

Mitochondrial DNA Col haplotype variation in sibling species of rough periwinkles

C. S. WILDING, J. GRAHAME & P. J. MILL

School of Biology, The University of Leeds, Leeds LS2 9JT

Three sibling species of rough periwinkles are currently recognized: *Littorina arcana*, *L. compressa* and *L. saxatilis*. Certain forms of *L. saxatilis* are also argued by some to deserve species status, such as the barnacle-dwelling '*L. neglecta*' and the lagoonal '*L. tenebrosa*'. Relationships between these taxa, and between and within representative populations, are investigated using sequence analysis and restriction fragment length polymorphism of a mitochondrial DNA fragment spanning the cytochrome oxidase I and cytochrome oxidase II gene boundary. These data show that there is some sharing of haplotypes between species, with *L. arcana* haplotypes paraphyletic with respect to *L. saxatilis* haplotypes, and *L. compressa* haplotypes paraphyletic to both *L. arcana* and *L. saxatilis* haplotypes. Such sharing of mtDNA haplotypes could be a consequence of either persistent hybridization or episodes of hybridization, or incomplete lineage sorting of ancestral polymorphisms. On the balance of evidence it is suggested that the latter, rather than hybridization events, is the more likely causal agent of the observed distribution. Intraspecific variation is extensive and it is suggested that the patterns of intraspecific polymorphism are explainable by a combination of historical factors (the impact of the Pleistocene ice-age) and contemporary restrictions to gene flow. It is argued that *Littorina* haplotypes evolved in at least two separate glacial refugia and became scattered by the subsequent range expansion around most of the coastline. Recent factors such as restricted gene flow and bottlenecks would then be capable of affecting the haplotype distribution, resulting in the pattern observed.

Keywords: hybridization, lineage sorting, *littorina*, mitochondrial DNA, phylogeography, post-glacial recolonization.

Introduction

Periwinkles of the genus *Littorina* are familiar components of rocky shore communities. Around northern European coastlines six species are recognized. *L. littorea* (L.), the edible periwinkle, is a distant relative to the other five species, probably sharing an ancestor 12–15 Ma (Reid *et al.*, 1996). These five are all members of the subgenus *Neritrema*, including two flat periwinkles *L. fabalis* (W. Turton) and *L. obtusata* (L.), and three sibling species of rough periwinkles, namely *L. saxatilis* (Olivi), *L. arcana* Hannaford Ellis and *L. compressa* Jeffreys. These five species have existed for a maximum of 3.5–4 Myr when their ancestral *Littorina* crossed from the Pacific into the Atlantic following the opening of the Bering Strait.

The age of the rough periwinkles is difficult to estimate, but reviewing allozyme evidence Reid *et al.*

(1996) suggest that a time period of approximately 1.73 Myr (SE = 0–3.55) is likely for this group.

The three rough periwinkle species show considerable allozyme polymorphism, which is apparently greatest in *L. saxatilis* (Ward, 1990). They lack a planktonic dispersal phase, and therefore show greater interpopulation genetic heterogeneity than do related snails with such dispersal (Janson, 1987). Further, their distribution on coasts differs: *L. arcana* is absent from the eastern two thirds of the south coast of England and *L. compressa* is particularly patchy in distribution; in contrast to these two, *L. saxatilis* is relatively ubiquitous on rocky coasts (Mill & Grahame, 1992). These distributions, together with the levels of inherent variation, population size and history, and low dispersal and gene flow, should have an influence on the genetic structure of the populations.

Allozyme electrophoresis studies of these species have failed to detect diagnostic loci, although there are frequency differences. Phylogenies based on allele frequencies are usually in agreement with morphological

*Correspondence. E-mail: bgyesw@leeds.ac.uk

data and analysis of reproductive mode in suggesting that *L. compressa* is the least derived member of the group (Ward, 1990). However, from allozyme evidence Backeljau & Warmoes (1992) concluded that *L. arcana* was basal. Whichever is correct, it is certain that the species are closely related, and this, coupled with the existence of contentious forms, has led to the term *L. saxatilis* species complex as a descriptor of the group as a whole. Their close relationships are corroborated by the ability of male *L. saxatilis* and female *L. arcana* to hybridize under laboratory conditions (Warwick *et al.*, 1990). Enzyme studies have been applied to the analysis of some of the contentious forms such as the barnacle-dwelling *L. neglecta* (Johannesson & Johannesson, 1990) and the lagoonal *L. tenebrosa* (Janson & Ward, 1985), with the conclusion that there is no evidence for species status, since neither diagnostic loci nor frequency shifts were evident. However, allozyme studies can be influenced by selection on the loci involved, moreover allozyme-inferred loci may require extensive time to reach a point at which separate gene pools are detectable, leading to low resolution. DNA based studies are potentially more powerful (Mitton, 1994), but few such studies have been undertaken on *Littorina*. Reid *et al.* (1996) established a mtDNA (cytochrome *b*, 12SrRNA and 16SrRNA) phylogeny of all extant *Littorina* but the resolution of the three rough periwinkle species was low, and there was no indication as to how inclusion of multiple individuals (and therefore potential intraspecific variation) affected the phylogeny. Kyle & Boulding (1998) reported on the use of mtDNA cytochrome *b* variation to examine population structure of *L. subrotundata* but, as yet, there have been no attempts to use mtDNA to examine population level variability in North Atlantic *Littorina* species. Evidence from maternally inherited mtDNA may allow analysis of introgression or of directionality of possible hybridization between closely related species, especially where comparable nuclear data sets are available.

Here we investigate the mitochondrial cytochrome oxidase I gene with respect to phylogeny, phylogeography, evidence for monophyly in the three currently recognized species, and the possibility of hybridization between them in the wild. We also ask what evidence these data provide concerning the taxonomic status of some contentious forms of *L. saxatilis* such as *L. neglecta* (e.g. Caley *et al.*, 1995) and *L. tenebrosa* (e.g. Barnes, 1993).

Materials and methods

Sample collection

Samples of *Littorina* were collected from locations around the United Kingdom and Eire over the period

Table 1 Locations of collection sites for *Littorina* samples used in the present study

Location	Grid reference	Species
Ursholmen, Sweden	58°50'10"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
Filey	OS: TA128816	N
Snettisham	OS: TF649319	S
Wells Next The Sea	OS: TF909458	
(2 samples)	OS: TF915456	S
Holkham	OS: TF886451	T
Cley	OS: TG062448	
(2 samples)	OS: TG067447	S
Alderton	OS: TM363419	S
St Margaret's at Cliffe	OS: TR368444	S
Folkestone	OS: TR245369	
(2 samples)	OS: TR244373	S
St Alban's Head	OS: SY959754	S
Portland Bill	OS: SY675683	S
East Fleet	OS: SY799635	S
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
Cape Cornwall	OS: SW349318	C
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
West Dale 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
Porth Towyn	OS: SH229376	C
Port St Mary, Isle of Man	OS: SC213680	S
Nynian's Cove,		
Galloway	OS: NW417364	S
Mizen Head	IGR: V 074023	C, S
Inismór, Aran Islands	IGR: L 103221	C, S
Ballynahown	IGR: L 992202	A, C, S
Golam Head	IGR: L 826214	S, T
Seapoint Beach,		
near Kittery, ME, USA		S

A, *Littorina arcana*; B, *L. saxatilis* 'b'; C, *L. compressa*; N, *L. neglecta*; S, *L. saxatilis*; T, *L. tenebrosa*; OS, Ordnance Survey grid reference; IGR, Irish grid reference.

between July 1996 and August 1998 (Table 1). Animals were collected from shores where locally abundant. Samples were taken over an area of at least 2 m² in an effort to avoid sampling of single families. A single specimen of *L. sitkana* from Akkeshi, Hokkaido, Japan was utilized as an outgroup.

DNA extraction and PCR

Total genomic DNA was extracted from head-foot tissue of individual animals using protocol 47 of Ashburner (1989). DNA sample concentrations were measured by fluorometry and adjusted to $10 \text{ ng } \mu\text{L}^{-1}$. PCR was performed in $50 \mu\text{L}$ volumes containing 50 mM KCl , $10 \text{ mM Tris-HCl pH } 9.0$, 1.5 mM MgCl_2 , 0.1% Triton X-100, 0.01% gelatin, $200 \mu\text{M}$ each dNTP, 25 pmol each primer, 25 ng DNA and 1 U Taq (Supertaq, HT Biotechnologies). The primer pair saxcoI and saxcoII, designed from sequence in Wilding *et al.* (1999) (accession number AJ132137), was used. SaxcoI covers bases 219–240 and saxcoII bases 1185–1166 of Wilding *et al.* (1999), and amplifies 967 bp (inclusive of primers) covering the 3' end of cytochrome oxidase I, 30 bp of apparent noncoding sequence and the first 38 bp of CoII. PCR conditions of $1 \times 94^\circ\text{C}$, 5 min ; $35 \times (94^\circ\text{C}, 1 \text{ min}, 55^\circ\text{C}, 30 \text{ s}, 72^\circ\text{C}, 1 \text{ min})$; $1 \times 72^\circ\text{C}$, 5 min were employed. Positive PCR reactions were cleaned by PEG precipitation (Ausubel *et al.*, 1990) and resuspended in $20 \mu\text{L}$. Subsequently $4 \mu\text{L}$ was used in an automated sequencing reaction with the amplification primer saxcoII, using Taq terminator sequencing mix, and run on an ABI377.

Sequence analysis

Haplotype sequences were aligned in CLUSTAL W (Thompson *et al.*, 1994). Identical haplotypes were identified through the construction of a distance matrix in MEGA v. 1.02 (Kumar *et al.*, 1993) and removed from the dataset. Sequence variation was examined with DnaSP v. 2.2 (Rozas & Rozas, 1997). Phylogenies were constructed using both distance and maximum likelihood (ML) methodologies with *L. sitkana* as an out-group. Haplotype relationships were inferred by both neighbour joining of Jukes–Cantor distances with 1000 bootstrapped data sets, implemented in MEGA v. 1.02 (Kumar *et al.*, 1993) and by ML analysis performed by quartet puzzling (Strimmer & von Haeseler, 1996) using PUZZLE with 1000 quartet puzzling steps (analysis performed at [HTTP://bioweb.pasteur.fr/seqanal/interfaces/Puzzle.html](http://bioweb.pasteur.fr/seqanal/interfaces/Puzzle.html)). A haplotype network was also manually constructed detailing the relationships of haplotypes and the number of mutational steps required to interconvert haplotypes.

RFLP analysis

From aligned sequences, polymorphic restriction enzymes were chosen for which the recognition sequence spanned a variable site. Three suitable enzymes (*DdeI*, *DraI* and *HindIII*) were identified and employed to

screen for variation. Digestions were performed in $15 \mu\text{L}$ volumes containing $5 \mu\text{L}$ PCR product, $1.5 \text{ mL } 10 \times$ reaction buffer and 1–4 units enzyme at 37°C . Digestion products were separated on 2% agarose gels.

Estimates of θ (F_{ST} analogue) were calculated using FSTAT (Goudet, 1995) with standard errors calculated using jackknifing procedures. AMOVA and F_{ST} were performed in ARLEQUIN v. 1.1 (Schneider *et al.*, 1997). For the AMOVA analysis, six groups were identified on the basis of regional differences in shell morphology and on distributional breaks: south-east, south-west, South Wales, Irish Sea, Ireland and East coast. Contingency tests were implemented in GENEPOP (Raymond & Rousset, 1995).

Results

Sequence analysis

PCR using the primer pair saxcoI–saxcoII generated a 967-bp fragment. The single primer saxcoII was then used to sequence one end of this fragment. Four hundred base pairs of unambiguous sequence were generated from each of 95 separate individual *Littorina* representing a range of recognized species including 89 rough periwinkles (*L. arcana*, *L. compressa* and *L. saxatilis* and including some representatives of the contentious taxa *L. neglecta*, *L. tenebrosa* and *L. saxatilis* 'b'), two *L. fabalis*, three *L. obtusata* and one specimen of *Littorina sitkana* (Table 2). Twenty-four distinct haplotypes were identified from these 95 individuals (Table 3). Sequences are deposited in Genbank with accession numbers AJ133311–133334. Of the 400 bp of sequence, 56 sites exhibited variation, all involving individual substitutions with the exception of a single base indel difference between the *L. sitkana* sequence and all other sequences. When *L. sitkana* is removed there are 37 variable sites of which 25 are parsimony informative. If only the rough periwinkles are considered these figures drop to 28 variable sites with 16 parsimony informative positions. All substitutions are silent with the exception of the $G \rightarrow A$ transition at position 938 which alters the amino acid from valine to isoleucine, and the $T \rightarrow G$ transversion at position 1070 which changes the amino acid from serine to alanine. Because of the low number of parsimony informative sites, parsimony-based phylogeny reconstruction methods were considered inappropriate and haplotype phylogenies were constructed using both a distance method and a maximum likelihood variant. The resultant trees (Fig. 1A and B) are essentially congruent, although the distance-based tree suggests some details of branching order that are not seen in the maximum likelihood tree. Bootstrap support values

Table 2 Numbers of haplotypes 1–24 encountered at various locations around the UK and Ireland through sequencing of 400 bp of mtDNA CoI from individual *Littorina*

Species	Sampling site	N	Haplotype																							
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
<i>L. saxatilis</i>	N. Berwick	1	1																							
<i>L. saxatilis</i>	St Abb's	1									1															
<i>L. saxatilis</i>	Old Peak	4									2				2											
<i>L. saxatilis</i>	Snettisham	2						2																		
<i>L. saxatilis</i>	St Margaret's	2	1			1																				
<i>L. saxatilis</i>	Folkestone	11	1		2			7							1											
<i>L. saxatilis</i>	Portland Bill	5								1				4												
<i>L. saxatilis</i>	East Fleet	1								1																
<i>L. saxatilis</i>	St Albans	2									1			1												
<i>L. saxatilis</i>	Pinhay Bay	5												5												
<i>L. saxatilis</i>	Cargreen	1						1																		
<i>L. saxatilis</i>	Lizard	5						2						3												
<i>L. saxatilis</i>	Porth Ysgo	4						1					1										2			
<i>L. saxatilis</i>	Gann	2											2													
<i>L. saxatilis</i>	Isle of Man	2											2													
<i>L. saxatilis</i>	Mizen Head	1						1																		
<i>L. saxatilis</i>	Aran Islands	1											1													
<i>L. saxatilis</i>	Maine, USA	1													1											
<i>L. neglecta</i>	N. Berwick	2											2													
<i>L. neglecta</i>	Peak Steel	2																					2			
<i>L. neglecta</i>	Filey	1									1															
<i>L. saxatilis</i> 'b'	Peak Steel	2											1									1				
<i>L. tenebrosa</i>	Golam	2											2													
<i>L. arcana</i>	St Abb's	2																					2			
<i>L. arcana</i>	Old Peak	2																					2			
<i>L. arcana</i>	Lizard	4											3										1			
<i>L. arcana</i>	Porth Towyn	1											1													
<i>L. arcana</i>	Ballynahown	1	1																							
<i>L. compressa</i>	N. Berwick	3																				2	1			
<i>L. compressa</i>	St Abb's	4											1									2			1	
<i>L. compressa</i>	Lizard	3																					2	1		
<i>L. compressa</i>	Cape Cornwall	1																					1			
<i>L. compressa</i>	Gann	1	1																							
<i>L. compressa</i>	Porth Llanllawen	3	1	1																			1			
<i>L. compressa</i>	Porth Towyn	2																					2			
<i>L. compressa</i>	Mizzen Head	1																					1			
<i>L. compressa</i>	Aran Islands	1																					1			
<i>L. fabalis</i>	Old Peak	2																						2		
<i>L. obtusata</i>	Old Peak	3																						1	2	
<i>L. sitkana</i>	Akkeshi, Japan	1																							1	

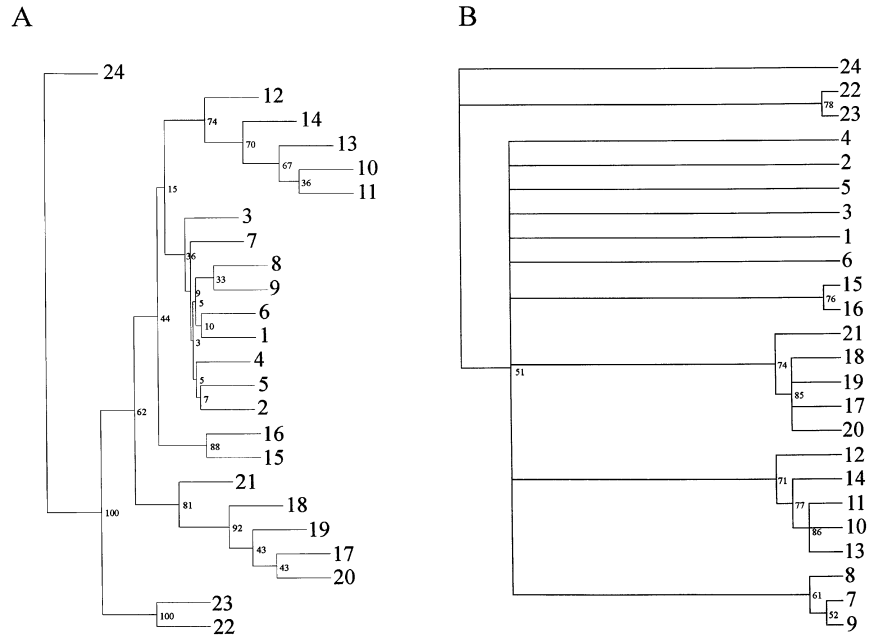


Fig. 1 (A) Neighbour-Joining tree of Jukes–Cantor distances for the 24 ColI haplotypes of *Littorina*. Numbers at nodes refer to percentage bootstrap support from 1000 replicates. (B). Quartet-puzzling maximum likelihood tree constructed using the HKY model in PUZZLE. Numbers at nodes indicate how many times the cluster to the right of the node was found in 1000 quartet puzzling steps.

are generally low, although there are some well supported nodes.

Disregarding the outgroup (*L. sitkana*, haplotype 24) and the sequences from flat periwinkles (haplotypes 22 and 23), there are four main groups discernible. The haplotypes limited to *L. compressa* (17–21) form a single group, and haplotypes 15 and 16 which are mainly limited to *L. arcana* (and some east coast *L. compressa*) also group together. The remaining *L. saxatilis* sequences (in a few instances these also occur in *L. arcana* or *L. compressa* — see later for details) form two main groups. One group (termed *L. saxatilis* group β) is well

supported in both phylogenies (haplotypes 10–14). The remaining sequences (1–9) form a single group in the distance method, whilst in the ML tree they are unresolved, with the exception of a group of three (7–9). Haplotypes 1–9 are labelled *L. saxatilis* group α . The lack of resolution in the phylogeny of haplotypes of group α is a consequence of these sequences differing from each other by only singleton base changes. Although the support values within the *L. saxatilis* haplotypes are low, the two distinct groups of *L. saxatilis* haplotypes are also revealed in the haplotype network (Fig. 2). Here the four separate haplotype

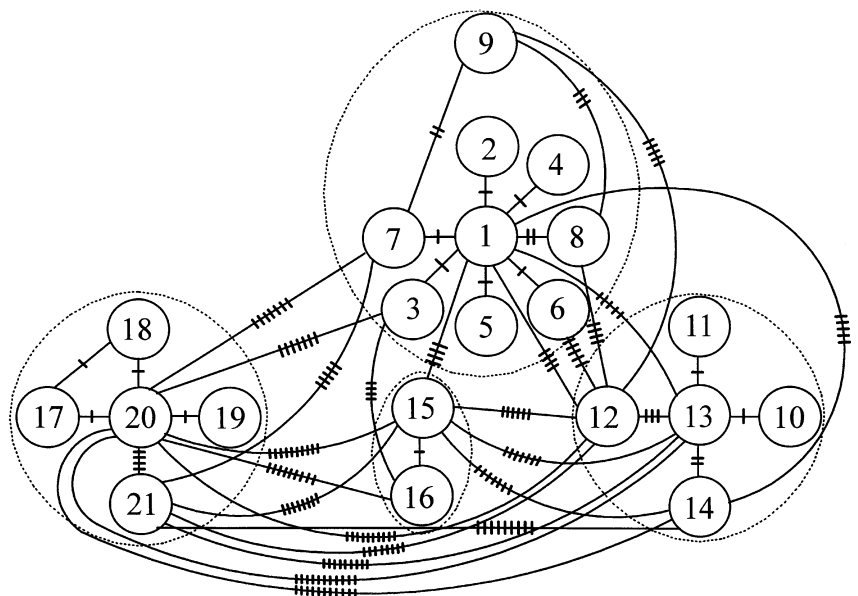


Fig. 2 Haplotype network showing the number of base changes between 24 ColI haplotypes of *Littorina*. Bars on interconnecting lines represent the number of mutational steps. Dashed lines encircle major groups (see text).

groups (the 'compressa' group, the 'arcana' group and *L. saxatilis* groups α and β) distinguishable in the phylogenies can be discerned (surrounded by the dashed lines). This haplotype network is not a minimum spanning tree in that individual haplotypes are not separated by single substitutions, and there are many ambiguous connections. This is probably a consequence of not having uncovered all existent haplotypes, and perhaps also due to homoplasies in the data (Besansky, 1997; Brown Gladden *et al.*, 1997), although such events seem less likely when haplotypes of closely related species are being compared.

Most haplotypes are unique to a particular species, with the exceptions of haplotypes 1 and 10, which are seen in *L. saxatilis*, *L. arcana* and *L. compressa*, and haplotype 15, which has been encountered in both *L. arcana* and *L. compressa*. Representative *L. neglecta*, *L. tenebrosa* and *L. saxatilis* 'b' did not have distinct haplotypes, but shared haplotypes with *L. saxatilis* (Table 2).

Restriction digestion

From sequence information three informative restriction enzymes were identified, *DdeI*, *DraI* and *HindIII*. These released 1–3 fragments per enzyme (Table 4). Individual haplotypes were designated by single letter codes and a three-letter composite haplotype was then generated for each animal. There were insufficient informative enzymes to differentiate every sequenced haplotype. However, the three enzymes used allowed resolution into the main groups obvious from the haplotype network and phylogenies. Six separate composite haplotypes were uncovered. The *L. compressa* group (17–21) are all BAB and the *L. arcana* haplotype group (15–16) all BAA. For *L. saxatilis* group α , haplotypes 2 and 5 are BCA and the remainder are BBA, whilst for group β , haplotypes 10 and 12–14 are

Table 4 Sizes of fragments (in base pairs) generated by restriction enzyme digestion with three separate restriction enzymes (*DdeI*, *DraI* and *HindIII*)

Restriction enzyme	Pattern	Fragment sizes		
<i>DdeI</i>	A	333	634	
	B	257	333	377
	C	195	333	439
	D	966		
<i>DraI</i>	A	60	907	
	B	967		
<i>HindIII</i>	A	967		
	B	190	777	

AAA and haplotype 11 is ADA. A stylised tree based on Fig. 1 A and B and summarizing the composite haplotype information is depicted in Fig. 3. Table 5 and Fig. 4 together describe the distribution of composite haplotypes within populations of the various species. Because sample sizes are much higher than those utilized for the sequencing study, this RFLP analysis gives a much better picture of the geographical distribution of haplotypes. In *L. arcana* there is a definite east–west split, with all east coast members having the BAA haplotype and the majority of west coast animals being AAA. Thus there is a geographical pattern to those haplotypes that are shared with *L. saxatilis*. East coast animals (St Abb's, North Berwick and Old Peak) all have the '*L. arcana*' haplotype BAA, whilst those on the west coast often have haplotypes typical of *L. saxatilis*, such as AAA. For *L. compressa* the picture is less clear and variation is higher than in *L. arcana*. In this species there seems to be no obvious pattern to those haplotypes that are shared with either *L. arcana* or *L. saxatilis*. *L. compressa* from the Lizard, Trevaunance and Inismór are fixed for the 'compressa' haplotype, but elsewhere animals with 'saxatilis' haplotypes (BCA, BBA, AAA) are common. Because of more intensive sampling, there is more information on the geographical variation of haplotypes in *L. saxatilis*. Two things are apparent. Firstly, as for *L. arcana* there is a difference between east coast and west coast animals, with the BBA haplotype much more frequent on the west coast, and a higher frequency of BCA and AAA in the east. Secondly, there

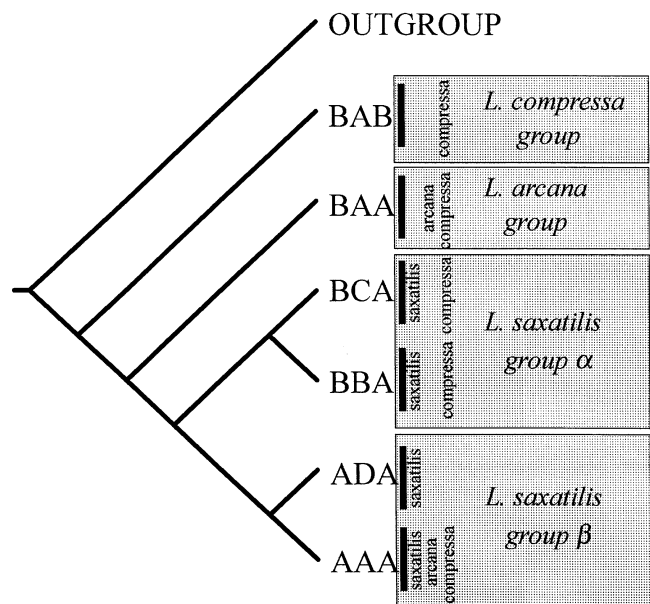


Fig. 3 Stylized dendrogram depicting relationships of RFLP composite haplotypes of *Littorina*.

Table 5 Frequency of composite haplotypes in populations of *Littorina* from locations around the UK and Ireland. Order of enzymes in composite haplotype is *Hind*III, *Dde*I, *Dra*I

	Sample site	RFLP haplotype						<i>N</i>
		AAA	BBA	ADA	BCA	BAB	BAA	
<i>L. saxatilis</i>	Portland Bill	0	3	6	0	0	0	9
<i>L. saxatilis</i>	East Fleet	0	8	2	0	0	0	10
<i>L. saxatilis</i>	Lizard	0	0	6	4	0	0	10
<i>L. saxatilis</i>	Golden Cap	0	0	20	0	0	0	20
<i>L. saxatilis</i>	St Margaret's	1	18	0	1	0	0	20
<i>L. saxatilis</i>	Pinhay Bay	0	0	5	0	0	0	5
<i>L. saxatilis</i>	Cargreen	5	0	0	5	0	0	10
<i>L. saxatilis</i>	Fleet gravel	0	10	0	0	0	0	10
<i>L. saxatilis</i>	St Albans	0	9	1	0	0	0	10
<i>L. saxatilis</i>	Trevaunance	1	7	1	4	0	0	13
<i>L. saxatilis</i>	Folkestone	1	5	0	13	0	0	19
<i>L. saxatilis</i>	N. Berwick	8	2	0	0	0	0	10
<i>L. saxatilis</i>	St Abb's	8	1	0	1	0	0	10
<i>L. saxatilis</i>	Galloway	11	1	0	6	0	0	18
<i>L. saxatilis</i>	Mumbles	3	12	0	5	0	0	20
<i>L. saxatilis</i>	Gann	15	3	0	2	0	0	20
<i>L. saxatilis</i>	St Ann's Head	37	1	0	2	0	0	40
<i>L. saxatilis</i>	Aran Islands	10	0	0	0	0	0	10
<i>L. saxatilis</i>	Old Peak	20	0	0	0	0	0	20
<i>L. saxatilis</i>	Porth Ysgo	5	15	0	0	0	0	20
<i>L. saxatilis</i>	Golam Head	2	8	0	0	0	0	10
<i>L. saxatilis</i>	Snettisham	3	0	0	7	0	0	10
<i>L. saxatilis</i>	Sweden	10	0	0	7	0	0	17
<i>L. saxatilis</i>	Porth Llanllawen	13	0	0	4	0	0	17
<i>L. saxatilis</i>	Trwyn Maen Melyn	4	6	0	0	0	0	10
<i>L. saxatilis</i>	Wells next the sea	9	0	0	11	0	0	20
<i>L. saxatilis</i>	Cley	42	0	0	0	0	0	42
<i>L. saxatilis</i>	Alderton	0	0	0	10	0	0	10
<i>L. saxatilis</i>	Ballynahown	1	19	0	0	0	0	20
<i>L. saxatilis</i>	West Dale Bay	0	10	0	0	0	0	10
<i>L. saxatilis</i> 'b'	Peak Steel	27	3	0	0	0	0	30
<i>L. neglecta</i>	Peak Steel	70	2	0	0	0	0	72
<i>L. neglecta</i>	N. Berwick	7	0	0	0	0	0	7
<i>L. neglecta</i>	West Dale Bay	0	12	0	7	0	0	19
<i>L. tenebrosa</i>	Holkham	0	0	0	10	0	0	10
<i>L. tenebrosa</i>	Golam	18	2	0	0	0	0	20
<i>L. arcana</i>	Lizard	6	0	0	0	0	4	10
<i>L. arcana</i>	Trevaunance	10	0	0	0	0	0	10
<i>L. arcana</i>	N. Berwick	0	0	0	0	0	10	10
<i>L. arcana</i>	St Abb's	0	0	0	0	0	10	10
<i>L. arcana</i>	Mumbles	5	0	0	0	0	0	5
<i>L. arcana</i>	Old Peak	0	0	0	0	0	15	15
<i>L. arcana</i>	Porth Llanllawen	5	0	0	0	0	0	5
<i>L. arcana</i>	Ballynahown	6	2	0	0	0	0	8
<i>L. arcana</i>	West Dale Bay	10	0	0	0	0	0	10
<i>L. compressa</i>	Lizard	0	0	0	0	10	0	10
<i>L. compressa</i>	Trevaunance	0	0	0	0	11	0	11
<i>L. compressa</i>	N. Berwick	3	0	0	0	1	6	10
<i>L. compressa</i>	St Abb's	1	0	0	0	5	4	10
<i>L. compressa</i>	Gann	2	0	0	0	2	2	6
<i>L. compressa</i>	Aran Islands	0	0	0	0	10	0	10
<i>L. compressa</i>	Porth Llanllawen	0	1	0	7	2	0	10
<i>L. compressa</i>	Ballynahown	0	0	0	0	9	1	10
<i>L. compressa</i>	West Dale Bay	0	2	0	1	7	0	10

N, sample size.

is a south-west enclave of *L. saxatilis* with a high frequency of the ADA composite haplotype. This haplotype is seen nowhere else in the country. F -statistics (θ and F_{ST}) indicate substantial population subdivision (Table 6) and this is corroborated by contingency tests which show significant differences among populations of all three species ($P=0.000 \pm 0.000$). AMOVA analysis (Table 6) using the preordained groups identified on the basis of shell morphology and known disjunctions in littorinid distribution, indicates varying degrees of evidence for mtDNA variability being partitioned among these groups. Although for *L. arcana* over 70% of the total variation is among these groups, the figure is somewhat misleading, since virtually all variation is partitioned solely on an east coast–west coast basis, and the choice of groups reflects this to an extent. For *L. compressa* and *L. saxatilis*, substantial variation is within populations, reflecting high levels of polymorphism. For *L. saxatilis*, little of the remainder of the variation is among groups, indicating that these do not explain the variation encountered. In *L. compressa* nearly 40% of the variation is among groups, but as for *L. arcana*, this is probably attributable to east–west division, whereby most west coast *L. compressa* show the ‘*compressa* haplotype’ whereas a higher proportion of east coast animals display the ‘*saxatilis*’ or ‘*arcana*’ type.

Discussion

The phylogeny of Littorina mitochondrial Col haplotypes

The deduced branching order of the *Littorina* species considered here is in agreement with current, consensus opinion (Reid, 1996). The low bootstrap support in the phylogeny at nodes leading to the three rough periwinkle species could be a consequence of a period of rapid speciation, with only short time periods between speciation events during which informative substitutions could accrue. Periods of rapid speciation have similarly affected phylogeny construction in finches (Fehrer, 1996) and corvids (Helm-Bychowski & Cracraft, 1993). Paraphyletic haplotype distributions are also apparent with *L. obtusata* haplotypes paraphyletic with respect to *L. fabalis*, *L. arcana* haplotypes paraphyletic with respect to *L. saxatilis*, and *L. compressa* haplotypes paraphyletic with respect to both *L. arcana* and *L. saxatilis* haplotypes. Such paraphyly may have arisen as a consequence of shared ancestral polymorphisms when (usually moderate to large) populations have not progressed to reciprocal monophyly during speciation. However introgressive hybridization may also be involved (Besansky, 1997; Schneider-Broussard *et al.*,

1998) and it is well documented that mtDNA can cross species boundaries through hybridization or introgression (Avice, 1994). If hybridization were the causal mechanism of the observed haplotype distribution, this would require viability of hybridization between all three species. Although Warwick *et al.* (1990) showed laboratory hybridization in one direction (female *L. arcana* with male *L. saxatilis*), there is no information on the mating preferences during back-crosses, and there were no successful hybridizations with other species crosses. We suggest that this asymmetrical haplotype distribution, with ‘*L. saxatilis*’ haplotypes observed in *L. arcana* and *L. compressa* but no ‘*L. compressa*’ haplotypes or ‘*L. arcana*’ haplotypes found in *L. saxatilis*, rules out general hybridization (i.e. occasional gene flow between the species). Although introgression of the mtDNA of one species into another is not precluded, it seems unlikely on the basis of existing evidence.

It is also the case that the animals with haplotypes typical of another species have a restricted distribution. For instance, *L. arcana* with the ‘*saxatilis*’ haplotype occur on the west coast of Britain, but have not been found on the east coast (Fig. 4, Table 5), and *L. compressa* with particular ‘*arcana*’ or ‘*saxatilis*’ haplotypes are patchy. In most cases these haplotypes are not the commonest ones, and in some cases are not present sympatrically in the other species, e.g. all West Dale Bay *L. arcana* have a ‘*saxatilis*’ haplotype AAA, but all West Dale Bay *L. saxatilis* are BBA. This again suggests that localized hybridization is not the cause of the observed haplotype sharing, unless the hybridization happened in the past when, perhaps, haplotype frequencies were very different.

The alternative explanation, retention of ancestral polymorphisms, has been documented for mtDNA genes in other groups (e.g. Helm-Bychowski & Cracraft, 1993; Fehrer, 1996; Schneider-Broussard *et al.*, 1998), in which shared haplotypes antedate speciation and are haphazardly maintained in certain populations through incomplete lineage sorting. Rapid and recent speciation (see above) is consistent with this since there will have been little time for sorting of ancestral haplotypes into the descendant taxa. This hypothesis also fits with the ‘directional’ nature of haplotype sharing. The least derived littorinid, *L. compressa*, displays haplotypes representative of all groups currently recognized as ‘*L. compressa* type’, ‘*L. arcana* type’ and ‘*L. saxatilis* type’. A more limited range of haplotypes were passed on to the descendant *L. arcana* and *L. saxatilis*, being enriched with respect to what has become the ‘*arcana* type’ and the ‘*saxatilis* type’ whilst still leaving historical signatures — ‘*arcana* type’ and ‘*saxatilis* type’ haplotypes in *L. compressa* and ‘*saxatilis* type’ haplotypes in

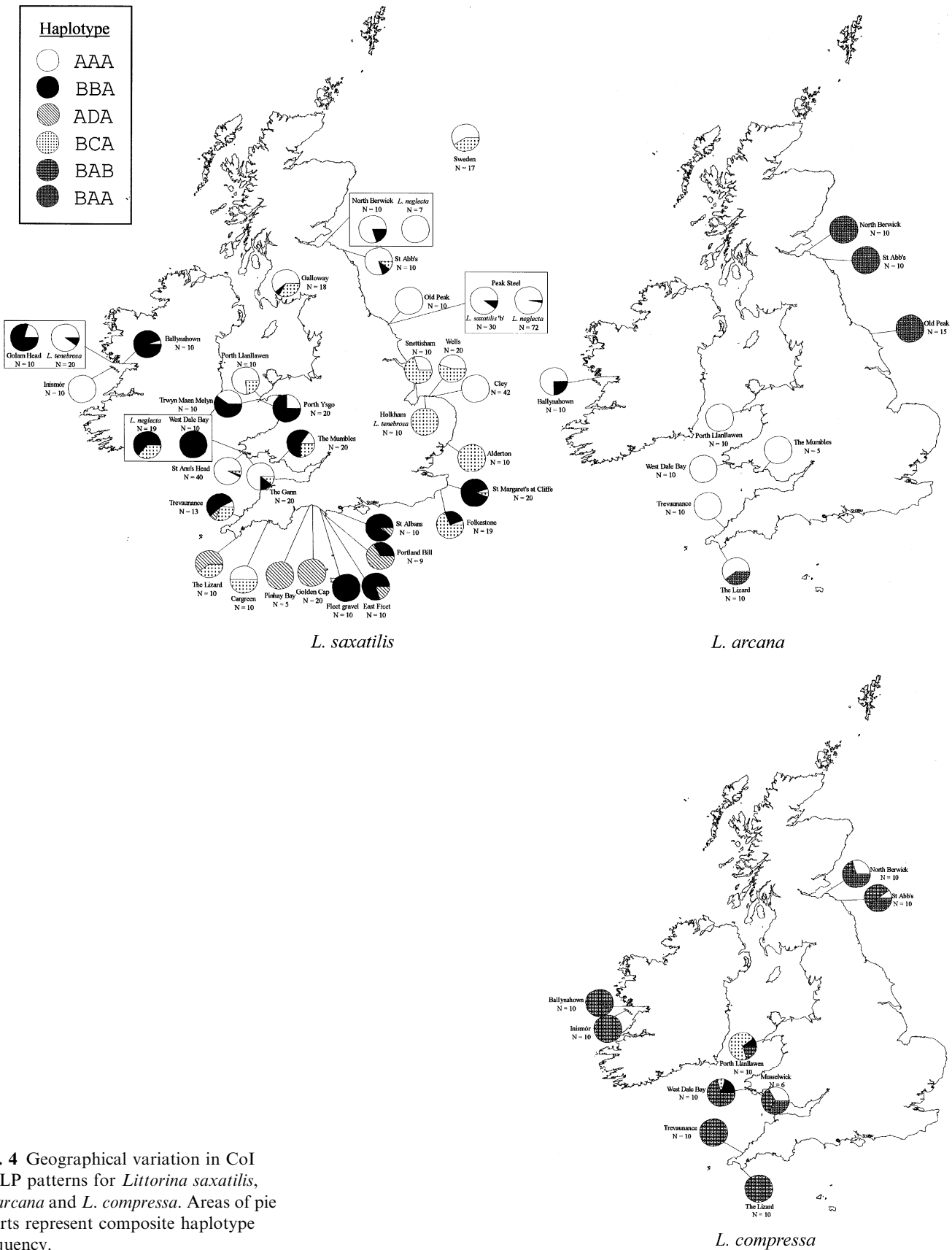


Fig. 4 Geographical variation in Col RFLP patterns for *Littorina saxatilis*, *L. arcana* and *L. compressa*. Areas of pie charts represent composite haplotype frequency.

L. arcana. The regional distribution of shared haplotypes is also consistent with this mechanism. Since all three species are either ovoviviparous or oviparous, and therefore undergo reduced gene flow, they will be more prone to bottlenecks and genetic drift, which would, over the time since speciation, have let some populations approach fixation for species-characteristic haplotypes ahead of others.

Watterson & Guess (1977) have argued that the most frequent alleles (haplotypes) may be the oldest; they will consequently be most likely to be shared. It is the case here that the shared haplotypes (i.e. numbers 1, 10 and 15; Table 3) are three of the six most frequent.

Lineage sorting should be quicker for mtDNA than for nuclear DNA and thus, if there is incomplete sorting of mtDNA haplotypes resulting in paraphyly, then this should be particularly evident in the case of nuclear alleles (Besansky, 1997). From data on nuclear DNA RFLPs at two calmodulin introns and two anonymous loci (Wilding *et al.*, unpubl. data) there is evidence that alleles at nuclear loci are incompletely sorted. They are not homogeneous between species, which again suggests that unconstrained gene flow between species is not occurring (Clarke *et al.*, 1996); neither are they fully diagnostic, suggesting mtDNA has not introgressed from a fully diagnostic nuclear background. In view of all these considerations, we suggest that the observed paraphyly is most likely due to incomplete lineage sorting.

Phylogeography of cytochrome oxidase I haplotypes

Population-level differences in the frequencies of composite haplotypes are evident within all three currently recognized species. Much of the variation is within populations, particularly for *L. saxatilis* and *L. compressa*, but there are significant differences between populations of all three species. Although there appear

to be regional differences, the exact nature of these, and the factors responsible are difficult to disentangle, although an east coast — west coast differentiation seems apparent, particularly so for *L. saxatilis* and *L. arcana*. The distribution of these haplotypes is particularly complex for the more intensively sampled *L. saxatilis*. From the haplotype network and phylogenies, two groups of *L. saxatilis* haplotypes are recognizable. Similar haplotype networks occur in anchovy (Magoulas *et al.*, 1996) and beluga whale (Brown Gladden *et al.*, 1997), and have been argued to result from the origination of haplotypes in different parts of the species range. However, haplotypes from both clusters in *L. saxatilis* can be found around the sampled range, and the current distribution of *L. saxatilis* in the range studied is largely continuous, with no obvious disjunctions on either side of which evolution could occur in relative isolation. However, historically, this may not have been so, particularly during ice ages. The concept of glacial refugia in which species could survive (Taberlet *et al.*, 1998) seems less appropriate to intertidal species, since it seems likely that ice sheets and permafrost should affect land biota more severely than species in the more buffered marine environment. However, it is without question that the distribution of *Littorina* spp. during the recent ice ages must have been considerably different from their current distribution, since sea levels were considerably lower, and glaciers extended from Scandinavia to Britain and through the Irish Sea (Bowen *et al.*, 1985). Indeed, virtually all of the present range of *Littorina* considered in this study was unavailable during the last glacial maximum; thus their distribution must have been radically different. The presence of two groups of *L. saxatilis* haplotypes (α and β) may therefore be explainable by contemporary *L. saxatilis* haplotypes having evolved in two distinct areas, from which animals have subsequently dispersed and intermingled. Indeed postglaciation scattering of haplotypes has been argued to explain the observation

Table 6 Genetic structure of British and Irish *Littorina*

	<i>L. arcana</i>	<i>L. compressa</i>	<i>L. saxatilis</i>	Overall
$\theta \pm \text{SE}$	0.836 + 0.127	0.410 + 0.109	0.592 + 0.068	0.648 + 0.051
AMOVA (percentage of variation)				
Among groups	70.98	38.62	15.13	14.76
Among populations within groups	7.51	4.38	44.88	53.39
Within populations	21.51	56.99	39.99	31.85
F_{ST}	0.785	0.430	0.600	0.682

$\theta = F_{\text{ST}} \pm \text{SE}$ jackknifed over populations. AMOVA calculated using ARLEQUIN v. 1.1 (Schneider *et al.*, 1997). Groups for analysis within individual species were south-east, south-west, South Wales, Irish Sea, Ireland, east coast, and overall were the individual species (including contentious forms), i.e. *L. arcana*, *L. compressa*, *L. neglecta*, *L. saxatilis* and *L. tenebrosa*. F_{ST} calculated using ARLEQUIN.

that mtDNA haplotypes in the song sparrow *Melospiza melodia* were geographically widespread, such that closely related haplotypes were not geographically proximal (Zink & Dittmann, 1993). Similarly, range expansion from different refugia would explain why *Littorina* haplotypes from the two divergent groupings are found virtually across the species range. The situation in *L. compressa* and *L. arcana* is complicated by the presence of shared haplotypes, but, given the similarity in biology and ecology with *L. saxatilis*, similar factors, both historical and current will have moulded the pattern of haplotype variation in these species. Uncovering further evidence for this refugium hypothesis requires additional sampling of *L. arcana*, *L. compressa* and *L. saxatilis*, particularly from around mainland Europe (the site of potential refugia) to unravel the effects of historical and other factors on the distributions of mtDNA variation in these, and related, species.

Although the observed distribution of haplotypes in these species undoubtedly has some historical basis, because there is no planktonic dispersal phase the population genetic structure of these animals will also be profoundly influenced by low gene flow and small population sizes. A probable example of the effects of the levels of gene flow is seen for the ADA haplotype (haplotype 11 in the sequence study) of *L. saxatilis*, which has a very restricted distribution (in this case to south-western Britain). This seems to be a haplotype that has arisen via mutation on the south coast, remaining restricted to this area due to low gene flow. Genetic drift and inbreeding may also be important considerations due to the life histories of these species, and this may explain the wide variance in haplotype frequencies seen even in neighbouring populations (Fig. 4).

The status of L. neglecta, L. saxatilis 'b' and L. tenebrosa

CoI data provide no evidence to suggest that *L. neglecta* and *L. saxatilis* 'b' are anything other than morphological or ecological variants of *L. saxatilis*. No diagnostic haplotypes are seen. However, there are some differences in haplotype frequencies between *L. tenebrosa* and neighbouring populations of *L. saxatilis*.

Acknowledgements

We wish to thank Dr David Reid of the Natural History Museum, London for providing the samples of *L. sitkana* and Cape Cornwall *L. compressa*, David Reid and Kerstin Johannesson for the Swedish *L. saxatilis*, and Dr Del Smith for provision of the Maine sample of *L. saxatilis*. This research was

supported by the MAST-3 programme of the European Commission under contract MAS3-CT95-0042 (AMB-IO5).

References

- ASHBURNER, M. 1989. *Drosophila: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- AUSUBEL, F., BRENT, R., KINGSTON, R., MOORE, D., SEIDMAN, J., SMITH, J. AND STRUHL, K. 1990. *Current Protocols in Molecular Biology*. Wiley, New York.
- AVISE, J. C. 1994. *Molecular Markers, Natural History, and Evolution*. Chapman & Hall, New York & London.
- BACKELJAU, T. AND WARMOES, T. 1992. The phylogenetic relationships of ten Atlantic littorinids assessed by allozyme electrophoresis. In: Grahame, J., Mill, P. J. and Reid, D. G. (eds) *Proc. 3rd Int. Symp. Littorinid Biol.*, pp. 9–24. Malacological Society of London, London.
- BARNES, R. S. K. 1993. On the nature of the coastal lagoon winkles attributed to *Littorina tenebrosa* and *Littorina saxatilis*. *Cahiers Biol. Mar.*, **34**, 477–495.
- BESANSKY, N. J. 1997. Patterns of mitochondrial DNA variation within and between African malaria vectors, *Anopheles gambiae* and *An. arabiensis*, suggest extensive gene flow. *Genetics*, **147**, 1817–1828.
- BOWEN, D. Q., ROSE, J. AND MCCABE, A. M. 1985. Correlation of quaternary glaciations in England, Ireland, Scotland and Wales. *Quart. Sci. Rev.*, **5**, 299–340.
- BROWN GLADDEN, J. G., FERGUSON, M. M. AND CLAYTON, J. W. 1997. Matriarchal genetic population structure of North American beluga whales *Delphiapterus leucas* (Cetacea: Monodontidae). *Mol. Ecol.*, **6**, 1033–1046.
- CALEY, K., GRAHAME, J. AND MILL, P. J. 1995. A geographically-based study of shell shape in small rough periwinkles. *Hydrobiologia*, **309**, 181–193.
- CLARKE, B., JOHNSON, M. S. AND MURRAY, J. 1996. Clines in the genetic distance between two species of island land snails: how 'molecular leakage' can mislead us about speciation. *Phil. Trans. R. Soc. B.*, **351**, 773–784.
- FEHRER, J. 1996. Conflicting character distribution within different data sets on cardueline finches: artifact or history? *Mol. Biol. Evol.*, **13**, 7–20.
- GOUDET, J. 1995. FSTAT (version 1.2): a computer program to calculate F-statistics. *J. Hered.*, **86**, 485–486.
- HELM-BYCHOWSKI, K. AND CRACRAFT, J. 1993. Recovering phylogenetic signal from DNA sequences: relationships within the corvine assemblage (Class Aves) as inferred from complete sequences of the mitochondrial DNA cytochrome-b gene. *Mol. Biol. Evol.*, **10**, 1196–1214.
- JANSON, K. 1987. Allozyme and shell variation in two marine snails (*Littorina*, Prosobranchia) with different dispersal abilities. *Biol. J. Linn. Soc.*, **30**, 245–256.
- JANSON, K. AND WARD, R. D. 1985. The taxonomic status of *Littorina tenebrosa* Montagu as assessed by morphological and genetic analysis. *J. Conch.*, **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.

- KUMAR, D., TAMURA, K. AND NEI, M. 1993. MEGA: *Molecular Evolutionary Genetic Analysis*. The Pennsylvania State University, University Park, PA.
- KYLE, C. J. AND BOULDING, E. G. 1998. Molecular genetic evidence for parallel evolution in a marine gastropod, *Littorina subrotundata*. *Proc. R. Soc. B.*, **265**, 303–308.
- MAGOULAS, A., TSIMENIDES, N. AND ZOUROS, E. 1996. Mitochondrial DNA phylogeny and the reconstruction of the population history of a species: the case of the European anchovy (*Engraulis encrasicolus*). *Mol. Biol. Evol.*, **13**, 178–190.
- MILL, P. J. AND GRAHAME, J. 1992. Distribution of the rough periwinkles in Great Britain. In: Grahame, J., Mill, P. J. and Reid, D. G. (eds) *Proc. 3rd Int. Symp. Littorinid Biol.*, pp. 305–307. Malacological Society of London, London.
- MITTON, J. B. 1994. Molecular approaches to population biology. *Ann. Rev. Ecol. Syst.*, **25**, 45–69.
- RAYMOND, M. AND ROUSSET, F. 1995. GENEPOP (version 1.2): *population genetics software for exact tests and ecumenicism*. *J. Hered.*, **86**, 248–249.
- REID, D. G. 1996. *Systematics and Evolution of Littorina*. The Ray Society, London.
- REID, D. G., RUMBAK, E. AND THOMAS, R. H. 1996. DNA, morphology and fossils: phylogeny and evolutionary rates of the gastropod genus *Littorina*. *Phil. Trans. R. Soc. B.*, **351**, 877–895.
- ROZAS, J. AND ROZAS, R. 1997. DnaSP, version 2.0: *a novel software package for extensive molecular population genetics analysis*. *CABIOS*, **13**, 307–311.
- SCHNEIDER, S., KUEFFER, J. M., ROESSLI, D. AND EXCOFFIER, L. 1997. ARLEQUIN, version 1.1: *a software for population genetic data analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- SCHNEIDER-BROUSSARD, R., FELDER, D. L., CHLAN, C. A. AND NEIGEL, J. E. 1998. Tests of phylogeographic models with nuclear and mitochondrial DNA sequence variation in the stone crabs, *Menippe adina* and *Menippe mercenaria*. *Evolution*, **52**, 1671–1678.
- STRIMMER, K. AND VON HAESLER, A. 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.*, **17**, 964–969.
- TABERLET, P., FUMAGALLI, L., WUST-SAUCY, A.-G. AND COS-SON, J.-F. 1998. Comparative phylogeography and post-glacial colonization routes in Europe. *Mol. Ecol.*, **7**, 453–464.
- THOMPSON, J. D., HIGGINS, D. G. AND GIBSON, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl. Acids Res.*, **22**, 4673–4680.
- WARD, R. D. 1990. Biochemical genetic variation in the genus *Littorina* (Prosobranchia: Mollusca). *Hydrobiologia*, **193**, 53–69.
- WARWICK, T., KNIGHT, A. J. AND WARD, R. D. 1990. Hybridisation in the *Littorina saxatilis* species complex (Prosobranchia: Mollusca). *Hydrobiologia*, **193**, 109–116.
- WATTERSON, G. A. AND GUESS, H. A. 1977. Is the most frequent allele the oldest? *Theor. Pop. Biol.*, **11**, 141–160.
- WILDING, C. S., MILL, P. J. AND GRAHAME, J. 1999. Partial sequence of the mitochondrial genome of *Littorina saxatilis*: relevance to gastropod phylogenetics. *J. Mol. Evol.*, **48**, 348–359.
- ZINK, R. M. AND DITTMANN, D. L. 1993. Gene flow, refugia, and evolution of geographic variation in the song sparrow (*Melospiza melodia*). *Evolution*, **47**, 717–729.