The genetic structure of *Beta vulgaris* ssp. *maritima* (sea beet) populations: RFLPs and isozymes show different patterns of gene flow

A. F. RAYBOULD*, R. J. MOGG & R. T. CLARKE

Institute of Terrestrial Ecology, Furzebrook Research Station, Wareham, Dorset BH20 5AS, U.K.

Genetic variation in 12 populations of sea beet was assessed at nine isozyme and seven RFLP loci. Mean observed heterozygosity, diversity index and number of alleles per locus were not significantly different between the two classes of marker. The genetic structure of 10 of the populations was analysed using seven of the isozymes and six of the RFLPs. F_{ST} values between all pairs of populations were calculated separately for the isozymes and RFLPs. F_{ST} values were converted to amounts of gene flow (*Nm*) between populations under the assumptions of an island model. A regression of log *Nm* against log distance was used to test for isolation by distance. Mantel tests showed a highly significant decrease in *Nm* with distance for RFLPs but not for isozymes. It is suggested that uniform balancing selection may operate to maintain approximately equal allele frequencies among populations at the isozyme loci. If this is true then isozymes may be unsuitable for modelling the spread of neutral transgenes.

Keywords: balancing selection, gene flow, genetic structure, isozymes, Mantel test, RFLPs.

Introduction

Over the past 30 years isozymes have been used extensively as markers to analyse the genetic structure of natural populations. Much of the theory underpinning these studies assumes that markers are neutral. There is increasing evidence, however, that some isozyme variation in both animals (Avise, 1993) and plants (Mitton, 1989; Lönn, 1993; Riddoch, 1993; Nevo *et al.*, 1994; Prentice *et al.*, 1995) is not neutral. Therefore it is of great interest to compare the patterns of genetic variation revealed by isozymes with those from other classes of marker that may be less subject to selection.

Single copy nuclear restriction fragment length polymorphisms (SCN RFLPs) offer a source of potentially neutral and highly variable markers (Clegg, 1989). Comparisons of the patterns of genetic structure revealed by isozymes and SCN RFLPs have been made in several animal species (see Avise, 1993 for a review) but there are few data from plants. Zhang *et al.* (1993) compared the variability of the two types of marker in 286 wild barley accessions and found that RFLPs showed more

between-population variation whereas isozymes had more within-population diversity. However, detailed spatial structure was not analysed. We are not aware of any other studies which compare isozymes with SCN RFLPs from natural plant populations (although there is work that compares isozymes with, for example, maternally inherited chloroplast RFLPs [McCauley, 1994]).

Here we analyse variation at seven isozyme and six SCN RFLP loci in 10 populations of wild sea beet, *Beta vulgaris* ssp. *maritima* (hereafter 'sea beet') from the Dorset coast. We compare the amounts of variation and patterns of gene flow detected by the two classes of marker.

Materials and methods

Collection of material

During the spring of 1994 leaf material of sea beet was collected from 10 sites on the Dorset coast (Fig. 1). Fifty plants were sampled from each population. All populations were approximately linear and therefore plants were sampled along transects, at 1 m intervals in regions of high plant density, and at minimum possible intervals in excess of 1 m in
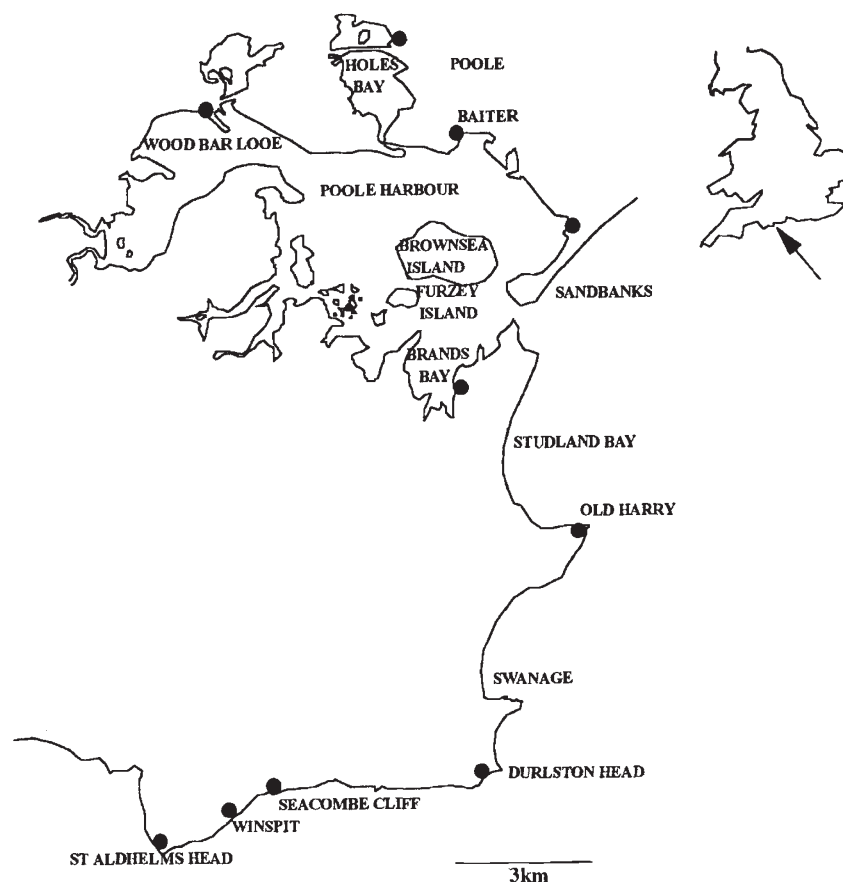
*Correspondence.

Fig. 1 Map of Poole Harbour and adjacent coast showing the location of the sea beet populations sampled.

populations with low plant density. Material was stored briefly (1–2 days) at 4°C prior to preparation of DNA and isozyme extracts.

Isozyme and RFLP analysis

Isozyme and RFLP analyses were carried out exactly as described by Raybould *et al.* (1996). The same loci were assayed, although problems with isozyme activity and probe labelling meant that malic enzyme (ME-2) and shikimic acid dehydrogenase (SkDH-1) and RFLP locus *L5* could not be included in this survey (Table 1). However, two additional isozyme loci were resolved; esterase (*Est-4*) on acrylamide gels and cathodal peroxidase (*Per-1*) on starch gels.

Statistical methods

The genetic variation detected by isozyme and RFLP markers was compared by calculating Nei's diversity index (H_t) (Nei, 1987), the observed heterozygosity (H_o) and the total number of alleles for each locus from the total sample collected in 1994 (a maximum of 500 plants). Previous work (Raybould

et al., 1996) had sampled some of these loci in 220 plants from two populations on Furzey Island (Fig. 1) in 1993. Therefore for several loci we present data from 12 populations with a maximum of 720 plants collected over two years. Two RFLP loci (*R1* and *R13*) had null alleles and therefore could not be used to calculate H_o. H_t values for these loci were obtained by estimating genotype frequencies corrected for inbreeding (see below). H_t, H_o and allele number were compared using *t*-tests with the degrees of freedom calculated assuming unequal variances (Bailey, 1959).

For genetic structure studies only plants from the 10 populations sampled in 1994 were considered. This avoided the problem of noncontemporaneous samples affecting the analysis. FSTAT (Goudet, 1995), a program which calculates Wright's *F*-statistics following Weir & Cockerham (1984), was used to calculate all 45 pairwise combinations of F_{ST} between the 10 populations over all isozyme loci and all RFLP loci. F_{ST} estimates were converted to amounts of gene flow (*Nm*) between all pairs of populations. *Nm* was calculated as $(1/4F_{ST}) - 0.25$.

As stated above, two RFLP loci (*R1* and *R13*) had

null alleles. They were used in calculating F_{ST} estimates (and H_t values) by adopting the following procedure. An F_{IS} estimate was obtained for each population using all the other 11 loci (RFLP and isozymes). (As we were testing for possible differences between RFLPs and isozymes this was a more conservative procedure than using the F_{IS} value from RFLPs only.) Then for each population the genotype proportions were estimated using Wright's (1921) method of correcting the Hardy–Weinberg formula for inbreeding. If the null allele ($-$) frequency is q and the detectable allele ($+$) frequency is p ($= 1-q$) then genotype proportions are:

$$+ +\quad p^2(1-F)+pF$$
$$+ -\quad 2pq(1-F)$$
$$- -\quad q^2(1-F)+qF,$$

where F is F_{IS} for the population in question. Thus the proportion of null homozygotes and F_{IS} for each population allow genotype proportions to be calculated. In practice, genotype frequencies estimated by this procedure are little different from Hardy–Weinberg frequencies.

The relationship between gene flow (Nm) and geographical distance between all pairs of populations was studied by a regression of log Nm (estimated from either isozymes or RFLPs) on log distance (Slatkin, 1993). As the 45 pairs of values are not independent of each other (they are calculated from all pairwise combinations of 10 populations), the significance of the correlation and hence regression slopes was tested using a Mantel randomization test (Mantel, 1967) as described by Manly (1985). It is important to realise that this approach correctly randomizes the n populations, not the $n(n-1)/2$ pairs of populations. For example, consider $n = 4$ populations (1–4) with pairwise log Nm and log distance values as follows:

Population i	Population j	Log Nm y_{ij}	Log distance x_{ij}	$x_{r(m)}$
1	2	1.20	0.49	$x_{k1,k2} = 0.78$
1	3	0.68	0.76	$x_{k1,k3} = 0.43$
1	4	0.29	0.94	$x_{k1,k4} = 0.49$
2	3	0.74	0.43	$x_{k2,k3} = 0.65$
2	4	0.03	0.78	$x_{k2,k4} = 0.94$
3	4	0.12	0.65	$x_{k3,k4} = 0.76$

Let Z denote the observed correlation between y_{ij} and x_{ij} based on all six ij population pairings.

Mantel's test involves keeping one column (say, y_{ij}) constant and then repeatedly randomly reordering the populations to derive random columns of x_{ij}. If the mth random reordering of the populations gives the order $k1$, $k2$, $k3$, $k4$, where $k1 = 2$, $k2 = 4$, $k3 = 3$, $k4 = 1$, then a new $x_{r(m)}$ column is derived as shown above, and the correlation $Z_{r(m)}$ between y and $x_{r(m)}$ is calculated. If, as here, the one-sided non-null hypothesis is that y is negatively correlated with x, the statistical significance of the observed correlation Z is the proportion of the N (say 10 000) values $Z_{r(m)}$ which are less than or equal to the observed value Z.

Results

Genetic variation

Table 1 gives H_t, H_o and the number of alleles for each locus. For each measure RFLPs showed more variation than isozymes, although in each case the difference was not significant. Hence these data provide no evidence that RFLPs reveal more variation than isozymes in sea beet, as might be expected from theory (Clegg, 1989). A similar result was found when only the Furzey data were analysed (Raybould *et al.*, 1996).

Gene flow

Gene flow rates estimated from isozyme and RFLP loci were compared using the regression method of Slatkin (1993) in which the log of gene flow between pairs of populations is regressed on the log of distance between the populations. A significant negative regression indicates isolation by distance. Gene flow (Nm) values for isozymes and RFLPs were obtained from the F_{ST} values (see methods). Distances in km were measured 'as the crow flies'.

Plots of log Nm on log distance show strikingly different patterns for the two groups of loci (Fig. 2). For isozymes there is no relationship between gene flow and distance ($b = 0.003$, Mantel test significance probability level $= 0.4913$ with 10 000 randomizations). However, for RFLPs there is a significant negative relationship ($b = -0.822$, Mantel test significance probability level $= 0.0006$ with 10 000 randomizations) with log distance explaining 39.8 per cent of the variation in log Nm values.

Strictly F_{ST} should be estimated between randomly mating groups (e.g. Raybould *et al.*, 1996). Randomization tests (Goudet, 1995) indicated that the mean F_{IS} values for the 10 populations calculated over all RFLP and all isozyme loci were significantly

different from zero ($P<0.001$ in both cases; 10 000 randomizations). This may lead to a general over-estimation of Nm (underestimation of F_{ST}). However, this does not invalidate the conclusion that Nm declines with distance for RFLPs but not for isozymes (high R^2 and highly significant Mantel test significance for RFLPs, low R^2 and nonsignificant Mantel test significance for isozymes), but only

Table 1 Diversity indices (H_t), observed heterozygosities (H_o) and number of alleles at 16 polymorphic loci in 12 pooled populations of sea beet from Dorset, U.K.

	Plants	H_t	H_o	Alleles
Isozyme loci				
Acph-3	718	0.487	0.284	4
*Est-4	498	0.140	0.110	5
Got-3	716	0.494	0.420	2
Got-4	719	0.148	0.067	3
†Me-2	220	0.530	0.446	3
*Per-1	500	0.190	0.178	3
6Pgdh-2	697	0.289	0.178	2
Pgi-2	718	0.108	0.081	2
†Skdh-1	220	0.031	0.014	3
Mean ± SE		0.268 ± 0.063	0.200 ± 0.053	3.000 ± 0.333
RFLP loci				
L3	716	0.238	0.212	2
†L5	220	0.018	0.018	2
L9	714	0.513	0.381	4
‡R1	720	0.513	N/A	3
R4	660	0.566	0.421	3
R7	714	0.574	0.209	6
‡R13	686	0.521	N/A	3
Mean ± SE		0.421 ± 0.080	0.270 ± 0.063	3.286 ± 0.522
P (isozyme = RFLP)		$t_{12} = 1.50$	$t_{11} = 0.86$	$t_{10} = 0.46$
		$P = 0.16$	$P = 0.41$	$P = 0.65$

*Loci not sampled from Furzey Island; †loci sampled from Furzey Island only; ‡loci with null alleles.

ISOZYMES

Y = 0.426 + 0.003X   R-squared = 0.000

RFLPs
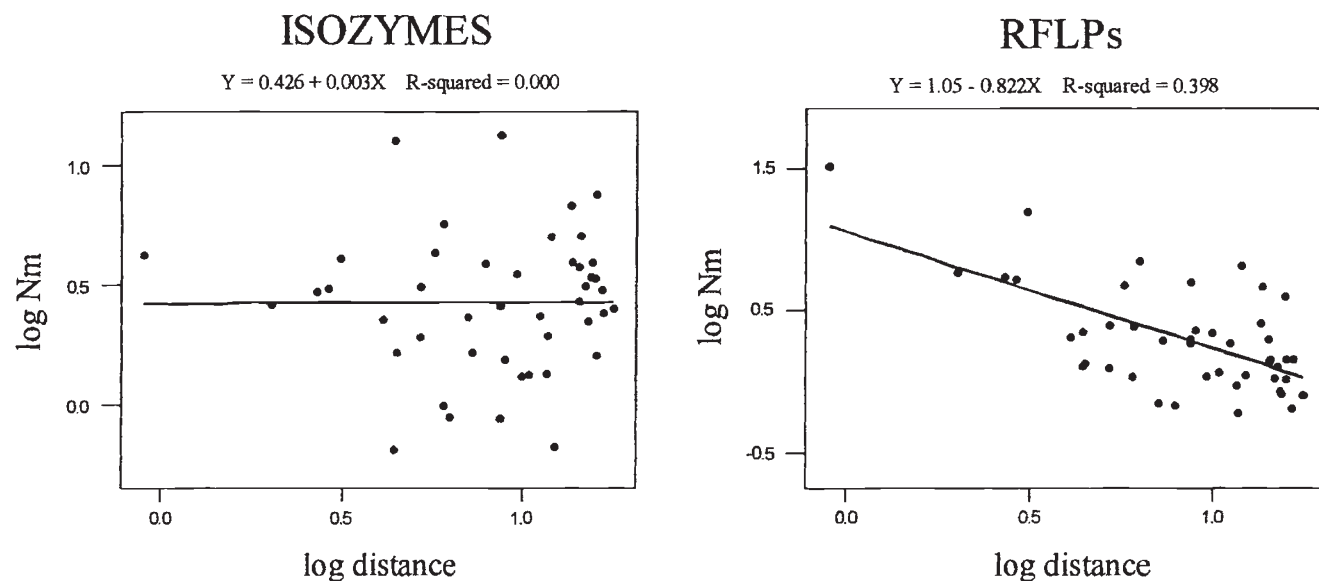
Y = 1.05 - 0.822X   R-squared = 0.398

Fig. 2 Linear regressions of log of gene flow (Nm), estimated from isozymes or RFLPs, against log of distance between all pairs of sea beet populations. P (Mantel test) = 0.4913 for isozymes and 0.0006 for RFLPs.

Table 2 Comparison of F_{IS} and F_{IT} for isozymes and RFLPs of sea beet

	Isozymes	RFLPs	P (Isozyme = RFLP)
F_{IS}	0.183 ± 0.053	0.079 ± 0.019	$t_9 = 1.85; P = 0.05$
F_{IT}	0.262 ± 0.057	0.214 ± 0.024	$t_9 = 0.78; P > 0.05$
P ($F_{IS} = F_{IT}$)(one-tailed)	$t_{12} = 1.01; P > 0.05$	$t_{10} = 4.41; P \simeq 0.001$	

Mean $\pm$ standard error obtained by jackknifing over loci.

indicates that the y intercepts and the slopes of the regressions may be overestimated.

Mean pairwise F_{ST} values for isozymes and RFLPs cannot be compared by conventional t-tests because of the nonindependence of F_{ST} values between pairs (see above). However, F_{IS} and F_{IT} values can be compared. It can be seen (Table 2) that F_{IS} for isozymes is higher than that for RFLPs, although the difference is not significant. F_{IT} values are also not significantly different. It is interesting to note that F_{IS} is not significantly different from F_{IT} for isozymes, whereas it is for RFLPs. This provides further evidence for isolation by distance in RFLPs but not isozymes.

Discussion

The data show that for three measures (H_t, H_o and number of alleles) RFLPs do not show significantly greater amounts of variation than isozymes (Table 1). Similar results were found when the Furzey data only were analysed (Raybould *et al.*, 1996) indicating that this effect seems consistent over different sampling scales. Zhang *et al.* (1993) also found that H_o and the number of alleles at RFLP loci were not significantly higher than the values for isozyme loci, although H_t was significantly higher.

It might be expected that RFLPs would show a higher number of alleles per locus than isozymes because (1) not all mutational changes at an isozyme locus cause a change in enzyme mobility, whereas, in theory, any mutational change can be detected by RFLP methods; and (2) RFLPs can detect polymorphism in noncoding DNA where there may be fewer functional constraints than in coding regions (Clegg, 1989). There is no evidence for this in the sea beet populations analysed here. One possible reason is that the probes used are cDNAs and therefore will detect polymorphisms in regions flanking coding DNA. These regions may contain control sequences and so be under functional constraints similar to the isozyme coding sequences.

Regressions of log Nm on log distance coupled with Mantel tests showed clear differences between RFLPs and isozymes. For neutral markers the genetic distance between pairs of populations (e.g. F_{ST}) should increase with geographical distance. An alternative way of expressing this is to say that gene flow (here denoted Nm) decreases with distance (Slatkin, 1993). Figure 2 shows clearly that this is true for RFLPs but not isozymes. Hence isozymes may not be acting as neutral markers in these populations.

This is very similar to the situation in the American oyster. Karl & Avise (1992) found uniform isozyme allele frequencies in populations throughout the Gulf of Mexico and the south-eastern United States. However, mitochondrial and SCN RFLPs showed clear allele frequency differences between the Gulf and Atlantic populations. Thus isolation by distance, caused by limited larval dispersal, appears to be operating in the oysters, but isozyme allele frequencies remain fairly constant. Karl & Avise (1992) suggested that several of the isozyme loci are under uniform balancing selection and so do not reflect the population subdivision resulting from limited dispersal.

In sea beet, therefore, isozyme variation may be under uniform selection which keeps allele frequencies approximately equal among populations in spite of limited pollen and/or seed dispersal. If this is true then isozymes may be unsuitable for modelling the spread of neutral transgenes, as discussed by Raybould *et al.* (1996). An alternative explanation is that isozymes give an accurate picture of the extent of gene flow, that is to say seeds and pollen are widely dispersed among the populations, with little variation in the rates of dispersal between pairs of populations different distances apart. With this scenario RFLPs could be under disruptive selection. Further work will aim to distinguish between these possibilities by comparing the relationship between Nm and distance in habitats where seed dispersal is likely to be extensive (drift lines) and more limited (cliffs).

Acknowledgements

References

AVISE, J. C. 1993. *Molecular Markers, Natural History and Evolution*. Chapman and Hall, New York and London.

BAILEY, N. J. T. 1959. *Statistical Methods in Biology*. English Universities Press, London.

CLEGG, M. T. 1989. Molecular diversity in plant populations. In: Brown, A. H. D., Clegg, M. T., Kahler, A. L. and Weir, B. S. (eds) *Plant Population Genetics, Breeding, and Genetic Resources*, pp. 98–115. Sinauer Associates, Sunderland, MA .

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.

LÖNN, M. 1993. Genetic structure and allozyme-microhabitat associations in *Bromus hordeaceus. Oikos*, **68**, 99–106.

McCAULEY, D. E. 1994. Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of *Silene alba*: Implications for the study of gene flow in plants. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 8127–8131.

MANLY, B. F. J. 1985. *The Statistics of Natural Selection on Animal Populations*. Chapman and 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 and Hall, London.

NEI, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

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.

RAYBOULD, A. F., GOUDET, J., MOGG, R. J., GLIDDON, C. J. AND GRAY, A. J. 1996. The genetic structure of a linear population of sea beet (*Beta vulgaris* ssp. *maritima*) revealed by isozyme and RFLP analysis. *Heredity*, **76**, 111–117.

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.

WEIR, B.S. AND COCKERHAM, C. C. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

WRIGHT., S. 1921. Systems of mating, I-V. *Genetics*, **6**, 111–178.

ZHANG, Q., SAGHAI MAROOF, M. A. AND KLEINHOFS, A. 1993. Comparative diversity of RFLPs and isozymes within and among populations of *Hordeum vulgare* ssp. *spontaneum. Genetics*, **134**, 909–916.