac{1}{n} \sum_{i=1}^n q_i$$

Inserting the relation into eqn (14) and equating $V_{\Delta q}$ to $\bar{q}(1 - \bar{q})/(2N_e)$ by definition, we obtain

$$N_e = \frac{nN_{eS}}{1 - F_{ST}}. \quad (15)$$

Equation (15) is essentially the same as that derived by Wright (1943) except for N_{eS} instead of

N . Therefore, the effective size of a subdivided population turns out to be very simple in terms of F -statistics and effective subpopulation size. All the complications within a subpopulation (such as nonrandom mating, variance in family size, etc.) can be taken into account by using the effective subpopulation size, which has been reviewed by Caballero (1994). For a monoecious population with partial selfing rate β , for example, N_{eS} is given by eqns (4) and (5), and the effective size of the metapopulation is

$$N_e = \frac{4nN}{(1 - F_{ST})[2(1 - F_{IS}) + S_k^2(1 + F_{IS})]}. \quad (16)$$

3.2 Variable contributions from subpopulations

Most previous analyses on metapopulations have assumed that all subpopulations contribute an equal number of offspring to the next generation and have the same constant size over generations. Natural populations are seldom so ideal, and all of their demographic parameters may change in space and time. In the extreme case, a particular subpopulation may become extinct and be replaced by migrants from other subpopulations. A general model has been proposed and analysed by Whitlock & Barton (1997) recently. They found that, when subpopulations are allowed to contribute differentially, subdivision may lead to a great decrease in N_e . This was also found by Gilpin (1991) and Hedrick & Gilpin (1997) using simulation studies incorporating both realistic features of population ecology (e.g. extinction and recolonization) and population genetics.

3.2.1 Effective size of monoecious species Whitlock & Barton (1997) considered a monoecious population subdivided into n subpopulations, whose sizes may differ and change over generations. Each subpopulation was assumed to be an ideal Wright-Fisher population but, as shown below, their model can easily be extended to non-ideal situations.

Consider a population subdivided into n subpopulations. In each discrete generation, reproduction is followed by migration. At generation t , the i th subpopulation has N_i individuals and contributes a total number of N'_i progeny to the next generation. Therefore, the fitness of the i th subpopulation at generation t is $w_i = N'_i/N_i$, with

$$\sum_{i=1}^n w_i = n$$

(the metapopulation size is constant). Let F_{ij} be the probability of identity by descent (PIBD) of two alleles chosen at random from subpopulations i and j , respectively, at generation t . After reproduction but before migration, the average PIBD for alleles taken from the i th subpopulation is increased to

$$F'_{ii} = \frac{1}{2N_{eSi}} + \left(1 - \frac{1}{2N_{eSi}}\right) F_{ii} \tag{17}$$

where N_{eSi} is the effective size of subpopulation i . The PIBD for alleles from different subpopulations remains the same,

$$F'_{ij} = F_{ij} \tag{18}$$

The overall average PIBD for alleles taken at random from the entire metapopulation is

$$\bar{F} = \sum_i \sum_j \frac{N_i N_j F_{ij}}{i^2 \bar{N}^2} \tag{19}$$

for generation t and

$$\bar{F}' = \sum_i \sum_j \frac{N'_i N'_j F'_{ij}}{i'^2 \bar{N}'^2} \tag{20}$$

for generation $t+1$, where \bar{N} is the mean number of individuals per subpopulation. Inserting eqns (17) and (18) into eqn (20) yields

$$\bar{F}' = \sum_i \frac{w_i^2 N_i^2 (1 - F_{ii})}{2i^2 \bar{N}^2 N_{eSi}} + \sum_i \sum_j \frac{w_i w_j N_i N_j F_{ij}}{i^2 \bar{N}^2} \tag{21}$$

Now, migration changes the PIBD for alleles within and among subpopulations, but the overall mean, \bar{F}' , remains the same. Therefore, it is needless to consider migration as far as \bar{F}' is concerned. Noting that the overall mean PIBD for alleles in generation t can be expressed as

$$\bar{F} = \sum_i \sum_j \frac{N_i N_j \bar{F}}{i^2 \bar{N}^2},$$

we obtain the rate of change in the overall mean PIBD from eqns (19) and (21),

$$\frac{\bar{F}' - \bar{F}}{1 - \bar{F}} = \sum_i \frac{w_i^2 N_i^2 (1 - F_{ST,ii})}{2i^2 \bar{N}^2 N_{eSi}} + \sum_i \sum_j \frac{w_i w_j N_i N_j F_{ST,ij}}{i^2 \bar{N}^2}, \tag{22}$$

where $F_{ST,ij} = (F_{ij} - \bar{F}) / (1 - \bar{F})$ is a generalization of Wright's F -statistic, and $F_{ST,ii}$ is the F_{ST} for the i th subpopulation. From the definition of \bar{F} , it is easy to see that

$$\sum_i \sum_j N_i N_j F_{ST,ij} = 0.$$

Equating eqn (22) to $1/2N_e$ yields the asymptotic effective size of a metapopulation

$$N_e = \frac{n\bar{N}}{\sum_i \frac{w_i^2 N_i^2 (1 - F_{ST,ii})}{n\bar{N}N_{eSi}} + 2\sum_i \sum_j \frac{w_i w_j N_i N_j F_{ST,ij}}{n\bar{N}}}. \tag{23}$$

When each subpopulation is an ideal Wright-Fisher population, $N_{eSi} = N_i$ and eqn (23) reduces to Whitlock & Barton's eqn (9). In principle, all the parameters in eqn (23) are measurable or predictable in a single generation. The $F_{ST,ij}$ could be estimated from the same genetic data as F_{ST} , using an analysis of covariance. The reproductive values (w_i) are for a single generation's transition (for more details, see Whitlock & Barton, 1997).

In the derivation of eqn (23), no assumption was made about the geographical structure or migration pattern of the population. The variation in fitness among subpopulations may come from individual variation of contributions and temporary or permanent quality differences among the patches of a habitat occupied by different subpopulations. Therefore, eqn (23) can be used to estimate N_e for a natural population if enough demographic data and genetic data are available for a single generation. Two special and important cases of eqn (23) are discussed next.

3.2.2 Constant and equal subpopulation size Consider the case that each subpopulation has the same constant size ($N_i = \bar{N} = N$) but contributes unequally to the next generation via migration. If $w_i w_j$ is not correlated with $F_{ST,ij}$, we have $F_{ST,ij} = -F_{ST} / (n - 1)$ for $i \neq j$ (Whitlock & Barton, 1997) where

$$F_{ST} = \frac{1}{n} \sum_i F_{ST,ii}$$

The effective subpopulation size is

$$N_{eSi} = \frac{2N_i w_i}{[S_{ki}^2 / (2w_i) + 2w_i] (1 + F_{IS,i}) - (1 + 3F_{IS,i})}$$

(Wang, 1996), where S_{ki}^2 is the variance in the number of gametes contributed per individual in subpopulation i . Inserting the above relations into eqn (23) yields

$$N_e = \frac{nN}{\frac{1}{4}(1 - F_{ST})[(1 + F_{IS})(S_k^2 + 4V) + 2(1 - F_{IS})] + 2NVF_{ST}n/(n - 1)}, \tag{24}$$

where V is the variance in fitness among subpopulations (variance of w_i),

$$F_{IS} = \frac{1}{n} \sum_i F_{IS,i}$$

and

$$S_k^2 = \frac{1}{n} \sum_i S_{ki}^2$$

With random mating and Poisson distribution of family size within subpopulations, we have $F_{IS} = 0$ and

$$S_k^2 = \frac{1}{n} \sum_i 2w_i = 2,$$

eqn (24) reduces to Whitlock & Barton's equation

$$N_e = \frac{nN}{(1 - F_{ST})(1 + V) + 2NVF_{ST}n/(n-1)}. \quad (25)$$

From eqn (25), it can be seen that, compared with an unsubdivided population of the same size nN , subdivision may increase or decrease N_e , depending simply on the variance in fitness among subpopulations. If subpopulations contribute equally to the next generation ($V = 0$), eqn (25) reduces to eqn (15) of the previous section with N replacing N_{es} , and N_e always being increased by subdivision. On the contrary, if $V \geq 1/(2N-1)$ approximately, subdivision will result in a decrease in N_e . A variance of $V \approx 1/(2N-1)$ refers to the situation in which there are no restrictions to the reproductive output of each subpopulation and no quality differences (temporary or permanent) among the patches of a habitat occupied by the metapopulation, and the distribution of reproductive success of individuals is Poisson.

Whereas the direction of the effect of subdivision on N_e depends on V , the magnitude of the effect relies on the genetic differentiation, F_{ST} . If $F_{ST} \rightarrow 0$, $N_e \rightarrow nN(1 + V) \approx nN$ because V is of the order $1/(2N)$; if $F_{ST} \rightarrow 1$, eqn (25) reduces to $N_e = (n-1)/(2V)$, which is equivalent to the N_e of an unsubdivided haploid population (Caballero, 1994, eqn 27) with the number of subpopulations taking the place of the number of individuals and the variance in the number of gametes being $2V$ instead of V . Because with $F_{ST} = 1$, each subpopulation is completely homozygous for a particular allele at a locus, it behaves as a haploid individual, in that no genetic drift results from Mendelian segregation, and as a diploid individual, in that the average contribution of gametes is two for stable population size. If each subpopulation contributes equally to the next generation, there is no genetic drift and the effective size is infinite.

For equal contributions from different individuals within a subpopulation ($S_k^2 = 0$) and from different subpopulations ($V = 0$), eqn (24) reduces to

$$N_e = \frac{2nN}{(1 - F_{ST})(1 - F_{IS})} = \frac{2nN}{1 - F_{IT}}, \quad (26)$$

indicating that inbreeding at any level (F_{ST} or F_{IS}) has an equal effect on increasing N_e . If, on the other hand, $S_k^2 = 2 + 2\beta$ (Poisson distribution of the reproductive success per individual when there is a proportion β of selfing) and $V = 1/N$ (i.e. twice the accumulated Poisson distribution variance), using eqn (5), eqn (24) reduces to

$$N_e = \frac{nN}{(1 + F_{ST})(1 + F_{IS}) - 2F_{IS}F_{ST}} = \frac{nN}{1 + F_{IT}} \quad (27)$$

approximately, indicating that inbreeding at any level has an equal effect on decreasing N_e . If $V = 0$ and $S_k^2 = 2 + 2\beta$, eqn (24) reduces to

$$N_e = \frac{nN}{(1 - F_{ST})(1 + F_{IS})} \quad (28)$$

approximately. Equations (27) and (28) were also derived by Nunney (1998) for a dioecious population using another inbreeding approach.

3.2.3 Extinction and recolonization The effective size of a subdivided population subject to extinction and recolonization was also derived by Whitlock & Barton (1997). The model assumes that a population is subdivided into n demes, each being an ideal population with N diploid individuals. The migration rate among demes is m , and each deme becomes extinct with probability e per generation. Each extinct deme is recolonized immediately by $2k$ gametes in which each pair has a probability ϕ of having come from the same deme. The newly colonized deme grows to N individuals before reproduction in the next generation (Slatkin, 1977; Whitlock & McCauley, 1990). With $\phi = 0$, the contribution of a deme to the next generation (N') is either 0 with probability e or $N(1-e)$ with probability $1-e$. Therefore, the variance in fitness among demes is $V = e/(1-e)$, and from eqns (23) or (25) we obtain that

$$N_e = \frac{nN(1-e)}{1 - F_{ST} + 2eF_{ST}N}. \quad (29)$$

Whitlock & Barton (1997) obtained from eqn (29) that

$$N_e = \frac{nN}{4N(m+e)F_{ST}} \quad (30)$$

approximately, which is also true relaxing the assumption about ϕ . For $\phi = 1$ and $k = N$, eqn (30) reduces to the equation given by Maruyama & Kimura (1980) without mutation. With no extinction ($e = 0$), eqn (30) reduces to eqn (15) with N taking the place of N_{eS} , as expected.

Although extinction and recolonization may increase or decrease F_{ST} (Whitlock & McCauley, 1990), they always decrease N_e . If $k \ll N$ and extinction is not rare relative to migration, N_e can be reduced substantially (Whitlock & Barton, 1997). For example, if $e = m = 0.1$, $N = 1000$ and $k = 1$, then the effective size is 0.0062 of what it would have been without local extinction.

3.3 The pattern of inheritance

3.3.1 Dioecious species The effective size of a subdivided population with separate sexes was derived by Nunney (1998) using an inbreeding approach. He assumed an equal size of subpopulations but with unequal contributions to the next generation. The subpopulation was allowed to have different numbers of males and females with a constant sex ratio over generations, partial sib-mating and arbitrary distributions of family size. A more general equation can be obtained using a derivation similar to the above monoecious case. The expression derived is essentially similar to eqn (23). When each subpopulation consists of N individuals (half in each sex) and the fitness of subpopulations is not correlated with F -statistics, the expression reduces to

$$N_e = \frac{nN}{\frac{1}{4}(1 - F_{ST})[(1 + 3F_{IS})(S_k^2 + 4V) + 2(1 - F_{IS})] + 2NVF_{ST}n/(n-1)}. \quad (31)$$

When the subpopulations have an equal size and are ideal populations except for unequal numbers of males (N_m) and females (N_f) and variable contributions to the next generation, the general expression is then simplified to

$$N_e = \frac{nN}{(1 - F_{ST})(1 + V)N/N_{eS} + 2NVF_{ST}n/(n-1)}, \quad (32)$$

where $N = N_m + N_f$ and $N_{eS} = 4N_mN_f/N$. Compared with an unsubdivided population of the same size, subdivision will decrease the effective size when $V \geq 1/(2N_{eS} - 1)$.

3.3.2 Sex-linked loci Similar to the case of a single unsubdivided population (Caballero, 1994), the effective size of a metapopulation for sex-linked loci or haplo-diploid species relates to the inbreeding rate of the homogametic sex or to the variance of gene frequency at a neutral sex-linked locus.

Assuming all subpopulations have an equal and constant size in the island model, Wang (1998) derived exact recurrence equations for the PIBD of two homologous genes from a female (the homogametic sex) and from separate females within and among subpopulations. Different numbers and separate migration rates of males (d_m) and females (d_f) were incorporated, with random mating within subpopulations. The derived expression for effective size is complicated, but to a good approximation it reduces to eqn (3) using an appropriate effective subpopulation size N_{eS} (see eqn 30 in Caballero, 1994) and the total migration rate $m = \frac{1}{3}(d_m + 2d_f)$. Female migration rate has a larger effect on N_e than male migration rate, because a female carries two genes and a male carries only one gene at a sex-linked locus.

When there are fitness differences among subpopulations, it can be shown that eqn (23) also applies for sex-linked loci, with N and F_{ST} referring to the homogametic sex only. Therefore, results and discussions above about autosomal loci are applicable to the sex-linked case.

3.3.3 Haploid species For haploid species, genetic drift in gene frequency or an increase in PIBD comes only from differential reproductive successes of individuals and different fitnesses among subpopulations. In the extreme case of equal contributions from different individuals and different subpopulations, the gene frequency remains the same, and the effective size is infinite.

Using a procedure similar to the monoecious case, an equation identical to eqn (23), except for the factor 2 of the second term in the denominator being dropped, can be derived. If subpopulations have an equal size but contribute unequally to the next generation via migration, the equation reduces to

$$N_e = \frac{nN}{(1 - F_{ST})(V + S_k^2) + NVF_{ST}n/(n-1)}, \quad (33)$$

where S_k^2 is the within-subpopulation variance in the number of gametes contributed per individual averaged over subpopulations. The effective size for a haploid species depends on the variance in genetic contributions at both the individual and the

subpopulation levels. The effect of the subpopulation level variation is amplified by the differentiation among subpopulations. If the differentiation is complete ($F_{ST} = 1$), each subpopulation behaves as a single individual, and the effective size is determined by the number of subpopulations and their differential contributions [$N_e = (n-1)/V$]. If the subpopulation level variance comes only from the individual level ($V = S_k^e/N$), then subdivision has little effect on the effective size of a metapopulation.

3.3.4 Non-nuclear gene inheritance Birky *et al.* (1989) investigated organelle gene diversity in subdivided populations, and Chesser & Baker (1996) considered the asymptotic effective size of a subdivided population for non-nuclear gene inheritance. If organelle genes (e.g. mtDNA) are inherited uniparentally and are always homologous within individuals, then they are similar to haploid inheritance, and the effective size for organelle genes can be predicted from the corresponding haploid equations (e.g. eqn 33), considering only the relevant sex.

Although homoplasmy for organelles is considered to be near ubiquitous in higher vertebrates (Avise, 1991), an increasing number of species are found to be heteroplasmic (Hartl & Clark, 1997). In this case, the genetic drift theory for organelles becomes more complex than haploid or diploid nuclear inheritance. An individual cell may have different organelles that are partitioned among daughter cells. Thus, the equations for either diploid or haploid populations are not applicable to the organelle genes.

4 The effects and implications of population subdivision

4.1 The effects

The effect of subdivision is similar to nonrandom mating in a single population, with a subpopulation corresponding to a family and F_{ST} to F_{IS} . The effective size for a monoecious population with partial selfing and a dioecious population with partial full-sib mating can be summarized as

$$N_e = \frac{N_T}{\frac{1}{2}(1 - F_{IS}) + V[(2N - 1)F_{IS} + 1]} \quad (34)$$

[Caballero, 1994, eqns (11) and (20), with $S_k^e/4$ replaced by V], where N_T is the total population size, N is the number of individuals per family ($N = 1$ and 2 for selfing and full-sib mating, respectively) and V is the variance in fitness among families. For a subdivided population with random

mating within subpopulations ($F_{IS} = 0$), eqns (24) or (31) reduce to

$$N_e = \frac{N_T}{\frac{1}{4}(1 - F_{ST})(S_k^e + 2) + V[(2N - 1)F_{ST} + 1]} \quad (35)$$

approximately. Note that, to be comparable with eqn (34), in which individuals within a family contribute equally to the next generation, the individual variance S_k^e should be set to zero. Thus, eqn (35) reduces to eqn (34) with F_{ST} taking the place of F_{IS} .

Knowing the similarity between subdivision and partial inbreeding, we can understand the effect of subdivision on effective size from the point of view of genetic drift. The original effect of subdivision is nonrandom mating; reproduction occurs mainly within subpopulations, resulting in a decrease in heterozygosity. Genetic drift in gene frequency can come from two sources in a finite population: one is the differential reproduction among subpopulations and among individuals within subpopulations; and the other is the Mendelian segregation of heterozygotes. A decrease in heterozygosity from subdivision will definitely lead to a decline in genetic drift resulting from segregation. With a given gene frequency, however, a decrease in heterozygosity also implies an increase in the variance of gene frequencies among subpopulations and among individuals within subpopulations, and therefore an increase in genetic drift because of differential reproduction. The net effect of a decrease in heterozygosity on genetic drift depends on the relative magnitudes of the counteracting effects from the two sources. When the variance in reproduction success and, thus, its effect on drift are small, segregation is the leading source of genetic drift and, therefore, subdivision results in an increase in effective size. In the extreme case of equal reproductive success ($V = S_k^e = 0$), genetic drift comes solely from segregation and is therefore minimized. Thus, subdivision gives the greatest increase in effective size, $N_e = 2N_T/(1 - F_{ST})$ from eqn (35), and $N_e \rightarrow \infty$ when $F_{ST} \rightarrow 1$. On the contrary, when the variance in reproductive success is large [$V > 1/(4N - 2)$], genetic drift mainly comes from differential reproduction among subpopulations and, therefore, subdivision always leads to a decrease in N_e . There exists, in fact, a balance (equilibrium) between the counteracting effects from the two sources, at which a change in heterozygosity resulting from subdivision has no influence on effective size. The condition for the balance can be found from eqn (35) as $V = 1/(4N - 2)$. With this value of V , the effects of segregation and differential reproduction cancel out

and $N_e = n(2N-1)$ from eqn (35), irrespective of the nonrandom mating (F_{ST}).

When ecological factors are considered, subdivision may well enlarge the variance in reproductive output among demes (V) in natural populations. Without local density regulation, V comes from three sources. First, individual variance in offspring number contributes an amount of $S_i^2/(4N)$ to V . This is the variance in reproductive output of arbitrarily defined groups of individuals. Secondly, temporary environmental differences among demes contribute to V . These two components of V do not cause correlation in the change of gene frequency over generations. Thirdly, there may be permanent quality differences among demes. If a subpopulation fortunately (unfortunately) inhabits an area of high (low) quality, it will contribute more (fewer) offspring consistently over generations. Unless migration is complete ($m=1$), the permanent difference in habitat quality that influences reproductive success will tend to increase the long-term genetic drift and inbreeding.

The permanent difference in fitness among subpopulations is similar to the inherited variation under selection (Santiago & Caballero, 1995). They both result in correlated changes in gene frequency or gene identity over generations and, therefore, N_e is greatly reduced. The patterns of correlated changes in fitness associated with an allele across generations are different, however, for inherited fitness and for permanent habitat difference. In the former case, the correlated change between two successive generations is gradually dissipated by recombination, whereas in the latter case, the correlated change is kept completely as long as the allele remains in the same subpopulation and is totally lost once it moves to a new subpopulation. Recombination plays a role similar to migration in reducing the correlated changes, the magnitude of effect lying somewhere between the island model and the stepping stone model (Whitlock & Barton, 1997). Subdivision with permanent difference in fitness among subpopulations is also analogous to hitchhiking or background selection, in which a neutral allele finds itself associated with one of the genetic backgrounds under directional selection.

Although subdivision usually tends to increase variance in genetic contributions among demes in natural populations, it might also be used to decrease the variance in long-term contribution from different individuals in controlled populations. This is best illuminated by the following example. To minimize the inbreeding and genetic drift in a small population with N_m males and N_f females (sex ratio

$r = N_f/N_m \geq 2$), Gowe *et al.* (1959) proposed a selection scheme in which each male has one son and r daughters, and each female has one daughter and a probability of $1/r$ of contributing one son. The selection scheme combined with random mating (each male mates at random with r females) has been known as minimal inbreeding (Falconer & Mackay, 1996, p. 69) and considered ideal for both control populations and conserved populations. Under minimal inbreeding, the variance in paternal contribution to the next generation is always zero. However, the paternal contribution to the gene pool after two or more generations is still variable. For example, males may have different numbers of grandsons resulting from the differential contributions of their daughters. If the population is subdivided into N_m herds, each consisting of N_f females and one male, and males migrate randomly among herds and females do not migrate, then the variance in paternal long-term contribution is also zero under Gowe *et al.*'s (1959) selection scheme (Wang, 1997c). Therefore, the subdivision scheme can increase the effective size by as much as 12.5% for autosomal loci and 50% for sex-linked loci, compared with random mating (Wang, 1997c). The subdivision scheme is very efficient in minimizing inbreeding, not only in the long term but also in the short term (especially for large r), in contrast to the usual subdivision with equal sizes and contributions of all subpopulations (section 3.1) or circular pair mating (Kimura & Crow, 1963b), which results in an increased N_e with a sacrifice of a higher rate of inbreeding at initial generations.

4.2 Implications

The above results imply that subdivision in natural populations is much more likely to result in a decrease rather than an increase in N_e . Classical models assume a constant size and an equal contribution of different subpopulations, resulting in an increase in N_e because of subdivision, particularly at low levels of gene flow (section 3.1). The assumption is, however, unlikely to be realistic in natural populations in which subpopulations may contribute unequally to the next generation. There are generally quality differences across the habitat that increase the variation in subpopulation productivity, and permanent quality difference or extinction and recolonization can substantially amplify the reproductive variance among subpopulations and thus decrease N_e enormously. The implication is that N_e/N of natural populations should be smaller than one, and the lower the level of gene flow in a

species, the smaller the estimate of its N_e/N . A review of 192 estimates of N_e/N from 102 species of plants and animals (Frankham, 1995) shows that the estimates vary dramatically, ranging from 10^{-6} in the pacific oyster (but see Nunney, 1996) to 0.99 in humans, with an overall average of only 0.1. These estimates, however, do not account for subdivision and gene flow, and some of them may therefore be biased.

In conservation, if the habitat of a species is permanently fragmented and migration among subpopulations is not possible except by carefully planned management, N_e of the species could be increased compared with a panmictic population of the same size. In this case, effort should be made to ensure that each subpopulation does not extinguish, and a certain level of artificially aided migration (Mills & Allendorf, 1996) is carried out to avoid inbreeding depression. The results reviewed in this paper suggest that subdivision and thus differentiation at different levels have a similar effect on N_e , the direction of the effect depending on the variance in reproductive success at the corresponding levels. To increase N_e , gene flow at a particular level of subdivision should be low, if reproductive variation at this level can be minimized by management, but gene flow should be high if reproductive variation cannot be minimized. However, even if effective size can be increased by restricted migration, it should be noted that this is realized only after many generations (see below), during which the inbreeding rate might be increased compared with panmixia.

At the planning stage of a conservation programme, whether a population should be subdivided or not depends on many factors. If the reproduction and migration of the population and the ecological factors could be managed intensively, then subdivision may be beneficial for conserving the genetic variation for a given population size. This is the case for domesticated species, in which a large number of local breeds are endangered by the worldwide spread of a few highly productive commercial breeds and their crosses. For wild species in a more or less natural habitat, it is difficult to practise intensive management. Therefore, it is generally safer to conserve the species in a single large population rather than in a number of subpopulations. Because of its small size, a subpopulation without careful management is more likely to become extinct because of demographic and environmental stochasticity (Lande, 1988) and inbreeding depression (Saccheri *et al.*, 1998), resulting in a drastic decrease in N_e of the metapopulation.

5 Some other considerations of subdivision

5.1 Complete subdivision

In the above, each subpopulation is not completely isolated from the rest; it receives immigrants from and contributes emigrants to the entire population. If a subpopulation only receives immigrants from but does not contribute emigrants to the metapopulation, the rate of inbreeding or genetic drift in this subpopulation will be smaller than that of the rest of the metapopulation. If, on the contrary, a subpopulation only contributes emigrants to but does not receive immigrants from the rest of the metapopulation, its effective size is determined by itself as a single population.

If all the subpopulations are isolated from each other, no asymptotic effective size for the entire metapopulation exists. In this case, the rate of inbreeding or genetic drift in each subpopulation will reach an asymptotic value, corresponding to the eigenvalue effective size of the subpopulation. The inbreeding effective size of the metapopulation (N_{ei}), signifying the rate of decrease in the average heterozygosity of individuals in the whole population, is a weighted average of the effective subpopulation sizes. The equilibrium variance effective size of the metapopulation (N_{ev}) is different from N_{ei} and is not reached until the differentiation is complete ($F_{ST} = 1$). With $F_{ST} = 1$, different alleles are fixed in different subpopulations, and the genetic drift is determined only by the differential contributions from subpopulations. For the case of an equal size and unequal contributions of subpopulations, we have $N_{ev} = (n-1)/(2V)$ for diploid or $N_{ev} = (n-1)/V$ for haploid species, and N_{ev} is infinite if subpopulations contribute equally to the next generation ($V = 0$).

5.2 More than two levels of subdivision

A population may be subdivided at more than two levels in a hierarchy. For simplicity, let us consider a monoecious population of constant size subdivided into s lines, each line subdivided into n sublines, and each subline containing N individuals. Different lines (sublines) have an equal size but may contribute unequally to the next generation, migration among sublines within lines being more frequent than that among lines. No line or subline is completely isolated from the rest of the metapopulation. It can easily be shown that the F -statistics satisfy the relation $(1 - F_{IT}) = (1 - F_{LT})(1 - F_{SL})(1 - F_{IS})$, where F_{IS} , F_{SL} , F_{LT} and F_{IT} are similarly defined as Wright's F -statistics with

subscripts I, S, L and T referring to individuals, sublines, lines and the total metapopulation, respectively. The effective size of the *i*th line can be obtained as

$$N_{eLi} = \frac{nN\mu_i}{\frac{1}{4}(1 - F_{SL,i})[(4V_{Si} + S_{ki}^2)(1 + F_{IS,i}) + 2\mu_i(1 - F_{IS,i})] + 2NV_{Si}F_{SL,i}n/(n-1)} \quad (36)$$

using a procedure similar to the derivation of eqn (24), where μ_i and V_{Si} are the average and variance of fitness of sublines within line *i*, $F_{IS,i}$ and $F_{SL,i}$ are the average F_{IS} and F_{SL} over sublines in line *i*, S_{ki}^2 is the variance in the number of gametes contributed per individual averaged over sublines within the *i*th line. The effective size of the metapopulation can be derived as

$$N_e = \frac{snN}{(1 - F_{LT})\{\frac{1}{4}(1 - F_{SL})[(1 + F_{IS})(S_k^2 + 4V_S) + 2(1 + V_L)(1 - F_{IS})] + 2NV_S F_{SL}n/(n-1)\} + 2nNV_L F_{LT}s/(s-1)} \quad (37)$$

where V_L is the variance in fitness among lines, F_{LT} , F_{SL} , F_{IS} , S_k^2 and V_S are averages of $F_{LT,b}$, $F_{SL,b}$, $F_{IS,b}$, S_{ki}^2 and V_{Si} over lines. If there is no variation in reproductive success at any level of subdivision ($S_k^2 = V_S = V_L = 0$), eqn (37) reduces to

$$N_e = \frac{2snN}{(1 - F_{LT})(1 - F_{SL})(1 - F_{IS})} = \frac{2snN}{1 - F_{IT}} \quad (38)$$

If individual reproductive success follows a Poisson distribution [$S_k^2 = 2 + 2\beta$, where β is the selfing proportion], V_S and V_L are twice the accumulated Poisson distribution variance ($V_S = 1/N$, $V_L = 1/nN$), and eqn (37) reduces to

$$N_e = \frac{snN}{(1 + F_{LT})(1 + F_{SL})(1 + F_{IS}) - 2F_{LT}F_{SL} - 2F_{LT}F_{IS} - 2F_{SL}F_{IS}} \\ = \frac{snN}{1 + F_{IT}} \quad (39)$$

approximately. In both cases, the effective size of the metapopulation depends on the total size (snN) and the total differentiation (F_{IT}); the number of levels of subdivision in the hierarchy is irrelevant. More generally, however, the importance of each level of subdivision in determining N_e depends on the differentiation and variance in fitness at the level, as can be seen from eqn (37).

5.3 Realization of the asymptotic effective size

In a Wright–Fisher ideal population with completely random union of gametes, the equilibrium rate of inbreeding and genetic drift or the asymptotic effective size is attained immediately (in one generation). Any form of nonrandom mating will retard the realization of the asymptotic effective size, inbreeding lagging some generations behind genetic drift when close inbreeding is avoided or vice versa (Caballero, 1994). In both cases, inbreeding and variance effective sizes are different and variable over the first few generations before they converge gradually to the same asymptotic value. The larger the departure from a random union of gametes, the greater the number of generations necessary to realize it. In the initial generations, neither inbreeding nor genetic drift can be predicted accurately by the asymptotic effective size. Other forms of nonrandom mating, such as age structure (Choy & Weir, 1978) and population subdivision (Chesser *et al.*, 1993; Wang, 1997a,b), and factors other than nonrandom mating, such as artificial selection, hitchhiking and permanent difference in fitness among subpopulations, also retard the realization of asymptotic effective size. Here, we consider population subdivision as a form of nonrandom mating and its effect on the time required to realize the equilibrium *F*-statistics and effective size.

For simplicity, let us consider Wright’s ideal infinite island model. The genetic differentiation among subpopulations will increase over generations until it reaches the equilibrium value (F_{ST}) asymptotically. The number of generations for the instantaneous differentiation ($F_{ST,i}$) to reach $F_{ST,i} = \lambda F_{ST}$ (where $0 < \lambda < 1$) is

$$t_i \approx -2NF_{ST} \ln(1 - \lambda) \quad (40)$$

when $N \gg 1$ and $m \ll 1$ (Whitlock, 1992). The length of the retardation is proportional to the subpopulation size and the equilibrium amount of differentiation. Equation (40) is also valid for a metapopulation with non-ideal subpopulations when the census size of a subpopulation is replaced by its effective size (N_{es}). The term $2N_{es}F_{ST}$ (that we may call the ‘effective genetic differentiation’) determines the length of time to reach the equilibrium rate of inbreeding and genetic drift. With $2N_{es}F_{ST} = 50$ for example, 35 and 150 generations are required for the subdivided population to realize 50% and 95% of its equilibrium F_{ST} , respectively.

The effective size depends on demographic parameters and *F*-statistics, and its asymptotic value is realized once all the *F*-statistics reach their equi-

brum values. Compared with F_{ST} , F_{IS} reaches its equilibrium value much faster and is little affected by factors other than nonrandom union of gametes within subpopulations (Wang, 1997a). Therefore, the number of generations required to reach the asymptotic effective size is essentially determined by F_{ST} . If there are more levels of subdivision in a hierarchy, it is the effective differentiation at the highest level of subdivision ($2N_{eL}F_{ST}$ for the case described in section 5.2) that determines the time. At initial generations, the variance effective size is generally larger than the inbreeding effective size. The former decreases steadily, and the latter generally increases with possible fluctuations over generations, eventually converging to the same asymptotic value (Wang, 1997a,b).

With subdivision, mating is more likely between individuals within subpopulations than among subpopulations. This is somewhat like mating between individuals of higher than average co-ancestry in a single population. Compared with completely random union of gametes, subdivision results in a higher initial rate of inbreeding, which decreases over generations. Therefore, the fixation probability of favourable genes and the rate of shifts between adaptive peaks (Barton & Rouhani, 1991) depend on population structure differently from neutral variability. For species with less frequent migration among subpopulations, a large number of generations is necessary to reach its asymptotic effective size. A highly differentiated population may never attain its asymptotic effective size, because it is highly unlikely that both its demographic and geographical properties remain constant for many generations.

Acknowledgements

We are grateful to N. H. Barton, W. G. Hill, L. Nunney and M. C. Whitlock for helpful discussions on the topic and useful comments on the manuscript. This work was supported by a Biotechnology and Biological Sciences Research Council grant 15/G09540 (J.W.) and by grant PB96-0304 (DGES, MEC from Spain) (A.C.).

References

- AVISE, J. C. 1991. Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Ann. Rev. Genet.*, **25**, 45–69.
- BARTON, N. H. AND ROUHANI, S. 1991. The probability of fixation of a new karyotype in a continuous population. *Evolution*, **45**, 499–517.
- BIRKY, A. J., JR, FUERST, P. AND MARUYAMA, T. 1989.

- Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approaches to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, **121**, 613–627.
- CABALLERO, A. 1994. Developments in the prediction of effective population size. *Heredity*, **73**, 657–679.
- CHESSER, R. K. AND BAKER, R. J. 1996. Effective sizes and dynamics of uniparentally and diparentally inherited genes. *Genetics*, **144**, 1225–1235.
- CHESSER, R. K., RHODES, O. E., JR, SUGG, D. W. AND SCHNABEL, A. F. 1993. Effective sizes for subdivided populations. *Genetics*, **135**, 1221–1232.
- CHOY, S. C. AND WEIR, B. S. 1978. Exact inbreeding coefficients in populations with overlapping generations. *Genetics*, **89**, 591–614.
- CROW, J. F. AND KIMURA, M. 1970. *Introduction to Population Genetics Theory*. Harper & Row, New York.
- EWENS, W. J. 1982. On the concept of the effective population size. *Theor. Pop. Biol.*, **21**, 373–378.
- FALCONER, D. S. AND MACKAY, T. F. C. 1996. *Introduction to Quantitative Genetics*, 4th edn. Longman, New York.
- FISHER, R. A. 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- FRANKHAM, R. 1995. Effective population size/adult population size in wildlife: a review. *Genet. Res.*, **66**, 95–107.
- GILPIN, M. 1991. The genetic effective size of a metapopulation. *Biol. J. Linn. Soc.*, **42**, 165–175.
- GOWE, R. S., ROBERTSON, A. AND LATTER, B. D. H. 1959. Environment and poultry breeding problems. 5. The design of poultry control strains. *Poultry Sci.*, **38**, 462–471.
- HARTL, D. L. AND CLARK, A. G. 1997. *Principles of Population Genetics*, 3rd edn. Sinauer, Sunderland, MA.
- HEDRICK, P. W. AND GILPIN, M. E. 1997. Genetic effective size of a metapopulation. In: Hanski, I. and Gilpin, M. (eds) *Metapopulation Dynamics: Ecology, Genetics, and Evolution*, pp. 166–182. Academic Press, New York.
- HILL, W. G. 1985. Effects of population size on response to short and long term selection. *J. Anim. Breed. Genet.*, **102**, 161–173.
- KIMURA, M. 1953. 'Stepping-stone' model of population. *Ann. Rep. Natl. Inst. Genet. Japan*, **3**, 62–63.
- KIMURA, M. AND CROW, J. F. 1963a. The measurement of effective population number. *Evolution*, **17**, 279–288.
- KIMURA, M. AND CROW, J. F. 1963b. On the maximum avoidance of inbreeding. *Genet. Res.*, **4**, 399–415.
- LANDE, R. 1988. Genetics and demography in biological conservation. *Science*, **241**, 1455–1460.
- LANDE, R. AND BARROWCLOUGH, G. F. 1987. Effective population size, genetic variation, and their use in population management. In: Soulé, M. E. (ed.) *Viable Populations for Conservation*, pp. 87–123. Cambridge University Press, Cambridge.
- LYNCH, M., CONERY, J. AND BURGER, R. 1995. Mutation accumulation and extinction of small populations. *Am. Nat.*, **146**, 489–518.
- MARUYAMA, T. 1970. On the rate of decrease of heterozygosity in circular stepping stone models of populations.

- Theor. Pop. Biol.*, **1**, 101–119.
- MARUYAMA, T. 1971. The rate of decrease of heterozygosity in a population occupying a circular or a linear habitat. *Genetics*, **67**, 437–454.
- MARUYAMA, T. 1972. Rate of decrease of genetic variability in a two-dimensional continuous population of finite size. *Genetics*, **70**, 639–651.
- MARUYAMA, T. AND KIMURA, M. 1980. Genetic variability and effective population size when local extinction and recolonisation of subpopulations are frequent. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 6710–6714.
- MILLS, L. S. AND ALLENDORF, F. W. 1996. The one-migrant-per-generation rule in conservation and management. *Conserv. Biol.*, **10**, 1509–1518.
- NAGYLAKI, T. 1995. The inbreeding effective population number in dioecious populations. *Genetics*, **139**, 473–485.
- NEI, M. AND TAKAHATA, N. 1993. Effective population size, genetic diversity and coalescence time in subdivided populations. *J. Mol. Evol.*, **37**, 240–244.
- NUNNEY, L. 1996. The influence of variation in female fecundity on effective population size. *Biol. J. Linn. Soc.*, **59**, 411–425.
- NUNNEY, L. 1998. The effective size of a hierarchically-structured population. *Evolution* (in press).
- ROBERTSON, A. 1960. A theory of limits in artificial selection. *Proc. R. Soc. B*, **153**, 234–249.
- SACCHERI, I., KUUSSAARI, M., KANKARE, M., VIKMAN, P., FORTELIUS, W. AND HANSKI, I. 1998. Inbreeding and extinction in a butterfly metapopulation. *Nature*, **392**, 491–494.
- SANTIAGO, E. AND CABALLERO, A. 1995. Effective size of populations under selection. *Genetics*, **139**, 1013–1030.
- SLATKIN, M. 1985. Gene flow in natural populations. *Ann. Rev. Ecol. Syst.*, **16**, 393–430.
- SLATKIN, M. 1991. Inbreeding coefficient and coalescence times. *Genet. Res.*, **58**, 167–175.
- SLATKIN, M. 1977. Gene flow and genetic drift in a species subject to frequent local extinctions. *Theor. Pop. Biol.*, **12**, 253–262.
- WANG, J. 1996. Inbreeding and variance effective sizes for nonrandom mating populations. *Evolution*, **50**, 1786–1794.
- WANG, J. 1997a. Effective size and *F*-statistics of subdivided populations. I. Monoecious species with partial selfing. *Genetics*, **146**, 1453–1463.
- WANG, J. 1997b. Effective size and *F*-statistics of subdivided populations. II. Dioecious species. *Genetics*, **146**, 1465–1474.
- WANG, J. 1997c. More efficient breeding systems for controlling inbreeding and effective size in animal populations. *Heredity*, **79**, 591–599.
- WANG, J. 1998. Effective size and *F*-statistics of subdivided populations for sex-linked loci. *Theor. Pop. Biol.* (in press).
- WHITLOCK, M. 1992. Temporal fluctuations in demographic parameters and the genetic variance among populations. *Evolution*, **46**, 608–615.
- WHITLOCK, M. AND BARTON, N. 1997. The effective size of a subdivided population. *Genetics*, **146**, 427–441.
- WHITLOCK, M. AND McCAULEY, D. E. 1990. Some population genetic consequences of colony formation and extinction: genetic correlations within founding groups. *Evolution*, **44**, 1717–1724.
- WRIGHT, S. 1946. Isolation by distance under diverse systems of mating. *Genetics*, **31**, 39–59.
- WRIGHT, S. 1931. Evolution in Mendelian populations. *Genetics*, **16**, 97–159.
- WRIGHT, S. 1943. Isolation by distance. *Genetics*, **28**, 114–138.
- WRIGHT, S. 1951. The genetical structure of populations. *Ann. Eugen.*, **15**, 323–354.