Molecular Ecology (2001) **10**, 2541–2554

Phenotypic evolution and hidden speciation in *Candidula unifasciata* ssp. (Helicellinae, Gastropoda) inferred by 16S variation and quantitative shell traits

MARKUS PFENNINGER* and FRÉDÉRIC MAGNIN†
*Abteilung Ökologie und Evolution, Zoologisches Institut der J.W. Goethe-Universität, Bio-Campus Siesmayerstraße, D-60054 Frankfurt/Main, Germany, †IMEP ESA 6116 Case 461, Faculté des Sciences et Techniques de St. Jerome, 13997 Marseille Cedex 20, France

Abstract

In an effort to link quantitative morphometric information with molecular data on the population level, we have analysed 19 populations of the conchologically variable land snail *Candidula unifasciata* from across the species range for variation in quantitative shell traits and at the mitochondrial 16S ribosomal (r)DNA locus. In genetic analysis, including 21 additional populations, we observed two fundamental haplotype clades with an average pairwise sequence divergence of 0.209 ± 0.009 between clades compared to 0.017 ± 0.012 within clades, suggesting the presence of two different evolutionary lineages. Integrating additional shell material from the Senckenberg Malacological Collection, a highly significant discriminant analysis on the morphological shell traits with fundamental haplotype clades as grouping variable suggested that the less frequent haplotype corresponds to the described subspecies *C. u. rugosiuscula*, which we propose to regard as a distinct species. Both taxa were highly subdivided genetically ($F_{ST} = 0.648$ and 0.777 $P < 0.001$). This was contrasted by the partition of morphological variance, where only 29.6% and 21.9% of the variance were distributed among populations, respectively. In *C. unifasciata*, no significant association between population pairwise F_{ST} estimates and corresponding morphological fixation indices could be detected, indicating independent evolution of the two character sets. Partial least square analysis of environmental factors against shell trait variables in *C. u. unifasciata* revealed significant correlations between environmental factors and certain quantitative shell traits, whose potential adaptational values are discussed.

Keywords: 16S mtDNA, adaptation, Helicellinae, hidden speciation, population structure, quantitative shell variation

Received 17 April 2001; revision received 18 July 2001; accepted 19 July 2001

Introduction

The evolutionary processes shaping the phenotypic variation among and within subdivided natural populations have only recently started to become a focus of research interest (e.g. Bonnin *et al*. 1996; Yang *et al*. 1996; Lynch *et al*. 1999). This is probably due to the multitude of factors shaping phenotypic traits in natural populations. The phenotypic appearance is influenced in a complex fashion by phylogenetic history, genetic drift, gene-flow, selection, environment and developmental processes. The relevant evolutionary processes are supposed to take place in natural populations and consequently should be studied there (Ritland 2000). To gain information about phylogenetic relations and population processes on the species level and beyond, effective neutral genetic markers such as microsatellites, restriction fragment length polymorphisms (RFLPs) or mitochondrial DNA (mtDNA) sequences have been widely used (reviewed in Avise 2000). This information can be used as a background against which the phenotypic variation can be compared in order to gain insight into the processes shaping the phenotypic appearance.

A necessary prerequisite for such a study is a detectable amount of genetic and phenotypic variation among

Correspondence: Dr Markus Pfenninger. Fax: +49 69 798 24910; E-mail: Pfenninger@zoology.uni-frankfurt.de

populations. A strong population structure is most likely found in species with poor dispersal capacity. Land snails with their restricted vagility, therefore, appear very suited organisms for the study on the issue of phenotypic evolution. Phenotypic variability of shell traits within and among populations of geographically subdivided taxa is often observed in land snails. While qualitative differences in shell colour and banding within Helicid snail species is well studied in conjunction with molecular markers (e.g. Arter 1990; Davison & Clarke 2000), quantitative variation has not yet been explored and causes much uncertainty about the evolutionary and taxonomical relevance (Kerney & Cameron 1979). Substantial variation in shell morphology can be observed in the small land snail *Candidula unifasciata* ssp. (Poiret 1801) throughout the species range (Gittenberger 1993). This has led to the description of at least four subspecies, based on differences in shell traits: *C. u. unifasciata* (Poiret 1801), *C. u. rugosiuscula* (Michaud 1831), *C. u. soosiana* (Wagner 1933) and *C. u. acosmia* (Bourguignat 1882). Studying the original descriptions of the subspecies (Michaud 1831; Bourguignat 1882; Wagner 1933), it becomes clear that the authors refer to quantitative rather than to qualitative differences to distinguish the taxa. Because this makes unequivocal identification difficult, there is much confusion about the distribution of each subspecies, though most authors agree that there is a 'geographical component' in the range of each type (Gittenberger 1993).

In this study, we use the genetic information from mitochondrial 16S haplotype analysis as a background to study quantitative morphologic differences at the population level throughout the core species range of all described *C. unifasciata* taxa. In particular, we focus on three questions: (i) does the observed phenotypic variation correspond to different evolutionary lineages? (ii) How is the phenotypic variation distributed within and among populations in relation to the 16S variation? And (iii) Can we identify environmental variables that co-vary with population differences in morphology?

Materials and methods

Samples and DNA extraction

Samples for this study were collected from 40 locations, comprising the core species range of *Candidula unifasciata* (Kerney & Cameron 1979). Morphologic analyses were performed on individuals from a subset of 19 populations (Fig. 1 and Table 1). Within these populations, we tried to analyse all available samples both morphologically and genetically. However, for some individuals the 16S analysis failed at some point [DNA extraction, polymerase chain reaction (PCR) or sequencing] and some individuals were subadults and not suitable for morphologic analysis. The overlap of both data sets was therefore not complete.

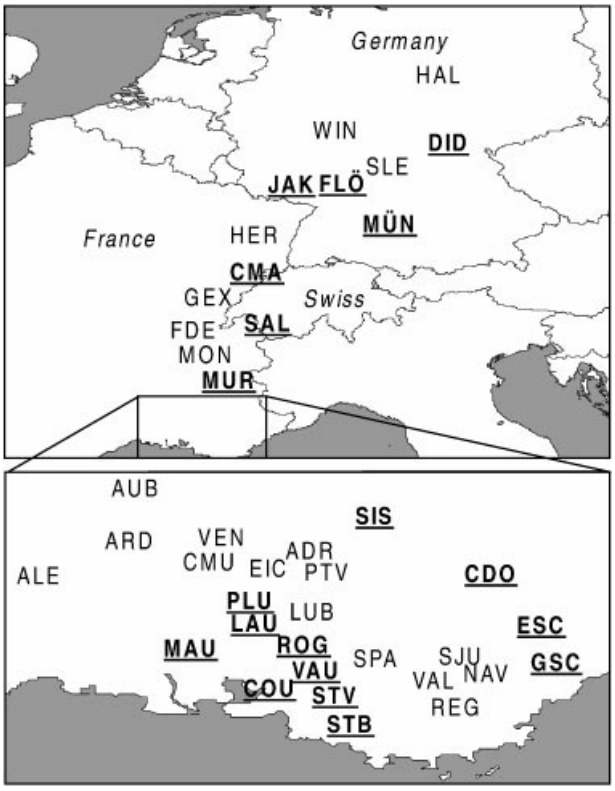

Fig. 1 Geographical location of sampled populations. The framed part of the upper map corresponds to the lower map. Populations where both morphological and genetic analysis was performed are bold and underlined.

Individuals destined for morphological analyses were photographed (see below) prior to DNA extraction. In total, 220 individuals were crushed with their shells in 10% w/v laundry detergent solution for storage at room temperature and tissue digestion following the protocol of Bahl & Pfenninger (1996). Additionally, we tried to extract DNA from museum specimens that were apparently sampled alive and contained dried remains of tissue. To preserve the shells, they were soaked in laundry detergent solution overnight and treated for 5 min with supersonic to detach the tissue from the shell. All samples were shaken for 24 h at 37 °C in the laboratory prior to phenol/chloroform extraction of total DNA following a standard protocol (Sambrook *et al*. 1989).

Amplification of 16S ribosomal DNA, sequencing and alignment

The 16S target-DNA was amplified for all samples by PCR with universal primers of the sequence 16S1 5′ > CGCAGTACTCTGACTGTGC < 3′ and 16S2 5′ > GTCCGGTTTGAACTCAGATC < 3′. Amplification was performed with

Table 1 (a) Locations, regions and countries sampled for *Candidula unifasciata* ssp., ordered from South to North, their abbreviations used, geographical position, number of individuals sampled for 16S rDNA (N_{gen}) and morphologic analysis (N_{morph}), respectively. (b) *C. unifasciata* ssp. Samples from the Malacological collection of the Senckenberg Museum/Frankfurt am Main, their origin as far as indicated on the collection label, the reference number of the collection, number of specimen used in morphologic and 16S rDNA analysis

(a) Location	Region	Abbr.	Position	N_{gen}	N_{morph}
1 Sainte Baume	Bouches du Rhône	STB	43°20′37″ N 05°38′87″ E	4	13
2 Sainte Victoire	Bouches du Rhône	STV	43°32′18″ N 05°34′66″ E	3	8
3 Vauvenargue	Bouches du Rhône	VAU	43°33′14″ N 05°35′06″ E	5	9
4 Coudoux	Bouches du Rhône	COU	43°33′89″ N 05°14′58″ E	4	13
5 Regasse	Var	REG	43°39′30″ N 06°08′00″ E	2	—
6 Saint Paul	Var	SPA	43°40′30″ N 05°41′15″ E	5	—
7 Grotte de Saint Cezaire	Alpes Maritimes	GSC	43°40′82″ N 06°48′57″ E	10	18
8 Rognes	Bouches du Rhône	ROG	43°41′18″ N 05°18′78″ E	5	14
9 Maussane	Bouches du Rhône	MAU	43°41′71″ N 04°50′55″ E	4	4
10 Escragnolles	Alpes Maritimes	ESC	43°43′78″ N 06°47′58″ E	11	19
11 Lauris	Vaucluse	LAU	43°45′35″ N 05°15′31″ E	5	12
12 Grand Luberon	Vaucluse	LUB	43°47′55″ N 05°26′10″ E	6	—
13 Petit Luberon	Vaucluse	PLU	43°48′18″ N 05°17′02″ E	5	21
14 Valensole	Alpes de Hte Provence	VAL	43°49′40″ N 05°58′50″ E	4	—
15 Naverre	Alpes de Hte Provence	NAV	43°52′55″ N 06°13′50″ E	5	—
16 Saint Jurs	Alpes de Hte Provence	SJU	43°53′55″ N 06°12′40″ E	8	—
17 Col d'Orme	Alpes de Hte Provence	CDO	43°54′40″ N 06°12′36″ E	2	5
18 Col de Murs	Vaucluse	CMU	43°58′50″ N 05°13′45″ E	5	—
19 Plateau de Vaucluse 1	Vaucluse	PTV	44°00′10″ N 05°32′05″ E	5	—
20 Les Eicharettes	Vaucluse	EIC	44°04′50″ N 05°26′45″ E	6	—
21 Plateau de Vaucluse 7	Vaucluse	ADR	44°04′50″ N 05°26′45″ E	5	—
22 Le Ventouret	Vaucluse	VEN	44°07′45″ N 05°22′20″ E	5	—
23 Alès	Gard	ALE	44°07′45″ N 04°12′40″ E	4	—
24 Sisteron	Drôme	SIS	44°12′15″ N 05°54′39″ E	4	12
25 Saint Martin d'Ardèche	Ardèche	ARD	44°18′30″ N 04°33′22″ E	5	—
26 Aubenas	Ardèche	AUB	44°35′30″ N 04°25′10″ E	1	—
27 La Mure	Isère	MUR	44°52′55″ N	15	19

Table 1 *Continued*

(a) Location	Region	Abbr.	Position	N_{gen}	N_{morph}
			05°50′62″ E		
28 Montalieu	Isère	MON	45°08′30″ N 05°23′25″ E	5	—
29 Mons Salève	Haute Savoie	SAL	46°04′40″ N 06°07′30″ E	6	6
30 Fort d'Ecluse	Ain	FDE	46°07′25″ N 05°53′15″ E	5	—
31 Gex	Ain	GEX	46°21′05″ N 06°03′00″ E	5	—
32 Col de Marchaidruz	Fribourg	CMA	46°32′62″ N 06°15′15″ E	6	9
33 Herticourt	Doubs	HER	47°15′30″ N 06°45′10″ E	5	—
34 Münsingen	Baden-Württemberg	MÜN	48°25′00″ N 09°30′00″ E	10	4
35 Jakobsberg	Rheinland-Pfalz	JAK	49°75′71″ N 07°59′08″ E	3	14
36 Flörsheim	Hessen	FLÖ	50°00′01″ N 08°23′10″ E	7	9
37 Schlüchtern/Elm	Hessen	SLE	50°21′40″ N 09°33′20″ E	8	—
38 Winterscheid	Hessen	WIN	50°55′85″ N 09°01′91″ E	4	7
39 Dierdorf	Thüringen	DID	51°10′66″ N 10°17′34″ E	5	11
40 Halle	Sachsen-Anhalt	HAL	51°42′35″ N 11°53′15″ E	5	—

(b) Taxon	Location	Region / Country	SNG reference no.	N_{gen}	N_{morph}
C. u. unifasciata	Grasse	Alpes Maritimes/F	60827	—	6
C. u. unifasciata	?	Bayern / D	60737	—	6
C. u. acosmia	Carpentras	Vaucluse / F	97053	1	6
C. u. soosiana	Budapest	Hungary	217549	1	7
C. u. soosiana	Budapest	Hungary	278316	—	6
C. u. rugosiuscula	Cariès	Var/F	97130	—	5
C. u. rugosiuscula	Castres	Provence/F	60212	—	5
C. u. rugosiuscula	Bois de Pires	Provence/F	97134	—	7

Boehringer *Taq*-polymerase in 12.5 μL total reaction volume with standard reaction conditions. Samples were amplified for 10 cycles (92 °C for 50 s, 44 °C for 50 s and 72 °C for 40 s) and 36 cycles (92 °C for 30 s, 48 °C for 40 s and 72 °C for 40 s) after initial incubation of 90 °C for 2 min 30 s. Both strands of the purified amplification products were cycle-sequenced with the Perkin Elmer Taq DyeDeoxy™ Terminator Cycle Sequencing Kit following the supplier's protocol and read automatically on the ABI Prism 377® sequencing device from the same manufacturer. Sequences were deposited in GenBank (Accession nos AF407841–AF408058) and were aligned manually after an initial alignment with the CLUSTAL option (Thompson *et al*. 1994) in the computer program SEQUENCENAVIGATOR (Perkin Elmer, Applied Biosystems).

16S Pairwise distance statistics and maximum likelihood analysis

The 16S sequences were collapsed to haplotypes with the COLLAPSE 1.0 tool provided by David Posada (http://bioag.byu.edu/zoology/crandall_lab). For the maximum likelihood (ML) analyses, the best-fit model of nucleotide substitution (TrNef + G) was selected using the hierarchy of likelihood ratio tests implemented in MODELTEST 3.0 (Posada & Crandall 1998). ML heuristic search was conducted with 100 random sequence addition replicates. Nodal support was estimated using the bootstrap approach (Felsenstein 1985) with 1000 replicates. ML analyses were performed with PAUP 4.03b (Swofford 1998). A pairwise sequence

divergence frequency distribution, including 16S rDNA sequences from the Helicellinae species *Trochoidea geyeri* (Soós 1926) (79 sequences, Accession nos AF407751–AF407829) and *C. gigaxii* (L. Pfeiffer 1850) (10 sequences Accession nos AF407830–AF407840), was calculated. Analysis of Molecular Variance (AMOVA) was performed on a matrix of pairwise Euclidean differences between haplotypes, grouped after population of origin. The calculations were performed using AMOVA 1.55 (Excoffier *et al*. 1992).

Morphological analysis

Morphological analysis was performed on adult individuals of a subset from 19 locations (226 individuals) and eight samples (38 individuals), from the Malacological Collection of the Senckenberg-Museum Frankfurt am Main, of the four presumed subspecies (Table 1). The data set for morphological and genetic analysis overlapped in 105 individuals for sampled populations and two individuals for the museum samples. The locations of first descriptions for the subspecies *C. u. unifasciata* and *C. u. rugosiuscula* are not very well described in literature, but the approximate areas were included in our sampling scheme (see Table 1). A lectotype locality exists for *C. u. acosmia* (Gittenberger 1993) that was not sampled. However, in the original description (Bourguignat 1882), neither type locality nor distribution range were given. *C. u. soosiana* was first described from Budapest/Hungary (Wagner 1933) and is represented by two Museum samples. Individuals were photographed through a Zeiss ZV8 binocular with 10× enlargement from the front and 30× from above, always including a ruler on the image. The paper prints were digitalized with a resolution of 300 pixel / inch.

Variables for the morphological analysis were chosen based on the original descriptions of the presumed subspecies of *C. unifasciata* (Michaud 1831; Bourguignat 1882; Wagner 1933; Gittenberger 1993). Eight variables representing three character sets were measured by digital image analysis from the individuals. Shell sculpture traits as the first character set were measured through *rib-spacing* as an average distance between ribs, *coarseness* as an average distance between a base line and perimeter of ribs, and *regularity of ribs* as the coefficient of variation of inter-rib distances. Size independent shell shape (*shape 1*, *2* and *3*) was assessed through the use of the first three relative warp scores of a Thin Plate Spline (TPS) analysis (Rohlf 1995). This geometrical analysis of shape was based on 55 landmarks applied to an electronic image of the individual. The *height* and *breadth* measurements of the shell were taken.

These standardized shell morphology variables were used to perform a principal component analysis (PCA) and discriminant analysis (DA), with STATISTICA (StatSoft, Inc. 1995). The mitochondrial DNA (mtDNA) type (A or B) of the respective individual resulting from 16S analysis was applied as grouping variable in DA. Museum specimens and not genetically typed individuals were then classified a posteriori with equal a priori probability to belong to one of the respective groups.

Euclidean squared distance matrices between all individuals were calculated for all standardized shell variables together and for each of the 10 variables separately. These matrices were used to perform an analysis of morphological variance in the same way as described above for the distance matrix derived from haplotypes. Standard deviations and confidence intervals of the *M*-statistics were obtained by bootstrapping 100 times over populations. Note, however, that these derived *F*-statistics (named here M_{ST} for distinction) cannot directly be compared to Spitze's (1993) Q_{ST} and similar coefficients (e.g. Yang *et al*. 1996), because we have no information about the heritability of the measured traits. Estimates of heritability of quantitative traits were obtained in these studies from elaborated breeding designs or a priori information about relatedness. Such information was not available for *C. unifasciata*. Hence, the presented population fixation indices comprise not only their *genetic* component, but are also indications of the *total* morphologic variance.

Associations between genetic, morphological and geographical distance between populations were explored through Mantel-tests on the pairwise population fixation indices derived from AMOVA analysis to test for parallel evolution between genetic and morphological data and isolation by distance, respectively.

Partial least square (PLS) analysis was performed using the PLS option of the package NTSYS-pc version 2.2f to assess the correlations between the average population values of quantitative shell traits on the one hand and environmental variables for each population on the other.

Environmental variables

Information about long-term climatical conditions at the sites sampled for morphological analysis was obtained from the web sites of MeteoFrance, Deutscher Wetterdienst, and Bessemoulin (1989). We used annual mean temperature, annual mean precipitation, and mean precipitation during the months of June, July and August (summer precipitation), and altitude, latitude and longitude as variables for each sampling site. The precipitation during the summer months was considered to be important, because most of the shell growth is achieved during this time.

Results

Haplotype analyses

The 220 sequences were collapsed to 60 haplotypes (see Table 2), of which 53 (88%) were present in individuals used

Table 2 Distribution of *C. unifasciata* haplotypes (columns) at each sampled location or museum sample (rows)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
STB																																																	1	1	1					1		
STV																					2																					1																						
VAU		2																			2									1																																
MAU		3																																										1																				
COU																																																						1	1	1	1						
SPA																					3	1	1																																										
REG																	1	1																																																
VAL																					1				3																																									
SJU															3									5																																									
NAV															4	1																																																		
CDO																																									2																							
GSC								8																							1							1																								
ROG																																																																		
ESC										1															7		1																																						
SIS	2	2																																																														
VEN		4		1																																																											
EIC		3		3																																																											
PTV		2		3																																																											
ADR				3							1	1																																																				
LUB		2		1																						1	1	1																																				
CMU				1	1	1	1	1																																																									
PLU				1																												2	1	1																														
LAU																					1	1														1	1	1																										
ALE		4																																																															
ARD		5																																																														
AUB		1																																																														
MUR										8									3						2																																									
MON		3										2																																																				
GEX				5																																																												
FDE		4							1																																																							
SAL		2		2															1	1																																												
HER		5																																																													
CMA		6																																																														
MÜN		8												2																																																	
FLÖ		7																																																													
SLE		4																							3					1																																
WIN		1	3																																																											
HAL		1	4																																																											
JAK		3																																																													
DID		1	4																																																											
aco				1																																																											
soos																																																																	

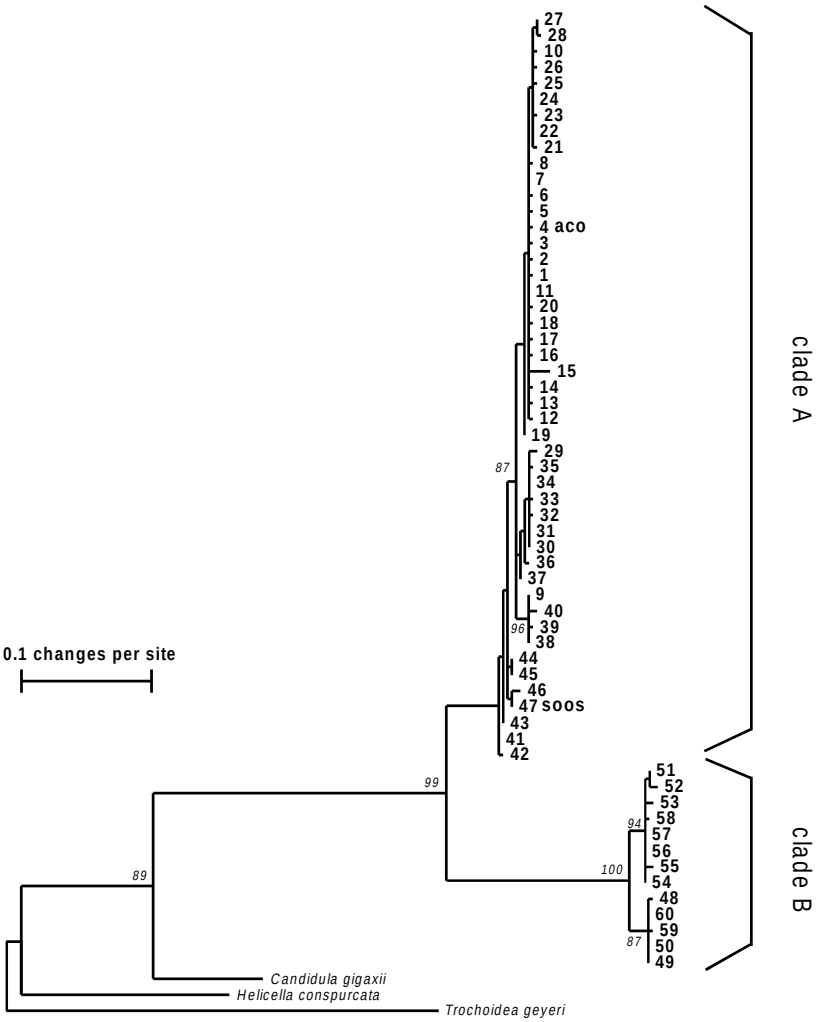

Fig. 2 Maximum likelihood tree of 16S haplotypes. Branch tips are labelled with the haplotype number. Italicised numbers indicate percentage bootstrap support of nodes from 1000 replicates; only bootstrap values in excess of 80% are given. The position of the sequence obtained from the museum specimen classified as *Candidula unifasciata acosmia* was labelled with 'aco', from *C. u. soosiana* with 'soos'. The two fundamental haplotype clades have been designated as clade A and clade B, respectively. Clade B consists exclusively of haplotypes from locations COU, ROG and STB. Haplotypes from these locations are not found in the clade A and vice versa.

for the morphologic analysis. Therefore, our morphological analyses comprised a large part of the genetic variation present in *Candidula unifasciata* ssp. Base changes and indels were observed exclusively in regions known to be variable in molluscs (Lydeard *et al*. 2000). Accidental amplification of nuclear copies of 16S target fragments is therefore unlikely. The average pairwise sequence divergence between haplotypes was 0.038 (0.065) (mean ± SD) changes per site. The topology of the ML tree strongly suggested the presence of two fundamental haplotype clades named A and B (see Fig. 2). Analysing each clade separately, the sequence divergence between haplotype clades A and B was 0.209 (0.009) changes per site compared to 0.014 (0.010) and 0.022 (0.016) changes within clade A and clade B, respectively. The 16S sequences derived from the museum shells assigned to the subspecies *C. u. acosmia* and *C. u. soosiana* clustered both in haplotype clade A. Clade B haplotypes occurred only in populations COU, STB and ROG, where no clade A haplotype was found.

The frequency distribution of pairwise sequence divergence estimates including sequences from the Helicellinae species *C. gigaxii* and *Trochoidea geyeri*, showed that the level of divergence between sequences belonging to clades A and B correspond to between rather than within species (Fig. 3).

Morphological analyses

The coarseness of the ribs was 0.036 (0.032) mm (mean ± SD), their average spacing 0.147 (0.057) mm and the regularity 0.274 (0.105). The first relative warp score of TPS analysis turned out to be a descriptor of a depressed vs. globular shell shape (24.95% of total shape variability), the second opposed roundish vs. elliptical apertures (17.87%) and the third score represented large apertures and a round periphery vs. a small aperture with an angular periphery of the last whorl (13.93%). The values for the shell size parameters were 3.828 (0.554) mm for the height, and 6.333 (0.854) mm for breadth.

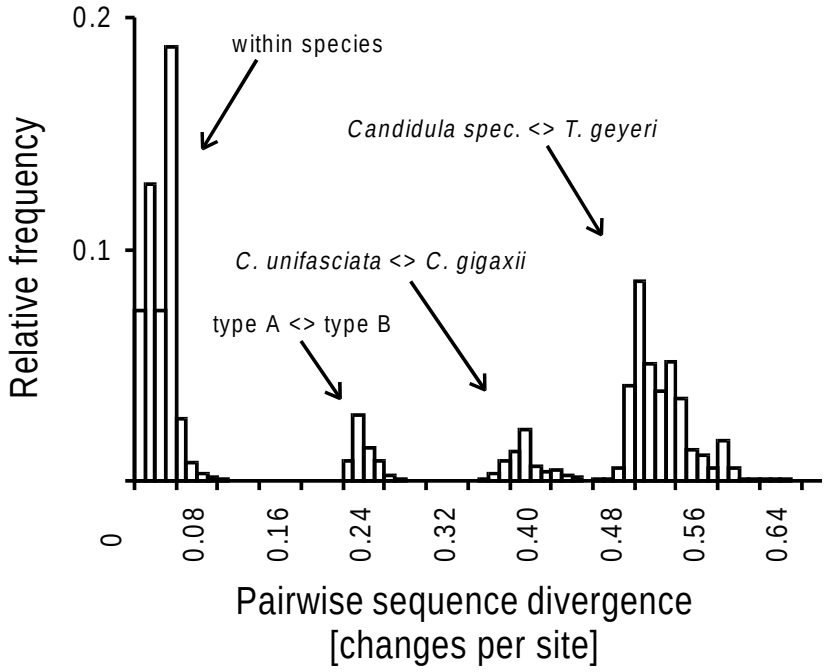

Fig. 3 Frequency distributions of pairwise sequence divergence between 16S rDNA sequences of *Candidula unifasciata*, *C. gigaxii* and *Trochoidea geyeri*.

The first two axes of the PCA on all eight standardized variables accounted for 26.5% and 24.1% of total conchological variation. The plot of the factor scores on the first two PCA axes shows that the morph-spaces of the museum specimens assigned to *C. u. unifasciata*, *C. u. acosmia* and *C. u. soosiana* overlap, whereas *C. u. rugosiuscula* forms a distinct cluster (Fig. 4). However, the morph-spaces of all museum taxa are interconnected or enclosed by sampled individuals, thus forming a more or less continuous range of conchological variation.

DA

Using the observed fundamental haplotypes A or B as grouping variables in DA yielded a highly significant discrimination model (Wilks' $\lambda = 0.168$, appr. $F_{[8,107]} = 65.58$, $P < 0.0000$). All eight variables were retained in the model, but the single root extracted discriminated the groups mainly in correlation to the coarseness of the shell. The posterior classifications of all individuals to clades in the analysis all had a probability greater than 95%. The result of the DA is in concordance with the results from haplotype analysis insofar as no morphotype associated with haplotype clade A was found in a population with phenotypes associated with haplotype clade B and vice versa. The haplotype clade B morphs are found only in the populations COU, ROG and STB in the south of the investigated area. Representative individuals of each clade are depicted in Fig. 4.

The museum specimens classified as *C. u. unifasciata*, *C. u. acosmia* and *C. u. soosiana* were all predicted to belong to the phenotype associated with haplotype clade A; the specimen named *C. u. rugosiuscula* corresponded to the phenotype exhibited by individuals of clade B.

Partition of molecular and morphological variance

Most of the 16S variation (89.47%) is partitioned between clade A and clade B. There is moderate variation (6.92%) between locations within the types, but the populations tend to be rather uniform (3.61%). In contrast, the morphological variance component between the clades accounts for only 18.04% of the total variance. Most of the morphological variation (70.63%) is due to differences between individuals within populations, leaving 11.33% of conchological variation for differences between locations within type A and B. All variance components were significantly different from zero.

In clade A alone, the F_{ST} between populations is as high as 0.648 (0.038), indicating high phylogenetic relatedness of haplotypes within and very little gene flow among populations. This is not matched by overall phenotypic similarity of the shells within populations, where M_{ST} was estimated to be 0.292 (0.042). This difference is similar in type B, where the F_{ST} was 0.777 and the corresponding morphological fixation index 0.219. All the above-mentioned variance components were highly significantly different from zero; standard deviations for clade B are not given because bootstrapping populations makes no sense with only three populations.

When looking at the measured traits separately, the picture becomes more differentiated in haplotype clade A. Except for the regularity of rib spacing, all fixation indices were significantly different from zero, indicating population

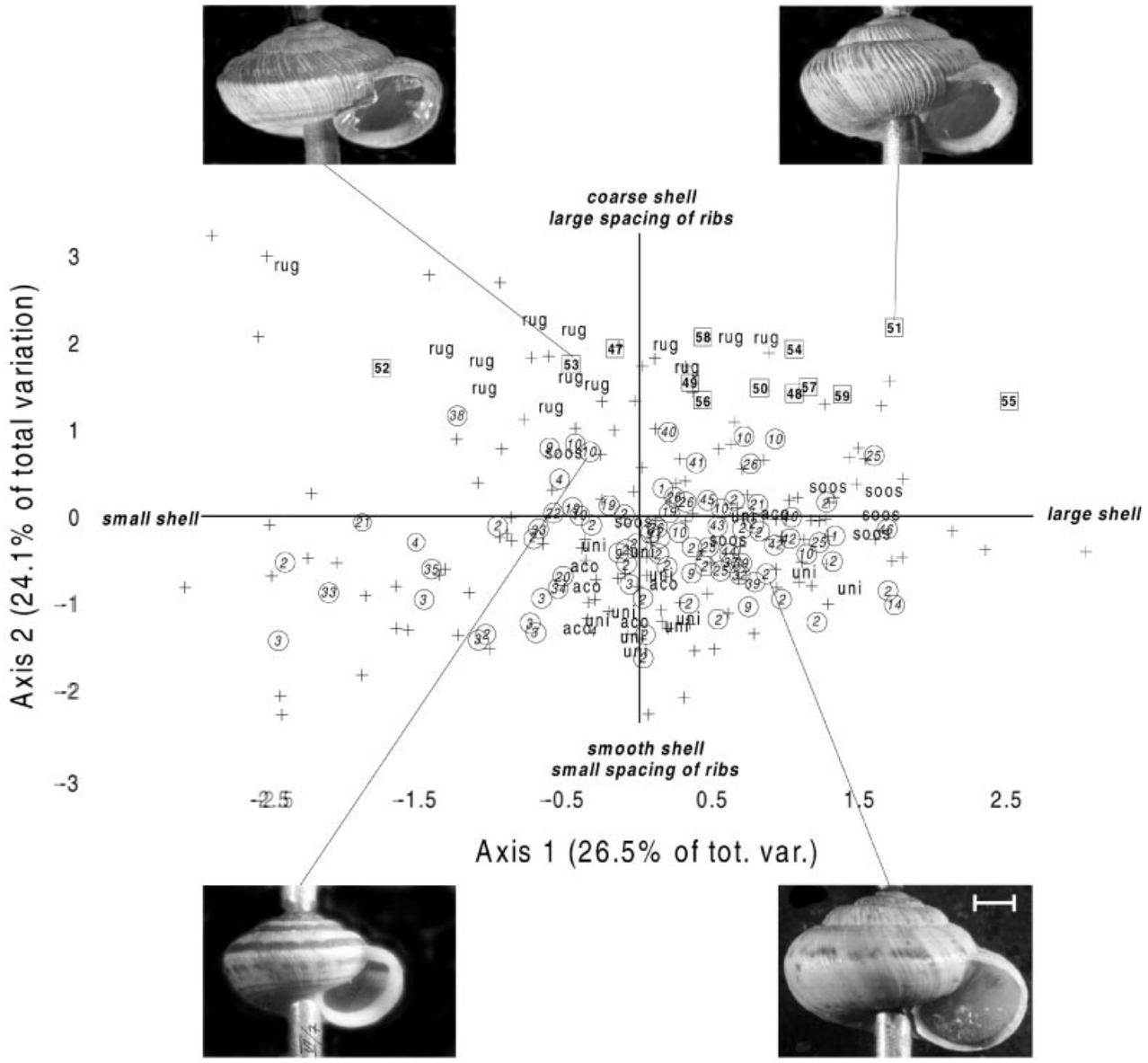

Fig. 4 Scatterplot factor scores on the first two axes of Principal component analysis on shell morphological variables. Axis 1 accounts for 26.5%, axis 2 for 24.1% of total variance. Museum specimen are labelled according to the Senckenberg Malacological Collection, where uni = *Candidula unifasciata unifasciata*, aco = *C. u. acosmia*, rug = *C. u. rugosiuscula* and soos = *C. u. soosiana*. Sampled individuals not typed genetically are represented by a cross, individuals that were genetically typed with their haplotype number (see Fig. 2). Haplotypes from clade A are encircled and from clade B in a square, respectively. Two individuals of each haplotype clade are shown exemplarily, representing the gradient of variation in shell coarseness along axis 2. The scale on each photograph represents 1 mm.

differentiation. The range of M_{ST} estimates lied between 0.150 (0.062) for *shape2* and 0.691 (0.078) for the *coarseness* of the shell (Table 3). Because regularity of rib spacing was not significantly different between populations, this variate was omitted from subsequent analyses.

Mantel tests

The Mantel test for association between the population pairwise F_{ST} and M_{ST} was nonsignificant in all cases. There was no correlation between geographical distance separating populations and pairwise F_{ST} either. By contrast, geographical distance proved to co-vary significantly ($P < 0.05$) with the overall phenotypic similarity, as well as with the variable *height* (Table 3).

PLS analysis

PLS analysis detected a significant correlation between the average population values of shell traits and environmental

Table 3 Fixation indices from analysis of molecular and morphologic variance and correlation coefficients from Mantel tests between population pairwise M_{ST}, F_{ST} and geographical distance between populations for haplotype clade A (*Candidula unifasciata*)

Variable	F_{ST}; M_{ST} ± (s.d.)	P	r (F_{ST})	r (geo. dist.)
16S haplotypes	0.648 (0.038)	< 0.001	—	0.131
Overall phenotypic similarity	0.296 (0.039)	< 0.001	0.034	0.281*
Shell shape				
Shape1 (depressed <> globular shell)	0.195 (0.095)	< 0.001	−0.059	0.164
Shape2 (round <> elliptical aperture)	0.150 (0.062)	< 0.001	−0.214	−0.009
Shape3 (large <> small aperture)	0.381 (0.112)	< 0.001	0.134	0.225
Shell size				
Breadth	0.399 (0.088)	< 0.001	−0.030	0.116
Height	0.446 (0.071)	< 0.001	0.024	0.244*
Shell sculpture				
Spacing of ribs	0.338 (0.048)	< 0.001	0.083	0.029
Coarseness of shell	0.691 (0.078)	< 0.001	−0.067	0.021
Regularity of rib spacing	0.053 (–)	0.057	−0.038	−0.032

s.d. = standard deviation obtained from 100 bootstraps, P = probability of obtaining a random F_{ST} equal or greater than the observed estimated from 9999 randomizations, r = Pearson's moment product correlation coefficient between pairwise among population M_{ST} for all variables and F_{ST}, and geographical distance between populations. *Indicates a significant correlation at the 5% level.

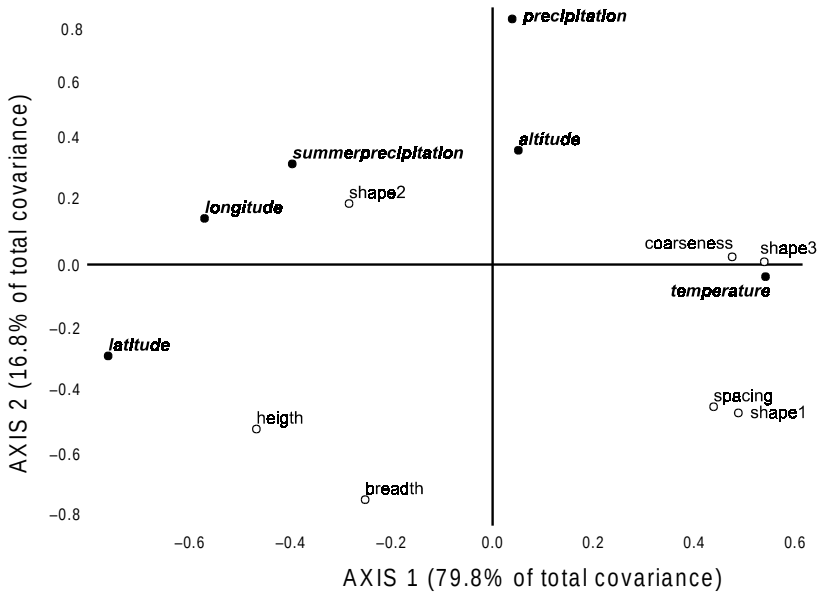

Fig. 5 Partial least square analysis of mean shell variables against environmental variables for all populations belonging to haplotype clade A (*Candidula unifasciata*).

variables for the respective sampling site (approximately $\chi^2_{[40]} = 60.8$, $P = 0.028$). The first two canonical variates accounted for 79.8% and 16.8% of the total squared covariance between the character sets. The first canonical axis showed that populations with a high annual mean temperature lying in the south of the species range tend to be composed of individuals showing a coarse shell surface with largely spaced ribs and a depressed shell shape with a small aperture. The second axis revealed a negative correlation between the annual amount of precipitation and shell trait variables measuring the size of the specimen (see Fig. 5).

Discussion

Two genetically and morphologically distinct evolutionary lineages in Candidula unifasciata *ssp*

Substantial morphologic and genetic variation within and among populations was detected in the present survey over the greatest part of the range of the taxon that was described as *C. unifasciata* ssp. The PCA on the morphometric data showed that the field samples used in the present study comprised most of the phenotypic variance described in this taxon, as evidenced by the

inclusion of samples from the Malacological Collection of the Senckenberg Museum (see Fig. 4). Even though the PCA revealed a more or less continuous distribution of individuals in morph-space, the two highly divergent haplotype clades in 16S analysis indicated the presence of two independently evolving lineages. The sequence divergence between the clades was almost as high as between other Helicellinae species and was in the magnitude of those reported between genera of other gastropod species (Koufopanou *et al*. 1999; Medina & Walsh 2000; but see Thomaz *et al*. 1996).

Despite of an almost continuous morph-space when looking at the overall phenotypic similarity, DA yielded a highly significant discrimination model with the haplotype clade as classification variable (Table 3). This shows that the haplotype clade of an individual is a good predictor of its morphology. Moreover, DA predicted a posteriori the clade membership of all individuals not screened for 16S variation with probabilities greater than 95%. This is in concordance with results from the analysis of variance that detected a significant variance component among the clades that accounted for approximately 20% of the total morphological variance.

The presence of two highly divergent haplotype clades or evolutionary lineages is therefore in concordance with morphological differences (Table 3, Fig. 4). This led us to the conclusion that there are two separately evolving lineages in *C. unifasciata* ssp. This view is strengthened by the observation that clade B, found only in populations ROG, COU and STB from the southeast of France, occurs in parapatry to populations of clade A (Fig. 1), but was not found together in the same population (Table 2). This obvious absence of gene-flow between neighbouring populations of clade A and clade B suggested reproductive isolation of the clades. Additionally, initial surveys with allozymes and random amplified polymorphic DNAs indeed showed divergence in the nuclear genome as well (data not shown). The average sequence divergence among clade A and B equalled 0.209 changes per site. Even if we consider a very high mutation rate of 10% per million years, we have to assume that the clades diverged at least 2 million years ago. Most species concepts (reviewed in Hull 1997) would recognize clade B as separate species in the light of the presented evidence. However, additional evidence, like crossbreeding experiments and analyses of habitat differences, are needed to confirm the species status of the inferred clades.

Taxonomic considerations

Having established the presence of two genetically and morphologically distinct evolutionary lineages, we wanted to know if there were arguments that allowed assigning an existing taxonomic name to them. DA predicted that the rarer haplotype clade B corresponded to museum specimen named by the collectors *C. u. rugosiuscula* as described by Michaud (1831). More evidence for the association of the identified evolutionary lineage (clade B) with this taxon arises from the correspondence of the morphotype with the original description and the fact that populations of clade B were found exclusively in the type area ('around Aix-en-Provence', Michaud 1831). We will therefore refer to clade B individuals as *C. rugosiuscula* hereafter. Specimen of each population have been deposited in the Senckenberg-Museum, Frankfurt/Germany (Collection nos 322884–322886).

The museum specimens assigned to *C. u. unifasciata*, *acosmia* and *soosiana* have overlapping morph spaces in PCA, which proposes that either the traits chosen by the original describers did not define statistically distinguishable entities in the first place or that they were not precise enough to allow successive collectors to distinguish them unequivocally. Corroboration of this view arises from a DA that failed to yield a significant morphologic discrimination model while using the taxonomic designation of the museum specimens as classification variable (data not shown). In addition, neither PCA nor 16S analysis suggested that there were more than two objectively distinguishable groups of individuals present in our sample. The 16S sequences obtained from museum shell material assigned to *C. u. soosiana* and *C. u. acosmia* belonged both to clade A and were, in case of the latter individual, identical to the widespread haplotype 4 (see Table 2). However, the collectors of the museum material might have misinterpreted their findings and because the holotype locality for *C. acosmia* is unknown and the lectotype locality was not sampled, we cannot exclude that more than the two identified evolutionary lineages exist in the southeast of France.

We will refer to all individuals with haplotypes of clade A as *C. unifasciata*. All sequences from the 20 populations that were not screened for quantitative shell variation, because no shell material was available, belonged to haplotype clade A. Since the haplotype clade has proven to be a good predictor for the two morphological types, it was likely that these populations corresponded equally to *C. unifasciata*. This confirmed our subjective impression of the spatial distribution of this morphotype during the sampling.

The assignment of existing taxonomic names to genetically identified evolutionary lineages is difficult and arbitrary to a certain degree, especially when holotype material for morphological comparison is not available. However, additional support for our taxonomical interpretation of the presented results arises from Kerney & Cameron (1999), where *C. u. rugosiuscula* is described on the basis of shell characteristics as only a subspecies to *C. unifasciata*. The description closely matches the morphotype associated with clade B.

Associations between 16S haplotypes and morphological variability within C. unifasciata

Portions of the molecular and phenotypic variation could be shown to be due to an obviously long lasting separation of different evolutionary lineages that we can regard as separate evolutionary entities *C. unifasciata* and *C. rugosiuscula*. As this was a rather unexpected result, we could not adjust the sampling scheme towards the inclusion of more populations of *C. rugosiuscula* into this study. This lack of data prevented us from exploring further the spatial distribution of phenotypic variation in this taxon and to compare it to the organization in *C. unifasciata*. An initial survey of the geographical distribution of *C. rugosiuscula* suggests however, that there are differences in habitat preference between the two lineages. In the area of parapatric occurrence, *C. rugosiuscula* can be found in more Mediterranean lowland habitats, whereas *C. unifasciata* is more frequently found on mountainous grasslands. We are working on a study to quantify these differences.

The sequence variation exhibited by mitochondrial markers like the 16S rRNA gene is presumably evolving in a selectively neutral fashion. The spatial distribution of this variation is therefore shaped mainly by recurrent population processes like neutral genetic drift and gene flow as well as singular historic events like fragmentation of populations or range expansions (Avise 1994). These issues have been addressed in a phylogeographic framework elsewhere; here we wanted to study how the phenotypic population structure relates to the background pattern of selectively neutral evolution at the 16S locus.

The results in Table 3 show that the M_{ST} estimates for quantitative shell traits were significantly smaller than the corresponding F_{ST}, except for the coarseness of the shell, which was indistinguishable from the latter. This shows that the populations did not diverge phenotypically to a level that should be expected from the estimation of neutral drift among populations. If this variation is heritable, the large proportion of phenotypic variation within populations of *C. unifasciata* could be maintained by a fine-scaled spatial population structure within land snail populations (e.g. Bahl *et al.* 1996; Pfenninger *et al*. 1996; Arnaud *et al*. 1999) that tends to maintain genetic diversity compared to an equal sized panmictic population (Lacy 1987). Phenotypic plasticity, common in land snails (Goodfriend 1986), could be another explication for the observed level of intrapopulation phenotypic variation.

As mentioned above, the presented phenotypic fixation indices termed M_{ST} cannot be used to infer directly the action or direction of natural selection, because they are obtained from the total morphologic variance of the quantitative traits and not only their genetic component (see Podolsky & Holtsford 1995). This has the consequence that we cannot distinguish between the action of natural selection, direct influences of the environment and individual random effects on the exhibited phenotype. The presented estimates, nevertheless, put an upper limit to the possible genetic contribution to the phenotypic among-population variances. There are several approaches to estimate the relative contribution in natural populations. But they are only feasible with elaborated breeding designs (Spitze 1993; Lynch *et al*. 1999) or independent estimates of within population relatedness, using nuclear markers (Yang *et al*. 1996) to estimate the heritability of the investigated traits. Both approaches were, however, not practicable here because *C. unifasciata* does not breed in the laboratory and information about the relatedness of individuals in populations is not available. Recently, Ritland (2000) reviewed approaches to jointly estimate both heritability and among population genetic variances of quantitative traits with nuclear markers that promise much potential for future studies.

The results from the Mantel-tests indicated that the null hypothesis of no association between population pairwise F_{ST} and among population differences in shell traits could not be rejected within *C. unifasciata* (Table 3). This suggested that the observed phenotypic differences have not evolved in concordance with the mitochondrial genome and resulting phenotypic population structure does not reflect the phylogenetic history of the populations. The significant phenotypic divergence among populations must therefore have other explanations.

Associations between the environment and morphological variability within C. unifasciata *(haplotype clade A)*

The results from PLS analysis showed that differences between populations in some traits co-varied significantly with long-term climatic conditions and information about the geographical position of the sampling site. Axis 1 in Fig. 5 shows that shells found in climatically warmer sampling sites in the south tend to have a coarser shell, larger spaced ribs, smaller apertures and more depressed shells. This association allows formulating hypotheses about the causes for these characteristics: in areas with a more Mediterranean climate, there are, despite of considerable mean annual precipitation, prolonged phases of drought that prevents activity of the snails in order to avoid desiccation. Like most other Helicidae (McQuaid *et al*. 1979), *C. unifasciata* manages to survive such periods aestivating at the roots of bushes, under stones or attached to the vegetation. A small aperture, minimizing the area of exposed surface, and a flat shell, allowing to penetrate deeper into the vegetation or under stones, could thus constitute a selective advantage by minimizing the loss of humidity under water stress conditions, as discussed in Goodfriend (1986). Association of shell shape with

physical habitat conditions was found as well by Johnson & Black (2000) in the littorine snail *Bembicium vittatum*. Because of the independence of gene flow patterns and the high heritability of shell shape in this species, they conclude that this association is due to local adaptation.

An adaptive explanation for shell sculpture characteristics is less readily found. Welter-Schultes (2000) found an altitude dependence of rib density in *Albinaria idaea*, that he attributed to the temperature gradient co-varying with altitude. Kemperman & Gittenberger (1988) proposed humidity and temperature related effects like surface water adhesion and shell permeability as factors influencing shell surface adaptation in these species. However, the formation of coarse, large spaced ribs on the shell surface could be influenced, at least in part, by the irregular feeding opportunities forced by aestivation periods (Cameron 1970) and thus probably be shaped mainly by environmental forces. This view was corroborated by our observation that the shell of *C. unifasciata*, after transferring them in a subadult stage to constant high humidity in the laboratory, grows much smoother.

The second axis revealed that sampling sites with more annual precipitation were associated with small snails (see Fig. 5). This was astonishing at first sight because a humid climate allows more feeding activity and snails should therefore attain a larger size. A similar tendency of increasing size with lower rainfall in small snail species was noted in Goodfriend (1986), and explained as a potential adaptation of the area/volume ratio to reduce the loss of water in drier areas. Johnson & Black (2000) on the other hand found that the growth rate was highly plastic in *B. vittatum*. Life history differences related to reproductive traits may as well be responsible for differences in shell size, as shown by Madec *et al*. (2000) for *Helix aspersa* populations. It is possible that *C. unifasciata* reaches sexual maturity, after which there is no further shell growth, earlier in wetter places in order to reproduce earlier. This would offer the next generation the possibility to reach a bigger size before winter. Early reproduction could thus constitute a selective advantage because winter mortality in juveniles has been shown to be size dependent in *H. aspersa* (Madec *et al*. 2000). Further experiments are needed to determine whether the size/precipitation correlation in *C. unifasciata* is due to phenotypic plasticity in response to the prevailing conditions or whether it has an adaptational significance.

Conclusions

We consider the phenotypic and phylogenetic divergence between the two identified lineages as sufficiently large to propose the presence of two distinct evolutionary entities described as *Candidula unifasciata* and *C. rugosiuscula*, though additional studies are necessary to confirm a species status for the latter. In the absence of obvious qualitative characters to distinguish the two lineages, only the joint statistical analysis of molecular genetic variation and quantitative shell traits was able to reveal this hidden divergence.

On the intraspecific level, the correlation of among population divergence to environmental variables in *C. unifasciata* allowed formulating hypotheses about their adaptive value. These hypotheses can now be tested in future studies. Integrating environmental, phenotypic and genetic data was thus a major step from observing and describing patterns to the understanding of the underlying evolutionary processes.

Acknowledgements

Part of this study was performed during the postdoctorate of M.P. at IMEP in Marseille, financed by a grant of the DAAD (No). We thank Dr R. Janssen for providing the shell material from Malacological Collection of the Senckenberg museum/Frankfurt am Main. Discussions with Anne Pfenninger, Klaus Schwenk and Tom Wilke improved the manuscript. Comments of three referees improved the manuscript further. We thank John Majstrick for his patience during his trials to amplify 16S from museum material.

References

Arnaud JF, Madec L, Bellido A, Guiller A (1999) Microspatial genetic structure in the land snail *Helix aspersa* (Gastropoda: Helicidae). *Heredity*, **83**, 110–119.

Arter HE (1990) Spatial relationship and gene flow paths between populations of the alpine snail *Arianta arbustorum* (Pulmonata: Helicidae). *Evolution*, **44**, 966–980.

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

Avise JC (2000) *Phylogeography — the History and Formation of Species*. Harvard University Press, London.

Bahl A, Pfenninger M (1996) A rapid method for isolation of DNA using laundry detergent. *Nucleic Acids Research*, **24**, 1587–1588.

Bahl A, Pfenninger M, Bamberger H, Frye M, Streit B (1996) Survival of snails in fragmented landscapes. In: *Species Survival in Fragmented Landscapes* (eds Settele J, Margules CR, Poschlod P, Henle K), pp. 329–343. Kluwer Academic Publishers, Dortrecht.

Bessemoulin J (1989) *Atlas climatique de la France: édition réduite*. Ministère des transports, Direction de la météorologie, Paris.

Bonnin I, Prosperi JM, Olivieri I (1996) Genetic markers and quantitative genetic variation in *Medicago truncatula* (Leguminosae): a comparative analysis of population structure. *Genetics*, **143**, 1795–1805.

Bourguignat JR (1882) In: *Prodome de la Malacologie française, Catalogue Mollusques terr., eaux douces* (ed. Locard A). J.-B. Bailliere, Paris.

Cameron RAD (1970) The effect of temperature on the activity of three species of Helicid snail (Mollusca: Gastropoda). *Journal of Zoology*, **162**, 303–315.

Davison A, Clarke B (2000) History or current selection? A molecular analysis of 'area effects' in the land snail *Cepaea nemoralis*. *Proceedings of the Royal Society London, Series B*, **267**, 1399–1405.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular

variance from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.

Gittenberger E (1993) On *Trochoidea geyeri* (Soós, 1926) and some conchological similar taxa (Mollusca: Gastropoda Stylommatophora: Hygromiidae). *Zoologische Medelingen*, **67**, 303–320.

Goodfriend GA (1986) Variation in land-snail shell form and size and its causes: a review. *Systematic Zoology*, **35**, 204–223.

Hull DL (1997) The ideal species concept — and why we can't get it. In: *Species: the Units of Biodiversity* (eds Claridge MF, Dawah HA, Wilson MR), pp. 357–380. Chapmann & Hall, London.

Johnson MS, Black R (2000) Associations with habitat versus geographic cohesiveness: size and shape of *Bembicium vittatum* Philippi (Gastropoda: Littorinidae) in the Houtman Abrolhos Islands. *Biological Journal of the Linnean Society*, **71**, 563–580.

Kemperman TCM, Gittenberger E (1988) On morphology, function and taxonomic importance of the shell ribs in Clauseliidae (Mollusca: Gastropoda, Pulmonata), with special reference to those in *Albinaria*. *Basteria*, **52**, 77–100.

Kerney MP, Cameron RAD (1979) *A Field Guide to the Land Snails of Britain and North-West Europe*. Collins Publishers, London.

Kerney MP, Cameron RAD (1999) *Guide Des Escargots et Limaces d'Europe* (ed. Bertrand A). Delachaux et Niestlé, Lausanne.

Koufopanou V, Reid DG, Ridgway SA, Thomas RH (1999) A molecular phylogeny of the Patellis Limpets (Gastropoda: Patellidae) and its implications for the origin of their antitropical distribution. *Molecular Phylogenetics and Evolution*, **11**, 138–156.

Lacy RC (1987) Loss of genetic diversity from managed populations: interacting effects of drift, mutation, immigration, selection and population subdivision. *Biological Conservation*, **1**, 143–158.

Lydeard C, Holznagel WE, Schnare MN, Gutell RR (2000) Evolutionary patterns and phylogenetic content of the molluscan mitochondrial LSU rDNA sequences and rRNA secondary structure. *Molecular Phylogenetics and Evolution*, **15**, 83–102.

Lynch M, Pfrender M, Spitze K *et al*. (1999) The quantitative and molecular genetic architecture of a subdivided species. *Evolution*, **53**, 100–110.

Madec L, Desbuquois C, Coutellec-Vreto MA (2000) Phenotypic plasticity in reproductive traits: importance in the life-history of the land snail *Helix aspersa* in a recently colonized habitat. *Biological Journal of the Linnean Society*, **69**, 25–39.

McQuaid CD, Branch GM, Frost PGH (1979) Aestivation behaviour and thermal relations of the pulmonate *Theba pisana* in a semi-arid environment. *Journal of Thermal Biology*, **4**, 47–55.

Medina M, Walsh PJ (2000) Molecular systematics of the order Anaspidea based on mitochondrial DNA sequences (12S, 16S and COI). *Molecular Phylogenetics and Evolution*, **15**, 41–58.

Michaud ALG (1831) *Complèment de l'Histoire Naturelle des Mollusques terrestres et fluviatile de la France, par Draparnaud*. Verdun.

Pfenninger M, Bahl A, Streit B (1996) Isolation by distance in a population of a small land snail *Trochoidea geyeri*: Evidence from direct and indirect methods. *Proceedings of the Royal Society, London B*, **263**, 1211–1217.

Podolsky RH, Holtsford TP (1995) Population structure of morphological traits in *Clarkia dudleyana*. I. Comparison of F_{ST} between allozymes and morphological traits. *Genetics*, **140**, 733–744.

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**, 817–818.

Ritland K (2000) Marker-inferred relatedness as a tool for detecting heritability in nature. *Molecular Ecology*, **9**, 1195–1204.

Rohlf FJ (1995) Statistical analysis of shape using partial warp scores. In: *Proceedings in Current Issues in Statistical Shape Analysis* (eds Mardia KV, Gill CA), pp. 154–158. Leeds University Press, Leeds.

Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Spitze K (1993) Population structure in *Daphnia obtusa* — Quantitative genetic and allozymic variation. *Genetics*, **135**, 367–374.

Swofford DL (1998) PAUP*. *Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4*. Sinauer Associates, Sunderland.

Thomaz D, Guiller A, Clarke B (1996) Extreme divergence of mitochondrial DNA within species of pulmonate land snails. *Proceedings of the Royal Society, London B*, **263**, 363–368.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Wagner J (1933) Egy íy Helicella-faj Magyarország faunájában. *Állattani Közlemények*, **30**, 3–4.

Welter-Schultes FW (2000) The pattern of geographical and altitudinal variation in the land snail *Albinaria idaea* from Crete (Gastropoda: Clausiliidae). *Biological Journal of the Linnean Society*, **71**, 237–250.

Yang RC, Yeh FC, Yanchuk AD (1996) A comparison of isozyme and quantitative genetic variation in *Pinus contorta* ssp. *latifolia* by F_{ST}. *Genetics*, **142**, 1045–1052.

The presented work constitutes part of a suite of studies on phylogeography, morphologic evolution and speciation of the *Candidula* genus. More information about the research interest of M.P. can be found at http://www.rz.uni-frankfurt.de/~streit/SCHNECKEN/Helicoidea.html. F. Magnin's interest are the ecology and palaeoecology of snail populations and communities.
