

The genetic structure of *Beta vulgaris* ssp. *maritima* (sea beet) populations. II. Differences in gene flow estimated from RFLP and isozyme loci are habitat-specific

A. F. RAYBOULD*, R. J. MOGG & C. J. GLIDDON†

Institute of Terrestrial Ecology, Furzebrook Research Station, Wareham, Dorset BH20 5AS, and †School of Biological Sciences, University of Wales, Bangor, Gwynedd LL57 2UW, U.K.

In a previous paper, isolation by distance at RFLP loci was demonstrated in 10 populations of sea beet growing on the Dorset (U.K.) coast. In the same populations there was no evidence of isolation by distance at isozyme loci. Here, the data are reanalysed to reduce biases resulting from nonrandom mating within the populations; the contrast between RFLPs and isozymes remains very clear. Separate analysis of cliff-top populations shows strong isolation by distance at both RFLP and isozyme loci, whereas in populations on sea walls around the perimeter of Poole Harbour there is isolation by distance at RFLP loci but not isozyme loci. Thus, whatever the cause of the difference between RFLP and isozyme loci, it occurs in the harbour populations only. There was no evidence that founder effects are a significant source of variation between populations on the perimeter of the harbour. This is very different from a previous finding of strong founder effects between populations on Furzey Island (in Poole Harbour).

Keywords: founder effects, gene flow, isolation by distance, isozymes, RFLPs, sea beet.

Introduction

Despite the widespread use of isozymes as molecular markers, their value in studies of population genetics and evolution has been questioned. Lewontin (1991) considered that isozymes have done little to resolve the fundamental question of how much genetic variation is adaptive. Nevertheless, our knowledge of genetic structure and gene flow in hundreds of species is based solely on isozyme variation, with the assumption that this variation is neutral. It is essential, therefore, that the potential for adaptive variation at isozyme loci should be considered when inferring genetic structure from isozyme variation only (e.g. Karl & Avise, 1992). For example, if balancing selection operates, gene flow between populations may be overestimated, whereas directional or disruptive selection can lead to underestimation of gene flow.

Koehn (1978) lists four kinds of data necessary to confirm the adaptive significance of an isozyme

polymorphism. They were summarized by Mitton (1989) as:

- 1 functional differences among genotypes;
- 2 physiological differences among genotypes;
- 3 physiological variation must be important in the ecological context of the population;
- 4 variation in fitness as a consequence of the variation in physiology.

Watt (1994) discusses studies of several animal species in which these criteria are met and, hence, appear to demonstrate selection at particular isozyme loci. However, there is little incentive to study enzyme function and physiology in species that do not show obvious clinal or habitat-associated variation. Therefore, although selection at individual isozyme loci has been demonstrated, its extent is not known.

Mitton (1989) regards no study of plants as fulfilling the criteria of Koehn. However, enzyme polymorphism has been shown to influence plant physiology (e.g. Rainey *et al.*, 1987), and there are many examples of isozyme allele frequency variation

*Correspondence. E-mail: a.raybould@ite.ac.uk

being correlated with habitat heterogeneity (e.g. Lönn, 1993; Nevo *et al.*, 1994; Prentice *et al.*, 1995). It is probable, therefore, that variation of at least a few isozyme loci in a few plant species is adaptive.

It is clear that it is difficult to prove that isozyme variation is neutral or under selection. However, in some circumstances it is possible to test whether isozymes are behaving as though they are neutral. One test of neutrality is whether variation obeys the assumptions of isolation by distance models (Wright, 1943; Slatkin, 1993). If isozymes behave as neutral markers, isolation by distance leads to the correlation of genetic and geographical distances between pairs of populations. Isolation by distance can be demonstrated independently of isozyme variation through studies of either pollen and seed dispersal or the patterns of variation in other markers. For example, Karl & Avise (1992) demonstrated isolation by distance (restricted larval dispersal) in the American oyster using single-copy nuclear (SCN) RFLPs. In the same populations, there was no relationship between geographical and genetic distance at several isozyme loci. From this Karl & Avise (1992) inferred balancing selection at these loci.

Recently, we demonstrated this effect in 10 populations of wild sea beet (*Beta vulgaris* ssp. *maritima*) growing on the Dorset (U.K.) coast (Raybould *et al.* 1996a). There was a significant negative regression in log-log plots of gene flow and geographical distance (Slatkin, 1993) when gene flow was estimated from six SCN RFLP loci, whereas there was no relationship between distance and gene flow estimated from seven isozyme loci.

In this paper, this situation is investigated in more detail. First, because the beet populations show significant F_{IS} , the data are reanalysed using a method for reducing the biases in gene flow estimates that arise through nonrandom mating. Secondly, this method is used to compare the effects of isolation by distance in two different habitats, namely cliff-tops on the Dorset coast and sea walls and shingle ridges in Poole Harbour, as patterns of selection and/or gene flow may be habitat-specific. Finally, founder effects are tested for in the harbour populations.

Materials and methods

Full details of the collection sites and isozyme and RFLP analysis are given in Raybould *et al.* (1996a, b). Briefly, 50 plants of sea beet were sampled from each of five cliff-top populations and five populations around the perimeter of Poole Harbour in Dorset. Four of the harbour populations grew on

sea walls and one on a sheltered shingle spit. Plants were scored for seven polymorphic isozyme loci and six polymorphic RFLP loci. Weir & Cockerham's (1984) unbiased estimators of F -statistics at these loci were calculated using F_{STAT} (Goudet, 1995). Variation at RFLP and isozyme loci was compared by log-log regressions of gene flow ($Nm = (1/4F_{ST}) - 0.25$) between all pairs of populations and the corresponding geographical distances between them (Slatkin, 1993).

In this paper, we adopt a similar approach, but the pairwise F_{ST} values, and hence Nm values, are derived in a different way. The mean F_{IS} of the sea beet populations analysed by Raybould *et al.* (1996a) was significantly higher than zero. For the purposes of this paper, this is assumed to result from substructure within the populations. To reduce the effect of this substructure, the sea beet plants were divided on the basis of their spatial distribution into smaller groups ('patches'). These patches were of approximately the same sizes as were determined to be randomly mating groups of sea beet on Furzey Island (Raybould *et al.*, 1996b). Genotypes at the two RFLP loci with null alleles were allocated on the basis of Hardy-Weinberg proportions, as locus-by-locus estimates of F_{IS} for the other four RFLP loci were not significant, although the over all RFLP loci F_{IS} was significant (Table 1). (In practice, correction for nonrandom mating does not change genotype numbers in patches this small.) F_{ST} estimates were then obtained between all pairs of patches in different populations. F_{ST} between two populations was taken to be the mean of the F_{ST} values between patches in the respective populations. Nm was then calculated from the mean F_{ST} . Deriving Nm from mean F_{ST} is more appropriate than calculating a mean Nm directly, because F_{ST} is always defined (negative F_{ST} gives an undefined Nm) and low F_{ST} values are not converted into very high Nm values that would have a disproportionate effect on the mean. Log-log regressions of these Nm values and distance were then calculated in the normal way and their significance tested by a Mantel test (Mantel, 1967; Raybould *et al.*, 1996a). To examine gene flow in different habitats, separate regressions were also calculated for the five cliff-top populations and the five harbour populations (fig. 1 in Raybould *et al.*, 1996a for locations).

As a test for founder effects and as a further (more powerful) test for isolation by distance, a partial Mantel test (Manly, 1991) was carried out on the patches in the harbour populations. This test uses all 351 pairwise F_{ST} estimates between the 27 patches. Partial regression coefficients are obtained

Table 1 Estimates of the mean and standard error of F_{IS} for Dorset sea beet populations ('10 populations') and for the same populations split into groups of the same size as those found to be randomly mating units on Furzey Island in 1993 (Raybould *et al.*, 1996b)

	10 Populations			53 Patches		
	Mean $F_{IS} \pm SE$	$P(F_{IS} = 0)$ (<i>t</i> -test)	$P(F_{IS} \neq 0)$ (randomization)	Mean $F_{IS} \pm SE$	$P(F_{IS} = 0)$ (<i>t</i> -test)	$P(F_{IS} \neq 0)$ (randomization)
All loci	0.138 ± 0.036	0.01–0.001	<0.001	0.085 ± 0.032	0.05–0.01	<0.001
All isozymes	0.183 ± 0.053	0.05–0.01	<0.001	0.117 ± 0.053	NS	<0.001
All RFLPs	0.079 ± 0.019	0.05–0.01	<0.001	0.049 ± 0.015	0.05–0.01	0.017

Means and standard errors from jackknifing over loci.

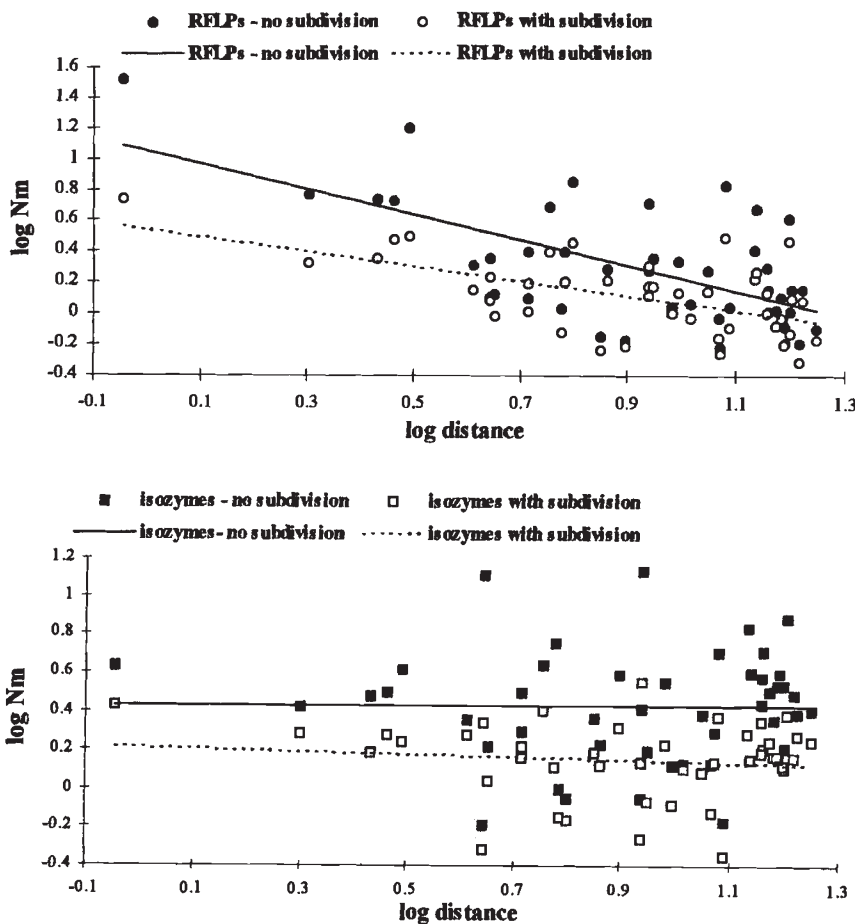


Fig. 1 Log–log regressions of gene flow against distance for RFLPs and isozymes where gene flow is estimated between populations of sea beet as a whole (no subdivision) and where gene flow is estimated from mean F_{ST} between patches within populations (subdivision).

for F_{ST} and geographical distance (removing the possible effect of correlation between patches in the same population) and F_{ST} and population membership (tests whether pairwise F_{ST} estimates within populations are different from those between populations, removing distance effects). The former coefficient tests for isolation by distance and the latter tests for founder effects.

Results

The effect of substructure

Subdivision of the 10 populations gave 53 patches (27 in the harbour and 26 on the cliffs). Table 1 shows how estimates of mean F_{IS} change with increased subdivision. All loci, all RFLP loci (excluding *RI* and *R13* as they have null alleles) and

Table 2 Summary of regression equation statistics for the slopes in Fig. 1

	Isozymes		RFLPs	
	No subdivision	With subdivision	No subdivision	With subdivision
<i>b</i>	0.003	-0.068	-0.822	-0.479
<i>A</i>	0.426	0.215	1.052	0.548
<i>r</i> ²	0.000	0.010	0.398	0.332
<i>P</i> (Mantel test)	0.491	0.185	0.0006	0.0007

b, regression slope of log *Nm* vs. log distance; *A*, y-intercept; *r*², proportion of variation in log *Nm* explained by log distance.

all isozyme loci estimates are still significant, despite the reduction in group size. Thus, the effect of substructure is not removed, but is reduced. It is not realistic to subdivide the populations further, as there are no spatially distinct groups within those already defined and the number of plants in each group would be very small. This would leave *F*_{IS} undefined at several loci in many groups.

Figure 1 shows the log-log regressions of *Nm* and distance calculated with and without subdivision of the populations. The details of the regression equations are given in Table 2. It is clear that for both isozymes and RFLPs the intercept of the regression slope is lower for the subdivided populations; also the slope of the RFLP regression slope is less negative with subdivision. This is to be expected, as reduction in *F*_{IS} reduces the overestimation of *Nm*. Crucially, however, the differences in the *r*² values for isozymes and RFLPs are little changed and the RFLP regression is still highly significant, while the isozyme regression is nonsignificant. Given that further subdivision of the populations is not possible, the regression slopes based on subdivided populations represent the least biased estimates of *Nm*. Therefore, as the RFLPs and isozymes are still different, we can be confident that, although substructure inflates the values of *b* and *A* in log *Nm*-log distance regressions, detection of different patterns through the *r*² and Mantel test values is robust and relatively insensitive to substructure.

Cliff and harbour habitats

Figure 2 and Table 3 summarize the contribution of isolation by distance to the pattern of variation within the cliff and harbour habitats. In log *Nm*-log distance regressions of cliff populations, isozymes and RFLPs behave very similarly. Although it is not possible to test the significance of the difference

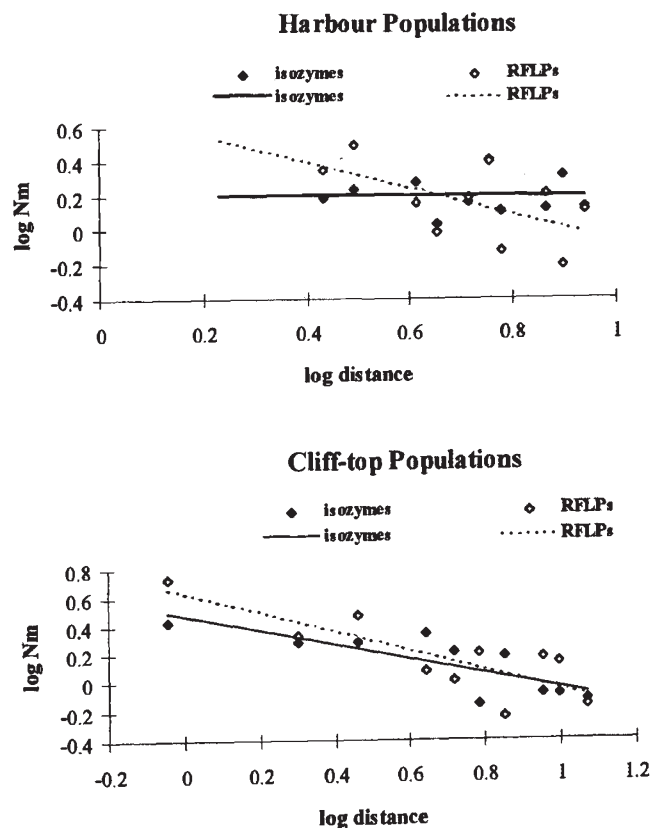


Fig. 2 Log-log regressions of gene flow against distance for RFLPs and isozymes among sea beetle populations in Poole Harbour (top) and adjacent cliff populations. Gene flow estimated from mean pairwise *F*_{ST} between patches (see text).

between the slopes or the intercepts, they appear very alike. Also, the *r*² values are both high and both slopes are significant. It appears, therefore, that isolation by distance acts almost identically on RFLPs and isozymes in cliff populations.

The pattern in the driftline habitats is very different. There is clearly no isolation by distance at

Table 3 Summary of regression equation statistics for the slopes in Fig. 2

	Cliff populations		Harbour populations	
	Isozymes	RFLPs	Isozymes	RFLPs
<i>b</i>	-0.5093	-0.6846	-0.0199	-0.7594
<i>A</i>	0.4727	0.6348	0.2129	0.6997
<i>r</i> ²	0.6753	0.6800	0.0009	0.3289
<i>P</i> (Mantel test)	0.0171	0.0334	0.4791	0.0768

b, regression slope of log *Nm* vs. log distance; *A*, y-intercept; *r*², proportion of variation in log *Nm* explained by log distance.

Table 4 Regression coefficients and significance levels from partial Mantel tests for relationship between pairwise *F*_{ST} and distance and pairwise *F*_{ST} and population membership for Harbour populations only of sea beet

	Partial regression of <i>F</i> _{ST} and distance (removes population effect)	Partial regression of <i>F</i> _{ST} and population membership (removes distance effect)
Isozymes	<i>b</i> = 0.0014 <i>P</i> = 0.6700 Not significant	<i>b</i> = -0.0356 <i>P</i> = 0.1571 Not significant
RFLPs	<i>b</i> = 0.0178 <i>P</i> < 0.0001 Significant at $\alpha = 0.1\%$	<i>b</i> = 0.0076 <i>P</i> = 0.7996 Not significant
All loci	<i>b</i> = 0.0115 <i>P</i> = 0.0003 Significant at $\alpha = 0.1\%$	<i>b</i> = -0.0104 <i>P</i> = 0.6502 Not significant

Significance levels derived by the sequential Bonferroni technique (Rice, 1989).

isozyme loci with a low *r*² value and nonsignificant Mantel test probability. The situation with the RFLPs is less clear. The *r*² of the regression is nearly 33 per cent, but the Mantel test gives a nonsignificant value. Although the Mantel test is robust, it is not particularly powerful and, therefore, is susceptible to Type II errors. Thus, isolation by distance is suggested, but not proven.

Because of the uncertainty surrounding the Slatkin test for isolation by distance at RFLP loci, a partial Mantel test was carried out on the patches in the harbour populations. The results are shown in Table 4. It is clear that there is significant isolation by distance at RFLP loci, but not at isozyme loci. Also, there is no effect of population membership for any combination of loci (that is to say, there is no significant difference between the *F*_{ST} values between patches in the same population and the *F*_{ST} values between patches in different populations). This is the opposite of the result found by Raybould *et al.* (1996b) on two transects of sea beet on Furzey

Island in Poole Harbour where transect membership rather than distance explained the *F*_{ST} values between patches.

Discussion

Log-log regressions of gene flow against distance to detect isolation by distance need careful interpretation. The first problem, which is common to all pairwise comparisons of genetic distance estimators, is that of nonrandom mating (significant *F*_{IS}) within populations. *F*_{ST} between two populations of which at least one has nonrandom mating is underestimated in comparison with *F*_{ST} between two panmictic groups. *Nm* is likewise overestimated. This could lead to the false inference of isolation by distance, if the nonpanmictic groups happen to be clustered. Therefore, in this paper the populations were subdivided to reduce the likelihood of this effect.

Secondly, it is not possible to use conventional tests of the significance of regressions because of the

nonindependence of population pairs. Thus, regressions need to be tested by a permutation method, such as that of Mantel (1967). It is not yet possible to test the difference between two regressions by this means. This is important because we need to be able to test whether the difference between the log Nm -log distance regression slopes has changed with reduced substructure.

Clearly, the regression slope and intercepts are not appropriate tests of the difference, because they change with reduced Nm as F_{ST} is less systematically underestimated (Table 2). The Mantel test probabilities and r^2 values, on the other hand, allow the effect of reduced substructure to be examined. Table 2 shows that r^2 values for both markers are little changed and that the Mantel test probabilities for the RFLPs are almost identical. Although the probability for the isozyme regression is lower, it is still not significant. These results indicate that the results of Raybould *et al.* (1996a) showing differences between RFLPs and isozymes were valid, despite the substructure within populations. The results in this paper also indicate that, providing there is no clustering of highly structured populations, the detection of isolation by distance among substructured populations by means of high r^2 and low Mantel test probabilities should be robust despite overestimation of Nm .

The results in Table 3 indicate that the effects of isolation by distance in sea beet populations are habitat-dependent. Isolation by distance among cliff populations is clearly shown by both isozymes and RFLPs. The situation is more complex in the harbour populations. Considering the Slatkin test first, isozymes show no evidence of isolation by distance with a very shallow regression slope (compared with the cliff populations), a high Mantel test probability and low r^2 . The RFLPs have a relatively high r^2 and a steep regression slope, whereas the Mantel test is just nonsignificant ($P \approx 0.08$). In view of the r^2 value, it is possible that the lack of power of the Mantel test gives a false acceptance of the null hypothesis of no isolation by distance. Thus, with the Slatkin test we can demonstrate strong differences between RFLPs and isozymes in the total sample that seem to be caused by effects within the harbour. It is not certain, however, whether the difference results from founder effects or from isolation by distance at RFLPs but not isozymes.

The lack of significant population effects in the partial Mantel test (Table 4) shows that founder effects are not the reason for the differences between isozymes and RFLPs. The significant partial regression coefficient for RFLPs shows that isolation

by distance does operate among the populations around the edge of the harbour but that its effect cannot be detected at isozyme loci. The result also confirms that a nonsignificant Mantel test of a 'Slatkin-type' regression does not necessarily disprove isolation by distance.

Balancing selection has been invoked to explain differences in gene flow at RFLP and isozyme loci in oysters (Karl & Avise, 1992). Salinity differences are known to be a cause of selection maintaining polymorphism at isozyme loci (Koehn & Hilbish, 1987; Riddoch, 1993), so it is not unreasonable to suppose that selection may operate on harbour sea beet populations. If balancing selection operates in the harbour, it does not do so in an obvious way. There is no evidence for heterozygote advantage at isozyme loci in the harbour populations compared with those on the cliffs (harbour $H_o = 0.205$, SE 0.029; cliffs $H_o = 0.220$, SE 0.014). Also, the gene diversity at isozyme loci of the harbour populations is not greater than that of the cliff populations (harbour $H = 0.239$, SE 0.021; cliff $H_t = 0.275$, SE 0.013), which might be expected if similar patterns of niche variation (Van Valen, 1965) within harbour populations was the source of selection.

It is interesting to note that founder effects were the main determinant of genetic structure of sea beet populations on driftlines on Furzey Island in Poole Harbour (Raybould *et al.*, 1996b), but that no founder effects could be detected among the populations around the harbour edge. The driftline populations are likely to be submerged by spring tides and to be affected by saltmarsh erosion occurring in Poole Harbour (A. J. Gray, personal communication). Thus, they may behave more as a metapopulation, with repeated colonization and extinction (Levins, 1970), compared with a more stable population structure on sea walls and shingle ridges, which are not eroding and are rarely inundated by spring tides. Detailed observations of population demography could verify this hypothesis but a brief survey of one of the Furzey Island transects in 1996 found a much smaller sea beet population than was present in 1993 (A. J. Gray, personal communication).

In summary, the difference in gene flow measured indirectly by RFLPs and isozymes in sea beet is not an artefact resulting from substructure within populations. Partial Mantel tests confirm that the source of the difference is in the harbour populations, in which RFLPs but not isozymes show isolation by distance. Partial Mantel tests also show that founder effects are not a significant factor in determining the genetic structure of sea beet populations on sea

walls and shingle ridges. This is a reflection of the greater stability of these habitats compared with the driftline habitat on Furzey Island. Further work will aim to test this conclusion with demographic studies and to find the source of the difference between RFLPs and isozymes by detailed analysis of variation at single loci.

Acknowledgements

This work was carried out under Contract PECD/7/8/190 of the Department of the Environment's Genetically Modified Organisms Research Programme. The views expressed are those of the authors and not necessarily those of the Department of the Environment.

References

- GOUDET, J. 1995. Fstat V-1.2: A computer program to calculate F-statistics *J. Hered.*, **86**, 485–486.
- KARL, S. A. AND AVISE, J. C. 1992. Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. *Science*, **256**, 100–102.
- KOEHN, R. K. 1978. Physiology and biochemistry of enzyme variation: the interface of ecology and population genetics. In: Brussard, P. (ed.) *Ecological Genetics: The Interface*, pp. 51–72. Springer-Verlag, New York.
- KOEHN, R. K. AND HILBISH, T. J. 1987. The adaptive importance of genetic variation. *Am. Sci.*, **75**, 134–141.
- LEVINS, R. 1970. Extinction. *Lect. Math. Life Sci.*, **2**, 75–77.
- LEWONTIN, R. C. 1991. Electrophoresis in the development of evolutionary genetics: milestone or millstone? *Genetics*, **128**, 657–662.
- LÖNN, M. 1993. Genetic structure and allozyme-microhabitat associations in *Bromus hordeaceus*. *Oikos*, **68**, 99–106.
- MANLY, B. F. J. 1991. *The Statistics of Natural Selection*. Chapman & Hall, London.
- MANTEL, N. 1967. The detection of disease clustering and a generalised regression approach. *Cancer Res.*, **27**, 209–220.
- MITTON, J. B. 1989. Physiological and demographic variation associated with allozyme variation. In: Soltis, D. E. and Soltis, P. E. (eds) *Isozymes in Plant Biology*, pp. 126–145. Chapman & Hall, London.
- NEVO, E., KRUGMAN, T. AND BEILES, A. 1994. Edaphic natural selection of allozyme polymorphisms in *Aegilops peregrina* at a Galilee microsite in Israel. *Heredity*, **72**, 109–112.
- PRENTICE, H. C., LÖNN, M., LEFKOVITCH, L. P. AND RUNYEON, H. 1995. Associations between allele frequencies in *Festuca ovina* and habitat variation in the alvar grasslands on the Baltic island of Öland. *J. Ecol.*, **83**, 391–402.
- RAINEY, D. Y., MITTON, J. B. AND MONSON, R. K. 1987. Associations between enzyme genotypes and dark respiration in perennial ryegrass, *Lolium perenne* L. *Oecologia*, **74**, 335–338.
- RAYBOULD, A. F., MOGG, R. J. AND CLARKE, R. T. 1996a. The genetic structure of *Beta vulgaris* ssp. *maritima* (sea beet) populations: RFLPs and isozymes show different patterns of gene flow. *Heredity*, **77**, 245–250.
- RAYBOULD, A. F., GOUDET, J., MOGG, R. J., GLIDDON, C. J. AND GRAY, A. J. 1996b. Genetic structure of a linear population of *Beta vulgaris* ssp. *maritima* (sea beet) revealed by isozyme and RFLP analysis. *Heredity*, **76**, 111–117.
- RICE, W. B. 1989. Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.
- RIDDOCH, B. J. 1993. The adaptive significance of electrophoretic mobility in phosphoglucose isomerase. *Biol. J. Linn. Soc.*, **50**, 1–17.
- SLATKIN, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.
- VAN VALEN, L. 1965. Morphological variation and niche width. *Am. Nat.*, **99**, 377–390.
- WATT, W. B. 1994. Allozymes in evolutionary genetics: self-imposed burden or extraordinary tool? *Genetics*, **136**, 11–16.
- WEIR, B. S. AND COCKERHAM, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.
- WRIGHT, S. 1943. Isolation by distance. *Genetics*, **28**, 114–138.