Wilding, CS, Grahame, J and Mill, PJ

 Nuclear DNA restriction site polymorphisms and the phylogeny and population structure of an intertidal snail species complex (Littorina)

http://researchonline.ljmu.ac.uk/id/eprint/2711/

Article

For more information please contact researchonline@ljmu.ac.uk

Nuclear DNA restriction site polymorphisms and the phylogeny and population structure of an intertidal snail species complex (*Littorina*)

CRAIG S. WILDING, JOHN GRAHAME and PETER J. MILL

The School of Biology, University of Leeds, Leeds, U.K.

Primers for amplification of four novel, unlinked nuclear DNA loci, the first reported for the rough periwinkles of the genus *Littorina*, are described. Patterns of restriction site polymorphism for these loci are detailed within the rough periwinkles. RFLPs are not found to be diagnostic for any of the currently accepted species within this group, nor for any of the contentious subspecies, or forms, whose taxonomic status is uncertain. However, there are important differences in allele frequencies between these taxa and certain of these mirror differences detected in a previous study of the mitochondrial DNA. These allele frequency data are used to construct a phylogeny in which groupings of the three recognised species are obvious when either Nei's genetic distances or Reynold's distances are clustered. Contentious forms (*L. neglecta*, *L. saxatilis* 'b' and *L. tenebrosa*) do not cluster as distinct taxa, although populations of *L. neglecta* have important allele frequency differences from *L. saxatilis*. These four loci have confirmed the consensus view of *Littorina* phylogeny and provided important information on population structure-however four loci is insufficient for reaching definitive conclusions. Since analysis of nuclear DNA polymorphisms such as these is invaluable for analysis of phylogeny, population structure and phylogeography, identification of additional loci is considered imperative.

Craig S. Wilding, The School of Biology, University of Leeds, Leeds, LS2 9JT, U.K. E-mail: bgycsw@leeds.ac.uk

Reliable estimates of genetic differentiation between populations or closely related species require analysis of multiple, independent loci. As a consequence of the wide availability of PCR primers and the high variability of the molecule, mitochondrial DNA (mtDNA) has become the mainstay of much population genetic and phylogenetic research (AVISE 1994). However, since all genes on this molecule are linked, even analysis of multiple mitochondrial loci will possibly not yield data typical of the genome as a whole. In order to assay genetic differentiation more effectively it is crucial that additional unlinked (nuclear) loci are analysed. For the majority of species there is little sequence information available on any nuclear DNA loci; thus the design of specific primer pairs is not possible. Some degenerate primers allow a way into the genome, for instance through amplification of nuclear DNA introns utilising primers placed in conserved exon regions (CÔRTE-REAL et al. 1994), but the alternative is to design primers from sequences obtained from anonymous DNA regions generated from, for instance, RAPD (XU et al. 1996).

Snails belonging to the rough periwinkles form a group of closely related species. The current consensus is that three species form this group: *Littorina arcana* Hannaford Ellis, *L. compressa* Jeffreys and *L. saxatilis* (Olivi), although *L. saxatilis* is known to be extremely variable, and ovoviviparous animals purported to be separate species by some are argued to be ecological variants of *L. saxatilis* by others

(REID 1993; REID 1996). On the basis of allozyme analyses (e.g. WARD 1990) *L. compressa* is argued to be basal in the phylogeny, and *L. saxatilis* and *L. arcana* to be derived species. However, this is not supported in all phylogenies (BACKELJAU and WARMOES 1992). The uncertainty surrounding this phylogeny is probably attributable to a burst of speciation, and our own studies of mtDNA suggest speciation in this group was indeed rapid (WILDING et al. 2000). Such rapid speciation results in an inconclusive tree topology of mtDNA haplotypes with low bootstrap support at some nodes. Nevertheless, the overall consensus phylogeny from mtDNA data (*L. compressa*, (*L. arcana*, *L. saxatilis*)) fits the accepted view. However, the existence, in some instances, of identical haplotypes in different species, suggests that these shared haplotypes predate the speciation event. Since mtDNA variation may not be representative of the true species phylogeny (DEGNAN 1993), additional loci are required to examine relationships of these species. Although RAPD (CROSSLAND et al. 1996) studies have been brought to bear upon the study of phylogeny in these species, there have been no attempts to use neutral, co-dominant loci. Such loci may be critical for the examination of species and population relationships in cases such as this where mtDNA haplotypes are non-diagnostic due to shared ancestral polymorphism.

Here restriction fragment length polymorphism (RFLP) is examined at four purportedly neutral loci

to investigate firstly whether such data provide insights into the controversial phylogeny of the rough periwinkles, and secondly if they provide information on population structure and gene flow within these species.

MATERIALS AND METHODS

Sample collection

Specimens of *Littorina* (Total number = 789; N = 5–72 per population) were collected from 32 locations around the UK and Republic of Ireland, with an additional sample from Ursholmen in the Koster Archipelago, Sweden (Table 1). Representatives of each of the rough periwinkle species were sampled where possible. However *L. arcana* and *L. compressa* are patchily distributed and missing from the coastline of the English Channel east of Lyme Bay (MILL and GRAHAME 1992) and from some other locations. These sampling sites do not represent the full range of these species. However, much of the range south of North Berwick on the east coast of Britain and south of the Lleyn peninsula (Wales) has been sampled particularly intensively. Additional samples of the contentious taxa *L. neglecta* (GRAHAME et al. 1995; JOHANNESSON and JOHANNESSON 1990), *L. saxatilis* 'b' (CALEY et al. 1995) and *L. tenebrosa* (BARNES 1993; JANSON and WARD 1985) were made at suitable locations. Although the taxonomic status of these is still uncertain, for clarity they are referred to here with the species names and not as *L. saxatilis* subspecies, ecotypes or other varieties.

DNA extraction

Extraction of genomic DNA followed the protocol (number 47) of ASHBURNER (1989). DNA samples were resuspended in 1/10 TE and adjusted to 10 ng.μl^{-1} concentration.

Choice of loci studied

Four polymorphic loci were examined. Two were intron 3 (CÔRTE-REAL et al. 1994) of separate calmodulin genes. Sequences of the introns and flanking exons of these genes in *Littorina* will be described elsewhere (WILDING et al. unpublished). The primers CAD1F (TGC ACA TCA TCA TGC CAA A) and CAD1RB (TCT TCA GAG CAG GGT TCC ATT) were used to amplify the third intron of what has been called CAL-1 (WILDING et al. unpublished data), and CAD2F (CTG CAG ATG GTG ACG CAA) and CAD2R (CTG ACG GTG AGT GAC AAT CG) were used to amplify the third intron of CAL-2. Two other polymorphic loci were identified from cloning and sequencing of RAPD bands (GRA-

HAME et al. 1997). One of these (X-80) was identified from sequences generated from RAPD-PCR using primer RAPD-X (≡ OPY-01, GTG GCA TCT C). Primers X80a (CAT CTC TGT GTT GAA AGA GGG T) and X80b (CAG AAC TAA ACT GAA GAA ACC CG) were derived from this sequence and shown to amplify a length-invariant band. Primers for the DELETION locus (RAPDRPT3 [AGT AAC GGC AGA CGC CAT] and RAPDRPT4 [TTA TTC TTG TCT TCT CTG CC]) were based on sequence neighbouring an extremely variable microsatellite repeat identified from a screen using primer RAPD-H

Table 1. *Grid references of collection sites for Littorina samples used in the present study. A = Littorina arcana. B = L. saxatilis 'b'. C = L. compressa. N = L. neglecta. S = L. saxatilis. T = L tenebrosa*

Location	Grid reference	Species
Ursholmen, Sweden	58'5"10°N 0'59"4°E	S
North Berwick	OS: NT555857	A, C, N, S
St Abb's	OS: NT907692	A, C, S
Old Peak*	OS: NZ984021	A, B, N, S
Snettisham	OS: TF649319	S
Wells next the sea	OS: TF909458	
	OS: TF915456	S
Holkham	OS: TF886451	T
Cley	OS: TG062448	
	OS: TG067447	S
Alderton	OS: TM363419	S
St Margaret's at Cliffe	OS: TR368444	S
Folkestone	OS: TR245369	
	OS: TR244373	S
St Alban's Head	OS: SY959754	S
Portland Bill	OS: SY675683	S
East Fleet	OS: SY799635	S
The Fleet (gravel)	OS: SY758665	S
Golden Cap	OS: SY407918	S
Pinhay Bay	OS: SY318907	S
Cargreen	OS: SX436627	S
The Lizard	OS: SW699114	A, C, S
Trevaunance	OS: SW725519	A, C, S
The Mumbles	OS: SS632873	A, S
The Gann flats	OS: SM812069	S
Musselwick (Gann)	OS: SM819065	C
St Ann's Head	OS: SM809028	S
Westdale Bay	OS: SM797056	A, C, N, S
Porth Ysgo	OS: SH207265	S
Porth Llanllawen	OS: SH266267	A, C, S
Trwyn Maen Melyn	OS: SH138251	S
Nynian's Cove, Galloway	OS: NW417364	S
Inismór, Aran Islands	IGR: L 103221	C, S
Ballynahown	IGR: L 992202	A, C, S
Golam Head	IGR: L 826214	S, T

* Peak Steel (collection site for *L. neglecta* and *L. saxatilis* 'b') is a flat rocky outcrop at Old Peak.

(GCC GTG GTT A). Regardless of target, primer pairs were employed in 25 μl PCR reactions containing 50 mM KCl, 10 mM Tris-HCl pH 9.0, 1.5 mM MgCl₂, 0.1% Triton X-100, 0.01% gelatin, 200 μM each dNTP, 12.5 pmol each primer, 12.5 ng DNA and 0.5U *Taq* (Super*taq*, HT Biotechnologies). PCR cycling conditions were 1 × 94°C, 5 min; 35 × (94°C, 1 min, 55°C, 30 secs, 72°C, 1 min); 1 × 72°, 5 min, with the exception of CAL-1 in which the annealing temperature was 57°C, and DELETION in which annealing was at 57°C and the extension time was reduced to 30 seconds.

Identification of alleles

A number of restriction enzymes were initially employed to screen for variation in the amplification products from CAL-1, CAL-2 and X-80. Polymorphisms in CAL-1 and CAL-2 were uncovered through digestion with *Taq*I and *Dde*I respectively and in X-80 with *Msp*I. Restriction digestion was performed at the appropriate temperature for 2 hours and products were separated on 2% agarose gels. For the deletion PCR, alleles differed due to length variation and this was recognised after PCR products were run on 3% agarose gels.

Analysis of results

Allele frequencies and genetic distances were generated in BIOSYS v1.7 (SWOFFORD and SELANDER 1989) and phylogenetic analysis performed in PHYLIP (FELSENSTEIN 1993). Nei's distance (NEI 1987) can be efficient for use in phylogeny construction from gene frequencies, but does have some failings when used for the study of phylogeny of closely related species. Thus other statistics are recommended for such analyses (HILLIS 1998). Both Nei's genetic distance and Reynold's distance were therefore calculated after bootstrapping and clustered using Neighbour-Joining within the PHYLIP package using the SEQBOOT, GENDIST (with Jumble option) and NEIGHBOR packages. Additionally, multi-dimensional scaling (MDS) of Prevosti distances was performed in NTSYS (the Prevosti distance produced essentially identical results when clustered to both Nei's and Reynold's distances). Twenty repetitions were performed and the run with the lowest iterative stress used for plotting. Gene diversity statistics (G_{ST}, H_S and H_T) were also generated in DISPAN and significance calculated according to WAPLES (1987). Exact tests for conformation to Hardy-Weinberg equilibrium and analysis of linkage disequilibrium were computed in GENEPOP v3.1c (RAYMOND and ROUSSET 1995).

RESULTS

Three to four alleles were uncovered per locus based on either length variation or restriction digest pattern. Allele frequencies are given in Table 2 and variation at individual loci is described below.

CAL-1

Three alleles were identified at this locus. Alleles A and B were uncovered from sequences of this intron and result from the disruption of a *Taq*I site by the absence of an indel. Both these alleles are found in all three species, although allele A is at higher frequency in *L. compressa* and allele B at higher frequency in *L. saxatilis*. Allele C was an additional variant uncovered during restriction digest screening, the exact sequence basis of which has not yet been investigated. It has only been uncovered in three of nine populations of *L. arcana* investigated and not in any other species.

CAL-2

CAL-2 is the least variable of the four studied loci although four alleles were identified. *L. saxatilis* harbours all four alleles across a range of populations and is the only species in which allele A is seen. *L. arcana* is virtually monomorphic for allele B, with the exception of the Mumbles population in which one animal was a BD heterozygote. Allele D is seen at the highest frequency in *L. compressa* along with allele B. For *L. saxatilis*, all populations have alleles A and B at varying frequencies, however allele A is seen only in certain populations on the South Coast (St Alban's, Portland Bill, Fleet gravel, Golden Cap, Pinhay Bay, The Lizard) and also at North Berwick and Old Peak on the East Coast and, at low frequency, at the Gann flats. Allele A is also seen in the small, barnacle dwelling animals (*L. neglecta* and *L. saxatilis* 'b') from Peak Steel and in the *L. neglecta* from Westdale Bay. It has not been seen in *L. neglecta* from North Berwick.

X-80

All three alleles at locus X-80 are found in *L. saxatilis* and *L. arcana* but allele C is absent from *L. compressa*. Allele C in *L. saxatilis*, A in *L. arcana* and B in *L. compressa* are all at low frequency.

DELETION

Alleles A and B differ in length due to an indel event in the sequence flanking a GTT microsatellite repeat. Alleles C and D also differ through length polymorphism but are much rarer. They are not seen in *L. arcana* and are absent or only at low frequency in *L. saxatilis* and *L. compressa*. Some differences in the

Table 2. *Allele frequencies at four loci in various populations of Littorina. Populations are organised clockwise around the British coastline from St Margarets at Cliffe (south-east) to Alderton (east)*

L. saxatilis

		St Margarets at Cliffe	Folkestone	St Albans	Portland Bill	East Fleet	Fleet Gravel	Golden Cap	Pinhay Bay	Cargreen	Lizard	Trevaunance	Mumbles	Gann	St Ann's Head	Westdale Bay	Porth Ysgo	Trwyn Maen Melyn	Porth Llanllawen	Galloway	N. Berwick
CAL-1	A	0.5	0.158	0.6	0.333	0.25	0.4	0.325	0.8	0.85	0.25	0.538	0.4	0.55	0.35	0.2	0.325	0.15	0.265	0.528	0.3
	B	0.5	0.842	0.4	0.667	0.75	0.6	0.675	0.2	0.15	0.75	0.462	0.6	0.45	0.65	0.8	0.675	0.85	0.735	0.472	0.7
	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CAL-2	A	0	0	0.4	0.056	0	0.1	0.5	0.8	0	0.1	0	0	0.05	0	0	0	0	0	0	0
	B	1	1	0.6	0.944	1	0.9	0.5	0.2	0.95	0.9	1	0.925	0.925	0.9	0.95	1	0.95	0.971	0.861	1
	C	0	0	0	0	0	0	0	0	0.05	0	0	0.075	0.025	0.087	0	0	0.05	0.029	0.139	0
	D	0	0	0	0	0	0	0	0	0	0	0	0	0	0.013	0.05	0	0	0	0	0
X-80	A	0.15	0.026	0.15	0	0.1	0	0.275	0	0.1	0.55	0.308	0	0.05	0.112	0.1	0.15	0.65	0.382	0.222	0.5
	B	0.85	0.974	0.85	1	0.9	1	0.725	1	0.9	0.45	0.654	0.975	0.95	0.85	0.9	0.85	0.35	0.618	0.778	0.5
	C	0	0	0	0	0	0	0	0	0	0	0.038	0.025	0	0.038	0	0	0	0	0	0
Deletion	A	0.45	0.237	0.35	0.389	0.45	0.3	0.1	0.4	0.25	0	0.038	0.3	0.15	0.338	0.2	0.125	0	0.294	0.361	0.1
	B	0.55	0.763	0.65	0.611	0.55	0.7	0.9	0.6	0.75	1	0.962	0.7	0.85	0.663	0.8	0.875	1	0.706	0.583	0.7
	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.056	0.2
	D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(N)	20	19	10	9	10	10	20	5	10	10	13	20	20	40	10	20	10	17	18	10
	Het	0.3205	0.1742	0.4395	0.2714	0.2763	0.2842	0.3891	0.3444	0.2382	0.2763	0.272	0.2788	0.2529	0.341	0.2408	0.234	0.2118	0.3434	0.414	0.3632
	Het S.E.	0.1219	0.0885	0.0574	0.1266	0.1147	0.1167	0.0714	0.0548	0.0625	0.1147	0.1361	0.1079	0.0917	0.0687	0.0584	0.0923	0.1049	0.0968	0.0693	0.1223

		L. saxatilis (cont.)										*L. arcana*								
		St Abb's	Old Peak	Snettisham	Wells	Cley	Alderton	Inismór	Ballynahown	Golam Head	Sweden	Lizard	Trevaunance	Mumbles	Westdale Bay	Porth Llanllawen	N. Berwick	St Abb's	Old Peak	Ballynahown
CAL-1	A	0.4	0.375	0.35	0.45	0.31	0.35	0.2	0.375	0.15	0.324	0.1	0.2	0.6	0.45	0.2	0.15	0.15	0	0.563
	B	0.6	0.625	0.65	0.55	0.69	0.65	0.8	0.625	0.85	0.676	0.85	0.55	0.4	0.5	0.8	0.85	0.85	1	0.438
	C	0	0	0	0	0	0	0	0	0	0	0.05	0.25	0	0.05	0	0	0	0	0
CAL-2	A	0.05	0.075	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	B	0.95	0.925	1	1	0.988	1	0.8	1	0.95	1	1	1	0.9	1	1	1	1	1	1
	C	0	0	0	0	0.012	0	0.2	0	0.05	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0	0	0	0	0
X-80	A	0.05	0.2	0.1	0.225	0.012	0.1	0.222	0.375	0.35	0.147	0	0	0	0	0	0	0	0.067	0.063
	B	0.95	0.8	0.9	0.775	0.988	0.9	0.778	0.625	0.65	0.853	0.35	0.25	0.8	0.65	0.1	0.65	0.9	0.933	0.5
	C	0	0	0	0	0	0	0	0	0	0	0.65	0.75	0.2	0.35	0.9	0.35	0.1	0	0.438
Deletion	A	0.55	0.25	0.15	0.5	0.476	0.4	0.3	0.15	0.4	0.353	0.7	0.6	0.6	0.8	0.9	0.85	0.35	0.633	0.375
	B	0.45	0.625	0.85	0.5	0.524	0.6	0.7	0.85	0.6	0.647	0.3	0.4	0.4	0.2	0.1	0.15	0.65	0.367	0.625
	C	0	0.125	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	(N)	10	20	10	20	42	10	10†	20	10	17	10	10	5	10	5	10	10	15	8
	Het	0.3066	0.374	0.2342	0.3446	0.2465	0.2934	0.3704	0.3058	0.3382	0.2951	0.2737	0.3816	0.4056	0.3474	0.1889	0.2539	0.2342	0.1525	0.4038
	Het S.E.	0.1193	0.0896	0.0991	0.1204	0.1292	0.1212	0.0249	0.1143	0.0955	0.1094	0.1004	0.1357	0.1357	0.1256	0.0803	0.0729	0.0981	0.0991	0.136

		L. compressa									*L. neglecta*			*L. saxatilis* 'b'	*L. tenebrosa*	
		Lizard	Trevaunance	Gann	Westdale Bay	Porth Llanllawen	N. Berwick	St Abb's	Inismór	Ballynahown	Westdale Bay	N. Berwick	Peak Steel	Peak Steel	Golam Head	Holkham
CAL-1	A	0.7	0.227	0.583	0.55	0.6	0.05	0.15	0.9	0.7	0.472	0.143	0.326	0.367	0.425	0.4
	B	0.3	0.773	0.417	0.45	0.4	0.95	0.85	0.1	0.3	0.528	0.857	0.674	0.633	0.575	0.6
	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
CAL-2	A	0	0	0	0	0	0	0	0	0	0.028	0	0.451	0.417	0	0
	B	1	1	0.667	0.65	1	0.45	0.55	1	1	0.972	1	0.549	0.583	0.9	1
	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0
	D	0	0	0.333	0.35	0	0.55	0.45	0	0	0	0	0	0	0	0
X-80	A	1	1	1	1	1	1	0.9	1	0.9	0.694	0.929	0.458	0.55	0.275	0.2
	B	0	0	0	0	0	0	0.1	0	0.1	0.306	0.071	0.542	0.45	0.725	0.8
	C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Deletion	A	0.95	0.5	0.5	0.4	0.45	1	0.45	0.35	0.25	0.361	0.214	0.09	0.017	0.275	0.15
	B	0.05	0.227	0.417	0.5	0.05	0	0.55	0.6	0.7	0.639	0.786	0.507	0.7	0.725	0.85
	C	0	0	0	0	0	0	0	0	0	0	0	0.403	0.283	0	0
	D	0	0.273	0.083	0.1	0.5	0	0	0.05	0.05	0	0	0	0	0	0
	(N)	10	11	6	10	10	10	10	10	10	18	7	72	30	20	10
	Het	0.1355	0.2553	0.409	0.4026	0.2698	0.1553	0.375	0.1829	0.275	0.375	0.1921	0.5044	0.4768	0.376	0.2776
	Het S.E.	0.1049	0.1586	0.1392	0.137	0.1564	0.1242	0.0859	0.1278	0.1112	0.107	0.0783	0.0275	0.0148	0.0674	0.1051

N = number of alleles sampled (2 × number of individuals); Het. = heterozygosity; Het. S.E. = standard error of calculated heterozygosity.
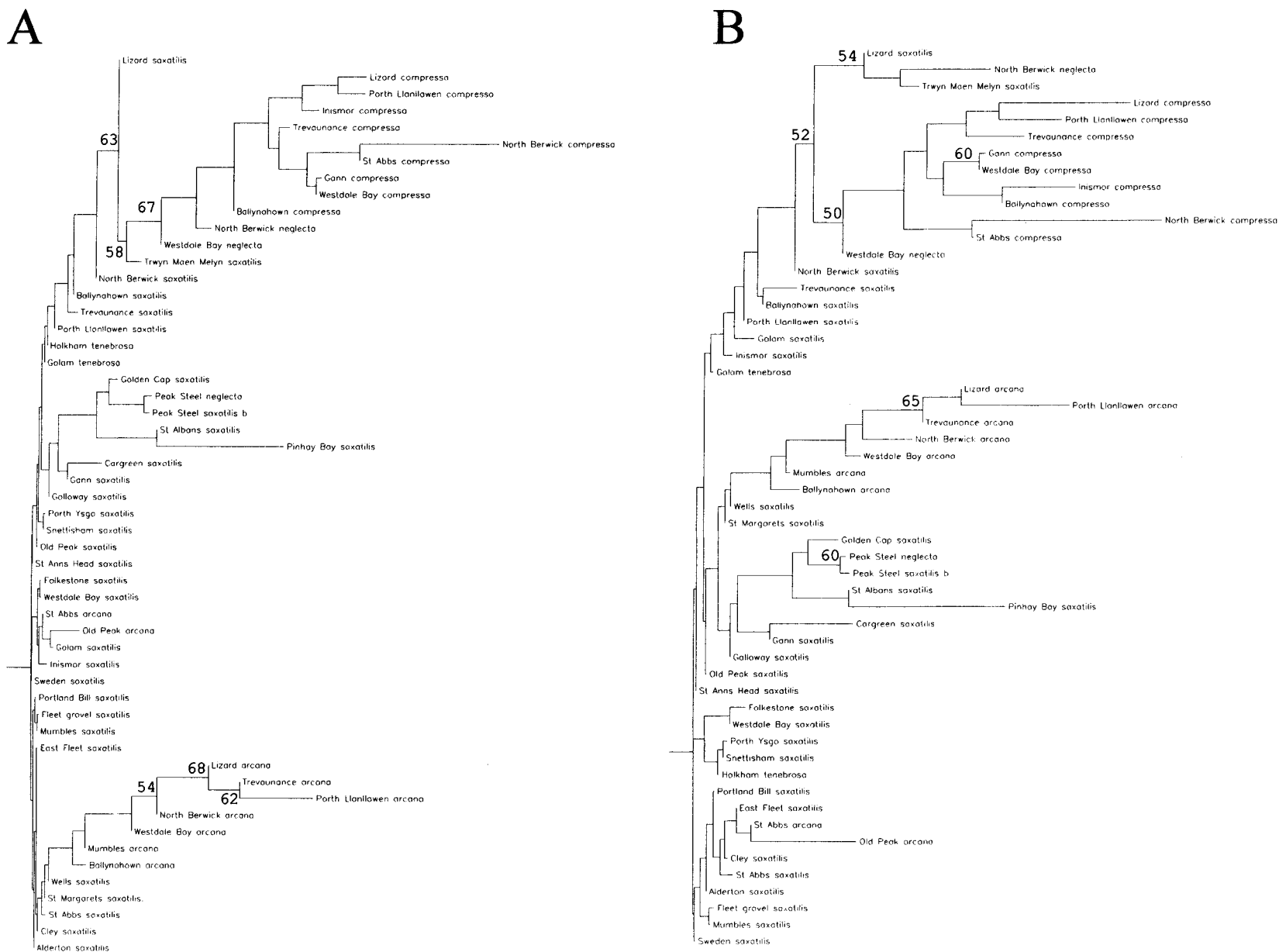
Fig. 1A and B. **A** Neighbour-joining tree from Nei's genetic distance, implemented in PHYLIP (FELSENSTEIN 1993). **B** Neighbour-joining tree from Reynold's distances implemented in PHYLIP. Bootstrap values over 50% are indicated.
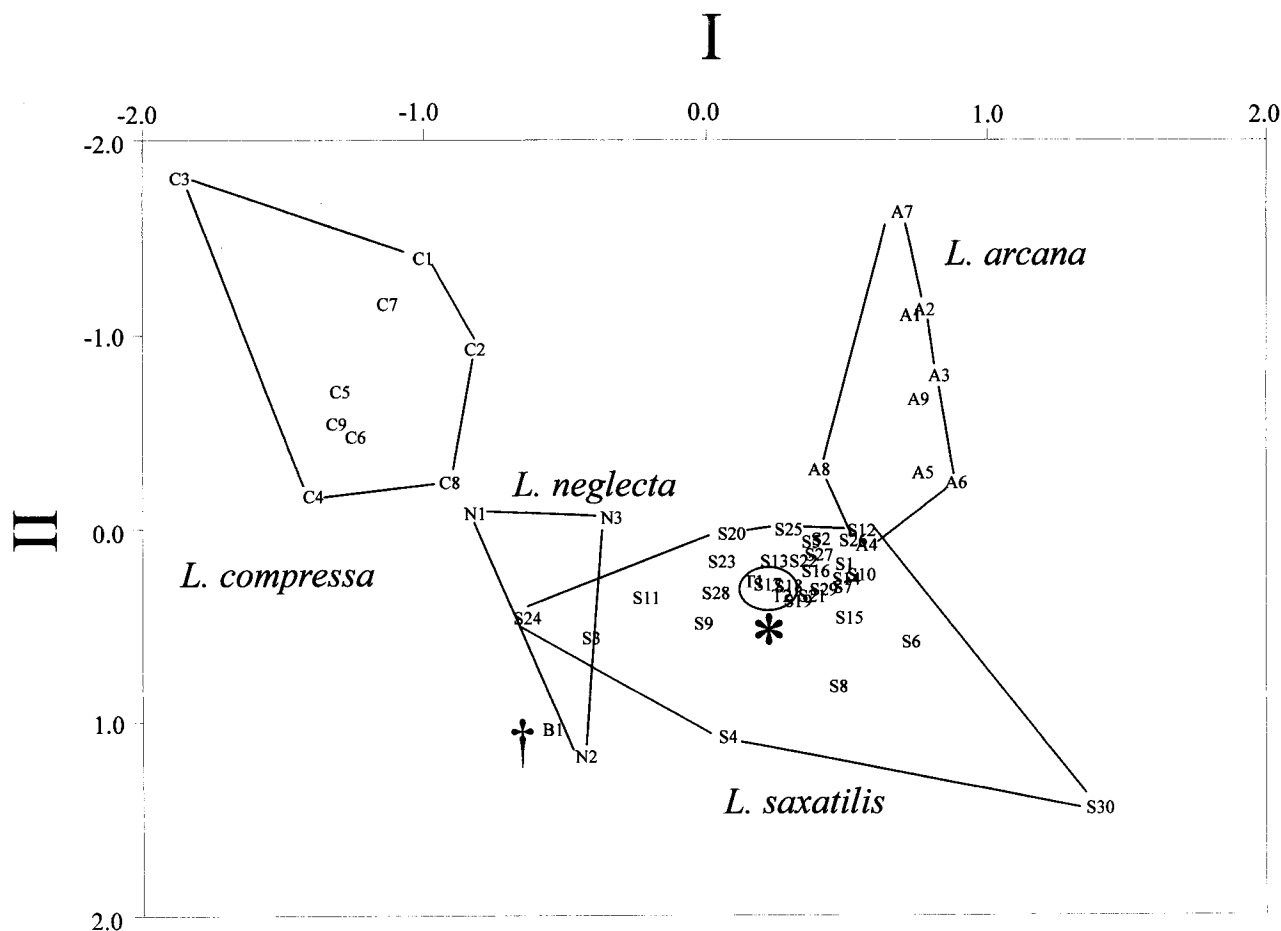
I

Fig. 2. MDS plot of Prevosti distances constructed in NTSYS. Borders are drawn around populations of *L. arcana*, *L. compressa*, *L. neglecta* and *L. saxatilis*. * Marks ellipsoid drawn around *L. tenebrosa* populations. † = *L. saxatilis* 'b'.

strength of amplification between putative alleles were noted. Where this did occur repeat amplifications from the DNA isolates were always found to be identical.

Phylogenetic analysis

Clustering of populations using both Nei's and Reynold's distance calculated from allele frequencies at these four loci reveals clusters which are made up almost exclusively of populations of single species (Fig. 1). When the three recognised species are considered (*L. arcana*, *L. compressa* and *L. saxatilis*) only two populations appear to be placed 'incorrectly'-the Old Peak and St Abb's populations of *L. arcana*, which cluster with *L. saxatilis* and not their conspecifics. When MDS is used on the Prevosti distances (Fig. 2), it is only the St Abb's population which overlaps with the '*saxatilis* cluster' (stress value for MDS = 0.27175). The status of *L. neglecta* is unclear from these data. From the dendrograms (Fig. 1) two populations of *L. neglecta* cluster with *L. compressa*, and one within *L. saxatilis*. However, in the MDS

plot all three *L. neglecta* populations cluster closely together and at the extreme of the *L. saxatilis* cluster, indicating some degree of differentiation. The *L. saxatilis* 'b' population are inseparable from the sympatric Peak Steel *L. neglecta* suggesting no differences are detectable between these forms at this site. Both forms however, do appear different from large *L. saxatilis* at this site. *L. tenebrosa* plot within *L. saxatilis* using MDS and cluster with *L. saxatilis* in the dendrograms, indicating that *L. tenebrosa* is simply an ecomorph of *L. saxatilis*.

Population substructure

These loci are shown to differ in their inherent level of variation and the differences between populations that they reveal as quantified by locus level G_{ST} statistics (Table 3). To some extent this is likely to be a result of the small sample sizes for some populations which may result in allele frequencies that are not reflective of the true values. However it is also the case that, in those situations where sample sizes are reasonable (N = 20 or above), there appear to be

Table 3. *Gene diversity statistics for samples of Littorina*

Locus	G_{ST}	χ^2	d.f.	H_T	H_S
All populations					
CAL-1	0.1685	531.79	106	0.4777	0.3972
CAL-2	0.3578	1692.87	159	0.1951	0.1253
X-80	0.5314	1654.80	106	0.5255	0.2463
Deletion	0.2410	1168.35	159	0.5146	0.3905
All loci	0.3232	6626.02	689	0.4282	0.2898
L. saxatilis					
CAL-1	0.1182	222.24	58	0.4705	0.4149
CAL-2	0.3519	992.23	87	0.1790	0.1160
X-80	0.1840	345.92	58	0.3093	0.2524
Deletion	0.1076	303.43	87	0.4200	0.3748
All loci	0.1601	1956.42	377	0.3447	0.2895
L. arcana					
CAL-1	0.2016	66.93	16	0.4462	0.3562
CAL-2	0.0899	44.77	24	0.0220	0.0200
X-80	0.3140	104.25	16	0.5020	0.3444
Deletion	0.1448	72.11	24	0.4578	0.3915
All loci	0.2212	477.35	104	0.3570	0.2780
L. compressa					
CAL-1	0.2928	101.89	16	0.5000	0.3536
CAL-2	0.3096	161.61	24	0.3041	0.2099
X-80	0.0795	27.67	16	0.0435	0.0400
Deletion	0.2620	136.76	24	0.5824	0.4298
All loci	0.2774	627.48	104	0.3575	0.2583

G_{ST} = coefficient of gene differentiation; H_T = total gene diversity; H_S = gene diversity within species.

differences in allele frequencies between populations. Inspection of the allele frequencies (Table 2) indicates that, within a species, there is little pattern to the differences between populations. However, some particular features are evident such as the higher frequency of CAL-2[A] in *L. saxatilis* populations on the South central-west coast (from St Alban's to the Lizard – see Table 2), only occurring elsewhere at St Abb's, Old Peak and the Gann. There are also differences between barnacle dwelling animals (*L. neglecta* and *L. saxatilis* 'b') at Peak Steel (a flat rocky outcrop at Old Peak) from large *L. saxatilis* at this site, particularly at the X-80 and DELETION loci. Here, large *L. saxatilis* have a much higher frequency of X-80[B] and a lower frequency of the DELETION[C] than in the barnacle dwelling animals. This seems not to be due to low sample sizes since these samples were very large. However, this does not seem to be typical of *L. neglecta*, since neither the Westdale Bay nor North Berwick populations show such differences from samples of large saxatilis taken at those sites.

Conformation to Hardy-Weinberg equilibrium

Most populations conform to Hardy-Weinberg equilibrium with limited exceptions. After adjustment of the table-wide significance level ($\alpha = 0.05$), using the sequential Bonferroni method (RICE 1989), only the

Trevaunance *L. saxatilis* at the X-80 locus, and the St Ann's Head *L. saxatilis* and Peak Steel *L. neglecta*, both at the DELETION locus, depart from Hardy-Weinberg equilibrium (all values are not reproduced but can be communicated upon request).

Linkage disequilibrium

There was no evidence of linkage in the dataset. No population showed evidence of linkage between loci, and over all populations this was also the case (Table 4). Although linkage is highly unlikely (MAYNARD SMITH 1989), since CAL-1 and CAL-2 represent introns of duplicated genes (WILDING et al. unpublished data) there is the potential for linkage to be present.

Table 4. *Genotypic disequilibrium analysis for four loci in Littorina across all populations (Fisher's method)*

Locus pair	χ^2	d.f.	P-value
CAL-1 & CAL-2	43.506	56	0.888
CAL-1 & X-80	54.840	86	0.996
CAL-2 & X-80	26.242	44	0.985
CAL-1 & DELETION	56.172	100	0.100
CAL-2 & DELETION	20.634	50	0.100
X-80 & DELETION	86.407	84	0.407

DISCUSSION

The four loci studied represent the first presumed neutral, co-dominant, nuclear DNA polymorphisms reported for the rough periwinkles. Using this restriction analysis technique, substantial polymorphism has been uncovered at these four loci and the data suggest that analysis of such nDNA-RFLPs can provide important information for the study of phylogeny and population genetics in these species.

As is the case for previous studies at allozyme loci (WARD 1990; BACKELJAU and WARMOES 1992) there are no loci diagnostic for any of the studied species. However, certain alleles are limited to particular species and, overall, allele frequencies appear quite different when species are compared. Despite this, many alleles are shared between the species. Such sharing of alleles between species is a phenomenon often found when recently speciated taxa are examined, e.g. for the *Drosophila melanogaster* group (COYNE and KREITMAN 1986) and for Lake Victoria cichlid fish (NAGL et al. 1998), and is indicative of shared ancestral polymorphisms which predate speciation. Theoretical studies have shown that $4N_e$ (effective population size) generations are required before ancestral variation is sorted, through genetic drift, to a state of reciprocal monophyly. Thus, unless populations are small, ancestral variation will persist in daughter species for some considerable time (COYNE and KREITMAN 1986). NAGL et al. (1998) have demonstrated how extensive nuclear polymorphisms have persisted in species of cichlids that have had 12,000 years available for diversification and lineage sorting. These *Littorina* have probably had much longer separation times, yet still show ancestral alleles.

This sharing of alleles between species could also have resulted from episodes of hybridisation rather than sharing of ancestral alleles. In a parallel study of mtDNA polymorphisms WILDING et al. (2000) argue that hybridisation was less likely to be the cause of shared mtDNA haplotypes in *Littorina* than incomplete lineage sorting of ancestral variation, and this seems also to be true for these nDNA data. Hybridisation should result in allele sharing irrespective of locus. Thus alleles should not be restricted to certain species if they all do occasionally interbreed. However this is not the case; for example CAL-1^C is found only in *L. arcana*, and X-80^C is common in *L. arcana*, but absent in *L. compressa* and found only very rarely in *L. saxatilis*. Lineage sorting, however, will vary between loci since it is a stochastic process and heavily dependent upon the variability of each locus in the ancestral lineages.

In spite of the observed shared polymorphisms, allele frequencies do vary considerably between spe-

cies, and this is reflected in the population phylogenies and MDS plot based on both distance values. This phylogeny is, to some extent, congruent with current opinion in suggesting that *L. arcana* and *L. saxatilis* are more closely related to each other than to *L. compressa*. Despite these results, the use of only four loci is insufficient for the production of a truly robust population phylogeny. Thus additional loci are needed, and the identification of additional polymorphic loci and suitable primers for their amplification is imperative. The contentious taxa *L. tenebrosa*, *L. neglecta* and *L. saxatilis* 'b' samples do not form monophyletic groups in these phylogenies. Because of the ovoviviparous nature of these taxa, it seems likely that they would be most closely related to *L. saxatilis* whatever their taxonomic status since ovoviviparity is a derived state in these animals (REID, 1996). With the number of loci available, the ability to resolve the branching order of the three recognised species is limited, thus resolution at a finer scale than this (the status of *L. tenebrosa*, *L. neglecta* and *L. saxatilis* 'b' with respect to *L. saxatilis*) is unlikely until further polymorphic loci are available. However, frequency differences at some loci between *L. saxatilis* and the barnacle dwelling *L. saxatilis* 'b' and *L. neglecta* suggests that gene flow between them is low (see below).

Within the recognised species at the population level, the clustering structure in these phylogenies does not seem to have a geographic basis. However, there is substantial intra-specific variability, as attested by the significant G_{ST} estimates, and in at least one case (for CAL-2), single-locus allele frequency differences seem to have a geographic basis. CAL-2^A is only found on the south central-west coast of Britain and at Old Peak and St Abb's. The distribution of this allele on the south central-west coast corresponds to the distribution of a mtDNA haplotype which is also only ever seen in this region (WILDING et al. 2000). Correlation between mtDNA and nDNA data is also found in comparisons of the large *L. saxatilis* at Old Peak with small animals (*L. neglecta* and *L. saxatilis* 'b') found in the barnacle zone at Peak Steel, a flat rocky outcrop at Old Peak. Haplotype frequencies for the cytochrome-*b* locus show significant differences between these samples (WILDING et al. unpublished data), suggesting that there is little gene flow between them. In this study, substantial differences are also seen between these taxa at the CAL-2, X-80 and DELETION loci, and these involve the largest sample sizes. This is strong evidence that nDNA loci such as these are capable of providing vital information on population structure and gene flow.

Populations on the whole did conform to Hardy-Weinberg predictions. In the cases for which this is not so, it is not known whether this is a consequence of

processes acting at the population level, or due to factors associated with the PCR, such as inefficient amplification of certain alleles leading to mis-scoring of heterozygotes (HARE et al. 1996). However, the observation that the larger (although often non-significant after Bonferonni correction) deviations from equilibrium occur at the deletion locus for which amplification was often the weakest, suggests that the latter possibility has had at least some effect.

Although it has been demonstrated here that this procedure (analysis of nDNA PCR restriction site polymorphisms) can be applied successfully, identification of additional loci and increase in sample sizes will enhance the value of these polymorphic loci for the analysis of *Littorina* phylogeny and population structure. Allelic variation at these loci was uncovered by the analysis of restriction site polymorphisms, although this technique is not ideal since it has low resolution of allelic variation. As is the case for allozyme studies when different buffers (BEAUMONT and BEVERIDGE 1983) or gel matrices (SINGH et al. 1976) are used additional allelic variants within a certain allele 'class' are revealed, it is possible that the application of more restriction enzymes would uncover additional variants at the studied loci. High resolution techniques such as single stranded conformational polymorphism (SSCP) analysis (e.g. FRIESEN et al. 1997) would be particularly suitable for a study such as this, since sequence information could be gathered for each allele identified, thus allowing additional analyses of the population data such as AMOVA.

ACKNOWLEDGEMENTS

REFERENCES

Ashburner M, (1989). Drosophila: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Avise JC, (1994). Molecular Markers, Natural History, and Evolution. Chapman and Hall, London.

Backeljau T and Warmoes T, (1992). The phylogenetic relationships of ten Atlantic littorinids assessed by allozyme electrophoresis. In: Proceedings of the Third International Symposium on Littorinid Biology, (eds J Grahame, PJ Mill and D Reid), The Malacological Society of London, London, p. 9-24.

Barnes RSK, (1993). On the nature of the coastal lagoon winkles attributed to Littorina tenebrosa and Littorina saxatilis. Cahiers de Biologie Marine 34: 477-495.

Beaumont AR and Beveridge CM, (1983). Resolution of phosphoglucomutase isozymes in Mytilus edulis L. Marine Biology Letters 4: 97-103.

Caley K, Grahame J and Mill PJ, (1995). A geographically-based study of shell shape in small rough periwinkles. Hydrobiologia 309: 181-193.

Côrte-Real HBSM, Dixon DR and Holland PWH, (1994). Intron-targeted PCR: a new approach to survey neutral DNA polymorphism in bivalve populations. Marine Biology 120: 407-413.

Coyne JA and Kreitman M, (1986). Evolutionary genetics of two sibling species, Drosophila simulans and D. secheliia. Evolution 40: 673-691.

Crossland S, Coates D, Grahame J and Mill PJ, (1996). The Littorina saxatilis species complex - interpretation using random amplified polymorphic DNAs. In: Origin and evolutionary radiation of the Mollusca, (ed. J Taylor), Oxford University Press, Oxford, p. 205-209.

Degnan SM, (1993). The perils of single gene trees-mitochondrial versus single-copy nuclear DNA variation in white-eyes (Aves: Zosteropidae). Mol. Ecol. 2: 219-225.

Felsenstein J, (1993). PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

Friesen VL, Congdon BC, Walsh HE and Birt TP, (1997). Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. Mol. Ecol. 6: 1047-1058.

Grahame J, Hull SL, Mill PJ and Hemmingway R, (1997). Discovering unrecognised diversity among marine molluscs. In: Marine Biodiversity: Patterns and Processes, (eds RFG Ormond, JD Gage and MV Angel), Cambridge University Press, Cambridge, UK, p. 293-318.

Grahame J, Mill PJ, Hull SL and Caley KJ, (1995). Littorina neglecta Bean: ecotype or species? J. Nat. History 29: 887-899.

Hare MP, Karl SA and Avise JC, (1996). Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. Mol. Biol. Evol. 13: 334-345.

Hillis DM, (1998). Phylogenetic analysis of frequency data in molecular ecological studies. In: Advances in molecular ecology, (ed. GR Carvalho), IOS Press, Amsterdam, p. 25-38.

Janson K and Ward RD, (1985). The taxonomic status of Littorina tenebrosa Montagu as assessed by morphological and genetic analysis. J. Conchology 32: 9-15.

Johannesson K and Johannesson B, (1990). Genetic variation within Littorina saxatilis (Olivi) and Littorina neglecta Bean: is L. neglecta a good species? Hydrobiologia 193: 89-97.

Maynard Smith J, (1989). Evolutionary Genetics. Oxford University Press, Oxford.

Mill PJ and Grahame J, (1992).Distribution of the rough periwinkles in Great Britain. In: Proceedings of the Third International Symposium on Littorinid Biology, (eds J Grahame, PJ Mill and DG Reid), The Malacological Society, London, p. 305-307.

Nagl S, Tichy H, Mayer WE, Takahata N and Klein J, (1998). Persistence of neutral polymorpisms in Lake Victoria cichlid fish. Proc. Natl. Acad. Sci. USA 95: 14238-14243.

Nei M, (1987). Molecular Evolutionary Genetics. Columbia University Press, New York.

Raymond M and Rousset F, (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J. Heredity 86: 248–249.

Reid DG, (1993). Barnacle-dwelling ecotypes of three British Littorina species and the status of Littorina neglecta Bean. J. Molluscan Studies 59: 51–62.

Reid DG, (1996). Systematics and evolution of Littorina. The Ray Society, London.

Rice WR, (1989). Analyzing tables of statistical tests. Evolution 43: 223–225.

Singh RS, Lewontin RC and Felton AA, (1976). Genetic heterogeneity within electrophoretic "alleles" of xanthine dehydrogenase in Drosophila pseudoobscura. Genetics 84: 609–629.

Swofford DL and Selander RB, (1989). BIOSYS-1. A computer program for the analysis of allelic variation in population genetics and biochemical systematics. Illinois Natural History Survey.

Waples RS, (1987). A multispecies approach to the analysis of gene flow in marine shore fishes. Evolution 41: 385–400.

Ward RD, (1990). Biochemical genetic variation in the genus Littorina (Prosobranchia: Mollusca). Hydrobiologia 193: 53–69.

Wilding CS, Grahame J and Mill PJ, (2000). Mitochondrial DNA CoI haplotype variation in sibling species of rough periwinkles. Heredity 84: 62–74.

Xu H, Wilson DJ, Arulsekar S and Bakalinsky AT, (1996). Sequence-specific polymerase chain-reaction markers derived from randomly amplified polymorphic DNA markers for fingerprinting grape (Vitis) rootstocks. J. Am. Soc. Horticultural Science 120: 714–720.