Genetic diversity and phylogeography of the Apennine yellow-bellied toad *Bombina pachypus*, with implications for conservation

DANIELE CANESTRELLI, ROBERTA CIMMARUTA, VERA COSTANTINI and GIUSEPPE NASCETTI
Dipartimento di Ecologia e Sviluppo Economico Sostenibile, Università della Tuscia, Via San Giovanni Decollato 1, 01100 Viterbo, Italy

Abstract
Genetic variation was investigated in 17 populations of the Italian endemic Apennine yellow-bellied toad using both mitochondrial (598 bp of the cytochrome *b* gene) and nuclear (21 allozyme loci) markers. Populations from central Calabria (southern Italy) showed the highest levels of intrapopulation genetic variation, whereas samples located north of this region were nearly lacking in variation. This appears to be a typical pattern of ‘southern richness and northern purity’, usually attributed to the prolonged population stability within southern refugia coupled with the loss of variation during postglacial northward expansion. However, the overall pattern of genetic variation observed has a strong geographical component, suggesting two Calabrian plains, Catanzaro and Crati-Sibari, as historical barriers to dispersal separating three population groups. These findings cannot be explained by the prolonged stability of southern populations alone, and suggest that the southern richness has been at least in part shaped by allopatric differentiation within the refugial range, followed by intermixing of previously differentiated lineages. From a conservation standpoint, Calabria is the major genetic diversity reservoir for this species, thus deserving particular conservation efforts. Furthermore, although the low intrapopulation genetic variation outside Calabria appears to be of clear historical origin, evidence of a current reduction of gene flow suggests that human disturbance has also played a part, particularly in the anthropogenic impacted Volturno river drainage basin.

Keywords: allozymes; *Bombina pachypus*, genetic diversity, glacial refugia, Italy, mtDNA

Received 17 April 2006; revision accepted 9 June 2006

Introduction
Climatic oscillations during the Pleistocene have greatly affected the pattern of distribution of many species in the Western Palaearctic region, as well as their demographic history and patterns of population genetic differentiation (for reviews see Avise 2000; Hewitt 2004a, b). The ranges of many organisms were fragmented during glacial maxima, with geographical isolates persisting in glacial refugia in the Mediterranean and western Asia, and particularly in the southern portions of the Mediterranean peninsulas of Italy, Iberia, and the Balkans. Following environmental amelioration during interglacials, several taxa underwent a northward range expansion from distinct southern refugia, recolonizing northern environments, and in several cases establishing secondary contact zones between previously isolated evolutionary lineages (Taberlet et al. 1998; Hewitt 1999; but see also Stewart & Lister 2001; Defontaine et al. 2005 and references therein). One of the major implications of this general scenario is the pattern of ‘southern richness and northern purity’ (Hewitt 1996, 1999, 2000). That is, the higher population genetic diversity found in the south of the Mediterranean peninsulas (i.e. in the former glacial refugia) compared with populations located further north. The ‘northern purity’ was mainly attributed to the progressive loss of genetic variation due to serial bottlenecking during the process of rapid northward range expansions (e.g. Hewitt 1996, 2000). The high genetic diversity found in the south would instead be mainly attributed to the prolonged demographic stability of the populations in these areas (Hewitt 1996). More recently, an alternative hypothesis has been proposed to explain the ‘southern richness’ pattern, based on the phylogeographic
studies concerning the Iberian Peninsula (Guillaume et al. 2000; Sanz et al. 2000; Gómez & Lunt in press). In fact, a strong population differentiation was found within this glacial refugium for a wide variety of taxa, suggesting the existence of multiple refugia. This pattern of ‘refugia-within-refugia’ (Gómez & Lunt, in press) has important implications in interpreting the distribution patterns of genetic diversity within the southern peninsulas. First, the ‘southern richness’ could have arisen from at least two mutually nonexclusive causes: (i) the prolonged demographic stability of populations from these areas (Hewitt 1996), and (ii) the existence of a high geographical structuring of populations due to allopatric differentiation (Guillaume et al. 2000; Sanz et al. 2000). Second, if a strong population structure exists, only a subset of the overall southern diversity would contribute to the northward colonization, causing an even more pronounced loss of genetic diversity (Gómez & Lunt, in press and references therein). Growing evidence supports this scenario for the Iberian Peninsula (e.g. Jaarola & Searle 2004; Vila et al. 2005).

As for the Italian refugium, prolonged demographic stability of populations (sometimes coupled with a northward range expansion during interglacials) was assumed as a likely common history for many taxa from this geographical area, by authors emphasizing its role either as a glacial refugium and source-area for successive recolonization of northern habitats, or as an area of high endemicity (Hewitt 1996; Bilton et al. 1998; Petit et al. 2003). However, the possible contribution of a ‘refugia-within-refugia’ scenario (and more broadly of allopatric differentiation) in shaping the pattern of genetic diversity in the south of the Italian Peninsula, still remains substantially unexplored (but see Santucci et al. 1996; Podnar et al. 2005).

The first aim of this study is to investigate whether and to what extent the ‘southern richness’ within the putative refugial area in southern Italy may be attributed to the existence of population subdivisions, rather than to the effects of prolonged demographic stability. With this aim we focused on the Apennine yellow-bellied toad Bombina pachypus, and here present an assessment of the population genetic structure of this species, as revealed by the analysis of variation at both nuclear (21 allozyme loci) and mitochondrial markers (a 598-bp fragment of the cytochrome b gene). This species is endemic to peninsular Italy, and some authors have suggested that it originated during the Quaternary ice ages, when this species and its sister taxon Bombina variegata were confined to their glacial refugia in southern Italy and the Balkans, respectively (Nascetti et al. 1982; Szymura et al. 1985; Szymura 1993). Since preliminary allozyme data also suggested a higher genetic diversity in the southern portion of the range than in the northern one (Nascetti et al. 1982), this species appeared a suitable case study to investigate the factors underlying the so-called ‘southern richness’. The Apennine yellow-bellied toad, formerly deemed a subspecies of B. variegata is now regarded a distinct species (after Lanza & Vanni 1991; Lanza & Corti 1993) on the basis of both allozyme (Nascetti et al. 1982) and morphological (Vaccaneo 1931) evidence (but see Veith 1996 and Sarrocco & Bologna 2000). It breeds in temporary pools and other shallow and unshaded freshwater environments (Lanza 1983), and is mainly distributed at mid-altitude areas along the Apennine chain (e.g. Caputo et al. 1985; Doria & Salvвидio 1994; Mazzotti et al. 1999; Sarrocco & Bologna 2000). The species’ conservation status is a cause for concern, and it is protected (still under the name of B. variegata) by several international conventions and regional laws (it is listed in Annex II of the Bern Convention and in Annexes II and IV of the EU Council Directive for the Conservation of Natural Habitats and of Wild Fauna and Flora). The species’ range appears fragmented, and several authors have recently suggested that many of its populations could be declining and disappearing, and that the extant populations often consist of no more than a few individuals (e.g. Caputo et al. 1985; Doria & Salvвидio 1994 and references therein; Sarrocco & Bologna 2000; Barbieri et al. 2004). Since genetic data are relevant for the assessment of both the species’ conservation status and management priorities (for extensive reviews see Avise & Hamrick 1996; Young & Clarke 2000; Frankham et al. 2002; Ferrière et al. 2004), in this study we also discuss the implications of our results for the conservation of the species.

Materials and methods

Sampling

We collected 161 individuals of Bombina pachypus from 17 localities ranging from the northern Apennines to the tip of Calabria. Geographical location of sampling sites and sample size are shown in Table 1 and Fig. 1a. Live animals were anaesthetized in the field by submersion in a 0.02% solution of MS222 (3-aminobenzoic acid ethyl ester). Tissue samples were obtained by toe-clipping, and each specimen was then released in the same collection place. Collected tissue samples were transported to the laboratory in a 75% solution of EtOH (samples 6, 8, 10 and 13) or in liquid nitrogen containers and then stored at −80°C until further analyses (all other samples).

DNA extraction, amplification and sequencing

Total DNA was extracted using the cetyltrimethyl ammonium bromide (CTAB) procedure (Doyle & Doyle 1987). Partial sequences of the mitochondrial DNA (mtDNA) gene encoding for the cytochrome b were obtained through polymerase chain reaction (PCR) amplification. Amplification was performed in a volume of 50 µL, containing MgCl₂
Table 1 Geographical location and sample size for the 17 populations sampled of Bombina pachypus

<table>
<thead>
<tr>
<th>Sampling locality</th>
<th>Altitude (m a.s.l.)*</th>
<th>Geographical coordinates</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bagno di Romagna</td>
<td>465</td>
<td>43°51′ N</td>
<td>13</td>
</tr>
<tr>
<td>Rio a Buti</td>
<td>200</td>
<td>43°56′ N</td>
<td>11</td>
</tr>
<tr>
<td>Firenze</td>
<td>100</td>
<td>43°47′ N</td>
<td>8</td>
</tr>
<tr>
<td>Rocca Caramanico</td>
<td>1050</td>
<td>42°09′ N</td>
<td>6</td>
</tr>
<tr>
<td>Stipes</td>
<td>880</td>
<td>42°15′ N</td>
<td>6</td>
</tr>
<tr>
<td>San Biagio</td>
<td>400</td>
<td>41°37′ N</td>
<td>5</td>
</tr>
<tr>
<td>San Lupo</td>
<td>330</td>
<td>41°16′ N</td>
<td>11</td>
</tr>
<tr>
<td>Carife</td>
<td>700</td>
<td>41°01′ N</td>
<td>4</td>
</tr>
<tr>
<td>S Angelo a Fasanella</td>
<td>280</td>
<td>40°27′ N</td>
<td>9</td>
</tr>
<tr>
<td>S. Severino Lucano</td>
<td>870</td>
<td>40°05′ N</td>
<td>8</td>
</tr>
<tr>
<td>Monte Cocuzzo</td>
<td>1200</td>
<td>39°14′ N</td>
<td>9</td>
</tr>
<tr>
<td>Macchia Longa</td>
<td>1300</td>
<td>39°22′ N</td>
<td>12</td>
</tr>
<tr>
<td>Taverna</td>
<td>520</td>
<td>39°01′ N</td>
<td>5</td>
</tr>
<tr>
<td>Cardinale</td>
<td>560</td>
<td>38°44′ N</td>
<td>5</td>
</tr>
<tr>
<td>Girifalco</td>
<td>300</td>
<td>38°49′ N</td>
<td>10</td>
</tr>
<tr>
<td>Stilo</td>
<td>300</td>
<td>38°30′ N</td>
<td>11</td>
</tr>
<tr>
<td>Gambrie</td>
<td>1300</td>
<td>38°10′ N</td>
<td>15</td>
</tr>
</tbody>
</table>

* m a.s.l., metres above sea level.

(2.5 mm), the reaction buffer (1X; Promega), the four dNTPs (0.2 mM each), the two primers (0.2 µM each), the enzyme Taq polymerase (2 U; Promega) and 2 µL of DNA template. Preliminary amplifications were performed using the generic primers L14841 (Kocher et al. 1989) and MVZ16 (Moritz et al. 1992). The sequences obtained with these primers were used to design the specific primers 494BOMOD (5′-GCATGGTGAAATTTTGGCTCC-3′) and CYTbBOMOD (5′-CCTAGTAGGTTTGGGGTGAATAGC-3′), which were then used to screen all studied specimens. PCR cycling procedure was: 95 °C for 5 min followed by 33 cycles of 93 °C for 1 min, 52 °C for 45 s, 72 °C for 1 min 30 s and a single final step at 72 °C for 10 min. Sequences were obtained using an ABI PRISM 377 DNA sequencer (PE Applied Biosystems) following the ABI PRISM BigDye Terminator Cycle Sequencing protocol. Both strands were sequenced in each individual analysed.

Multiple sequence alignments were made using CLUSTAL X (Thompson et al. 1997) and checked by eye.

mtDNA data analysis

The genealogical relationships between the haplotypes found were represented by a network, estimated using the statistical parsimony algorithm described by Templeton et al. (1992), and implemented by the software tcs 1.13 (Clement et al. 2000).

Populations’ nucleotide (π) and haplotype diversity (h) were estimated following Nei (1987), using the software DNASP 4.0 (Rozas et al. 2003). In order to detect signatures of past population growth, the same software was also used to calculate the statistics Fc (Fu 1997) and R2 (Ramos-Onsins & Rozas 2002), which in a recent study appeared the most powerful for the detection of recent demographic expansions (Ramos-Onsins & Rozas 2002). The statistical significance of these statistics was evaluated by coalescent simulations (10 000 replicates).

Genetic differentiation between populations was assessed by estimating pairwise values of FST, and the associated significances (by 100 000 permutations), using the software ARLEQUIN 2.0 (Schneider et al. 1999). This software was also used to perform an analysis of molecular variance (AMOVA; Excoffier et al. 1992; significances assessed by 100 000 permutations). Both the pairwise FST and AMOVA analyses were carried out incorporating information about divergence among haplotype pairs. With this aim, they were run using Tamura & Nei (1993) genetic distance, which is the best approximation available in ARLEQUIN of the best-fit model of sequence evolution (HKY), selected through the Akaike Information Criterion as implemented by the program MODELTREE 3.6 (Posada & Crandall 1998).

In order to ascertain the existence of a migration–drift equilibrium at a regional scale, we used the approach proposed by Hutchinson & Templeton (1999). Under equilibrium conditions, both genetic distance measures and their variability are expected to increase with increasing geographical distance between populations. The lack or weakness of one or both of these relationships is indicative
of a lack of regional equilibrium. When gene flow is more influential than drift (case II of Hutchinson & Templeton 1999), a pattern reflecting panmixia is expected, with no relationships between genetic and geographical distance and little variation in estimates of genetic divergence. This variation is instead expected to be large when genetic drift is strong, relative to gene flow, as in cases where habitat discontinuity has led to extensive population isolation (case III of Hutchinson & Templeton 1999). The relationship between geographical (in km) and genetic ($F_{ST}$) distances was evaluated through a Mantel test (with 10,000 permutations), as implemented by the software ibd 1.52 (Bohonak 2002). The reduced major axis regression was used to evaluate the strength of the relationship and to calculate regression statistics. Moreover, to determine whether the variability of the $F_{ST}$ estimates increases with geographical distance, the residuals obtained from the regression of the geographical distances vs. $F_{ST}$ estimates were correlated with the geographical distances, using the same statistical procedure as above.

**Allozyme electrophoresis**

Standard horizontal starch gel (10%) electrophoresis was performed in order to screen the samples for their allozyme variation. The following enzyme systems ($n = 16$) and loci ($n = 21$) were resolved: Glycerol-3-phosphate dehydrogenase ($G3pdh$; EC 1.1.1.8), Lactate dehydrogenase ($Ldh-1$ and $Ldh-2$; EC 1.1.1.27), Malate dehydrogenase ($Mdh-1$ and $Mdh-2$; EC 1.1.1.37), Malate dehydrogenase NADP+-dependent ($Mdhp-1$ and $Mdhp-2$; EC 1.1.1.40), Isocitrate dehydrogenase ($Icdh-1$ and $Icdh-2$; EC 1.1.1.42), 6-Phosphogluconate dehydrogenase ($6Pgdh$; EC 1.1.1.44), Superoxide dismutase ($Sod$; EC 1.15.1.1), Aspartate trans-
aminase (Aat-1 and Aat-2; EC 2.6.1.1), Creatine kinase (Ck; EC 2.7.3.2), L-LeucylGlycylGlycine Peptidase (Pep-B; EC 3.4.11.23), L-Phenylalanyl-L-proline Peptidase (Pep-D; EC 3.4.13.9), Carbonic anhydrase (Ca; EC 4.2.1.1), Aconitase (Aco; EC 4.2.1.3), Mannose phosphate isomerase (Mpi; EC 5.3.1.8), Glucose phosphate isomerase (Gpi; EC 5.3.1.9), Phosphoglucomutase (Pgm-2; EC 5.4.2.2). Zymograms were visualized using enzyme-specific staining procedures, following techniques in Harris & Hopkinson (1976) and Richardson et al. (1986). Alleles at each locus were designated by their mobility (in mm, standardized conditions) relative to the most common allele (named 100) from a reference population (Bagno di Romagna).

**Allozyme data analysis**

Estimates of allele frequencies and population genetic variability — mean observed heterozygosity, Nei’s (1978) unbiased estimate of expected heterozygosity, and the proportion of polymorphic loci — were computed with the software biosys-2 (Swofford & Selander 1999). Average number of alleles per locus was obtained for each population from the allelic richness (a measure of the number of alleles independent of sample size), estimated with fstat-2.9 (Goudet 2001). Exact significance probabilities for departures from the Hardy–Weinberg equilibrium were computed for each locus in each sample.

A hierarchical cluster analysis of populations was carried out using the neighbour-joining method (NJ) with Cavalli-Sforza & Edwards (1967; CSE) chord distances. Nodal support was assessed by 1000 bootstrap replicates, obtained using the subroutines with the support was assessed by 1000 bootstrap replicates, obtained with the software bootdist (Swofford & Selander 1999). Average number of alleles per locus was obtained for each population from the allelic richness (a measure of the number of alleles independent of sample size), estimated with fstat-2.9 (Goudet 2001). Exact significance probabilities for departures from the Hardy–Weinberg equilibrium were computed for each locus in each sample.

A hierarchical cluster analysis of populations was carried out using the neighbour-joining method (NJ) with Cavalli-Sforza & Edwards (1967; CSE) chord distances. Nodal support was assessed by 1000 bootstrap replicates, obtained with the bootstrap option in biosys-2. The consensus NJ tree was then obtained using the subroutines NEIGHBOUR and CONSENSE in PHYLIP 3.5c (Felsenstein 1993). Because tree-building methods force populations into a dichotomic branching pattern, we also computed a principal component analysis (PCA) of allele frequencies by means of the software PCAGEN 1.2 (Goudet 1999), in order to detect potentially intergraded populations relative to the main lineages. The statistical significance of each axis was evaluated over 10 000 randomizations.

The amount of genetic differentiation between populations was investigated by calculating pairwise $F_{ST}$ values as estimated by the parameter $\theta$ of Weir & Cockerham (1984), using the software fstat 2.9 (Goudet 2001). The significance of the $F_{ST}$ estimates was assessed by 1000 permutations. As for mitochondrial data, the hierarchical partitioning of the genetic variance was evaluated by an AMOVA analysis (Excoffier et al. 1992), as implemented by the software ARLEQUIN 2.0 (Schneider et al. 1999).

The relationships between geographical and genetic distances were investigated using the same methodological framework as for mtDNA data.

### Results

**mtDNA**

**Sequence variation and population genetic diversity.** For all the 148 specimens studied, we obtained a fragment of 598 base pairs of the mitochondrial cytochrome $b$ gene, corresponding to the positions 14286–14884 of the previously published Bombina bombina mitochondrial genome (GenBank Accession no. NC 006402). Thirteen distinct haplotypes were found (GenBank Accession nos DQ320148–DQ320160), showing a sequence divergence ranging from 0.2 to 1.0% (the overall average being 0.5%). These were identified by 13 variable positions, with all substitutions being transitions, nine located at the third codon position, three at the second position and one at the first position. Substitutions located at the first and second position gave rise to four amino-acidic replacements.

Estimates of haplotype ($h$) and nucleotide diversity ($\pi$) for each sampled population are given in Table 2. Twelve out of 17 populations analysed showed a complete lack of variability at the mtDNA gene fragment analysed and are all located north of the Crati-Sibari plain (samples 1–10), except for the two southernmost samples (16–17). The only populations showing appreciable mtDNA diversity are those located in central Calabria (samples 11–15). Moreover, both $h$ and $\pi$ were particularly high in samples 13 and 14, located, respectively, slightly north and slightly south of the Catanzaro plain.

**Haplotype network and geographical distribution of haplotypes.** The minimum spanning network is presented in Fig. 1(b). A reticulation between haplotypes h5 and h7, caused by an inferred haplotype, was removed following the criteria of Crandall & Templeton (1993).

The geographical distribution of haplotypes is presented in Fig. 1(a) and Table 2. The haplotype showing the highest number of connections in the network was h6, which was restricted to the three populations (11–13) sampled in the geographical area between the Crati-Sibari plain and the Catanzaro plain (hereon CS- and CA-plain, respectively), where it is also the most common haplotype (42.3% of the specimens analysed from this area). Within this geographical area, five other haplotypes were found (h7, h8, h9, h10 and h11), three of which were never found elsewhere (h7, h8 and h10). Among the populations sampled along the Apennine chain north of the CS-plain (1–10), the most common haplotype was h1, which was fixed in six out of 10 populations (2–5, 9–10), but which was never found south of CS-plain. Each of the other four populations sampled north of this area (the sample 1 from the northern Apennines and the three samples from the Volturno river basin 6, 7, 8) showed a single, fixed private haplotype. Among the four populations sampled south of the CA-plain (14–17), the
haplotype h11 was by far the most common, being shared by 92.7% of the specimens analysed. Of the other three haplotypes observed in this area (each from a single specimen), two were restricted to it, and the third (h9) was shared with an individual sampled north of the CA-plain (from sample 13).

Population genetic structure and demographic analyses. The overall pattern of population differentiation was strong and significant, with an $F_{ST} = 0.86$ ($P < 0.01$). The $F_{ST}$ estimates among population pairs (Table 3) showed the widest possible variation, ranging from 0.00 to 1.00. Only 0.00 and 1.00 values were found in pairwise comparisons between populations located north of the CS-plain (samples 1–10). Lower-than-average values of $F_{ST}$ were observed among samples located within the two main Calabrian plains (samples 11–13) and, particularly, among samples located south of the CA-plain (samples 14–17). The differentiation between the samples located on alternative sides of the two Calabrian plains was instead very pronounced, the majority of pairwise $F_{ST}$ values exceeding 0.60, with a few exceptions in the comparisons between samples located slightly north and south of the CA-plain (samples 11, 13 vs. 14, 15).

The AMOVA revealed that a large portion (63.65%) of the overall genetic variance observed in the mtDNA gene fragment analysed can be attributed to the difference between the three population groups, as defined by the two major
phylogeographic discontinuities observed: north of the CS-plain (samples 1–10), between the CS- and CA-plain (samples 11–13), south of the CA-plain (samples 14–17). Nevertheless, a non-negligible part of the overall variation is due to differences among populations within groups (25.71%) and to differences within populations (10.64%), all variance components being highly significant ($P < 0.01$).

The association between genetic ($F_{ST}$) and geographical (in km) distance matrices was investigated at the level of both the entire data set and by analysing separately the samples located on different sides of the CS-plain. A slight but significant correlation was observed at the level of the entire data set ($r = 0.24$, $P < 0.05$; see Fig. 2). The association become stronger and significant when restricting the analysis to the populations sampled south of the CS-plain ($r = 0.60$, $P < 0.01$), with no association being detected north of this area. Nevertheless, in no case was a significant correlation observed between residuals of the regressions and geographical distances.

Possible signatures of past demographic changes were investigated for the whole data set, as well as for samples north of the CS-plain, between the two plains, and south of the CA-plain separately. The statistics $F_{S}$ (Fu 1997) and $R_{2}$ (Ramos-Osins & Rozas 2002) in no case indicated the occurrence of significant past population expansions (all $P > 0.05$).

### Table 3
Pairwise estimates of $F_{ST}$ for the 13 samples of Bombina pachypus for which both mitochondrial and allozymes data were collected (samples are numbered as in Table 1). Allozyme estimates (according to Weir & Cockerham 1984) are above the diagonal; mitochondrial estimates (incorporating Tamura & Nei 1993 genetic distances) are below the diagonal.

<table>
<thead>
<tr>
<th>Samples</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
<th>9</th>
<th>11</th>
<th>12</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.04</td>
<td>0.19</td>
<td>0.00</td>
<td>0.12</td>
<td>0.06</td>
<td>0.76</td>
<td>0.67</td>
<td>0.52</td>
<td>0.54</td>
<td>0.46</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td>0.70</td>
<td>0.59</td>
<td>0.39</td>
<td>0.46</td>
<td>0.36</td>
</tr>
<tr>
<td>3</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.06</td>
<td>0.66</td>
<td>0.54</td>
<td>0.32</td>
<td>0.42</td>
<td>0.32</td>
</tr>
<tr>
<td>4</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.11</td>
<td>0.06</td>
<td>0.62</td>
<td>0.52</td>
<td>0.32</td>
<td>0.41</td>
<td>0.33</td>
<td>0.65</td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.04</td>
<td>0.00</td>
<td>0.65</td>
<td>0.54</td>
<td>0.34</td>
<td>0.41</td>
<td>0.33</td>
<td>0.66</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.67</td>
<td>0.57</td>
<td>0.35</td>
<td>0.43</td>
<td>0.34</td>
<td>0.64</td>
</tr>
<tr>
<td>9</td>
<td>0.92</td>
<td>0.84</td>
<td>0.81</td>
<td>0.78</td>
<td>0.78</td>
<td>0.91</td>
<td>0.82</td>
<td>—</td>
<td>0.00</td>
<td>0.25</td>
<td>0.25</td>
<td>0.26</td>
<td>0.46</td>
</tr>
<tr>
<td>12</td>
<td>0.84</td>
<td>0.76</td>
<td>0.72</td>
<td>0.70</td>
<td>0.70</td>
<td>0.83</td>
<td>0.74</td>
<td>0.49</td>
<td>—</td>
<td>0.15</td>
<td>0.16</td>
<td>0.15</td>
<td>0.39</td>
</tr>
<tr>
<td>14</td>
<td>0.90</td>
<td>0.83</td>
<td>0.79</td>
<td>0.75</td>
<td>0.75</td>
<td>0.89</td>
<td>0.81</td>
<td>0.40</td>
<td>0.56</td>
<td>—</td>
<td>0.00</td>
<td>0.00</td>
<td>0.29</td>
</tr>
<tr>
<td>15</td>
<td>0.92</td>
<td>0.88</td>
<td>0.86</td>
<td>0.84</td>
<td>0.84</td>
<td>0.91</td>
<td>0.86</td>
<td>0.60</td>
<td>0.67</td>
<td>0.00</td>
<td>—</td>
<td>0.00</td>
<td>0.19</td>
</tr>
<tr>
<td>16</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.80</td>
<td>0.76</td>
<td>0.17</td>
<td>0.01</td>
<td>—</td>
<td>0.19</td>
</tr>
<tr>
<td>17</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.83</td>
<td>0.79</td>
<td>0.24</td>
<td>0.04</td>
<td>0.00</td>
<td>—</td>
</tr>
</tbody>
</table>

* $P < 0.05$ after 1000 permutations. Values ≤ 0 were in all cases set to 0.

**Hardy–Weinberg equilibrium** were detected across samples or loci. At all the loci polymorphic in more than one sample (all but Mdh-1 and Mpi), alleles were found to be completely (Mdh-2157, Idh-2107, Pgm-2194, Gpi214) or almost completely (Ldh-172, Pep-B110) confined to Calabria, south of the CS-plain. Among these, the frequency of allele Ldh-172 appeared to gradually decrease from south to north. When regressed over the geographical distance (in km) from the southernmost sample, the frequency of this allele appeared highly and significantly correlated with geography (Spearman $r = -0.81$; $P < 0.05$). At the locus Gpi, high frequencies of the allele Gpi214 were observed in the two samples surveyed in central Calabria (11–12), whereas its frequency was never above 0.23 among samples located south of the CA-plain.

Estimates of population genetic variability are given in Table 2. All the samples surveyed along the Apennines north of Calabria completely or almost completely lack genetic variation at all the parameters estimated. For instance, $H_{E}$ observed among samples from this area was never above 0.02 (sample 9). By contrast, among samples from Calabria, $H_{E}$ was never below 0.06, a value observed in the southernmost sample. The highest values of $H_{E}$ were observed among the samples located slightly south of the CA-plain (14–16).

**Population genetic structure and differentiation.** The neighbour-joining analysis conducted on the CSE genetic distances (Fig. 3) showed that two main population clusters can be identified at the allozyme level, comprising samples located alternatively north and south of the CS-plain. The same two population groups were also evident along the
first principal component (explaining 74.6% of the total genetic variance), resulting from the PCA (see Fig. 4). Along the second principal component (explaining a further 19.5% of the total genetic variance), the samples located south of the CS-plain are arranged approximately according to their geographical distribution, with the more geographically distant populations at the extremes and the populations from central Calabria being intermediate.

The overall pattern of population differentiation was found to be strong and highly significant, with an $F_{ST}$ value among all populations of 0.47 ($P < 0.01$). Pairwise estimates of $F_{ST}$ are given in Table 3. A wide variation was found in pairwise comparisons, with $F_{ST}$ ranging from 0 to 0.76. Almost all comparisons among samples from the northern and central Apennines (samples 1–9) were very low and not significant, whereas $F_{ST}$ values between these samples and those located in Calabria south of the CS-plain were never below 0.32. With the single exception of comparisons involving the southernmost sample, pairwise comparisons among Calabrian samples located on the same side of the CA-plain always gave rise to $F_{ST} = 0$, whereas comparisons between samples located on different sides of the CA-plain ranged between 0.15 and 0.46.

The hierarchical amova, based on the three groups defined for mtDNA data, revealed that 52.44% of the overall allozyme genetic variance can be attributed to differences between groups, 43.85% to the within-population level of variation, and 3.71% to differences between populations within groups, all variance components being significant ($P < 0.01$).

At the level of the entire data set, a significant association was found between the genetic and geographical distance matrices ($r = 0.56; P < 0.01$; see Fig. 2). This association was, however, stronger when restricting the analysis only to the populations located south of the CS-plain ($r = 0.85; P < 0.01$). By contrast, no significant association was found when analysing separately those populations surveyed north of the CS-plain. In no case were significant relationships observed between the residuals of the regressions and geographical distances.
Discussion

Genetic diversity and population structure

In their anecdotal work, Nascetti et al. (1982) suggested a higher genetic diversity in populations of the Apennine yellow-bellied toad from Calabria compared to those from the northern portion of the species’ range. This pattern was interpreted as evidence of a postglacial northward range expansion from a refugial range in Calabria. Our analyses concerning both mtDNA and allozymes support this general scenario, and provide additional data to gain further insight into the evolutionary processes that have shaped the present population genetic structure of Bombina pachypus.

Rapid range expansions are expected to produce a pattern of reduced genetic variation and population structure within the recently colonized areas because of serial founder events, particularly when long-distance colonization occurs (Hewitt 1996, 1999; Ibrahim et al. 1996; Hewitt & Ibrahim 2001). As a consequence, the pattern obtained when plotting pairwise genetic vs. geographical distances would reflect a lack or weakness of correlation between the two distance measures. The degree of scattering will be more pronounced if the habitat fragmentation within the newly colonized areas is strong, as will be also the isolation among the newly established populations (Hutchinson &
Templeton 1999). For the populations north of Calabria, allozyme data indicated an absence of correlation between pairwise genetic vs. geographical distances and a narrow degree of scattering, as expected from the paucity of genetic differentiation. Following the above arguments, this genetic pattern appears fully consistent with a scenario of recent and rapid colonization of the northern portion of the Italian Peninsula by refugial populations in Calabria. This scenario is also supported by mitochondrial data. From the network topology (Fig. 1), the haplotype h1 appears to be derived when compared to h6 (which is confined to central Calabria) and ancestral compared to the other haplotypes found north of the CS-plain (Crandall & Templeton 1993; Posada & Crandall 2001). The presence of this haplotype fixed in six out of 10 populations, also comprising the southernmost samples from this area (9 and 10), and its wide distribution spanning the northern and central Apennines (see Fig. 1 and Table 2), both accord with a scenario of a recent founder event followed by rapid colonization of the northern habitats (e.g. Cann et al. 1987; Austerlitz et al. 1997; Edmonds et al. 2004). The fact that the statistics $F_S$ (Fu 1997) and $R_2$ (Ramos-Onsins & Rozas 2002) failed to identify traces of a recent expansion may be due to the very high degree of population subdivision shown by the mtDNA (Ray et al. 2003).

The only populations showing appreciable levels of genetic diversity were found in Calabria south of the CS-plain. Furthermore, all the analyses concur in showing that the genetic variation in these populations has a strong geographical component, and a north–south intergradation (see Figs 2 and 4). Distinguishing between primary and secondary intergradation can be difficult, particularly when extensive intermixing has occurred and there is weak palaeogeographic evidence. However, in the present case both genetic and historical information suggest the occurrence in Calabria of an allopatric differentiation, followed by secondary contact. The CA-plain appears as the source of a major phylogeographic break, according to both the geographical distribution of haplotypes (Table 2 and Fig. 1) and the pattern of population differentiation (Table 3). A significant pattern of isolation by distance was detected among Calabrian samples, according to both mtDNA and allozymes. Nevertheless, the lack of increased variability among differentiation estimates with increasing geographical distance allows us to reject the migration–drift equilibrium under the Hutchinson & Templeton (1999) criteria. Clinal variation along a north–south axis was observed at the single locus Ldh-1, whereas at the locus Gpi a significant variation of the allele frequencies was observed among samples located on different sides of the CA-plain. Simulation studies have shown that primary intergradation does not produce clinal variations at neutral loci, whereas it is yielded by secondary contacts between previously differentiated lineages (Nichols & Hewitt 1994; Durrett et al. 2000). The absence of a clear, straightforward linear transect (see also Santucci et al. 1996), and the small number of samples from Calabria do not allow us to perform classical statistical cline analyses. Moreover, inferences from single-locus clines are hardly reliable, particularly in relation to the neutral/non-neutral nature of the cline (e.g. Endler 1977; Kruuk et al. 1999). A closer sampling in southern Italy and the application of new molecular markers (microsatellites) are therefore in progress, and could help clarify this issue. Finally, samples located slightly south of the CA-plain are among the most variable of the entire data set (Table 2), a pattern consistent with a history of past fragmentation, followed by secondary contact at the level of the CA-plain. This finding of a high genetic variability also allows us to discard an alternative hypothesis of a recent colonization of southern Calabria.
possibly suggested by the wide distribution of the haplotype h11 south of the CA-plain.

From the overall genetic pattern reported, it seems that two geographical areas, the CS- and CA-plains, have played a major role as historical barriers to dispersal, with the latter also giving rise to allopatric differentiation. The present knowledge of palaeogeographic history in southern Italy strongly supports this hypothesis (Ghisetti 1979, 1981; Tortorici 1981; Caloi et al. 1989), depicting the CA- and CS-plains as two of the main grabens in southern Italy. During the Pleistocene, they underwent alternate rising and sinking movements and were repeatedly marine-flooded as a consequence of glacio-eustatic sea level fluctuations, which rendered environmental conditions unfavourable to the dispersal of terrestrial species along the north–south axis until the late Pleistocene (Caloi et al. 1989). Central and northern Calabria were connected during the early Pleistocene by a narrow land bridge (the emerging Catena Costiera), owing to the submersion of the CS-plain. The gradual process of erosion and deposition from the adjacent lands led the two plains to finally emerge during the middle-late and late Pleistocene. In the case of the Apennine yellow-bellied toad, it is plausible that unfavourable environmental conditions persisted even during the early phases after the plains’ emergence, as suggested by the altitudinal distribution of the species, presently absent from the main Italian lowlands (e.g. Guarino et al. 2006). If we apply the clock calibration of 1.3–1.6% per million years proposed by Garcia-Paris & Jockusch (1999) for cytochrome b of Discoglossid frogs to the 0.5% of divergence among B. pachypus lineages, the coalescence time would be dated at 313 000–385 000 years. Although we are aware of the inaccuracies of the molecular clocks (e.g. Ayala 1997, 1999; Gibbons 1998), these estimates would suggest that the Pleistocene, in agreement with the palaeogeographic

B. pachypus

Gibbons 1998), these estimates would suggest that the

inaccuracies of the molecular clocks (e.g. Ayala 1997, 1999;

at 313 000–385 000 years. Although we are aware of the

B. pachypus

b

proposed by Garcia-Paris & Jockusch (1999) for cytochrome

of Discoglossid frogs to the 0.5% of divergence among

B. pachypus

lineages, the coalescence time would be dated at

313 000–385 000 years. Although we are aware of the

inaccuracies of the molecular clocks (e.g. Ayala 1997, 1999;

Gibbons 1998), these estimates would suggest that the

above-mentioned events in the history of populations of

B. pachypus

have occurred approximately since the Middle

Pleistocene, in agreement with the palaeogeographic

scenario so far described. Calabria south of the CS-plain

has been identified as the refuge for B. pachypus during

glacial stages. The crossing of the CA-plain, with the

occurrence of the secondary contact, probably occurred

after the Eemian sea regression (~110 000 years bp; Van

Andel & Tzedakis 1996), whereas the recolonization of

northern habitats occurred later, at the end of the last

glaciation (the Würm).

Interestingly, growing evidence confirms the main role

of the two Calabrian plains in shaping the patterns of

distribution of species and of evolutionary lineages within

them. According to the classical biogeography, several
taxa found in Calabria south of the CA-plain have greater

affinities with Sicilian taxa than with those from central

and northern Calabria, and the two plains represent the

southern or northern range edge of many species (e.g.

Audisio 1984; Pignatti 1984; Caloi et al. 1989). More recently,
genetic investigations have provided further evidence that

these plains have repeatedly acted as barriers to

dispersal (Santucci et al. 1996 and maybe Podnar et al.

2005). These findings therefore suggest that the two plains,

and particularly the CA-plain, could be sites of suture

zones (Remington 1968), although further investigations

focused on more taxa will be necessary to strengthen this

hypothesis.

At first glance, the genetic pattern we observed within

B. pachypus

could be regarded as a typical pattern of ‘southern

richness, northern purity’ (Hewitt 1999). However, our

results also indicated that this southern richness is at least

in part due to allopatric differentiation within the southern

glacial refugium followed by intermixing of previously

differentiated lineages. A prolonged demographic stability

of southern populations (e.g. Hewitt 1996, 2000) cannot

explain the observed pattern of genetic diversity by itself,

although neither can it be completely ruled out. Most

probably, both allopatric differentiation and demographic

stability have played a role, and this situation could be

more widespread than previously thought. In particular,
allopatric differentiation may have played an important

role in shaping genetic diversity also within the southern

Italian refugium, as previously proposed for the Iberian

one (Guillaume et al. 2000; Sanz et al. 2000; see Gómez &

Lunt, in press for a review of evidence), thus suggesting

that this pattern could be a common feature of refugial

ranges. Finally, a point that deserves particular attention is

the necessity for an appropriate sampling scheme so that

this pattern and its strength may be properly appreciated

(see also Taberlet 1998; Gómez & Lunt, in press). In fact,

several case studies showing higher genetic variability in

populations from southern Italy have been carried out

with only one or very few samples from that region (e.g.

Ragghianti & Wake 1986; Capula & Ceccarelli 2003), or

with samples from northern Calabria and Sicily, but not

from central and southern Calabria (e.g. Randi et al. 2003;

Fritz et al. 2005). With such sampling schemes, it is difficult

to disentangle the complex evolutionary dynamics that

could have shaped genetic diversity within this southern

refugium.

Conservation implications

The conservation status of B. pachypus has recently been a

cause for concern. The main threats to this species include

habitat loss, degradation and fragmentation, with con-

sequent population isolation, and the recent finding of

chytridiomycosis in samples from the northern Apennines

(e.g. Caputo et al. 1985; Doria & Salvadò 1994 and refer-

ences therein; Sarrocco & Bologna 2000; Stagni et al. 2002;

Barbieri et al. 2004).

The importance of genetic diversity for the long-term

persistence of populations, species and (more recently)
ecosystem functions, has long been acknowledged (e.g. Young & Clarke 2000; Frankham et al. 2002; Frankham 2005). Our results unequivocally indicate Calabria south of the CS-plain as the major reservoir of genetic diversity for B. pachypus. This geographical area acted as a refugium during the ice ages, and has been the site of various evolutionary processes that are among the most influential in the history of this species, such as past fragmentation, range expansion and secondary contact. Therefore, protecting this genetic diversity reservoir is of the greatest importance for the conservation of the species. Unfortunately, although a large portion of the region is devoted to protected areas, the effectiveness of this protection is in many cases doubtful. For instance, the southernmost population we sampled (17) is located on the southern edge of the Aspromonte National Park, but the imminent activation of a dam will soon erase the sampling site and its surroundings. Moreover, within the Serre National Park (a little south of the CA-plain), several breeding sites have recently disappeared due to activities related to tree-cutting and water-draining practices (Bagnoli C. personal communication).

As discussed above, our data show that the low genetic variability of samples located north of the CS-plain is of historical origin. Nevertheless, a role for a current reduction of gene flow/dispersal is suggested by the finding of fixed private haplotypes in relatively close samples, as for example in the three populations from the Voltorno river basin (Fig. 1, Tables 2 and 3). Interestingly, this area has been heavily affected by human activities in recent decades, and the poor conservation status of the species from this region has recently been reported (Barbieri et al. 2004 and references therein).

It is worth noting that, in the categorization of the conservation status of this species made by Barbieri et al. (2004; basing on demographic and ecological data and according to the IUCN categories), the Apennine yellow-bellied toad was judged ‘not threatened’ only in Calabria, whereas north of this region it was categorized as ‘vulnerable’.

On the whole, our results suggest the need for careful planning of conservation practices for the Apennine yellow-bellied toad. These ought to include: (i) an accurate assessment of the patterns of population connectivity and its enhancement where necessary, (ii) the long-term monitoring of trends in population demography and fitness-related traits and, where also suggested by this latter investigation, and (iii) the restoration of population genetic variation. This latter management practice has the potential to reverse negative population trends, as shown by studies in both natural and laboratory-reared populations (e.g. Madsen et al. 1999; Hedrick & Kalinowski 2000). However, before this practice could be carried out, fitness-related phenotypic differences among populations ought to be properly evaluated, and the appropriate source population should be identified.

Acknowledgements

We are grateful to Francesca Zangari for insightful discussions and suggestions. We also thank Trevor J. Beebee and three other anonymous reviewers, who greatly improved an early version of the manuscript. Sincere thanks are also due to Pietro Bagalà, Claudio Bagnoli, Paola Bellini, Fulvio Chiacchiera and Marco Puleo for their kind help during field sampling and/or manuscript preparation, and to Mark Ebbert who reviewed the English. This work was funded by MIUR (Italian Ministry of University and of Scientific and Technological Research).

References


© 2006 The Authors

Journal compilation © 2006 Blackwell Publishing Ltd
Garcia-Paris M, Jockusch EL (1999) A mitochondrial DNA per-

Goudet J (2001)

Goudet J (1999)

Gómez A, Lunt DH (in press) Refugia within refugia: patterns of


Fu YX (1997) Statistical tests of neutrality of mutations against

Ghisetti F (1981) L’evoluzione strutturale del bacino plio-

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

data.


3.5c. University of Washington, Seattle.

Ferreire R, Dieckmann U, Couvet D (2004) Evolutionary Conserva-


Frankham R, Ballou JD, Briscoe DA (2002) Introduction to Conser-

Fritz U, Fatuzzo T, Guicking D et al. (2005) A new cryptic species of
pond turtle from Southern Italy, the hottest spot in the range of the genus Emys (Reptilia, Testudines, Emydidae). Zoologica Scripta, 34, 351–371.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection.

Genetics, 147, 915–925.


Ghisetti F (1981) L’evoluzione strutturale del bacino plio-


Science, 279, 38–39.


Goudet J (1999) PCaGen, a computer package which performs principal component analysis (Pca) on gene frequency data. Available at http://www2.unil.ch/izea/software/pcaogen.html.

Goudet J (2001) Fstat, a program to estimate and test gene diver-
sities and fixation indices (version 2.9.3). Available at http://

www.unil.ch/izea/software/fstat.html.

Guarino FM, Picariello O, Pellegrini M (2006) Apennine yellow-


Hedrick PW, Kalinowski ST (2000) Inbreeding depression in con-


Lanza B, Vanni S (1991) Notes on the biogeography of the Medi-


Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection.

Genetics, 147, 915–925.

Garcia-Paris M, Jockusch EL (1999) A mitochondrial DNA perspec-


Ghisetti F (1981) L’evoluzione strutturale del bacino plio-


Science, 279, 38–39.


Goudet J (1999) PCaGen, a computer package which performs principal component analysis (PCA) on gene frequency data. Available at http://www2.unil.ch/izea/software/pcaogen.html.

Goudet J (2001) FSTAT, a program to estimate and test gene diver-
sities and fixation indices (version 2.9.3). Available at http://

www.unil.ch/izea/software/fstat.html.

Guarino FM, Picariello O, Pellegrini M (2006) Apennine yellow-

Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics, 85, 583–589.


This page is part of PhD thesis of D. Canestrelli and of MsC thesis of V. Costantini, carried out at the laboratory of ecology of the Tuscia University, headed by professor G. Nascetti. The main interest of D. Canestrelli, now research associate, is the investigation of factors underlying geographical distribution of genetic diversity among Italian amphibians. R. Cimmerata is senior researcher in the same lab and she is mainly interested in biodiversity conservation and relationships between environmental stress and population genetic structure. G. Nascetti is involved in ecological genetics, with his interests encompassing co-evolution, speciation, phylogeny and conservation genetics.