Wilding, CS, Butlin, RK and Grahame, J

Differential gene exchange between parapatric morphs of Littorina saxatilis detected using AFLP markers

http://researchonline.ljmu.ac.uk/id/eprint/2717/

Citation (please note it is advisable to refer to the publisher’s version if you intend to cite from this work)

Differential gene exchange between parapatric morphs of *Littorina saxatilis*
detected using AFLP markers

C. S. Wilding, R. K. Butlin and J. Grahame*

Centre for Biodiversity and Conservation, School of Biology, The University of
Leeds, Leeds LS2 9JT

*Author for correspondence: Tel. 0113 233 2852, FAX 0113 233 2835, Email
j.w.grahame@leeds.ac.uk

Word count: 6066 including tables and references.

Running head: AFLP markers of differential gene exchange

Key words: speciation, disruptive selection, clines, AFLP, *Littorina.*
Abstract

Speciation requires the acquisition of reproductive isolation, and the circumstances under which this could evolve are of great interest. Are new species formed after the acquisition of generalized incompatibility arising between physically separated populations, or may they arise as a result of the action of disruptive selection beginning with the divergence of a rather restricted set of gene loci? Here we apply the technique of Amplified Fragment Length Polymorphism (AFLP) analysis to an intertidal snail whose populations display a cline in shell shape across vertical gradients on rocky shores. We compare the $F_{ST}$ values for 306 AFLP loci with the distribution of $F_{ST}$ estimated from a simulation model using values of mutation and migration derived from the data. We find that about 5% of these loci show greater differentiation than expected, providing evidence of the effects of selection across the cline, either direct or indirect through linkage. This is consistent with expectations from non-allopatric speciation models that propose an initial divergence of a small part of the genome driven by strong disruptive selection while divergence at other loci is prevented by gene flow. However, the pattern could also be the result of differential introgression after secondary contact.
Introduction

The process of speciation requires the acquisition of reproductive isolation. If populations are separated by a physical barrier to dispersal, speciation may follow: the acquisition of intrinsic reproductive isolation is then an incidental consequence of the accumulation of genetic differentiation (Mayr, 1963). Increasingly, attention has shifted to the possibility that reproductive barriers might arise in populations not separated by major physical features (Bush & Howard, 1986), i.e. that speciation might begin with genetic diversification in spite of some gene exchange between constituent populations. Empirical evidence shows, for example, that a single founder population in a lake may diversify and undergo speciation following use of different niches (Schliewen et al., 1994; Schluter, 1996; Wilson et al., 2000), and theoretical work suggests that gene flow can be less of a cohesive force than previously thought (Barton, 1988).

Barton (1988) and Rice & Hostert (1993) have reviewed the literature on speciation mechanisms, showing that there are plausible and simple models of non-allopatric speciation. In these models, genetic divergence may be initiated by disruptive selection without a period of extrinsic isolation. This requires strong selection and either pleiotropy or linkage of the genes involved in the adaptive polymorphism with those affecting the probability of gene exchange. For parapatric populations, where gene exchange is restricted, an initial level of differentiation may be modified to increase isolation by the accumulation of different alleles in the diverging genetic backgrounds. Strong selection also is needed here if gene flow is other than negligible. Nevertheless, Rice & Hostert (1993) concluded that laboratory experiments on the development of isolation strongly support the idea that
reproductive isolation can evolve between sympatric or parapatric populations if divergent selection is strong relative to gene flow.

Although the conclusion of Rice & Hostert (1993) is well supported by laboratory experiments, there is less evidence from natural populations. Host races provide the best examples, especially *Rhagoletis* (Feder et al., 1994; Feder et al., 1997). Host fidelity provides the major barrier to gene exchange, permitting further differentiation under selection on the alternative hosts. Some markers (presumably those linked to selected loci, or perhaps under selection themselves) show allele frequency differentiation, while others do not, suggesting that gene exchange is more restricted in some parts of the genome than others. This may be viewed as a signature of non-allopatric speciation and is in contrast to the generalised barrier to gene flow that results from physical isolation. The uniform divergence across the genome that evolves in allopatry may be maintained following secondary contact due to the accumulation of genetic incompatibility at many loci that is revealed in some hybrid zones (Barton & Hewitt, 1981; Szymura & Barton, 1991). However, it may be eroded by introgression.

We address the issue of uniform versus restricted differentiation using a system where divergent populations are parapatric. They are likely to be exchanging genes only in the region of contact, and the selection gradient on which they exist is imposed by the physical environment and by predation. *Littorina saxatilis* (Olivi) (the 'rough periwinkle') is widespread on North Atlantic shores, exhibits high morphological and allozyme variability, and is ovoviviparous and of low vagility - see Reid (1996) for review. In Britain it is found as two morphological forms ('H' and 'M') (Hull et al., 1996) that show good evidence of partial reproductive isolation. This interpretation was based on reduced fertility in females inferred to be hybrids, and is
supported by the observation of assortative mating (Hull, 1998; Pickles & Grahame, 1999). The observed differentiation could be attributed to secondary contact between populations that had been undergoing allopatric divergence. Alternatively we may be seeing divergence in situ due to strong selection, despite gene flow (Endler, 1977; Rice & Hostert, 1993). In either case, the current pattern of differentiation is probably maintained by a balance between gene flow and selection, where the selection is due, at least in part, to environmental pressures rather than genetic incompatibility.

Predation by crabs is thought to exert strong selection on periwinkle shell form (Heller, 1976; Raffaelli, 1978; Janson, 1983; Johannesson, 1986), and among molluscs more widely - see Vermeij (1987) for review. Both thickness and form of the periwinkle shell may vary adaptively in response to differing predation pressures, and inducible phenotypic responses are considered to be involved for thickness changes in at least some species (Trussell & Smith, 2000). However, there is abundant evidence that in L. saxatilis some of the variation is genotypic (Newkirk & Doyle, 1975; Grahame & Mill, 1993; Johannesson & Johannesson, 1996), and this is especially likely for shell shape. Because crab predation increases down the shore in most sites, clines in shell shape are often found (Grahame et al., 1997). In the upper shore, L. saxatilis H are thin-shelled, wide-apertured animals with relatively low spires. This shape may come about simply as a result of the constraints on shell shape when the aperture is large (Clarke et al., 1999) thus affording greater foot area (Grahame & Mill, 1986) for adhesion and leading to greater gravitational stability (Heller, 1976). Therefore, this is probably the optimum shape for maintaining a grip on wave or wind-affected substrates in the absence of crab predation. In the lower shore L. saxatilis M are thicker shelled, with relatively smaller apertures; these features are
likely to be adaptive in reducing the risk of crab predation (Johannesson, 1986; Boulding et al., 1999).

Primary and secondary origins of clines are notoriously difficult to distinguish (Barton & Hewitt, 1985). Wilding et al. (2000) considered it probable that the current distribution of mitochondrial haplotypes in *L. saxatilis* in the British Isles indicated expansion from different glacial refugia. However, the distribution of the H and M forms is quite different from that described for these haplotypes (Wilding et al., 2000), and Wilding et al. (2001) concluded that the current haplotype distribution was unrelated to whether populations were H or M morph. We tentatively suggest that the *L. saxatilis* H-M cline has evolved *in situ*.

Here we examine putative loci (hereafter, simply 'loci') revealed by the Amplified Fragment Length Polymorphism technique (AFLP) (Vos et al., 1995) in samples from four locations on the coast of Yorkshire, England. We compare observed $F_{ST}$ distributions across loci between populations of *L. saxatilis* H and M with $F_{ST}$ distributions in within-morph comparisons, and with expected distributions. These expected distributions were derived from simulations of $F_{ST}$ values in the absence of selection, using an approach analogous to that of Beaumont & Nichols (1996). We ask whether the barrier to gene exchange between H and M populations is uniformly effective across loci.
Materials and methods

Sampling

Periwinkles were collected from rocky shores at Thornwick Bay, Flamborough (British Grid reference TA 233724), Filey Brigg (TA 132815), Old Peak (NZ 982024) and Robin Hood's Bay (NZ 955055). The coast trends overall northwesterly in this region, the straight line distances between the sites are: Flamborough - Filey Brigg, 15 km (we estimate that 60% of the intervening shore represents suitable habitat for 

*L. saxatilis*); Filey Brigg - Old Peak, 26 km (80% suitable habitat); Old Peak - Robin Hood's Bay, 4 km (90% suitable habitat). At each site snails were collected from each of two locations (one in an area occupied by the H morph and one in an area occupied by the M morph, except at Robin Hood’s Bay), individual snails were taken from an area of about 2 m$^2$. H and M animals were characterized on the basis of sample location and shell form (by eye), and only brooding females were used to avoid contaminating the H samples with specimens of *Littorina arcana* Hannaford Ellis (which lay eggs on the shore). Sampling locations were 5 m apart at Flamborough, 15 m apart at Filey, 300 m apart at Old Peak, and 75 m apart at Robin Hood's Bay. In the first three instances, these distances were dictated by the presence of workable abundances of the animals, the aim being to sample from H and M populations which were as close to one another as possible. At Robin Hood’s Bay the samples were of M animals only, 75 m was chosen as a distance likely to be considerably in excess of migration distance (Janson, 1983).
DNA isolation

Genomic DNA was purified from head-foot tissue of individual *Littorina saxatilis* using a modified version of Winnepenninckx *et al.* (1993). Tissue was macerated in 300μl 60°C CTAB buffer (2% CTAB, 1.4M NaCl, 20mM EDTA, 100mM Tris-HCl pH 8, 0.2% β-mercaptoethanol) to which 20mg proteinase K was added and incubated at 60°C for 3-16hr. Subsequently, two extractions with chloroform:isoamyl alcohol (24:1) were performed, and the DNA further purified with Promega’s Wizard DNA Clean-Up System following the manufacturer’s instructions. Concentration was assessed by spectrophotometry and adjusted to 100 ng.μl⁻¹.

AFLP analysis

AFLP analysis was performed using a modified version of Vos *et al.* (1995). Adapter and primer sequences are given in Table 1. For each sample genomic DNA (500ng) was digested with 5U EcoRI (NEB) and 3U MseI (NEB) in 25μl total volume of 1x NEB buffer #2 supplemented with 100μg.ml⁻¹ BSA, for 3 hours at 37°C. Following enzyme inactivation at 65°C, 25μl of a solution containing 5pmol EcoRI adapter, 50pmol MseI adapter, 200U DNA ligase (NEB) and 5μl 10x ligase buffer (NEB), was added and samples incubated for 16 hours at 16°C. Preselective PCRs were then performed on 5μl diluted ligation (1:9 with 0.1xTE) in 50μl volumes containing 200μM each dNTP, 25pmol Eco+(C/A) primer, 25pmol Mse+(C/A) primer, 1.5mM MgCl₂ and 1U Taq in manufacturer’s buffer. PCR conditions were 20x(94°C 30secs, 56°C 1min, 72°C 1min). Selective Eco+3 primers were labeled in 0.5μl volumes containing 1xT4 PNK buffer, 0.2μl T4 PNK (Promega) 5ng Eco+3 primer and 0.1μl γ³³P ATP. Selective PCRs were undertaken in 20μl volumes containing 30ng Mse+3
primer (see Table 1), 5ng labeled Eco+3 primer, 200μM each dNTP, 1.5mM MgCl₂, 1x buffer (Promega) and 0.4U Taq. Cycling conditions in the first cycle were 94°C 30secs, 65°C 30sec, 72°C 1min with the annealing temperature reduced by 0.7°C over next 12 cycles, then 23x(94°C 30secs, 56°C 30sec, 72°C 1min). On completion, 20μl STOP solution (95% formamide, 10mM EDTA pH8.0, 0.025% w/v bromophenol blue, 0.025% w/v xylene cyanol) was added. AFLP products were separated on 6% polyacrylamide gels (Sequagel, Flowgen), for 2-2½ hours at 55W then fixed, and dried to the glass plate. Kodak Biomax MR-1 film was exposed to the gel for 48 hours. An initial study of reproducibility showed absolute consistency of banding patterns between repeated reactions. Subsequent monitoring where ≈5% reactions were repeated has confirmed this.

**Data analysis**

Gels were scored manually for band presence/absence. The frequency of the band presence allele was estimated from the band presence/absence matrix for each sample as $p = 1 - ((N-C)/N)^{0.5}$ where $N =$ sample size and $C =$ number of individuals with the band. This calculation assumes Hardy-Weinberg genotypic frequencies and dominance of band presence over absence.

We wish to use the allele frequency data for H and M samples to distinguish two possibilities: 1. That all loci reflect mutation/drift/dispersal balance, perhaps influenced by some general intrinsic barrier to gene exchange between H and M populations, or 2. That strong differentiation is maintained by selection at some proportion of loci, against a background of less-differentiated loci. We followed the approach developed by Bowcock *et al.* (1991) and Beaumont & Nichols (1996) by using simulations to predict the expected distribution of differentiation across loci for
a given average divergence. Differentiation is measured by $F_{ST}$, calculated for each
locus by the method of Nei (1977) with the correction suggested by Nei & Chesser
(1983). Simulation is necessary because the distribution of $F_{ST}$ across loci is
influenced by historical sampling in the natural populations (i.e. by genetic drift) and
by experimental sampling. Here there is the added complication that AFLP loci are
dominant and, therefore, the experimental sampling error of $F_{ST}$ is greater for high
mean allele frequencies (of the ‘presence’ allele) than for low frequencies. This is
because the allele frequencies have to be estimated from the proportion of ‘absence’
homozygotes and the errors are greatest when this proportion is low.

We have used a simple simulation of two populations of size $N$ diploid individuals,
with mutation rate $\mu$ and migration rate $m$, per generation. Allele frequencies for 500
simulated bi-allelic loci were initiated with a uniform random distribution, equal in
the two populations and then allowed to drift for $10N$ generations. Samples of 50
individuals were then taken from each simulated population and mean allele
frequencies and $F_{ST}$ values were calculated in exactly the same way as for the
observed data (with the band presence allele dominant to the absence allele). The
simulation was checked by comparing the $F_{ST}$ calculated in this way with both the $F_{ST}$
expected from theory and the $F_{ST}$ calculated from the whole simulated population (i.e.
without sampling effects). The theoretical $F_{ST}$ was calculated from

$$F_{ST} = \frac{1}{1 + 16Nm + 16N\mu}$$

since only two populations are considered and the mutation
rate may be high relative to the migration rate (see below) (Crow & Aoki, 1984). The
simulated values calculated from the whole population agreed precisely with this
expectation but the simulated sample values showed a consistent upward bias of
0.0093 over the range of values of $Nm$ relevant to this study. This bias is consistent
with previous simulation studies using Nei’s method for calculation of $F_{ST}$ (Slatkin & Barton, 1989).

For each comparison between observed samples, $Nm$ in the simulation was set to a value expected to return the observed mean $F_{ST}$ allowing for the estimation bias. The simulation was then repeated 50 times to generate a total of 25000 values of mean allele frequency and $F_{ST}$ (minus those loci that were monomorphic in the simulated samples, approximately 5%). Simulated mean $F_{ST}$ values differed from observed means by up to 6.77% but were always higher, making the test for loci with unexpectedly high levels of differentiation conservative. Observed $F_{ST}$ values were compared with the 0.99 quantile of the simulated values determined for each of 20 categories of mean allele frequency, because the distribution of $F_{ST}$ values is expected to vary with mean allele frequency (see below and Fig. 1).

**Results**

**Levels of polymorphism**

A total of 306 fragments (loci) were scored from five primer combinations for 50 individuals per sample (Table 2). Additional, variable fragments could not be scored unambiguously and were not considered further. Levels of polymorphism were particularly high with 94.8% of loci polymorphic (a locus was considered polymorphic if at least one individual showed a variant pattern). There was some variation in the number of scorable loci per primer combination with the $Eco$+CTC-$Mse$+CGA yielding 43 polymorphic bands and $Eco$+CAG-$Mse$+CGA yielding 80. This high level of polymorphism suggests a value for $N\mu$ of the order of $10^{-1}$, using Kimura’s (1968) formula for bi-allelic loci. This formula assumes symmetrical mutation, which may not be true for AFLP bands, and ignores the possible existence
of many loci that are monomorphic for the ‘absence’ allele. This may mean that \( N\mu \)
has been overestimated. We have used \( N\mu = 0.1 \) (\( N = 10^3, \mu = 10^{-4} \)) in the simulations
reported below but other runs have demonstrated that neither the mean nor the
variance of \( F_{ST} \) is sensitive to these parameters (as also observed by Beaumont &
Nichols 1996). We have also run simulations with the mutation rate from presence to
absence ten times greater than the reciprocal rate. This increases the proportion of loci
monomorphic for the absence allele but has no effect on the distribution of \( F_{ST} \).

Detection of differentiated loci

Ten loci had \( F_{ST} \) values higher than the 0.99 quantile of the initial simulation results
for all three individual H-M comparisons. Since these loci are implicated as being
under selection or linked to areas of the genome that are under selection, \( Nm \) was
recalculated after their removal, simulations were repeated, and the data compared
with new 0.99 quantiles. This process was carried out four times. At this stage, no
further locus showed observed values of \( F_{ST} \) lying above the 0.99 quantiles in all three
H-M comparisons, and 15 loci were identified as lying above the 0.99 quantile (Fig.
1). If the three H-M comparisons were independent, one would expect to see \(<1
locus falling outside the 0.99 quantile in all three cases (\( 0.01^3 \times 306 \)). However, gene
exchange between sites potentially means that allele frequencies do not vary
independently. Therefore, we repeated the analysis making the alternative extreme
assumption that the three H samples come from one population and the three M
samples from another. In this case, all 15 of the loci previously identified fell outside
the 0.99 quantile (now based on sample sizes of 150).

In all three H-M comparisons on the same shore, the same 15 loci lie above the
0.99 quantile, together with a much smaller number of other loci whose behaviour is
erratic. In comparisons within morphs, mostly also between shores, there are fewer loci above the 0.99 quantile, they are nearer to this limit, and rarely are any of the 15 loci identified above involved (see Fig. 1).

Table 3 shows that when $F_{ST}$ is calculated using all loci, values are usually higher for H-M comparisons than they are for H-H or M-M comparisons. The few within morph comparisons which are as large as the smallest between morph ones are from samples at or near the extremes of the sample range, e.g. Old Peak H-Thornwick Bay H (0.0318). Yet overall, $F_{ST}$ seems to be independent of distance, thus the $F_{ST}$ for H-M at Thornwick Bay is 0.0378 (spatial distance 5 m) while the values for H at Thornwick Bay compared with the two M samples at Robin Hood's Bay (distance 45 km) are 0.0350 and 0.0340. The lack of relationship between all $F_{ST}$ values and linear distance is further suggested by a randomization test (Manly, 1996; Manly, 1997) (1000 permutations) when the value of P for association was 0.3690. However if $F_{ST}$ is estimated after removal of the 15 loci considered to be differentiated between H and M (Fig. 2), there is evidence of association with distance, $P = 0.0020$. In the figure, and for the randomization tests, distance was transformed by taking base 10 logarithms, $F_{ST}$ by taking $F_{ST}/(1 - F_{ST})$ as recommended by Rousset (1997).

Two-sample randomization tests (Manly, 1996; Manly, 1997) were carried out on the $F_{ST}$ data in Table 3 either when the values were calculated with, or without, the 15 loci considered as likely to be differentiated. For values including these 15 loci, the probability that within morph and between morph $F_{ST}$ values were the same was $P = 0.001$. When these 15 loci were excluded from the $F_{ST}$ estimates, this probability became 0.1450, indicating no difference between the two groups of $F_{ST}$ estimates.

Mean $F_{ST}$ values after removal of these 15 loci imply that $Nm$ between H and M
mophs within shores is in the range 5.5 at Old Peak, 6.3 at Thornwick Bay and 308
individuals per generation (respective $Nm$ values were 1.9, 2.0 and 3.9 before
removal). $Nm$ between M morphs at Robin Hood's Bay is estimated as infinity ($F_{ST} =
0$).

**Genetic variation among *L. saxatilis* populations**

Nei’s genetic distances between samples of *L. saxatilis* H and M were used to
construct a neighbour-joining tree (Fig. 3a). The three samples of *L. saxatilis* H form
one cluster separated from the five samples of *L. saxatilis* M by the greatest internal
branch length and with high bootstrap support. When we omitted the data for the 15
loci identified as potentially under selection from the three comparisons of *L. saxatilis*
H and M, the revised tree showed radically altered structure (Fig. 3b). Now, instead of
populations clustering by morphotype (H and M), they cluster by site, with Filey H
and M clustering together, Old Peak H and M together, etc.

Fig. 3 about here
Discussion

This study asks whether the *Littorina saxatilis* H-M cline represents a general barrier
to gene exchange or reflects divergence at a limited number of loci under selection.

By generating a large number of marker loci using AFLPs, and using the analytical
approach of Beaumont & Nichols (1996), we have identified at least 15 loci (from a
total of 306 studied; 5%) that seem either to be under selection or (more likely) linked
to loci that are. However, none of the 306 loci is implicated as under selection when
two populations of *L. saxatilis* M are compared from the same shore (Robin Hood’s
Bay). It is interesting that our H-M comparisons show differentiation at these loci
regardless of whether they are spatially widely separate (300 m at Old Peak) or close
 together (5 m at Flamborough). Within morph comparisons do not show such
differentiation, and now there is evidence of isolation-by-distance. $F_{ST}$ values for
between morph comparisons are evidently higher than for within morph comparisons
when all loci are considered. The $F_{ST}$ values after removal of these exceptional loci
are more nearly similar, but still imply that there is a general barrier to gene exchange
between H and M populations that is greater than would be expected from their spatial
separation.

Our simulation assumes free recombination among loci. In reality, this is clearly
not the case with 300 loci randomly distributed across the genome. In the extreme,
some AFLP bands may be allelic or very tightly linked and so their levels of
differentiation will not be independent. This will be detectable in hybridizing
populations because it will generate strong disequilibrium between differentiated loci.

We are currently analyzing such populations. However, in the present analysis, any
effect of linkage would apply equally to all comparisons and so cannot explain the
difference in distribution of $F_{ST}$ between H-M and within-morph comparisons.
Thus, while there are no fixed differences between morphs in any of the populations we have investigated, in appropriate comparisons (H versus M populations), there is a small group of loci which show considerable differentiation against a background of a majority where differentiation is weak. We suggest that this is the most striking aspect of the data reported above: that there is a consistent group of loci apparently differentiated. This point is further supported by comparing trees in which the samples group by morphotype when the differentiated loci are included in the analysis, but by shore when they are excluded. From this we infer that the majority of the AFLP loci are in mutation/drift/dispersal equilibrium, although we cannot exclude the possibility of a general reduction in gene exchange between H and M populations relative to populations of the same morph. Against this background, we suggest that differentiation is being maintained for the small number of differentiated loci by selection on the loci themselves, or on closely linked loci. These findings are consistent with earlier work demonstrating morphological, ecological and behavioural differences between *L. saxatilis* H and M (Hull et al., 1996; Hull, 1998; Pickles & Grahame, 1999) but imply that the genetic differences underlying these characters involve only a small proportion of the genome. This is what would be expected in a case of non-allopatric speciation in progress. However, it could also be the result of differential introgression following secondary contact resulting in homogenization of allele frequencies at all loci except those under selection, or closely linked to loci under selection.

The H and M forms of *L. saxatilis* represent one of several cases of divergence in shell shape in this species. Similar variation is reported for shores in Sweden (Janson & Sundberg, 1983), where it is considered to be phenotypic. It has been shown that some allozyme loci are under selection, or linked to selected loci, in Swedish
populations (Johannesson et al., 1995a; Johannesson & Tatarenkov, 1997), although this has not been explicitly associated with shell form. On the Galician coast of Spain very different shell forms occur in populations between which there is some restriction of gene flow and evidence of selection on shell form (Johannesson et al., 1995b; Rolán-Alvarez et al., 1997). We do not have direct evidence of selection operating on H and M forms on the Yorkshire coast, but it seems reasonable to infer that it does. The findings from Britain and Spain suggest that a pervasive influence in habitat use and subsequent diversification in L. saxatilis is the vertical shore gradient. In turn, this suggests an unusually simple physical background (a spatially very restricted cline, limited by the extent of the intertidal zones occupied by the animals) against which to study speciation processes.

Whether the differentiation of the small proportion of loci between H and M is primary (the result of divergent selection) or secondary (the result of renewed contact), the main point is that differentiation is maintained for a small portion of the genome, while gene exchange continues to prevent divergence at the majority of loci. Detailed investigation of these loci in particular may provide important insights into the nature of the barrier between these two forms of intertidal snail, and into the evolution of barriers to gene exchange in general.
Acknowledgements

This work is supported by GR3/12528 from the NERC. We thank Paul Ashley for technical assistance, and Kerstin Johannesson and Richard Nichols for helpful discussions. We are grateful to two anonymous referees for their constructive criticisms.
Table 1. Adapters and selective primer sequences used for AFLP analysis.

<table>
<thead>
<tr>
<th>Primer/adapter</th>
<th>Adapters</th>
<th>5'–CTCGTAGACTGCGTACC–3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>EcoRI</td>
<td>3'–CATCTGACGCATGGTAA–5'</td>
<td></td>
</tr>
<tr>
<td>MseI</td>
<td>5'–GACGATGAGTCCTGAG–3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'–TACTCAGGACTCAT–5'</td>
<td></td>
</tr>
</tbody>
</table>

**Primers**

<table>
<thead>
<tr>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eco+1</td>
</tr>
<tr>
<td>Eco+A</td>
</tr>
<tr>
<td>Eco+C</td>
</tr>
<tr>
<td>Mse+1</td>
</tr>
<tr>
<td>Mse+A</td>
</tr>
<tr>
<td>Mse+C</td>
</tr>
<tr>
<td>Eco+3</td>
</tr>
<tr>
<td>Eco+ACG</td>
</tr>
<tr>
<td>Eco+CAG</td>
</tr>
<tr>
<td>Eco+CTC</td>
</tr>
<tr>
<td>Mse+3</td>
</tr>
<tr>
<td>Mse+AGT</td>
</tr>
<tr>
<td>Mse+ATC</td>
</tr>
<tr>
<td>Mse+CAA</td>
</tr>
<tr>
<td>Mse+CGA</td>
</tr>
</tbody>
</table>

**Combinations**

<table>
<thead>
<tr>
<th>A</th>
<th>Eco+CTC-Mse+CGA</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>Eco+CAG-Mse+CGA</td>
</tr>
<tr>
<td>D</td>
<td>Eco+CAG-Mse+CAA</td>
</tr>
<tr>
<td>E</td>
<td>Eco+ACG-Mse+ATC</td>
</tr>
<tr>
<td>F</td>
<td>Eco+ACG-Mse+AGT</td>
</tr>
</tbody>
</table>
Table 2. Levels of polymorphism of scored AFLP markers.

<table>
<thead>
<tr>
<th>Restriction site</th>
<th>A</th>
<th>B</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of variable bands</td>
<td>43</td>
<td>80</td>
<td>54</td>
<td>54</td>
<td>59</td>
<td>290</td>
</tr>
<tr>
<td>No. of fixed bands</td>
<td>0</td>
<td>3</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>43</td>
<td>83</td>
<td>62</td>
<td>55</td>
<td>63</td>
<td>306</td>
</tr>
</tbody>
</table>
Table 3. $F_{ST}$ (below diagonal) between populations of *Littorina saxatilis* (mean over 290 loci). Above diagonal, $F_{ST}$ following removal of 15 loci. Standard errors of $F_{ST}$ estimates range from 9.80-22.15% (below diagonal) and 10.19-22.20% (above diagonal) of the mean. TH, Thornwick Bay; OP, Old Peak; FY, Filey Brigg; RB, Robin Hood’s Bay (two samples, M only).

<table>
<thead>
<tr>
<th></th>
<th>TH-H</th>
<th>TH-M</th>
<th>OP-H</th>
<th>OP-M</th>
<th>FY-H</th>
<th>FY-M</th>
<th>RB1</th>
<th>RB2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TH-H</td>
<td>0.0190</td>
<td>0.0311</td>
<td>0.0372</td>
<td>0.0242</td>
<td>0.0227</td>
<td>0.0301</td>
<td>0.0298</td>
<td></td>
</tr>
<tr>
<td>TH-M</td>
<td>0.0378</td>
<td>0.0397</td>
<td>0.0293</td>
<td>0.0322</td>
<td>0.0217</td>
<td>0.0260</td>
<td>0.0255</td>
<td></td>
</tr>
<tr>
<td>OP-H</td>
<td>0.0318</td>
<td>0.0633</td>
<td>0.0204</td>
<td>0.0223</td>
<td>0.0256</td>
<td>0.0254</td>
<td>0.0247</td>
<td></td>
</tr>
<tr>
<td>OP-M</td>
<td>0.0489</td>
<td>0.0292</td>
<td>0.0396</td>
<td>0.0309</td>
<td>0.0241</td>
<td>0.0221</td>
<td>0.0212</td>
<td></td>
</tr>
<tr>
<td>FY-H</td>
<td>0.0247</td>
<td>0.0551</td>
<td>0.0238</td>
<td>0.0480</td>
<td>0.0095</td>
<td>0.0204</td>
<td>0.0207</td>
<td></td>
</tr>
<tr>
<td>FY-M</td>
<td>0.0328</td>
<td>0.0236</td>
<td>0.0402</td>
<td>0.0244</td>
<td>0.0247</td>
<td>0.0148</td>
<td>0.0142</td>
<td></td>
</tr>
<tr>
<td>RB1</td>
<td>0.0350</td>
<td>0.0308</td>
<td>0.0347</td>
<td>0.0241</td>
<td>0.0286</td>
<td>0.0156</td>
<td>0.0055</td>
<td></td>
</tr>
<tr>
<td>RB2</td>
<td>0.0340</td>
<td>0.0299</td>
<td>0.0339</td>
<td>0.0230</td>
<td>0.0291</td>
<td>0.0153</td>
<td>0.0052</td>
<td></td>
</tr>
</tbody>
</table>
References


Figure legends

Fig. 1

$F_{ST}$ values estimated from 290 variable AFLP loci plotted against mean allele frequency in three H-M comparisons (Thornwick Bay, Old Peak and Filey), one M-M comparison (Robin Hood’s Bay) and six representative between shores, within morph comparison. The solid line represents the 0.99 quantile estimated from a simulation model (see methods). The position of each of the 15 loci implicated as linked to a region under selection due to their presence outside the region defined by the 0.99 quantile consistently in all three H-M comparisons (see Results) is labeled with the locus identifier (where A-F = primer combination, see Table 1).

Fig. 2

Relationship of mean $F_{ST}$ with linear distance, 15 differentiating loci removed.

Fig. 3.

Neighbour-Joining tree calculated from allele frequency data in PHYLIP (Felsenstein, 1993). Data were bootstrapped (x100) using SEQBOOT and Nei’s genetic distance calculated using GENDIST. Distances were clustered with NEIGHBOR and the consensus tree constructed with CONSENSE. Bootstrap values are shown at the nodes of a representative non-consensus tree in order to retain branch length information. a), using all 290 AFLP loci; b), after removal of 15 differentiated loci.
Fig. 3.

(a)