few markers have patterns that reflect the populations' history, if indeed there is any clear meaning to this concept.

7.3. THE THEORY OF SELECTION WITH SPATIAL STRUCTURE

7.3.i. Introduction

So far in this chapter, we have assumed neutral variants. We previously assumed panmixis when examining the joint effects of selection and genetic drift, for the purpose of interpreting data on molecular evolution and variation (**Chapters 5** and **6**). We are now ready to put together models of selection, drift, and population structure, and to ask questions about how population structure affects variants under selection in several biologically important contexts. For instance, how are the properties of populations (such as mean fitness and the level of inbreeding depression, discussed in **Section 4.4**) affected by population structure? Can differences in the mean fitnesses of different local populations be an effective evolutionary force, driving a process of inter-deme selection?

In contrast to **Sections 2.2.iii** and **4.1**, we will mostly assume here that selection pressures are constant across the metapopulation, because we are interested in how to deal with the consequences of departures from panmixis, not in how selection leads to local adaptation (this topic is considered further in **Section 8.3.iv**). The problem of selection in spatially structured populations is comprehensively reviewed by Rousset (2004).

7.3.ii. Fixation probabilities in structured populations

Maruyama (1970c, 1974) and Nagylaki (1982) showed that, with semidominant or haploid selection, the fixation probability of a new mutation in a structured population consisting of a set of Wright–Fisher populations connected by conservative migration (**Section 7.2.ii.d**) is invariant with respect to migration patterns and rates. In this case, the product of the selection coefficient and the migration effective population size, N_{eS}, determines the fixation probability in the same way as the selection coefficient and effective population size in a single, panmictic population (**Equations B6.4.2a** and **B6.4.2b** in **Section 6.2.iii.a**). In this case, N_{eS} is the same as the effective size for a single population with the same number of breeding individuals as the metapopulation.

How applicable is this result when the relevant assumptions do not apply? Recent work suggests that an approximate diffusion equation can be derived for fairly general selection and migration models, using the fast time scale

approximation for large deme numbers discussed in **Section 7.2.v** (see **Box 7.7***). This has been rigorously established only for semidominant and haploid selection for the island model (Cherry and Wakeley 2003), and for a haploid extinction/colonization model with a propagule number of 1 (Cherry 2003b), but simulations suggest that it applies more generally.

Box 7. 7* DRIFT AND SELECTION WITH A LARGE DEME NUMBER

The approximation for the effects of drift and selection in a metapopulation with a large deme number depends on the fact that, under the conditions discussed in **Section 7.2.v**, the time scale of drift within a deme is negligible compared with that for drift in the metapopulation as a whole. The latter is given by the reciprocal of the coalescence probability, P_{Cc}, for a pair of alleles sampled from two different demes, i.e., $2N_{eM}$, where N_{eM} is the effective size of the metapopulation (**Box 7.6***). We also assume that the selection coefficient s is of the same order as P_{Cc}, so that the time scale of the change under selection is very long.

The effects of drift and selection on the state of the metapopulation with respect to the frequencies of two alternative variants at a site, A_1 and A_2, are represented by the expected change in $\bar{q}$, the mean over demes of the frequency of A_2, and the variance of the change in $\bar{q}$ (Wakeley 2003; Roze and Rousset 2003). (As described in **Section 7.2.v.c**, any convenient method of weighting the contributions of individual demes to the mean over demes can be used in this context.) An argument similar to that used in **Section 5.1.i.c** implies that the variance of the change in $\bar{q}$ is equal to $\bar{q}(1-\bar{q})/2N_{eM}$ (Whitlock and Barton 1997). We can thus make use of the standard diffusion equations for a panmictic population (**Sections 5.3.ii.c** and **6.2.iii.a**), with $\bar{q}$ replacing the single population frequency q.

We can now write the following equation:

$$\sum_i \omega_i q_i (1 - q_i) = \bar{q} - \left(\bar{q}^2 + V_q\right) = \bar{q}(1-\bar{q})\left(1 - F_{ST}\right) \qquad \text{(B7.7.1)}$$

where q_i is the variant frequency for deme i and and ω_i is the weight given to deme i when taking the mean of the q_i (where $\Sigma_i \omega_i = 1$). The mean of the selection terms for each deme with semidominant selection, $sq_i(1 - q_i)/2$, is thus equal to $s\bar{q}(1-\bar{q})(1-F_{ST})/2$. The assumption of weak

selection implies that F_{ST} is given by the same equation as in the neutral case (**Equation 7.16** of **Section 7.2.v.d**), to an accuracy of the order of s^2.

In order to remove any contribution to the expected change in $\bar{q}$ from migration or extinction, we should use the weighting $\omega_i = \nu_i$, where ν_i is the probability that an ancestral allele is present in the ith deme (**Section 7.2.v.b**). Using the value of F_{ST} with this weighting of deme contributions, we then have $N_{eM} = N_{eS}/(1 - F_{ST})$ under the large deme number approximation of **Section 7.2.v.d**.

Only the ratio of the selection and variance terms appears in the diffusion equation formula for fixation probability (see **Box 6.4*** of **Section 6.2.iii.a**), so that the common factor $\bar{q}(1 - \bar{q})(1 - F_{ST})$ cancels, implying that the fixation probability is determined by $N_{eS}s$ and the initial value of $\bar{q}$. It is easily shown that this leads to the same formula for the fixation probability as **Equation B6.4.2b**. If deme sizes do not differ much, the exact choice of weight will have only a second-order effect on the fixation probability.

This result means that the effective population size that determines the mean within-deme nucleotide site diversity under the infinite sites model can be used to determine fixation probabilities in general migration or extinction/colonization models with semidominant or haploid selection, at least as a first-order approximation. This is very useful, because it implies that spatial structure may not have much effect on the fixation probabilities of weakly selected mutations, so that the models used in **Chapter 6** to interpret data on molecular evolution apply even to highly subdivided populations. Tests of predictions of the effects of differences in species' effective population sizes on rates of sequence evolution for species (e.g., Woolfit and Bromham 2005) should therefore use estimates of N_e based on mean within-population diversities.

With dominance, however, population structure can cause important departures from the panmictic results for selection (Slatkin 1981; Barton 1993; Whitlock 2002, 2003; Cherry 2003a; Roze and Rousset 2003). In particular, the increased local frequencies of homozygotes caused by drift within demes enhance the effectiveness of selection, similar to the effect of inbreeding within a population (**Section 3.1.v.c**). Thus, even with conservative migration, fixation probabilities are reduced for recessive or partly recessive deleterious mutations, and increased for recessive or partly recessive advantageous mutations, relative to the value for a panmictic population with an effective size of N_{eT}, as

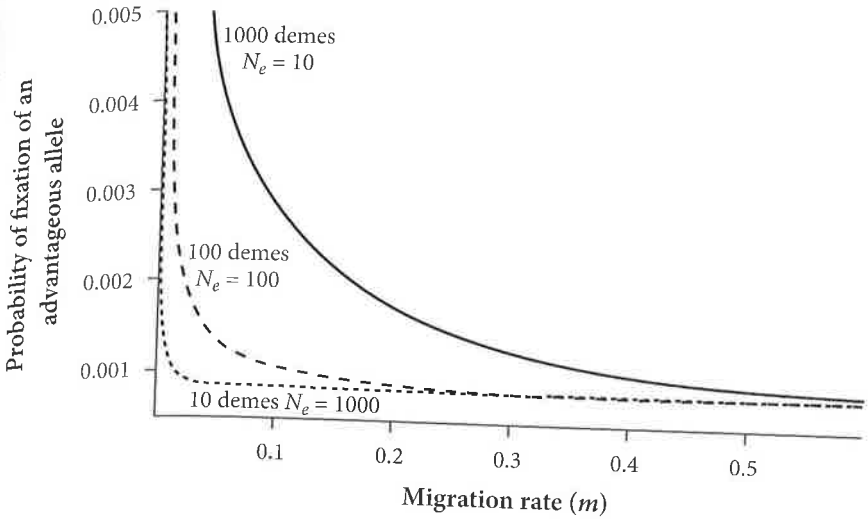

FIGURE 7.14 Effect of population structure on the probability of fixation of a recessive advantageous mutation (dominance coefficient $h = 0$) that arises as a single copy in a population with N adult individuals in each of the stated numbers of demes. In the example, the selection coefficient (s) was set equal to 0.01. [Adapted from Figure 4 of Roze and Rousset (2003).]

shown in **Figure 7.14**. The reverse is true for dominant or partially dominant mutations. The overall effect of population subdivision on the rate of DNA sequence evolution thus depends on both the level of dominance of new mutations and the extent to which advantageous or deleterious mutations contribute to sequence evolution. Few data are available to determine whether such effects occur in nature.

7.3.iii. Variant frequency distributions with selection and spatial structure

We next examine the distribution of variant frequencies at selected sites in a subdivided population, which is important for the tests of neutrality described in **Sections 6.4.iii** and **6.4.iv**.

7.3.iii.a. *The infinite-deme island model*

The problem of determining the distribution of variant frequencies was solved by Wright (1940a) for an island model with an infinitely large number of demes with equal effective population sizes. **Box 7.8*** derives a general formula for the probability density of allele frequency q in a given deme, with different demes being treated as independent draws from this distribution. It is assumed that selection is strong enough to hold allele frequencies close to their equilibrium values under mutation–selection balance or balancing selection.

Box 7.8* APPROXIMATIONS FOR THE DISTRIBUTION OF ALLELE FREQUENCIES UNDER SELECTION IN THE ISLAND MODEL

B7.8.i. General treatment

We will show that the probability density of allele frequency q for a single deme, $\phi(q)$, can be approximated by a beta distribution (**Section 5.3.iii.c**) when the expected change in q, $M_{\delta q}$, is a linear function of the departure of q from an equilibrium value, q^*. This is the case if q is always close to its equilibrium value, so that we can use a Taylor's series (**Appendix A1.ii.a**) to approximate $M_{\delta q}$ by $\kappa(q - q^*)$, where κ is the derivative of $M_{\delta q}$ with respect to q at q^* (**Section 2.1.ii.c**). If the equilibrium is stable, $\kappa < 0$. We can then write:

$$M_{\delta q} \approx \kappa(q - q^*) = \kappa(1 - q^*)q - \kappa q^*(1 - q)$$

This can be substituted into **Equation B5.9.1** in **Box 5.9** of **Section 5.3.iii.c** to give the exponential term in the expression for $\phi(q)$ as:

$$\exp -4N_e\kappa[(1 - q^*)\ln(1 - q) + q^* \ln q] \qquad \text{(B7.8.1)}$$

Evaluating this and substituting the result into **Equation B5.9.1**, we get:

$$\phi(q) \approx Cp^{-(4N_e\kappa p^* + 1)}q^{-(4N_e\kappa q^* + 1)} \qquad \text{(B7.8.2)}$$

This has the same form as **Equation 5.19**, and is thus also a beta distribution, with parameters $\alpha = -4N_e\kappa p^*$ and $\beta = -4N_e\kappa q^*$.

From the properties of this distribution, the mean of q is equal to $q^* = \beta/(\alpha + \beta)$, and the variance is $V_q = p^*q^*/(1 + \alpha + \beta) = p^*q^*/(1 - 4N_e\kappa)$.

B7.8.ii. Large deme number approximation

Under the large deme number approximation used in **Section 7.2.v**, the above result can be applied to the mean of q across demes, $\bar{q}$, using the argument in **Box 7.7*** of **Section 7.3.ii**, so that the distribution of $\bar{q}$ also approaches a beta distribution with mean q^*. With an indefinitely large number of demes, $\bar{q}$ can be approximated by q^*. This implies that, for a deme with frequency q, we can write the expected change in frequency due to migration as:

$$(1 - m)q + mq^* - q = -m(q - q^*) = m(1 - q)q^* - mq(1 - q^*)$$

This has the same form as the linearized selection term, so that κ can be replaced by $\kappa - m$ in **Equations B7.8.1** and **B7.8.2**, thus giving the final expression for the distribution of q among demes as:

$$\phi(q) \approx Cp^{4N_e(m-\kappa)(1-q^*)-1}q^{4N_e(m-\kappa)q^*-1} \qquad \text{(B7.8.3)}$$

Use of **Equation (B7.8.3)** and the formula for the variance of a beta distribution gives:

$$F_{ST} \approx \frac{1}{1 + 4N_e(m - \kappa)} \qquad \text{(B7.8.4)}$$

Since κ is negative, a comparison with the neutral formula (**Equation 7.6** of **Section 7.2.ii.b**) shows that F_{ST} is reduced below the neutral value; this is because selection tends to restore allele frequencies to their equilibrium, in opposition to the effects of drift.

B7.8.iii. Mutation, selection, and migration

Here, we assume that mutations occur from A_1 to A_2 at rate u, with selective disadvantage hs to A_2 in heterozygotes. In this case, use of **Equation 4.2** of **Section 4.2.ii.a** gives the approximate change in frequency of A_2 as $-(qhs - u)$. The equilibrium value of q is $q^* = u/hs$ (**Equation 4.3**), so that $(qhs - u) = hs(q - q^*)$, giving $\kappa = -hs$.

B7.8.iv. Heterozygote advantage with drift

Using the notation of **Section 2.1.ii.c**, the equilibrium frequency q^* is $s/(s + t)$. From **Equation B2.4.3** of **Section 2.1.ii.c**, we have $\kappa \approx -st/(s + t)$, assuming that selection is weak.

The box also gives explicit formulae for F_{ST} for the cases of mutation–selection balance and heterozygote advantage. In these cases, selection always reduces F_{ST} below the neutral value, a result that holds more widely than for this particular model of population structure (e.g., Malécot 1969, Section 3.3). Differences in F_{ST} among loci may therefore provide evidence for the action of selection on some of them, and methods for testing whether there is more variation among loci in F_{ST} than expected under neutrality have been devised (Lewontin and Krakauer 1973; Beaumont and Nichols 1996; Beaumont and

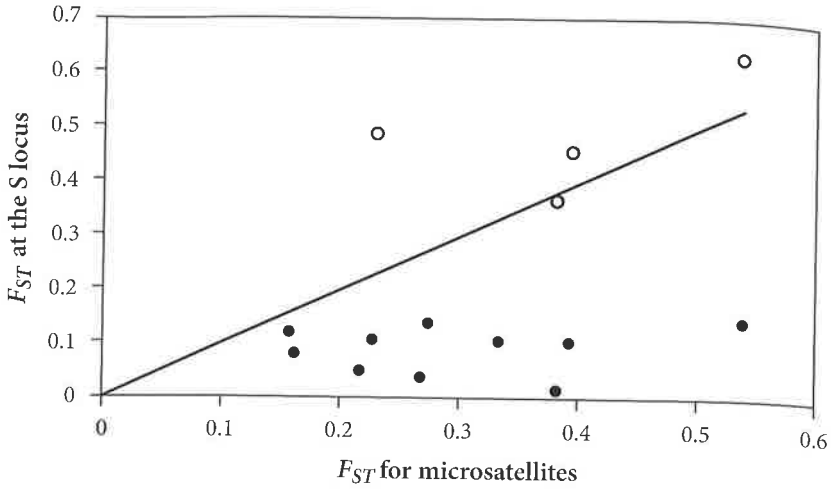

FIGURE 7.15 Comparison between F_{ST} estimated from microsatellite alleles in populations of the plant *Brassica insularis*, and from alleles at the self-incompatibility locus, which is under strong balancing selection (see text). Open circles are F_{ST} values estimated from microsatellite markers, whereas black ones are estimates from the self-incompatibility alleles. [Adapted from Figure 5 of Glémin et al. (2005).]

Balding 2004). In **Section 8.3.iv** we discuss how F_{ST} for presumptively neutral sites can be used to detect the effects of selection at linked sites.

There is evidence that some polymorphisms maintained by balancing selection have lower F_{ST} values than other loci. A study of self-incompatibility alleles (see **Section 2.2.i.b**) in a wild *Brassica* species found that alleles are only mildly differentiated between five populations in Corsica (Glémin et al. 2005). These populations are strongly ecologically isolated, and **Figure 7.15** shows that pairwise F_{ST} values for microsatellite markers are much higher than for the SI locus (mean pairwise $F_{ST} = 0.28$).

CASE STUDY
Lethal mutations in *Drosophila pseudoobscura*

A classic example of how this theory can be applied to real populations, to ask interesting questions about selection and migration, is the extraordinarily thorough investigation of the spatial distribution of lethal mutations on chromosome 3 of *Drosophila pseudoobscura* in southern California (Dobzhansky and Wright 1941; Wright et al. 1942). The data were collected using crosses

with balancer chromosomes, as described in **Section 1.1.iii.b**. Two different localities separated by 200 miles, Death Valley and Mt. San Jacinto, were sampled, with sites separated by several miles within each of the two localities, and with several different nearby collecting stations at each site.

The results were summarized in terms of the frequencies of lethal-bearing chromosomes and the frequency of allelism between different lethal chromosomes found in a sample, i.e., the frequency with which a cross between carriers of two different lethal chromosomes yields a lethal combination (**Section 1.1.iii.b**). Chromosome 3 is homologous to the right arm of chromosome 2 of *D. melanogaster*. Accordingly, the mean frequency of lethal chromosomes was about half that for *D. melanogaster* second chromosomes (a mean of about 15.5% over the two localities). For lethal chromosomes collected at the same station, the mean allelism rate (with standard error) was 2.6% ± 0.5%, versus 0.41% ± 0.08% for chromosomes from different sites. Different stations from the same site had intermediate allelism rates of about 0.88% ± 0.20%. These results show that individual lethal mutations differ in their frequencies between the different locations, despite the fact that the flies can migrate over several kilometers in their lifetime (Dobzhansky and Wright 1947; Coyne et al. 1982).

As shown in **Box 7.9**, several parameters of interest can be estimated by combining the data on natural populations with an estimate of the rate at which lethal mutations arise on chromosome 3 (0.31% ± 0.04% per generation; Dobzhansky and Wright 1941). First, from **Equation B7.9.2b**, F_{ST} values among stations at San Jacinto and Death Valley are 2.39×10^{-3} and 5.01×10^{-3}, respectively. **Box 7.9** also outlines the reasons for inferring that lethals are removed mainly through their heterozygous effects on fitness (h): the mean h value was estimated to be 0.019 over the two localities. This inference was confirmed by later experiments that showed heterozygous fitness effects of lethals in *Drosophila* (**Section 4.2.ii.a**).

Second, the data allow N_e for a deme to be estimated, despite the possibility of migration. Using **Equation B7.8.4** and setting $\kappa = -h$, the estimated values of $N_e(m + h)$ for San Jacinto and Death Valley are 104 and 50, respectively. This is much higher than the $N_e m$ value of about 2 estimated from DNA sequence variation at the *Adh* locus for North American populations of this species (Schaeffer and Miller 1992). The term involving h is therefore the major contributor to $N_e(m + h)$. After subtracting the $N_e m$ value (estimated from the *Adh* data), and using the h estimate of 0.019 derived above, N_e can be estimated. The N_e values are 5358 for San Jacinto and 2400 for Death Valley. Further discussions of these results are given by Wright et al. (1942) and Wright (1978, pp. 157–170).

Box 7.9 BEHAVIOR OF LETHAL MUTATIONS IN SUBDIVIDED POPULATIONS

Let the frequency of lethal mutations for the *i*th gene on the chromosome be q_i. Using an approach similar to that of **Section 4.2.ii.c**, the expected number of lethal mutations per chromosome is $\Sigma_i q_i$, where the sum is taken over all genes capable of mutating to recessive lethal mutations. If the mean frequency of lethal mutations per gene is $\bar{q}$, and there are n genes capable of mutating to lethals per chromosomes, $\Sigma_i q_i = n\bar{q}$. If mutations in different genes are distributed independently of each other, the number of lethal mutations per chromosome follows a Poisson distribution with mean $n\bar{q}$, so that the observed frequency of lethal chromosomes is $P = 1 - \exp(-n\bar{q})$. This enables $n\bar{q}$ to be estimated from $-\ln(1 - P)$. For small P, $n\bar{q}$ is close to P.

The frequency with which two lethal chromosomes carry a mutation in gene i is equal to $p_i = (q_i/n\bar{q})^2$, because the frequency of lethals at locus i, relative to the overall frequency of lethals, is $q_i/n\bar{q}$. The frequency with which two lethal chromosomes carry a mutation in the same gene is thus $p_a = \Sigma_i p_i^2$. From the relation between p_i and q_i, we obtain:

$$p_a = \frac{\left(\bar{q}^2 + V_q\right)}{\left(n\bar{q}^2\right)} \qquad \text{(B7.9.1)}$$

where V_q is the variance in q among loci.

V_q has two components—the variance in lethal allele frequencies due to differences in selection coefficients and mutation rates among different loci, and the variance due to the effect of genetic drift discussed in **Box 7.8***. For chromosomes collected from remote localities, with allelism rate p_{ar}, only the first component needs to be considered, and we can write $V_q = V_{qd}$, where d indicates the variance among loci caused by the variance in deterministic effects. For chromosomes from the same collecting station, the allelism rate is p_{as} and $V_q = V_{qd} + V_{qc}$, where c indicates the stochastic component. Substituting these into the versions of **Equation B7.9.1** for remote populations and the same station, we obtain:

$$V_{qc} = \left(p_{as} - p_{ar}\right)n\bar{q}^2 \qquad \text{(B7.9.2a)}$$

For a structured population, V_{qc} can be equated to the variance in allele frequencies among demes due to restricted migration and finite population size. Since $\bar{q}$ is expected to be $<< 1$, this gives:

$$F_{ST} = (p_{as} - p_{ar})n\bar{q} \qquad \text{(B7.9.2b)}$$

As shown above, $n\bar{q}$ can be estimated from the frequency of lethal chromosomes, and the other two variables are obtainable from the relevant allelism frequencies, allowing F_{ST} for lethal chromosomes to be estimated.

Knowing the rate of mutation to lethals allows further conclusions to be drawn. Let this rate be U_l per chromosome per generation; for chromosome 3 of *D. pseudoobscura*, this is 0.0031. For allelism between remote populations, **Equation B7.9.1** gives:

$$n = (1 + V_{qd}/\bar{q}^2)/p_{ar} \qquad \text{(B7.9.3)}$$

A lower bound estimate of n is thus provided by $1/p_{ar}$. This is equal to $1/0.0035 = 285$ in the case of the California populations of *D. pseudoobscura* discussed in the main text. If lethals are completely recessive in their fitness effects and mating is random, **Equation 4.4** of **Section 4.2.ii.a** implies that $\bar{q} = \sqrt{U_l/n}$, i.e., $n\bar{q} = \sqrt{nU_l}$. For the *D. pseudoobscura* data, the predicted value of $n\bar{q}$ is thus greater than 0.939, more than five times the observed value. This implies that lethals are either eliminated mainly as a result of their heterozygous fitness effects (h, since $s = 1$), or that populations are partially inbred. The latter is very unlikely, however, given the breeding biology of this species. We can therefore infer that this discrepancy is due to selection against the heterozygous effects of lethals. On this basis, the expected value of $n\bar{q} = U_l/h$, so that $h = U_l/n\bar{q}$. This gives an average h for the two localities of about 0.019.

7.3.iii.b. *Quantitative traits and Q_{ST} values*

Variability in quantitative traits can also be partitioned into total and within-population components of genetic variance. The difference between these components, relative to the total genetic variance, provides a way of partitioning the variance, similar to F_{ST} (Prout and Barker 1993), which is often called Q_{ST} (Lynch et al. 1999). Under neutrality, F_{ST} and Q_{ST} should be similar, but if stabilizing selection is acting on the quantitative trait in the same way in each deme, $Q_{ST} < F_{ST}$. If different trait values are favored in different locations, we expect the reverse. Strictly speaking, this result assumes additive inheritance of the trait (**Section 3.3.ii.b**), and dominance effects are likely to lead to $Q_{ST} < F_{ST}$ for a subdivided populations, but the bias is small if many loci affect the trait (Goudet and Martin 2007; Santure and Wang 2009). Methods have been developed to test the statistical significance of differences between Q_{ST} and F_{ST} (Prout and Barker 1993; Lynch et al., 1999).

The amount of differentiation for molecular markers, such as SNPs or microsatellites, is commonly much lower than that estimated for phenotypic characters, suggesting that phenotypic differences between populations have often evolved by adaptation to local environments (**Section 4.1**), despite the flow of neutral or weakly selected variants between the populations. The first such test was used to detect selection in natural populations of *Drosophila buzzatii* (Prout and Barker 1993), and some other animal studies are reviewed by Lynch et al. (1999) and McKay and Latta (2002). Most uses of this test so far are in plants, where there is already much evidence for local adaptation from classic approaches, such as the reciprocal transplantation experiments reviewed in Linhart and Grant (1996) (see **Figure 7.16**). However, Whitlock (2008) has pointed out that the statistical tests used in these studies overlook the fact that F_{ST} for neutral loci has a probability distribution generated by the coalescent process in a structured population, and describes a method for overcoming this problem.

7.3.iii.c. *The sampling properties of alleles under the large deme number approximation*

The approach outlined in **Section 7.3.iii.a** above ignored the fact that the mean allele frequency for the metapopulation is not fixed, but itself evolves under

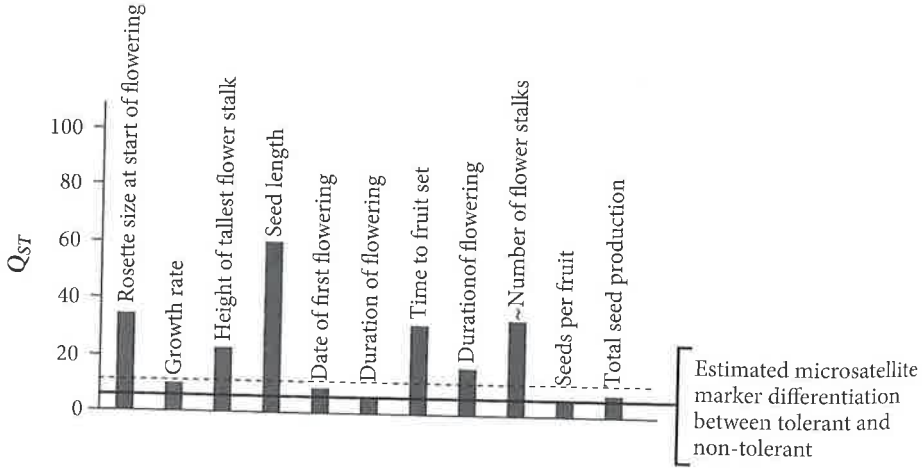

FIGURE 7.16 Comparison between F_{ST} estimated from microsatellite alleles in populations of the plant *Thlaspi caerulescens*, and Q_{ST} values for a number of phenotypic characters. The figure compares metal-tolerant and non-tolerant plants, showing that there is differentiation for several phenotypic characters, but not for the microsatellite markers (as can be seen from their mean and upper 95% confidence interval, indicated by the black line and the dashed line, respectively; the lower confidence interval coincides with the x axis). [Adapted from Figure 3 of Jimenez-Ambriz et al. (2007).]

selection, mutation, and drift. This complicates the analysis of the distribution of variant frequencies. In general, it is necessary to use simulations to determine the probability distribution of variant frequencies for a sample of alleles from a set of populations. A powerful approximate method for doing this for a pair of selected variants at a single site has recently been devised (Fearnhead 2006).

If the large deme number approximation is valid, however, approximate analytical results can be derived for the island model (Wakeley 2003). These results are particularly simple when each allele in the sample comes from a different deme. The large deme number approximation then implies that the distribution of the mean frequency of A_2 for the metapopulation, $\phi(\overline{q})$, is approximately the same as that for a panmictic population, with the effective population size for the metapopulation, N_{eM} (**Section 7.2.ii.b**), replacing N_e for a single population (Wakeley 2003). If k alleles are sampled, each from a different deme, these can be considered as k independent draws from a population with frequency $\overline{q}$. The distribution of variant frequencies discussed in **Section 6.3.ii.d** can then be used, with N_{eM} instead of N_e for a single population. These results suggest that tests for selection using frequencies of variants should use samples of this kind when there is evidence for significant effects of population structure.

7.3.iv. Effects of population structure on mean fitness and selection among demes

7.3.iv.a. *The effects of drift on population mean fitness*

The effect on population mean fitness of the random differentiation of allele frequencies caused by population structure is similar to the effects of inbreeding modeled in **Section 4.4.ii**. To the level of accuracy of the approximations used in **Box 7.8*** of **Section 7.3.iii.a**. we can simply replace f in **Equations 4.14–4.16** by F_{ST} to obtain the mean over all demes of the mean fitness of a deme. The reduction in this mean fitness caused by a set of loci with deleterious alleles whose fitness effects combine multiplicatively, relative to that for a panmictic population, is thus BF_{ST}, where B is given by **Equation 4.16.b**. This is true for any type of population structure, but simple expressions incorporating the effect of selection on F_{ST} have been derived only for the island model (**Box 7.8***). If selection is weak relative to migration, the neutral value of F_{ST} is a good approximation, but this will considerably overestimate F_{ST} for strongly deleterious variants. If there is a large reservoir of very weakly selected deleterious variants in natural populations, as suggested by the analyses of data on amino acid polymorphisms described in **Section 6.4.iv**, the neutral value may be a good approximation.

Populations formed by intercrossing local populations should thus experience an increase in mean fitness of approximately BF_{ST}, and this prediction is

supported by more rigorous theoretical analyses (Whitlock 2002; Roze and Rousset 2003, 2004; Glémin 2005). Interpopulation hybrids do indeed often exhibit heterosis, which can be substantial (Escobar et al. 2008), although it is unclear whether it can be explained by this simple mechanism alone. One of the first such studies demonstrated heterosis for survival and fecundity in crosses between different *D. pseudoobscura* populations (Vetukhiv 1953, 1956). Heterosis was also found in a water flea (*Daphnia magna*) metapopulation living in small pools of water in Finland, in an experiment in which pools were emptied and then seeded with equal numbers of animals previously collected from the pool, and animals collected from a different pool. In most pools, hybrid genotypes (detected using markers) rapidly appeared and increased in frequency in successive samples, while the previously resident genotypes became rare (4 pools) or disappeared (11 pools); in only 2 pools were the initial genotypes still common at the end of the experiment (Ebert et al. 2002).

Interpopulation heterosis has also been observed in many experiments with plants (Van Treuren et al. 1993; Ouborg and Van Treuren 1994; Willi and Fischer 2005). In the plant *Silene latifolia*, for example, seed germination rates were significantly higher in interpopulation crosses than those generated on the same maternal plants by within-population outcrossing (Richards 2000).

7.3.iv.b. *Inter-deme selection*

Many biologists who have not been trained in evolutionary biology, and even some who have, assume that a trait will evolve if it enhances the fitness of the population or the species as a whole, without asking whether it can spread through a population by the type of selection that we have considered up to now (including kin selection, as discussed in **Section 3.1.v.d**). A famous example is the claim by V. C. Wynne-Edwards (1962) that animals have evolved mechanisms to reduce their reproductive rate, because this prevents them from exhausting their resources, thereby causing local populations to become extinct. Appeals to group- or species-level selection have also been made in relation to the evolution of sexual reproduction (**Section 10.2**) and the mating systems of sexual populations (**Section 9.2**).

If such arguments are to be made, we need to consider carefully how selection at higher levels than that of the individual can be effective, especially when it opposes selection involving fitness differences between individuals (if there is no such opposition, group selection might accelerate the rate of evolution, but adds nothing essentially new). For example, altruistic behavior increases the fitness of other members of the population, at the expense of the fitness of the individuals who exhibit the behavior (**Section 3.1.v.d**).

We therefore focus on how selection might act at the level of demes in a subdivided population of the type we have been considering in this chapter. There are two ways in which such *inter-deme selection* can act. First, a deme that