A three-way contact zone between forms of *Patella rustica* (Mollusca: Patellidae) in the central Mediterranean Sea

ALEXANDRA SÁ-PINTO1*, STUART J. E. BAIRD1,2, CATARINA PINHO1,3, PAULO ALEXANDRINO1,4* and MADALENA BRANCO1

1Centro de Investigação em Biodiversidade e Recursos Genéticos (CIBIO/UP), Campus Agrário de Vairão, 4485-661 Vairão, Portugal
2Centre de Biologie et de Gestion des Populations (CBGP), Campus International de Baillarguet, CS 30 016, 34988 Montferrier/Lez cedex, France
3Department of Genetics, Rutgers, The State University of New Jersey, 604 Allison Road, Piscataway, NJ 08854, USA
4Departamento de Zoologia-Antropologia, Faculdade de Ciências da Universidade do Porto, 4099-002 Porto, Portugal

Received 9 August 2009; accepted for publication 16 November 2009

Previous studies have reported the occurrence of three differentiated mtDNA lineages within *Patella rustica* in the Mediterranean Sea. Two hypotheses have been proposed to explain these observations: (1) the maintenance of ancestral polymorphism within a single species; (2) the occurrence of cryptic species not identified previously. To distinguish between these hypotheses, we screened the genetic variability at nine allozyme loci, an intron from the α-amylase gene and a mitochondrial gene for 187 individuals of *P. rustica* sampled from seven Mediterranean localities. Eight additional localities were screened for the last two markers to place the differentiated lineages in a clear geographic context. Our results demonstrate that the three mtDNA lineages correspond to three distinct nuclear genotype clusters and provide further details on their distribution: the cluster corresponding to the mtDNA lineage from the Atlantic and western Mediterranean extends as far as the south coast of Italy, whereas the remaining two clusters occur in sympatry in the eastern Mediterranean. One of the eastern Mediterranean clusters is highly differentiated and seems to be reproductively isolated from the codistributed form; we therefore suggest that it corresponds to a new species. The remaining two clusters are less differentiated and form a contact zone across south Italian shores. This three-way contact zone constitutes an interesting model for the study of speciation in the marine realm. © 2010 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2010, 100, 154–169.


INTRODUCTION

Limpets are common intertidal grazers on much of the world’s coasts and mostly belong to the family Patellidae (Gastropoda: Patellogastropoda). Historically, the species-level taxonomy of this family was based on morphological features of the shell which, despite their recognized utility in species’ identification, are highly variable (Fischer-Piette & Gaillard, 1959) and are known to be influenced by environmental conditions (Branch, 1981). This led to an enormous profusion of species’ descriptions and great taxonomic confusion, with different authors considering different numbers of species (Fischer-Piette & Gaillard, 1959; see revision in Titselaar, 1998). As an example, the revision of Christiaens (1973) reduced the 240 species described for the genus *Patella* (*sensu lato*) to 32.

*Corresponding authors. E-mails: xanasapinto@gmail.com; palexan@mail.icav.up.pt
Linnaeus described the taxon *Patella rustica*, in 1758, from type specimens collected from an unknown locality. Limpets with this description occur in the Mediterranean Sea and on the north-east Atlantic coast from Biarritz (France) to Mauritania (Christiaens, 1973). Such limpets are also present in Cape Verde, Canary Islands, Selvagens, Desertas and Madeira (Cörte-Real, Hawkins & Thorpe, 1996a), although some authors argue that the Macaronesian forms should be considered as a different species – *Patella piperata* (Christiaens, 1973; Titselaar, 1998). Recent studies involving mtDNA sequencing have revealed high genetic substructuring both between and within the Macaronesian and continental forms (Koufopanou et al., 1999; Sá-Pinto et al., 2005). Three mtDNA lineages of *P. rustica* have been found to occur on continental shores (Sá-Pinto et al., 2005): one described from the Atlantic and Western Mediterranean; a second described from the Aegean Sea and Eastern Mediterranean Basin; and the third from a single individual collected from the Mediterranean shores of south-east Turkey, in the same locality in which individuals from the second lineage were also found (Sá-Pinto et al., 2005). The differentiation between this single haplotype and those found in sympathy was greater than that reported between the Macaronesian and continental lineages, yet it clearly belongs in the *P. rustica* mtDNA clade. Two alternative explanations for the occurrence of this highly divergent haplotype are: the maintenance of ancestral polymorphism within a single species or the occurrence of an unidentified cryptic species (Sá-Pinto et al., 2005).

The likely occurrence of cryptic marine species has long been recognized (Knowlton, 1993, 2000). Knowlton (1993) pointed out several factors that limit the ability of taxonomists to detect species’ differences in the marine environment: (1) difficult access to organisms for sampling and for direct observation in their natural habitat; (2) traditional preservation methods for marine organisms which destroy important morphological characters; (3) large phenotypic plasticity; and (4) the dominant role of chemical recognition systems involved in the choice of mates and egg-sperm recognition, as well as in the settlement preferences of larvae. The first two factors limit the quantity and/or quality of morphological information available for taxonomic work, and the last two factors limit the utility of this information for species’ delimitation. Molecular tools are therefore particularly important in the taxonomy of marine species, allowing morphologically indistinguishable species to be recognized (Davidson & Haygood, 1999; Tarjuelo et al., 2001; McGovern & Hellberg, 2003; Remerie et al., 2006; see also revisions in Knowlton, 1993; 2000). Molecular markers have been used to distinguish interspecific morphological differences from high intraspecific variability and phenotypic plasticity, resulting in either taxonomic splitting (Weber & Hawkins, 2005; Nakai et al., 2006) or lumping of forms (Cörte-Real, Hawkins & Thorpe, 1992; Simison & Lindberg, 1999; Weber & Hawkins, 2002; Tsang et al., 2007). Molecular tools have also been used to study the degree of hybridization between species, its consequences for their genetic composition and the relative importance of pre- and post-zygotic barriers to gene flow, opening a window on the speciation process in the marine environment (Quesada et al., 1998; Comesana et al., 1999; Jones et al., 2006).

The morphological similarity of cryptic species may mask very different population histories, habitat requirements, ecological characteristics and physiologies (see, as an example, Davidson & Haygood, 1999; McGovern & Hellberg, 2003). Confounding cryptic species will therefore mislead scientific research and compromise management and conservation measures. Taking this into account, and motivated by the presence of three mtDNA lineages in the Mediterranean Sea, we investigated the genetic structure of *P. rustica* using both mitochondrial and nuclear markers and a fine-scale geographic sampling of the region. We aimed to: (1) describe the geographic substructuring of *P. rustica*; (2) test whether mtDNA and nuclear markers show congruent geographic patterns and, if so, (3) assess the degree of gene flow between differentiated forms; and (4) discuss the taxonomic and evolutionary implications of the results obtained.

**METHODOLOGY**

**SAMPLING**

Samples of *P. rustica*, identified by the analysis of morphological characters (Christiaens, 1973; Titselaar, 1998), were collected from the localities listed in Table 1 and shown in Figure 1. A subset was transported to the laboratory in dry ice where the samples were dissected and stored at −80 °C and analysed for allozyme markers, for a portion of the mitochondrial gene cytochrome *c* oxidase subunit I (COI) and for a portion of an intron of the α-amylase (α-AMY) nuclear gene. The remaining samples were kept in 100% ethanol (precluding allozyme analysis) and used to increase the resolution of geographic sampling for the three mtDNA lineages reported by Sá-Pinto et al. (2005) and for the nuclear gene α-AMY. Vouchers are available from the authors.

**DATA COLLECTION**

**Allozyme electrophoresis**

Nine allozyme loci, corresponding to eight enzymatic systems, were analysed by starch gel electrophoresis:
malate dehydrogenase (MDH, EC 1.1.1.37), malic enzyme (ME, EC 1.1.1.40), isocitrate dehydrogenase (IDH, EC 1.1.1.42), glutamate-oxaloacetate transaminase (GOT, EC 2.6.1.1), phosphoglucomutase (PGM, EC 2.7.5.1, two detectable loci hereafter referred to as PGM1 and PGM2), peptidase D (PEP D, EC 3.4.13.9), glucosephosphate isomerase (GPI, EC 5.3.1.9) and 6-phosphogluconate dehydrogenase (PGD, EC 1.1.1.44). Electrophoretic and staining procedures were carried out according to Sá-Pinto, Alexandrino & Branco (2007).

**Single-strand conformation polymorphism (SSCP) analysis**

The extraction of total genomic DNA from a portion of foot muscle was carried out using a standard high-salt protocol (Sambrook, Fritsch & Maniatis, 1989). A portion of 255 base pairs (bp) of the COI gene (previously shown to distinguish the three continental lineages) was amplified using the primers LCO1490 (Folmer et al., 1994) and PR_COIR (ACTAATCA ATTACCAAATCCTCC). Approximately 110 intronic bp from α-AMY were amplified using the primers PR_AMYF (AGAATTCATCAGTATGGTTCAACG) and PR_AMYR (ATAYGAACTGAATTAAAGACG; α-AMY sequence data from J.-L. Da Lage, CNRS, Paris, unpubl. data). For both fragments, polymerase chain reactions (PCRs) were performed with a total volume of 12.5 μL, using 0.5 U of Taq polymerase, 0.2 mM of the primers, 0.2 mM of each deoxynucleoside triphosphate (dNTP) and 2.0 mM of MgCl₂. PCR amplifications were performed for 3 min at 94 °C, 35 cycles of 45 s at 94 °C, 45 s at 50 °C and 45 s at 72 °C, followed by a 2-min extension at 72 °C. For each marker, 5 μL of the PCR product was mixed with 20 μL of denaturing loading buffer (95% deionized formamide, 10 mM NaOH, 0.01% bromophenol blue and 0.01% of

**Figure 1.** Map displaying the frequency of each mtDNA clade in every sampling site (location codes refer to Table 1). The area of each circle is proportional to the sample size and the area filled with a given colour is proportional to the number of individuals exhibiting the corresponding mtDNA clade (see Fig. 2): grey, mtDNA clade C; white, mtDNA clade D; black, mtDNA clade A.
xylene cyanol), denatured for 5 min at 95 °C and kept on ice until gel loading. Eight microlitres of each mixture were independently run on a 12% polyacrylamide gel (29 : 1 acrylamide : methylenebisacrylamide) with 1 x Tris Borate EDTA buffer on a vertical electrophoresis system (BIORAD Protein II). Electrophoreses were performed at a constant 220 V, 10 °C, for 18 h and visualized by silver staining. The identity of the SSCP conformers was confirmed by sequencing at least one individual per allele per sampling locality. For COI, sequencing for the identification of SSCP conformers was carried out for a longer fragment (PCR conditions followed Folmer et al., 1994). All sequencing was performed using an ABI Prism BigDye Terminator Cycle sequencing protocol on an ABI Prism 310 automated sequencer (Applied Biosystems). A ONE-WAY MEDITERRANEAN CONTACT ZONE IN PATELLA RUSTICA

DATA ANALYSES

Phylogenetic analysis of mtDNA

COI gene sequences were aligned manually with Bioedit (Hall, 1999) and trimmed to a 577-bp fragment. Maximum likelihood (ML) and Bayesian phylogenetic analyses were performed on a dataset that included all distinct P. rustica COI haplotypes found in the present and previous work (present work; Sá-Pinto et al., 2005; GenBank accession numbers GU205449–GU205471; DQ089602–DQ089612, DQ089624, DQ089627; EU073897–EU073913). Given the unresolved phylogenetic relationship between P. ferruginea Gmelin, 1791 and P. rustica mtDNA lineages, all the haplotypes available for the former species were included in the analysis (Sá-Pinto et al., 2005; Espinosa & Ozawa, 2006; accession numbers AB201478–AB201525; DQ089622–DQ089623). Sequences of P. vulgata Linnaeus, 1758, P. pellucida Linnaeus, 1758, P. ulysseponisens Gmelin, 1791, P. caerulea Linnaeus, 1758, P. depressa Pennant, 1777, P. candei d'Orbigny, 1839 and P. lugubris Gmelin, 1791 were used as outgroups (Sá-Pinto et al., 2005; accession numbers DQ089619, DQ089620, DQ089596, DQ089590, DQ089584, DQ089614, DQ089564, DQ089571, DQ089572, DQ089576).

ML analysis was performed with GARLI v0.95 (Zwickl, 2006) which uses an algorithm that simultaneously searches over tree topology, branch lengths and model parameters. Sequence evolution model parameters were co-estimated assuming a general time reversible (GTR) model with a proportion of invariable sites and a discrete approximation to a γ distribution for four rates (GTR + I + G; Rodriguez et al., 1990; Yang, 1994). Twenty independent runs were performed to avoid trapping in local optima. Bootstrap values were obtained with the same software using 100 pseudo-replicates. Bayesian phylogenetic analysis was carried out with MrBayes v3.1.2 (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). This program was run allowing the parameters of a GTR + I + G model to be searched independently for three partitions corresponding to the first, second and third codon positions. This procedure has been shown to result in better estimates of ln L and posterior probabilities (Brandley, Schimtz & Reeder, 2005). To avoid trapping in local optima, we performed two independent runs using parallel Metropolis-coupled Markov chain Monte Carlo (MCMCMC) with three hot chains and one cold chain each (Altekar et al., 2004). Both analyses started with a random tree and ran for 11.0 x 10⁸ generations, thinned to one sample each 1000 generations. The convergence of posterior probabilities in each chain and between the chains was assessed with AWTY (Nylander et al., 2008). As the two runs converged, both were used to estimate the 50% majority rule consensus tree, after discarding the trees produced before stationary posterior probabilities were reached (burn-in; 6000 trees).

Testing for the existence of three nuclear genotype clusters

The total dataset was made up of information for each individual's mtDNA clade, allozyme markers and α-AMY intron states. Traditional signals of population differentiation [deviations from Hardy–Weinberg and linkage (HWL) equilibria expectations] were tested using the exact test of Fst (Raymond & Rousset, 1995). To avoid trapping in local optima, we performed two independent runs using parallel Metropolis-coupled Markov chain Monte Carlo (MCMCMC) with three hot chains and one cold chain each (Altekar et al., 2004). Both analyses started with a random tree and ran for 11.0 x 10⁸ generations, thinned to one sample each 1000 generations. The convergence of posterior probabilities in each chain and between the chains was assessed with AWTY (Nylander et al., 2008). As the two runs converged, both were used to estimate the 50% majority rule consensus tree, after discarding the trees produced before stationary posterior probabilities were reached (burn-in; 6000 trees).

© 2010 The Linnean Society of London, Biological Journal of the Linnean Society, 2010, 100, 154–169
native predefined subdivisions of the total dataset (above), and to estimate the proportion of each individual's markers derived from each of a number of potential clusters. STRUCTURE was run for two and three clusters \((K = 2, K = 3)\) using the admixture model with correlated allele frequencies. For each analysis, the program was run three times for \(1 \times 10^6\) steps with a burn-in period of \(1 \times 10^5\). No prior information was provided regarding the origin of the individuals. STRUCTURE was run on both the total dataset, and with the mtDNA excluded (nuclear dataset). The mtDNA information was coded as diplo-type 1 0 for clade C, 2 0 for clade D and 3 0 for clade A (zero meaning missing data, see Fig. 2 for the clades’ nomenclature).

Testing the existence of reproductive barriers between differentiated clusters

As pointed out by Jiggins & Mallet (2000), for two sympatric taxa, a bimodal distribution of genotypes with a predominance of pure genotypes and deficiency of intermediates suggests restrictions to gene flow, whereas a unimodal distribution is indicative of a free admixture. In this study, we inspected the distribution of the frequency of individuals estimated to have a given proportion of their genotype assigned to one of three STRUCTURE clusters, for the analyses of both total and nuclear datasets. These point estimates for the proportion of the genotype origin of individuals must be treated with caution because intermediate point estimates may arise not only from truly hybrid individuals, but also from individuals whose particular genotypes simply contain little information about their source; the former have well-supported intermediate proportion estimates, the latter poorly supported. We therefore also considered the confidence in the point estimates of ancestry inferred for each individual by STRUCTURE. Confidence in (or support for) individual point estimates was expressed as 1 – CR, where CR (credible region) was defined as the span of the 95% probability interval output by STRUCTURE. In addition, we compared the number of ‘intermediate’ individuals (defined as those having between 20 and 80% of their genotype assigned to a cluster relevant for each analysis) in the real data with the number observed in simulations, assuming the free admixture of the inferred clusters. The number of ‘intermediate’ individuals in real locality samples was determined by running STRUCTURE for the total dataset, three times, for \(K = 3\), for \(1 \times 10^5\) steps with a burn-in period of \(1 \times 10^4\), under the assumptions of admixture and correlated allele frequencies with prior information regarding the origin of individuals from Peschici, Santa Maria di Leuca and Vibo Valentia. Migration between these populations was set to 0.01, a conservative assumption as our results suggest very low levels of effective migration between these sampling localities. The admixture simulation procedure was based on that described by Pinho et al. (2009) and for each sympatric sampling location involved: (1) the construction of a ‘generation zero’ population including 10 000 pure individuals belonging to two different clusters (simulated from Santa Maria di Leuca if pure A, Vibo Valentia if pure C or Peschici if pure D) in the same proportions as those of the mtDNA carrier classes observed in that sampling location; (2) the simulation of 100 generations of random mating within this mixed population; (3) the replacement of the individuals from the sympatric locality sample by an equal number of individuals randomly sampled from each simulated population in the total dataset; (4) a run of each dataset in STRUCTURE as described previously for determining the number of ‘intermediates’ in real locality samples. This procedure was repeated 100 times for each of the two admixture scenarios \((A|D\) mitochondrial carrier contact and \(C|D\) mitochondrial carrier contact) being analysed, resulting in a total of 200 datasets of simulated data. All simulations, construction of simulated datasets and posterior data summaries were performed with computer programs written in Python 2.6.1 (available from the authors).

Testing geographic clinal variation of cluster membership frequencies

None of the previously mentioned analyses are suitable for the capture of information on the gradual change in allele frequencies with geography. For this purpose, we used likelihood analyses of sigmoid clines (Barton & Gale, 1993), which are expected under a wide range of admixture circumstances (Kruuk et al., 1999). Likelihoods were calculated given data on mitochondrial carrier class frequencies (weighted by mitochondrial sample size) and, for nuclear markers, cluster assignment probabilities (weighted by the total assignment probability to each cluster within a locality sample). Observations were assumed to be binomially distributed around the cline expectations for each locality. ML estimates and support intervals were calculated using likelihood profiles, as implemented in Analyse (Barton & Baird, 1996). Profiles were calculated at a resolution of 10 km for centre and width parameters. A hierarchy of nested hypotheses regarding the nature of allele frequency change were compared in order of increasing complexity, with more complex hypotheses being accepted if their gain in likelihood was significant, given the increased number of parameters: H0, the data are explained by finite sampling from a panmictic system; H1, the data are explained by clinal change, mitochondrial and nuclear clines share the same centre and width; H2a, mitochondrial and nuclear clines share the same...
Figure 2. Maximum likelihood tree based on sequences of cytochrome c oxidase subunit I. Bootstrap values and posterior probabilities are shown above and below the branches respectively and are represented by * when lower than 50%. Sampling localities are coded as follows: ○, Azores; ●, Cape Verde; ○, Madeira and Desertas; ○, Canary Islands and Selvagens; , Continental Atlantic shores and Western Mediterranean Basin; III, Eastern Mediterranean Basin; 1, haplotypes reported in previous studies (Sá-Pinto et al., 2005, 2008; Espinosa & Ozawa, 2006); 2, haplotypes detected in the present work.
centre, but differ in their width; H2b, the two clines have the same width but differ in their centre; H3, the two clines differ in both centre position and width.

RESULTS

PHYLOGENY AND GEOGRAPHIC DISTRIBUTION OF mtDNA CLADES

By SSCP analysis, 22 conformers were detected and their subsequent sequencing revealed 28 haplotypes in a 577-bp fragment. New haplotypes were submitted to GenBank (accession numbers GU205449–GU205471).

Figure 2 shows the ML tree (A; log likelihood = −2829.6301) inferred for the haplotypes described for P. rustica and P. ferruginea. For each clade, bootstrap support and posterior probabilities higher than 50% are shown above and below the branches, respectively. According to both analyses, P. ferruginea is the sister species of all the genetically differentiated forms currently recognized as P. rustica. Within P. rustica, five differentiated clades were recovered by ML analysis: clade B groups the samples from Desertas Islands; clade E comprises the samples from Canary and Selvagens islands; clade C groups samples mostly from continental Atlantic shores and the western Mediterranean Basin; clade D groups samples from the eastern Mediterranean Basin; clade A also occurs in the eastern Mediterranean Basin and includes the highly divergent haplotype Pr10 described previously (Sá-Pinto et al., 2005). The Bayesian analysis recovered a similar tree topology with the exception of clade E (posterior probability lower than 50%). The clade E haplotypes were instead placed in a basal polytomy relative to all other P. rustica forms. Although low bootstrap support values were obtained for most P. rustica clades, their posterior probabilities were usually higher than 90%. Phylogenetic relationships between the clades usually displayed low bootstrap support and posterior probabilities. The exceptions were the high posterior probabilities obtained for the clade joining all the forms of P. rustica and for the group joining clades C and D. The differences between bootstrap values and posterior probabilities obtained for each clade were consistent with the tendency for the latter to be higher (Suzuki, Glazko & Nei, 2002). It has been shown that bootstrap analyses yield a greater underestimation of clade support than do posterior probabilities, although high posterior probabilities for untruth clades are sometimes obtained when oversimplified evolutionary models are used (Erixon, Svennblad & Britton, 2003).

The proportions of individuals carrying mtDNA from clades A, C and D are shown for each locality in Figure 1. Clade C, previously described from Atlantic continental shores (Sá-Pinto et al., 2005) and Sicily (Mauro, Arculeo & Parrinello, 2003), reaches its easternmost limit in Crotone. Clades A and D seem to be restricted to the eastern Mediterranean Basin and are found to occur in sympathy in six locations widespread throughout the sampled area. Clades C and D are found in sympathy only in Crotone. The simultaneous presence of the three mtDNA continental clades was not detected in any of the sampling localities.

NUCLEAR GENETIC POLYMORPHISM AND DISTINCTION BETWEEN FORMS

The survey of nine allozyme loci revealed an average of 3.8 alleles per locus, ranging from two alleles in IDH to seven alleles in PGM1 and PGM2. For the α-AMY fragment, six alleles were detected by SSCP. Significant linkage disequilibria were detected between the GPl and α-AMY loci in Savelletri and Taranto (P < 0.01), locations in which significant deviations from HWL expectations were also found (P < 0.01). By contrast, no significant deviations from HWL equilibria were detected within any of the A, C or D mtDNA carrier classes (P > 0.01). Significant differentiation was found between mitochondrial carrier classes (see Table 2). The highest Fst values were observed between mitochondrial carrier classes C and A (0.51) and mitochondrial carrier classes D and A (0.39; see Table 2). Loci GPl, PGM2 and α-AMY were the most informative for identifying the nuclear genotypes of individuals in mitochondrial carrier class A, whereas α-AMY, PGM2 and PEP D were informative for discriminating between the nuclear genotypes of mitochondrial carrier classes C and D.

The nuclear genetic differentiation observed between individuals belonging to different mtDNA carrier classes was reflected in the results of the clustering analyses. HWL clustering merged localities and separated mitochondrial carrier classes. When the nuclear dataset (mtDNA information excluded) was run in STRUCTURE for K = 2 (ln Probability of data (LnPD) of best run = −1146.8), two clusters were recovered, cluster A mostly matching individuals carrying mtDNA type A and cluster CD mostly matching individuals carrying mtDNA types C or D. Individuals were considered to be assigned to a cluster if they had at least 80% of their genotype assigned to that cluster. A total of 92% of the individuals in mitochondrial carrier class A were assigned to nuclear genetic cluster A, and 98% and 99% of carrier classes C and D, respectively, were assigned to cluster CD. The same two clusters were recovered when the total dataset was run in STRUCTURE for K = 2 (LnPD of best run = −1224.7), with 98% of mitochondrial carrier class A being assigned to nuclear cluster A and 100% of classes C and D being assigned to cluster CD.
The runs of the nuclear dataset for \( K = 3 \) had much higher values of LnPD than those obtained for \( K = 2 \) (LnPD of best run = -1095.1; difference in log likelihood, +50.7). For \( K = 3 \), one of the clusters recovered corresponded to cluster A obtained for \( K = 2 \), whereas the CD cluster of \( K = 2 \) was split into clusters C and D, mostly corresponding to mitochondrial carrier classes C and D, respectively (see Fig. 3). The same three clusters were recovered when the total dataset was run in STRUCTURE for \( K = 3 \) (LnPD of best run = -1119.4; difference in log likelihood, 105.3). An individual was considered to be consistently assigned if it carried mtDNA type A, C or D and was assigned, respectively, to clusters A, C or D according to its nuclear genotype. The vast majority of individuals with mtDNA from clade A were consistently assigned (see Table 3). Lower percentages of consistent assignments were obtained for individuals carrying mtDNA from clades C and D, with an asymmetric tendency to assign individuals from mitochondrial carrier class D to nuclear cluster C. For both clusters C and D, the percentages of individuals consistently assigned were higher in Vibo Valentia and Peschici, where all individuals carried mtDNA from clades C or D (see Table 3). Particularly low percentages of consistently assigned individuals were found in Crotone, where mtDNA from clades C and D occurred in sympatry.
Hybridization and reproductive isolation between forms

Figure 4 shows the frequency of individuals with a given proportion of their genotype assigned to cluster A (results from the analysis of the total dataset for $K = 3$). The vast majority of individuals coming from locations in which mtDNA clade A was absent had less than 10% of their genotype assigned to cluster A, whereas the majority of those coming from locations in which only mtDNA type A was present had more than 90% of their genotype assigned to this cluster. In the sampling locations of Savelletri and Taranto, strong bimodal distributions were obtained, with the vast majority of individuals having less than 20% or more than 80% of their genotype assigned to this cluster. The intermediate estimates cannot therefore be interpreted as evidence of hybrids, merely that some individuals have no clear assignment. Similar results were obtained from the analysis of the nuclear dataset only (data not shown). The proportions of ‘intermediate’ individuals (those having between 20 and 80% of their genotype assigned to cluster A) observed in Taranto and Savelletri (0.034 and 0, respectively) were lower than those obtained in all simulated populations after 100 generations of random mating between individuals originating from localities with samples ‘pure’ for mitochondrial carrier classes A and C (minimum proportion of ‘intermediate’ individuals in the 100 simulated populations was 0.41 for Taranto and 0.45 for Savelletri). Thus, comparison with the simulations shows that the bimodal distributions observed in Taranto and Savelletri have a real hybrid deficit over and above what could be expected under free admixture, indicating the existence of barriers to gene flow between clusters A and D. Alongside this evidence for a barrier to gene flow, some evidence of A–D introgression was found in two individuals from Taranto. These mtDNA D carriers had 79% and 89% of their genotypes attributed to cluster D, but were heterozygous for the $\alpha$-AMY locus by carrying alleles $\alpha$-AMY*2 and $\alpha$-AMY*6, respectively. These two alleles carry an
insertion that is probably a derived state (as it is absent from *P. ulyssiponensis* and *P. vulgata*, data not shown) and, except for these two individuals, is restricted to cluster A. Accordingly, the presence of these alleles in individuals from clade D in Taranto supports the occurrence of introgression against the alternative hypothesis of maintenance of ancestral polymorphism.

To analyse the hybridization between clusters C and D, the frequency of individuals with a given proportion of genotypes assigned to cluster C was calculated for each sampling location after removing all mitochondrial carrier class A individuals (results from the analysis of the total dataset for \( K = 3 \); see Fig. 5). According to these results, all individuals from Vibo Valentia had more than 80% of their genotype assigned to cluster C, whereas the majority of individuals from Peschici and Savelletri had less than 10% of their genotype assigned to this cluster. In Taranto and Crotone, a higher frequency of individuals with intermediate proportions of genotype assigned to cluster C was observed, but all intermediate point estimates were poorly supported (Fig. S2, see Supporting Information). Despite the bimodal distribution observed in Crotone, the proportion of individuals with intermediate ancestry estimates (0.33) was similar to or higher than that obtained in 34% of the simulated populations after 100 generations of random mating between individuals from pure mitochondrial carrier class D and C localities. Accordingly, the bimodal distribution at Crotone cannot be considered as evidence for a barrier to gene flow between clusters C and D.

**Geographic variation of cluster C and D frequencies**

The possibility of clinal change between C and D carrier classes was explored after excluding carrier class A individuals from all locality samples. Sigmoid clines explained the geographic variation of C vs. D mtDNA carrier class frequency and of nuclear cluster assignment to C or D significantly better than sampling noise within a panmictic system (H0 rejected with respect to H1 with \( P < 0.001 \)). Mitochondrial and nuclear data were better explained by two distinct sigmoid clines differing in both their widths and centre locations (H1 rejected with respect to H2a with \( P = 0.01 \), but not with respect to H2b (\( P = 0.06 \)); H2a rejected with respect to H3 with \( P = 0.02 \)). Figure 6 shows the ML estimate for mitochondrial and nuclear clines under the H3 hypothesis. The mtDNA cline has an estimated centre at 270 km along the coastline transect [2 unit support interval (CI), 230–280 km] and has a width of 70 km (CI, 30–260 km). The nuclear cline is estimated to be centred east of the mtDNA cline centre, at 300 km along the coastline transect (CI, 280 and 320 km) with a width of 330 km (CI, 270 and 400 km). The position of Crotone (the locality in which both mitochondrial lineages are present) is 277 km along the coastline transect.

---

**Figure 4.** Plot of the frequency of individuals with a given proportion of genotype assigned to cluster A recovered by STRUCTURE for the total dataset for \( K = 3 \). Black, individuals sampled from Peschici, Crotone and Vibo Valentia; dark grey, individuals sampled from Savelletri; light grey, individuals sampled from Taranto; white, individuals sampled from Crete and Santa Maria di Leuca.
DISCUSSION

The study of the genetic variation of mtDNA and 10 nuclear markers in several locations across the Mediterranean Sea revealed three distinct HWL clusters within *P. rustica*. One was highly differentiated and corresponds to a previously unidentified species. The other two clusters were much less differentiated and there appears to be ongoing gene flow between them.

CRYPTIC SPECIES IN THE EASTERN MEDITERRANEAN

Our results show that the highly differentiated haplotype of *P. rustica* found in the Eastern Mediterranean Sea (Sá-Pinto et al., 2005) belongs to a mtDNA clade (clade A) which is present in numerous locations in the Eastern basin, where it may occur in sympatry with another mtDNA clade (clade D). Several lines of evidence support the existence of strong barriers to

**Figure 5.** Plot of the frequency of individuals from each sampling location with a given proportion of genotype assigned to cluster C recovered by STRUCTURE for the total dataset for $K = 3$: black, Peschici; grey, Savelletri; orange, Taranto; green, Crotone; white, Vibo Valentia. All individuals carrying mtDNA from clade A were removed from this analysis.

**Figure 6.** Observed frequencies (full circles) and ML sigmoid clines (solid lines) estimated for mtDNA carrier classes C (black) and D (grey), and weighted average proportion of C (blue) and D (red) nuclear ancestry point estimates over individuals when $K = 3$, in every sample location as a function of its distance to Vibo Valentia (X axis; Kms). All mtDNA clade A carriers were removed from these analyses.
genetic integrity in sympatry, their genetic, ecological, behavioural, chemical and demographic properties may be different and evolve independently, even if some hybridization occurs. We consider that the two groups of individuals are able to maintain their genetic integrity in sympatry. In fact, if some hybridization are still possible; (2) at least some of the hybrids between these two entities are fertile, as the highlighted individuals must have resulted from backcrosses of F1 with cluster D individuals.

Despite the evidence for some introgression between clusters A and D, hybridization between these two entities must be rare as they are able to maintain their genetic integrity in sympatry. Although the hard line of the biological species' concept (BSC; Mayr, 1942) is that total reproductive isolation is necessary for species' status to be recognized, many defenders of this concept now admit that some hybridization might still occur between distinct species (Coyne & Orr, 2004). To deal with the problem of hybridization, a more practical species' concept was proposed by Mallet (1995), called the 'genotypic clustering species' concept' (GCSC). GCSC puts the emphasis not on reproductive isolation itself, but on the ability to remain distinct in sympatry. In fact, if two groups of individuals are able to maintain their genetic integrity in sympatry, their genetic, ecological, behavioural, chemical and demographic properties may be different and evolve independently, even if some hybridization occurs. We consider that the two entities shall, in this case, be considered as independent species. As our results clearly suggest, clusters A and D are able to maintain their genetic integrity in sympatry, thus fulfilling the requirements to be considered as different species according to GCSC. As the two entities differ in mtDNA as well as in more than one nuclear locus randomly selected from the genome, we believe that the maintenance of their cluster integrity is a result of the occurrence of strong reproductive isolation and not just selection in one particular locus.

Although it was not our aim to perform a morphological study, after assigning genetically each individual to one of the two entities present in the Eastern Mediterranean, we were able to recognize two distinct groups on the basis of shell morphology. Class A individuals clearly fall within the description made by Pallary (1938) for Patella (Patellastra) lusitanica (= Patella rustica) variety orientalis, which has a denser, thinner and more regular sculpture, with a less granulose and thinner ornamentation, than the western variety. Unfortunately, the shells used by Pallary (1938) to describe variety orientalis of P. rustica (deposited in the Natural History Museum of Paris) have apparently been lost, precluding a direct comparison with our samples. We suggest that comparisons of limpets from the type locality of Pallary's variety orientalis (Beirut, Pallary, 1938) might be sufficient for taxonomists to confirm Pallary's naming for the species we have identified using molecular tools.

In our study, the two Eastern Mediterranean species were found to occur sympatrically in six of 12 sampling localities and, in some of these locations, both were present in high frequency. If reproductive isolation between them is a result of post-zygotic mechanisms only, a fraction of descendants in each generation would be nonviable or less fertile; this is a severe constraint on the conditions for sympatric coexistence, as either would benefit from the local loss of the other. We therefore hypothesize that pre-zygotic mechanisms are playing an important role. Mechanisms such as asynchronous spawning (Palumbi, 1994; Clifton, 1997; Knowlton et al., 1997; Clifton & Clifton, 1999) and differences in mechanisms of male–female gamete recognition and attraction (Palumbi, 1994; Howard & Berlocher, 1998; Biern et al., 2002) have already been shown to play an important role in maintaining the integrity of marine species, and may well be involved in the reproductive isolation observed between clusters A and D.

A CONTACT ZONE ACROSS THE SHORES OF SOUTHERN ITALY

Another outcome of the present work is the description of an interesting contact zone across the southern coast of the Italian Peninsula, where the ranges of mtDNA clades C and D meet. The mitochondrial carrier classes C and D were recovered by HWL clustering of nuclear data only, although with lower fidelity than the reconstruction of clade A (Fig. 3 and
Table 3). Both the relative frequency of mtDNA from clades C and D and the relative proportion of nuclear genotype assigned to clusters C and D were shown to vary cliionly from west to east (Fig. 6), with the centres of both clines being located around Crotone, the sole location in which mtDNA clades C and D occur sympatrically. There are almost no rocky shores between Taranto and Crotone, and, although we carefully prospected the few rock outcrops in this region and the man-made structures that could support limpets (such as harbours and piers), we could not find any P. rustica in this area. The fact that the centres of both mitochondrial and nuclear clines are located near or within Taranto Bay thus matches theoretical expectations for a tension zone (Key, 1968), maintained by endogenous selection, and thus free to move across the environment. These tension zones tend to move towards, and become trapped by, barriers to dispersal or density troughs (Barton & Hewitt, 1985), such as the Taranto–Crotone gap in Patella habitat.

Two questions need to be addressed in order to understand the degree of reproductive isolation between these two genetically differentiated forms of P. rustica: (1) is there evidence of gene flow between eastern and western forms?; (2) if so, can these two forms freely admix when in sympathy or is there evidence for restricted gene flow?

Two different lines of evidence suggest ongoing gene flow between these two divergent forms. First, the percentage of individuals consistently assigned to clusters C and D is much lower in the sympatric location of Crotone than in Vibo Valentia (only mitochondrial class C carriers) or Peschici and Savelletti (no mitochondrial class C carriers), suggesting admixture in the former sampling location. The differences in the width and centre of nuclear and mitochondrial clines (Fig. 6) also suggest the occurrence of fertile hybrids between the two forms, with higher introgression of the western form’s nuclear genome into the eastern form, a pattern consistent with migration being driven by oceanic currents which flow from the Western into the Eastern Mediterranean Basin (Rohling et al., 2009).

The second question regarding the extent of hybridization between the two forms when in sympathy cannot be clearly answered from our results. Although a small number of individuals with intermediate point estimates for ancestry are found in Crotone, this number is consistent with the level we would expect to detect with STRUCTURE under random mating between members of mitochondrial carrier classes C and D at this sampling location. The point estimates of these ‘intermediate’ individuals are also unsupported, as the 95% confidence interval includes both zero and 100% of cluster C ancestry (see Fig. S2). Taken together, these results suggest that insufficient information is contained in the genetic markers used in the present work to fully address the second question. The narrower mtDNA cline could be explained by the occurrence of high selective pressures against introgression of the nonrecombining mitochondrial genome (see Geraldes et al., 2006 for the effect of the recombination rate on the degree of introgression). Such an effect has been reported for other hybrid zones and would imply some degree of hybrid unfitness. However, steeper mtDNA clines can occur even if hybrids are completely fertile and viable, as drift is expected to more strongly affect mtDNA because of its lower effective population size (Sequeira et al., 2005). Other alternative explanations, such as the higher dispersal abilities of males when compared with females, are unlikely to account for the observed pattern, as dispersal is via sea currents during the (nonsexually dimorphic) larval phase and thus very unlikely to differ between the sexes. Different mobility for male and female gametes is also an unlikely explanation for the lower mtDNA introgression, as fertilization must take place near the spawning location given the dilution effect of gametes and their short span of viability (see Le Quesne, 2005; Hodgson et al., 2007).

Taken together, our results support hybridization between eastern and western forms of P. rustica, but do not allow us to measure any intrinsic barriers restricting gene flow between them. Other questions remain to be addressed. As an example, epistasis between traits can cause clines with sharp central steps (Barton & Gale, 1993) – it is possible that the clines of some nuclear loci show such a sharp step at the mitochondrial cline centre, but such questions can only be addressed with sampling at a finer geographic resolution and with screening a higher number of loci. It would also be interesting to study the possibility of genetic substructuring for each of the three differentiated forms, but this requires samples from other areas of both Eastern and Western Mediterranean Basins. In the Eastern Mediterranean Basin, this additional sampling should focus outside the Adriatic Sea, as differentiated populations have been reported for several marine organisms in this shallow basin (see, for example, Borsa, Blanquer & Berrebi, 1997 and revision in Pérez-Losada et al., 2007).

The study of zones in which differentiated forms occur in sympathy may shed light on factors involved in the development and maintenance of reproductive isolation (Harrison, 1993; see examples in Jiggins et al., 1996 and Payseur, Krenz & Nachman, 2004). The occurrence over a small geographic scale of three forms which apparently show different degrees of differentiation and reproductive isolation allows us to look at different stages of the speciation process.
suggesting this species’ complex as an ideal model for marine speciation studies.

ACKNOWLEDGEMENTS

We are very grateful to Jean-Luc Da Lage who kindly provided unpublished sequence data for the α-AMY gene. We would also like to thank the following for their help in the collection of samples: Marco Arculeo, Pedro Cardia, Diogo Sayanda, António Múrias, Raquel Ribeiro, Miguel Carretero, Diana Barbosa and Cilia Antoniou. Pedro Ribeiro, António Múrias and Fernando Lima are acknowledged for discussions regarding morphological differentiation. Pedro Cardia, Rui Faria and Sara Rocha are acknowledged for discussions, suggestions and important comments on the manuscript. This work was financed by Fundação para a Ciência e Tecnologia (FCT), through the project POCTI/BSE/42300/2001. Alexandra Sá Pinto, Madalena Branco and Catarina Pinho were supported by FCT grants SFRH/BPD/48750/2008, SFRH/BPD/40073/2007 and SFRH/BPD/28869/2006, respectively. Stuart Baird is financed by the FCT’s programme Ciência 2007. Three anonymous referees are also acknowledged for important suggestions made on a previous version of the manuscript.

REFERENCES


© 2010 The Linnean Society of London, Biological Journal of the Linnean Society, 2010, 100, 154–169


**SUPPORTING INFORMATION**

Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Plot of proportion of genotype attributed to cluster A in each individual (x axis) and the respective confidence for this estimated proportion (y axis), obtained by calculating 1 – (95% credibility range): white diamonds, individuals sampled from Peschici, Crotone and Vibo Valentia; black circles, individuals sampled from Savelletri; black triangles, individuals sampled from Taranto; white squares, individuals sampled from Crete and Santa Maria di Leuca.

**Figure S2.** Plot of proportion of genotype attributed to cluster C in each individual (x axis) and the respective confidence for this estimated proportion (y axis), obtained by calculating 1 – (95% credibility range). All individuals carrying mtDNA from clade A were removed from this analysis. White triangles, individuals sampled from Vibo Valentina; black circles, individuals sampled from Crotone; grey diamonds, individuals sampled from Taranto, Savelletri and Peschici.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.