Newts in time and space: the evolutionary history of Triturus newts at different temporal and spatial scales

Espregueria Themudo, G.

Citation

Espregueria Themudo, G. (2010, March 10). *Newts in time and space: the evolutionary history of Triturus newts at different temporal and spatial scales*. Retrieved from https://hdl.handle.net/1887/15062

Version: Corrected Publisher's Version

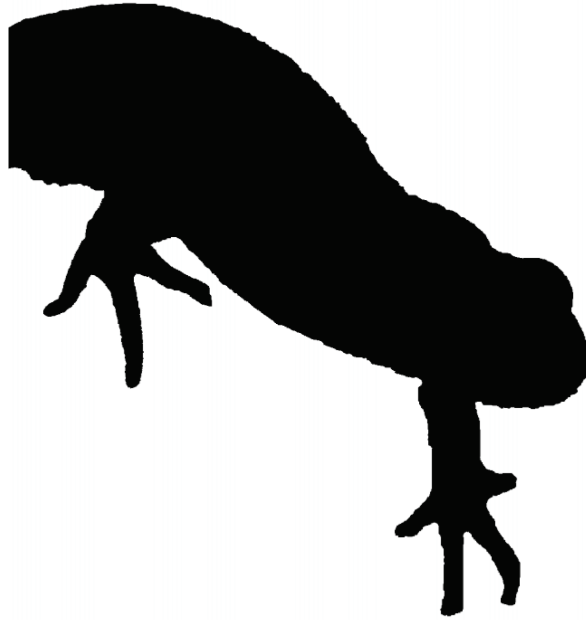
License: Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden

Downloaded from: https://hdl.handle.net/1887/15062

Note: To cite this publication please use the final published version (if applicable).

Newts in time and space:

The evolutionary history of *Triturus* newts at different temporal and spatial scales



Gonçalo Espregueira Themudo

Espregueira Themudo, Gonçalo

Newts in time and space: The evolutionary history of *Triturus* newts at different temporal and spatial scales

This thesis was financially supported by a Fundação para a Ciência e para a Tecnologia (FCT) PhD grant (ref. SFRH/BD/16894/2004)

Thesis, Leiden University

Cover illustration: © Nick Poyarkov

Photo credits (introduction): Figs. 2, 4 - Piet Spaans; Fig. 3 - Christian Herzog; Fig. 7 - Clara Cartier; Licensed under a Creative Commons Attribution ShareAlike 2.5 License
(http://creativecommons.org/licenses/by-sa/2.5/)
Fig. 5 - Ben Wielstra; Fig. 8 - Ricardo Pereira.
© All rights reserved. Reproduced with permission.

Newts in time and space: The evolutionary history of *Triturus* newts at different temporal and spatial scales

Proefschrift
ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Prof. Mr. P. F. van der Heijden,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 10 maart
klokke 16:15 uur

door

Gonçalo Espregueira Themudo

geboren te Vila Nova de Gaia, Portugal, in 1979

Promotiecomissie

Promotor:	Prof. Dr. E. Gittenberger (Universiteit Leiden)
Co-promotor:	Dr. J. W. Arntzen (Nationaal Natuurhistorisch Museum *Naturalis*, Leiden)
Referent:	Prof. Dr. S. B. J. Menken (IBED, Universiteit van Amsterdam)
Overige leden:	Prof. Dr. P.M. Brakefield (Universiteit Leiden Prof. Dr. C.J. ten Cate (Universiteit Leiden) Dr. J. A. Godoy (Estación Biológica Doñana, CSIC, Sevilla, Spain)

Aos meus pais

Contents

	Nederlands samenvatting	
CHAPTER 1	Introduction and Summary	1
CHAPTER 2	The phylogeny of crested newts (*Triturus cristatus* superspecies): nuclear and mitochondrial genetic characters suggest a hard polytomy, in line with the paleogeography of the centre of origin. *Arntzen, J.W., B. Wielstra & G. Espregueira Themudo*	17
CHAPTER 3	A combination of techniques proves useful for the development of nuclear genes for the newt genus *Triturus* *Espregueira Themudo, G., W. Babik & J. W. Arntzen*	45
CHAPTER 4	Multiple nuclear and mitochondrial genes resolve the branching order of a rapid radiation of crested newts (Triturus, Salamandridae) *Espregueira Themudo, G., B. Wielstra & J. W. Arntzen*	55
CHAPTER 5	Newts in space: Geography helps in the distinction between gene flow and incomplete lineage sorting *Espregueira Themudo, G., A. Bickham-Baird & J.W. Arntzen*	79
CHAPTER 6	Current gene flow across a newt hybrid zone follows local ecological conditions *Espregueira Themudo, G. & J. W. Arntzen*	95
CHAPTER 7	Environmental parameters that determine species geographical range limits as a matter of time and space. *Arntzen, J. W. & G. Espregueira Themudo*	121
CHAPTER 8	Molecular identification of marbled newts and a justification of species status for *Triturus marmoratus* and *T. pygmaeus* *Espregueira Themudo, G. & J. W. Arntzen*	137
CHAPTER 9	Newts under siege: range expansion of *Triturus pygmaeus* isolates populations of its sister species *Espregueira Themudo, G. & J. W. Arntzen*	157

Acknowledgments 177

Curriculum Vitae 178

NEDERLANDSE SAMENVATTING

Dit proefschrift behandelt de evolutionaire geschiedenis van het geslacht *Triturus* (Amphibia, Salamandridae, Pleurodelinae), de kamsalamanders. Het omvat naast de Nederlandse en Engelse samenvattingen acht hoofdstukken, die inmiddels vrijwel allemaal afzonderlijk zijn gepubliceerd in internationale wetenschappelijke tijdschriften. Na het inleidende eerste hoofdstuk komt in de hoofdstukken 2-5 de fylogenie van de soorten aan bod, en worden de evolutionaire processen besproken die van invloed zijn (geweest) op de soortvorming. Hoofdstuk 6 vormt de verbinding met de hoofdstukken 7-9, waarin de verhoudingen tussen de verspreidingsgebieden van de diverse soorten worden besproken in het licht van bepaalde omgevingsfactoren. De nadruk ligt hierbij op de invloed van de omgeving op de onderlinge competitie, en de daarmee samenhangende grenzen van hun verspreidingsgebieden.

De fylogenetische verwantschappen tussen de *Triturus* soorten zijn nog niet geheel duidelijk. In hoofdstuk 2 wordt op basis van allozymen en mitochondriale DNA-kenmerken getracht hier meer helderheid in te brengen. De resultaten suggereren dat de soortvorming in deze groep in een korte periode aan het einde van het Midden Mioceen heeft plaatsgevonden. De fylogeografische reconstructies van het vermoedelijke 'centre of origin' (de Balkan) ondersteunen deze hypothese. De verspreiding en de relatief grote genetische afstand tussen *T. carnifex* en *T. macedonicus* zijn daarnaast aanleiding om deze taxa als zelfstandige soorten te beschouwen.

Om het hypothetische scenario dat in hoofdstuk 2 werd gepostuleerd nader te toetsen, werd onderzoek gedaan naar een aantal onafhankelijke genetische kenmerken. Hoofdstuk 3 beschrijft in dit verband een studie naar genetische 'markers' die voor dit doel geschikt zouden kunnen zijn. Van de vijftig geteste 'markers' bleken er vijf potentieel geschikt voor het onderzoek aan *Triturus*.

Uitgaande van hoofdstuk 2, en gebruik makend van de gegevens uit hoofdstuk 3, wordt in hoofdstuk 4 getracht de evolutionaire geschiedenis van *Triturus* te reconstrueren. De resultaten laten zien dat, met uitzondering van de mitochondriale 'markers', de vijf genetische kenmerken verschillende scenario's ondersteunen. Met behulp van fylogenetische netwerken worden de alternatieve scenario's nader bestudeerd. Tevens wordt met behulp van een Bayesiaanse hiërarchische methode de informatie uit de afzonderlijke genetische kenmerken gekombineerd in een fylogenie reconstructie. Deze reconstructie ondersteunt de in hoofdstuk 2 gepostuleerde hypothese van een snelle radiatie, en is bovendien aanleiding om de taxa *T. karelinii* en *T. arntzeni* als aparte soorten te beschouwen.

Een deel van de problemen bij het reconstrueren van de verwantschappen tussen de soorten kamsalamanders wordt veroorzaakt door hybridisatie en introgressie van genetisch materiaal van de ene soort in de andere. De oorspronkelijke verwantschappen worden hierdoor overschaduwd. In hoofdstuk 5 worden daarom de evolutionaire processen 'gene flow' en 'incomplete lineage sorting' met behulp van een nieuwe methode onderzocht. De duidelijke soortgrenzen en de beperkte mogelijkheden tot dispersie bij de *Triturus* soorten maken het mogelijk om met behulp van deze methode onderscheid te maken tussen beide evolutionaire processen.

De hoofdstukken 2-5 hebben duidelijk gemaakt dat de kamsalamanders complexe verspreidings- en verwantschapspatronen laten zien. De zustersoorten *T. marmoratus* en *T. pygmaeus* vormen een subgroep in het geslacht *Triturus* (zie hoofdstuk 2) en komen alleen voor op het Iberisch schiereiland. Omdat het hier slechts twee soorten betreft is het mogelijk om de evolutionaire processen in meer detail te bestuderen. In hoofdstuk 6 wordt onderzocht hoe verschillende ecologische omstandigheden in een hybridisatie zone, de structuur van deze zone en de mate van genetische uitwisseling kunnen beïnvloeden.

Volwassen salamanders zijn relatief eenvoudig te onderscheiden op basis van uiterlijke kenmerken, maar ze zijn lastig te verzamelen. De eieren en larven kunnen daarentegen

eenvoudig in grote aantallen verzameld worden, maar zijn niet gemakkelijk op basis van morfologische kenmerken te identificeren. In hoofdstuk 7 wordt daarom een goedkope en snelle manier beschreven om de grote aantallen monsters te identificeren die de basis vormen voor de studies in hoofdstukken 8-9.

Op basis van genetische, morfologische en verspreidingsgegevens van *T. marmoratus* en *T. pygmaeus* wordt in hoofdstuk 8 onderzocht welke ecologische factoren een rol spelen bij het bepalen van de soortgrenzen. Het onderzoek wijst uit dat de ecologische modellen die de verspreiding van de soorten kunnen verklaren, verschillen tussen vier geografische gebieden aangeven: (1) de contactzone in het noorden, dicht bij Aveiro in Portugal, (2) de rest van het kustgebied (zie ook hoofdstuk 9), (3) het gebied bij de rivier de Tejo en (4) het overige deel van de contactzone die oostwaarts doorloopt tot bij Madrid.

In het kustgebied van Portugal, dicht bij Caldas da Rainha, werd *T. marmoratus* gevonden in een gebied waar tot nu toe gedacht werd dat alleen *T. pygmaeus* voorkwam. De verspreiding van beide soorten in dit gebied werd daarom zeer nauwkeurig in kaart gebracht. De resultaten, beschreven in hoofdstuk 9, laten zien dat het om een kleine enclave van *T. marmoratus* gaat, midden in het verspreidingsgebied van *T. pygmaeus*. Deze populatie is vermoedelijk ontstaan bij een noordwaartse migratie van *T. pygmaeus*, waarbij *T. marmoratus* in een klein gebied is achtergebleven.

CHAPTER 1

INTRODUCTION & SUMMARY

Introduction

Species are confined in all four dimensions of space and time. But while geographical borders can be defined where no more individuals of a certain species can be found, temporal borders are more difficult to define, as they can not be determined directly, but rather inferred from the fossil record, palaeogeography, and genetics. It is difficult to determine when an ancestral species ceases to be and the derived species comes into existence (see, for example, DE QUEIROZ, 2007). Darwin, for example, considered species to be part of a continuum of diversification, without any real border.

Species' distributions are continuous in areas of, for example, favourable habitat, amenable ecological conditions or lack of competitors. Closer to the border, population density starts decreasing, and the distribution will pass from continuous to patchy, until no more individuals are found. In the case of two closely related parapatric species, these empty patches can be filled by related competing species (ARNTZEN, 2006).

Through time, the range of a species contracts and expands, populations split and merge, gene flow stops and restarts. This also follows the suitability of habitat through time. Climatic changes push populations into different areas, with expansions when the climate is more favourable and retractions when conditions are worse. Distribution becomes patchy, then continuous, and then patchy again, over and over in cycles. Given enough time between contractions and expansions, the populations that meet will be different enough from the populations that had split, and reproductive isolation will have developed.

This thesis is a contribution to unravel the phylogenetic history of a genus of newts, at different scales. It starts by taking a broad picture of the history of the genus, and will then *zoom in* into higher and higher detail, going to phylogeography and further into local ecological conditions that determine species range limits together with the

presence of a closely related competitor. Like with species, the limits between these approaches are somewhat arbitrary.

Newts

Newts are part of a family of salamanders (family Salamandridae; subfamily Pleurodelinae) that evolved from other amphibians around a hundred million years ago (STEINFARTZ *et al.*, 2006). The objects of this study are the large-bodied European newts: the crested and the marbled newts. They form the genus *Triturus*, a group that occurs all over Europe and western Asia. *Triturus* was once a larger genus, comprising not only the large-bodied newts, but also other medium and small sized European newts, now forming *Lissotriton*, *Mesotriton* and *Ommatotriton* (GARCÍA-PARÍS *et al.*, 2004); members of the crested newt group were once considered subspecies of *T. cristatus* (ARNTZEN and WALLIS, 1999), and the pygmy marbled newt, *Triturus pygmaeus*, was until just recently considered a subspecies of *Triturus marmoratus* (GARCÍA-PARÍS *et al.*, 2001).

The distribution of the *Triturus* species is essentially parapatric, their ranges only slightly overlap (Figure 1). This pattern repeats itself in every area where two or more members of this genus meet (ARNTZEN and WALLIS, 1991; see also CHAPTERS 2, 8 and 9). The largest area of overlap is between the great crested newt (*T. cristatus*) and

Figure 1 - Distribution of *Triturus* newts in Eurasia. Notice the area of overlap between the marbled newt (*Triturus marmoratus*) and the great crested newt (*T. cristatus*), in France.

the marbled newt (*T. marmoratus*), but in finer spatial detail, the two forms are well separated (ARNTZEN and WALLIS, 1991). The two species have different ecological requirements: marbled newts occur predominantly at forests and hilly terrain with scrub and hedges, while crested newts prefer flat and open areas (ARNTZEN and WALLIS, 1991; JEHLE and ARNTZEN, 2000).

Crested newts

Crested newts are present in most of Europe and western Asia. Their taxonomy has changed considerably over the last two centuries. All of the species now recognized used to be considered subspecies of the great crested newt, *Triturus cristatus* (Laurenti, 1768). Five species are currently recognized: the northern crested newt, *Triturus cristatus* (Laurenti, 1768), the Italian crested newt, *Triturus carnifex* (Laurenti, 1768), the Danube crested newt, *Triturus dobrogicus* (Kiritzescu, 1903), the southern crested newt, *Triturus karelinii* (Strauch, 1870), and the Macedonian crested newt, *Triturus macedonicus* (Karaman, 1922). The Macedonian crested newt was just recently raised to species level given its allopatric distribution to its sister species, the Italian crested newt, and the level of genetic differentiation (see CHAPTER 2). The two known subspecies of the southern crested newt, *T. k. karelinii* and *T. k. arntzenii* may someday be raised to species level as well, given their substantial genetic differentiation (see CHAPTER 4).

Morphologically the species are very similar. They are all large newts with heavily build, and warty skin. Their dorsal side is usually dark brown to black, while their sides are sometimes punctuated with small white spots. Males present a serrated

Figure 2 - *Triturus carnifex* from an introduced population in the region Veluwe, The Netherlands.

dorsal crest and a white band in their tale during the breeding season. Their belly is yellow to orange, with variable number of white and black spots. The Danube crested newt is the most slender and elongated, which might be an adaptation to its more aquatic habitat (ARNTZEN and WALLIS, 1999). The species can be distinguished by the number of rib-bearing vertebra: *T. karelinii* has 14, *T. carnifex* and *T. macedonicus* have 15; *T. cristatus* has 16; and *T. dobrogicus* has 17 or 18 (ARNTZEN and WALLIS, 1999).

Distribution

Triturus carnifex and T. macedonicus

The Italian crested newt, *Triturus carnifex* (Laurenti, 1768), is present south of the Alps and occupies Italy, Slovenia, Croatia, and Austria. It has been introduced in several places, like the Azores in Portugal (where it is the only newt present; MALKMUS, 1995), Geneva in the French-Swiss border (ARNTZEN, 2001), Birmingham and Surry in England (BEEBEE and GRIFFITHS, 2000), Veluwe in The Netherlands (BOGAERTS, 2002), and Bavaria, Germany (FRANZEN *et al.*, 2002). The Macedonian crested newt, *Triturus macedonicus* (Karaman, 1922) was until recently considered a subspecies of *T. carnifex*, but as a result of this thesis, this taxon has been raised to species level (see CHAPTER 2). It occurs in Macedonia, Greece, Serbia, Montenegro, Albania and southern Bosnia and Herzegovina.

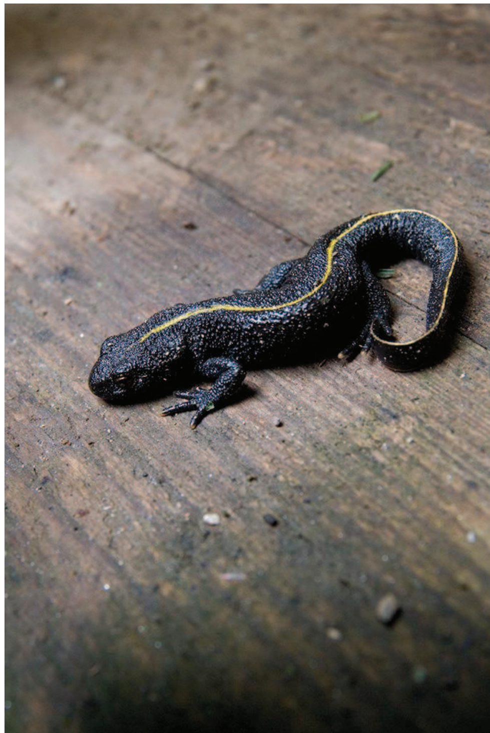


Figure 3 *Triturus carnifex* female.

Figure 4 Larva of *Triturus cristatus*. Notice the dark spots on the dorsal tail fin and the thin long fingers, characteristic of larval stages of *Triturus*.

Triturus cristatus

The great or northern crested newt is the most widespread species of crested newt. It occurs all the way from Great Britain to northern France, central and eastern Europe north of the Alps and the Carpathians, southern Scandinavia, and southwestern Siberia.

Triturus dobrogicus

The Danube crested newt is present along the Danube river basin, encompassing Slovakia, Hungary, Romania, Bulgaria and Moldova. Its distribution is divided into two regions: the Dobrogean and the Pannonian. The two populations are separated by the Carpathian Mountains, but seem to be connected through the Iron Gate in the Danube (ARNTZEN, 2003).

Triturus karelinii

T. karelinii occurs along the southern shore of the Iranian Caspian Sea, Georgia, Azerbeijan, the Russian Black Sea coast, Crimea (Ukraine), Turkey, Bulgaria, northern Greece, and Serbia. The southern crested newt has two recognized subspecies: *T. k. karelinii* from Iran, Azerbeijan, Georgia, Russia and Ukraine, and *T. k. arntzenii*, from the Balkans. The populations from Turkey have an unknown status (see CHAPTERS 2 and 4). There is a known enclave of *T. k. karelinii* in eastern Serbia that is completely surrounded by populations of three other crested newts (ARNTZEN and WALLIS, 1999).

Figure 5 - Male *Triturus karelinii* from Bozdag, Turkey. The dorsal crest, typical during the breeding season, is folded to the right, and so is not clearly visible.

Conservation and threats

T. dobrogicus is considered a near threatened species by the International Union for Conservation of Nature (IUCN), due to the rate of population decline caused by habitat loss throughout its distribution range. It is also threatened by hybridization with its neighbouring crested newt species, given its central position and limited distributional range (ARNTZEN *et al.*, 2006b).

T. macedonicus is not listed in the IUCN red list, as it was considered a subspecies of *T. carnifex* in the latest assessment. However, the entry for *T. carnifex* already mentions major decline of the Balkan populations due to decrease in spring precipitation, possibly a consequence of global climate change (ARNTZEN *et al.*, 2006a).

Although the trend in the other species is for population decrease due to deforestation and pollution of wetlands, the IUCN red list lists them as least concern, as they consider that given their wide distribution, the speed of the decline is not fast enough to include it in a more threatened category (ARNTZEN *et al.*, 2006c).

Figure 6 Typical breeding site for newts in Turkey. Nets seen in the left and right are used to capture larvae and breeding adults.

Marbled newts

This section was adapted from Espregueira Themudo & Arntzen (2009)

The marbled newts are two species of *Triturus*: the northern marbled newt, *Triturus marmoratus* (Latreille, 1800); and the pygmy marbled newt, *Triturus pygmaeus* (Wolterstorff, 1905). The pygmy marbled newt was just recently recognized as a full species, while it used to be considered a subspecies of *Triturus marmoratus* (García-París *et al.*, 2001). Arguments in favour of this position include the level of genetic differentiation between the two, diagnostic morphological characters and the lack of hybrids in Spain. However, the situation in Portugal is spatially more complex, as the two species are in contact, and some hybrids have been detected (see Chapters 6 and 9). *Triturus marmoratus* is clearly larger, with a strong build, rough skin, a more or less uniform dark ventral colouration with white stipples and a hard-green dorsal and lateral coloration in a course network. Characteristic features of *T. pygmaeus* are a small body size, elegant built, smooth skin, greyish and spotted ventral colouration and an olive-green dorsal and lateral coloration in a fine network.

Distribution

Triturus marmoratus

The range of *T. marmoratus* covers major parts of France, Spain and Portugal. In France, *T. marmoratus* is found in the southwestern part of the country, northwards to Normandy and Paris. In central France the range of *T. marmoratus* overlaps with that of *T. cristatus*, with interspecific hybridisation taking place (ARNTZEN and WALLIS, 1991). In Spain, *T. marmoratus* is found all over the northern part of the country, in the east southwards to the valley of the Ebro, in the centre as far south as the Sierra de Guadarrama and in the west as far south as the Sierra de Gata (ALBERT and GARCIA-PARIS, 2004). In Portugal, *T. marmoratus* is found all across the northern part of the country, excluding the coastal zone south of Aveiro. There is also evidence of an isolate around Caldas da Rainha surrounded by populations of *T. pygmaeus* (see CHAPTER 9). The southern border runs from the Serra de Gata at the Spanish border, southwards to reach but not cross the river Tejo in central Portugal, approximately following the line Castelo Branco - Abrantes - Leiria.

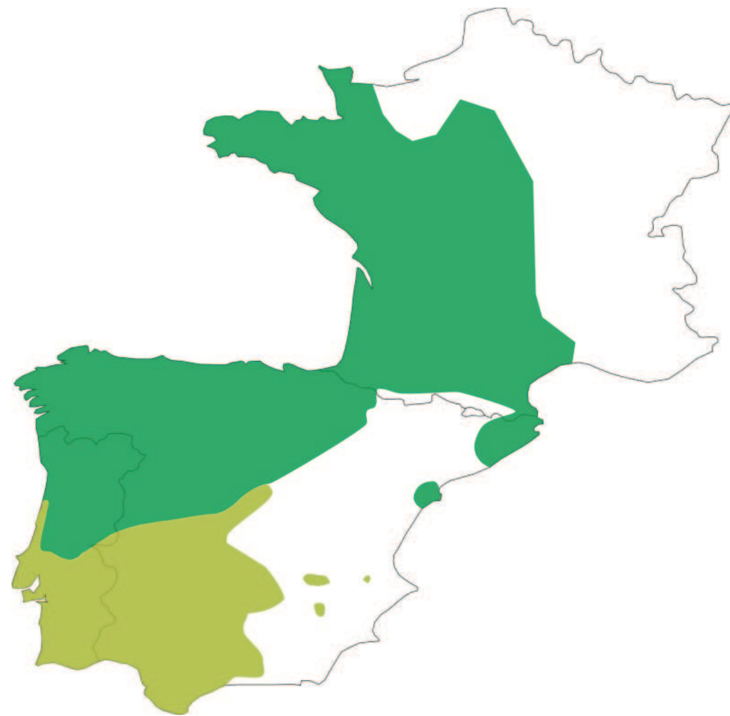


Figure 7 Distribution of *Triturus marmoratus* and *T. pygmaeus* in Western Europe. *Triturus marmoratus* is depicted in the darker shade of green, and *T. pygmaeus* in olive green.

Figure 8 *Triturus marmoratus* from Dordogne, France.

Triturus pygmaeus

The range of *T. pygmaeus* covers the southern part of the Iberian Peninsula, with the exception of the eastern and southeastern parts of Spain. The northernmost localities of *T. pygmaeus* are situated in the Portuguese coastal zone, as far north as Aveiro. In central Portugal the range of *T. pygmaeus* is contiguous with that of *T. marmoratus*. In Spain the range follows the southern slopes of the Sistema Central, including the Sierra de Gata, Sierra de Gredos and the Sierra de Guadarrama (GARCÍA-PARÍS, 2004). The shortest documented distance between populations of both species in the Madrid area is c. 6 km (GARCÍA-PARÍS *et al.*, 2001).

Conservation and threats

The significant loss of habitat in the south of the Iberian Peninsula, specially by the decrease in the number of temporary ponds caused by desertification led IUCN to list the pygmy marbled newt as 'Near Threatened' (STUART *et al.*, 2008). In Gerês National Park (northwest of Portugal), several amphibian species, including *T. marmoratus* are infected by an iridovirus that causes high mortality (ALVES DE MATOS *et al.*, 2002). Other more general causes of decline include the draining of temporary ponds as a consequence of the intensification of agriculture; increase urbanization; and predation by invasive species, such as the

Figure 9 Female *Triturus marmoratus* from Porto, Portugal. Notice the orange dorsal strip, warty skin, and tissue regeneration in the tailtip, a few weeks after the tailtip was cut for sampling.

Louisiana crayfish (*Procambarus clarkii*) and the sunfish (*Lepomis gibbosus*). These combined factors are causing the disappearance of several populations in southern Spain, contributing to the fragmentation of its distribution (García-París et al., 2001). The situation in *T. marmoratus* is more stable than in *T. pygmaeus*, despite its regression in the western coast of Portugal, as it was replaced by *T. m. pygmaeus* (CHAPTER 9).

Scope of the thesis

This thesis is roughly divided in two sections. The first concerns the phylogenetic history of the genus *Triturus*, how species are related to each other, and the effect of some evolutionary processes on the inferred phylogeny. The last chapter of this section, on the phylogeography of the marbled newt group, links this to the second section. The second half of the thesis concerns the relationship between the distribution ranges of species and the environmental conditions. The main focus was on the effect that the local environment has on the relative competitiveness of sibling species, and ultimately

on their distribution limits.

This thesis consists of eight scientific chapters, apart from this introduction, most of which have been published in (or, at least, submitted to) international peer reviewed scientific journals.

Summary

The interspecific relationships in the genus *Triturus* are incompletely known. In **Chapter 2**, we attempt to resolve them by using allozyme and mtDNA data. Despite the large number of markers used, relationships continue to elude us. The results suggest that speciation in the group occurred during a short time period (the end of the Middle Miocene). Paleogeographic reconstructions of the presumed centre of origin (the Balkans) support this hypothesis. We proposed here that *T. macedonicus* should be raised to full species given its allopatric distribution and high genetic divergence with *T. carnifex*.

The best way to test the scenario presented in Chapter 2 is to look at multiple independent markers that, unfortunately, were not readily available. **Chapter 3** describes the process through which dozens of markers were designed and tested for the genus *Triturus*. Out of more than fifty markers tested, five provided promising results with enough variability to study the phylogeny and phylogeography of the genus. This opened the door not only for Chapter 4, but also for 5 and 6.

Taking Chapter 2 as the starting point and with the tools developed in Chapter 3, **Chapter 4** attempts to decipher the history of the genus *Triturus*. The study includes samples from 15 individuals of the seven species of the group. Locations were selected to cover most of the variability in the group, with the exception of areas close to other species. Hybridization is known to occur in these areas, and could bias the inferences made. Results show that all the genes, except the two mtDNA ones, have incongruent phylogenetic signals. We used phylogenetic networks to visualize the alternative phylogenetic signals and have built a phylogenetic tree based on a Bayesian hierarchical method that obtains the species tree based on individual gene trees. This approach successfully resolved the branching order of the newts, although time intervals

are very narrow, confirming the near simultaneous speciation scenario of CHAPTER 2. We also found a high genetic differentiation between the two forms of the southern crested newt (*T. karelinii*) and proposed that they should be raised to full species (*T. karelinii* and *T. arntzeni*).

As described in CHAPTER 4, hybridization can have a confounding effect on phylogenetic inferences. Incomplete lineage sorting can also produce similar patterns as gene flow, further complicating matters. **CHAPTER 5** takes a new approach in distinguishing between gene flow and incomplete lineage sorting, only possible in species with very well defined species borders and limited dispersal capability as the newts, which limits gene flow to a narrow geographical region.

As can be seen in the previous chapters, the crested newts show complex (although interesting) patterns resulting from compound interactions, especially in the Balkans. The marbled newts, on the other hand, are only two species, and therefore we can cover their evolutionary history in more detail, as it is simpler. **CHAPTER 6** studies the phylogeography of the two species of marbled newts. Being just a pair and not a group of species, relationships are not problematic. Their sibling relationship is well established (see CHAPTER 2). This chapter also explores how differences in ecological conditions (see CHAPTER 8) along a hybrid zone affect its structure and the amount of gene flow between species.

The morphological distinction of adults of the two marbled species is relatively straightforward, even though variation is present. Eggs and larvae, on the contrary, are easy to spot and collect, but impossible to distinguish. **CHAPTER 7** describes a cheap and fast molecular technique that allowed the identification of the large number of samples used in CHAPTERS 8 and 9.

Based on a morphological and genetic identification of individuals of the two species of marbled newts and published distributional data, **CHAPTER 8** identified ecological factors associated with the range border of the two species of marbled newt. Ecological models defining the distribution of the two species differed in four main areas: the northern most region of contact close to Aveiro, the rest of the coastal area (see also CHAPTER 9), the region coinciding with the Tejo river, and the remainder contact zone going until Madrid.

In the coastal area of Portugal, close to Caldas da Rainha, the northern marbled newt was found where only pygmy marbled newts were thought to occur. This prompted a detailed study on the distribution of the two species in this area described in **Chapter 9**. This study revealed a small pocket of populations of the northern species surrounded by populations of its sister southern species. Given the distance of this pocket to the main distribution, we believe that the enclave was created by *T. pygmaeus* moving north, superseding *T. marmoratus*, rather than the latter species expanding southwards.

References

Albert, E. M. and M. Garcia-Paris (2004) *Triturus marmoratus* (Latreille, 1800) Tritón jaspeado. In: *Atlas y libro rojo de los anfibios y reptiles de España* (eds. Pleguezuelos, JM, R Márquez, M Lizana). Dirección General de Conservación de la Naturaleza - Associación Herpetologica Española, Madrid.

Alves de Matos, A. P., A. Loureiro, M. A. Carretero and C. Soares (2002) Iridovirus-like particles in high mortality episode of *Triturus marmoratus* (Amphibia) from Portugal.

Arntzen, J. W. (2001) Genetic variation in the Italian crested newt, *Triturus carnifex*, and the origin of a non-native population north of the Alps. *Biodiversity and Conservation* **10**, 971-987.

Arntzen, J. W. (2003) *Triturus cristatus* Superspecies–Kammolch-Artenkreis. In: *Handbuch der Reptilien und Amphibien Europas. Schwanzlurche IIA* (eds. Grossenbacher, K, B Thiesmeier), pp. 421-514. Aula-Verlag, Wiebelsheim.

Arntzen, J. W. (2006) From descriptive to predictive distribution models: a working example with Iberian amphibians and reptiles. *Frontiers in Zoology* **3**, 8.

Arntzen, J. W., M. Denoël, R. Jehle, F. Andreone, B. Anthony, B. Schmidt, W. Babik, R. Schabetsberger, M. Vogrin, M. Puky, P. Lymberakis, J. C. Isailovic, R. Ajtic and C. Corti (2006a) *Triturus carnifex*. In: *2008 IUCN Red List of Threatened Species.*

Arntzen, J. W., S. Kuzmin, R. Jehle, M. Denoël, B. Anthony, C. Miaud, W. Babik, M. Vogrin, D. Tarkhnishvili, V. Ishchenko, N. Ananjeva, N. Orlov, B. Tuniyev, D. Cogalniceanu and T. Kovács (2006b) *Triturus dobrogicus*. In: *2008 IUCN Red List of Threatened Species.*

ARNTZEN, J. W., T. PAPENFUSS, S. KUZMIN, D. TARKHNISHVILI, V. ISHCHENKO, B. TUNIYEV, M. SPARREBOOM, N. RASTEGAR-POUYANI, I. UGURTAS, S. ANDERSON, W. BABIK, C. MIAUD and J. C. I. (2006c) *Triturus karelinii*. In: *2008 IUCN Red List of Threatened Species*.

ARNTZEN, J. W. and G. P. WALLIS (1991) Restricted gene flow in a moving hybrid zone of the newts *Triturus cristatus* and *T. marmoratus* in western France. *Evolution* **45**, 805-826.

ARNTZEN, J. W. and G. P. WALLIS (1999) Geographic variation and taxonomy of crested newts (*Triturus cristatus* supersecies): morphological and mitochondrial data. *Contributions to Zoology* **68**, 181-203.

BEEBEE, T. J. C. and R. GRIFFITHS (2000) *Amphibians and reptiles: a natural history of the British herpetofauna* HarperCollins.

BOGAERTS, S. (2002) Italian crested newts, *Triturus carnifex*, on the Veluwe, Netherlands. *Zeitschrift fur Feldherpetologie* **9**, 217-226.

DE QUEIROZ, K. (2007) Species concepts and species delimitation. *Systematic Biology* **56**, 879-886.

FRANZEN, M., H. J. GRUBER and U. HECKES (2002) Eine allochthone *Triturus carnifex*-Population in Sudbayern (Deutschland). *Salamandra* **38**, 149-154.

GARCÍA-PARÍS, M. (2004) *Triturus pygmaeus* (Wolterstorff, 1905). Tritón pigmeo. In: *Atlas y libro rojo de los anfibios y reptiles de España* (eds. PLEGUEZUELOS, JM, R MÁRQUEZ, M LIZANA), pp. 70-72. Dirección General de Conservación de la Naturaleza - Associación Herpetologica Española (2nd edition), Madrid.

GARCÍA-PARÍS, M., B. ARANO and P. HERRERO (2001) Molecular Characterization of the contact zone between *Triturus pygmaeus* and *T. marmoratus* (Caudata: Salamandridae) in Central Spain and their taxonomic assessment. *Revista Española de Herpetología* **15**, 115-126.

GARCÍA-PARÍS, M., A. MONTORI and P. HERRERO (2004) Amphibia: Lissamphibia. In: *Fauna Ibérica*. Consejo Superior de Investigaciones Científicas.

JEHLE, R. and J. W. ARNTZEN (2000) Post-breeding migrations of newts (*Triturus cristatus* and *T. marmoratus*) with contrasting ecological requirements. *Journal of Zoology* **251**, 297-306.

MALKMUS, R. (1995) *Die Amphibien und Reptilien Portugals, Madeiras und der Azoren* Westarp Wissensehaften.

STEINFARTZ, S., S. VICARIO, J. W. ARNTZEN and A. CACCONE (2006) A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus* newts. *J Exp Zoolog B Mol Dev Evol* **308B**, 139-162.

STUART, S. N., M. HOFFMANN, J. S. CHANSON, N. A. COX, R. J. BERRIDGE, P. RAMANI and B. E. YOUNG (2008) Threatened Amphibians of the World, Lynx Edicions, Barcelona, Spain; IUCN, Gland, Switzerland and Conservation International, Arlington, Virginia, USA.

CHAPTER 2

THE PHYLOGENY OF CRESTED NEWTS (*TRITURUS CRISTATUS* SUPERSPECIES): NUCLEAR AND MITOCHONDRIAL GENETIC CHARACTERS SUGGEST A HARD POLYTOMY, IN LINE WITH THE PALEOGEOGRAPHY OF THE CENTRE OF ORIGIN

Arntzen, J. W.[1], G. Espregueira Themudo[1,2] and B. Wielstra[1,3]

[1] *National Museum of Natural History, P. O. Box 9517, 2300 RA Leiden, The Netherlands.*
[2] *CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661*
Vairão, Portugal.
[3] *Institute of Biology Leiden, Leiden University, P.O. Box 9516, 2300 RA Leiden, The Netherlands*

Contents

Abstract 18
Introduction 19
Material and methods 21
  Sampling strategy 21
  Allozyme data 22
  Mitochondrial DNA-sequence data 22
  Phylogenetic analysis 23
Results 25
  Allozymes 25
  Mitochondrial DNA-sequences 30
Discussion 33
  Genetic coherence of species and subspecies 33
  Phylogeny 34
  Historical biogeography 35
  Taxonomic considerations 38
Acknowledgements 39
References 39
Appendix 1 43
Appendix 2 43

Published in *Contributions to Zoology* **76** (4) 261-278 (2007).

Abstract
Newts of the genus *Triturus* (Amphibia, Caudata, Salamandridae) are distributed across Europe and adjacent Asia. In spite of its prominence as a model system for evolutionary research, the phylogeny of *Triturus* has remained incompletely solved. Our aim was to rectify this situation, to which we employed nuclear encoded proteins (40 loci) and mitochondrial DNA-sequence data (mtDNA, 642 bp of the ND4 gene). We sampled up to four populations per species covering large parts of their ranges. Allozyme and mtDNA data were analyzed separately with parsimony, distance, likelihood and Bayesian methods of phylogenetic inference. Existing knowledge on taxonomic relationships was confirmed, including the monophyly of the genus and the groups of crested newts (four species) and marbled newts (two species). The genetic coherence of species and subspecies was also confirmed, but not always with high statistical support (depending on taxon, characters under consideration, and method of inference). In spite of our efforts we did not obtain sufficient phylogenetic signal to prefer one out of twelve potential topologies representing crested newts relationships. We hypothesize that the lack of phylogenetic resolution reflects a hard polytomy and represents the (near)-simulateneous origin of crested newt species. Using a calibration point of 24 Ma (million years before present) for the most recent common ancestor of *Triturus*-species, the crested newt radiation event is dated at 7-6 Ma (scenario 1) or at 11-10 Ma (scenario 2), depending on the application of an allozyme versus mtDNA molecular clock. The first biogeographical scenario involves the spread of crested newts from the central Balkans into four compass directions. This scenario cannot be brought into line with potential vicariance events for south-eastern Europe. The second scenario involves the more or less simultaneous origin of four species of crested newts through large-scale vicariance events and is supported by the paleogeographical reconstruction for the region at the end of the Middle Miocene. The subspecies *Triturus carnifex macedonicus* carries in one large area the mtDNA that is typical for the neighbouring species *T. karelinii*, which is attributed to introgression and a recent range shift. It is nevertheless a long distinct evolutionary lineage and we propose to elevate its taxonomic status to that of a species, i.e., from *Triturus c. macedonicus* (Karaman, 1922) to *Triturus macedonicus* (Karaman, 1922).

Key words: allozymes, historical biogeography, mitochondrial DNA-sequences, *Triturus macedonicus*, *Triturus marmoratus*, vicariance.

Introduction

The newt genus *Triturus* Rafinesque, 1815 (Amphibia, Caudata, Salamandridae) is a model group for various lines of evolutionary research, on e.g. locomotion (GVOZDÍK, L., R. VAN DAMME, 2006), cranial ontogeny (IVANOVIĆ, A. *et al.*, 2007), reproductive behaviour (STEINFARTZ, S. *et al.*, 2007) and sexual size dimorphism (A. Ivanović, K. Sotiropoulos, M. Furtula, G. Džukić and M.L. Kalezić, submitted). Considering the efficient use of the 'comparative method' that aims to account for non-independence among characters due to shared evolutionary histories (HARVEY, P.H., M.D. PAGEL, 1991), it would seem important that the phylogenetic relationships within the group are well resolved. Despite considerable efforts, starting with the osteological study of Bolkay (1928), an unambiguous phylogenetic resolution for *Triturus* has not been obtained (ARNTZEN, J.W., 2003).

The genus *Triturus* encompasses six species of so-called 'large bodied' newts that are organized in two groups: the marbled newts with two species and the crested newts with four species. Marbled newts are characterized by a dorsal green colouration and a fairly uniform light grey or black underside. *Triturus marmoratus* (Latreille, 1800) occurs in central and southern France and the northern part of the Iberian Peninsula and *T. pygmaeus* (Wolterstorff, 1905) is confined to the south-western and southern parts of the Iberian Peninsula. Crested newts are characterized by an orange and black spotted ventral colouration and a dark backside and occupy most other parts of Europe and adjacent Asia. Four species are recognized: *T. cristatus* (Laurenti, 1768), that is distributed over central, western and northern Europe and eastwards deep into Russia, *T. carnifex* (Laurenti, 1768), that occurs on the Appenine Peninsula and part of the Balkans, *T. dobrogicus* (Kiritzescu, 1903), that is confined to the Pannonian and Dobrogean lowlands, and *T. karelinii* (Strauch, 1870), with a range from the southern Balkans to the Caspian Sea. *Triturus carnifex* has two allopatric subspecies, the nominotypical subspecies in the western part and *T. c. macedonicus* (Karaman, 1922) in the eastern part of the species range, separated by the Adriatic Sea and the area approximately coinciding with Bosnia-Herzegovina.

Given their contiguous parapatric ranges with limited hybridization (WALLIS, G.P., J.W. ARNTZEN, 1989), the four species of crested newt are conveniently referred to as a superspecies (or 'Artenkreis') *sensu* Rensch (1929).

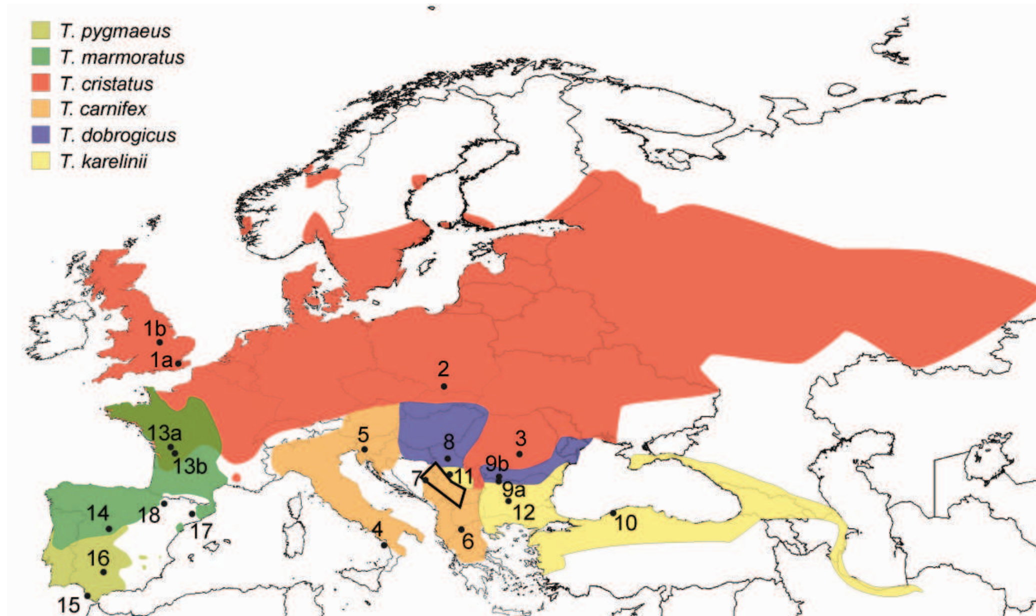


Figure 1 - Approximate geographical distribution of *Triturus* species, represented by the following colours: orange - *T. carnifex*, red - *T. cristatus*, blue - *T. dobrogicus*, yellow - *T. karelinii*, dark green - *T. marmoratus*, light green - *T. pygmaeus*. Numbers indicate the populations sampled in this study as listed in Appendix I. The concave polygon in the Balkans represents the area where karelinii-type mtDNA was observed in *T. carnifex macedonicus* (ARNTZEN, J.W., G.P. WALLIS, 1999; and present study). Note the zone of sympatry between *T. cristatus* and *T. marmoratus* in western France.

Molecular phylogenetic studies corroborate the monophyly of *Triturus* and the sistergroup relationship of crested and marbled newts (e.g. ZAJC, I., J.W. ARTZEN, 1999; STEINFARTZ, S. *et al.*, 2007). Approximate geographical distributions of *Triturus* species are shown in Figure 1.

The monophyly of the genus *Triturus* is strongly supported by a number of synapomorphic character states (ARNTZEN, J.W., 2003), including a remarkable genetic condition that kills off all embryos that are homomorphic for either the long or the short copy of chromosome- 1, the so-called 'chromosome-1 syndrome' (RIDLEY, M., 1993). The small- and medium-sized newt species (*alpestris*, *vittatus*, *vulgaris* and allies), traditionally included in *Triturus* have recently been placed in genera of their own (GARCÍA-PARÍS, M. *et al.*, 2004; STEINFARTZ, S. *et al.*, 2007).

Adult crested newts can be organized in a morphological series from a stocky built with well-developed appendages in *T. karelinii*, to a slender built with short appendages in *T. dobrogicus*, with *T. carnifex* and *T. cristatus* taking intermediate positions. This morphocline is expressed in the Wolterstorff Index, which is defined as forelimb length divided by inter-limb distance (WOLTERSTORFF,

W., 1923; FUHN, I.E., 1960; ARNTZEN, J.W., W.G. P., 1994). A correlated diagnostic character is the number of rib-bearing vertebrae (NRBV). The NRBV is typically 13 in *T. karelinii*, 14 in *T. carnifex*, 15 in *T. cristatus* and 16 or 17 in *T. dobrogicus* (ARNTZEN, J.W., G.P. WALLIS, 1999; ARNTZEN, J.W., 2003). In *T. marmoratus* that is stockier and possesses more developed appendages than either species of crested newt, the NRBV is typically 12. Taking the marbled newts as an outgroup, the parsimony-based phylogenetic hypothesis for crested newts is (root / *T. karelinii* (*T. carnifex* (*T. cristatus*, *T. dobrogicus*))) with, seen from the root, the first species to split off *T. karelinii*, than *T. carnifex* and followed by *T. cristatus* and *T. dobrogicus*. We will refer to this topology as the NRBV-constraint.

Newts are aquatic breeders. A relationship has been documented between body shape and the length of time the species annually spends in the water: c. two months for *T. marmoratus*, three months for *T. karelinii*, four months for *T. carnifex*, five months for *T. cristatus* and six months for *T. dobrogicus*. So, the longer the body, the longer the aquatic period (ARNTZEN, J.W., 2003). This suggests that body shape is under selection and raises the possibility that morphology reflects environment instead of phylogeny. Indeed, the phylogenetic hypothesis from the RFLP and sequence analysis of mitochondrial DNA (mtDNA) deviates from the phylogeny suggested by the morphocline (WALLIS, G.P., J.W. ARNTZEN, 1989; STEINFARTZ, S. *et al.*, 2007). Twelve tree topologies (enumerated in Table 3) are possible under the assumptions that i) the marbled newts form the sistergroup to the crested newts, i.e., the trees are rooted, and ii) populations and haplotypes fall within the species for which they are recognized. The aforementioned 'NRBV-constraint' equals to tree number 10. In the present study, we set out to elucidate the phylogenetic relationships of the genus *Triturus* and in particular the *Triturus cristatus* superspecies. For this purpose, we employ the independent datasets of nuclear encoded protein data and mtDNA-sequence data.

Material and methods

Sampling strategy

Adult and juvenile newts were caught by dip netting across Europe and Asiatic Turkey. Individuals were identified to species based on morphology (GARCÍA-PARÍS, M., HERRERO, P., MARTÍN, C., DORDA, J., ESTEBAN, M. & ARANO, B., 1993; ARNTZEN, J.W., G.P. WALLIS, 1999). For phylogeny reconstruction, two to four

populations per species were selected from a larger sample set, such that 1) identifications are in line with documented geographical distributions, 2) spatial coverage is large, and 3) samples from putative admixture zones are preferably excluded (Fig. 1). Two British (*T. cristatus*), two French (*T. marmoratus*) and two neighbouring Dobrogean (*T. dobrogicus*) populations were pooled. Amplification of DNA from the *T. pygmaeus* population Venta del Charco was not successful. Sample size ranges from 4 to 20 (average *N* = 10.0) individuals per population in the protein data set and from 3 to 10 in the mtDNA dataset (average *N* = 4.6). Following the most recent phylogenetic hypothesis on the Salamandridae (STEINFARTZ, S. *et al.*, 2007), four specimens of *Calotriton asper* (Dugès, 1852) from the eastern (N = 2) and western (N = 2) Pyrenean mountains wereselected as outgroup to the genus *Triturus*. Details on localities, sample size and voucher material are presented in Appendix I.

Allozyme data

Twenty-seven protein systems representing 40 presumptive gene loci were studied by means of starchand polyacrylamide-gel electrophoresis (Table 1). Laboratory protocols and nomenclature of gene products are as in Arntzen (2001). The population genetic parameters evaluated with HP-rare (KALINOWSKI, S.T., 2005) were heterozygosity as expected under conditions of Hardy-Weinberg equilibrium (*H*e), and allelic richness (*A*) under reference to a sample size of 20 genes. Genetic distances calculated with BIOSYS (SWOFFORD, D.L., R.B. SELANDER, 1981) were the Prevosti distance and the Cavalli-Sforza and Edwards chord distance for each locus and Nei's unbiased genetic distance (*D*N) over all loci.

Mitochondrial DNA-sequence data

Total DNA was extracted from small amounts of newt tissue using the DNeasy Tissue Kit (Qiagen). A segment of subunit 4 of the NADH dehydrogenase mitochondrial gene (ND4) was amplified by polymerase chain reaction (PCR) using the primer ND4 (light chain; 5'-CACCTATGACTACCAAAAGCTCATGTAGAAGC- 3'; ARÉVALO, E. *et al.*, 1994)) and the newly designed primers KARF4 (light chain; 5'-AGCGCCTGTCGCCGGGTCAATA-3'), ND4R2 (heavy chain; 5'-CCCTGAAATAAGAGAGGGTTTAA-3'), KARR1 (heavy chain; 5'-AACTCTTCTTGGTGCGTAG-3'), DOBR1 (heavy chain; 5'-GTTTCATAACTCTTCTTGGTGT-3') and DOBR2 (heavy chain; 5'-GTGTTTCATAACTCTTCTTGGT-3').

The polymerase chain reaction (PCR) was conducted in a total volume of 25 µL, containing 1.0 µL of DNA extraction, 2.5 µL 10x CoralLoad PCR Buffer (containing 15mM MgCl2; Qiagen), 0.8 µL of each primer (10 µM), 1µL dNTPs (10mM) and 0.2 µL of *Taq* DNA polymerase (5 U/µL; Qiagen), replenished with Milli-Q water. Reaction conditions were: initial denaturation for 3 min at 94°C; 35 cycles of 30 sec denaturation at 94°C, 30 sec annealing at 58°C, 1 min extension at 72°C; and 4 min final extension at 72°C. The PCR product was run on 1% TBE (tris-borate- EDTA) agarose gel by electrophoresis, stained with ethidium bromide, and visualised by exposure to UV light in order to check for quality. Negative controls were included to assure that PCR product was not from contaminated sources.

The PCR product was purified using the Wizard SV Gel and PCR Clean-up System (Promega). Cycle sequencing of both the heavy and the light strand was done commercially through Macrogen Inc. The forward and reverse sequences were checked by eye and a consensus sequence was made with Sequencher 4.5 (Gene Codes Corporation). Sequences were aligned manually in MacClade 4.08 (MADDISON, D.R., W.P. MADDISON, 2005) and identical sequences were merged into haplotypes. Indications that haplotypes would represent nuclear insertions (i.e., pseudogenes) were not found from either the translated amino-acid sequence (that had no stop codons or inferred insertions/deletions) or nucleotide composition (which was anti G-biased).

The mtDNA dataset contained 642 bp, homologous to position 10854-11495 of the mitochondrial genome of *Lyciasalamandra atifi* (Basoglu, 1967) (ZARDOYA, R. *et al.*, 2003). To check for the level of substitution saturation, we plotted uncorrected sequence divergence against the Kimura 2-parameter distance (d_{K80}). The relationship was near-linear, suggesting that loss of phylogenetic signal due to multiple hits was not an issue (data not shown). The level of saturation was quantified as 'little' with DAMBE software (XIA, X., Z. XIE, 2001).

Phylogenetic analysis

The last four decades have witnessed a plethora of approaches in which allozyme data are employed for phylogeny reconstruction. In order to avoid analytical bias, we followed the recommendations by Wiens (2000) and applied the following methods: maximum parsimony on single-locus Prevosti distances that were coded through step-matrices (MP), phenetic clustering by neighbour-joining on the Cavalli-Sforza and

Edwards chord distance (NJ) and continuous maximum likelihood (ML). Tree length comparisons were made under the MP-approach with the Kishino-Hasegawa and Templeton tests in PAUP 4.0* (SWOFFORD, D.L., 2003). For reasons of consistency, similar methods were used on the mtDNA-sequences, be it that not populations, but individual haplotypes were involved as Operational Taxonomic Units (OTU). The measure used in the NJ-analysis was the Kimura two-parameter distance. For the likelihood based analyses of DNA-sequence data, we used Modeltest 3.7 (POSADA, D., K.A. CRANDALL, 1998) to determine the best fitting model of sequence evolution as determined by the AIC-criterion. This was the TIM model with a gamma shape parameter of 0.32, a base composition of A = 0.30, C = 0.29 and G = 0.14, and the relative rate parameters 1.00, 11.56, 0.36, 0.36 and 6.88 (as in the output format of the program). The software programmes used for phylogenetic analysis were PHYLIP 3.573c (FELSENSTEIN, J., 1993) and PAUP 4.0* (SWOFFORD, D.L., 2003). To evaluate the strength of support by the underlying data to branches in the phylogenetic tree, we ran 2000 bootstrap replications for each of the three approaches. Bootstrap replication scores (brs) of > 0.80 were interpreted as strong support, 0.70 ≤ brs ≤ 0.80 were interpreted as moderate support of the data to the phylogenetic tree. Additionally, the mtDNA data were analyzed under Bayesian inference with MrBayes 3.1 (RONQUIST, F., J.P. HUELSENBECK, 2003). This involved the running of four Metropolis Coupled Monte Carlo Markov Chains (MCMC), one cold and three incrementally heated, starting from a random topology. Two separate runs of two million generations were conducted simultaneously and for each run, the cold chain was sampled every 1000 generations under the best-fit model of molecular evolution (HKY+G) selected with MrModel- Test (NYLANDER, J.A., 2004). The software Tracer 1.3 (RAMBAUT, A., A.J. DRUMMOND, 2007) was used to check for stabilization of overall likelihood within and convergence between runs. The first 10% of sampled trees was discarded as burn-in and the inference was drawn from the remaining pooled sample. Bayesian posterior probabilities (pp) of > 0.95 were interpreted as strong support and 0.90 ≤ pp ≤ 95 were interpreted as moderate support of the data to the tree.

We used linear interpolation from a set calibration point to estimate the divergence times for major clades. With the same aim, 2000 Bayesian trees from two runs and a 50% 'burn-in' were analyzed with the software r8s (SANDERSON, M.J., 2004), to reconstruct divergence times under the assumption of a molecular clock using the LF maximum likelihood method and the TN algorithm (see r8s manual).

This yielded the most likely estimates (mean and mode) and the 95% confidence interval of the mean.

Results

Allozymes

The allelic richness averaged over 40 enzyme loci ranged from low values in *T. marmoratus* and *T. pygmaeus* (*A* = 1.13 - 1.22), medium values in *T. cristatus*, *T. dobrogicus, T. c. macedonicus* and *T. karelinii* (*A* = 1.34 - 1.44), to a relatively high value in *T. c. carnifex* (*A* = 1.64). *Triturus c. macedonicus* and *T. cristatus* stand out from the other (sub)species by the marked difference between populations, with low values observed in Višegrad and the UK and Poland, respectively (Table 1). Within *T. cristatus* the spatial distribution of allozyme genetic variation is in line with a postglacial dispersal scenario from the southern Carpathians all over northern Europe, supporting similar data from RFLP- and MHC-based molecular analyses (WALLIS, G.P., J.W. ARNTZEN, 1989; W. Babik, personal communication). Average expected heterozygosity and AR are correlated across populations (rSpearman = 0.88, Table 1). Maximum parsimony analysis of the allozyme data yielded moderate or strong support for the grouping of populations within (sub)species for *T. c. macedonicus*, *T. dobrogicus*, *T. karelinii* and *T. marmoratus* (Figure 2, Table 2). Neighbour-joining and maximum likelihood analysis showed moderate or strong support for all taxa within *Triturus* with the exception of *T. c. carnifex* and *T. carnifex*. The grouping together of all crested newt taxa was supported throughout, while moderate support (brs = 0.71) for a sister group position of *T. cristatus* to the other crested newt species was obtained with maximum parsimony analysis only. The MP-tree obtained under the NRBV-constraint was 4.5 % longer than the most parsimonious solution. The difference in tree length is marginally significant under the Kishino- Hasegawa test ($P \approx 0.05$) and significant under the Wilcoxon signed-ranks tests ($P < 0.05$). Under the prerequisite that populations fall within the species for which they are recognized, tree number 2: (root /*T . cristatus* (*T. dobrogicus* (*T. carnifex*, *T. karelinii*))) is the most parsimonious and only tree number 4: (root / *T. carnifex* (*T. cristatus* (*T. dobrogicus*, *T. karelinii*))) is significantly longer than this best tree (Table 3).

Table 1. Allele frequency for 40 presumptive gene loci, surveyed in the *Triturus cristatus* superspecies and *T. marmoratus*. Alleles are enumerated alphabetically. Twenty-three different alleles not observed in the present study will be reported elsewhere (J.W. Arntzen, in prep.).

(Sub)species		*T. cristatus*			*T. c. carnifex*		*T. c. macedonicus*		*T. dobrogicus*
Population		Canterbury & Peterborough	Limanowa	Sinaia	Fuscaldo	Kramplje	Ano Kaliniki	Višegrad	Ečka
Country		UK	Poland	Romania	Italy	Slovenia	Greece	Bosnia-Herzegovina	Serbia and Montenegro
Population number		1	2	3	4	5	6	7	8
Sample size		10	10	10	11	10	11	15	14
Locus and alleles									
Acph-1	b	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Acph-2	a						0.05		
	b	1.00	1.00	1.00	1.00	1.00	0.86	1.00	0.96
	c						0.09		0.04
Ada	a	1.00	1.00	1.00	0.09	1.00	1.00	1.00	1.00
	b				0.91				
Alb	a								
	b				0.05				
	c	1.00	1.00	1.00	0.96	1.00	1.00	1.00	1.00
	d								
Adh	a						0.14		
	b	1.00	0.90	0.90	1.00	1.00	0.86	1.00	1.00
	c		0.10	0.10					
	d								
Ata-1	b	1.00	1.00	1.00	0.05	0.20	1.00	1.00	0.39
	c				0.96	0.80			0.61
Ata-2	a	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Est-1	a	1.00	1.00	1.00		0.15			1.00
	b								
	c				0.09	0.80	0.73	0.57	
	d				0.91	0.05	0.27	0.43	
	f								
Est-2	a			0.10					
	b	1.00	1.00	0.90	1.00	1.00	1.00	1.00	0.89
	c								0.11
	e								
GP	a								
	b								
	c	1.00	1.00	1.00	0.41	0.35	0.91	1.00	0.14
	d				0.59	0.65	0.09		0.86
Gdh	a	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	b								
Gpi	a								
	c						1.00	1.00	0.04
	d		0.25						
	e	0.80	0.75	1.00	1.00	0.65			0.96
	f								
	g					0.35			
	i	0.20							
G-6-pd	a						1.00	1.00	
	b	1.00	1.00	1.00	1.00	1.00			1.00
α-Gly	a	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	b								
Icd-1	a	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Icd-2	a	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	b								
Ldh-1	a				0.18				
	b					0.70	0.73	0.97	0.82
	c	1.00	1.00	1.00	0.73	0.25	0.14	0.03	0.18
	d				0.09	0.05	0.14		
Ldh-2	a								
	b							0.57	
	d	1.00	1.00	1.00	1.00	1.00	1.00	0.43	1.00
Lap	b	1.00	1.00	1.00	0.91	1.00	1.00	1.00	1.00
	c				0.09				

To the bottom: average allelic richness in reference to a standard population size of 20 genes (A) and average heterozygosity under the assumption of Hardy-Weinberg equilibrium (H_e).

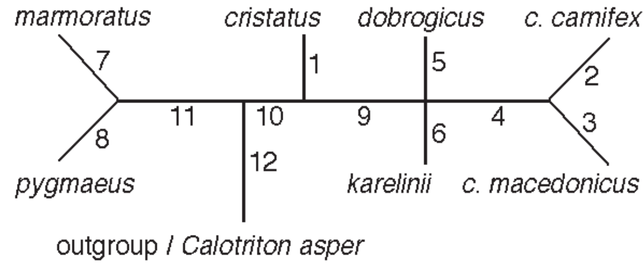
	T. karelinii			*T. marmoratus*		*T. pygmaeus*	
Svištov Bulgaria Zimnicea Romania	Bartin Turkey	Grivac Serbia	Rakovski Bulgaria	Confolens & Rochechouart France	El Berrueco Spain	Rio Alberito Spain	Venta del Charco Spain
9	10	11	12	13	14	15	16
4	10	9	4	5	7	20	10
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	0.93	0.95
						0.08	0.05
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
				1.00	1.00	0.95	1.00
						0.05	
1.00	0.85	1.00	1.00				
	0.15						
1.00	1.00	1.00	1.00			0.88	1.00
				1.00	1.00	0.13	
0.38	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.63							
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.25							
0.75			0.25				
	0.65	0.61	0.63				
	0.35	0.39	0.13				
				1.00	1.00	1.00	1.00
				1.00	1.00	1.00	1.00
1.00							
	1.00	1.00	1.00				
				1.00	1.00		
						1.00	1.00
0.25	1.00	1.00	1.00				
0.75							
1.00	0.95	1.00	1.00		0.07		
	0.05			1.00	0.93	1.00	1.00
0.13							
	0.05	0.11					
0.88	0.90	0.89	1.00				
				1.00	1.00	1.00	1.00
	0.05						
		0.11			0.14	0.25	
1.00	1.00	0.89	1.00	1.00	0.86	0.75	1.00
1.00	1.00	1.00	1.00				
				1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	0.88	1.00	1.00	1.00	1.00
			0.13				
0.88							
0.13	1.00	1.00	1.00	1.00	1.00	1.00	1.00
				1.00	1.00	1.00	1.00
	0.85	1.00	1.00				
1.00	0.15						
1.00	1.00	1.00	1.00				
				1.00	1.00	1.00	1.00

Table 1. Cont.

(Sub)species		*T. cristatus*			*T. c. carnifex*		*T. c. macedonicus*		*T. dobrogicus*
Population		Canterbury & Peterborough	Limanowa	Sinaia	Fuscaldo	Kramplje	Ano Kaliniki	Višegrad	Ečka
Country		UK	Poland	Romania	Italy	Slovenia	Greece	Bosnia-Herzegovina	Serbia and Montenegro
Mdh-1	b	1.00	1.00	0.95	1.00	0.95	1.00	0.83	0.32
	c								
	d			0.05		0.05		0.17	0.68
Mdh-2	b	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	c								
Me	a								
	b	1.00	1.00	1.00		0.05	0.27		
	c				1.00	0.95	0.73	1.00	1.00
	d								
Mpi	b				0.18				
	c	1.00	1.00	1.00	0.82	1.00	1.00	1.00	1.00
Nadhdh-1	b	1.00	1.00	0.55	1.00	1.00	1.00	1.00	1.00
	c			0.45					
Nadhdh-2	b								
	c			0.25	1.00	0.60	0.09	0.73	0.54
	d	0.95	1.00	0.60		0.40	0.91	0.27	0.46
	e	0.05		0.15					
NP	a				0.05				
	b			0.05	0.05				
	c	1.00	1.00	0.95	0.86	1.00	1.00	1.00	1.00
	d				0.05				
Pg-1	a		0.05	0.05		0.10			0.25
	b	1.00	0.95	0.95	1.00	0.90	1.00	1.00	0.75
Pg-2	b	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.93
	c								0.07
Pg-3	a		0.05					0.07	
	b	1.00	0.95	1.00	1.00	1.00	1.00	0.93	1.00
Pg-4	b	1.00	1.00	1.00	0.96	1.00	1.00	1.00	1.00
	c				0.05				
Pep-1	b				1.00				0.11
	d	1.00	1.00	1.00		0.50	0.86	1.00	0.89
	e					0.50	0.14		
Pep-2	a		0.05	0.15					
	b	1.00	0.95	0.85	1.00	1.00	1.00	1.00	1.00
	c								
Pgm-1	a						0.05		
	b	1.00	1.00	1.00	1.00	1.00	0.96	1.00	1.00
Pgm-2	a	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
6-Pgd	b								
	c	1.00	1.00	0.15					
	d			0.85	0.05	1.00	0.96	0.43	1.00
	e						0.05	0.57	
	g				0.96				
Sdh	a			0.15			0.64		0.07
	b	1.00	1.00	0.60	1.00	1.00	0.36	1.00	0.46
	c			0.25					0.46
Sod-1	a					0.10			
	b	1.00	1.00	1.00	1.00	0.90	1.00	1.00	1.00
Sod-2	b	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Trf	b			0.10					
	c	1.00	1.00	0.90	1.00	1.00	1.00	1.00	1.00
	d								
	f								
Xdh	b			0.05	0.55	0.15	0.14		0.11
	c	0.80	0.95	0.55	0.46	0.70	0.82	0.97	0.86
	d	0.20	0.05	0.40		0.15	0.05	0.03	0.04
A		1.08	1.15	1.38	1.36	1.38	1.37	1.18	1.35
Average		1.34			1.64		1.44		1.46
He (%)		1.93	2.46	9.15	7.33	10.54	8.43	6.20	9.83

	T. karelinii			*T. marmoratus*			*T. pygmaeus*
Svištov Bulgaria Zimnicea Romania	Bartin Turkey	Grivac Serbia	Rakovski Bulgaria	Confolens & Rochechouart France	El Berrueco Spain	Rio Alberito Spain	Venta del Charco Spain
0.50	0.75	1.00	0.75				
			0.25			1.00	0.40
0.50	0.25			1.00	1.00		0.60
1.00	1.00	0.06	0.25	1.00	1.00	1.00	1.00
		0.94	0.75				
				1.00	1.00		
		0.28				1.00	1.00
1.00							
	1.00	0.72	1.00				
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	0.45						
0.63	0.55	0.67	0.88				
0.38		0.33	0.13	1.00	1.00		
						1.00	1.00
	0.05	0.50					
0.88	0.95	1.00	0.50			0.03	
0.13				1.00	1.00	0.98	1.00
0.13				1.00	1.00	1.00	0.25
0.88	1.00	1.00	1.00			0.75	
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
				0.80	1.00	1.00	1.00
1.00	1.00	1.00	1.00	0.20			
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.13	0.05		0.75				
0.88	0.60	0.89	0.25			1.00	1.00
	0.35	0.11		1.00	1.00		
0.13							
0.88	0.85	1.00	1.00	1.00	1.00	1.00	1.00
	0.15						
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	0.55	1.00	0.88				
	0.45		0.13		0.14	0.08	
1.00				1.00	0.86	0.10	0.75
						0.83	0.25
	0.05	0.06	0.13				
1.00	0.85	0.94	0.88			1.00	1.00
	0.10			1.00	1.00		
				1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00				
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
1.00	1.00	1.00	1.00				
				1.00	1.00	0.98	1.00
						0.03	
0.13				0.90	1.00		
0.75	1.00	1.00	1.00	0.10			
0.13						1.00	1.00
1.33	1.38	1.20	1.25	1.05	1.08	1.16	1.10
	1.44			1.13		1.22	
11.16	9.84	5.62	8.66	1.39	1.68	3.15	3.49

Figure. 2. Hypothetical phylogeny for the genus *Triturus*. Branches are numbered as in Table 2.



Genetic distances ranged from *D*N = 0.001 among populations of *T. marmoratus* and *D*N ≈ 0.035 among populations of *T. karelinii* and *T. pygmaeus*, to *D*N = 0.11 among populations of *T. c. carnifex* (Table 2). The maximum observed genetic distance within a species was *D*N = 0.25 between *T. c. carnifex* from Fuscaldo, Italy and *T. c. macedonicus* from Ano Kaliniki, Greece. This single value exceeded the average value among crested newt species (*D*N = 0.19) and among marbled newt (*D*N = 0.19) species. The crested newt and marbled newt species groups showed genetic differentiation at the level of *D*N = 0.68. Linear interpolation of the pairwise genetic distances against a calibration point of 24 Ma (million years before present) for the most recent common ancestor (MRCA) of *T. cristatus* - *T. marmoratus* (STEINFARTZ, S. *et al.*, 2007) yielded estimates in the range of 7-6 Ma for the age of each of the six *Triturus* species and an estimate of 1.3 Ma for the Asian versus European populations of *T. karelinii* (Table 2).

Mitochondrial DNA-sequences

Fifteen different haplotypes were identified among the crested newt mtDNA-sequences, four different haplotypes were found for the marbled newts and three for *Calotriton asper* (Appendix II). The maximum observed sequence divergence ranged from d_{K80} = 0 (haplotypes identical) in *T. pygmaeus*, d_{K80} ≈ 0 in *T. cristatus* and *T. c. macedonicus*, d_{K80} ≈ 0.01 in *T. dabrogicus,* d_{K80} ≈ 0.02 in *T. c. carnifex* and *T. marmoratus* to d_{K80} ≈ 0.06 in *T. carnifex* and *T. karelinii* and d_{K80} = 0.09 in *T. dobrogicus* (Table 2). Interspecific comparisons yielded mean values of d_{K80} ≈ 0.05 among the marbled newts, d_{K80} ≈ 0.10 among crested newts and d_{K80} = 0.21 between marbled and crested newts. The mtDNA-sequence data also indicate a marked substructuring within *T. karelinii*, between the Asian (haplotypes h11-h13) and European populations (haplotypes h14 and h15) at the level of d_{K80} = 0.053.

Table 2. Left panel. Bootstrap replication scores (brs) and Bayesian posterior probabilities (pp) for branches in the phylogenetic tree of the genus *Triturus* as estimated under maximum parsimony (MP), neighbour-joining (NJ), maximum likelihood (ML) and Bayesian (B) methods of inference for allozyme data (top panel) and for mtDNA-sequence data (lower panel). Clades with consistently low statistical support are collapsed (bootstrap replication score (brs) not exceeding 0.70 or posterior probability (pp) not exceeding 0.90). The remaining branches are numbered as in Figure 2. Internal interspecific branches are shown in boldface type. Right panel: observed maximum and hierarchical averages of genetic distances and divergence times between species and subspecies estimated from an external calibration point at 24 Ma (STEINFARTZ, S. *et al.*, 2007) by linear interpolation and with the software r8s (SANDERSON, M.J., 2004; details see text). The calibration point refers to the crown node of branch 12 and represents the most recent common ancestor (MRCA) of all *Triturus* species. Abbreviations are: car = *T. c. carnifex*, cri = *T. cristatus*, dob = *T. dobrogicus*, kar = *T. karelinii*, mac = *T. c. macedonicus*, mar = *T. marmoratus*, pyg = *T. pygmaeus*, n.a. = not applicable and CI is 95% confidence interval.

Allozymes

		Method of phylogenetic inference			Nei's unbiased genetic distance		estimated MRCA in million years
Branch number	Clade	MP brs	NJ brs	ML brs	observed maximum	hierarchical average ✤	by linear interpolation
1	cristatus	<0.50	0,95	0,76	0,034		
2	c. carnifex	<0.50	<0.50	<0.50	0,111		
3	c. macedonicus	0,78	0,87	0,82	0,043		
4	carnifex	<0.50	<0.50	<0.50	0,245	0,192	6,8
5	dobrogicus	0,85	0,90	0,70	0,017		
6	karelinii	0,95	0,99	0,97	0,039	0,038	1,3
7	marmoratus	0,99	1,00	0,99	0,001		
8	pygmaeus	0,61	0,97	0,85	0,035		
9	**car+dob+kar+mac**	**0,71**	**<0.50**	**<0.50**		0,171	6,1
10+11	**cristatus superspecies**	**1,00**	**1,00**	**1,00**		0,186	6,6
10+11	**mar+pyg**	**idem**	**idem**	**idem**		0,194	6,9
12	**genus Triturus**	**n.a.**	**n.a.**	**n.a.**		0,678	24 #

mtDNA sequence data

		Method of phylogenetic inference				Kimura two-parameter distance		estimated MRCA in million years		
									with software r8s	
Branch number	Haplotype clade	MP brs	NJ brs	ML brs	B pp	observed maximim	hierarchical average	by linear interpolation	mean (mode)	CI
1	cristatus	1,00	1,00	0,99	1,00	0,002				
2	c. carnifex	0,91	0,99	0,94	0,99	0,022				
3	c. macedonicus $	1,00	1,00	1,00	1,00	0,002				
4	carnifex	0,78	0,96	0,88	1,00	0,059	0,051	5,8	6.8 (6.5)	4.3-9.3
5	dobrogicus	1,00	1,00	0,95	1,00	0,093				
6	karelinii $	0,99	1,00	0,98	1,00	0,056	0,053	6,1	6.2 (6.1)	4.0-8.4
7	marmoratus	1,00	0,98	0,77	0,96	0,017				
8	pygmaeus &	1,00	1,00	1,00	1,00					
9	**car+dob+kar+mac**	**0,51**	**<0.50**	**<0.50**	**<0.50**		0,091	10,4	*	
10	**cristatus superspecies**	**1,00**	**1,00**	**0,99**	**1,00**		0,095	10,9	11.6 (11.2)	8.9-14.2
11	**mar+pyg**	**1,00**	**1,00**	**1,00**	**1,00**		0,047	5,4	5.4 (5.0)	3.3-7.6
12	**genus Triturus**	**1,00**	**1,00**	**1,00**	**1,00**		0,210	24 #		

✤determined through UPGMA-clustering.
# calibration point following Steinfartz et al., 2007.
$ haplotypes from the Višegrad population included with T. karelinii.
& represented by one haplotype.
* tree not supported.

One hundred and ninety two nucleotide positions (29.9%) were variable and 178 (27.7%) parsimony informative at the level of the genus *Triturus* and 133 nucleotide positions (20.7%) were variable and 120 (18.7%) parsimony informative at the level of the *T. cristatus* superspecies. The four methods of phylogenetic inference (MP, NJ, ML and Bayesian) applied to the unique haplotypes yielded essentially equivalent results. This included, at best, weak support for a position of *T. cristatus* as sister-group to the other crested newt species (Table 2). A discrepancy was observed between signals from the nuclear and mitochondrial markers in the position of the population from Višegrad. This population classifies on the basis of allozymes as *T. c. macedonicus* whereas the mitochondrial haplotype that it carries is typical for *T.*

Table 3. Difference in tree length between the phylogeny showing maximum parsimony and 11 alternative topologies for the *Triturus cristatus* superspecies, for respectively allozyme and mtDNA-sequence data, tested with the Kishino-Hasegawa and Templeton tests (* - P < 0.05, and NS – not significant). The twelve topologies are constructed under the proviso that the marbled newts are the sister group to the crested newts and that OTU's (i.e., populations, respectively haplotypes) fall within the species for which they are recognized. Note however that *karelinii*-like haplotypes observed in *Triturus carnifex macedonicus* from Višegrad were legitimately included with *T. karelinii* (for details see text).

Tree	Species splitting off as seen from the root		Allozymes		mtDNA sequences	
number	first	second	Kishino-Hasegawa	Templeton	Kishino-Hasegawa	Templeton
1	cristatus	carnifex	NS	NS	best tree	best tree
2		dobrogicus	best tree	best tree	NS	NS
3		karelinii	NS	NS	*	NS
4	carnifex	cristatus	*	*	NS	NS
5		dobrogicus	NS	NS	*	NS
6		karelinii	NS	NS	NS	NS
7	dobrogicus	carnifex	NS	NS	NS	NS
8		cristatus	NS	NS	NS	NS
9		karelinii	NS	NS	NS	NS
10 #	karelinii $	carnifex	NS	NS	NS	NS
11		cristatus	NS	NS	*	*
12		dobrogicus	NS	NS	(*)	NS

# equivalent to NRBV-constraint (see text).
$ two haplotypes from the Vis'egrad population included with T. karelinii.

karelinii (Figure 3). This situation was discovered for several populations in northern Serbia before (WALLIS, G.P., J.W. ARNTZEN, 1989; ARNTZEN, J.W., G.P. WALLIS, 1999).

The MP-tree obtained under the NRBV-constraint was 1.0% longer than the most parsimonious solution. The difference in tree length is not significant under the Kishino-Hasegawa and the Wilcoxon signed-ranks tests (P > 0.05). Under the prerequisite that haplotypes fall within the species for which they are recognized (with the exception of haplotypes from Višegrad), the favoured tree was number 1: (root / *T. cristatus* (*T. carnifex* (*T. dobrogicus*, *T. karelinii*))). Trees 3, 4, 5 and 11 are statistically less parsimonious than this best tree, depending on the testing procedure (Table 3).

Interpolation of the d_{K80} distances against the 24 Ma calibration point for the most recent common ancestor of *T. cristatus* - *T. marmoratus* yielded estimates in the range of c. 11 Ma for the radiation of the *Triturus cristatus* superspecies, c. 6 Ma for the MRCA of *T. c. carnifex* and *T. c. macedonicus* and for the Asian versus European populations of *T. karelinii* and c. 5 Ma for the split between *T. marmoratus* and *T. pygmaeus* (Table 2).

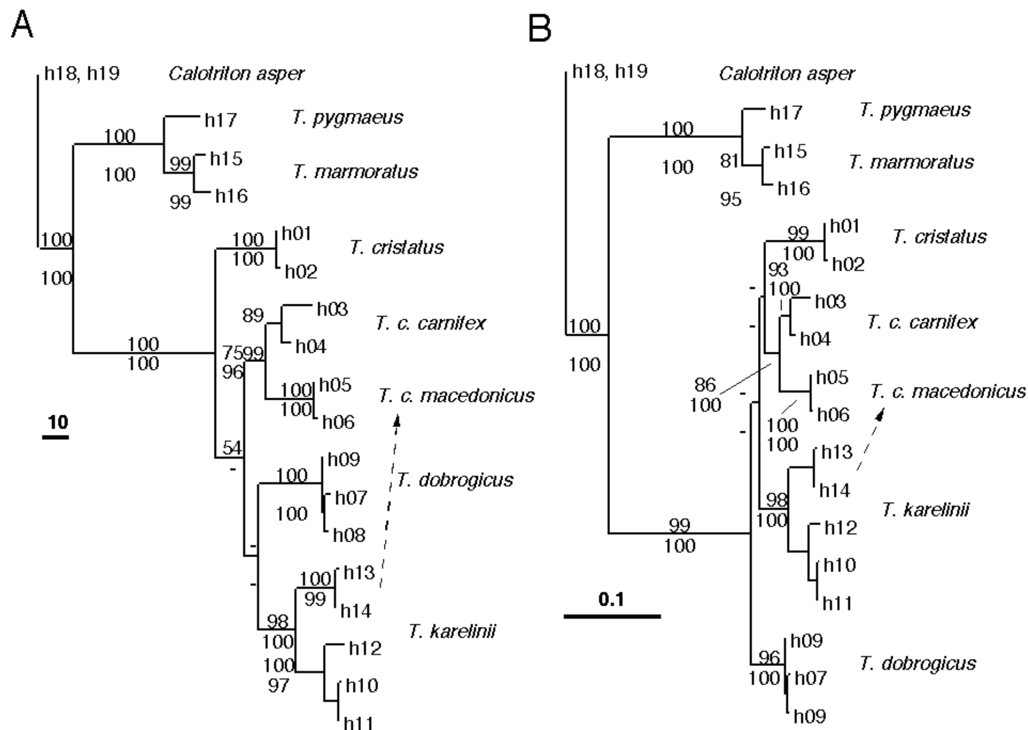


Figure 3. Phylogenetic trees for the genus *Triturus* on the basis of ND4 mtDNA haplotypes (h01-h19), with *Calotriton asper* (h20-h22) as outgroup, with the preferred solution under (a) maximum parsimony and (b) maximum likelihood. Haplotype numbers are as presented in Appendix II. Values over branches are percent bootstrap replications scores for maximum parsimony and maximum likelihood (above branches) and neighbour-joining and Bayesian posterior probabilities (below branches). Values < 50 are indicated by a hyphen (-). Note that four individuals of *T. carnifex macedonicus* from Višegrad carry haplotypes h14 or h15, which are otherwise characteristic of *T. karelinii.*

Discussion

Genetic coherence of species and subspecies

The monophyly of the *Triturus cristatus* superspecies and its sister-group relationship to the marbled newts are unambiguously supported by mtDNA and allozymes, through outgroup rooting and midpoint rooting, respectively. The two data sets generally agree with one another and with current taxonomy, in the recognition of two species of marbled newts, four species of crested newts plus the subspecies *T. c. macedonicus*. However, the allozyme data are less explicit than the mtDNA-sequence data and do not provide strong support for the monophyly of *T. carnifex* and its subspecies *T. c. carnifex* (all three methods of phylogenetic inference; see also ARNTZEN, J.W., 2001), nor for the genetic coherence of *T. cristatus* and *T. pygmaeus* under maximum parsimony (Table 2). Also the subdivision of *T. karelinii* over two European and one Asian populations is stronger supported by mtDNA data than with

allozymes. Conversely, the Višegrad population, which belongs to *T. c. macedonicus* on the basis of external morphology, NRBV and allozymes, would erroneously be classified as *T. karelinii* on the basis of mtDNA. The extensive asymmetric introgression of mtDNA across species boundaries has been reported manifold. An extreme case is documented for the newt *Lissotriton montandoni* (Boulenger, 1880). This species has turned 'invisible' from the mitochondrial perspective, due to the complete replacement of indigenous mtDNA by that of *L. vulgaris* (Linnaeus, 1758) (BABIK, W. *et al.*, 2005). In order to compare patterns of introgression between nuclear and mtDNA markers and to distinguish between recent introgression and ancients polymorphisms cq. incomplete lineage sorting, it would be required to conduct a wider spatial survey with populations at various distances to species borders (J.W. Arntzen, in prep.).

Phylogeny

We gathered data from 40 allozyme loci and DNA-sequence information from a stretch of a rapidly evolv- ing protein coding mitochondrial gene (ND4) to reconstruct the evolutionary history of the *T. cristatus* superspecies, employing a cross-section of phylogenetic methodologies (MP, NJ, ML and Bayesian). However, neither the allozyme nor the mtDNA data resolve the phylogeny of crested newts at the species level. The preferred topologies are diverse and have in common that internal branches at the species level gain low statistical support (Table 2, Fig. 3). The absence of phylogenetic resolution can not be attributed to either too high or too low a level of genetic differentiation at the average enzyme locus (see Table 1 and genetic distance information in Table 2). Also, homoplasy in the DNA-sequences is not a major issue (see results). Both data sets contradict the NRBV-based phylogenetic hypothesis and because of the correlation between NRBV character-states and the ecology of the species (in terms of the general slender–stout morphological series and the length of the annual aquatic phase) we suggest that NRBV is under natural selection and, hence, not an informative character for building a phylogeny. In the absence of phylogenetic resolution, we suggest that the polytomy recovered in the crested newts is 'hard' and that the suggested simultaneous origin of four lineages reflects reality (WHITFIELD, J.B., P.J. LOCKHART, 2007). However, 'absence of proof' shall not be confused with 'proof of absence'. In other words, this hard polytomy constitutes a

hypothesis that may be falsified, e.g., through the sequencing of the full mitochondrial DNA genome (c. 16,000 bp) and the sequencing of nuclear genes.

Historical biogeography

The Balkan Peninsula, where the four species of crested newts meet (Fig. 1), has been recognized as the centre of origin of crested newt diversity (CRNOBRNJA-ISAILOVIC, J. *et al.*, 1997). Our data support this interpretation, to which we add that the different species have arisen in close temporal proximity. Whether the evolutionary split of one into four lineages was truly simultaneous (i.e., a 'hard' polytomy with branch lengths of zero in the phylogenetic tree) or a nearsimultaneous 'soft' polytomy (with small but non-zero branch lengths in a topology that is currently unresolved) is immaterial for the purpose of our biogeographical analysis, given the rough resolution - in terms of both time and space - provided by the available paleogeographical reconstructions (see below).

A temporal calibration of divergence times within the Salamandridae was provided by Steinfartz *et al*. (2007), who used a fossil-based molecular dating approach, cross-validated with palaeogeographical data for the Holarctic region. The MRCA of the crested and marbled newts was dated at 24 Ma. Accordingly, we estimate the radiation of the proto-crested newt into four lineages (corresponding with the four species) by linear interpolation separately for the allozyme data and for the mtDNA-sequence data. This yields, however, two different solutions (Table 2) and the discrepancy is reflected in different biogeographical reconstructions (scenarios 1 and 2 below). Popov *et al*. (2004) present a series of ten paleogeographic maps for southeastern Europe that cover the period from 37 Ma to the present. Schematic versions of map 6 of the Mid Middle Miocene at 14-13 Ma, map 7 of the Late Middle Miocene at 12-11 Ma and map 8 of the Mid Late Miocene at 8.5-7 Ma are shown in Figure 4.

Scenario 1 – allozymes. Following the allozyme molecular clock, the as yet undifferentiated protocrested newt would be positioned on the central Balkan Peninsula in the Late Miocene. At 8.5-7 Ma a solid, unfragmented landmass encircled the Pannonian Basin and Europe and Asia were not disconnected (Figure 4c). This allows for the spread of crested newts over the region, as was proposed by Crnobrnja-Isailović *et al*. (1997: Figure 2). From one central stock positioned at approximately the present day city of Sofia, the crested newts dispersed in four compass directions:

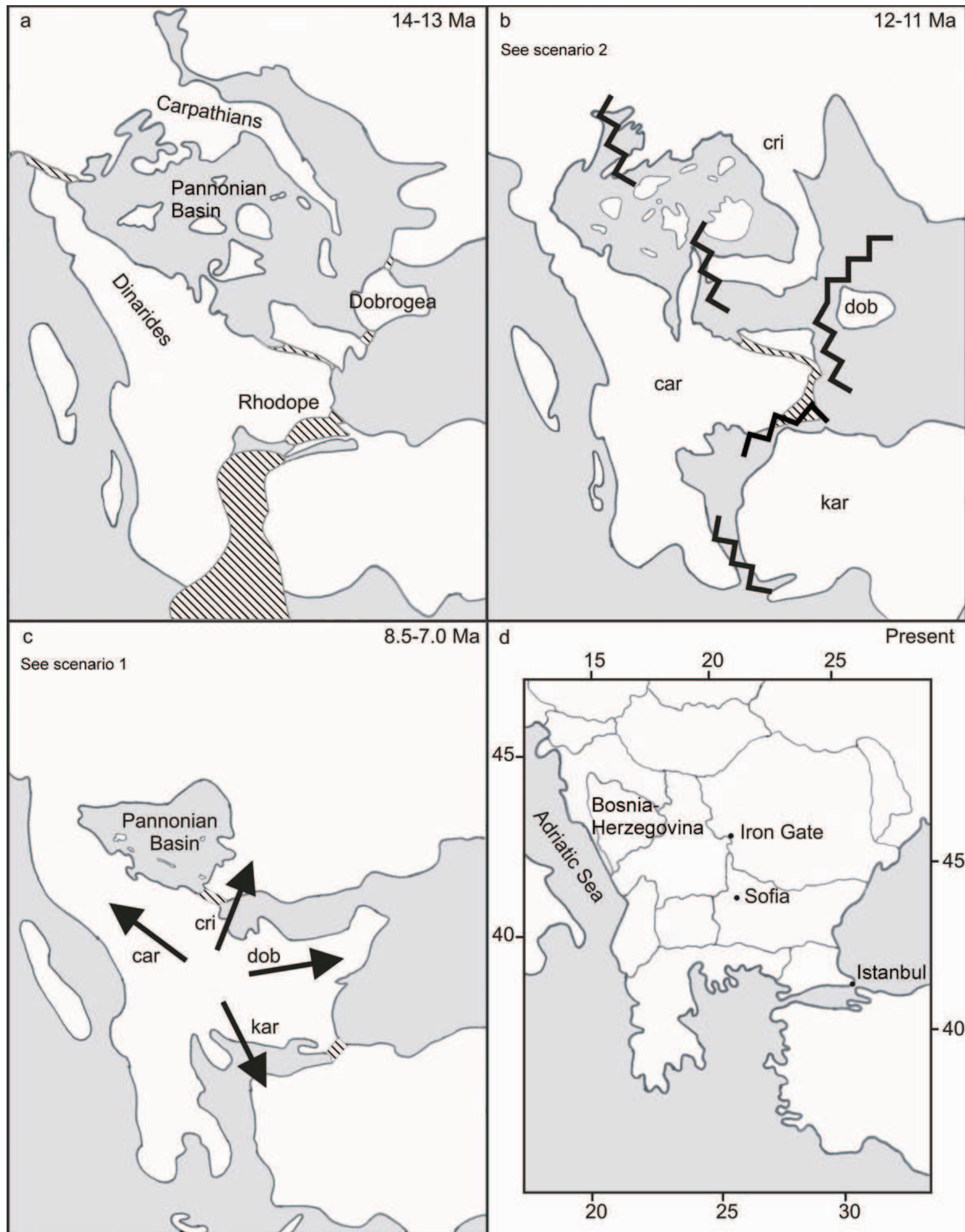


Figure 4. Paleogeographical reconstruction of south-eastern Europe according to Popov *et al*. (2004), with land-masses unshaded and seas shaded. Question marks in the original publication are represented by hatching. (a) Situation in the Mid Middle Miocene (14-13 Ma), (b) situation in the Late Middle Miocene (12-11 Ma), with biogeographical hypothesis according to scenario 2, (c) Mid Late Miocene (8.5- 7.0 Ma) with biogeographical hypothesis according to Crnobrnja-Isailović *et al.* (1997) (scenario 1), and (d) present day. Abbreviations are as in Table 2. Proposed vicariant effects are shown by zig-zag lines.

the lineage to the east gave rise to *T. dobrogicus*, the lineage to the south gave rise to *T. karelinii*, the lineage to the west dispersed along the Dinarides and formed *T. carnifex* and one lineage dispersed northwards along the Carpathians and formed *T.*

cristatus. However, no vicariant events were postulated that would promote speciation through genetic differentiation in allopatry. Indeed, no vicariant events potentially associated with cladogenetic events are apparent for the region from the paleogeography of southeastern Europe over the period 6.1 Ma to 1.8 Ma (POPOV, S.V. *et al.*, 2004: maps 9 and 10).

Scenario 2 – mitochondrial DNA. According to the mitochondrial DNA molecular clock, one would place the as yet undifferentiated proto-crested newt in the central Balkan Peninsula in the Middle Miocene. At 14-13 Ma the area was covered by a presumably continuous landmass encircling the Pannonian Basin. Turkey and Europe were connected (Figure 4a). In particular, Dobrogea was not yet an island. During this époque, no major barriers would seem to hamper the spread of the species over the area. Two million years on (12-11 Ma, Figure 4b), the different areas keeping the four species *in statu nascendi* became disconnected. The isolation was most pronounced for Dobrogea that harboured the *T. dobrogicus* lineage. Interestingly, *T. dobrogicus* is the first species to arise in the maximum likelihood and Bayesian analysis of the mtDNA data, be it with low statistical support (Figure 3b). The Carpathians were disconnected from the mountains on the Balkans (Dinarides, Rhodope) by a sea-straight in position of the present day Iron Gate. This sea-strait separated the *T. cristatus* and *T. carnifex* lineages and a sea-strait between Europe and Asia separated the *T. karelinii* lineage (Fig. 4b). At 8.5-7.0 Ma Dobrogea became reconnected to the continent and the sea-straits at the Iron Gate and north of present day Istanbul may have closed (Fig. 4c). This brought the four species back into contact and allows for the relatively small interspecific range adjustments that are required to explain the present day distribution. Scenario 2 supports the near-synchronous crested newt radiation dated at either c. 13.5 Ma (paleographic data) or c. 11 Ma (molecular data).

It invokes vicariant effects brought about by major changes in the land-water distribution over the area. The scenario also allows for a long period - between 14-13 Ma and 8.5-7 Ma - over which the four lineages started their independent evolutionary routes and acquired a large degree of reproductive isolation. To position the crested newt radiation at precisely 10.9 Ma (Table 2, Fig. 4b) would require extensive dispersal including the crossing of sea-straits to reach the various landmasses that - on the basis of the current distribution of crested newts - would logically be associated with the four species (Fig. 4b). Amphibians however have generally poor dispersal

capabilities and their propensity to cross major barriers such as sea-straits is small. Indeed, present day sea straights such as the Bosporus and the Straights of Dover and Gibraltar are effective barriers to dispersal (BEERLI, P. *et al.*, 1996; VEITH, M. *et al.*, 2004; RECUERO, E. *et al.*, 2007).

We prefer scenario 2 over scenario 1. Scenario 2 is detailed and explicit. It naturally suggests the near-simultaneous origin of the four species comprising the crested newt superspecies. Moreover, scenario 2 involves only a small discrepancy in the exact timing of the radiation as c. 11 Ma from molecular data versus > 12-11 Ma (Fig. 4b) and < 14-13 Ma (Fig. 4a) from paleogeographic data. By consequence, we prefer the mtDNA calibration of the molecular clock over the allozyme calibration. This calibration dates the origin of *T. c. carnifex* of the Apennine Peninsula and *T. c. macedonicus* of the Balkan Peninsula at 5.8 Ma. These two subspecies are allopatric, separated from one another by the karst region of Bosnia-Herzegovina. The date approximately coincides with the filling of the Mediterranean that followed the Messinian Event at 5.3 Ma (KRIJGSMAN, W. *et al.*, 1999). So, here we invoke yet another vicariance effect to explain the evolutionary history of the genus *Triturus* and suggest that *T. c. carnifex* and *T. c. macedonicus* started to diverge following the origin of the Adriatic Sea.

Taxonomic considerations

Handbooks of the past century either position a parapatric range border between *T. carnifex* and *T. karelinii* in the northern Balkans (THORN, R., 1968), or place their range border at the Bosporus (STEWARD, J.W.). It was not recognized that the large central part of the peninsula, approximately coinciding with Bosnia-Herzegovina, is devoid of *Triturus* newts. Arntzen and Wallis (1999) documented the eastern range limit of *T. carnifex* as running between Croatia and Bosnia- Herzegovina. They resurrected the taxon *Triturus c. macedonicus* (with a distribution over the Balkans south and east of Bosnia-Herzegovina, Fig. 1) that was originally described as *Molge karelinii* var. *macedonica* by S. Karaman in 1922. The subspecies was classified as *T. carnifex* (and not *T. karelinii*) on the basis of NRBV and mtDNA RFLP analysis (ARNTZEN, J.W., G.P. WALLIS, 1999) and our new sequence data support this interpretation. We observed a high level of genetic differentiation between *T. c. macedonicus* and the nominotypical subspecies (*D*N = 0.192, *d*K80 = 0.051), paralleling RFLP data (WALLIS, G.P., J.W. ARNTZEN, 1989; FAITH, D.P., P.S.

CRANSTON, 1991). Morphologically the taxa are distinguishable in colouration and spotting pattern (ARNTZEN, J.W., G.P. WALLIS, 1999; ARNTZEN, J.W., 2003: 459). This raises the question whether *T. c. macedonicus* had better be considered a species. Unfortunately, the biological species concept cannot be made operational due to the subspecies' allopatry. In the absence of a direct test we use an indirect line of reasoning. Firstly, the genetic differentiation between *T. c. carnifex* and *T. c. macedonicus* is comparable to that of *T. marmoratus* and *T. pygmaeus* (Table 2), two former subspecies for which species status has recently been corroborated (see CHAPTER 7). Secondly, the forms are old (we estimate >5 Ma) and may have followed independent evolutionary pathways ever since. We therefore elevate the taxonomic status of *T. c. macedonicus* to that of the species, i.e., *Triturus macedonicus* (Karaman, 1922).

Acknowledgements

We thank S. Carranza (Barcelona, Spain) for tissue material of *Calotriton asper*.

References

ARÉVALO, E., DAVIS S.K., SITES J.W., JR. (1994) Mitochondrial DNA sequence divergence and phylogenetic relatonships among eight chromossome races of Sceloporus grammicus complex (PHRYNOSOMATIDAE) in Central Mexico. *Systematic Biology* **43**, 387 - 418.

ARNTZEN, J.W., P. W.G. (1994) The 'Wolterstorff Index' and its value to the taxonomy of the Crested Newt superspecies. *Abhandlungen und Berichte fÓr Naturkunde* **17**, 57-66.

ARNTZEN, J.W., WALLIS G.P. (1999) Geographic variation and taxonomy of crested newts (*Triturus cristatus* supersecies): morphological and mitochondrial data. *Contribution to Zoology* **68**, 181-203.

ARNTZEN, J.W. (2001) Genetic variation in the Italian crested newt, Triturus carnifex, and the origin of a non-native population north of the Alps. *Biodiversity and Conservation* **10**, 971-987.

ARNTZEN, J.W. (2003) *Triturus cristatus* Superspecies–Kammolch-Artenkreis. In: *Handbuch der Reptilien und Amphibien Europas. Schwanzlurche IIA* (eds. GROSSENBACHER, K., THIESMEIER B.), pp. 421-514. Aula-Verlag, Wiebelsheim.

BABIK, W., BRANICKI W., CRNOBRNJA-ISAILOVIC J., *et al.* (2005) Phylogeography of two European newt species- discordance between mtDNA and morphology. *Molecular Ecology* **14**, 2475-2491.

BEERLI, P., HOTZ H., UZZELL T. (1996) Geologically dated sea barriers calibrate a protein clock for aegean water frogs. *Evolution* **50**, 1676-1687.

BOLKAY, J. (1928) Die Schädel der Salamandrinen, mit besonderer Rücksicht auf ihre systematische Bedeutung. *Anatomy and Embryology* **86**, 259-319.

CRNOBRNJA-ISAILOVIC, J., DZUKIC G., KRSTIC N., KALEZIC M.L. (1997) Evolutionary and paleogeographical effects on the distribution of the *Triturus cristatus* superspecies in the central Balkans. *Amphibia-reptilia* **18**, 321.

FAITH, D.P., CRANSTON P.S. (1991) Could a cladogram this short have arisen by chance alone?: On permutation tests for cladistic structure. *Cladistics* **7**, 1-28.

FELSENSTEIN, J. (1993) PHYLIP version 3.63. Dept of Genetics, Univ of Washington, Seattle.

FUHN, I.E. (1960) Amphibia. Fauna Republicii Populare Romine, Vol. XIV part 1. Academiei RP Romine, Bucuresti.

GARCÍA-PARÍS, M., MONTORI A., HERRERO P. (2004) Amphibia: Lissamphibia. In: *Fauna Ibérica*. Consejo Superior de Investigaciones Científicas.

GARCÍA-PARÍS, M., HERRERO, P., MARTÍN, C., DORDA, J., ESTEBAN, M. & ARANO, B. (1993) Morphological characterization, cytogenetic analysis, and geographical distribution of the Pigmy Marbled Newt Triturus marmoratus pygmaeus (Wolterstorff, 1905) (Caudata: Salamandridae). *Bijdraden tot de Dierkunde* **63**, 3-14.

GVOZDÍK, L., VAN DAMME R. (2006) *Triturus* newts defy the running-swimming dilemma. *Evolution* **60**, 2110-2121.

HARVEY, P.H., PAGEL M.D. (1991) *The comparative method in evolutionary biology* Oxford University Press Oxford, United Kingdom.

IVANOVIĆ, A., VUKOV T., DŽUKIĆ G., TOMAŠEVIĆ N., KALEZIĆ M. (2007) Ontogeny of skull size and shape changes within a framework of biphasic lifestyle: a case study in six Triturus species (Amphibia, Salamandridae). *Zoomorphology* **126**, 173.

KALINOWSKI, S.T. (2005) HP-rare 1.0: a computer program for performing rarefaction on measures of allelic richness. *Molecular Ecology Notes* **5**, 187-189.

KRIJGSMAN, W., HILGEN F.J., RAFFI I., SIERRO F.J., WILSON D.S. (1999) Chronology, causes and progression of the Messinian salinity crisis. *Nature* **400**, 652.

MADDISON, D.R., MADDISON W.P. (2005) MacClade 4: analysis of phylogeny and character evolution, version 4.08. *Sunderland (Massachusetts): Sinauer Associates*.

NYLANDER, J.A. (2004) MrModelTest2. Computer program distributed by the author. Evolutionary Biology Centre, Uppsala University, Sweden.

POPOV, S.V., RÖGL F., ROZANOV A.Y.*, et al.* (2004) Lithological-paleogeographic Maps of Paratethys: 10 Maps Late Eocene to Pliocene. *Courier Forschunginstitut Senckenberg* **250**, 1-46.

POSADA, D., CRANDALL K.A. (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817-818.

RAMBAUT, A., DRUMMOND A.J. (2007) Tracer v1.4, Available from http://beast.bio.ed.ac.uk/Tracer.

RECUERO, E., IRAOLA A., RUBIO X., MACHORDOM A., GARCÍA-PARÍS M. (2007) Mitochondrial differentiation and biogeography of *Hyla meridionalis* (Anura: Hylidae): an unusual phylogeographical pattern. *Journal of Biogeography* **34**, 1207-1219.

RENSCH, B. (1929) *Das Prinzip geographischer Rassenkreise und das Problem der Artbildung* Gebrüder Borntraeger.

RIDLEY, M. (1993) *Evolution* Blackwell Publishing, Oxford.

RONQUIST, F., HUELSENBECK J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

SANDERSON, M.J. (2004) r8s, version 1.70. *User's manual and software (available at: http://ginger.ucdavis.edu/r8s/index.html)*.

STEINFARTZ, S., VICARIO S., ARNTZEN J.W., CACCONE A. (2007) A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus* newts. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **308B**, 139-162.

STEWARD, J.W. *The tailed amphibians of Europe* David & Charles.

SWOFFORD, D.L., SELANDER R.B. (1981) BIOSYS-1: a FORTRAN program for the comprehensive analysis of electrophoretic data in population genetics and systematics. *Journal of Heredity* **72**, 281.

SWOFFORD, D.L. (2003) PAUP*: Phylogenetic Analysis Using Parsimony (and other methods) 4.0 Beta 10. Sunderland (MA): Sinauer Associates.

THORN, R. (1968) *Les Salamandres d'Europe, d'Asie et d'Afrique du Nord* Editions Paul Lechevalier.

VEITH, M., MAYER C., SAMRAOUI B., BARROSO D.D., BOGAERTS S. (2004) From Europe to Africa and vice versa: evidence for multiple intercontinental dispersal in ribbed salamanders (Genus *Pleurodeles*). *Journal of Biogeography* **31**, 159-171.

WALLIS, G.P., ARNTZEN J.W. (1989) Mithocondrial-DNA variation in the crested newt superspecies: limited cytoplasmic gene flow among species. *Evolution* **43**, 88-104.

WHITFIELD, J.B., LOCKHART P.J. (2007) Deciphering ancient rapid radiations. *Trends in Ecology & Evolution* **22**, 258-265.

WIENS, J.J. (2000) Reconstructing phylogenies from allozyme data: comparing method performance with congruence. *Biological Journal of the Linnean Society* **70**, 613-632.

WOLTERSTORFF, W. (1923) Übersicht den Unterarten und Formen des Triton cristatus Laur. *Blätter für Aquarien-und Terrarienkunde* **34**, 120-126.

XIA, X., XIE Z. (2001) DAMBE: Software Package for Data Analysis in Molecular Biology and Evolution. *J Hered* **92**, 371-373.

ZAJC, I., ARTZEN J.W. (1999) Phylogenetic relationships of the European newts (genus Triturus) tested with mitochondrial DNA sequence data. *Contributions to Zoology* **68**, 73-81.

ZARDOYA, R., MALAGA-TRILLO E., VEITH M., MEYER A. (2003) Complete nucleotide sequence of the mitochondrial genome of a salamander, Mertensiella luschani. *Gene* **317**, 17-27.

Appendix 1

Material used in the present study. Shown are taxon name, origin of the sample, sample size and voucher information. ZMA = Herpetology Department at the Zoological Museum Amsterdam, The Netherlands. N.a. = not available.

Population		Coordinates		Sample size		Voucher at ZMA Herpetology	
Taxon	Locality and country	latitude	longitude	allozymes	mtDNA	Collection number	Individual numbers
Triturus cristatus							
1a	Canterbury, United Kingdom	51°17'N	01°05'E	5	2	7802, 9200	594, 595, 967, 970, 971
1b	Peterborough, United Kingdom	52°35'N	00°15'W	5	3	9144, 9199	583, 972, 973, 968, 969
2	Limanowa, Poland	49°43'N	20°25'E	10	5	9135	584-593
3	Sinaia, Romania	45°20'N	25°33'E	10	10	9156	369-371, 698-704
Triturus carnifex carnifex							
4	Fuscaldo, Italy	39°25'N	16°02'E	11	5	7849, 9108	340, 405-414
5	Kramplje, Slovenia	45°44'N	14°30'E	10	4	9132	312-317, 618, 619, 785, 786
Triturus carnifex macedonicus							
6	Ano Kaliniki, Greece	40°52'N	21°26'E	11	6	9085	602-605, 658-661, 674-676
7	Višegrad, Bosnia and Herzegovina	43°47'N	19°20'E	15	4	9168	896-909, 962
Triturus dobrogicus							
8	Ečka, Serbia	45°18'N	20°27'E	14	5	9104, 9188	415-425, 510, 746, 747
9a	Svištov, Bulgaria	43°37'N	25°21'E	2	2	9160	459, 748
9b	Zimnicea, Romania	43°40'N	25°22'E	2	2	9171	336, 705
Triturus karelinii							
10	Bartin, Turkey	41°37'N	32°20'E	10	5	7564	116-125
11	Grivac, Serbia	43°58'N	20°40'E	9	5	9115	817-824, 893
12	Rakovski, Bulgaria	42°16'N	24°58'E	4	4	9149	515, 766-768
Triturus marmoratus							
13a	Confolens, France	46°01'N	00°40'E	1	1	9095	518
13b	Rochechouart, France	45°49'N	00°50'E	4	3	9151	535-538
14	El Berrueco, Spain	40°54'N	03°53'W	7	5	7614	161-167, 980-982
Triturus pygmaeus							
15	Rio Alberite, Spain	36°24'N	05°39'W	20	3	7552, 7618	77-95, 180
16	Venta del Charco, Spain	38°12'N	04°16'W	10	n.a.	7677	282-291
Calotriton asper							
17	Berga, Spain	42°06'N	01°50'E	n.a.	2	n.a.	
18	Valle Hecho, Spain	42°44'N	00°45'W	n.a.	2	n.a.	

Appendix 2

Twenty-two different ND4-haplotypes (h01-h22) were observed. Their corresponding GenBank accession numbers are AM900468-89. The haplotype distribution over populations and individuals is as follows: h01 - Canterbury, specimens 594 and 595; Limanowa 589-593; Sinaia 369-371, 698-701, 703, 704; h02 - Sinaia 702; h03 - Fuscaldo 405-409; h04 - Kramplje 312-314, 786; h05 - Ano Kaliniki 602, 658, 659, 675, 676; h06 - Ano Kaliniki 604; h07 - Ečka 415; h08 - Ečka 416; h09 - Ečka 417, 418, 419; Zimnicea 336, 705; Svištov 459; h10 - Svištov 748; h11 - Bartin 121; h12 - Bartin 118, 124; h13 - Bartin 116, 123; h14 - Grivac 821-824, 893; Rakovski 515, 766-768; Višegrad 898; h15 - Višegrad 896, 897, 899; h16 - Confolens 518; Rocheouart 535, 536, 538; h17 - El Berrueco 161; h18 - El Berrueco 162, 980-982; h19 - Rio Alberite 77, 80, 81; h20 and h21 - Berga and h22 - Valle Hecho. For voucher information see Appendix I.

CHAPTER 3

A COMBINATION OF TECHNIQUES PROVES USEFUL IN THE DEVELOPMENT OF NUCLEAR MARKERS IN THE NEWT GENUS *TRITURUS*

Espregueira Themudo, G.[1,2], W. Babik,[3,4] & J.W. Arntzen[1]

[1] National Museum of Natural History Naturalis, PO Box 9517, 2300 RA Leiden, the Netherlands
[2] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal
[3] Helmholtz Centre for Environmental Research - UFZ, Department of Community Ecology, Theodor-Lieser-Strasse 4, 06120 Halle (Saale), Germany
[4] Institute of Environmental Research, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland

Contents

Abstract 46
Main text 47
Acknowledgments 50
References 50
Appendix 1 53

Published in *Molecular Ecology Resources* **9** (3), 1160-1162 (2009)

Abstract

To increase the number of markers available for study of phylogeny and phylogeography in the newt genus *Triturus*, we developed and tested 59 primer pairs using three different techniques. Primers were obtained from published sources, by designing Exon-primed Intron-Crossing primers (EPIC) and from randomly cloned anonymous nuclear DNA fragments. Successful PCR products were cloned and sequenced. Five fragments were successfully amplified and sequenced for six species of *Triturus*: intron 7 of the β-fibrinogen gene (βfib*int*7), third intron of the calreticulin gene (Cal*int*C), the eleventh intron of the α-subunit of the platelet derived growth factor receptor (PDGFRα) and two anonymous markers (Cri1 and Cri4). The average percentage species divergence across all the markers is low (c. 3%), compared to what has been found in mitochondrial DNA (25-30%).

Keywords: Amphibia, anonymous markers, introns, newts, nuclear DNA markers, *Triturus*.

Main text

Mitochondrial DNA has been the prime tool used in both phylogenetic and phylogeographic studies due to its abundance in the cell, uniparental inheritance and (mostly) non-recombining nature. This translates into a relatively straightforward accessibility and availability of established analytical techniques (AVISE, 1994). In recent years, however, the wisdom of relying on only this molecule for drawing evolutionary inferences at the inter- and intra-specific levels have been repeatedly questioned (for example, BALLARD and WHITLOCK, 2004). Even if multiple mtDNA regions are studied, they do not provide independent information due to the lack of recombination at this molecule. Moreover, because individual loci in the genome have different histories, reflecting stochasticity of the coalescent process, multiple loci are essential for reconstructing historical evolutionary processes within species. (for example, FELSENSTEIN, 2006). The sampling of multiple unlinked loci will, by averaging out genealogical stochasticity, provide better estimation of population parameters which are usually the values of interest. For 'non-model' species, however, nuclear sequence markers are often unavailable. Genomic data from various genome projects provide information that can be employed for tackling problems in other species. The focus of our attention is the newt genus *Triturus*. These newts have huge genomes (ten times bigger than the human genome; GREGORY *et al.*, 2007) and using prior information is the genetic equivalent of using a magnet to find a needle in a hay-stack.

Triturus is a group of closely related species of newts (Salamandridae: Amphibia). It includes the members of the Crested newt group (*Triturus cristatus* superspecies) and the Marbled newt group (*Triturus marmoratus* species pair). The divergence of the group is estimated at a minimum of 24 Ma (million years before present; STEINFARTZ *et al.*, 2007). The level of mtDNA genetic differentiation is around 10% for the crested newts (ARNTZEN *et al.*, 2007), 5% for the marbled newts (GET and JWA, unpublished results) and 25-30% for the genus as a whole (STEINFARTZ *et al.*, 2007). We here describe the development of nuclear sequence markers for the study of the genus *Triturus* through three different strategies. Firstly, we tested published primers known to work in other salamanders, amphibians in general or fishes. Secondly, we developed Exon-primed Intron-crossing (EPIC) primers. Searching databases such as GenBank we downloaded relevant sequences and designed primers in conserved

Table 1 Primers that successfully produced PCR bands and sequences in the genus *Triturus*

Gene	Symbol	Primer	Sequence 5′–3′	N	Length (bp)	SNPs	Number of indels (length — bp)	π	H_d	H_O
β-fibrinogen intron 7	βfib*int*7	FIBX7[1]	GGAGANAACAGNACNATGACAATNCAC	25	500–514	38	1 (14)	0.022	0.960	0.56
		FIBX8[1]	ATCTNCCATTAGGNTTGGCTGCATGGC							
		BFXF[1]	CAGYACTTTYGAYAGAGACAAYGATGG							
		BFXR[1]	TTGTACCACCAKCCACCRTCTTC							
		BF CRI 1F	AAGTAGTGCTCCAGGCTTCATC							
		BF CRI 1R	GCACACTGTGTTAATCCTCCTG							
Platelet-derived growth factor receptor α	PDGFRα	PDGFRa F[2]	CGGGTCATTGAGTCCATCAGCC	24	617–662	66	9 (1, 12, 4, 8, 3, 1, 13, 2, 1)	0.031	0.978	0.50
		PDGFRa R[2]	CAGTGGGTTTTAACATTTTCACAG							
		PDGFRa Fa	GTCATTGAGTCCATCAGCCCTG							
		PDGFra 2F	AGCTGCCCTATGACTCCAGATG							
		PDGFra 2R	GCTCAAGCCATACGCTGTTCCT							
		PDGFra 2Ra	GCTCAAGCCATACGCTGTTCCTTC							
Calreticulin intron C	Cal*int*C	CalC 1F	GGMGACTCAGARTACAACATCAT	21	461–509	39	8 (1, 1–2, 2, 3, 3, 6–22, 3, 10)	0.026	0.957	0.31
		CalC 1Rb	GAATGTCYTTGTTGATCTGCATGT							
		CalC 3F	CGTTTGCGTCCAGTGTATTG							
		CalC 3R	GTCGGAGGTCCGCAGATGT							
		CalC 4R	GTCCTTGTTGATCTGCAGGTTT							
Cristatus anonymous locus 1	CRI1	CRI1 1F	ATCGCGACTGGGAGTCTTATT	24	493–511	51	5 (1, 1–2, 1, 1, 11–12)	0.025	0.976	0.50
		CRI1 1R	ATGTTCTATGCCCTCCCAGAGT							
		CRI1 1Fa	GCGACTGGGAGTCTTATTTCG							
		CRI1 1Ra	GTTCTATGCCCTCCCAGAGTGTG							
Cristatus anonymous locus 4	CRI4	CRI4 1F	AGCTCTTTGAAGACAGCATTCC	19	507–512	38	4 (2, 1, 1, 1)	0.024	0.942	0.19
		CRI4 1R	CGCTTTGTGAACTACCATACCA							
		CRI4 1Fb	CTCTTTGAAGACAGCATTCCAG							

Number of chromosomes screened (N), the length of amplified fragments, the number of single nucleotide polymorphism (SNPs), the number and length of indels in the alignment, nucleotide (p) and haplotype diversity (H_d) and observed heterozygosity (H_O).
References: [1], Sequeira *et al*. (2006); [2], Voss *et al*. (2001).

regions of adjacent exons, close to the intron-exon boundaries. And thirdly, we focussed on anonymous markers, i.e., random sequences of nuclear DNA from an unknown location in the genome. To obtain these we cloned unspecific bands co-amplified in other PCR's and then checked the sequences for single base repeats and base diversity. Note that we did not, as is more usual, construct a genomic library (JENNINGS and EDWARDS, 2005; KARL and AVISE, 1993).

Fifty-nine primer pairs were tested through PCR and sequencing (see Supplementary information). If PCRs yielded multiple bands, those of similar size were cut from the gel and purified using the Qiagen gel extraction kit (Qiagen) prior to cloning. Successful first-round PCR products were cloned with the pGEM T-Easy cloning kit (Promega). Plasmid DNA was extracted from overnight cultures of individual colonies and inserts were sequenced in both directions. The criteria to select fragments were size (>500 bp), the absence of large repeats, the presence of sufficient genetic variation as well as PCR and sequencing efficiency.

The sequences obtained were compared for similarity to sequences deposited in GenBank using the BLAST algorithm (ALTSCHUL *et al.*, 1990). Except for the anonymous markers the external fragments matched the exon regions (adjacent to exon/intron boundaries) of the respective genes whereas the anonymous markers did not show any BLAST with GenBank. Based on the new sequences, primers were then redesigned, to increase PCR efficiency and specificity and PCR products so obtained were sequenced directly. The basic PCR program consisted of four minutes at 95 ºC, followed by 35 cycles of successive denaturing (95 ºC) for 30'', annealing (57-68 ºC depending on the fragment) for 30'' and extension (72 ºC) for 90'', and a final extension (72 ºC) of three minutes. Reaction chemistry was 23 μL of H_2O, 3 μL of buffer (15 μM $MgCl_2$), 1.8 μL of 25 μM $MgCl_2$, 0.6 μL of dNTPs (10 mM), 0.2 μL of each primer (100 μM) and 0.2 μL (1 U) of Taq DNA Polymerase (Qiagen).

Sequences were obtained from intron 7 of the β-fibrinogen gene (βfib*int*7), intron C of the calreticulin gene (Cal*int*C), intron eleven of the α-subunit of the platelet derived growth factor receptor (PDGFRα*int11*) and for two anonymous markers (Cri1 and Cri4). Several sequences displayed length size polymorphisms or single nucleotide polymorphisms (SNP). To resolve length size polymorphisms we read the unphased chromatogram by eye comparing it to homozygous sequences from other individuals of the same species. For sequences with more than one SNP we used Phase v.2.1 (STEPHENS and DONNELLY, 2003; STEPHENS *et al.*, 2001) to reconstruct the

haplotypes. Sequences are available from GenBank (http://www.ncbi.nlm.nih.gov/Genbank) with accession numbers FJ526219-FJ526331. Polymorphisms for these five fragments are described in Table 1. The genes were successfully amplified for all six species of *Triturus*. Average percentage species divergence was c. 3% for the genus and c. 1% for the groups of crested newt species and marbled newts, respectively.

Acknowledgments

This work was funded by a PhD grant to GET (SFRH/BD/16894/2004) by Fundação para a Ciência e Tecnologia and by an Alexander von Humboldt Foundation grant (3-Fokoop-POL/1022634) to WB.

References

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403.

Arntzen, J. W., G. Espregueira Themudo and B. Wielstra (2007) The phylogeny of crested newts (*Triturus cristatus* superspecies): nuclear and mitochondrial genetic characters suggest a hard polytomy, in line with the paleogeography of the centre of origin. *Contributions to Zoology* **76**, 261-278.

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

Ballard, J. W. O. and M. C. Whitlock (2004) The incomplete natural history of mitochondria. *Molecular Ecology* **13**, 729-744.

Chiari, Y., M. Vences, D. R. Vieites, F. Rabemananjara, P. Bora, O. R. Ravoahangimalala and A. Meyer (2004) New evidence for parallel evolution of colour patterns in Malagasy poison frogs (<i>Mantella</i>). *Molecular Ecology* **13**, 3763-3774.

Crawford, A. (2003) Relative Rates of Nucleotide Substitution in Frogs. *Journal of Molecular Evolution* **57**, 636.

Felsenstein, J. (2006) Accuracy of Coalescent Likelihood Estimates: Do We Need More Sites, More Sequences, or More Loci? *Molecular Biology and Evolution* **23**, 691-700.

GREGORY, T. R., J. A. NICOL, H. TAMM, B. KULLMAN, K. KULLMAN, I. J. LEITCH, B. G. MURRAY, D. F. KAPRAUN, J. GREILHUBER and M. D. BENNETT (2007) Eukaryotic genome size databases. *Nucl. Acids Res.* **35**, D332-338.

HOEGG, S., M. VENCES, H. BRINKMANN and A. MEYER (2004) Phylogeny and Comparative Substitution Rates of Frogs Inferred from Sequences of Three Nuclear Genes. *Mol Biol Evol* **21**, 1188-1200.

JENNINGS, W. B. and S. V. EDWARDS (2005) Speciational history of Australian grass finches (*Poephila*) inferred from thirty gene trees. *Evolution* **59**, 2033-2047.

KARL, S. A. and J. C. AVISE (1993) PCR-based assays of Mendelian polymorphisms from anonymous single-copy nuclear DNA: techniques and applications for population genetics. *Molecular Biology and Evolution* **10**, 342-361.

PUTTA, S., J. SMITH, J. WALKER, M. RONDET, D. WEISROCK, J. MONAGHAN, A. SAMUELS, K. KUMP, D. KING, N. MANESS, B. HABERMANN, E. TANAKA, S. BRYANT, D. GARDINER, D. PARICHY and S. R. VOSS (2004) From biomedicine to natural history research: EST resources for ambystomatid salamanders. *BMC Genomics* **5**, 54.

SAN MAURO, D., D. J. GOWER, O. V. OOMMEN, M. WILKINSON and R. ZARDOYA (2004) Phylogeny of caecilian amphibians (Gymnophiona) based on complete mitochondrial genomes and nuclear RAG1. *Molecular Phylogenetics and Evolution* **33**, 413.

SEQUEIRA, F., N. FERRAND and D. HARRIS (2006) Assessing the phylogenetic signal of the nuclear β-Fibrinogen intron 7 in salamandrids (Amphibia: Salamandridae). *Amphibia-Reptilia* **27**, 409-418.

STEINFARTZ, S., S. VICARIO, J. W. ARNTZEN and A. CACCONE (2007) A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus* newts. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **308B**, 139-162.

STEPHENS, M. and P. DONNELLY (2003) A Comparison of Bayesian Methods for Haplotype Reconstruction from Population Genotype Data. *The American Journal of Human Genetics* **73**, 1162-1169.

STEPHENS, M., N. J. SMITH and P. DONNELLY (2001) A New Statistical Method for Haplotype Reconstruction from Population Data. *The American Journal of Human Genetics* **68**, 978-989.

VOSS, S. R., J. J. SMITH, D. M. GARDINER and D. M. PARICHY (2001) Conserved Vertebrate Chromosome Segments in the Large Salamander Genome. *Genetics* **158**, 735-746.

Appendix 1 - Primer pairs tested for usability in the genus *Triturus*. Published primers were retrieved from the literature, EPIC (Exon-Primed Intron-Crossing) primers were designed based on GenBank sequences of related groups and anonymous markers were based on sequences from unspecific PCR bands (see text for details). Primers in bold were selected for further testing.

Gene	Forward Primer	sequence 5' - 3'	Reverse Primer	sequence 5' - 3'	Reference
Published primers					
1H3	1H3F	GGC AAA TGC TGG TCC CAA CAC AAA	1H3R	CGA CAA CAC TGC CAA ATA CCA CAT	Putta *et al.* (2005)
Beta-fibrinogen intron 7 (βfib*int* 7)	**FIBX7**	**GGA GAN AAC AGN ACN ATG ACA ATN CAC**	**FIBX8**	**ATC TNC CAT TAG GNT TGG CTG CAT GGC**	**Sequeira *et al.* (2006)**
Beta-fibrinogen intron 7 (βfib*int* 7)	**BFXF**	**CAG YAC TTT YGA YAG AGA CAA YGA TGG**	**BFXR**	**TTG TAC CAC CAK CCA CCR TCT TC**	**Sequeira *et al.* (2006)**
Beta-fibrinogen intron 7 (βfib*int* 7)	**BF-CRI-1F**	**AAG TAG TGC TCC AGG CTT CAT C**	**BF-CRI-1R**	**GCA CAC TGT GTT AAT CCT CCT G**	**this study**
c-myc	cmyc1U	GAG GAC ATC CTG GAA RAA RTT	cmyc3L	GTC TTC CTC TTG TCR TTC TCY TC	Crawford (2003)
Collagen 1a	Col1a1-f	CAC CGA AGC CTC CCA AAA CAT CAC	Col1a1-r	GAG CCC TTC CAT CTT AGT CGT	Voss *et al.* (2001)
Homeo box a4 (HoxA4)	HoxA4-f	CTG CAG CAC TGG CAG GTC CTG CTG	HoxA4-r	TGG CGA GCG CAT CTT GGT GTT GG	Voss *et al*. (2001)
Platelet-derived growth factor receptor *alpha* (PDGFRα)	**PDGFRA f**	**CGG GTC ATT GAG TCC ATC AGC C**	**PDGFRA r**	**CAG TGG GTT TTA ACA TTT TCA CAG**	**Voss *et al.* (2001)**
Recombination activating gene 1 (Rag-1)	Mart FL1	AGC TGG AGY CAR TAY CAY AAR ATG	Amp-RAG1 R1	AAC TAC GCT GCA TTK CCA ATR TCA CA	Chiari *et al.* (2004)
Recombination activating gene 1 (Rag-1)	Amp F2	ACN GGN MGN CAR ATC TTY CAR CC	Amp-RAG1 R1	AAC TAC GCT GCA TTK CCA ATR TCA CA	Chiari *et al.* (2004)
Recombination activating gene 1 (Rag-1)	Amp-RAG1 F1	ACA GGA TAT GAT GAR AAG CTT GT	Mart R6	GTG TAG AGC CAR TGR TGY TT	Chiari *et al.* (2004)
Recombination activating gene 1 (Rag-1)	Amp-RAG1 F1	ACA GGA TAT GAT GAR AAG CTT GT	Amp R2	GGT GYT TYA ACA CAT CTT CCA TYT CRT A	Chiari *et al.* (2004)
Recombination activating gene 2 (Rag-2)	Rag-2A-F35	TGG CCN AAA MGN TCY TGY CCM ACW GG	Rag2.Lung.320R	AYC ACC CAT ATY RCT ACC AAA CC	Hoegg *et al.* (2004)
Recombination activating gene 2 (Rag-2)	Rag-2.Lung.35F	GGC CAA AGA GRT CYT GTC CNA CTG G	Rag2.Lung.320R	AYC ACC CAT ATY RCT ACC AAA CC	Hoegg *et al.* (2004)
Recombination activating gene 2 (Rag-2)	31 FN. Venk	TTY GGN CAR AAR GGN TGG CC	Lung.460R	GCA TYG RGC ATG GAC CCA RTG NCC	Chiari *et al.* (2004)
Rhodopsin	Rhod.ma	AAC GGA ACA GAA GGY CC	Rhod.md	GTA GCG AAG AAR CCT TC	Hoegg *et al.* (2004)
Rhodopsin (Rho)	Rho-f	CCA AGA GTT CTG CCA TCT ACA ATC CAG	Rho-r	CGC AGG AGA AAC CTG GCT GGA AGA CAC	Voss *et al.* (2001)
v-kit (KIT)	Kit f	TCC GTG TGG GAA TCC AGT CAC T	Kit r	AGA TGG CAT ATC TGG GAC ATA TTC	Voss *et al*. (2001)
EPIC primers					
Aldolase C (Ald) intron C	AldC2F	GGT GGA AAA CAC AGA GGA GAA C	AldC1R	CCA GAG GAA CGA CAC CTT TAT C	this study
Aldolase C (Ald) intron D	AldD1F	TTG ATA AAG GTG TCG TTC CTC TG	AldD1R	CTC ACT GAT CTT CAG CAC ACA AC	this study
Aldolase C (Ald) intron F	AldF1F	CTG ATG GAG ACC ATG ACY TGA A	AldF1R	ATG GCA ATC TCC TCA GGA CTG TA	this study
Aldolase C (Ald) intron G	AldG1F	GAC GCA CTG TAC CAC CTG CT	AldG1Ra	AGA GCA TTT GTT GAT GGC ATT C	this study
Aromatase P450	P450F	GAA ATA TTG AAC CCC ATG CAC TA	P450R	CCT GGT ATT GTT GAC GTT TCT TC	this study
Calreticulin (Cal) intron C	**CalC1F**	**GGM GAC TCA GAR TAC AAC ATC AT**	**CalC1Rb**	**GAA TGT CYT TGT TGA TCT GCA TGT**	this study
Calreticulin (Cal) intron C	**CalC3F**	**CGT TTG CGT CCA GTG TAT TG**	**CalC3R**	**GTC GGA GGT CCG CAG ATG T**	this study
Calreticulin (Cal) intron C	**CalC3F**	**CGT TTG CGT CCA GTG TAT TG**	**CalC4R**	**GTC CTT GTT GAT CTG CAG GTT T**	this study
Elongation Factor (EF) intron C	EFC1F	ACA TCA AGA AAA TCG GCT ACA AC	EFC1R	ATT TCC CTC CTT ACG GTC AAC	this study
Elongation Factor (EF) intron C	EFC1F	ACA TCA AGA AAA TCG GCT ACA AC	EFC2R	CAC TGG CAT TTC CCT CCT TMC	this study
Elongation factor alpha, intron E	EFE2F	GGT GAG TTG AGT GTT GCG TTT A	EFE3R	GAC CAG GGT GAT TCA GAA TAA TG	this study
Glyceraldehyde-3-phosphate dehydrogenase (Gapd) intron B	GapdB1F	AAG ATG AAA GTA GGA GTC AAT GG	GapdB1R	GAC TAC AGC GCG GGT CAC	this study
Glyceraldehyde-3-phosphate dehydrogenase (Gapd) intron B	GapdB1F	AAG ATG AAA GTA GGA GTC AAT GG	GapdB2R	AGT TGA CTA CAG CGC GGG TCA C	this study
Glyceraldehyde-3-phosphate dehydrogenase (Gapd) intron D	GapdD1F	CTG AGA ACG GCA AAC TTG TMA TC	GapdD1Ra	TTT GTC AAT GGT GGT GAA CAC T	this study
Growth Hormone (GH) intron C	GHC1Fa	ACA GCA TTC TGC TGC TCT GA	GHC1Ra	AGA CCG AAT GAG AGT CAA RGA GA	this study
Growth Hormone (GH) intron D	GHD1F	CTA CGA GAG GCT TAA GGA CYT GG	GHD1R	GTC TTT CTT GAA GCA GGA TAG CA	this study
Growth Hormone (GH) intron D	GHD2F	CAC ATG AGA TTC TTT CCC CAG T	GHD3R	GTT CCG TCT TCC AGT TCC TGT A	this study
Growth Hormone (GH) intron D	GHD3F	TCT CAT CAA GGT GAG TTT GAA CA	GHD3R	GTT CCG TCT TCC AGT TCC TGT A	this study
Growth Hormone (GH) intron D	GHD_int1F	CGA CAA GGA TTG TGG TTG TTG C	GHD_criR	GGC ATC TTC GTT TCT CTG GTT G	this study
Growth Hormone (GH) intron D	GHD3F	TCT CAT CAA GGT GAG TTT GAA CA	GHD_int1R	TCC CTT CAT GCA CAA AGG AGG T	this study
Growth Hormone (GH) intron D	GHD3F	TCT CAT CAA GGT GAG TTT GAA CA	GHD_CRI2R	GGA AGA GAA GGC CCC AAG AGT A	this study
Platelet-derived growth factor receptor *alpha* (PDGFRA)	PDGFRA Fa	GTC ATT GAG TCC ATC AGC CCT G	PDGFRA r	CAG TGG GTT TTA ACA TTT TCA CAG	this study
Platelet-derived growth factor receptor *alpha* (PDGFRA)	**PDGFRA 2f**	**AGC TGC CCT ATG ACT CCA GAT G**	**PDGFRA 2r**	**GCT CAA GCC ATA CGC TGT TCC T**	this study
Recombination activating gene 1 (Rag-1)	Rag1 250F	GAC ATG GAR GAC ATY ATY TTG	Rag1 1460R	ACT TAG ACT GCC TGG CAT TCA TTT	this study
Signal Recognition Particle 54 kD Protein (SRP54) intron C	SRP54C2F	GCG GAT GTG AAT ATT AAG CTT GT	SRP54C1R	GAC AAG CTC TTT GAA GAC AGC A	this study
Triosephosphate isomerase (Tpi) intron A	TpiA1F	AGT TCT TTG TCG GAG GCA AYT	TpiA1R	CAA AGT CGA TGT AGA TGG MWG GT	this study
Triosephosphate isomerase (Tpi) intron B	TpiB1F	GAG CCT TCA CTG GAG AGA TCA	TpiB1R	GAC TCT CCA AAG ACA TGC CTY CT	this study
Triosephosphate isomerase (Tpi) intron B	TpiB3F	CGC AGT TTT ACA AGC TTT GAT G	TpiB3R	CAA AGA CAT GCC TCC TCT CAG	this study
Triosephosphate isomerase (Tpi) intron B	TpiB3F	CGC AGT TTT ACA AGC TTT GAT G	TpiB4R	CTA TGG GAT AAA GCC TCA GGT G	this study
Triosephosphate isomerase (Tpi) intron B	TpiB3F	CGC AGT TTT ACA AGC TTT GAT G	TpiB5R	GTG CTA TGG GAT AAA GCC TCA G	this study
Vitelogenin	VTG F	ACC TCA ACT ACA TTC AGA CC	VTG R	GAG CTA TAT CCC AAG CAG G	C. Pinho (personal communication)
Anonymous markers					
CRI1	**CRI1 1F**	**ATC GCG ACT GGG AGT CTT ATT**	**CRI1 1R**	**ATG TTC TAT GCC CTC CCA GAG T**	this study
CRI2	CRI2 1F	GAA ATC TCT CTT CAG GGA AGC A	CRI2 1R	AAA CGG TTT GAA AGG AGT ACG A	this study
CRI2	CRI2 1F	GAA ATC TCT CTT CAG GGA AGC A	CRI2 2R	ATG AGC ATG AAG CAT TTG TCT C	this study
CRI3	CRI3 1F	CGA CTT TGA GAA AGC CTT TGA T	CRI3 1R	TCA ATT CTA TAA GCC GGG TCA G	this study
CRI3	CRI3 2F	ACT TGG TCC ACT CTG ACA CTC A	CRI3 2R	CCC AGT GGA TTG AGA GGT AGT T	this study
CRI4	**CRI4 1F**	**AGC TCT TTG AAG ACA GCA TTC C**	**CRI4 1R**	**CGC TTT GTG AAC TAC CAT ACC A**	this study
CRI4	**CRI4 1F**	**AGC TCT TTG AAG ACA GCA TTC C**	**CRI4 2R**	**CTC CAC ATC TGC TGA CAT GAT T**	this study
TVA4	TVA4 1F	ACA GTG CAA ATG CGT ACA ATT C	TVA4 1R	AGC AAG GAT CTG CTC AAG AAA C	Nadachowska & Babik (submitted)
TVA6	TVA6 1F	CTG CAT CAA ATG AGA GTC AAG C	TVA6 1R	ATC ATA TCC CCG ATT GGT GTA G	this study
TVA6	TVA6 1F	CTG CAT CAA ATG AGA GTC AAG C	TVA6 2R	GAT TGG TGT AGT CCC CAA GAA G	this study

CHAPTER 4

MULTIPLE NUCLEAR AND MITOCHONDRIAL GENES RESOLVE THE BRANCHING ORDER OF A RAPID RADIATION OF CRESTED NEWTS (*TRITURUS*, SALAMANDRIDAE)

Espregueira Themudo, G.[1,2], B. Wielstra[1,3] & J.W. Arntzen[1]

1 - National Museum of Natural History – Naturalis, PO Box 9517, 2300 RA Leiden, the Netherlands.
2 - CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal.
3 - Institute of Biology Leiden - IBL, PO Box 9516, 2300 RA Leiden, the Netherlands.

Contents

Abstract 56
Introduction 57
Material and methods 59
Sampling design 59
DNA extraction and sequencing 59
Data analysis 60
Phylogeny reconstruction 60
Molecular dating 62
Results 62
Polymorphism and level of variation 62
Phylogenetic reconstruction 65
Non-hierarchical Bayesian phylogeny 65
Phylogenetic networks 65
Hierarchical Bayesian phylogeny 67
Divergence times 67
Discussion 69
Rapid radiations, the coalescence, mitochondrial versus nuclear genes, and concatenation 69
Multiple independent markers solve the phylogeny of Triturus newts 70
Phylogenetic position and taxonomic status of T. karelinii 72
Acknowledgments 73
References 73

Published in *Molecular Phylogenetics and Evolution* **52** (2), 321-328 (2009).

Abstract

Newts of the genus *Triturus* are parapatrically distributed across Europe. Within this group, the crested newts (*Triturus cristatus* superspecies) radiated in a short temporal interval. Given the relatively short timespan in between branching events and to address the gene tree – species tree problem, we sequenced two mitochondrial and five nuclear genes from populations representing the distribution range of all the five crested newt species. We built gene trees using non-hierarchical Bayesian phylogenetics and phylogenetic networks, and a species tree with a recently developed method, which uses a hierarchical Bayesian approach. While the single gene trees did not provide resolution, the hierarchical Bayesian method yielded an almost fully resolved species tree, even though branching events followed one another closely. Results show a previously undetected basal dichotomy between *T. karelinii* and the other four species and a deep differentiation of *T. karelinii* in two lineages, here raised to full species status.

Keywords: Evolutionary radiation, introns, nuclear markers, phylogeny, phylogenetic network, polytomy, species tree, *Triturus cristatus* superspecies, *Triturus karelinii*, *Triturus arntzeni*

Introduction

The task of molecular phylogenetics is to reconstruct the order by which species have originated. However, when the different lineages split over short time spans or when the studied markers contain insufficient information, phylogenies can remain unresolved. One way forward is to collect large amounts of independent data. Independence is essential to resolve the gene tree – species tree problem, as linked *loci* (for example, from mitochondrial DNA) will have the same history. Because of stochasticity in the lineage sorting process, incomplete lineage sorting and horizontal gene transfer, gene trees generally differ from one another and from the actual species tree (which is tried to be recovered) (AVISE and WOLLENBERG, 1997). This also brings up analytical problems. Over the last decades, phylogenetic analyses have mostly followed three strategies (reviewed in HUELSENBECK *et al.*, 1996): the total evidence approach (KLUGE, 1989), the separate analysis for each marker followed by an analysis for congruence (MIYAMOTO and FITCH, 1995), and conditional data combination (DE QUEIROZ, 1993). More recently, phylogenetic methods have been proposed that incorporate biologically more realistic models of evolution, by taking into account processes like incomplete lineage sorting or horizontal gene transfer. Two new methods are phylogenetic networks (see HUSON and BRYANT, 2006) and an approach that calculates species trees explicitly, (hereafter called 'the hierarchical method', EDWARDS *et al.*, 2007).

The genus *Triturus* is composed of two groups, with two and five species, respectively (the marbled newts - *T. marmoratus* and *T. pygmaeus* - from south-western Europe and the crested newts – or *T. cristatus* superspecies - from western, central and eastern Europe). Support for a monophyletic origin is strong for either of the groups, i.e., crested newts, marbled newts and *Triturus* (ARNTZEN *et al.*, 2007; STEINFARTZ *et al.*, 2007). The five crested newt species have mutually exclusive, parapatric ranges across Europe: the Italian crested newt (*T. carnifex*), the northern crested newt (*T. cristatus*), the Dobrogean crested newt (*T. dobrogicus*), the southern crested newt (*T. karelinii* –with two subspecies: *T. k. karelinii* and *T. k. arntzeni*) and the Macedonian crested newt (*T. macedonicus*). All species meet in the Balkans which is considered to be the centre of origin of the group (CRNOBRNJA-ISAILOVIC *et al.*, 1997, see Fig. 1). Crested newts come in four morphologically distinct lineages

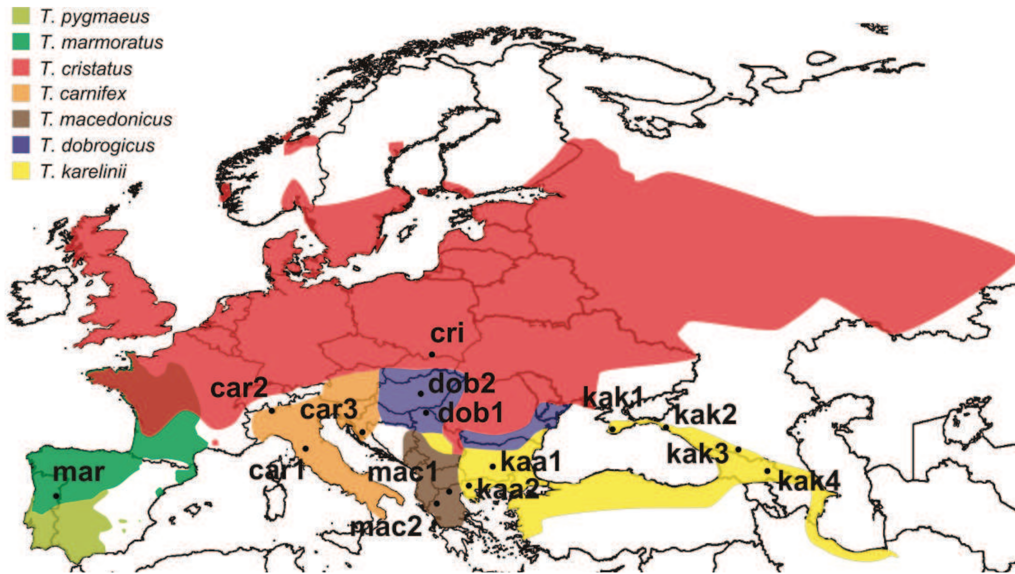


Figure 1- Geographical distribution of *Triturus* species, represented by the following colours: orange - *T. carnifex* (car), red - *T. cristatus* (cri), blue - *T. dobrogicus* (dob), yellow - *T. karelinii* (*T. k. arntzenii* - kaa and *T. k. karelinii* - kak), brown – *T. macedonicus* (mac), dark green - *T. marmoratus* (mar), light green - *T. pygmaeus*. Sample localities are as described in Appendix I.

that differ in the number of presacral vertebrae and relative appendage size (e.g., forelimb length divided by interlimb distance - a ratio known as the 'Wolterstorff-Index' and negatively correlated with the number of presacral vertebrae). Modal presacral vertebrae counts are 14 in *T. karelinii*, 15 in *T. carnifex* and *T. macedonicus*, 16 in *T. cristatus* and 17 or 18 in *T. dobrogicus*, with little intraspecific variation (ARNTZEN, 2003). *Triturus marmoratus* and *T. pygmaeus* have a modal presacral vertebrae count of 13. Note however that the evolutionary tree suggested by this morphological series, as seen from the root (*T. karelinii*, ((*T. carnifex*, *T. macedonicus*), (*T. cristatus*, *T. dobrogicus*))), is not supported by molecular genetic analyses (ARNTZEN *et al.*, 2007).

As a result of that analyses, we forwarded the hypothesis that the speciation of crested newts was near-simultaneous at around 11-10 Ma and occurred in the Balkans (ARNTZEN *et al.*, 2007). The temporal calibration was obtained from a 24 Myr old fossil record ascribed to *T. marmoratus* (ESTES, 1981; STEINFARTZ *et al.*, 2007). We found external support for this hypothesis in the observations that genetic and species diversity is highest in the Balkans and that a polytomous speciation scenario matches with the palaeogeography of the Balkan region (POPOV *et al.*, 2004). It is not expected that more precise paleogeographic reconstructions will

become available soon, so we would suggest, on the contrary, that an established phylogeny - either resolved or polytomous - might actually assist the reconstruction of the paleogeography of south-eastern Europe.

In our previous work on the phylogeny and historical biogeography of the genus *Triturus* (Caudata: Salamandridae) we failed to resolve the basal node of the crested newt radiation (ARNTZEN *et al.*, 2007). Indeed, no firm phylogenetic resolution has been found among five crested newt lineages, other than the sister-species status of *T. carnifex* and *T. macedonicus* (see below). We here apply some of the above methods with a set of seven nuclear and mitochondrial genes, in an attempt to break the crested newt polytomy.

Material and methods

Sampling design

We sampled the five species of crested newts (including both subspecies of *T. karelinii*; *T. k. karelinii* and *T. k. arntzeni*) and one species of marbled newt as outgroup, with altogether 15 individuals from as many populations (Appendix 1). Because each species of crested newt is in spatial contact with two or more other species of the group (Fig. 1) and because hybridization and introgression is known to occur in limited areas (WALLIS and ARNTZEN, 1989), we sampled localities at least 100 km away from documented contact zones (ARNTZEN, 2003). This was done to minimize confounding effects of interspecific gene flow on the phylogenetic reconstruction. Sampling effort reflected the level of intraspecific genetic variation known for the species, which is low in *T. cristatus* (one locality sampled) and high in *T. karelinii* (six localities sampled).

DNA extraction and sequencing

Tail-tip tissue samples were stored in -80 ºC and later transferred to 95% ethanol. Total genomic DNA was extracted with the Qiagen DNeasy tissue kit. PCR conditions consisted of an initial denaturation step at 94 ºC for 4 min, followed by 35 cycles of 30 s denaturation at 94 ºC, 45 s annealing at 57-68 ºC (depending on the fragment) and 90 s extension at 72 ºC and a final extension step of 4 min at 72 ºC. The molecular markers employed are two protein-coding mitochondrial genes (subunits 2 and 4 of the NADH dehydrogenase gene complex - ND2 and ND4) and

five nuclear genes, namely three introns (β-Fibrinogen intron 7 - βfib*int*7, Calreticulin intron C - *Calint*C, Platelet-derived growth factor receptor α - Pdgfrα) and two anonymous markers (Cri1 and Cri4) (ESPREGUEIRA THEMUDO *et al.*, in press). The primers used are described in Table 1. To increase accuracy of the results, both DNA strands were sequenced. Sequences were obtained commercially through Macrogen Inc. BLAST searches at Genbank confirmed the homology of the nuclear introns and the mtDNA genes with published data, while the anonymous nuclear markers remained unidentified.

Data analysis

Chromatograms were checked for basecalling errors with Chromas software (Technelysium Pty Ltd). Sequences were aligned in BioEdit (HALL, 1999). Insertion/deletion polymorphisms (indels) reflecting heterozygosity were reconstructed by comparing chromatograms for the forward and reverse primers (following FLOT *et al.*, 2006). Indels were excluded from the phylogenetic analysis. Because of difficulties in analyzing individuals which are heterozygous for more than one nuclear gene, phased haplotypes were combined into a consensus haplotype by coding polymorphic sites with the IUPAC ambiguity codes (IUPAC, 1974). The partition homogeneity test (or Incongruence length difference test, FARRIS *et al.*, 1994) as implemented in PAUP 4.0b10 (SWOFFORD, 2003) was used to check for congruence in phylogenetics signal among genes. Phylogenetic congruence was rejected for all pairs of markers, except the two mitochondrial genes. Nucleotide composition, transition/transversion ratio and sequence divergence under the Kimura 2-parameter model were calculated with MEGA4 (TAMURA *et al.*, 2007). Translation of coding sequences to amino acids did not reveal any stop codons, validating the authenticity of the mtDNA sequences. To quantify the amount of nucleotide substitution saturation and test if the data was compatible with a molecular clock model, we used DAMBE (XIA and XIE, 2001). We used the Phi test (BRUEN *et al.*, 2006) to check for intragenic recombination. For each of the genes the most appropriate model of nucleotide substitution was selected with MrModelTest (NYLANDER, 2004) as evaluated with the Akaike Information Criterion.

Phylogeny reconstruction

Phylogeny reconstruction was carried out with phylogenetic network, non-

hierarchical and hierarchical Bayesian approaches.

Under the non-hierarchical Bayesian approach, phylogenetic trees were estimated with MrBayes v.3.1. (RONQUIST and HUELSENBECK, 2003). Following the conditional combination method described in the introduction, only the two mitochondrial genes should be concatenated. However, in order to be able to compare the results over the three methods, we still estimated trees from all five nuclear genes concatenated, and from the combined nuclear and mitochondrial genes. For each analysis, the software ran for 25 million generations, with a sample taken every 1000 generation, in two parallel runs of four Metropolis Coupled Monte Carlo Markov Chains (one cold and three incrementally heated). Stabilization of the likelihood occurred well before the runs were terminated. Convergence between runs was good, with an average standard deviation of split frequencies < 0.01 (RONQUIST and HUELSENBECK, 2003). The first quarter of the trees was discarded as burn-in and the remaining trees were summarized under the majority-rule consensus criterion.

Phylogenetic networks were constructed using the Neighbor-Net algorithm as implemented in Splitstree v.4.10 (HUSON and BRYANT, 2006) for 1) each of the five nuclear genes, 2) the five nuclear genes combined, 3) the two mtDNA genes combined, and 4) all seven genes combined. The statistical confidence in inferred nodes was assessed through 1000 bootstrap pseudo-replicates.

A hierarchical Bayesian method was performed with the software BEST 2.1 (LIU and PEARL, 2007). This programme implements the search strategy of Rannala and Yang (2003), which aims to reconcile phylogenetic signal across gene trees into a species tree. Individuals were assigned to a particular species or subspecies, a requirement of the software, based on morphology and geographical origin. Analyses were conducted for only the nuclear genes, and for the complete dataset. BEST was run with topology, branch lengths, gene mutation rates, gamma shape parameter, character state frequencies, transition/transversion ratio, substitution rates under the GTR model, and the proportion of invariable sites set to be unlinked across genes. The chain length used was 300 million generations, with one sample taken every 1000 generations. The last 10,000 trees sampled from the stationary phase were summarized under the 50% majority rule consensus criterion.

In this paper, we will use the term 'species trees' to define a phylogenetic

tree calculated from a method that acknowledges heterogeneity in gene trees (EDWARDS, 2009), whereas 'concatenation' is used to define the construction of a supermatrix containing all available data.

Molecular dating

Molecular clock estimates for salamandrids are available for the ND2 gene, but vary considerably. They are contained between the extremes 0.37 (MUELLER, 2006) and 0.64 (WEISROCK *et al.*, 2005) substitutions/site/100 Ma. Several other papers have used fossil calibrations and/or biogeographical data to estimate divergence times in the family Salamandridae. Steinfartz et al (2007) used a fossil calibration of 24 million years before present (Ma) at the root of *Triturus* (their calibration point C4). Zhang et al (2008) estimated the divergence of Salamandridae based on two sets of calibration: one with similar fossil calibrations as Steinfartz and another with an additional biogeographical calibration point for the split between Corsico-Sardinian *Euproctus* and continental *Triturus*. These two sets of calibrations gave an estimate for the root of *Triturus* of 26.6 and 11.3 (ZHANG *et al.*, 2008). We primarily followed the calibration point of 24 Ma of Steinfartz to date divergences on the species tree, and then compared the results to the other possible calibrations. The divergence times of the various lineages were determined from the relative branch lengths in the consensus tree by linear interpolation.

Results

Polymorphism and level of variation

The aligned sequence matrix (with indels removed) has a length of 4336 bp, with 694 variable positions of which 373 are parsimony-informative (Table 2). With 288 (77%) versus 85 (23%) variable positions, phylogenetic information content is higher in the two mitochondrial genes than in the five nuclear genes. The G-content was higher in nuclear genes (17.4 – 26.3 %) than in mitochondrial genes (13.3 – 16.4 %), as expected from the anti-G bias in mtDNA (ZHANG and HEWITT, 1996). Cri4 showed the lowest level of variation and ND2 and ND4 the highest. The level of nucleotide substitution saturation was classified as 'little' for each gene. No

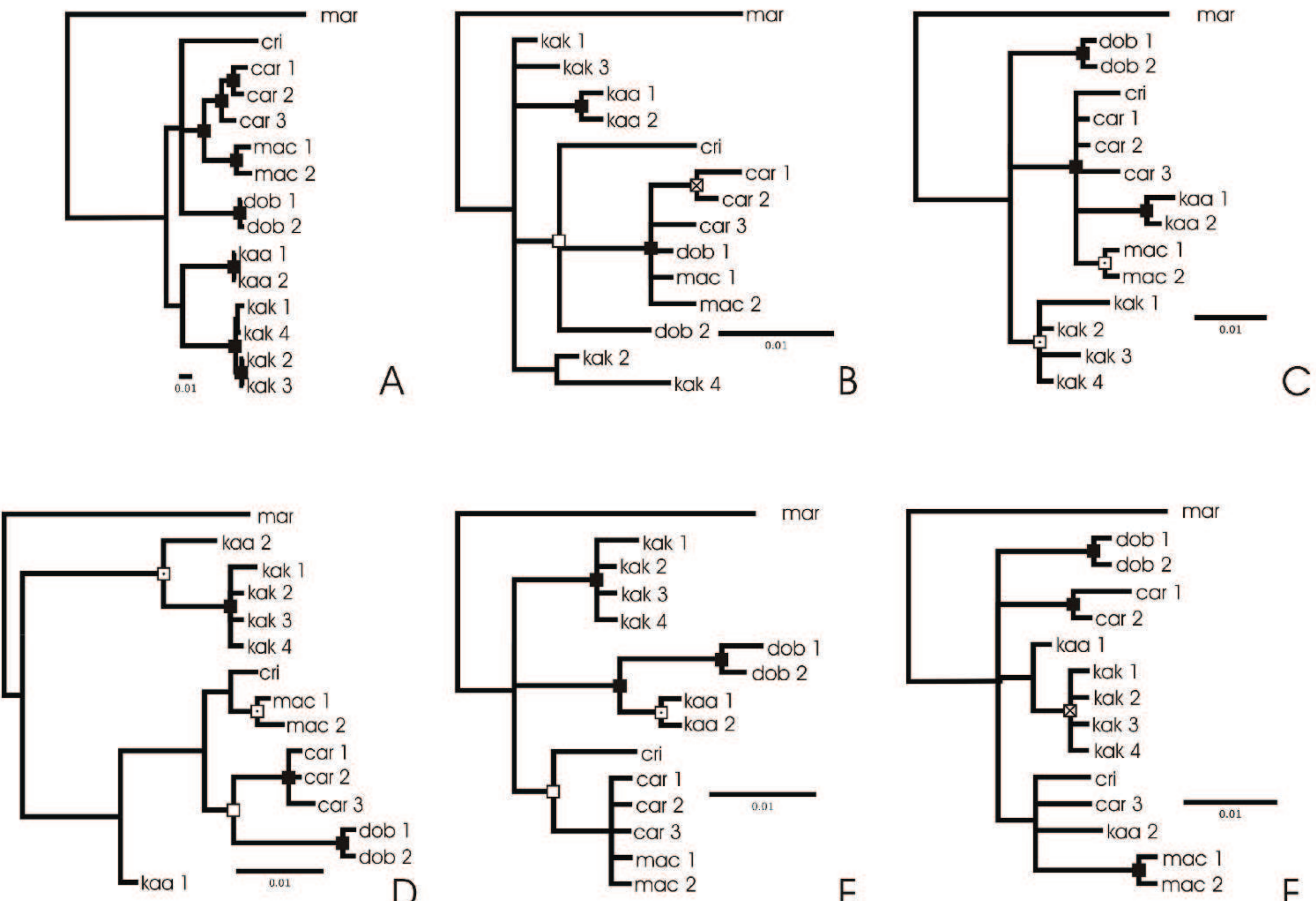


Figure 2 Results of a non-hierarchical Bayesian phylogenetic analysis for the genus *Triturus* from DNA sequence data with *T. marmoratus* as outgroup. The genes studied are of either mitochondrial origin, NADH dehydrogenase subunits 2 and 4 (A) or nuclear origin, β-fibrinogen intron 7 (B), calreticulin intron C (C), platelet-derived growth factor receptor α (D), cristatus anonymous locus 1 (E) and cristatus anonymous locus 4 (F). Posterior probabilities are classified as 1.0 (solid squares), no less than 0.99 (crossed squares), no less then 0.95 (dotted squares) and no less then 0.90 (open squares). Scale bar is in expected changes per site.

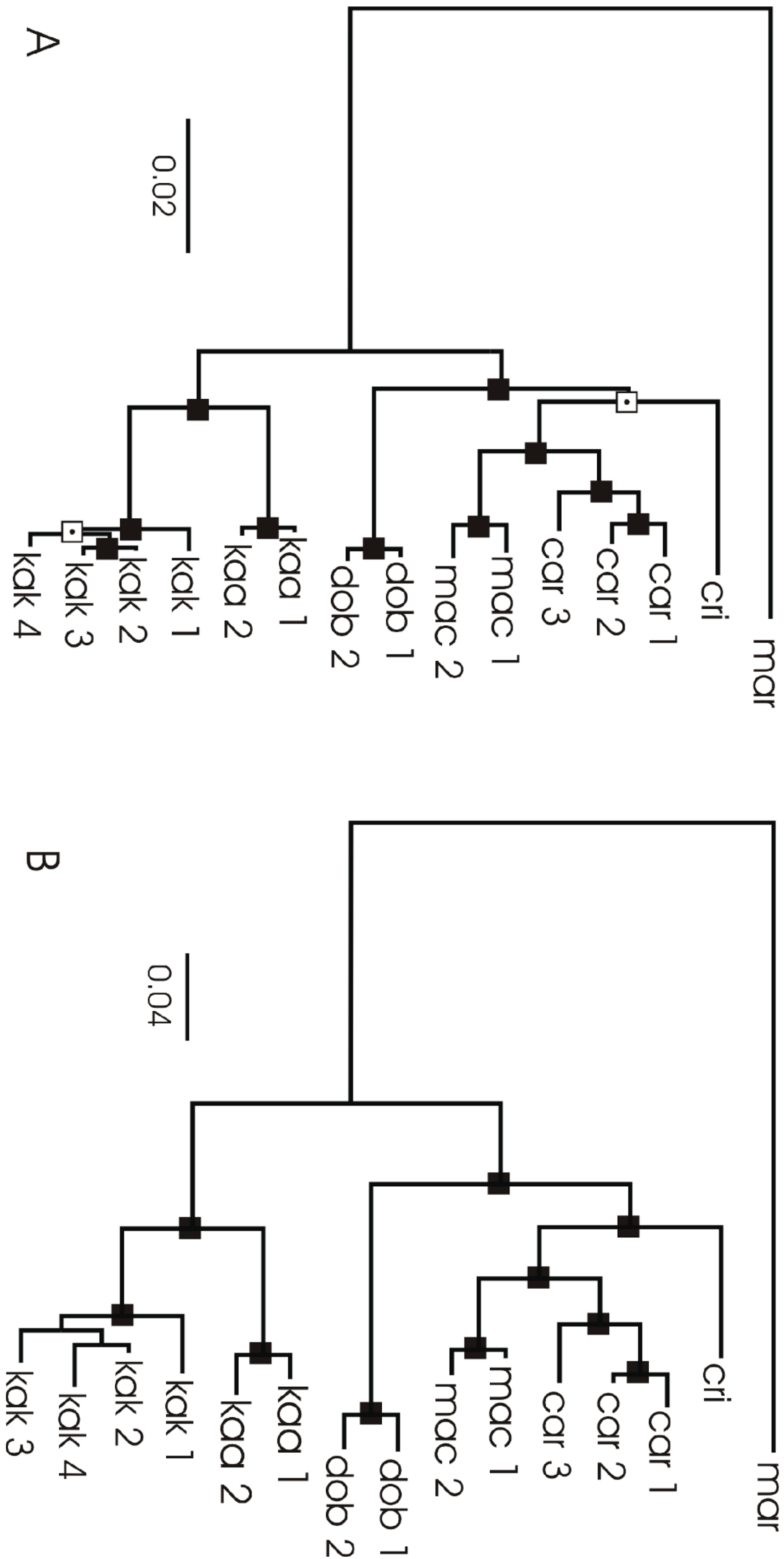


Figure 3 Results of a non-hierarchical Bayesian phylogenetic analysis of *Triturus* with a Total Evidence approach (A) and with a supermatrix of the nuclear genes (B). For details, see text. Scale bar is in expected changes per site.

significant signal for recombination was detected. The selected substitution models were F81+I for βfib*int*7, HKY+G for Cal*int*C, PDGFRα, Cri1 and Cri4, and a mixed model for the combined mtDNA genes: GTR+I+G for ND2 and GTR+I for ND4. Across nuclear genes heterozygosity ranged from 13 to 47%.

Phylogenetic reconstruction

Non-hierarchical Bayesian phylogeny

In the mtDNA gene tree all individuals are grouped into their nominal species and subspecies, with a Bayesian posterior probability (pp) of 1.0 (Fig. 2A). The mtDNA data gives strong support to the sister-group relationship of *T. carnifex* and *T. macedonicus* (pp=1.0) and weak to the two subspecies of *T. karelinii* (pp=0.60). *Triturus karelinii* is the sister species to all other crested newts, but the support for this configuration is weak (pp=0.78).

Across the nuclear genes, the results are not easy to reconcile, neither at the level of recognized species and subspecies, nor in terms of phylogenetic structure (Fig. 2 B-F). Nodes with high posterior probability (pp≥0.95) that also support nominal (sub)species are found for *T. carnifex* (Pdgfrα and Cri4), *T. macedonicus* (*Calint*C, Pdgfrα and Cri4), *T. k. arntzeni* (βfib*int*7, *Calint*C and Cri1) and *T. dobrogicus* and *T. k. karelinii* (all genes except βfib*int*7). Inter(sub)specific phylogenetic relationships supported with pp≥0.95 are *T. dobrogicus* and *T. k. arntzeni* by Cri1 (with exclusion of *T. k. karelinii*) and *T. carnifex*, *T. cristatus*, *T. macedonicus* and *T. k. arntzeni* by Cal*int*C (with exclusion of *T. k. karelinii*).

The combined nuclear tree (Fig. 3B) shows all nodes resolved with a pp of 1.0. *Triturus carnifex* and *T. macedonicus* are sister species. These two species are sister to *T. cristatus*, and *Triturus dobrogicus* is sister to this group. *Triturus. k. karelinii* and *T. k. arntzenii* are sister taxa. The tree resulting from the complete dataset differs from the one in Fig. 3B only by a lower support (pp=0.95) for the ((*T. carnifex* + *T. macedonicus*) + *T. cristatus*) clade (Fig. 3A).

Phylogenetic networks

Networks for the five nuclear genes separately show mixed results, with the occasional high support for the predefined taxa, such as *T. carnifex*, *T. dobrogicus* and *T. macedonicus* by Pdgfrα (results not shown). Conversely, individuals representing a nominal species might also be found scattered over the

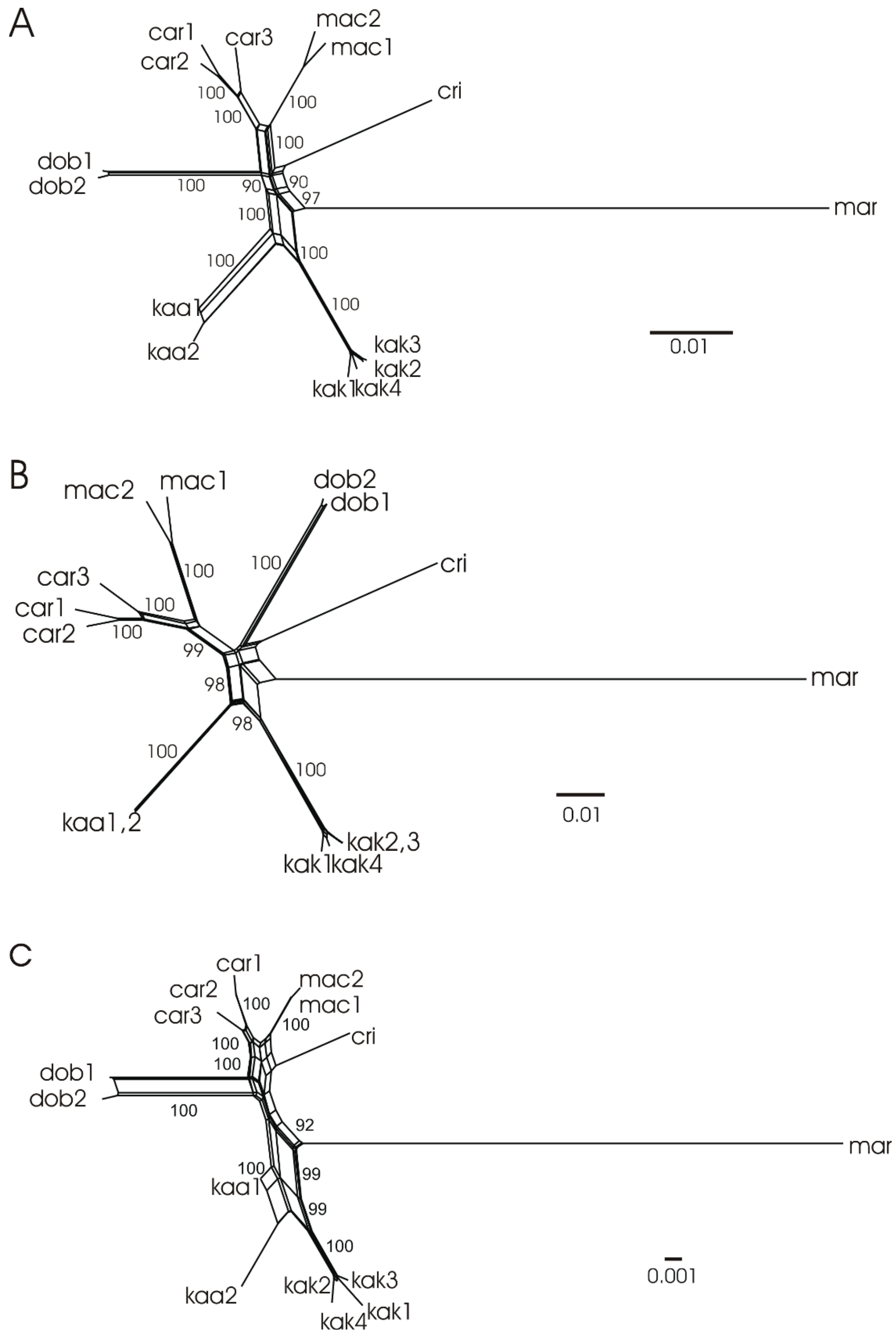


Figure 4 Results of a phylogenetic network analysis for the genus *Triturus* from DNA sequence data with *T. marmoratus* as outgroup. Results are shown for the concatenation of five nuclear genes (A), two mitochondrial genes (B) and all seven genes (c). Values on branches are bootstrap replication scores. Phylogenetic networks do not require congruence between fragments and conflict in the dataset is represented graphically through closed networks (or 'rings'), indicating that interspecific relationships are not fully resolved (Huson and Bryant, 2006).

network, such as *T. carnifex* for Cri4, *T. dobrogicus* for βfib*int*7 and *T. karelinii* for Cal*int*C, Cri1 and Cri4. The networks show generally low phylogenetic structure at the (sub)species level, with the exception of the gene PDGFRα, which supports a clade consisting of *T. carnifex*, *T. cristatus*, *T. dobrogicus* and *T. macedonicus* with a low bootstrap replication scores (brs) = 0.65 (results not shown). The phylogenetic network derived from the combined dataset resolves five species of crested newts in line with current taxonomy (Fig. 4A). Substructuring is found for *T. karelinii* with two groups that concur with current subspecies designations (*T. k. karelinii* and *T. k. arntzeni*). Bootstrap replication scores that would support phylogenetic structure among these taxa vary. *Triturus karelinii* is either monophyletic and differentiated from the other taxa (brs = 1.0), or paraphyletic (brs = 0.97) with just *T. k. karelinii* differentiated from all the other taxa. Moderate support is found for a clade consisting of *T. carnifex*, *T. cristatus*, *T. macedonicus* and *T. dobrogicus* (brs = 0.90). Within this clade, only the sister-group relationship of *T. carnifex* and *T. macedonicus* is strongly supported (brs = 1.0). A network based on the two mtDNA genes has the same topology but differs slightly in bootstrap replication scores (Fig. 4B). Based on the five nuclear genes, the network has the same branching order, with high support for the (sub)species, but low for interspecific relationships (Fig. 4C). The overall distances among taxa are shorter, due to the lower number of informative characters.

Hierarchical Bayesian phylogeny

The species tree is fully resolved (pp > 0.99), except for the position of *T. cristatus* (Fig. 5A). The supported clades are i) *T. carnifex* and *T. macedonicus*, ii) group i and *T. dobrogicus*, and iii) *T. k. karelinii* and *T. k. arntzeni*. When only the nuclear genes are used to build the species tree, the topology is identical, but pp values are lower (Fig. 5B).

Divergence times

The molecular clock test in DAMBE does not reject the null hypothesis of equidistance of all tips to the root of the tree. Linear interpolation on the species tree (Fig. 5A) over the 24 Myr window that covers the radiation of the genus *Triturus*, based on the fossil calibration (Fig. 5A) yields the following estimates of divergence times within the crested newt group: 11.2 Ma for the MRCA of the crested newts

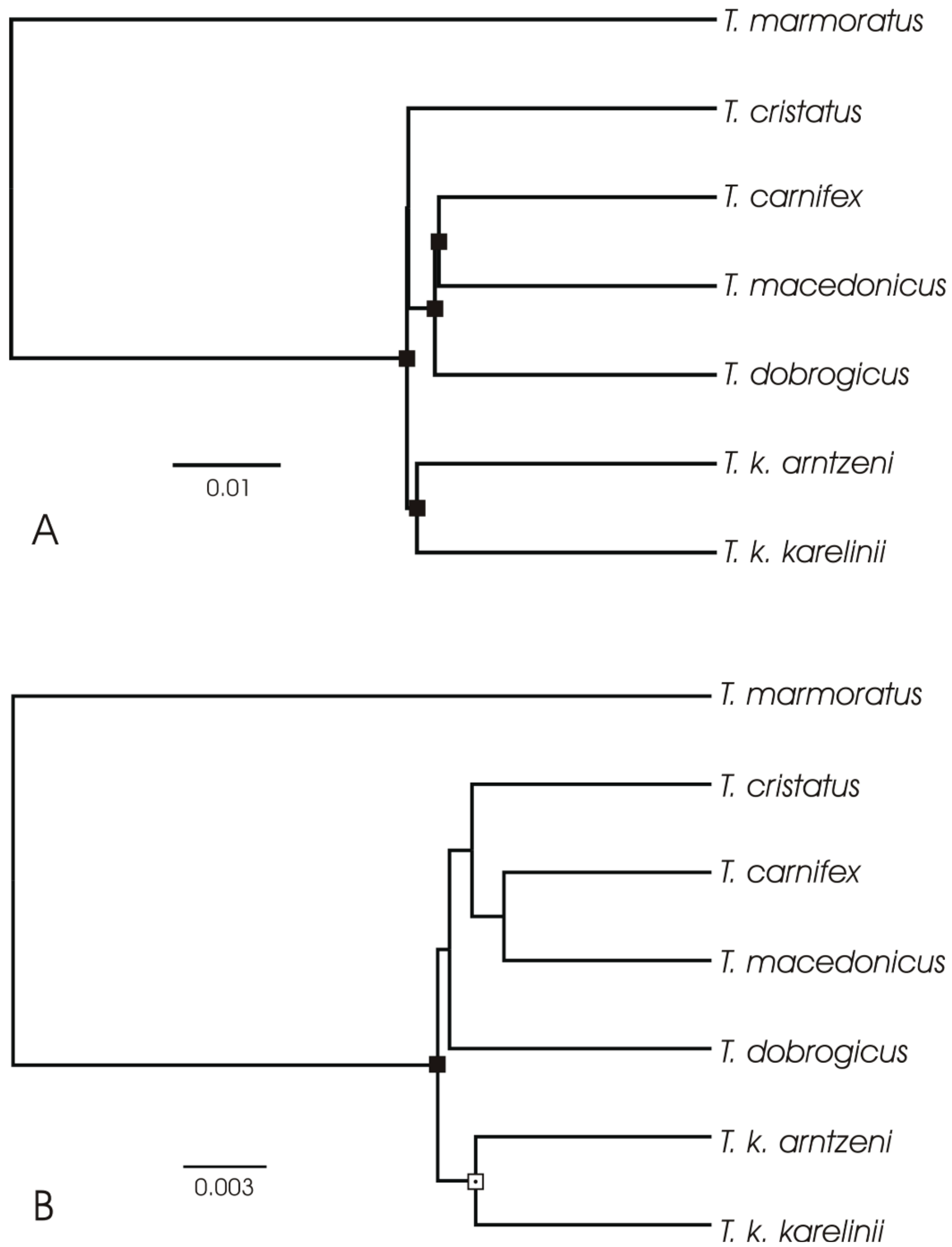


Figure 5 Results of a hierarchical Bayesian phylogenetic analysis for the genus *Triturus*, based upon DNA sequence data from two mitochondrial and five nuclear genes with *T. marmoratus* as outgroup (A); and an analysis with only the five nuclear genes (B). For details, see text. Posterior probabilities are coded as in Fig. 2. Scale bar is in expected changes per site.

(and, by consequence, for the branching off of *T. cristatus*, which stems from the root), 11.1 Ma for the divergence between *T. k. karelinii* and *T. k. arntzeni*, 8.0 Ma for the origin of *T. dobrogicus* and 7.9 Ma for the split between *T. carnifex* and *T. macedonicus*. As molecular clock calibrations for ND2 vary considerably, the time to the MRCA of crested newts is estimated at 14 – 28 Ma under these calibrations, which would multiply each of the above temporal estimates by a factor 1.25-2.5.

Discussion

Rapid radiations, the coalescence, mitochondrial versus nuclear genes, and concatenation

The evolutionary history of rapid radiations is difficult to resolve. During the short timespan between branching events, there is limited opportunity for phylogenetic signal to accumulate. Furthermore, the older the radiation, the more difficult it will be to uncover the order of branching events, as the relative amount of informative phylogenetic change decreases through time (i.e. anagenesis masks cladogenesis). In order to resolve such apparent polytomies, large amounts of data will generally be required.

Coalescence is the merging of genetic lineages backwards in time to a most recent common ancestor. Forward in time, the sorting process taking place during diversification from the ancestral lineages to the pattern observed today is a chance process. Because of this stochasticity, and because lineage sorting may be incomplete, gene trees will generally be different from one another and frequently not represent the species tree (AVISE and WOLLENBERG, 1997). The only way out of this dilemma is to study multiple genes because the more genes involved, the more convincing the convergence towards the species tree will be. By studying many genes together, the ‘gene tree versus species tree’ problem should ultimately be eliminated.

Animal molecular phylogenetics, in former times, largely relied on just the mitochondrial genome, with two main drawbacks. First, because of its small size (c. 16 kb in vertebrates) it is quickly exhausted in terms of data yield compared to the nuclear genome (for comparison Triturus newts have a nuclear genome size of c. 25 Gb, GREGORY, 2008, even though a large portion is repetitive DNA and phylogenetically not-informative). Second, the mitochondrial genes are linked. The

molecule is inherited as a single unit and mitochondrial gene trees are, hence, not independent. Indeed, recent studies on the phylogeny of the Salamandridae provide similar results whether based upon a small number of representative genes (STEINFARTZ *et al.*, 2007) or on the complete mitochondrial genome (ZHANG *et al.*, 2008). In practice, this means that nuclear genes have to be used to get independent estimations. Unfortunately, primers for nuclear genes will often have to be newly developed (for Triturus see CHAPTER 3).

Multiple independent markers solve the phylogeny of Triturus *newts*

The non-hierarchical Bayesian phylogenetic reconstruction yielded five nuclear gene trees and one mitochondrial tree that were all different from each other, to the extent that and a consensus tree would not be informative. A striking result is that individual sequences do group according to predefined species and subspecies in the mtDNA-based tree, while for nuclear genes this is rarely the case. Possible explanations are i) better phylogenetic resolution due to a larger number of informative characters and ii) lineage sorting has progressed more rapidly, due to a fourfold smaller effective population size (MOORE, 1995). Either way, the observed species level consistency in the individual phylogenetic placement suggests the absence or low frequency of horizontal gene exchange, which we had aimed for through a thoughtful sampling scheme, relying on the well documented species ranges (ARNTZEN, 2003).

The total evidence tree is completely resolved, and is identical to the combined nuclear tree, except for the lower statistical support for the position of *T. cristatus* (Fig. 3B). Compared to the total evidence tree, the mtDNA tree has the basal branches unresolved (Fig. 2A). This means that the nuclear data contributes significantly to the resolution of the species tree, even though this is not obvious from a gene by gene perspective.

The phylogenetic relationships supported by our 1747 bp of mtDNA data are the sister-group status of *T. carnifex* and *T. macedonicus* and *T. k. karelinii* and *T. k. arntzeni* (with pp > 0.99), while moderate support was for the clade including all crested newts that are not *T. karelinii* (pp=0.78). The basal dichotomy between *T. karelinii* versus non-*T. karelinii* was not found by earlier, smaller molecular datasets (e.g. WALLIS and ARNTZEN, 1989; MACGREGOR *et al.*, 1990). Based upon 839 bp of cytochrome-b sequence data, Steinfartz et al. (2007) described the paraphyly of *T.*

karelinii, whereas Arntzen et al. (2007) found no resolution at the basal *Triturus* node, from either 40 allozyme loci or 642 bp of ND4 mtDNA. The two *T. karelinii* subspecies (*arntzeni* and *karelinii*) are almost as old as the *T. karelinii* lineage itself.

In the phylogenetic network analysis, pronounced ambiguities in the total evidence approach are the unresolved branching order (*T. dobrogicus*, *T. cristatus,* (*T. carnifex*, *T. macedonicus*) and the position of the *T. karelinii* subspecies as either monophyletic or paraphyletic (Fig. 4C). The next highest uncertainty refers to the monophyly of the group *T. carnifex*, *T. cristatus*, *T. dobrogicus* and *T. macedonicus* and, following that, the monophyly of *T. carnifex* and *T. macedonicus*.

In contrast with the above results, the phylogenetic tree obtained with the hierarchical Bayesian method is fully resolved, except for the position of *T. cristatus* (Fig. 5A). The newly obtained phylogenetic resolution supports the notion that the radiation of crested newts lineages occurred over a relatively short period. This window of radiation, dated at 11-8 Ma (or 14-10 Ma to 28-20 Ma using the varying ND2 molecular clocks cited above), yielded all six different lineages that we observe today.

This study provides empirical support for the claim (EDWARDS *et al.*, 2007; EDWARDS, 2009) that the method, by accounting for the differential coalescent process across independent markers, estimates the phylogenetic relationships between recently diverged taxa with high precision. Statistical support expressed by posterior probability values is high. This is a remarkable improvement over competing methods (Fig. 2 and 4). We only found a fully bifurcating and supported tree with the non-hierarchical method by concatenating incongruent markers, thus violating the prerequisites of conditional combination. A potential drawback of the hierarchical method is that additional information should be provided in the form of species identification of the individuals included. Crested newt species show distinct morphological characteristics, making species assignment straightforward. Furthermore, our sampling scheme was designed to exclude areas if interspecific gene flow. Indeed, none was detected in the mtDNA marker (Fig. 2A). Despite this, distinguishing between interspecific gene flow and incomplete lineage sorting is a complex issue (AVISE *et al.*, 1983; SANG and ZHONG, 2000; see also HEY and NIELSEN, 2004). We are currently exploring this area of research in *Triturus* newts, by sampling from areas potentially subject to introgression, i.e., within 100 km of

described range borders (G. Espregueira Themudo, A. Bickham Baird & J. W. Arntzen, in prep.).

As predicted by Edwards (2009), statistical support in the species tree from the hierarchical method is less than the with concatenation. The theoretical advantages of the hierarchical method over data concatenation, however, make us believe that the outcome of the hierarchical Bayesian method is more reliable.

Concatenation is also prone to "data swamping", where one or a few of the partitions provides most of the information retained in the phylogenetic tree. This does not seem to be problematic here, as we have shown that the nuclear genes are contributing significantly to the resolution of the total evidence tree (Fig. 3A and B). In the hierarchical method, the effect of the number of informative sites per locus is still not clear (EDWARDS, 2009).

Phylogenetic position and taxonomic status of T. karelinii

Marked levels of genetic differentiation within *T. karelinii* were first apparent from mtDNA, as either monophyletic (WALLIS and ARNTZEN, 1989) or paraphyletic (STEINFARTZ *et al.*, 2007). Interestingly, the phylogenetic network gives high bootstrap support to both the monophyly and the paraphyly of *T. karelinii* (Fig. 4A). This type of uncertainty, expressed in so called phylogenetic 'spider-webs', may be indicative of reticulate evolution, c.q. hybridization (HUSON and BRYANT, 2006).

Based on differences in genome size, protein variation and morphological characteristics, Litvinchuk et al. (1999) described a Balkan population as the representative of a distinct subspecies, *Triturus karelinii arntzeni*. The present study suggest that the two taxa are genetically distinct, at the same level or deeper than other *Triturus* species are to each other, with a date of origin close to the basal radiation of the crested newt group (Fig. 4). Considering their long independent evolutionary history, we suggest that the two taxa should be raised to full species status, i.e., *T. karelinii* (Strauch, 1870) and *T. arntzeni* Litvinchuk, Borkin, Džukić and Kalezić, 1999 (in LITVINCHUK *et al.*, 1999).

Triturus arntzeni and *T. karelinii* may have their range limits at either side of the Bosporus (RAFFAËLLI, 2007), but geological information does not support this view, since the current connection of the Black Sea to the Mediterranean opened a mere 5000 years ago (RYAN *et al.*, 2003). Other data suggest that the species range border could be either in Europe or in Asia. A population from Istanbul (at the

European side of the Bosporus) is classified as *T. karelinii* on the basis of genome size, that is larger in *T. karelinii* than in *T. arntzeni* (LITVINCHUK *et al.*, 1999), whereas mtDNA data suggest that the ranges are in contact around Adapazari in northwestern Anatolia (WALLIS and ARNTZEN, 1989). This region is a highly active tectonic area, with major fault lines going through the Sakarya river valley, called the Adapazari – Karasu Fault, that developed since the Miocene (ELMAS, 2003).A wider analysis on the spatial distribution of morphological and genetic variation in the *T. karelinii* group is underway (K. Olgun et al., in prep.; B. Wielstra et al., in prep.)

Acknowledgments

This work was funded by a PhD grant (Ref. SFRH/BD/16894/2004) to GET from FCT (Fundação para a Ciência e a Tecnologia, Portugal). We thank W. Babik, Ö. Güçlü, K. Olgun, N. Poyarkov and G. Wallis for help in collecting material, and A. Bickham Baird, A. Larson, G. Wallis and two anonymous reviewers for useful comments to earlier versions of this paper.

References

ARNTZEN, J.W. (2003) *Triturus cristatus* Superspecies–Kammolch-Artenkreis. In: *Handbuch der Reptilien und Amphibien Europas. Schwanzlurche IIA* (eds. GROSSENBACHER, K., THIESMEIER B.), pp. 421-514. Aula-Verlag, Wiebelsheim.

ARNTZEN, J.W., ESPREGUEIRA THEMUDO G. and WIELSTRA B. (2007) The phylogeny of crested newts (*Triturus cristatus* superspecies): nuclear and mitochondrial genetic characters suggest a hard polytomy, in line with the paleogeography of the centre of origin. *Contributions to Zoology* **76**, 261-278.

AVISE, J.C., SHAPIRA J.F., DANIEL S.W., AQUADRO C.F. and LANSMAN R.A. (1983) Mitochondrial DNA differentiation during the speciation process in *Peromyscus*. *Molecular Biology and Evolution* **1**, 38-56.

AVISE, J.C. and WOLLENBERG K. (1997) Phylogenetics and the origin of species. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 7748-7755.

BININDA-EMONDS, O.R.P., CARDILLO M., JONES K.E., *et al.* (2007) The delayed rise of present-day mammals. *Nature* **446**, 507-512.

BRUEN, T.C., PHILIPPE H. and BRYANT D. (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**, 2665-2681.

CRNOBRNJA-ISAILOVIC, J., DZUKIC G., KRSTIC N. and KALEZIC M.L. (1997) Evolutionary and paleogeographical effects on the distribution of the *Triturus cristatus* superspecies in the central Balkans. *Amphibia-Reptilia* **18**, 321.

DE QUEIROZ, A. (1993) For consensus (sometimes). *Systematic Biology* **42**, 368-372.

EDWARDS, S.V., LIU L. and PEARL D.K. (2007) High-resolution species trees without concatenation. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 5936-5941.

EDWARDS, S.V. (2009) Is a new and general theory of molecular systematics emerging? *Evolution* **63**, 1-19.

ELMAS, A. (2003) Late Cenozoic tectonics and stratigraphy of northwestern Anatolia: the effects of the North Anatolian Fault to the region. *International Journal of Earth Sciences* **92**, 380.

ESPREGUEIRA THEMUDO, G., BABIK W. and ARNTZEN J.W. (in press) A combination of techniques proves useful in the development of nuclear markers in the newt genus *Triturus*. *Molecular Ecology Notes*.

ESTES, R. (1981) Gymnophiona, Caudata. In: *Handbuch der Palaoherpetologie, Part 2* (ed. WELLNHOFER, P.), pp. 1-15. Gustav Fischer, Stuttgart.

FARRIS, J.S., KALLERSJO M., KLUGE A.G. and BULT C. (1994) Testing significance of incongruence. *Cladistics* **10**, 315-319.

FLOT, J.-F., TILLIER A., SAMADI S. and TILLIER S. (2006) Phase determination from direct sequencing of length-variable DNA regions. *Molecular Ecology Notes* **6**, 627-630.

GREGORY, T.R. (2008) Animal Genome Size Database. http://www.genomesize.com.

HALL, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95-98.

HEY, J. and NIELSEN R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747-760.

HUELSENBECK, J.P., BULL J.J. and CUNNINGHAM C.W. (1996) Combining data in

phylogenetic analysis. *Trends in Ecology & Evolution* **11**, 152-158.

HUSON, D.H. and BRYANT D. (2006) Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**, 254-267.

IUPAC (1974) Abbreviations and symbols for nucleic acids, polynucleotides and their constituents (Rules approved 1974). *Pure and Applied Chemistry* **40**, 277-331.

KLUGE, A.G. (1989) A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Systematic Zoology* **38**, 7-25.

LITVINCHUK, S.N., BORKIN L.J., DŽUKIĆ G.*, et al.* (1999) Taxonomic status of *Triturus karelinii* on the Balkans, with some comments about other crested newt taxa. *Russian Journal of Herpetology* **6**, 153-163.

LIU, L. and PEARL D.K. (2007) Species trees from gene trees: reconstructing bayesian posterior distributions of a species phylogeny using estimated gene tree distributions. *Systematic Biology* **56**, 504-514.

MACGREGOR, H.C., SESSIONS S.K. and ARNTZEN J.W. (1990) An integrative analysis of phylogenetic relationships among newts of the genus *Triturus* (family Salamandridae), using comparative biochemistry, cytogenetics and reproductive interactions. *Journal of Evolutionary Biology* **3**, 329-373.

MIYAMOTO, M.M. and FITCH W.M. (1995) Testing species phylogenies and phylogenetic methods with congruence. *Systematic Biology* **44**, 64-76.

MOORE, W.S. (1995) Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49**, 718-718.

MUELLER, R.L. (2006) Evolutionary rates, divergence dates, and the performance of mitochondrial genes in bayesian phylogenetic analysis. *Systematic Biology* **55**, 289-300.

NYLANDER, J.A. (2004) MrModelTest2. Computer program distributed by the author. Evolutionary Biology Centre, Uppsala University, Sweden.

POPOV, S.V., RÖGL F., ROZANOV A.Y.*, et al.* (2004) Lithological-paleogeographic maps of Paratethys: 10 maps Late Eocene to Pliocene. *Courier Forschunginstitut Senckenberg* **250**, 1-46.

RAFFAËLLI, J. (2007) *Les Urodèles du monde* Penclen édition.

RANNALA, B. and YANG Z. (2003) Bayes estimation of species divergence times and

ancestral population sizes using DNA sequences from multiple loci. *Genetics* **164**, 1645-1656.

RONQUIST, F. and HUELSENBECK J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.

RYAN, W.B.F., MAJOR C.O., LERICOLAIS G. and GOLDSTEIN S.L. (2003) Catastrophic Flooding of the Black Sea. *Annual Review of Earth and Planetary Sciences* **31**, 525-554.

SANG, T. and ZHONG Y. (2000) Testing Hybridization Hypotheses Based on Incongruent Gene Trees. *Systematic Biology* **49**, 422 - 434.

STEINFARTZ, S., VICARIO S., ARNTZEN J.W. and CACCONE A. (2007) A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus* newts. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **308B**, 139-162.

SWOFFORD, D.L. (2003) PAUP*: Phylogenetic Analysis Using Parsimony (and other methods) 4.0 Beta 10. Sunderland (MA): Sinauer Associates.

TAMURA, K., DUDLEY J., NEI M. and KUMAR S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* **24**, 1596.

WALLIS, G.P. and ARNTZEN J.W. (1989) Mithocondrial-DNA variation in the crested newt superspecies: limited cytoplasmic gene flow among species. *Evolution* **43**, 88-104.

WEISROCK, D.W., HARMON L.J. and LARSON A. (2005) Resolving deep phylogenetic relationships in salamanders: analyses of mitochondrial and nuclear genomic data. *Systematic Biology* **54**, 758-777.

XIA, X. and XIE Z. (2001) DAMBE: Software Package for Data Analysis in Molecular Biology and Evolution. *Journal of Heredity* **92**, 371-373.

ZHANG, D.X. and HEWITT G.M. (1996) Nuclear integrations: challenges for mitochondrial DNA markers. *Trends in Ecology & Evolution* **11**, 247-251.

ZHANG, P., PAPENFUSS T.J., WAKE M.H., QU L. and WAKE D.B. (2008) Phylogeny and biogeography of the family Salamandridae (Amphibia: Caudata) inferred from complete mitochondrial genomes. *Molecular Phylogenetics and Evolution* **49**, 586–597.

CHAPTER 5

Geography helps in the distinction between gene flow and incomplete lineage sorting in the crested newt phylogeny

Espregueira Themudo, G.[1, 2] and J.W. Arntzen[1]

[1] National Museum of Natural History – Naturalis, PO Box 9517, 2300 RA Leiden, the Netherlands.

[2] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal.

Contents

Abstract 80
Introduction 81
Material and methods 83
    Sampling 83
    DNA extraction and sequencing 83
    Sequence analysis 85
    Distinguishing incomplete lineage sorting and current gene flow 85
Results 88
    DNA sequence data 88
    Linkage disequilibrium 89
    Phylogenetic analysis 89
Discussion 89
    Hybridization 90
Acknowledgments 92
References 92

Unpublished manuscript

Abstract

In molecular phylogenetics, gene flow and incomplete lineage sorting are difficult to distinguish and can bias reconstructions of species history. In species with limited dispersal capability, gene flow will be limited to the vicinity of species border, whereas incomplete lineage sorting will be randomly present across the complete range. Therefore, the geographical location of samples should help distinguish the two effects.

The crested newts have multiple, parapatric, well delimited contact zones in the Balkans and are known to hybridize. We sequenced five nuclear and two mitochondrial genes from 31 crested and 2 marbled newts, from populations close to ('parapatric') and away from ('allopatric') contact zones in order to test if it was possible to distinguish the two phenomena.

We first tested for linkage disequilibrium in 'allopatric' and 'parapatric' populations. Then, we used a known *Triturus* phylogeny as a guide tree to which we allocated the individuals studied according to their morphological identification and / or documented range. After that, we checked if the likelihood of this tree would increase if we moved - one allele at the time - from the species in which it was observed to another species.

Results indicate higher linkage disequilibrium in 'parapatric' populations, although the number of tree "improvements" is not significantly higher there according to a G-test of independence. This seems to indicate that both processes are present to some extent in our dataset and that it is possible to identify gene flow based on the level of linkage disequilibrium.

Keywords: gene flow, incomplete lineage sorting, nuclear DNA, *Triturus*

Introduction

Interspecific gene flow can take place between species of many different taxonomic groups (MALLET, 2005). Nevertheless, phylogenetic inference usually assumes that lineages split and do not meet again, while there is accumulating evidence that the speciation process can proceed while the "species" are still exchanging genes (for example, NOSIL, 2008). While the increase of accessibility to nuclear gene sequencing has allowed an examination of the process of speciation and gene flow in greater detail, an unfortunate aspect is that incomplete lineage sorting is frequently encountered, which can cause serious difficulties for phylogenetic inference (MADDISON and KNOWLES, 2006).

Incomplete lineage sorting, a locus-specific stochastic process, is one of the phenomena that underlies the topological and branch-length differences between phylogenies derived from different loci (Edwards, 2009). It can produce phylogenetic patterns similar to those produced by interspecific hybridization. Incomplete lineage sorting is especially problematic in groups that have undergone a recent and rapid radiation (for example BUCKLEY *et al.*, 2006; BELFIORE *et al.*, 2008). We here argue that, under certain conditions, the geographical location of the populations studied may help to distinguish between these essentially ‘horizontal’ (interspecific gene flow) and ‘vertical’ (incomplete lineage sorting) processes.

In deme-structured species with low dispersal ability, such as amphibians, contemporary gene flow will be restricted to neighbouring populations. Thus, hybridization and introgression will be naturally localized near the species border in parapatric hybrid zones. As individuals that picked up alien genes move away from the border area, they will mate with genetically pure individuals of either their own or the other species. Over the generations recombination will break down linkage disequilibrium, eventually diluting the genetic signature of gene introgression. Incomplete lineage sorting will, in contrast, be expressed in parts of the genome randomly and there will be no association with geographic location.

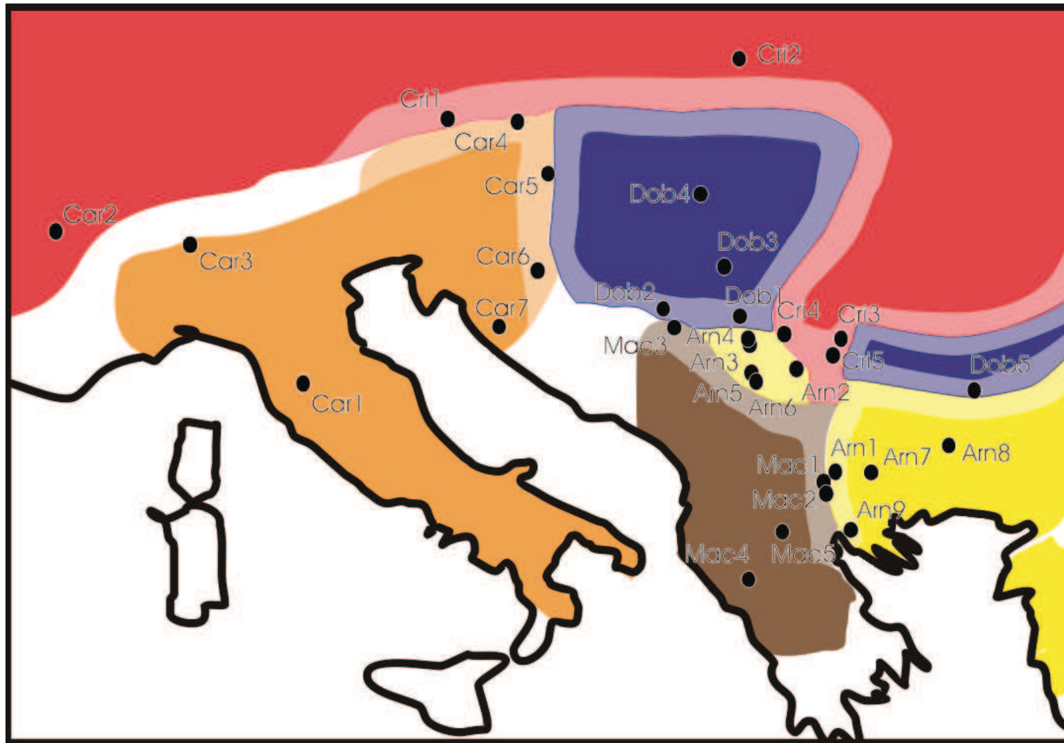


Figure 1 Distribution of *Triturus* species in the Balkans, represented by the following colours: orange - *T. carnifex*, red - *T. cristatus*, blue - *T. dobrogicus*, yellow - *T. arntzenii*, brown – *T. macedonicus*. Lighter colours represent the areas within 100 km of neighbouring species. Sample localities are numbered as described in Table I.

The *Triturus cristatus* superspecies was chosen as a study group because their populations are deme-structured and their parapatric range borders are long, well-documented (Figure 1; Arntzen, 2003) and subject to hybridization and introgression (Wallis and Arntzen, 1989). The superspecies is composed of six closely related 'crested' newt species, placed in the genus *Triturus* and the family Salamandridae, and has a wide Palearctic distribution. *T. arntzeni* (arn) is native to the southeastern Balkans and Turkey, *Triturus carnifex* (car) is found south of the Alps and in the northeastern Balkans, *T. cristatus* (cri) has a wide range from France to Russia, to the north of the other species in the group, *T. dobrogicus* (dob) occurs in the Pannonian and Dobrogean basins and *T. macedonicus* (mac) is restricted to the southern Balkans. Finally, *T. karelinii* is strictly Asian occurring in Anatolia, the Caucasus and the southern Caspian sea-shore. Its only contact zone with another crested newt (*T. arntzeni*) has an undetermined location (see Chapter 4) and so it was not included in this study. Interspecific gene flow is limited to areas closer than 100 km to the range border (J.W. Arntzen *et al*., unpublished data). We use the

species as a model to investigate the genetic signatures of hybridization and incomplete lineage sorting by comparing five nuclear and two mitochondrial DNA gene fragments. With these seven genes and *T. marmoratus* as outgroup, we solved the phylogeny of an otherwise problematic clade (CHAPTER 2) as seen from the root: (*T. arntzeni*, (*T. cristatus*, (*T. dobrogicus*, (*T. carnifex*, *T. macedonicus*)))) (CHAPTER 4). All branches in this tree were supported with a posterior probability of 1.0, with the exception of the branch leading to *T. cristatus*. Arguably, the phylogenetic reconstruction worked because only populations from the central part of the species' ranges were considered, that were unaffected by interspecific gene flow. If some alleles from certain species grouped with the "wrong" species clade the preferred explanation was incomplete lineage sorting and not interspecific gene flow. We here extend this study through the inclusion of populations close to the parapatic range borders.

Material and methods

Sampling

Tail-tip tissue was sampled in 33 newts from as many populations covering the Balkan area where five crested newt species meet, and including *T. marmoratus* and *T. pygmaeus* as outgroups (Fig. 1; Table 1). We did not include *T. karelinii* from Asia, as its only contact zone with another crested newt (*T. arntzeni*) has an unknown location (CHAPTER 4). Localities were classified as either "fringe" or "central", when the distance to the range of another species was lower or higher than 100 km, respectively (Fig. 1).

DNA extraction and sequencing

Tissue samples were stored in -80 ºC and later transferred to 95% ethanol. We extracted total genomic DNA from muscle tissue, using the Qiagen extraction kit. The nuclear markers examined include three introns: β-Fibrinogen intron 7 (βfib*int*7), Calreticulin intron C (Cal*int*C), intron 11 of the Platelet-derived growth factor receptor α (Pdgfrα); and two anonymous markers (Cri1 and Cri4) (CHAPTER 3). We also examined the mitochondrial genes NADH dehydrogenase subunits 2 and 4 (ND2 and ND4). PCR conditions consisted of a denaturation step at 94 ºC for 4 minutes, followed by 35 cycles of a denaturing step (30'' at 94 ºC), an annealing step (45'' at

57-68 ºC depending on the fragment) and an extension step (90’’ at 72 ºC) with a final extension step of 4’ at 72 ºC. The primers used are listed in CHAPTER 3 (Table 1). Sequences were obtained commercially at Macrogen Inc. Both DNA strands were sequenced to increase the accuracy of the results.

Table 1 *Triturus* individuals analysed, with geographical coordinates and population qualification as either parapatric or allopatric with other species.

Taxon Individual	Population number	Locality and country	Coordinates latitude (N)	longitude (E)	Voucher numbers	Location of population
Triturus cristatus						
cri 1	1	Ottenstein, Austria	48¼ 28'	14¼ 17'	691	Parapatric
cri 2	2	Limanowa, Poland	49¼ 43'	20¼ 25'	590	Allopatric
cri 3	3	Kladovo, Serbia	44¼ 36'	22¼ 33'	462	Parapatric
cri 4	4	Klokocevac, Serbia	44¼ 20'	21¼ 12'	446	Parapatric
cri 5	5	Stubik, Serbia	44¼ 18'	22¼ 22'	452	Parapatric
Triturus carnifex						
car 1	6	Florence, Italy	43¼ 47'	11¼ 15'	780	Allopatric
car 2	7	Geneva, Switzerland	46¼ 10'	6¼ 00'	471	Parapatric
car 3	8	Locarno, Switzerland	46¼ 10'	8¼ 48'	749	Allopatric
car 4	9	Etzmandorf, Austria	48¼ 39'	15¼ 45'	678	Parapatric
car 5	10	Lackenbach, Austria	47¼ 36'	16¼ 27'	382	Parapatric
car 6	11	Belovar Moravce, Croatia	45¼ 51'	16¼ 10'	788	Parapatric
car 7	12	Sinac, Croatia	44¼ 49'	15¼ 22'	876	Allopatric
Triturus macedonicus						
mac 1	13	Probistip, Macedonia	41¼ 59'	22¼ 10'	862	Parapatric
mac 2	14	Karbinci, Macedonia	41¼ 46'	22¼ 14'	860	Parapatric
mac 3	15	Gorne Cadavici, Bosnia	44¼ 45'	19¼ 05'	833	Parapatric
mac 4	16	Monodendri, Greece	39¼ 49'	20¼ 43'	2104	Allopatric
mac 5	17	Ano Kaliniki, Greece	40¼ 52'	21¼ 26'	602	Allopatric
Triturus dobrogicus						
dob 1	18	Belgrade, Serbia	44¼ 50'	20¼ 30'	297	Parapatric
dob 2	19	Jamena, Serbia	44¼ 54'	19¼ 02'	850	Parapatric
dob 3	20	Senta, Serbia	45¼ 55'	20¼ 06'	511	Allopatric
dob 4	21	Albertirsa, Hungary	47¼ 14'	19¼ 36'	337	Allopatric
dob 5	22	Svishtov, Bulgary	43¼ 37'	25¼ 21'	748	Parapatric
Triturus arntzeni						
arn 1	23	Bigla, Macedonia	41¼ 56'	22¼ 40'	794	Parapatric
arn 2	24	Cicevac, Serbia	43¼ 42'	21□27'	827	Parapatric
arn 3	25	Arandjelovac, Serbia	44¼ 19'	20¼ 35'	390	Parapatric
arn 4	26	Djurinci, Serbia	44¼ 30'	20¼ 38'	931	Parapatric
arn 5	27	Grivac, Serbia	43¼ 58'	20¼ 40'	821	Parapatric
arn 6	28	Vitanovac, Serbia	43□28'	21□16'	869	Parapatric
arn 7	29	Bansko, Bulgaria	41¼ 23'	22¼ 46'	867	Parapatric
arn 8	30	Karlovo, Bulgaria	42¼ 38'	24¼ 49'	719	Allopatric
arn 9	31	Dafnohori, Greece	40¼ 57'	22¼ 48'	816	Allopatric
Triturus marmoratus						
mar	32	Castelo Mendo, Portugal	40□ 35'	6□ 56'		Allopatric
Triturus pygmaeus						
pyg	33	Sagres, Portugal	37□ 04'	8□ 54'		Allopatric

Sequence analysis

Chromatograms were read with Chromas software (Technelysium Pty Ltd.) to check for unmistaken basecalling errors and the obtained sequences were imported into BioEdit (HALL, 1999). Insertion/deletion polymorphisms (indels) reflecting heterozygosity were reconstructed by hand through sequence comparison with a locally homozygous individual. Sequences polymorphic for indels were identified based on the presence of overlapping peaks downstream from particular areas of the sequence using both forward and reverse primers. Individual alleles were inferred with the software Phase (STEPHENS *et al.*, 2001). Nucleotide base composition, the transition–transversion ratio and the percentage sequence divergence were calculated with MEGA4 software (TAMURA *et al.*, 2007). Translating the mtDNA sequence to amino acids did not reveal any stop codons in coding regions.

Distinguishing incomplete lineage sorting and current gene flow

Linkage disequilibrium (LD) for nuclear markers was examined with Arlequin version 3.1 (EXCOFFIER *et al.*, 2005) under default parameter settings, for 'parapatric' populations, 'allopatric' populations and 'all' populations of a species. A Bonferroni correction was applied to take into account the non-independence of tests at each locus. We conducted a *t-test* on the Bonferroni corrected values to assess significance ($p<0.05$) between the LD values of parapatric/allopatric populations.

The phylogenetic tree described in the Introduction (Fig. 5a in CHAPTER 4) was accepted as representing the 'true' phylogeny at both the species and gene level. For each of the seven investigated genes, alternative trees were built by swapping an allele from its 'correct position' to each of the other clades in turn. Since we studied 33 individuals in five species, this yielded 330 trees for the nuclear genes and 165 trees for the mtDNA genes (Fig. 2). For each tree the likelihood score was calculated in PAUP (SWOFFORD, 2003) and compared with the baseline score of the 'true' phylogeny, as either more or less likely, under the rationale that better scores obtained for 'non-true' gene trees would indicate an incorrect phylogenetic signal of the considered allele. We then performed a *G-test* for independence to check if tree with improved likelihoods occur independently from the location of the population (parapatric vs allopatric).

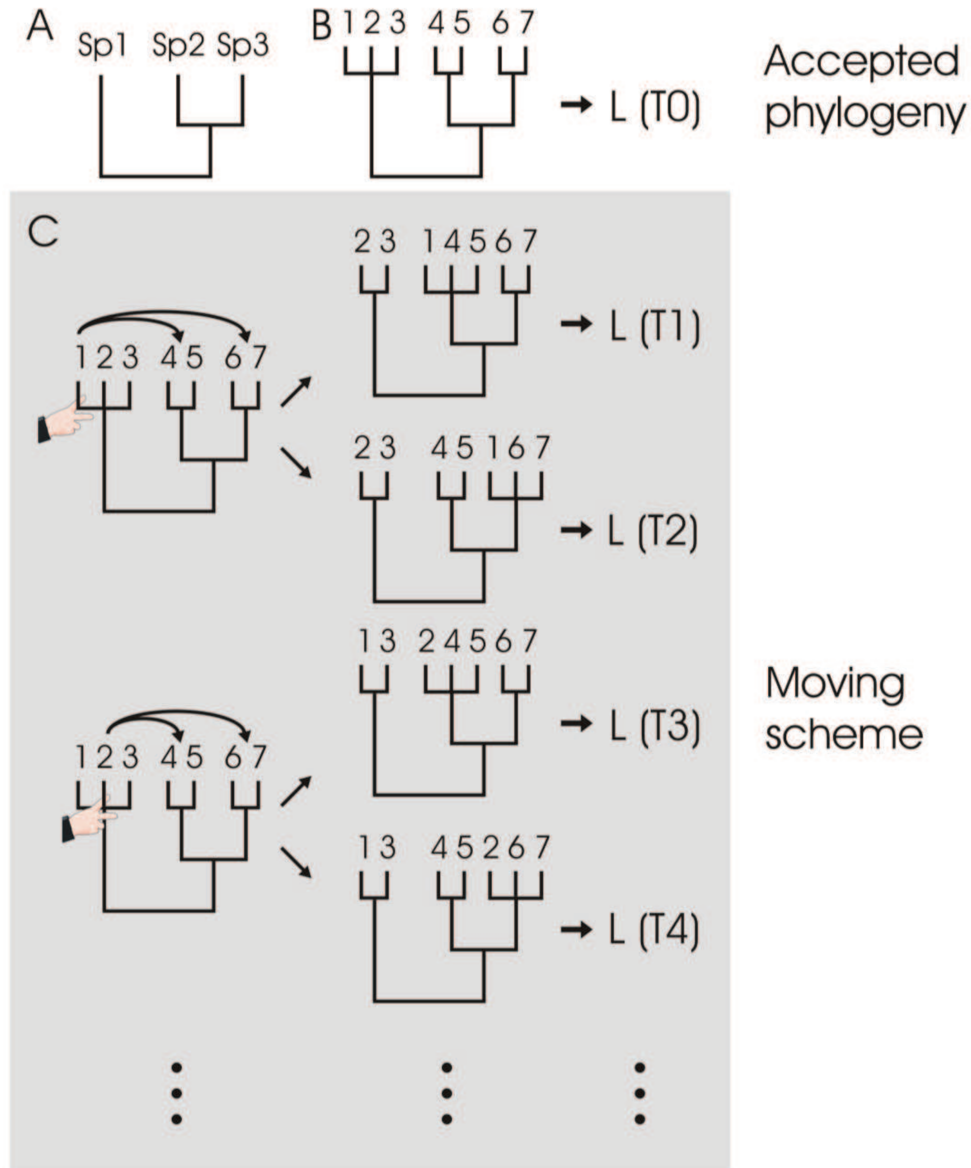


Figure 2 Schematic representation of the allele swapping procedure for a simple case with three species, with three, two and two individuals, respectively, starting with the 'true' or source phylogeny of Figure A, with an allelic species identity based on independent data such as morphology. Genetic substructuring within species is ignored (B). Alleles are swapped between groups yielding one tree for every swap. With three species and seven individuals the number of trees generated is 21 for mtDNA and 42 for nuclear alleles. The resulting tree likelihood (L (Tn)) qualifies the fit of the allele to the species tree as increased, equal or worse than the source phylogeny.

Table 2 Results from the *t*-test showing where the likelihood was higher if the allele was moved to another species in the tree. (A) ND2; (B) ND4; (C) βfib*int*7; (D) Cal*int*C; (E) Pdgfrα; (F) Cri1 and (G) Cri4. Star symbols represent that the allele would be better placed in the neighbouring population; while solid circles mean that the populations is not close to any other species or that it would be better placed in a non-neighbouring species.

Taxon	Used for constraint tree	Increased likelihood if sample moved to another group						
		βfib*int*7	Cal*int*C	Pdgfrα	Cri1	Cri4	ND2	ND4
Triturus cristatus								
cri 1		●		*	*		*	*
cri 2	Yes			●	●		●	●
cri 3			*	*	●	*	*	*
cri 4							●	*
cri 5							●	*
Triturus carnifex								
car 1	Yes							
car 2							●	●
car 3	Yes							
car 4		*	*		*	*	●	●
car 5		●	*	*	*	●	*	*
car 6		*						
car 7	Yes	●				●		
Triturus macedonicus								
mac 1							*	*
mac 2			*				*	*
mac 3				*			*	*
mac 4	Yes						●	●
mac 5	Yes						●	●
Triturus dobrogicus								
dob 1		*						
dob 2		●		*		*		
dob 3	Yes	●						
dob 4	Yes				●			
dob 5		●		●	*			
Triturus arntzeni								
arn 1							*	*
arn 2				*			*	*
arn 3				*		*	●	●
arn 4				*			●	●
arn 5				*			*	*
arn 6					●		*	*
arn 7							●	●
arn 8	Yes			●			●	●
arn 9	Yes						●	●

● - Moving the sample to another group improves the likelihood of the tree
* - Moving the sample to the neighbour species group improves the likelihood of the tree

Results

DNA sequence data

The sequencing of 66 chromosomes for five nuclear genes, and 33 haploid mitochondria, yielded 23 haplotypes in βfib*int*7, 24 in Cal*int*C, 26 in Pdgfrα, 29 in Cri1, 21 in Cri4, 23 in ND4 and 22 in ND2. Average heterozygosity in the nuclear genes was 0.33, 0.24, 0.33, 0.45 and 0.18, respectively. At least one indel was inferred in each of the nuclear markers, identical to the ones described in CHAPTER 3.

Table 4 Number of sites at linkage desiquilibrium (LD) in 'allopatric', 'parapatric' and total number of populations of crested newt species. Results are shown without any sort of correction for multiple tests and with Bonferroni correction. Values represent number of sites in linkage desiquilibrium as calculated by Arlequin. Asteriscs indicate a significant difference in number of sites at linkage disequilibrium between parapatric and allopatric populations after a Bonferroni correction. The bottom row for each species represents results from a double-sided *t*-test between parapatric and allopatric values of LD (* significant at p<0.05; NA Not applicable).

Species			BF	CalC	Cri1	Cri4	PDG
T. cristatus							
	parapatric	no correction	6*	211*	36*	21*	80*
		Bonferroni	6*	118*	21*	21*	70*
	allopatric	no correction	0*	0*	0*	0*	0*
		Bonferroni	0*	0*	0*	0*	0*
	all	no correction	6	196	50	22	80
		Bonferroni	6	79	7	0	4
	t-test		0.014*	0.000*	0.000*	0.000*	0.000*
T. carnifex							
	parapatric	no correction	7	173*	67	5	120*
		Bonferroni	1	126*	54	1	120*
	allopatric	no correction	0	1*	40	1	0*
		Bonferroni	0	0*	40	1	0*
	all	no correction	15	234	107	65	120
		Bonferroni	12	182	77	24	120
	t-test		0.317	0.000*	0.149	1.000	0.000*
T. macedonicus							
	parapatric	no correction	0	36*	10*	0	0
		Bonferroni	0	36*	10*	0	0
	allopatric	no correction	0	0*	0*	0	0
		Bonferroni	0	0*	0*	0	0
	all	no correction	0	15	0	0	780
		Bonferroni	0	2	0	0	0
	t-test		NA	0.000*	0.002*	NA	NA
T. dobrogicus							
	parapatric	no correction	0	0	37*	0	4*
		Bonferroni	0	0	37	0	4*
	allopatric	no correction	0	0	0*	0	0*
		Bonferroni	0	0	0*	0	0*
	all	no correction	8	0	21	55	21
		Bonferroni	6	0	18	10	18
	t-test		NA	NA	0.000*	NA	0.046*
T. arntzeni							
	parapatric	no correction	0	0	55*	28*	955*
		Bonferroni	0	0	55*	28*	441*
	allopatric	no correction	0	0	0*	0*	0*
		Bonferroni	0	0	0*	0*	0*
	all	no correction	0	0	55	13	956
		Bonferroni	0	0	0	2	800
	t-test		NA	NA	0.000*	0.000*	0.000*

Linkage disequilibrium

Each of the five nuclear genes examined exhibited a higher number of sites with linkage disequilibrium in parapatric than in allopatric populations (exact test, Table 4). We also observed significant differences in the amount of linkage disequilibrium between parapatric and allopatric populations in all species (*t*-test). *Triturus cristatus*, *T. arntzeni* and the remaining species showed significant differences at all loci, at three loci and at two loci, respectively.

Phylogenetic analysis

Moving an allele to a group different from the one in which it was classified yielded improved likelihood scores for the βfib*int*7 tree (25 times), Cal*int*C (10 times), Pdgfrα (52 times), Cri1 (26 times), Cri4 (40 times), ND2 (22 times) and ND4 (37 times) (table 2). Improved likelihood scores were more frequently observed for 'parapatric' than for 'allopatric' populations, but the results were not statistically significant (Table 3).

Table 3 Results of a *G*-test for independence (Williams) for the number of times the likelihood of the phylogenetic tree is improved by moving a sample to another group in parapatric and allopatric populations. In the parapatric column, only movements to neighbouring populations were included. None of the values is significant at p = 0.05.

Locus	Parapatric		Allopatric		*G*-test
	increased	not increased	increased	not increased	
βfib*int*7	3	18	2	8	0.141
Cal*int*C	4	17	0	10	1.002
Pdgfrα	9	12	2	8	1.535
Cri1	4	17	2	8	0.004
Cri4	4	17	1	9	0.392
ND2	10	11	5	5	0.015
ND4	12	9	5	5	0.132

Discussion

We set out to distinguish gene flow from incomplete lineage sorting by two different methods, based upon population genetics and phylogenetics, respectively. The analysis of linkage disequilibrium in nuDNA showed that all species have higher LD closer to range borders. Linkage disequilibrium is caused by the mixing of gene pools that have different allele frequencies (GOODMAN *et al.*, 1999) and is

indicative of hybridization having occurred relatively recently. Because our data show a marked contrast in the presence of LD in populations near species borders compared to populations far from these borders (Table 2), we conclude that hybridization is occurring, or has occurred in the last several generations, to some degree in all species pairs that we examined. The effect seems to be the greatest in *T. cristatus* - *T. carnifex* and weakest in *T. macedonicus* - *T. dobrogicus*, although there is evidence at some loci to suggest hybridization in the latter two species also.

The phylogeny based method did not reveal a convincing pattern, in which phylogenetic signal was improved more often in 'parapatric' than in 'allopatric' populations. Moreover, one might argue that our line of argument is circular in part, because of overlap in sequence data used for the current analysis and for phylogeny reconstruction (CHAPTER 4). Results from the combined phylogenetic/geographic analysis showed that "misplacements" happen in parapatric as well as in allopatric populations. Possible explanations for the misplacements in allopatric populations (and the fact that some parapatric "misplacements" are not with neighbouring species) in mtDNA include: 1) mtDNA has the ability to introgress faster and further relative to nuDNA, due to selection (RUEDI *et al.*, 1997); 2) the high degree of variability in mtDNA makes the species (constraint) tree we used a poor fit to that data (see mtDNA phylogeny in ARNTZEN *et al.*, 2007); 3) male-biased dispersal preserves "alien" genes in mtDNA (AVISE, 1994; GARCÍA-PARÍS *et al.*, 2003). Arntzen and Wallis (1999) and Arntzen et al (2007) argue that the presence of *T. arntzeni* alleles in bordering populations of *T. macedonicus* resulted from a range shift displacing *T. arntzeni* rather than positive selection of *T. arntzeni* alleles. There is no evidence supporting hypothesis 1), but it is possible that the observed pattern results from a combination of 2) and 3).

Hybridization

Species borders in this genus are permeable. Most contact zones in the Balkans are melting pots, where individuals from across taxa borders meet and reproduce. This seems to be occurring at an especially high degree between *T. macedonicus* and *T. arntzeni*. In Gorne Cadavici (Bosnia), the individual identified morphologically as *T. macedonicus* has great part of its DNA from other sources (*T. dobrogicus* or *T. arntzeni*). Even other populations of *T. macedonicus* from locations in Macedonia

have mtDNA haplotypes characteristic of the *T. arntzeni*. The finding that mtDNA has introgressed into *T. macedonicus* from *T. arntzeni* is not new (see ARNTZEN and WALLIS, 1999; ARNTZEN *et al.*, 2007), but the previously known area of introgression only included southern Bosnia and western Serbia. Our findings here show that this phenomenon extends to eastern Macedonia. A possible explanation for this and the apparent lack of *T. macedonicus* DNA in neighbouring *T. arntzeni* was advanced by Artnzen and Wallis (1999) involving the range shift of *T. arntzeni* and *T. macedonicus*. Previously, *T. arntzeni* was present further to the West than now, and *T. macedonicus* subsequently expanded its range eastward and progressively replaced *T. arntzeni*. This species' mtDNA remains in *T. macedonicus* as a genetic footprint, a relic of a past presence. The existence of a *T. arntzeni* exclave in the northwestern edge of its distribution (see Fig. 1 and ARNTZEN, 1978; ARNTZEN and WALLIS, 1999), is also indicative that *T. arntzeni* has retreated while the other species have expanded around the enclave. Formation of exclaves has also happened in the close relatives, *T. marmoratus* and *T. pygmaeus* (CHAPTER 9) and in other amphibians, including the fire-bellied toads, *Bombina bombina* and *B. variegata* (ARNTZEN, 1978). It is possible that this phenomenon is actually more common than previously thought in species with low dispersal ability and deme-structured populations.

In the northern limit of *T. macedonicus*, where it meets both *T. dobrogicus* and *T. arntzeni*, it is apparent that the three species are mixing, resulting in individuals that are morphologically *T. macedonicus*, with *T. dobrogicus* mtDNA and some *T. arntzeni* nuDNA. In the northeast of the *T. carnifex* distribution this species was able to cross the Alps, and there it meets with both *T. cristatus* and *T. dobrogicus*. The presence of DNA of both these species in *T. carnifex* could also be explained by *T. carnifex* expanding north and eastward, but that would only explain the presence of 'alien' mtDNA and not of nuDNA. We therefore suggest that this pattern is not the result of range displacement but of dispersal of *T. dobrogicus* and *T. cristatus* to the area. Comparing these results with published data on allozymes, we find that the populations of Kladovo (Serbia) and Etzmandorf (Austria) show the same pattern of introgression as in DNA, with a mixture of *T. cristatus*/*T. dobrogicus* and *T. carnifex*/*T. dobrogicus* respectively, in the population (ARNTZEN and WALLIS, 1999).

There are many areas where introgression is detected. In the introduced *T. carnifex* Geneva population, both the mtDNA as well as the nuDNA markers show that *T. cristatus* DNA has introgressed into *T. carnifex* since this population was introduced. The two samples of *T. macedonicus* show signs of introgression at the mtDNA level but not in nuDNA. Similarly, the *T. cristatus* populations of Klokocevac and Stubik in Serbia show the same pattern of introgression in mtDNA but not nuDNA. All these populations are close to the border with *T. arntzeni*. No signs of introgression are evident in *T. arntzeni* on the opposite side of the range limits. In the proximity of *T. dobrogicus* similar introgression events can be spotted: in the *T. cristatus* population of Kladovo (Serbia) it is the nuDNA that show signs of introgression and not mtDNA; the *T. carnifex* populations of Lackenbach and Etzmandorf (Austria), and Gorne Cadavici (Bosnia) mtDNA is typical of *T. dobrogicus*, however nuDNA is either typical of *T. arntzeni* (Gorne Cadavici), *T. cristatus* (Etzmandorf) or indeed *T. dobrogicus* (Lackenbach).

Taken together, our results suggest that there is a greater frequency of observed "alien" genes and linkage disequilibrium in populations close to the borders of other species. These phenomena do not occur as often in populations near the centre of the species' ranges. This indicates that the phylogenetic patterns we have observed are more likely to be explained by a hypothesis of hybridization between *Triturus* species, rather than incomplete lineage sorting of the loci we sampled. With these results, we have been able to show that in species with relatively stable distributions, taking geography into account can help resolve the underlying biological processes leading to observed phylogenetic patterns.

Acknowledgments

This work was funded by a PhD grant to GET from Fundação para a Ciência e para a Tecnologia (SFRH/BD/16894/2004).

References

ARNTZEN, J.W. (1978) Some hypotheses on postglacial migrations of the fire-bellied toad, *Bombina bombina* (Linnaeus) and the yellow-bellied toad, *Bombina variegata* (Linnaeus). *Journal of Biogeography* **5**, 339-345.

ARNTZEN, J.W. and WALLIS G.P. (1999) Geographic variation and taxonomy of crested newts (*Triturus cristatus* supersecies): morphological and

mitochondrial data. *Contribution to Zoology* **68**, 181-203.

ARNTZEN, J.W. (2003) *Triturus cristatus* Superspecies–Kammolch-Artenkreis. In: *Handbuch der Reptilien und Amphibien Europas. Schwanzlurche IIA* (eds. GROSSENBACHER, K., THIESMEIER B.), pp. 421-514. Aula-Verlag, Wiebelsheim.

ARNTZEN, J.W., ESPREGUEIRA THEMUDO G. and WIELSTRA B. (2007) The phylogeny of crested newts (*Triturus cristatus* superspecies): nuclear and mitochondrial genetic characters suggest a hard polytomy, in line with the paleogeography of the centre of origin. *Contributions to Zoology* **76**, 261-278.

AVISE, J.C. (1994) *Molecular Markers, Natural History and Evolution* Chapman & Hall.

BELFIORE, N.M., LIU L. and MORITZ C. (2008) Multilocus Phylogenetics of a Rapid Radiation in the Genus *Thomomys* (Rodentia: Geomyidae). *Systematic Biology* **57**, 294 - 310.

BUCKLEY, T., CORDEIRO M., MARSHALL D. and SIMON C. (2006) Differentiating between hypotheses of lineage sorting and introgression in New Zealand Alpine cicadas (*Maoricicada dugdale*). *Systematic Biology* **55**, 411.

EXCOFFIER, L., LAVAL G. and SCHNEIDER S. (2005) Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.

GARCÍA-PARÍS, M., ALCOBENDAS M., BUCKLEY D. and WAKE D.B. (2003) Dispersal of viviparity across contact zones in Iberian populations of fire salamanders (Salamandra) inferred from discordance of genetic and morphological traits. *Evolution* **57**, 129-143.

GOODMAN, S.J., BARTON N.H., SWANSON G., ABERNETHY K. and PEMBERTON J.M. (1999) Introgression Through Rare Hybridization: A Genetic Study of a Hybrid Zone Between Red and Sika Deer (Genus Cervus) in Argyll, Scotland. *Genetics* **152**, 355-371.

HALL, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95-98.

MADDISON, W.P. and KNOWLES L.L. (2006) Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology* **55**, 21-30.

MALLET, J. (2005) Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* **20**, 229-237.

NOSIL, P. (2008) Speciation with gene flow could be common. *Molecular Ecology* **17**, 2103-2106.

RUEDI, M., SMITH M.F. and PATTON J.L. (1997) Phylogenetic evidence of mitochondrial DNA introgression among pocket gophers in New Mexico (family Geomydae). *Molecular Ecology* **6**, 453-462.

STEPHENS, M., SMITH N.J. and DONNELLY P. (2001) A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics* **68**, 978-989.

SWOFFORD, D.L. (2003) PAUP*: Phylogenetic Analysis Using Parsimony (and other methods) 4.0 Beta 10. Sunderland (MA): Sinauer Associates.

TAMURA, K., DUDLEY J., NEI M. and KUMAR S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* **24**, 1596.

WALLIS, G.P. and ARNTZEN J.W. (1989) Mithocondrial-DNA variation in the crested newt superspecies: limited cytoplasmic gene flow among species. *Evolution* **43**, 88-104.

CHAPTER 6

CURRENT GENE FLOW ACROSS A NEWT HYBRID ZONE FOLLOWS LOCAL ECOLOGICAL CONDITIONS

G. Espregueira Themudo[1, 2] and J. W. Arntzen[1]

[1] *National Museum of Natural History, P. O. Box 9517, 2300 RA Leiden, The Netherlands.*
[2] *CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661*
Vairão, Portugal.

Contents

Abstract 96
Introduction 97
Material and methods 100
  Sampling 100
  DNA protocol 102
  Sequence analysis 102
  Mutation and recombination rates for IM analysis 105
  IM analysis 106
Results 106
  Gene diversity 106
    mtDNA diversity 107
    Beta fibrinogen diversity 107
    Calreticulin intron c diversity 108
    Platelet derived growth factor receptor alpha diversity 108
  Mutation rates 108
  Isolation with migration analysis 108
    Coastal area 112
    Tejo basin 112
    Idanha 112
  Migration over time 112
Discussion 113
  Species divergence and reticulation, assumptions of the method 113
  Hybrid zone position and width 114
Acknowledgments 116
References 116

Manuscript in preparation

Abstract

Conditions of the environment can influence the number of genes that pass through an interspecific hybrid zone. In long hybrid zones, different patterns may emerge depending on local conditions. We studied gene flow across a ca. 500 km long hybrid zone between two newt species, *Triturus marmoratus* and *T. pygmaeus,* in the Iberian Peninsula. We sequenced one mitochondrial and three nuclear genes in 355 individuals from 73 populations to test *a priori* hypotheses on the level and symmetry of gene flow, under an *Isolation with Migration* framework. Four different environmental settings associated with the position and structure of the hybrid zone (identified with ecogeographical modeling; Arntzen & Espregueira Themudo, J. Biogeogr. 35, 2008) included i) a lowland - mountain transition at Sierra de Gata in western Spain and ii) the river Tejo in central Portugal forming a wide and a narrow barrier to gene flow, respectively. In between Tejo and Gata (iii) the position of the zone is determined by a minimum spanning distance with no environmental signature, as predicted by hybrid zone theory. The Portuguese coastal area (iv) has been subject to the rapid expansion of one of the species (*T. pygmaeus*) over the other, following edaphic and climatological features of the environment. There we observed a unidirectional pattern of gene flow and symmetry elsewhere. The prediction that gene flow across a barrier (Tejo river) would be lower than in an area without such a barrier was not confirmed. However, a temporal analysis indicated that most local gene flow is ancient, indicating that the barrier is stronger at present then it was in the past.

Key words: dispersal, isolation with migration, non-equilibrium, parapatry, *Triturus marmoratus*, *Triturus pygmaeus*

Introduction

Closely related species may 'meet, mate and produce hybrids' (Barton and Hewitt, 1985), forming more or less narrow hybrid zones that can span for hundreds of kilometers. The study of these areas of limited intergradation has provided new insights into the process of speciation (Coyne and Orr, 2004). Across the hybrid zone populations will have progressively fewer genotypes of the first species until only pure genotypes of the second species are found, forming a cline of variable width and shape. The width of the cline depends on the balance of two opposing forces: dispersal of parental individuals into the hybrid zone and endogenous selection against hybrids, caused by the disruption of intra-genomic coadaptation. This clinal model is also referred to as the 'tension zone' model (a term coined by Key, 1969). In strict tension zones, the frequency of alternative genotypes in a population is only dependent to the distance to the centre of the hybrid zone (Barton, 1979; Szymura and Barton, 1986).

On another extreme, the 'mosaic' model of hybrid zones postulates that two species will be distributed according to habitat preference, and distance to the centre of the zone is not a good predictor for the proportion of parental genotypes in a population (Rand and Harrison, 1989). Mosaic hybrid zones are usually bimodal hybrid zones, with a low percentage of hybrid genotypes (Jiggins and Mallet, 2000). The two models can be seen as part of a continuum, and can be observed in the same system, dependent on spatial scale (Arntzen, 1996). Indeed, the position of a clinal hybrid zone may be determined by the distribution of landscape elements and the presence of geographical barriers, such as altitude (or perhaps relief) in the fire-bellied toads, *Bombina bombina / B. variegata*, hybrid zone (Arntzen, 1978), or the Strait of Gibraltar in the long-eared bats, *Myotis myotis / M. punicus*, hybrid zone (Castella et al., 2000).

Width and shape of the clinal hybrid zone may be affected by smaller landscape elements, such as forestation, the availability of a particular breeding habitat, or the distribution of relief (e.g. KRUUK et al., 1999; SZYMURA, 1988) or different degrees of complexity of the habitat (MACCALLUM *et al*., 1998; YANCHUKOV *et al*., 2006). Hybrid zones are not necessarily stable neither (ARNTZEN and WALLIS, 1991; BRITCH *et al.,* 2001; FITZPATRICK et al., 2004). Changes in the environment can favor one of the species, allowing it to expand into the area and

eventually supersede the other species (ARNTZEN, 1978; ARNTZEN and WALLIS, 1991; DASMAHAPATRA *et al.*, 2002; PEARSON, 2000).

Large rivers and mountain ranges may pose barriers to amphibian dispersal. Evidence supporting this notion is mostly derived from the analysis of distribution patterns in which the edge of the range coincides with the feature considered (GASCON, 1996; GARCÍA-PARÍS *et al.*, 1998). Alternatively, phylogeographic patterns may coincide with intraspecific genetic differentiation, such as the Mondego river in central Portugal, that was implicated in the maintenance of a hybrid zone in the salamander *Chioglossa l. lusitanica* and *C. l. longipes* (ALEXANDRINO *et al.*, 2000). Once crossed, obviously, neither a river nor a mountain range will form a barrier any longer. In terms of their barrier function, the essential difference between a mountain range and a river is in their width. Considering its limited width, a river is best qualified as a semi-permeable barrier, the strength of which will depend on the swimming and survival capability of the species during its various life-stages.

In long hybrid zones, environmental conditions may be specific to a part of the system, like differences in land use. These different conditions may produce opposing patterns in separate areas of the hybrid zone, such as in the *Mus musculus* hybrid zone in the Czech Republic and Germany, where introgression went in opposite directions (Božíková et al., 2005). The analysis of species distributions as a function of the environment gives, in parapatrically distributed and hybridizing species, rise to testable hypothesis in the amount of gene flow they exhibit. We here study the amount of gene flow across a bimodal hybrid zone of two species of newts in the Iberian Peninsula. We have a good overall picture of how the environment regulates the distribution of the species (Chapter 8), with detailed ecogeographical models for the entire hybrid zone. This zone is particularly well-suited for such an approach, because it is long (ca. 500 km), with various environmental settings determining its position (soil type, river, mountains). We here focus on areas where interspecific interactions are subject to clearly different ecogeographical signatures (CHAPTER 8) and we, thus, expect the genetic interaction of the species to vary with space.

Triturus marmoratus and *T. pygmaeus* are two European newt species (although sometimes still referred to as subspecies, LOUREIRO *et al.*, 2008) with a parapatric distribution. The contact zone runs along an east-west axis from Madrid in Spain to Abrantes in Portugal, where the zone bends northwards, with *T. pygmaeus*

present in the south and along the Atlantic coast up to Aveiro and *T. marmoratus* inland and to the north (Figure 1). Southwest of the bend, we observed a residual population of *T. marmoratus* (CHAPTER 9). Since the population is entirely surrounded by *T. pygmaeus* it effectively constitutes an enclave. A detailed study in the area revealed a single possible hybrid among ca. 400 individuals studied (Chapter 9). North of here, about 150 km of dune habitat is occupied by *T. pygmaeus* (Fig. 1, area A; 'windows 1-13' in Chapter 8). It is important to note that *T. marmoratus* is widely distributed in the coastal dunes north of Aveiro. The presence of an enclave, the near absence of interspecific hybridization and the remarkable distribution pattern along the coast, suggests the fast advance of *T. pygmaeus* up to the Vouga estuary in Aveiro, while superseding *T. marmoratus*. We predict that gene flow is absent or very limited and when it occurs, unidirectional from the advancing species, *T. pygmaeus* into the receding species, *T. marmoratus*,and localized at the edge of the wave of advance.

On the opposite side of joint *T. marmoratus* – *T. pygmaeus* distribution at the eastern fringe of the Central System, the contact zone appears residual (sensu SZYMURA, 1993) with a low density of populations and no documented hybridization (GARCÍA-PARÍS et al., 2001). However, a dense network of populations spanning both species was found at the Sierra de Gata mountains, at the western fringe of the Central System (Figure 1, area D). The species divide is positioned along the southern mountain slopes and is fairly abrupt. Locally it is difficult to disentangle between the multiple climatological parameters that coincide with the species border (CHAPTER 8). Because of the barrier function of (or associated with) the mountain/lowland transition, our prediction here is that the hybrid zone is stable with little or no interspecific gene flow. Note however that the genetic data on the Gata transect are as yet incomplete.

In central Portugal, at the Tejo Basin east of Abrantes (Figure 1, area B), the position of the hybrid zone coincides with the river, that seems to be working as a barrier to dispersal (Chapter 8; 'windows 14-20'). We predict that local gene flow will be low, given the presence of the barrier (but not as low as at the mountains) and symmetrical as there is no evidence for large scale movement of the parapatric range border.

Finally, around Idanha-a-Nova, in eastern Portugal (Figure 1, area C), the position of the contact zone appears not associated with features of the environment

(Chapter 8; 'windows 21-26'). Rather, the position of the zone seems to be dependent on fixed points to the west (B) and east (D). Following tension zone theory (BARTON and HEWITT, 1985), the situation in Idanha (C) represents a minimum-spanning distance between the Tejo river and the Gata mountains. Because stable at either side, at Idanha the zone will be moving neither, yielding the prediction of gene-flow symmetry. Because of the absence of a physical barrier to dispersal, we predict that gene flow across the zone will be relatively high. Our predictions are summarized in Table 1.

Table 1- Hypothesis regarding symmetry in gene flow ($2N_1m_1=2N_2m_2$), presence of gene flow from Triturus pygmaeus to *T. marmoratus* ($2N_1m_1>0$) and in the opposite direction ($2N_2m_2>0$). The predicted level of gene-flow is D < B < A < C.

Area	$2N_1m_1=2N_2m_2$	$2N_1m_1>0$	$2N_2m_2>0$
A - Coast	no	yes	no
B - Tejo	yes	yes	yes
C - Idanha	yes	yes	yes
D - Gata	not applicable	no	no

Material and methods

Sampling

We sampled 355 marbled newts in 73 populations across the range of both species in Portugal, Spain and France between 1998 and 2008 (DNA protocol

DNA extraction followed standard protocols from the DNeasy extraction kit (Qiagen). The three introns examined were: □-Fibrinogen intron 7 (□fibint7), Calreticulin intron

C (CalintC), and Platelet-derived growth factor receptor □ (Pdgfr□) intron 11 (Chapter 3). We also examined the mitochondrial gene NADH dehydrogenase subunits 4 (ND4). PCR conditions consisted of a denaturation step at 94 ºC for 4 minutes, followed by 35 cycles of a denaturing step (30'' at 94 ºC), an annealing step (45'' at 57-68 ºC depending on the fragment) and an extension step (90'' at 72 ºC) with a final extension of 4' at 72 ºC. The primers used are presented in Chapter 3. Reaction chemistry was 23 µL of H2O, 3 µL of buffer (15 µM MgCl2), 1.8 µL of 25 µM $MgCl_2$, 0.6 µL of dNTPs (10 mM), 0.2 µL of each primer (100 µM) and 0.2 µL (1 U) of Taq DNA Polymerase (Qiagen). Sequences were obtained commercially at

Macrogen Inc (Seoul, Korea). Both DNA strands were sequenced to increase the accuracy of the results.

Sequence analysis

DNA sequences were edited using Chromas (Technelysium Pty Ltd.) to check for basecalling errors. Insertion/deletion polymorphisms (indels) reflecting heterozygosity were reconstructed by hand through sequence comparison with a homozygous individual. Sequences polymorphic for indels were identified based on the presence of overlapping peaks downstream from particular areas of the sequence, using both forward and reverse primers. Individual alleles were inferred with the software Phase (STEPHENS *et al.*, 2001). The obtained sequences were aligned in BioEdit (HALL, 1999). Multi-base indels were pruned leaving the first base and any polymorphic bases, to avoid overestimating the differences between haplotypes. Nucleotide base composition, the transition–transversion ratio and the percentage sequence divergence (p-distance) were calculated with MEGA4 software (TAMURA *et al.*, 2007). Translating the coding regions of the sequences to amino acids did not reveal any unexpected stop codons. We calculated haplotype diversity (h), nucleotide diversity (π), and Watterson's θ for both species using DNAsp (ROZAS *et al.*, 2003).

Table 1; Figure 1). Adult newts and larvae were captured using dip nets under the appropriate licences from 'Instituto da Conservação da Natureza' (Portugal), 'Consejarias del Medio Ambiente' (Spain) and 'Ministére de l'Environment' (France). Adults were identified based on morphology, and larvae were identified by allozymes (see CHAPTERS 7 and 9). To preserve the DNA, samples were immediately stored in liquid nitrogen and later transferred to a -80º C freezer.

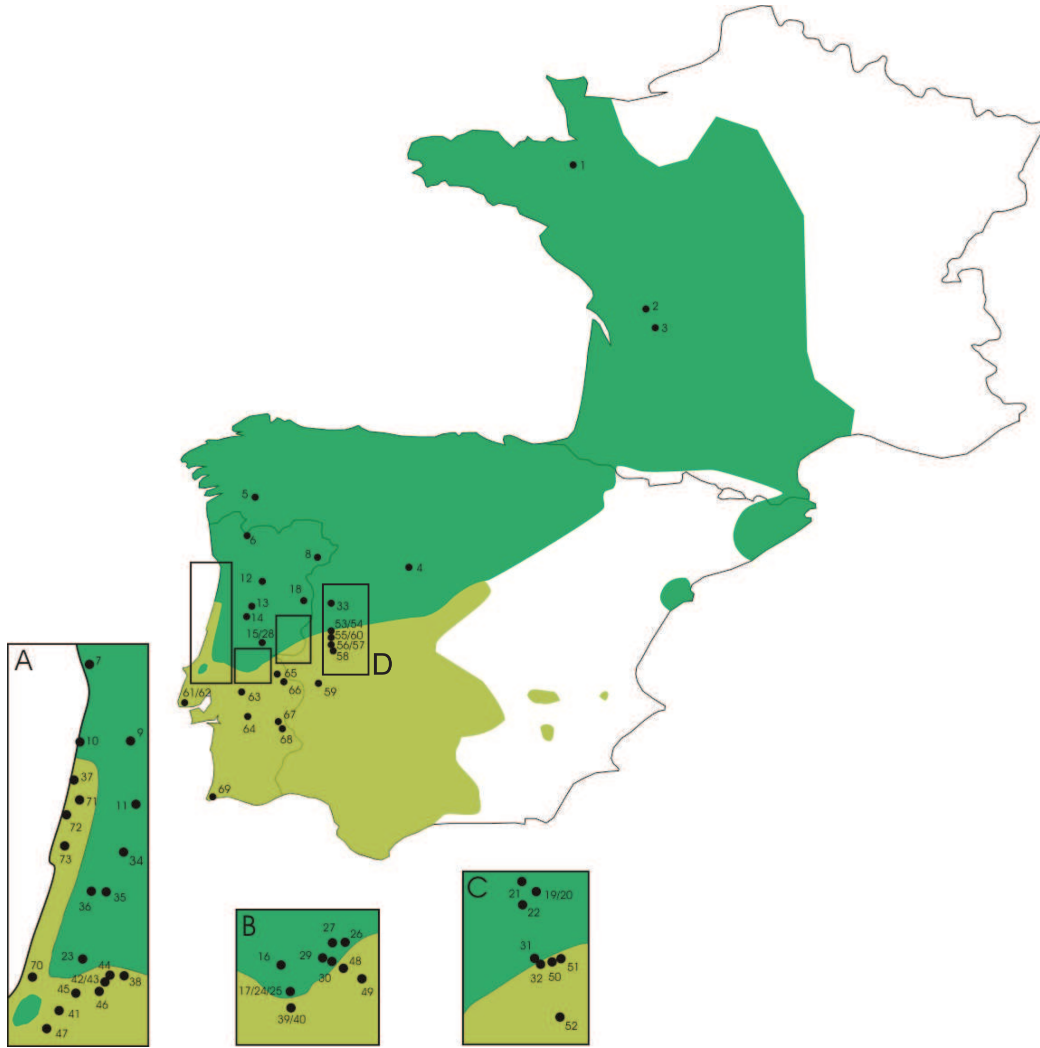


Figure 1 - Distribution map showing the range of *Triturus marmoratus* (in dark green) and *T. pygmaeus* (in light green) in Western Europe. Black dots represent sampling localities. Boxes labelled from A to C indicate areas sampled in more detail (see text). A – Coastal area, B – Tejo basin, C – Idanha-a-Nova and D - Sierra da Gata.

DNA protocol

DNA extraction followed standard protocols from the DNeasy extraction kit (Qiagen). The three introns examined were: β-Fibrinogen intron 7 (βfib*int*7), Calreticulin intron C (Cal*int*C), and Platelet-derived growth factor receptor α (Pdgfrα) intron 11 (CHAPTER 3). We also examined the mitochondrial gene NADH dehydrogenase subunits 4 (ND4). PCR conditions consisted of a denaturation step at 94 ºC for 4 minutes, followed by 35 cycles of a denaturing step (30'' at 94 ºC), an annealing step (45'' at 57-68 ºC depending on the fragment) and an extension step (90'' at 72 ºC) with a final extension of 4' at 72 ºC. The primers used are presented in CHAPTER 3. Reaction chemistry was 23 μL of H_2O, 3 μL of buffer (15 μM $MgCl_2$), 1.8 μL of 25 μM $MgCl_2$, 0.6 μL of dNTPs (10 mM), 0.2 μL of each primer (100 μM) and 0.2 μL (1 U) of Taq DNA Polymerase (Qiagen). Sequences were obtained commercially at Macrogen Inc (Seoul, Korea). Both DNA strands were sequenced to increase the accuracy of the results.

Sequence analysis

DNA sequences were edited using Chromas (Technelysium Pty Ltd.) to check for basecalling errors. Insertion/deletion polymorphisms (indels) reflecting heterozygosity were reconstructed by hand through sequence comparison with a homozygous individual. Sequences polymorphic for indels were identified based on the presence of overlapping peaks downstream from particular areas of the sequence, using both forward and reverse primers. Individual alleles were inferred with the software Phase (STEPHENS *et al.*, 2001). The obtained sequences were aligned in BioEdit (HALL, 1999). Multi-base indels were pruned leaving the first base and any polymorphic bases, to avoid overestimating the differences between haplotypes. Nucleotide base composition, the transition–transversion ratio and the percentage sequence divergence (p-distance) were calculated with MEGA4 software (TAMURA *et al.*, 2007). Translating the coding regions of the sequences to amino acids did not reveal any unexpected stop codons. We calculated haplotype diversity (h), nucleotide diversity (π), and Watterson's θ for both species using DNAsp (ROZAS *et al.*, 2003).

Table 1- List of sample localities per species with information on the country (F=France, S=Spain, P=Portugal), geographical coordinates and sample size (N) for three nuclear and one mitochondrial genes. Top panel *Triturus marmoratus* and bottom panel *T. pygmaeus.*

Triturus marmoratus Locality	Population	Country	Latitude (N)	Longitude (W)	βfib*int7*	Cal*intC*	Pdgfrα	mtDNA
Mayenne	1	F	48º 17' 60"	0° 37' 01"	4	4	4	
Confolens	2	F	46º 01’ 00"	- 0º 40' 00"		1	1	
Rochechouart	3	F	45º 49’ 00"	- 0º 50’ 00"		3	3	
El Berrueco	4	S	40º 54’ 00"	3º 53' 00"		3	4	
Cabreiras	5	S	42º 26' 49"	7º 54' 20"	4	5	4	5
Serra do Gerês	6	P	41º 49' 04"	8º 02' 12"	4	5	7	5
Porto	7	P	41º 09' 12"	8º 38' 31"	5	5	5	5
Mogadouro	8	P	41º 20' 48"	6º 36' 33"	5	5	5	5
Nespereira de Baixo	9	P	40º 45' 58"	8º 22' 19"	5	5	5	5
Torreira	10	P	40º 45' 51"	8º 42' 34"	2	3	3	3
Saide	11	P	40º 27' 02"	8º 19' 40"	2	2	2	5
Mezio	12	P	40º 56' 31"	7º 48' 43"	1	5	5	5
Nelas	13	P	40º 29' 30"	7º 50' 36"	5	5	5	10
Madeirã	14	P	39º 56' 38"	8º 02' 24"				5
Pedra do Altar	15	P	39º 42' 37"	7º 49' 14"				4
Santos	16	P	39º 36' 03"	7º 57' 45"	5	4	4	5
Domingos da Vinha	17	P	39º 31' 24"	7º 56' 48"				5
Castelo Mendo	18	P	40º 35' 15"	6º 56' 38"	3	3	4	5
Monte do Conde	19	P	40º 10' 08"	7º 15' 00"				3
Aldeia de João Pires	20	P	40º 06' 48"	7º 08' 45"				3
São João do Deserto	21	P	40º 08' 28"	7º 11' 31"	5	4	5	5
Aldeia Stª Margarida	22	P	40º 03' 08"	7º 15' 53"				5
Santa Catarina	23	P	39º 27' 02"	9º 00' 54"	3	4	4	4
Areia 2	24	P	39º 30' 50"	7º 55' 54"	4	5	5	5
Areia 3	25	P	39º 30' 50"	7º 56' 02"				5
Vila Velha de Ródão	26	P	39º 40' 12"	7º 40' 06"	4	4	5	4
Vilar de Boi	27	P	39º 40' 18"	7º 44' 12"	5	5	5	5
Vale da Figueira	28	P	39º 40' 52"	7º 44' 50"				5
Carepa	29	P	39º 37' 33"	7º 46' 23"	5	5	5	5
Fratel	30	P	39º 37' 06"	7º 44' 18"	5	5	5	5
Idanha-a-Nova	31	P	39º 56' 02"	7º 11' 51"	5	5	5	4
Senhora do Almurtão	32	P	39º 54' 23"	7º 09' 55"	4	5	5	5
Serradilla del Llano	33	S	40º 29' 39"	6º 20' 42"	5	5	5	5
Coimbra	34	P	40º 12' 28"	8º 24' 44"				5
Sicó	35	P	40º 01' 09"	8º 31' 52"				5
	36	P	40º 01' 10"	8º 37' 30"				10

Table 1 (cont.)

Triturus pygmaeus Locality	Population	Country	Latitude (N)	Longitude (W)	βfib*int*7	Cal*intC*	Pdgfrα	mtDNA
Praia de Vagueira	37	P	40º 34' 32"	8º 45' 23"	10	10	10	10
Tomar	38	P	39º 36' 15"	8º 25' 03"	5	5	5	5
Gavião 1	39	P	39º 28' 36"	7º 56' 19"	3	5	6	8
Gavião 2	40	P	39º 27' 32"	7º 55' 55"				5
Mosteiro Alcanene	41	P	39° 25' 47"	8° 50' 26"	4	4	4	
Rexaldia 1	42	P	39º 34' 40"	8º 32' 03"	1	1	1	1
Rexaldia 2	43	P	39º 34' 38"	8º 31' 59"	2	3	4	4
Assentiz - Fungalvaz	44	P	39º 35' 44"	8º 30' 39"	5	5	5	5
Serra de Stº António	45	P	39º 30' 55"	8º 43' 44"	5	5	5	5
Alqueidão	46	P	39º 32' 17"	8º 34' 51"	1	5	5	5
Rio Maior	47	P	39º 20' 46"	8º 54' 55"	5	5	5	6
Velada	48	P	39º 34' 49"	7º 42' 04"	5	5	2	5
Nisa	49	P	39º 33' 01"	7º 35' 34"	1	1	1	5
Zebreira West	50	P	39º 50' 32"	7º 06' 59"	5	5	5	5
Zebreira	51	P	39º 55' 56"	7º 05' 32"	2	3	4	4
Rosmaninhal	52	P	39º 45' 08"	7º 05' 50"	5	5	5	5
Casar de Palomero 1	53	S	40° 16' 44"	6° 14' 17"	4	4	4	5
Casar de Palomero 2	54	S	40° 16' 32"	6° 14' 27"	1	1	1	
Palomero	55	S	40º 15' 02"	6º 19' 51"	5	5	5	5
El Bronco	56	S	40º 12' 52"	6º 19' 56"	5	5	5	5
Stª Cruz de Paniagua	57	S	40º 11' 48"	6º 19' 59"	3	4	4	5
Aceitunas	58	S	40º 09' 05"	6º 19' 23"	4	4	4	4
Alcuescar	59	S	39º 05' 21"	6º 21' 49"	4	5	5	5
Pedro Muñoz 3	60	S	40º 15' 32"	6º 20' 46"	1	1	1	4
Sintra 1	61	P	38º 47' 38"	9º 25' 33"	3	3	2	2
Sintra 2	62	P	38º 47' 16"	9º 23' 17"	5	5	5	4
Mora	63	P	38º 57' 46"	8º 08' 45"	4	5	4	4
Mitra	64	P	38º 32' 17"	8º 00' 06"	5	5	5	5
Esperança	65	P	39º 10' 05"	7º 10' 57"	7	9	10	8
Sovrete	66	P	39º 16' 07"	7º 17' 03"				3
Mourão	67	P	38º 24' 47"	7º 18' 58"	2	1	1	2
Granja	68	P	38º 17' 59"	7º 15' 18"	5	4	4	5
Vila do Bispo	69	P	37º 04' 56"	8º 54' 35"	4	5	5	5
Valado dos Frades	70	P	39º 35' 51"	9º 00' 33"				5
Calvão	71	P	40º 28' 13"	8º 42' 18"				3
Mira	72	P	40º 23' 54"	8º 47' 16"				4
Quiaios	73	P	40º 14' 42"	8º 48' 00"				4

To represent relationships between haplotypes, we built phylogenetic networks for each gene, using the Median Joining algorithm as implemented in the program Network (BANDELT *et al.*, 1999). For intraspecific data, polymorphism is scarce and relationships between alleles may not be hierarchical, so networks will display the information from these genes more adequately than trees (POSADA and CRANDALL, 2001).

<u>Mutation and recombination rates for IM analysis</u>

In the absence of estimates on the mutation rates of the nuclear genes here analyzed, they were estimated with Beast 1.4.8 (DRUMMOND and RAMBAUT, 2007), under a strict clock model, HKY substitution model and a normal prior on the separation of *T. cristatus / T. marmoratus* with a mean of 24 Ma (millions of years), following (STEINFARTZ *et al.*, 2007). We then calculated the geometric mean of mutation rates, and mutation rate per generation by multiplying the per year rate with generation time. Since generation times vary geographically from 1-4 years in *T. pygmaeus* (CAETANO and CASTANET, 1993; DIAZ-PANÍAGUA, 1996) and 2-6 years in *T. marmoratus* (JEHLE *et al.*, 2001; CAETANO and CASTANET, 1993; JAKOB *et al.*, 2003; FRANCILLON-Vieillot, 1990), we accepted a generation time of 4 years, by excluding the population used in Diaz-Paníagua (1996) that is aberrantly small-sized compared to other populations of *T. pygmaeus*. Mutation rates and generation time are required to convert model parameter estimates from IMa (see below) into demographic quantities of number of migrants, effective population sizes and time of cladogenesis.

Isolation with Migration (IM) analyses assumes the absence of intralocus recombination. We tested this assumption by calculating the Φ_w statistic with PhiPack (BRUEN *et al.*, 2006). This statistic was less sensitive in simulation studies to mutation rate correlation than other similar statistics in simulation studies, making it less prone to falsely infer recombination when levels of homoplasy are high. PhiPack calculates *P*-values under the null hypothesis of no recombination. We used both available methods in the software to obtain these *P*-values: the analytical approach and a permutation test with 1000 permutations.

IM analysis

We carried out estimations of demographic parameters and dispersal rates with the coalescent based program IMa (HEY and NIELSEN, 2004) on three datasets from: the coastal area of Portugal, the Tejo basin in central Portugal, and the Idanha-a-Nova region close to the Spanish border (see areas A, B and C in Figure 1).

To determine appropriate parameter priors, we ran some initial runs with large priors for all the parameters, under the HKY substitution model. Runs started with 10,000 burnin steps after which they were allowed to continue until the minimum effective sample size (ESS) was >100. The Isolation with Migration (IM) model (Nielsen and Wakeley, 2001) assumes that two populations (or species) diverged from one ancestral populations at a certain time t and have since possibly been exchanging genes at migration rates m1 (from unit 2 to unit 1) and m2 (from 1 to 2). The program estimates posterior probability distributions of these three parameters (t, m1 and m2) and additionally the effective population size of the ancestral population (Θ_A) and of the two descendant units (Θ_1 and Θ_2). The program also yields estimates for the number of migration events through time.

Results

Gene diversity

Table 2 shows summary statistics for the nucleotide base composition, transition / transversion bias and percentage sequence divergence. Table 3 shows haplotype and nucleotide diversity, Watterson's θ, Tajima's D and Fu and Li's D for the complete dataset, and for each species.

Table 2 - Nucleotide base composition (A, C, G and T), transition/transversion bias (R), and intra (mar – *T. marmoratus*, and pyg – *T. pygmaeus*) and interspecific percentage sequence divergence (p-distance; ± s.e.) for the four *loci*.

	Nucleotide base composition (%)				R	% Divergence [± s.e.]		
Locus	A	C	G	T		mar	pyg	inter
βfib *int*7	29.7	19.1	18.2	33.1	0.537	0.007 ± 0.002	0.013 ± 0.003	0.014 ± 0.004
Cal *int*C	25.3	17.0	24.5	33.1	1.188	0.004 ± 0.002	0.004 ± 0.001	0.014 ± 0.005
Pdgfr α	28.3	21.2	18.3	32.1	3.364	0.003 ± 0.001	0.002 ± 0.001	0.004 ± 0.002
ND 4	29.4	29.8	14.9	25.9	6.681	0.015 ± 0.002	0.016 ± 0.003	0.038 ± 0.006

Table 3 – Sample size (N), length of fragment (L), haplotype diversity (Hd), nucleotide diversity (π), Theta (θ), and significance of neutrality tests (Tajima's D, and Fu and Li's D for *Triturus marmoratus*, *T. pygmaeus* and the total dataset. ** - significant at α=0.01.

Species	Locus	N	L	Hd	π	θ	D (Tajima)	D (Fu & Li)
Triturus marmoratus	βfib *int*7	111	500	0.607	0.006	0.006	NS	NS
	Cal *int*C	133	472	0.567	0.004	0.006	NS	NS
	Pdgfr α	124	672	0.517	0.002	0.003	NS	NS
	ND 4	160	658	0.886	0.015	0.019	NS	NS
Triturus pygmaeus	βfib *int*7	172	500	0.924	0.013	0.010	NS	NS
	Cal *int*C	198	472	0.743	0.003	0.008	NS	NS
	Pdgfr α	214	672	0.746	0.003	0.003	NS	NS
	ND 4	165	658	0.911	0.016	0.024	NS	- 3.336 **
Total	βfib *int*7	283	500	0.886	0.012	0.010	NS	NS
	Cal *int*C	331	472	0.815	0.009	0.009	NS	- 3.521 **
	Pdgfr α	338	672	0.769	0.003	0.004	NS	NS
	ND 4	325	658	0.933	0.027	0.029	NS	- 4.190 **

mtDNA diversity

The haplotype network for 658 bp of the ND4 gene for 324 individuals from 67 populations shows two clear groups separated by nineteen mutational steps (Figure 2). The first group (M) consists of 35 haplotypes present in 144 samples of *T. marmoratus* and 23 of *T. pygmaeus*. The 39 haplotypes of the second group (P) are present in 16 samples of *T. marmoratus* and 141 of *T. pygmaeus*. The outgroup, *T. cristatus*, connects with the M group at 95 mutational steps from H35, which is in turn four mutational steps away from H32, the M-group haplotype closest to the P-group. Some *T. marmoratus* individuals from Areia, Coimbra, Senhora do Almurtão and Vila Velha de Rodão show P haplotypes, while some *T. pygmaeus* from Casal Palomero, Esperança, Gavião, Mora, Nisa, Palomero, Zebreira and Zebreira West exhibit M haplotypes.

Beta fibrinogen diversity

The beta fibrinogen intron 7 dataset consists of 270 sequences (500 bp), from 208 individuals in 53 populations. The haplotype network shows 65 haplotypes divided into three major groups, separated by 2-4 mutational steps (Figure 3). The M-group of 14 haplotypes consists of 83 *T. marmoratus* samples and 12 *T. pygmaeus*, while the other two groups together represent 156 *T. pygmaeus* and 25 *T. marmoratus* samples.

Calreticulin intron c diversity

We obtained 321 sequences of the intron c of calreticulin (472 bp), from 235 individuals of 56 populations. The haplotype network (Figure 4) presents two distinct groups (M and P) separated by five mutational steps. Six haplotypes forming the M group are present in 115 *T. marmoratus* and 4 *T. pygmaeus*, while the 22 P group haplotypes were found in 18 *T. marmoratus* and 183 *T. pygmaeus*.

Platelet derived growth factor receptor alpha diversity

For Pdgfrα, we obtained 332 sequences from 242 individuals of 56 populations (672 bp). The haplotype network consists of 23 haplotypes (Figure 5). There are two groups of haplotypes (M and P) separated by one mutational step. The M group consists of six haplotypes present in 106 *T. marmoratus* and 21 *T. pygmaeus*, while the P group has 17 haplotypes and is represented by 30 *T. marmoratus* and 175 *T. pygmaeus*.

Mutation rates

Mutation rates were 4.528×10^{-3} subst/site/Ma for the mitochondrial gene ND4, 8.758×10^{-4} subst/site/Ma for βfib*int*7, 1.351×10^{-3} subst/site/Ma for Cal*int*C and 1.13×10^{-3} subst/site/Ma for Pdgfrα. The geometric mean of the substitution rates is 1.568×10^{-3} subst/site/Ma or 0.892 subst/Ma.

Isolation with migration analysis

The IM method assumes that the loci considered have not been subject to recombination or to directional or balancing selection. The null hypothesis of no recombination was not rejected by the phi test in any of the markers and neutrality was rejected for Cal*int*C and ND4 in the combined dataset and for ND4 in *T. pygmaeus* (Table 3).

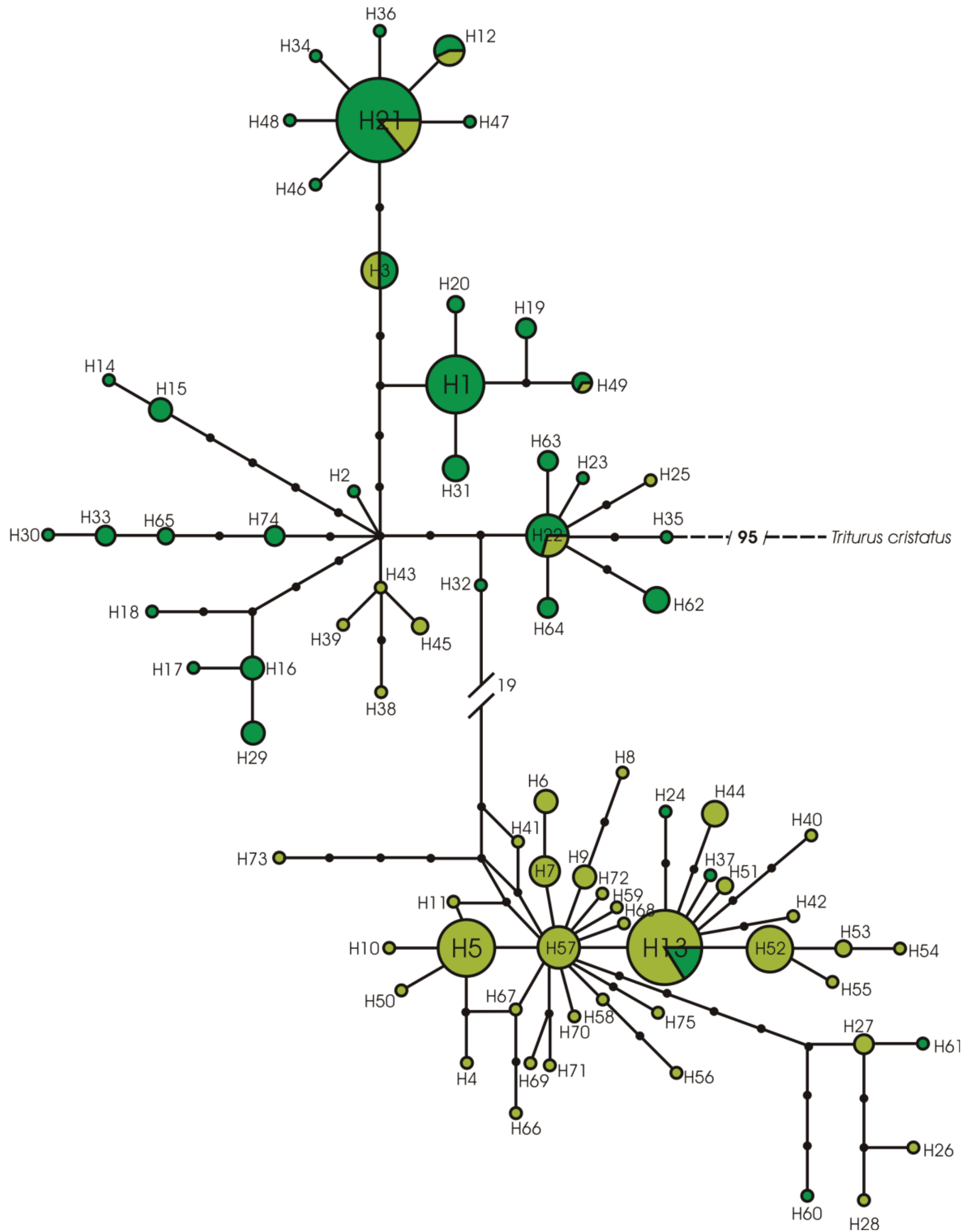


Figure 2 - Haplotype network representing relationship between alleles of the mitochondrial gene ND4 from marbled newts. Size of circles is proportional to frequency of haplotypes. Black circles are missing haplotypes and pie slices represent frequency of the haplotype in *T. marmoratus* (dark colour) and *T. pygmaeus* (light). The number nineteen surrounded by the two slashes in the centre of the figure represents the number of mutations separating the two haplogroups (M and P, see text).

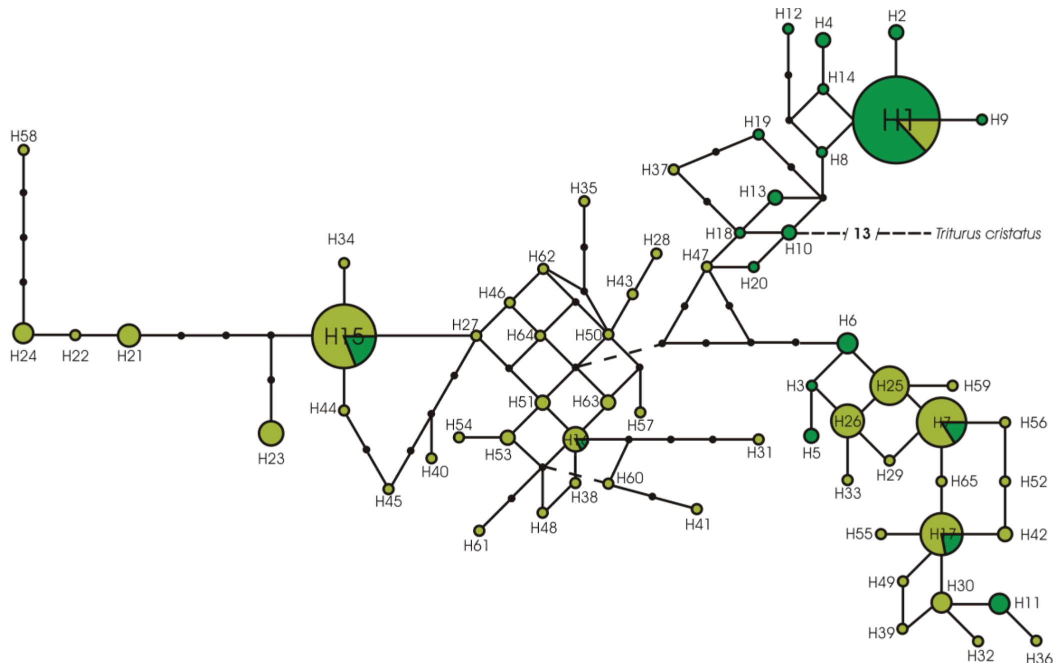


Figure 3 - Haplotype network representing relationship between alleles of the nuclear intron βfib *int*7 from marbled newts**.** Representation follows figure 2. Interrupted line at the centre of the network represents a zero-length connection.

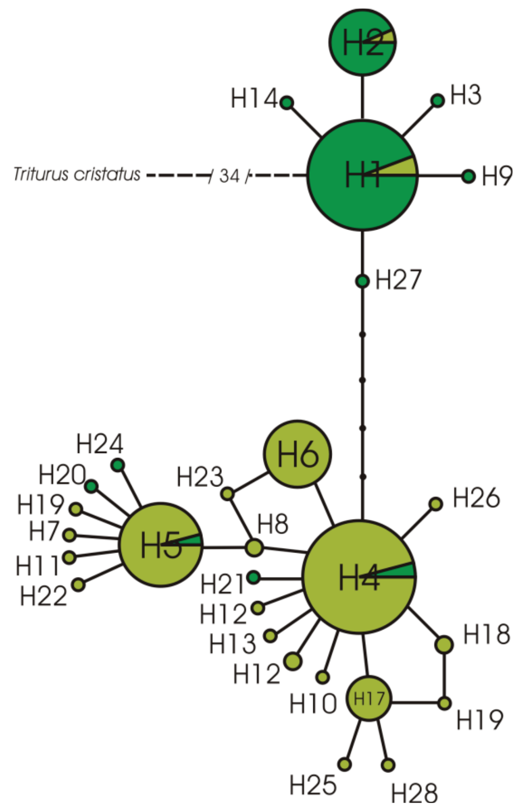


Figure 4 - Haplotype network representing relationship between alleles of the nuclear intron Cal*int*C from marbled newts. Representation follows figure 2.

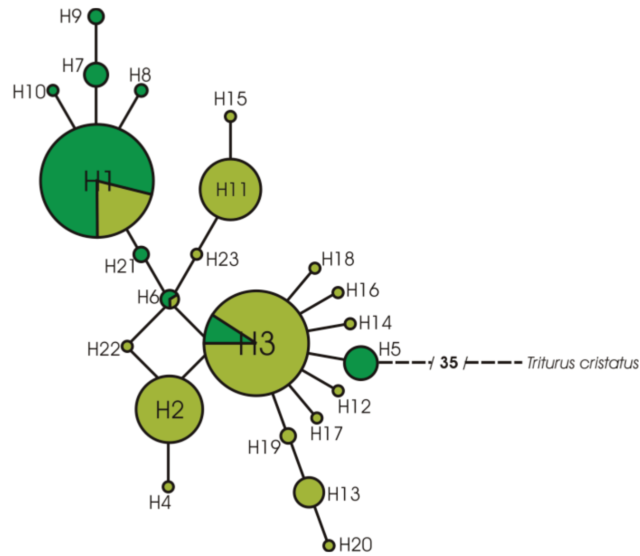


Figure 5 - Haplotype network representing relationship between alleles of the nuclear intron PDGFRα from marbled newts. Representation follows figure 2.

Table 4 summarizes the marginal posterior density curves and maximum likelihood estimates for all parameters in the complete dataset, and the three areas studied. For the ancestral population size and time of divergence, flat curves of the marginal posterior densities without distinct peaks indicate that these parameters could not be reliably estimated from our data. In Table 5, we present the conversion of parameter estimates to demographic quantities.

Table 4 - Results of IMa analysis with effective population size for *T. marmoratus* (q1), *T. pygmaeus* (q2) and the ancestral population (qa), migration rate from *T. pygmaeus* to *T. marmoratus* (m1 migration rate from *T. marmoratus* to *T. pygmaeus* (m2) and the divergence time between them (t).

	Θ1	95% CI	Θ2	95% CI	Θa	95% CI
Coastal	4.321	2.967 - 6.716	7.341	5.362 - 10.256	3.593	0.885 - 88.563
Tejo	1.287	0.585 - 2.833	3.629	1.944 - 7.235	4.987	1.194 - 45.402
Idanha	0.712	0.289 - 2.138	2.562	1.329 - 5.528	8.533	1.483 - 37.351

	m1	95% CI	*m2*	95% CI	*t*	95% CI
Coastal	0.206	0.085 - 0.482	0.001	0.001 - 0.086	2.78	1.74 - 23.18
Tejo	1.625	0.755 - 7.435	0.675	0.205 - 3.685	1.42	1.18 - 37.86
Idanha	0.755	0.385 - 8.825	0.235	0.055 - 2.605	1.26	0.70 - 37.42

Table 5 – Conversion of parameter estimates in Table 4 to demographic quantities.

	N1	N2	Na	2N1m1	2N2m2	t
Coastal	106 364	180 692	88 424	0.444	0.002	1 094 806
Tejo	28 433	86 834	148 308	0.973	1.173	118 144
Idanha	10 746	67 643	187 754	0.180	0.378	527 712

Coastal area

Dispersal estimates (2N*m*) in the coastal area (Table 4) show near-zero (0.002) migration from *T. marmoratus* to *T. pygmaeus* (the hypothesis of $m2 = 0$ was not rejected in an analysis of nested models), and migration significantly different from zero (0.444) in the opposite direction ($m1 = 0$ was rejected at $p<0.05$). Other nested models that are not rejected involve equal values for effective population sizes. Estimated time of divergence shows a distinct peak at approximately 1 094 806 years with a large confidence interval. Effective population (*Ne*) sizes are 106 364 for *T. marmoratus*, 180 692 for *T. pygmaeus* and 88 424 for the ancestral population.

Tejo basin

Migration estimates in both directions are significantly different from zero, with similar migration rates (0.973 vs. 1.173). Examination of nested models of migration rejects models with m1 or m2 = 0, but not models with m1=m2. The estimated time of divergence is 118 144 years. Effective population sizes are estimated at 28 433 for *T. marmoratus*, 86 834 for *T. pygmaeus* and 148 308 for the ancestral population.

Idanha

Migration rates in Idanha are significantly different from zero in both directions. Estimates of migration rates are $m1 = 0.180$ and $m2 = 0.378$. Examination of nested models of migration show the rejection of models where *m1* and/or $m2 = 0$, but not of models with $m1=m2$. Estimated time of divergence is 527 712 years. Effective population sizes are 10 746 for *T. marmoratus*, 67 643 for *T. pygmaeus* and 187 754 for the ancestral population.

Migration over time

Figure 6 shows the distribution of migration events backwards in time in the three areas. Migration takes off at around the same time for the three areas (2.5 Ma). In the coastal area, a plateau is reached around 1.1 Ma followed by an increase of migration in the last 100 ka (thousand years) for m1, while m2 is zero. In the Tejo Basin, migration is the highest until ca. 500 ka when it reduces to similar levels than the other areas. m1 is consistently higher than m2, but the difference is not significant. In Idanha, there is a gradual increase of migration until it reaches a plateau around 600 ka that is maintained until very recently, with m1 consistently higher than m2, but the difference is not significant.

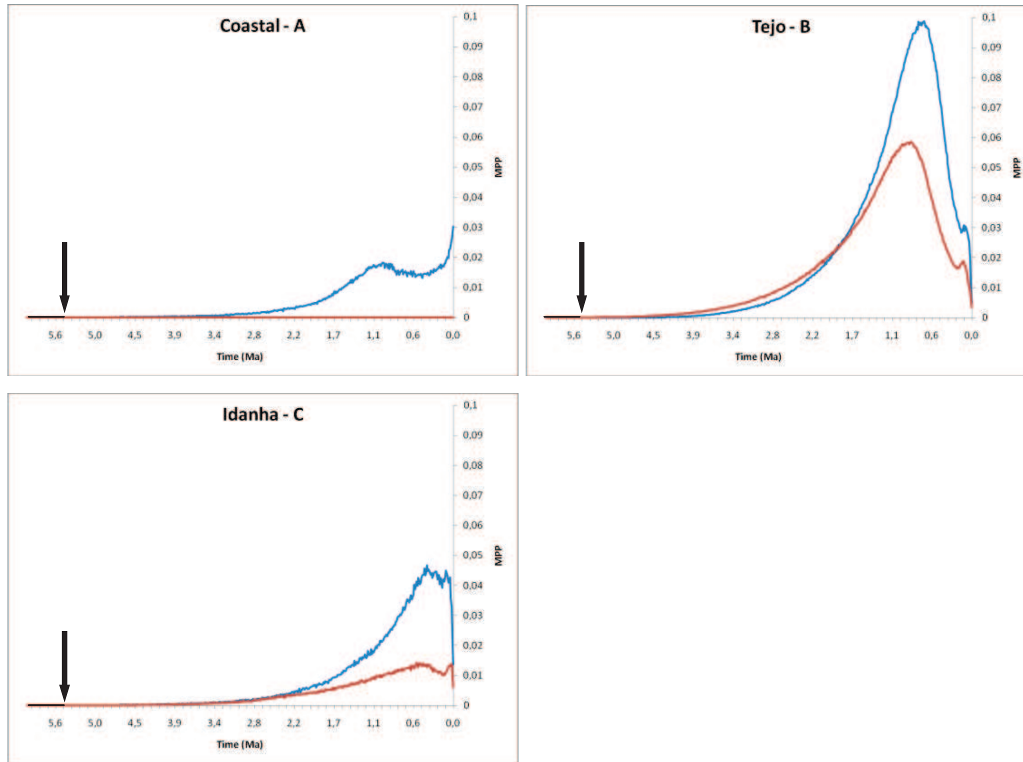


Figure 6 - Temporal distribution of migration times in the three study areas. The horizontal axis indicates time in millions of years (Ma) and the vertical axis denotes marginal posterior probability (MPP). Blue curve indicates *m1* (migration from *T. pygmaeus* to *T. marmoratus*) and red line *m2* (migration in the opposite direction). In the coastal region m2 is not significantly different from zero and hence represented by a straight line. Arrows indicate divergence time based on mtDNA (5.4 Ma; Arntzen *et al.*, 2007).

Discussion

Species divergence and reticulation, assumptions of the method

The estimated time of divergence between *T. pygmaeus* and *T. marmoratus* has large confidence intervals, with a 95% upper bound of the confidence interval of ca. 10 Ma for the coastal region. A phylogenetic analysis of mtDNA estimated that the two species separated around 5.4 ± 2.2 Ma (ARNTZEN *et al.*, 2007). The order in which the interspecific contact was established is estimated as A (Coastal), B (Tejo) and then C (Idanha). Given the large uncertainty we will refrain from making temporal inferences in an absolute sense.

One assumption of IMa is that there are no populations more closely related to the studied populations than the two sampled populations are to one another (Nielsen and Wakeley, 2001). We believe that breaking the assumption will not affect the

results, as in interspecific datasets the genetic substructure within the species is significantly less than between them.

Hybrid zone position and width

Ample evidence suggests that the position of a hybrid zone may be determined by ecological conditions. At smaller scale, also the width of the zone may be affected, suggesting levels of dispersal across the zone to vary with the environment. We studied interspecific gene flow across a newt hybrid zone in Portugal and Spain, for which we selected four areas representing different environmental settings and with different predictions on gene flow. Moreover, in non-equilibrium conditions such as we reconstructed in the coastal area of Portugal where the newt *T. pygmaeus* has been superseding its sister species *T. marmoratus*, gene flow is predicted to be asymmetric. The estimates of dispersal rates and symmetry are indeed different in three out of four selected study areas (data on one area are as yet incomplete).

According to theory, a tension hybrid zone will become fixed in a place such that its length is minimized (BARTON and HEWITT, 1995). Such a situation we encountered at Idanha with no obvious ecological parameter associated to the zone's position other than the Tejo river at ca. 50 km to the southwest and the western fringe of the Central System mountain range at ca. 100 km to the east (CHAPTER 8). The amount of gene flow was expected to be higher here, with no physical barrier, than at either side, with riverine and mountain barriers, respectively, but this was not confirmed by the results. The time plot of migration events, however, shows that the *current* level of gene flow for Tejo and Idanha is similar (Figure 6; no data for area D yet). The origin of the Tejo at its present position started with the lifting of the area and the subsequent incision by the Tejo of its own depositions at the Middle Pliocene, with a second lifting phase at the Late Pliocene (G.-J. Vis, pers. comm.). This would suggest gene flow between the species to be unhampered by the river up to ca. 2 Ma before present. More gene flow between populations at either side of the river then at present may also have been possible in periods of decreased fluvial discharge during periods of climate change in the Pleistocene.

A similar situation was observed in two species of chorus frogs (*Pseudacris ferarium / P. triseriata*) with a hybrid zone at the Ohio basin, USA. The frog species formation is not related to the Ohio river (that is more recent than the frogs) but to

another river more to the north that disappeared in the Middle Pleistocene (Lemmon *et al.*, 2007).

We have reported the mosaic distribution of *T. marmoratus* and *T. pygmaeus* western Portugal, including an isolated patch of *T. marmoratus* near Caldas da Rainha that is characterized by a higher prevalence of terrain used for orchards than the surrounding terrain (Figure 1A; see CHAPTER 9). A straightforward explanation for such a pattern, at least in organisms with low dispersal capability such as amphibians, is the superseding by one species (*T. pygmaeus*) of the other (*T. marmoratus*). The near-absence of genetic footprints of *T. marmoratus* in the areas that changed species occupancy further indicates that the process did not involve much hybridization, while the peculiar distribution pattern of *T. pygmaeus* suggest that the advance has been particularly prominent over the dunal systems of coastal Portugal. With no barrier in place other than the newly reached one at the Vouga estuary in Aveiro, the advance may have been fast and relatively recent. It is therewith interesting to note that the temporal profile of gene-flow shows a recent increase, different from the pattern observed in the other study areas (Fig. 6).

A common feature among the three transects studied so far is the higher level of gene-flow from *T. pygmaeus* into *T. marmoratus* than the other way round (although the difference is only statistically significant in the coastal area). This observation is in line with a general, historical northward advance of *T. pygmaeus*, not stopped by *T. marmoratus*, but by barriers posed by the environment such as the mountains of the Central System in Spain, the river Tejo in central Portugal and also, perhaps more recently, the Vouga estuary at the Atlantic coast. A principal difference between the species is the higher general population size of *T. pygmaeus* than in *T. marmoratus*. This is expressed through generally larger population sizes and denser networks of *T. pygmaeus* than in *T. marmoratus* (unpublished field observations). It is also expressed by the effective population sizes estimates that are consistently higher for *T. pygmaeus* than in *T. marmoratus* (Tables 5 and 6). Typical breeding sites are large and temporary ponds for *T. pygmaeus* and smaller, more permanent ponds, tanks and wells for *T. marmoratus* (CHAPTER 8) that may be more widely distributed over lowland and mountainous terrain, respectively. We predict the northward advance of *T. pygmaeus* to be continued once the riverine barriers of Tejo and Vouga have been taken. Ecogeographical models suggest that in central Portugal the zone will find a

stable position at ca. 40 km to the north (CHAPTER 8). Along the coast there may not be any impediment other than the absence of coastal dunes.

Acknowledgments

This work was financed by Fundação para a Ciência e para a Tecnologia with project grant POCTI/BSE/34110/99, and a PhD grant to GET (SFRH/BD/16894/2004). We thank Geert-Jan Vis (Free University, Amsterdam) for information and discussion on the geological history of the Tejo river.

References

ALEXANDRINO, J., FROUFE, E. , ARNTZEN, J. W. and FERRAND, N. (2000) Genetic subdivision, glacial refugia and postglacial recolonization in the golden-striped salamander, *Chioglossa lusitanica* (Amphibia : Urodela). *Molecular Ecology* **9**, 771-781.

ARNOLD, M.L. (1992) Natural hybridization as an evolutionary process. *Annual Review of Ecology and Systematics* **23**, 237-261.

ARNTZEN, J. W. (1978) Some hypotheses on postglacial migrations of the Fire-bellied Toad, *Bombina bombina* (Linnaeus) and the Yellow-bellied Toad, *Bombina variegata* (Linnaeus). *Journal Biogeography* **5**, 339-345.

ARNTZEN, J. W. (1996) Parameters of ecology and scale integrate the gradient and mosaic models of hybrid zone structure in *Bombina* toads and *Triturus* newts. *Israel Journal of Zoology* **42**, 111-119.

ARNTZEN, J. W. and WALLIS, G. P. (1991) Restricted gene flow in a moving hybrid zone of the newts *Triturus cristatus* and *T. marmoratus* in western France. *Evolution* **45**, 805-826.

BANDELT, H.-J., FORSTER P. and RÖHL A. (1999) Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16**, 37-48.

BARTON, N. H. and HEWITT, G. M. (1985) Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics* **16**, 113-148.

BOŽÍKOVÁ, E., MUNCLINGER, P., TEETER, K. C., TUCKER, P. K., MACHOLÁN, M. and PIÁLEK, J. (2005) Mitochondrial DNA in the hybrid zone between *Mus musculus musculus* and *Mus musculus domesticus*: a comparison of two transects. *Biological Journal of the Linnean Society* **84**, 363-378.

BRUEN, T.C., PHILIPPE H. and BRYANT D. (2006) A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**, 2665-2681.

CAETANO, M. H. and CASTANET, J. (1993) Variability and microevolutionary patterns in *Triturus marmoratus* from Portugal: age, size, longevity and individual growth *Amphibia-Reptilia* **14**, 117-129.

CASTELLA, V., RUEDI, M., EXCOFFIER, L., IBÁÑEZ, C., ARLETTAZ, R. and HAUSSER, J. (2000) Is the Gibraltar Strait a barrier to gene flow for the bat *Myotis myotis* (Chiroptera: Vespertilionidae)? *Molecular Ecology* **9**, 1761-1772.

COSTA, J. C., AGUIAR C., CAPELO J. H., LOUSÃ M. and NETO C. (1998) Biogeografia de Portugal continental. *Quercetea* **1**, 5–56.

COYNE, J. A. and ORR, H. A. (2004) *Speciation*. Sinauer Associates, Inc., Sunderland, MA.

DASMAHAPATRA, K. K., BLUM, M. J., AIELLO, A., HACKWELL, S., DAVIES, N., BERMINGHAM, E. P. and MALLET, J. (2002) Inferences from a rapidly moving hybrid zone. *Evolution* **56**, 741-753.

DIAZ-PANÍAGUA, C., MATEO, J. A. and ANDREU, A. C. (1996) Age and size structure of populations of small marbled newts (*Triturus marmoratus pygmaeus*) from Doñana National Park (SW Spain). A case of dwarfism among dwarfs. *Journal of Zoology* **239**, 83-92.

DRUMMOND, A.J. and RAMBAUT A. (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**: 214.

BENJAMIN M. FITZPATRICK, H. BRADLEY SHAFFER, G. WALLIS (2004) Environment-dependent admixture dynamics in a tiger salamander hybrid zone *Evolution* **58**, 1282-1293.

FRANCILLON-VIEILLOT, H., ARNTZEN, J. W. and GÉRAUDIE, J. (1990) Age, growth and longevity of sympatric *Triturus cristatus*, *T. marmoratus* and their hybrids (Amphibia, Urodela): a skeletochronological comparison. *Journal of Herpetology* **24**, 13-22.

GARCÍA-PARÍS, M., ALCOBENDAS, M. and ALBERCH, P. (1998) Influence of the Guadalquivir River Basin on mitochondrial DNA evolution of *Salamandra salamandra* (Caudata: Salamandridae) from southern Spain. *Copeia* **1998,** 173-176.

GARCÍA-PARÍS, M., ARANO B. and HERRERO P. (2001) Molecular Characterization of the contact zone between *Triturus pygmaeus* and *T. marmoratus* (Caudata:

Salamandridae) in Central Spain and their taxonomic assessment. *Revista Española de Herpetología* **15**, 115-126.

Gascon, C. (1996) Amphibian litter fauna and river barriers in flooded and non-flooded Amazonian rain forest. *Biotropica* **28**, 136-140

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95-98.

Hey, J. and Nielsen R. (2004) Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* **167**, 747-760.

Hayes, F. E. and Sewlal, J. N. (2004) The Amazon River as a dispersal barrier to passerine birds: effects of river width, habitat and taxonomy. *Journal of Biogeography* **31**, 1809-1818.

Jakob, C., Miaud, C., Crivelli, A. J. and Veith, M. (2003) How to cope with periods of drought? Age at maturity, longevity, and growth of marbled newts (*Triturus marmoratus*) in Mediterranean temporary ponds. *Canadian Journal of Zoology* **81**, 1905-1911.

Jehle, R., Arntzen J.W., Burke T., Krupa P. and Hödl W. (2001) The annual number of breeding adults and the effective population size of syntopic newts (*Triturus cristatus*, *T. marmoratus*). *Molecular Ecology* **10**, 839-850.

Jiggins, C. D. and Mallet, J. (2000) Bimodal hybrid zones and speciation *Trends in Ecology and Evolution* **15**, 250-255.

Key, K. H. L. (1968) The concept of stasipatric speciation. *Systematic Zoology* **17**, 14-22.

Kruuk, L. E. B., Baird, S. J. E., Gale, K. S., Barton, N. H. (1999) A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids *Genetics*, **153**, 1959-1971.

Lemmon, E. M., Lemmon, A. R. and Cannatella, D. C. (2007) Geological and climatic forces driving speciation in the continentally distributed trilling chorus frogs (*Pseudacris*). *Evolution* **61,** 2086-2103.

Loureiro, A., Ferrand de Almeida N., Carretero M.A. and Paulo O.S. (2008) *Atlas dos Anfíbios e Répteis de Portugal* Instituto da Conservação da Natureza e da Biodiversidade, Lisboa.

MacCallum, C. J., Nurnberger, B., Barton, N. H. and Szymura, J. M. (1998) Habitat preference in the *Bombina* hybrid zone in Croatia. *Evolution* **52**, 227-239.

Nielsen, R. and, Wakeley, J. (2001) Distinguishing migration from isolation: a Markov Chain Monte Carlo approach. *Genetics* **158**, 885-896.

Pearson, S. F. (2000) Behavioral asymmetries in a moving hybrid zone. *Behavioral Ecology* **11**, 84-92.

Posada, D. and Crandall K.A. (2001) Intraspecific gene genealogies: trees grafting into networks. *Trends in Ecology and Evolution* **16**, 37-45.

Rand, D. M. and Harrison, R. G. (1989) Ecological genetics of a mosaic hybrid zone: mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution* **43,** 432-449.

Rozas, J., Sanchez-DelBarrio J.C., Messeguer X. and Rozas R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**, 2496-2497.

Sequeira, F., Alexandrino, J., Rocha, S., Arntzen, J. W. and Ferrand, N. (2005) Genetic exchange across a hybrid zone within the Iberian endemic golden-striped salamander, *Chioglossa lusitanica*. *Molecular Ecology*, **14**, 245-254.

Slatkin, M. (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics* **16**, 393-430.

Slechtova, V., Bohlen, J., Freyhof, J., Persta, H. and Delmastro, G. B. (2004) The Alps as barrier to dispersal in cold-adapted freshwater fishes? Phylogeographic history and taxonomic status of the bullhead in the Adriatic freshwater drainage. *Molecular Phylogenetics and Evolution* **33**, 225-239.

Steinfartz, S., Vicario S., Arntzen J.W. and Caccone A. (2007) A Bayesian approach on molecules and behavior: reconsidering phylogenetic and evolutionary patterns of the Salamandridae with emphasis on *Triturus* newts. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **308B**, 139-162.

Stephens, M., Smith N.J. and Donnelly P. (2001) A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics* **68**, 978-989.

SZYMURA, J. M. (1993) Analysis of hybrid zones with *Bombina*. in *Hybrid zones and the evolutionary process* (eds. HARRISON, R. G.). Oxford University Press, New York.

SZYMURA, J. M. and BARTON, N. H. (1991) The genetic structure of the hybrid zone between the fire-bellied toads *Bombina bombina* and *B. variegata*: comparisons between transects and between loci. *Evolution* **45**, 237-261.

TAMURA, K., DUDLEY J., NEI M. and KUMAR S. (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599.

YANCHUKOV, A., HOFMAN, S., SZYMURA, J. M., MEZHZHERIN, S. V., MOROZOV-LEONOV, S. Y., BARTON, N. H., NURNBERGER, B. and HARRISON, R. (2006) Hybridization of *Bombina bombina* and *B. variegata* (Anura, Discoglossidae) at a sharp ecotone in western Ukraine: comparisons across transects and over time. *Evolution* **60**, 583-600.

CHAPTER 7

MOLECULAR IDENTIFICATION OF MARBLED NEWTS AND A JUSTIFICATION OF SPECIES STATUS FOR *TRITURUS MARMORATUS* AND *T. PYGMAEUS*

Espregueira Themudo, G.[1,2] and J.W. Arntzen[1]

[1] National Museum of Natural History, P. O. Box 9517, 2300 RA Leiden, The Netherlands.
[2] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661
Vairão, Portugal.

Contents

Abstract 122
Introduction 123
Material and methods 123
Results 125
Discussion 128
Acknowledgements 132
References 132

Published in *Herpetological Journal* **17** 24-30 (2007).

Abstract

The marbled newts *Triturus marmoratus* and *T. pygmaeus* are common and well diagnosable species occurring in central Portugal, but difficult to survey in large and deep ponds. Conversely, embryos of both species are easy to locate but morphologically indistinguishable. We studied a panel of nuclear genetic loci by starch gel electrophoresis (the enzymes PEP-A, PEP-B and PEP-D) and isoelectric focusing techniques (the enzyme LDH-2, post-embryonic stages only) that together yield a species-specific signature (Cohen's kappa = 1.00). On a locus by locus basis the scores for correct classification range from kappa 0.12 to kappa 0.97. The method allows the reliable, fast and cheap identification of both species across life-stages, with a better behaviour and performance than mtDNA sequencing (i.e. bar-coding) and nuclear DNA microsatellite profiling. The observed distribution of *T. marmoratus* and *T. pygmaeus* over 25 aquatic breeding sites in the Caldas da Rainha area in western Portugal is parapatric, with no mixed populations and no F_1 interspecific hybrids. This demonstrates that *T. marmoratus* and *T. pygmaeus* are genetically isolated, even when populations are within the 'dispersal distance per generation' range of one another. We consider the data adequate for supporting the species status of *T. marmoratus* and *T. pygmaeus* under the Biological Species Concept.

Keywords: allozymes; Amphibia, principal coordinate analysis

Introduction

Newts are characterized by long annual periods (2-6 months) of pond breeding and dip-netting provides a mostly efficient and unbiased sampling technique for adults (Plethodon, ARNTZEN, 2002a; 2002b). This allows the reliable surveying over large areas for the purpose of, for example, the improved understanding of distribution patterns and the assessment of conservation status. In Portugal, we noted the frequent occurrence of newts in 'albercas'. These albercas are deep (> 3 m) and sizable (diameter 3-8 m), mostly circular stone structures for agricultural and household water supply [for typical examples of 'albercas', see Malkmus (1982; 2004: Figure 61). Albercas are exceedingly difficult to investigate. The dip-netting approach fails, because at the first sweep the adult newts dive beyond reach and hide in the crevices of the stone walls.

The presence of adults however may be revealed by the easy to spot eggs that are individually attached to the leaves of submerged, floating and overhanging vegetation (MIAUD, 1993). For the eggs to be useful in surveying, a method is required for species identification.

Aims of the present paper are, first, to develop an efficient and reliable method for marbled newts (*Triturus marmoratus* and *T. pygmaeus*) egg identification. Since the eggs of both species are morphologically indistinguishable these will be molecular genetic tools. On a technical note, most eggs in the field will be fertilized and should be referred to as embryos, but for convenience we will use the terms interchangeably. Secondly, we apply and test the new method to qualify the distribution of both species in central Portugal over and across a contact zone between them. Thirdly, we test the hypothesis of Garcia-Paris *et al*. (2001) that *Triturus marmoratus* and *T. pygmaeus* are full species under the Biological Species Concept.

Material and methods

Research was carried out in an area of *c.* 4000 km^2 around Caldas da Rainha to the north of Lisboa, Portugal, where *T. marmoratus* and *T. pygmaeus* have both been observed in a pilot study (JWA & E. Froufe, unpublished data). Ponds and other potential newt breeding sites were located by motorized field searches, assisted by military topographical maps and with help and information by local inhabitants. On

the spot we checked the submerged, floating and marginal vegetation for the presence of marbled newt eggs, which are clearly distinguishable from that of other amphibians, by size, structure and the way they are deposited.

The eggs of the sympatric small bodied newt *Triturus boscai* (placed in the genus Lissotriton by García-París *et al.*, 2004; and in the genus Lophinus by Litvinchuk *et al.*, 2005) are readily distinguished from marbled newt eggs on account of their small size, bipolar pigmentation and the round (as opposed to ovoid) shape of the jelly capsule around them. Eggs were collected from the vegetation all over the accessible parts of the water body and placed in Eppendorf vials. Occasionally, adult and larval marbled newts were caught by dip-netting. The adults were identified as *T. marmoratus* or *T. pygmaeus* on the basis of size, colour and colour-pattern, whereas larvae remained unidentified. Tail tips were removed and placed under buffer in individual Eppendorf vials. All vials were placed in liquid nitrogen for transportation to the laboratory and then stored in an -80º C freezer for future electrophoresis.

In the laboratory, the tail tips and entire embryos were homogenized in an aliquot amount of ice cold buffer (100 mM Tris, 1 mM EDTA, 0.05 mM NADP, adjusted to pH=7.0 with HCl) and centrifuged for 15 minutes at 13,000 rpm at 4 ºC. The supernatant was treated with dithiotreitol (120 μM DTT) for 1 hour at 37 ºC prior to electrophoresis and staining on starch gels for three peptidases (PEP-A, PEP-B and PEP-D) and on acrylamide gels with isoelectric focussing for the enzyme Lactate dehydrogenase (LDH-2), following standard protocols (e.g., Pinho *et al.*, 2003). Electromorphs were interpreted as alleles at the corresponding genetic locus. We used the program Genepop (Raymond and Rousset, 1995) to test for population genetic differentiation by Fisher's exact test, to calculate expected heterozygosity (H_e) and to test for departure from Hardy-Weinberg equilibrium (HW) under standard Bonferroni correction.

Principal Coordinate Analysis (PCA) was performed on a binary data set with alleles as characters and presence (1) or absence (0) of alleles as character states. Character states were assumed to be independent, although in reality limited to a maximum of two scores of 1 per locus. Homozygotes were not distinguished from heterozygotes (i.e., they were represented by a single score of 1). The subroutine SIMQUAL of the program NTSYS 1.7 (Rohlf, 1992) was used to compare the enzyme profiles and to calculate a matrix with pairwise similarity coefficients. We

choose the Jaccard similarity coefficient because it ignores joint absences. The subroutine DCENTER was used to transform the similarity matrix into scalar product form, after which it was factored using the subroutine EIGEN.

Results

The presence of marbled newts was confirmed in 25 aquatic sites, from which we sampled 101 eggs and embryos, 41 larvae and 84 adults. In 14 sites we failed to catch any adults, either because of timing (adults had left the water) or because the site had inaccessible parts. Tissue samples from adults and larvae were scored for four loci (with few exceptions) and embryos were scored for the three peptidase loci. The observed number of alleles was four at *Pep-A*, three at *Pep-B*, eight at *Pep-D* and three at the *Ldh-2* locus. The allele frequencies are presented in Table I. Observed genetic heterozygosity averaged at 0.20±0.11. A significant departure from Hardy-Weinberg equilibrium was observed for *Pep-D* in the sample from Valado dos Frades (population 4, P<0.05). No significant genetic differentiation was found between cohorts in any population. *Pep-A* showed significant population differentiation within *T. marmoratus* and *Pep-B* and *Pep-D* showed significant population differentiation within *T. pygmaeus* (P<0.001 in all three cases).

The PCA scores fell in two non-overlapping groups with values <-0.22 and >0.11, that we term the 'M'-group and the 'P'-group, respectively. Seven adult *T. marmoratus* had a genetic make-up that placed them in the M-group and 77 adult *T. pygmaeus* had a genetic make-up that placed them in the P-group. Common alleles with diagnostic properties are *Pep-D*^{d} and *Pep-D*^{f} and, to a lesser extent, *Pep-A*^{a} and *Pep-A*^{d}. Individuals heterozygous for the *Pep-D* diagnostic alleles were observed in the populations from Alqueidão (population 3), São Bartolomeu dos Galegos (5), Genrinhas (15), Santa Susana (16) and Fonte da Pena da Couvinha (23) (n=1 in all five cases). Less common and rare alleles associated with either group are *Pep-A*^{b}, *Pep-A*^{e}, *Pep-D*^{a}, *Pep-D*^{b} and *Ldh-2*^{f} in the M-group and *Pep-D*^{e}, *Pep-D*^{g}, *Pep-D*^{h}, *Ldh-2*^{c} in the P-group. Alleles shared between the groups are *Pep-B*^{bde}, *Pep-D*^{c} and *Ldh-2*^{b} (Table I). Correct classification on a locus-by-locus basis is very good for *Pep-D* (k=0.97) and *Pep-A* (k=0.92), moderate for *Ldh-2* (k=0.56) and poor for *Pep-B* (k=0.12), in the terminology of Altman (1991).

Table I.
Allele frequencies over four loci in Marbled newts from the Caldas da Rainha area, Portugal.
Values in parentheses are the electrophoretic mobilities relative to the most common allele that is shown as 100.
Six alleles present elsewhere in Portugal were not encountered in the Caldas da Rainha area.
Analysis of the results indicate the existence of two separate genetic units that correspond to *T. pygmaeus* (17 populations) and *T. marmoratus* (eight populations) respectively (details see text)..
Panels at the bottom present heterozygosity on the assumption of Hardy-Weinberg equilibrium and average score at the first PCA axis (details see text), with and without the locus Ldh-2.

		Triturus pygmaeus																	T. marmoratus							
Population	Total	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Sample size	226	13	19	12	22	6	8	8	9	8	1	7	10	8	2	7	3	4	9	8	10	21	10	7	4	10
adults	84	0	19	12	22	0	5	0	0	1	1	7	10	0	0	0	0	0	4	0	0	1	0	2	0	0
larvae	41	5	0	0	0	6	0	0	0	0	0	0	0	0	0	0	3	0	5	0	1	20	0	1	0	0
embryos	101	8	0	0	0	0	3	8	9	7	0	0	0	8	2	7	0	4	0	8	9	0	10	4	4	10
Locus and allele																										
Pep-A																										
a	(122).																		0,56	0,83	0,85	0,91	0,95	1,00	1,00	1,00
b	(107).																						0,05			
d	(100).	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	0,44	0,17	0,15	0,05				
e	(78).																					0,05				
Pep-B																										
b	(113).		0,13	0,29	0,07	0,08	0,19						0,15				0,17		0,11	0,06		0,07				
d	(100).	0,38	0,76	0,46	0,66	0,92	0,75	0,81	1,00	0,94	1,00	0,93	0,60	1,00	1,00	0,79	0,83	1,00	0,89	0,94	0,95	0,88	1,00	0,93	0,88	0,75
e	(88).	0,62	0,11	0,25	0,27		0,06	0,19		0,06		0,07	0,25			0,21					0,05	0,05		0,07	0,13	0,25
Pep-D																										
a	(125).																		0,06			0,14		0,07		
b	(121).																		0,06							
c	(117).								0,50										0,06							
d	(112).			0,04		0,08										0,07	0,17		0,78	1,00	1,00	0,86	1,00	0,93	1,00	1,00
e	(106).		0,11	0,04	0,18		0,25			0,50																
f	(100).	0,96	0,84	0,71	0,71	0,92	0,56	0,94	0,50	0,50	1,00	1,00	1,00	0,94	1,00	0,79	0,83	1,00	0,06							
g	(95).	0,04	0,05		0,11			0,06						0,06		0,14										
h	(88).			0,21			0,19																			
Ldh-2																										
b	(100).	1,00	1,00	0,96	1,00	1,00	1,00			1,00		1,00	1,00				0,83		0,56		1,00	0,45		1,00		
c	(72).			0,04																						
f	(36).																0,17		0,44			0,55				
Three loci																										
He		0,19	0,23	0,38	0,32	0,11	0,35	0,15	0,18	0,22	0,00	0,05	0,20	0,04	0,00	0,25	0,22	0,00	0,38	0,14	0,12	0,22	0,03	0,10	0,08	0,13
SE on He		0,15	0,12	0,20	0,16	0,06	0,18	0,10	0,18	0,16	0,00	0,05	0,20	0,04	0,00	0,13	0,11	0,00	0,09	0,09	0,08	0,02	0,03	0,05	0,08	0,13
Weighted average										0.214±0.132												0.163±0.063				
Four loci																										
He		0,05	0,09	0,25	0,19	0,08	0,25				0,00	0,04	0,15				0,25		0,39			0,26		0,13		
SE on He		0,05	0,06	0,15	0,13	0,05	0,15				0,00	0,04	0,15				0,08		0,11			0,07		0,13		
Weighted average										0.150±0.100												0.284±0.084				
Average score along first PCA axis																										
Three loci		0,34	0,32	0,28	0,29	0,28	0,28	0,32	0,25	0,25	0,33	0,32	0,33	0,33	0,33	0,28	0,23	0,33	-0,44	-0,60	-0,66	-0,68	-0,73	-0,72	-0,72	-0,69
Four loci		0,41	0,37	0,36	0,35	0,33	0,37				0,37	0,36	0,39				0,23		-0,36			-0,57		-0,51		

The distribution of the two groups was spatially structured to the extent that, first, all ponds yielded either M- or P-group individuals and not both and second, ponds in the centre of the study area had M-group individuals whereas ponds at the fringe had P-group individuals. The average distance to the nadir point of the study area was 12.8±4.3 km for M-ponds and 14.0±7.5 km for P-ponds. Note that aspects of the spatial distribution of *T. marmoratus* and *T. pygmaeus* in western Portugal will be dealt with separately (CHAPTER 9).

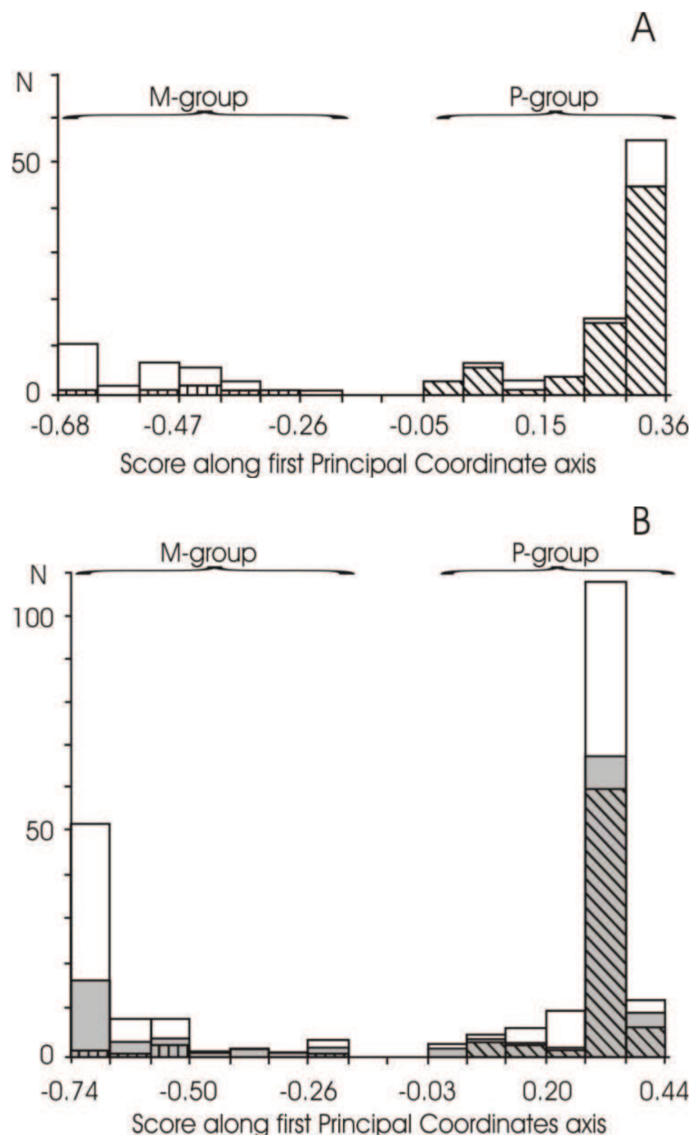


Figure 1. Histograms representing the scores along the first principal coordinate axis for enzyme genetic markers in marbled newts from the Caldas da Rainha area in western Portugal, with four enzyme loci studied in 39 larvae and 82 adults (A) and three enzyme loci studied for an additional 101 embryos, two larvae and two adults (B). Individuals in B also figuring in A are marked by grey shading. Adults identified from morphology are shown by horizontal hatching (*Triturus marmoratus*, *n*=6) and diagonal hatching (*T. pygmaeus*, *n*=76). Note that these fall into different, nonoverlapping M- and P- groups, respectively.

Discussion

We are interested in the distribution and ecology of amphibians from the Iberian Peninsula, with the particular aim to elucidate those environmental correlates that help to reconstruct, explain, predict and understand species ranges (e.g., TEIXEIRA *et al.*, 2001; ARNTZEN, 2006). This requires extensive surveying based upon reliable species identification. In Portugal and Spain, as in most other temperate regions, amphibians gather in ponds and streams for breeding which, by choosing the time and place of the fieldwork advantageously, facilitates the gathering of data. In practice, most surveys concentrate on offspring, because adult pond presence may be short, as in 'explosive breeders' (e.g. *Rana temporaria*), species that mate on land and only come to the water for offspring deposition (e.g. *Salamandra salamandra*), or species that are especially secretive (e.g. *Pelodytes punctatus*). As a rule, however, the earlier the life stage, the more problematic identification in the field proves to be. Classical keys for identification of eggs, spawn and larvae (HÉRON-ROYER and BAMBEKE, 1889; BOULENGER, 1891) have recently been upgraded (MIAUD and MURATET, 2004) and most modern field guides will include identification keys for adults, larvae and eggs (e.g., FERRAND DE ALMEIDA *et al.*, 2001; DUGUET and MELKI, 2003).

We set out to develop a molecular marker technique for the unambiguous species identification of marbled newt eggs that would allow all ponds to be investigated, including technically problematic ones such as albercas and lakes. In ponds with aquatic vegetation absent, rare or out of reach, collecting may be facilitated by the introduction of strings of thin plastic liner available for egg-deposition, cut from e. g. garbage bags. With the phenotypic identification of adult *T. marmoratus* and *T. pygmaeus* as a reference and acknowledging the equivalent allelic expression among embryos, larvae and adults, the observed 'M' and 'P' enzyme profiles can be equated with *T. marmoratus* and *T. pygmaeus*, respectively. The single case of departure from Hardy-Weinberg equilibrium, with a lower than expected number of heterozygotes, may well be attributed to a less than optimal resolution of PEP-D zymogram, perhaps caused by interference of the enzyme and the mucous components in larval tissue.

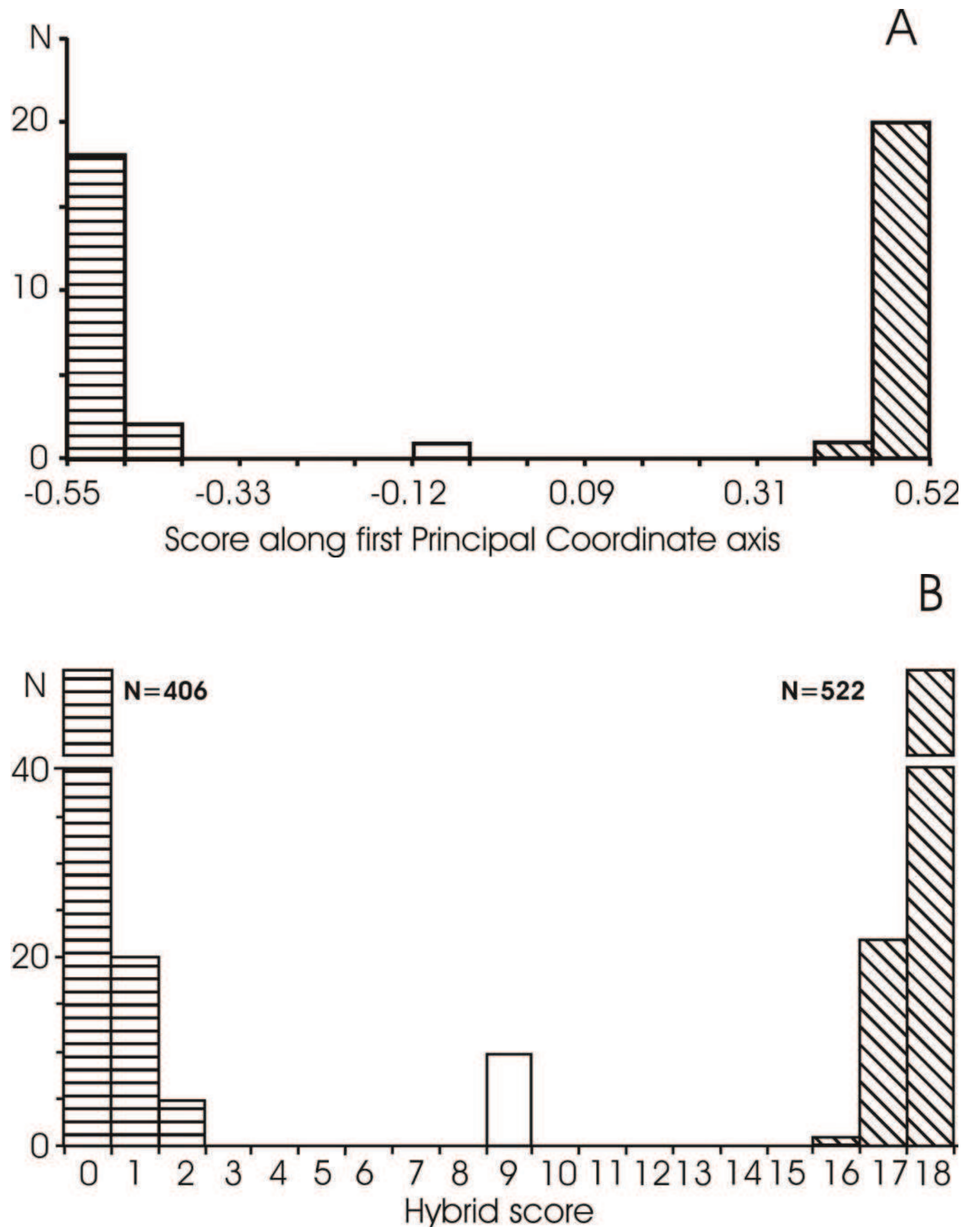


Figure 2. Histograms representing the scores along the first principal coordinate axis derived from 87 alleles in a panel of 30 variable enzyme genetic markers in the newts *Triturus helveticus* and *T. vulgaris* in Mayenne, western France (A: after ARNTZEN *et al.*, 1998) and hybrid index based on nine diagnostic enzyme genetic markers for the newts *T. cristatus* and *T. marmoratus* in the same area (B: after ARNTZEN and WALLIS, 1991).

The method of *T. marmoratus* versus *T. pygmaeus* identification through expressed proteins loci is lethal when applied to embryos. However, considering the high fecundity of marbled newts (ARNTZEN and HEDLUND, 1990) we presume that the effect will be negligible at the population level. Tissue sampling by clipping tail-tips has beenshown no ill effect in adult big-bodied newts (ARNTZEN *et al.*, 1999). The

effect of tail-tip sampling of larvae has not yet been studied. For applying our method of species identification in other parts of the Iberian Peninsula a note of warning is in place, since we observed geographic variation in the level of discrimination achieved by Pep-B and Pep-D (unpublished data). This reservation is in line with the observed population genetic differentiation in the Caldas da Rainha area for both species. An alternative technique for species identification would be the DNA sequencing of a mitochondrial gene such as COI, currently known as 'bar-coding' (HEBERT *et al.*, 2003; VENCES *et al.*, 2005). An important shortcoming of this marker is that it - in a phylogenetic sense - may not represent the species from which it is isolated correctly, due to hybridisation or incomplete lineage sorting, in combination with maternal inheritance and low effective population size compared to nuclear genetic markers. This phenomenon, that ultimately may involve the complete 'mtDNA-capture' by one species from the other, appears fairly frequent in salamanders. Discordance between the signature of mitochondrial and nuclear genetic markers has been found in various genera of plethodontid (e.g., Batrachoseps, WAKE and JOCKUSCH, 2000) and salamandrid salamanders (e.g., Salamandra, GARCÍA-PARÍS *et al.*, 2003; Taricha, KUCHTA and TAN, 2005; Chioglossa, SEQUEIRA *et al.*, 2005). Well-documented cases in big-bodied newts (genus *Triturus*) include i) the near-complete bias for *T. cristatus* mothered hybrids in *T. marmoratus* x *T. cristatus* interspecies hybrids in western France (ARNTZEN and WALLIS, 1991) and ii) the presence of mtDNA typical of *T. karelinii* in *T. carnifex* and *T. dobrogicus* over a large area of northern Serbia (WALLIS and ARNTZEN, 1989; ARNTZEN and WALLIS, 1999). An example among small bodied newts, genus *Triturus* (or *Lissotriton* or *Lophinus*) is the replacement across the entire Carpathian mountain range of the original *T. montandoni* mtDNA by that of *T. vulgaris* (BABIK *et al.*, 2005). Under the notion that flawed inferences from mtDNA may not be infrequent, its choice as a species marker was in this study *a priori* rejected. Nuclear microsatellite DNA markers have been successfully used to uncover genetic variation in *T. marmoratus* (JEHLE *et al.*, 2001; KRUPA *et al.*, 2002; JEHLE *et al.*, 2005) but in our experience, it is not easy to isolate and amplify nuclear DNA from freshly deposited embryos. This technique may require a larger number of copies of nuclear DNA than is available in this life stage, the one most frequently observed in the field and its application would involve raising the embryos in the laboratory.

The contact zone between *T. marmoratus* and *T. pygmaeus* runs over c. 600 km across central to western Iberia, from approximately Madrid in central Spain to north of Lisboa in Portugal. None of the 25 ponds around Caldas da Rainha had a mixed population and individuals with intermediate enzyme profiles were not found (Fig. 1). This suggests the absence in our sample of F_1-interspecies hybrids. On the other hand, the pattern of allozyme discrimination here revealed is flat U-shaped (Fig. 1), rather than more sharply |_|- shaped as in the *T. helveticus* - *T. vulgaris* and *T. cristatus* - *T. marmoratus* situations (Fig. 2). The extent to which the shape of the curve represents incomplete diagnosticity of the enzyme genetic markers (as in *T. marmoratus* and *T. pygmaeus* and *T. helveticus* - *T. vulgaris*) versus gene flow has yet to be determined.

The minimum observed distance between *T. marmoratus* and *T. pygmaeus* populations in this study was 3.3 km. This contrasts to observations in Spain where the recorded minimum distance between the species was c. 26 km between Cilleros and Zarza la Mayor in western Spain and c. 6 km between Hoyo de Manzanares and Villalba in central Spain (García-París *et al.*, 2001). With just three localities of *T. marmoratus* and eight localities of *T. pygmaeus* recorded in the province of Madrid (García-París *et al.*, 1993) both species are locally rare and the contact zone between them has presumably deteriorated, effectively forming a residual contact zone, sensu Szymura (1993).

In the section of the *T. marmoratus* - *T. pygmaeus* distribution here considered, the contact zone between the species is firmly parapatric. Moreover, the data point to the complete or near-complete genetic isolation of the taxa, therewith supporting their specific status under the Biological Species Concept. The taxonomic change was first put forward by García-París *et al.* (2001) and implemented by e.g. Frost (2004); Montori *et al.* (2005). Our study differs from theirs in the following respects: i) a fully diagnostic panel of nuclear genetic markers, ii) utilization of the mitochondrial genetic marker rejected, iii) larger number of populations (25 versus 6), iv) small minimum inter-pond distances (3.3 km versus 6.0 km) and v) not part of a residual contact zone. All too frequent taxonomic and nomenclatorial change is proposed on the basis of a single type of data, including cases that rely on mtDNA data exclusively (e.g. Carlia, Couper *et al.*, 2005; Salamandrina, Mattoccia *et al.*, 2005; Plethodon, Mead *et al.*, 2005). This contrasts with studies that integrate evidence from two or more sources, such as morphology, allozymes, mtDNA and

nuclear DNA (e.g. Calotriton, CARRANZA and AMAT, 2005; Scaphiophryne, GLOS *et al.*, 2005; Salamandrina, NASCETTI *et al.*, 2005; Hyla,SALDUCCI *et al.*, 2005) .

Acknowledgements
We thank N. Ferrand for support and discussion. The work was carried under licence from the ICN (Instituto de Conservação da Natureza) in accordance to National law for capturing wild fauna and was financed by FCT (Fundação para a Ciência e Tecnologia) research project POCTI/34110/99.

References

ALTMAN, D. G. (1991) *Practical Statistics for Medical Research* Chapman & Hall/CRC.

ARNTZEN, J. W. (2002a) Seasonal variation in sex ratio and asynchronous presence at ponds of male and female *Triturus* newts. *Journal of Herpetology* **36**, 30-35.

ARNTZEN, J. W. (2002b) Testing for Equal Catchability of Triturus Newts by Dip Netting. *Journal of Herpetology* **36**, 272-276.

ARNTZEN, J. W. (2006) From descriptive to predictive distribution models: a working example with Iberian amphibians and reptiles. *Frontiers in Zoology* **3**, 8.

ARNTZEN, J. W., P. DE WIJER, R. JEHLE, E. SMIT and J. SMIT (1998) Rare hybridization and introgression in smooth and palmate newts (Salamandridae: *Triturus vulgaris* and *T. helveticus*). *Journal of Zoological Systematics & Evolutionary Research* **36**, 111-122.

ARNTZEN, J. W. and L. HEDLUND (1990) Fecundity of the newts *Triturus cristatus*, *T. marmoratus* and their natural hybrids in the relation to species coexistance. *Holarctic Ecology* **13**, 325-332.

ARNTZEN, J. W., A. SMITHSON and R. S. OLDHAM (1999) Marking and tissue sampling effects on body condition and survival in the newt *Triturus cristatus*. *Journal of Herpetology* **33**, 567-576.

ARNTZEN, J. W. and G. P. WALLIS (1991) Restricted gene flow in a moving hybrid zone of the newts *Triturus cristatus* and *T. marmoratus* in western France. *Evolution* **45**, 805-826.

ARNTZEN, J. W. and G. P. WALLIS (1999) Geographic variation and taxonomy of crested newts (*Triturus cristatus* supersecies): morphological and mitochondrial data. *Contributions to Zoology* **68**, 181-203.

BABIK, W., W. BRANICKI, J. CRNOBRNJA-ISAILOVIC, D. COGALNICEANU, I. SAS, K. OLGUN, N. A. POYARKOV, M. GARCIA-PARIS and J. W. ARNTZEN (2005) Phylogeography of two European newt species- discordance between mtDNA and morphology. *Molecular Ecology* **14**, 2475-2491.

BOULENGER, G. A. (1891) A synopsis of the tadpoles of the European batrachians. *Proceedings Zoological Society* **16**, 593-627.

CARRANZA, S. and F. AMAT (2005) Taxonomy, Biogeography and evolution of *Euproctus* (Amphibia: Slamandridae), with the resurrection of the genus *Calotriton* and the description of a new endemic species from the Iberian Peninsula. *Zoological Journal of the Linnean Society* **145**, 555-582.

COUPER, P. J., J. W. WILMER, L. ROBERTS, A. P. AMEY and G. R. ZUG (2005) Skinks currently assigned to Carlia aerata (Scincidae: Lygosominae) of north-eastern Queensland: a preliminary study of cryptic diversity and two new species. *Australian Journal of Zoology* **53**, 35-49.

DUGUET, R. and F. MELKI (2003) *Les amphibiens de France, Belgique et Luxembourg* Editions Biotope, Mèze.

FERRAND DE ALMEIDA, N., P. FERRAND DE ALMEIDA, H. GONÇALVES, F. SEQUEIRA, J. TEIXEIRA and F. FERRAND DE ALMEIDA (2001) *Guia FAPAS Anfíbios e Répteis de Portugal* Edićões FAPAS, Porto, Portugal.

FROST, D. R. (2004) Amphibian Species of the World: an Online Reference. Version 3.0 (22 August, 2004). Electronic Database accessible at http://research.amnh.org/herpetology/amphibia/index.html. American Museum of Natural History, New York, USA.

GARCÍA-PARÍS, M., M. ALCOBENDAS, D. BUCKLEY and D. B. WAKE (2003) Dispersal of viviparity across contact zones in Iberian populations of fire salamanders (Salamandra) inferred from discordance of genetic and morphological traits. *Evolution* **57**, 129-143.

GARCÍA-PARÍS, M., B. ARANO and P. HERRERO (2001) Molecular Characterization of the contact zone between Triturus pygmaeus and T. marmoratus (Caudata: Salamandridae) in Central Spain and their taxonomic assessment. *Revista Española de Herpetología* **15**, 115-126.

GARCÍA-PARÍS, M., P. HERRERO, C. MARTÍN, J. DORDA, M. ESTEBAN and B. ARANO (1993) Morphological characterization, cytogenetic analysis, and geographical distribution of the Pigmy Marbled Newt Triturus marmoratus pygmaeus (Wolterstorff, 1905) (Caudata: Salamandridae). *Bijdraden tot de Dierkunde* **63**, 3-14.

GARCÍA-PARÍS, M., A. MONTORI and P. HERRERO (2004) Amphibia: Lissamphibia. In: *Fauna Ibérica*. Consejo Superior de Investigaciones Científicas.

GLOS, J., F. GLAW and M. VENCES (2005) A new species of *Scaphiophryne* from Western Madagascar. *Copeia* **2005**, 252-261.

HEBERT, P. D. N., A. CYWINSKA, S. L. BALL and J. R. DEWAARD (2003) Biological identifications through DNA barcodes **270**, 512.

HÉRON-ROYER, L. F. and C. V. BAMBEKE (1889) Le vestibule de la bouche chez les tètards des batraciens anoures d'Europe: sa structure, ses caratéres chez les diverses espéces. *Archives de Biologie* **9**, 185-309.

JEHLE, R., J. W. ARNTZEN, T. BURKE, P. KRUPA and W. HÖDL (2001) The annual number of breeding adults and the effective population size of syntopic newts (*Triturus cristatus*, *T. marmoratus*). *Molecular Ecology* **10**, 839-850.

JEHLE, R., G. A. WILSON, J. W. ARNTZEN and T. BURKE (2005) Contemporary gene flow and the spatio-temporal genetic structure of subdivided newt populations (*Triturus cristatus*, *T. marmoratus*). *Journal of Evolutionary Biology* **18**, 619-628.

KRUPA, A. P., R. JEHLE, D. A. DAWSON, L. K. GENTLE, M. GIBBS, J. W. ARNTZEN and T. BURKE (2002) Microsatellite loci in the crested newt (Triturus cristatus) and their utility in other newt taxa. *Conservation Genetics* **3**, 87-89.

KUCHTA, S. R. and A. N. M. TAN (2005) Isolation by distance and post-glacial range expansion in the rough-skinned newt, Taricha granulosa. *Molecular Ecology* **14**, 225-244.

LITVINCHUK, S. N., A. ZUIDERWIJK, L. J. BORKIN and J. M. ROSANOV (2005) Taxonomic status of Triturus vittatus (Amphibia: Salamandridae) in western Turkey: trunk vertebrae count, genome size and allozyme data. *Amphibia-Reptilia* **26**, 305-323.

MALKMUS, R. (1982) Die Bedeutung der Brunnen für den Amphibien-Bestand Portugals. *Salamandra* **18**, 205-217.

MALKMUS, R. (2004) Amphibians and reptiles of Portugal, Madeira and the Azores-Archipelago, 448p. ARG Gantner Verlag KG, Ruggel.

MATTOCCIA, M., A. ROMANO and V. SBORDONI (2005) Mitochondrial DNA sequence analysis of the spectaculed salamander, *Salamandrina terdigitata* (Urodela: Salamandridae), supports the existance of two distinct species. *Zootaxa* **995**, 1-19.

MEAD, L. S., D. R. CLAYTON, R. S. NAUMAN, D. H. OLSON and M. E. PFRENDER (2005) Newly discovered populations of salamanders from Siskiyou County California represent a species distinct from *Plethodon stormi*. *Herpetologica* **61**, 158-177.

MIAUD, C. (1993) Predation on newt eggs (Triturus alpestris and T. helveticus): identification of predators and protective role of oviposition behaviour. *JOURNAL OF ZOOLOGY* **231**, 575-575.

MIAUD, C. and J. MURATET (2004) *Identifier les oeufs et les larves des amphibiens de France* Institut National de la Recherche Agronomique, Paris.

MONTORI, A., G. A. LLORENTE, M. A. ALONSO-ZARAZAGA, Ó. ARRIBAS, E. AYLLÓN, J. BOSCH, S. CARRANZA, M. A. CARRETERO, P. GALÁN, M. GARCÍA-PARÍS, D. J. HARRIS, J. LLUCH, R. MÁRQUEZ, J. A. MATEO, P. NAVARRO, M. ORTIZ, V. P. MELLADO, J. M. PLEGUEZUELOS, V. ROCA, X. SANTOS and M. TEJEDO (2005) *Lista patrón actualizada de la Herpetofauna Española: Conclusiones de nomenclatura y taxonomía para las especies de anfibios y reptiles de España* Asociación Herpetológica Española, Barcelona.

NASCETTI, G., F. ZANGARI and D. CANESTRELLI (2005) The spectacled salamanders, *Salamandrina terdigitata* (Lacépède, 1788) and *S. perspicillata* (Savi, 1821): 1) genetic differentiation and evolutionary history. *Rendiconti Lincei Scienze Fisiche e Naturali* **9**, 159-169.

PINHO, C., D. J. HARRIS and N. FERRAND (2003) Genetic polymorphism of 11 allozyme loci in populations of wall lizards (Podarcis sp.) from the Iberian Peninsula and North Africa. *Biochemical Genetics* **41**, 343-359.

RAYMOND, M. and F. ROUSSET (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity* **86**, 248-249.

ROHLF, F. J. (1992) NTSYS-pc numerical taxonomy and multivariate analysis system: version 1.7. *New York: Exeter Software*.

SALDUCCI, M. D., C. MARTY, A. FOUQUET and A. GILLES (2005) Phylogenetic relationships and biodiversity in Hylids (Anura: Hylidae) from French Guiana. *Comptes rendus-Biologies* **328**, 1009-1024.

SEQUEIRA, F., J. ALEXANDRINO, S. ROCHA, J. W. ARNTZEN and N. FERRAND (2005) Genetic exchange across a hybrid zone within the Iberian endemic golden-striped salamander, *Chioglossa lusitanica*. *Molecular Ecology* **14**, 245-254.

SZYMURA, J. M. (1993) Analysis of hybrid zones with *Bombina*. In: *Hybrid zones and the evolutionary process* (ed. HARRISON, RG). Oxford University Press, New York.

TEIXEIRA, J., N. FERRAND and J. W. ARNTZEN (2001) Biogeography of the golden-striped salamander *Chioglossa lusitanica*: a field survey and spatial modelling approach. *Ecography* **24**, 618-624.

VENCES, M., M. THOMAS, R. M. BONETT and D. R. VIEITES (2005) Deciphering amphibian diversity through DNA barcoding: chances and challenges. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**, 1859-1868.

WAKE, D. B. and E. L. JOCKUSCH (2000) Detecting species borders using diverse data sets: examples from plethodontid salamanders in California. *The Biology of Plethodontid Salamanders. Plenum Press, New York, New York, USA*, 95–120.

WALLIS, G. P. and J. W. ARNTZEN (1989) Mithocondrial-DNA variation in the crested newt superspecies: limited cytoplasmic gene flow among species. *Evolution* **43**, 88-104.

CHAPTER 8

ENVIRONMENTAL PARAMETERS THAT DETERMINE SPECIES GEOGRAPHICAL RANGE LIMITS AS A MATTER OF TIME AND SPACE

Arntzen, J. W.[1] and G. Espregueira Themudo[1,2]

[1] *National Museum of Natural History, P. O. Box 9517, 2300 RA Leiden, The Netherlands.*
[2] *CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661*
Vairão, Portugal.

Contents

Abstract 138
Introduction 139
Materials and methods 140
  Biological data 140
  Environmental data 140
  Modelling and analysis 142
Results 144
  Distribution of *Triturus marmoratus* and *T. pygmaeus* in the Iberian Peninsula and characterization of the contact zone from field observations 144
  Ecogeographical models 146
Discussion 146
  If ecological parapatry is the case, have we been able to identify a more or less sharp ecological interface (ecotone) associated with the mutual species border? 148
  Can the results from the more or less intuitive field observations and the formal ecogeographical analysis be reconciled? 149
  Is the species contact zone fixed or moving? 151
Acknowledgements 152
References 153

Published in *Journal of Biogeography* **35** (7), 1177–1186 (2008).

Abstract

Aim To identify the ecological factors associated with the range borders of two amphibian species, to investigate geographical variation in environmental parameters that determine species range limits, and to develop a testable hypothesis on the species' biogeographical history.

Methods We documented the distribution of two species of marbled newts (*Triturus marmoratus* and *T. pygmaeus*) from atlas data, survey data and six latitudinal transects. Countrywide ecogeographical models were constructed from presence only data and a suite of 19 environmental parameters. Local ecogeographical models were constructed along the contact zone.

Results *Triturus marmoratus* and *T. pygmaeus* share a parapatric range border from coastal Portugal to central Spain. Ecogeographical models performed better in descriptive mode (i.e., for the country where it was developed) than in predictive mode (i.e., when extrapolated for the neighbouring country). This suggests the existence of spatial variation for parameters that determine the species range borders. Local models showed good performance in Spain and western Portugal. Parameters dominating the models were climatic (precipitation, temperature and solar radiation), geographical (altitude) and geophysical (soil permeability). Enclaves in the distribution of *T. marmoratus*/*T. pygmaeus* also corresponded with environmental parameters. Poor-fitting local models were found in central Portugal where the species' mutual range border coincides with the River Tejo.

Main conclusions The fit and predictive performance of the ecogeographical models suggest that equilibrium conditions have been reached over most of the contact zone. Parameter selection, field observations and natural history knowledge led us to identify, in terms of reproductive strategy, a stable species with regular, but low, annual recruitment (*T. marmoratus*) versus an opportunistic species with fluctuating 'all or nothing' annual recruitment (*T. pygmaeus*). Ecological parapatry derives from the use of permanent ponds for breeding in *T. marmoratus* versus ephemeral ponds used by *T. pygmaeus*. Enclaves in the distribution show that *T. pygmaeus* has been displacing *T. marmoratus* historically. In central Portugal, where model fit is poor, the advance of *T. pygmaeus* is currently impeded by the River Tejo. We predict that once this barrier has been crossed, the contact zone will move ca. 40 km to the north. Since *T. marmoratus* and *T. pygmaeus* are hybridizing species, we predict the occasional presence of *T. marmoratus* genes in areas that have changed occupancy.

Keywords: contact zone, ecological parapatry, enclave, geographical information system, Iberian Peninsula, newts, presence-only ecogeographical models, *Triturus*

Introduction

Species ranges are finite. They are delimited on all sides by a more or less sudden decrease in population density up to a point where no more individuals are present. One of the most important objectives in ecology is to understand why species exist in one place and not in another. Distributional models are based on (usually) limited records of presence and (presumed) absence, and they predict where species will occur given the correlation with one or a few ecological variables. However, the factors that determine these limits will not be the same over the entire edge of the distribution. A species may be bound by, for example, the lack of appropriate habitat in the north, a competing species in the east, an ocean in the south and a mountain range in the west. Such factors act on a local scale by influencing reproductive success, survivorship or the dispersal ability of individuals and populations. This is especially true when species ranges are large, when the habitat is heterogeneous or when the species have high genetic diversity. A corollary of these observations is that ecological models should not be extrapolated to other areas when significant ecological variability is present.

Many researchers have focused on the change in species distribution through time. This is notable in environmental research where the possible effects of man-made changes to the environment are assessed and predicted (TEIXEIRA and ARNTZEN, 2002; ARAUJO and RAHBEK, 2006; ACEVEDO and RESTREPO, 2008). However, few studies have focused on geographical variation in environmental correlates of species distributions. Those that have done this have reported that geographical variation is limited to, for example, differences between a small number of populations (AYRES and SCRIBER, 1994) or between the northern and southern edge of the range (GROSS and PRICE, 2000). In a study exploring data from a Portuguese atlas of amphibians and reptiles (GODINHO *et al.*, 1999), Arntzen (ARNTZEN) compared the precision of descriptive ecogeographical models for Portugal with the accuracy of the predictions in neighbouring Spain. Well-fitting descriptive models tended to predict distributions rather well. Hence, these models appear to convey biologically meaningful information that improve our understanding of species' ecological requirements and can be used in conservation management (ARNTZEN, 2006). An issue that has not yet been addressed is the presence of regional variation in environmental parameters that determine the geographical range limits of

species. With geographical variation present, model fit in predictive mode would be relatively low, but that does not necessarily indicate poor modelling conditions.

The present paper deals with geographical variation in environmental parameters that determine species range limits. Firstly, we describe the distribution of two species of marbled newt (genus *Triturus*) in the Iberian Peninsula. The species' territories are parapatric (i.e., contiguous and only narrowly overlapping in relation to the vagility of the individuals). This makes it possible to contrast the presence of one species with that of the other without reference to absence data. Secondly, we document the existence of spatial variation in the environmental parameters determining a species' range. Thirdly, we present and discuss a hypothesis on the historical biogeography of Iberian *Triturus* species.

Materials and methods

Biological data

The current knowledge of the distribution of Iberian amphibians is organised in atlases for Spain (PLEGUEZUELOS *et al.*, 2004) and Portugal (LOUREIRO *et al.*, 2008). Both employ the Universal Transverse Mercator grid with a spatial resolution of 10 x 10 km (UTM10). Marbled newts (*T. marmoratus* and *T. pygmaeus*) are recorded in > 2000 grid cells. The Spanish data suggest that the distributions of *T. marmoratus* and *T. pygmaeus* are mutually exclusive and that their ranges are parapatric. For Portugal this is unclear because the data are pooled across species, due to the fact that *T. pygmaeus* was until recently considered a subspecies of *T. marmoratus*. In order to clarify the distributions of *T. marmoratus* and *T. pygmaeus*, we studied six approximately north to south orientated transects, and surveyed the areas around Aveiro and Caldas da Rainha (Fig. 1A). Species identification was based on adult morphology, and the molecular genetic analysis of embryo and larval tissue samples (CHAPTER 7).

Environmental data

The environmental variables considered for ecogeographical modelling are those provided by Teixeira & Arntzen (2002), Real *et al.* (2005) and the EROS program

(Earth Resources Observation and Science - Global Land Cover Characterization, Version 2.0. available at http://eros.usgs.gov/products/landcover.html). Variables that

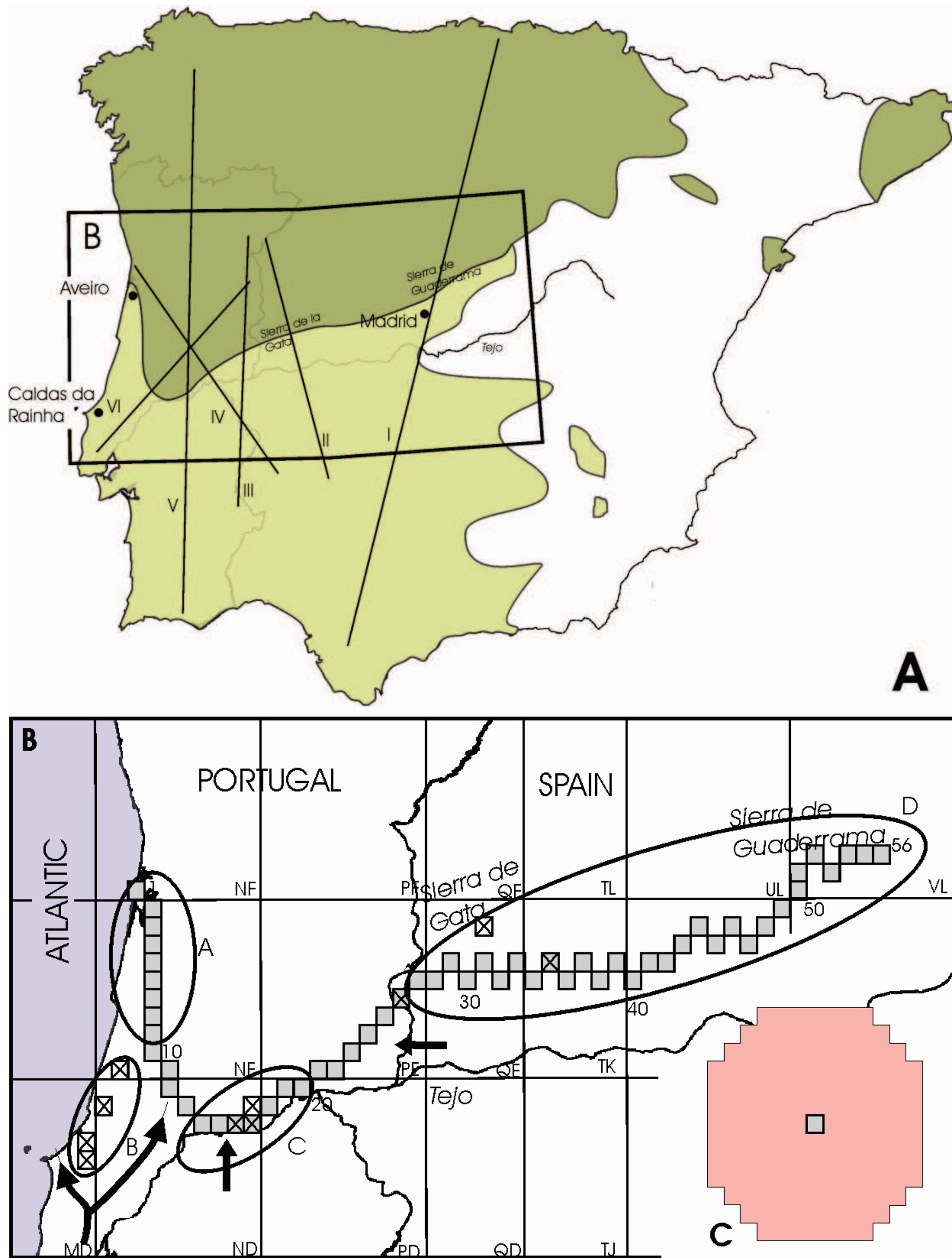


Figure 1 (a) Geographical distribution of the marbled newts Triturus marmoratus (dark shading) and Triturus pygmaeus (light shading) in the Iberian Peninsula. Six latitudinal transects perpendicular to the mutual species boundary (I–VI, continuous lines) showed sharp morphological and genetic transitions. (b) The inferred parapatric range border is shown by shaded UTM10 grid cells. Cells for which both species were actually observed are marked by a cross, including five such cells outside the continuous contact zone. Ellipses A–D refer to areas discussed in the text. (c) Shape and size of a window, the central point of which is moved over the string of shaded grid cells, in 56 steps. The 145 grid cells enclosed in each of the 56 steps were used to build a series of distribution models (see text for details).

were deemed biologically uninformative (e.g., distance to urban centres and highways) or not available for both countries were deleted. A further selection was made with UPGMA cluster analysis, using the criterion of parameter redundancy at the level of Spearman's rank correlation coefficient, $r_s > 0.8$, as follows : the 'mean annual number of frost days' (minimum temperature ≤ 0.5 ºC; code DFRO) was chosen over 'mean temperature in January'; the 'mean relative air humidity in July at 07h00' (HJUL, in %) was chosen over the 'annual relative air humidity range'; the 'mean annual precipitation' (PREC, in mm) was chosen over 'mean annual actual evapotranspiration'; the 'mean annual temperature' (TEMP, in ºC) was chosen over 'mean annual potential evapotranspiration'; and the 'mean temperature in July' (TJUL, in ºC) was chosen over 'mean annual insolation'. Other selected variables were 'altitude' (ALTI, in m a.s.l.); 'mean annual number of days with precipitation ≥ 0.1 mm' (DPRE); 'mean relative air humidity in January at 07h00' (HJAN, in %); 'permeability of the soil' in three categories of increasing permeability (PERM); 'relative maximum precipitation' (RMP, equalling MP24/PREC); 'maximum precipitation in 24 h' (MP24, in mm); 'slope' in degrees, calculated from ALTI (SLOP); 'mean annual solar radiation' (SRAD, in $kWh.m^{-2}.day^{-1}$); and 'annual temperature range' (TRAN, in ºC, equalling 'mean temperature in July' minus its equivalent for January). From EROS we chose the 'Running Vegetation Lifeforms Classification' (RUNNING *et al.*, 1994 7 and 8 pooled) and 12 monthly vegetation index composites from April 1992 to March 1993. With UPGMA, the 12 NDVIs clustered in three seasonal groups from which we chose the April (NDVI04), August (NDVI08) and December (NDVI12) images. The two most dissimilar images were used to calculate an annual NDVI range, as dNDVI = NDVI03 - NDVI12. The 19 variables listed in Table 1 were introduced into the GIS software ILWIS 3.0 (ITC, 2001) as raster layers with 1 km spatial resolution. Mean values for UTM10 grid cells were obtained by averaging the data (modal values were used for the categorical variable RVLC).

Modelling and analysis

Logistic regression was performed with SPSS 14 (SPSS, 2005) with a forward stepwise addition of explanatory variables and with Bonferroni correction to the

initial α=0.05 (HOLM, 1979). The latter, guarded against overfitting the models. For specific models, sets of variables were entered simultaneously. The presence of *T.*

Table 1 Environmental parameters used to model the distribution of Marbled newts on the Iberian Peninsula.

Code	Variable description	Source
ALTI	altitude, in m a.s.l.	1
DFRO	mean annual number of frost days, minimum temperature ≤ 0.5 ºC	1
dNDVI	annual NDVI range, NDVI March 1993 minus NDVI December 1992	2
DPRE	mean annual number of days with precipitation ≥ 0.1 mm	1
HJAN	mean relative air humidity in January at 07h00, in %	1
HJUL	mean relative air humidity in July at 07h00, in %	1
MP24	maximum precipitation in 24 h, in mm	1
NDVI04	monthly vegetation index composite April 1992	2
NDVI08	monthly vegetation index composite August 1992	1
NDVI12	monthly vegetation index composite December 1992	1
PERM	permeability of the soil in three categories of increasing permeability	1
PREC	mean annual precipitation, in mm	1
RMP	relative maximum precipitation, equalling $MP24*PREC^{-1}$	1
RVLC	Running Vegetation Lifeforms Classification	2
SLOP	slope in degrees, calculated from ALTI	1
SRAD	mean annual solar radiation, in $kWh*m^{-2}*day^{-1}$	1
TEMP	mean annual temperature, in ºC	1
TJUL	mean temperature in July, in ºC	1
TRAN	annual temperature range, in ºC, equalling mean temperature in July minus its equivalent for January	1

References : (1) Real *et al*. (REAL *et al.*), (2) Earth Resources Observation and Science Program into Global Land Cover Characterization, Version 2.0. The rare classes 6, 7 and 8 were pooled with class 3 (RUNNING *et al.*, 1994).

marmoratus versus that of *T. pygmaeus* was the (binary) response variable. The numerical impact of the recorded presence of one species versus the other was equilibrated through a weighting variable, using the 'weight cases' option in SPSS, resulting in a balanced dataset with a fifty/fifty representation. Ten cells with both species recorded (<1% of the data) were deleted from the analysis. The continuous and ordered variables were standardized to an average of zero and standard deviation of unity, to increase the comparability of their effects. To test their spatial predictive performance, models were constructed from data in one country and applied to the neighbouring country (from Portugal to Spain and *vice versa*). A significant change in model performance from descriptive to predictive mode would indicate spatial variation in the ecological determinants of the contact zone position. To investigate this phenomenon in detail, a circular window with a surface of 145 grid cells was

moved with the nadir point from cell 1 to cell 56 along the reconstructed contact zone of the species (Fig. 1). Inside the window, marbled newts were presumed present (including 'pseudo-presences') in all cells north (*T. marmoratus*) and south (*T. pygmaeus*) of the contact zone. The strength of agreement among distribution models and distribution data was summarized with 'Area Under the Curve' statistics (PEARCE and FERRIER, 2000; FIELDING and BELL, 2002) and their asymptotic standard errors (SE).

Results

Distribution of *Triturus marmoratus* and *T. pygmaeus* in the Iberian Peninsula and characterization of the contact zone from field observations

We observed a steep transition from *T. marmoratus* in the north to *T. pygmaeus* in the south in each of six transects (Fig. 1A). The abruptness of the transition suggested the absence of wide areas of intergradation and confirmed the essentially parapatric species distribution suggested by the Spanish atlas data. Both species together were observed in five grid cells of the contact zone and in five grid cells outside the contact zone (Fig. 1B: MD95, MD96, ND08, ND87, ND97, ND98, NE10, PE84, QE38 and TK16). Spatial interpolation of the contact zone was straightforward in Portugal. It was somewhat equivocal in data deficient areas of Spain, which is reflected in a sub-optimally resolved ‘zig-zag’ pattern. The overall result allows an analysis in which the presence of one species is contrasted with that of another, without reference to so-called ‘soft’, potentially problematic absence data.

Around Aveiro, *T. pygmaeus* occurs inside and *T. marmoratus* outside the coastal dunes (Fig. 1B, area A). In western Portugal, an exception to strict parapatry was observed north of Caldas da Rainha in area B, where one or more pockets of *T. marmoratus* were found surrounded by *T. pygmaeus*, although this was not resolved at the UTM10 grid scale (The fine scale distribution and environmental analysis is dealt with in CHAPTER 9). In central Portugal, the position of the contact zone appears to be associated with the River Tejo (area C). The local distribution is, however, asymmetric, with *T. pygmaeus* occurring over the river northwards and *T. marmoratus* not crossing it southwards. In Spain, the contact zone runs along the southern edge of the Central System, from the Sierra de Gata (and adjacent Serra da Gata in Portugal) in the west to the Sierra de Guadarrama in central Spain (area D).

The species transition is situated at ca. 1000 m a.s.l. in transect I and at ca. 500 m a.s.l. in transect II.

Table 2 Distribution models for the parapatric range border of Marbled newts (*Triturus marmoratus* and *T. pygmaeus*) by logistic regression of presence-only data.

	Model *						
	1	2	3	4	5	6	7
Distribution data from	Iberia	Portugal	Spain	window			
				1-4	5-13	23-27	28-56
Parameter code							
Constant	9,504	-16,649	33,695	-14,160	-20,287	-1,319	1,831
ALTI				-13,498			0,201
DFRO	2,824	6,729	17,172		-16,468		-4,906
dNDVI	2,309	4,081	3,427				
DPRE	4,106		12,765				
HJAN			-2,679				
HJUL			7,565				
MP24	-5,297		-10,396				-0,124
NDVI04	1,441						
NDVI08							
NDVI12							
PERM	-1,731	-2,969			1,797		1,669
PREC	9,011	7,381	19,495				
RMP	5,301	3,806	11,746	8,837			-1,662
RVLC1	1,578	29,697					
RVLC2	-0,884	30,123					
RVLC4	0,200	31,425					
RVLC5	2,063	30,978					
RVLC3678	0,000	0,000					
SLOP							
SRAD	-12,114	-7,937	-39,183			20,207	
TEMP	2,508		9,913				
TJUL							
TRAN							
Model fit							
Descriptive mode				window			
	Iberia	Portugal	Spain	1-4	5-13	23-27	28-56
AUC	0,996	0,996	1,000	0,995	0,936	0,986	0,970
SE	0,001	0,002	0,000	0,004	0,017	0,006	0,007
Predictive mode	Iberia	Spain	Portugal	all 56 windows			
AUC	not	0,932	0,898	0,822	0,824	0,969	0,825
SE	applicable	0,007	0,016	0,015	0,015	0,006	0,016

Environmental data are standardized, except for the categorical variable RVLC that is represented by four binary variables. The fit of the models to the distribution data is expressed by the 'Area Under the Curve in Receiver Operating Characteristic' plots (AUC) and asymptotic standard error (SE).
* Example - considering data over window 23-27, the probability of occurrence of *T. marmoratus* is estimated as (1/(1+exp(20.207*SRAD-1.319))).

Ecogeographical models

The minimal adequate model (sensu CRAWLEY, 2002) for the Iberia-wide differential *T. marmoratus–T. pygmaeus* distribution is described as model 1 in Table 2. For countries separately, the models were slightly better fitting for Spain (model 3) than for Portugal (model 2). Extrapolating the models to the neighbouring country showed that model fit is poorer in predictive mode than in descriptive mode in both directions – Spain to Portugal and Portugal to Spain. This suggests spatial variation in the variables that determine the species' geographical range limit. We therefore proceeded with 56 models that were constructed by moving a circular window over the contact zone, with the focal point stepping from position 1 to 56 (Fig. 1B and 1C). In those models, the variables ALTI and RMP dominated in the coastal zone of Aveiro. Variables frequently selected were DFRO (29 times) and SRAD (22 times). DFRO was often accompanied by MP24 (10 times) or PERM (11 times). Explanatory variables other than these six entered the models occasionally and were spatially incoherent. We therefore restricted the variable evaluation to the following sets: ALTI+RMP, DFRO, DFRO+MP24, DFRO+PERM and SRAD (Fig. 2). Four models were calculated over areas where variable sets performed best and had high values of fit: ALTI+RMP over windows 1-4 (model 4); DFRO+PERM over windows 5-13 (model 5); SRAD over windows 23-27 (model 6); and ALTI, DFRO, MP24, PERM and RMP over windows 28-56 (model 7; Table 2). No model was formalized for area C, where model fit was low. Models 4-7 performed better in local mode than in spatial extrapolation (Table 2). The *T. marmoratus–T. pygmaeus* interspecific range limit described by the models is shown in Fig. 3.

Discussion

We document the distribution of *T. marmoratus* and *T. pygmaeus* from field surveys and atlas data and from the study of six latitudinal transects. *Triturus marmoratus* and *T. pygmaeus* are shown to have parapatric ranges. No further analysis is required to conclude that the single factor affecting the mutual geographical range limit in central Iberia is the presence of the sister species. This raises the question of what keeps the species separate. Key (1981) distinguished 'hybridization parapatry' and 'ecological parapatry'. In hybridization parapatry, the lack of admixture results primarily from the populations [species] mating more or less freely with each other, but either

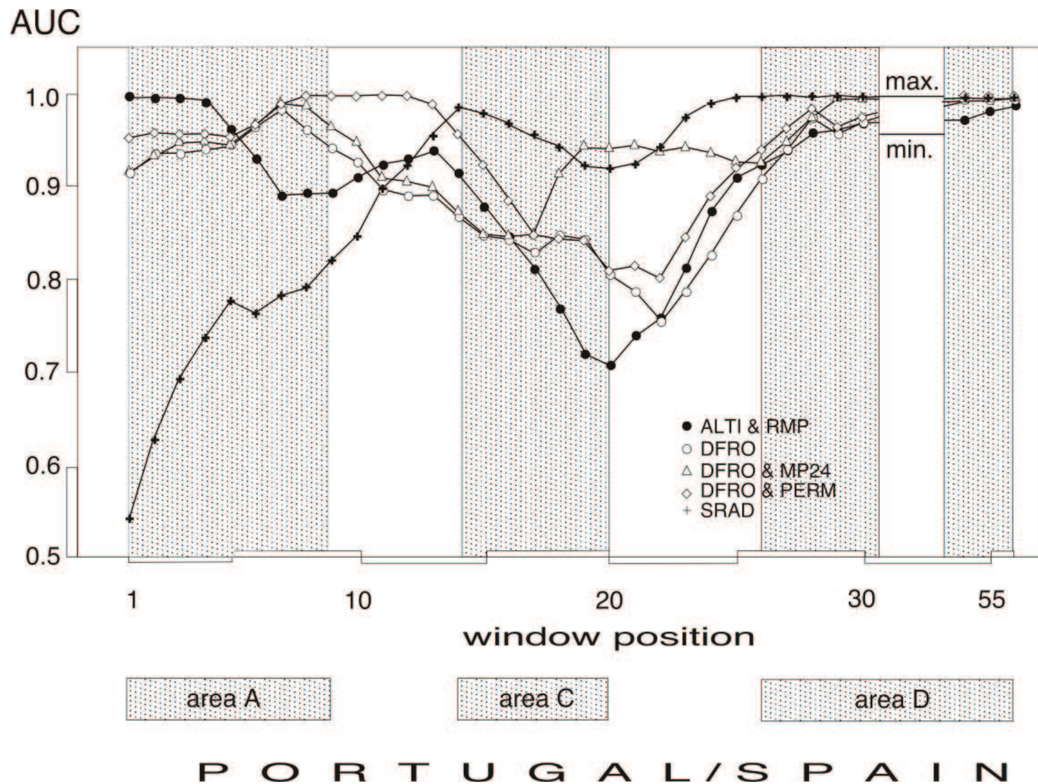


Figure 2. Area under the curve values describing the fit of 56 models constructed with a window (Fig. 1c) moving over the contact zone of Triturus marmoratus and Triturus pygmaeus, from position 1 near Aveiro to position 56 north of Madrid (Fig. 1b). Ecogeographical models were generated with logistic regression for the independent variable Triturus with character states marmoratus range vs. pygmaeus range and selected environmental variables in five sets (see text for details).

leaving no fertile progeny or leaving progeny of reduced fertility. In ecological parapatry, the lack of admixture results primarily from a sharp ecological interface between the habitats of the two populations [species], perhaps reinforced by competition between them. Hybridization between *T. marmoratus* and *T. pygmaeus* has so far been documented in central Portugal (J. W. Arntzen & G. Espregueira Themudo, unpublished data) and perhaps in eastern Portugal also (MALKMUS and LOUREIRO, 2007). One marbled newt of mixed *T. marmoratus–T. pygmaeus* ancestry was found in a survey around Caldas da Rainha in western Portugal (area B in Figure 1; CHAPTER 9). Considering the coverage of the studies, we conclude that hybridization between *T. marmoratus* and *T. pygmaeus* is rare and that the observed distribution pattern does not qualify as hybridization parapatry.

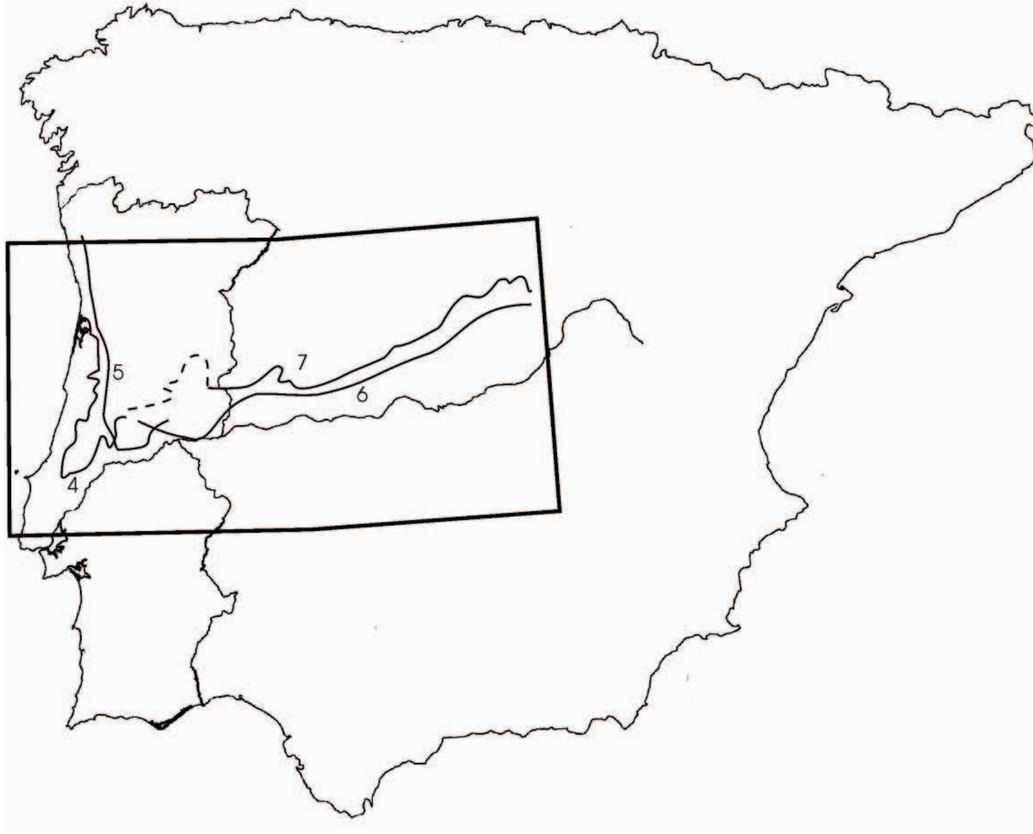


Figure 3. Local limit of the Triturus marmoratus–Triturus pygmaeus differential distribution as modelled from variable sets ALTI + RMP (model 4) and DFRO + PERM (model 5) in western Portugal, SRAD (model 6) in eastern Portugal and adjacent Spain, and ALTI, DFRO, MP24, PERM and RMP in Spain (model 7). The variables are described in Table 1 and the models in Table 2. The dashed line shows the predicted mutual range limits of the species assuming no impeding effect of the River Tejo on the northward range extension of T. pygmaeus. This links model 4 and model 7 and is based on the Iberia-wide model 1.

<u>If ecological parapatry is the case, have we been able to identify a more or less sharp ecological interface (ecotone) associated with the mutual species border?</u>

The field observations and ecogeographical analyses provide us with a clear positive answer, but, interestingly, the variables associated with the position of the mutual species border vary from one region to the other. The primary factors suggested by field observation are soil composition (coastal dunes versus other soil types inland) in area A, the River Tejo in area C, and the southern slopes of the Central Iberian Mountain System over area D. In area B the primary ecological correlate of *T. marmoratus* presence is the agricultural practice of fruit growing (CHAPTER 9). No obvious correlates of the differential *T. marmoratus –T. pygmaeus* distribution were apparent for the regions flanking area C.

From the statistical perspective, the dominating explanatory parameters are ALTI + RMP (model 4) and DFRO + PERM (model 5) in area A. Model 4 has stronger appeal than model 5 because: 1) its local descriptive performance is better (Table 2); 2) it makes significantly better predictions at the Iberian-wide scale (Table 2); 3) it does not describe the presence of *T. pygmaeus* along the Atlantic coast north of Aveiro as does model 5 (Fig. 3); and 4) it is the only model that predicts the presence of *T. marmoratus* over area B (Fig. 3; remember that records for the grid cells MD95, MD96, ND08 and NE10 were excluded from the analysis on the basis of species syntopy at the UTM10 scale). As yet no UTM10 grid cells are known south of the River Tejo with *T. marmoratus* present and *T. pygmaeus* absent. No explanatory parameters are strongly associated with the position of the contact zone in area C (Fig. 2). Model 6 with SRAD is the only model that performs well over windows 23-27 (for which it was designed), but otherwise has little appeal because in western Portugal it performs no better than random data (Fig. 2). The differential *T. marmoratus–T. pygmaeus* distribution along the Central Iberian Mountain System is captured by SRAD in model 6 and by five variables in model 7. These variables are locally strongly correlated and will be difficult to disentangle.

<u>Can the results from the more or less intuitive field observations and the formal ecogeographical analysis be reconciled?</u>

In area A the species breeding localities are drastically different. The dune ponds with *T. pygmaeus* are usually wide, shallow and ephemeral whereas *T. marmoratus* breeding sites outside the dunes are mostly smaller, deeper and more permanent. The dunes represent an unstable system for amphibian breeding. In wet years reproductive success could be great, because of low predation pressure by fish and aquatic insects, but in dry years many ponds desiccate before the larvae reach metamorphosis or do not form at all. We predict that the breeding success of *T. pygmaeus* inside the dunes is subject to large fluctuations and that the annual reproductive output of *T. marmoratus* is comparatively steady. This hypothesis can be tested through a demographic study, for example using the skeletochronological approach for determining individual age (FRANCILLON-VIEILLOT *et al.*, ; CAETANO and CASTANET, 1993; JAKOB *et al.*, 2003). Soil type also differentiates the terrestrial shelter it provides to the (postmetamorphic) juveniles and adults. Shelter may be readily available through active burrowing in sand, but not through the availability of cracks and fissures in hilly and mountainous terrain. The alternative model 4 is composed of

ALTI and RMP. We interpret low altitude as a proxy to the coastal zone and, hence, ALTI as an indirect descriptor of soil structure that locally differentiates between the dune and inland areas. The alternative model 5, composed of DFRO and PERM, also distinguishes the conditions inside the dunes (with a higher than average temperature and permeable soil) and outside the dunes. The steady moisture conditions described by high RMP values are in line with the field based interpretation that point to more constant (reproductive and foraging) conditions for *T. marmoratus* than for *T. pygmaeus*. This variable was particularly successful in predicting the presence of *T. marmoratus* in area B. Area C is characterized by the absence of a strong ecogeographical signal associated with the differential species distribution. The River Tejo (not an explanatory variable in the GIS analysis) locally separates the species, which suggests that the river acts as a barrier to dispersal. In area D, the distribution of *T. pygmaeus* extends northwards into the foothills of the Central Iberian mountain system where *T. marmoratus* occupies the higher slopes. The altitude at which the species transition takes place varies and we observed more permeable soils in the south than in the north. Since we are dealing with a latitudinal divide running along a mountain range, it is no surprise that the species transition is also described by climatic conditions of rainfall (MP24 and RMP), temperature (DFRO) and solar radiation (SRAD). The observed species differential ecology is similar to that in area A, with relatively small and permanent ponds in the mountains, and large, often temporary ponds in the plains accompanied by a soil structure and opportunities for hiding that are also similar to those in area A. The essential difference between the species appears to be an opportunistic life history (r-strategy) of *T. pygmaeus* and a more stable life history of *T. marmoratus* (K-strategy), consistent with their occurrence in habitats providing more fluctuating 'all-or-nothing' and more predictable opportunities for reproduction, respectively. The difference in life histories does not seem to be important in central Portugal where the mutual species border is bounded by the River Tejo. We suggest that reciprocal transplant experiments and laboratory experiments would be required to test our interpretations. The agreement between the field-raised biological interpretation in areas A and D would further suggest that the selected ecogeographical parameters only indirectly relate to the differential *T. marmoratus*–*T. pygmaeus* geographical range limit. It follows that the intuitive field approach and the formal ecogeographical approach have complementary strengths, with respect to a straightforward interpretation vs. a

formal statistical evaluation and spatial extrapolation, respectively. The joint results indicate that in some areas the environment determines the outcome of interspecific competition whereas elsewhere it is a physical barrier (a river) that explains the current distributions. Other studies have dealt with variation in environmental parameters that govern the geographical range limit of species. However, these only compare the northern vs. southern edge of the range (GROSS and PRICE, 2000), coastal vs. oceanic, or presence vs. absence of a parapatric competitor (BULLOCK *et al.*, 2000). Competition is the most likely cause for reciprocal range boundaries in parapatric species, but here we see that local ecological conditions are determining which species has the competitive advantage.

Is the species contact zone fixed or moving?

Species distributions change over time and a contemporary distribution map represents no more, so to speak, than one frame in a film covering the lifetime of a species. Thus, it is difficult to predict the direction and rate of change through time. Nevertheless, the isolated pockets in the easternmost part of the *T. marmoratus* and *T. pygmaeus* ranges have been interpreted as the signal of a receding range (GARCÍA-PARÍS *et al.*). This interpretation is reasonable, because drought conditions in these areas have been increasing (ESTRELA *et al.*, 1996) and newts have limited dispersal abilities that do not allow them to cross large areas of unsuitable habitat. An effective dispersal rate of one $km.y^{-1}$ is a fair estimate for *Triturus* newts (ARNTZEN and WALLIS, 1991; JEHLE and SINSCH, 2007). We used this line of reasoning to explain the isolated occurrence of *T. marmoratus* in western Portugal (area B in Fig. 1B), but the argument is stronger here because the surroundings are occupied by the sister species *T. pygmaeus* (ESPREGUEIRA THEMUDO and ARNTZEN, 2007). The isolated occurrence is effectively an enclave to *T. pygmaeus* (and an exclave to *T. marmoratus*). In species with low dispersal rates such as amphibians, enclaves are a signal of significant changes in the parapatric range border. The existence of this particular enclave is best explained by *T. pygmaeus* displacing *T. marmoratus*. The enclave is a remnant and appears to be associated with more favourable habitat for one species (*T. marmoratus*) and less favourable habitat for the other (*T. pygmaeus*). The most important environmental factor in the enclave is the extent to which the land is used for orchards. The growing of fruit trees in turn is selected for by higher moisture than in the surrounding areas (A. da Costa Pais, Caldas da Rainha, personal

communication). Other examples of enclaves/exclaves for closely related, parapatric species are in *Bombina* toads in central Europe, and *Triturus* newts in western France and the northern Balkans. Ecological correlates of the distribution pattern are altitude and slope in *Bombina variegata* that is surrounded by *B. bombina* (ARNTZEN, 1978;1996), forestation and slope in *T. marmoratus* that is surrounded by *T. cristatus* (ARNTZEN and WALLIS, 1991), and as yet undetermined for a *T. karelinii* enclave that is sandwiched by *T. cristatus* and *T. carnifex* (ARNTZEN and WALLIS, 1999; ARNTZEN, 2003). The best explanation in each of these cases is the movement of the contact zone or hybrid zone in which the surrounded and isolated species is receding.

The distribution pattern of marbled newts on the Iberian Peninsula suggests that *T. pygmaeus* has been advancing at the expense of *T. marmoratus*. Given the change in range border, it further suggests that in central Portugal the advance of *T. pygmaeus* is stopped or impeded by the River Tejo. An essential difference between a river and other landscape elements that could act as barriers to dispersal is the linear aspect of the former. It is reasonable to assume that the strength of the barrier is related to the width of the river and that rivers are semi-permeable barriers. The actual strength of the barrier will be difficult to determine because it is possible that *T. marmoratus* in area C is not displaced from movements of *T. pygmaeus* across the River Tejo, but from the west and from the east (see arrows in Fig. 2B). Yet another possibility is that the newts do not cross the river, but that 'the river crosses the newts' through the redirection of its course or oxbow formation. Once the river has been passed and populations are settled, there is no impediment to further range expansion until the next barrier is encountered. Following this scenario of a northward advance of *T. pygmaeus* we can estimate from the models the position at which its advance will be stopped (Fig. 3, dashed line). If it was not for the barrier effect of the River Tejo, the position of the contact zone would be located ca. 40 km to the north. A corollary of this interpretation is that the *de novo* observation of geographical variation in the ecological parameters determining the range limit of species might reflect the non-equilibrium condition of the species and their environments.

Acknowledgements
We thank Márcia Barbosa (Málaga, Spain) for making available digital copies of environmental data, Elsa Froufe, Armando Geraldes (Porto, Portugal) and Annie Zuiderwijk (Amsterdam, the Netherlands) for assistance in the field, the Walen

family (Nisa, Portugal) for logistic support, Rudolf Malkmus (Wiesthal, Germany) for locality information, the 'Fundação para a Ciência e a Tecnologia' for financial support (project number POCTI/BSE/34110/99) and the 'Instituto de Conservação da Natureza' (Portugal) and the 'Consejarias del Medio Ambiente' (Caceres and Madrid) for permission to handle legally protected newts.

References

ACEVEDO, M.A. and RESTREPO C. (2008) Land-cover and land-use change and its contribution to the large-scale organization of Puerto Rico's bird assemblages. *Diversity and Distributions* **14**, 114-122.

ARAUJO, M.B. and RAHBEK C. (2006) How does climate change affect biodiversity? *Science* **313**, 1396-1397.

ARNTZEN, J.W. (1978) Some hypotheses on postglacial migrations of the fire-bellied toad, *Bombina bombina* (Linnaeus) and the yellow-bellied toad, *Bombina variegata* (Linnaeus). *Journal of Biogeography* **5**, 339-345.

ARNTZEN, J.W. and WALLIS G.P. (1991) Restricted gene flow in a moving hybrid zone of the newts Triturus cristatus and T. marmoratus in western France. *Evolution* **45**, 805-826.

ARNTZEN, J.W. (1996) Parameters of ecology and scale integrate the gradient and mosaic models of hybrid zone structure in Bombina toads and Triturus newts. *Israel Journal of Zoology* **42**, 111-119.

ARNTZEN, J.W. and WALLIS G.P. (1999) Geographic variation and taxonomy of crested newts (*Triturus cristatus* supersecies): morphological and mitochondrial data. *Contribution to Zoology* **68**, 181-203.

ARNTZEN, J.W. (2003) *Triturus cristatus* Superspecies–Kammolch-Artenkreis. In: *Handbuch der Reptilien und Amphibien Europas. Schwanzlurche IIA* (eds. GROSSENBACHER, K., THIESMEIER B.), pp. 421-514. Aula-Verlag, Wiebelsheim.

ARNTZEN, J.W. (2006) From descriptive to predictive distribution models: a working example with Iberian amphibians and reptiles. *Frontiers in Zoology* **3**, 8.

AYRES, M.P. and SCRIBER J.M. (1994) Local Adaptation to Regional Climates in Papilio Canadensis (Lepidoptera: Papilionidae). *Ecological Monographs* **64**, 465-482.

BULLOCK, J.M., EDWARDS R.J., CAREY P.D. and ROSE R.J. (2000) Geographical separation of two Ulex species at three spatial scales: does competition limit species' ranges? *Ecography* **23**, 257-271.

CAETANO, M.H. and CASTANET J. (1993) Variability and microevolutionary patterns in Triturus marmoratus from Portugal: age, size, longevity and individual growth. *Amphibia-reptilia* **14**, 117-129.

CRAWLEY, M.J. (2002) *Statistical Computing: An Introduction to Data Analysis Using S-Plus* John Wiley and Sons.

ESPREGUEIRA THEMUDO, G. and ARNTZEN J.W. (2007) Newts under siege: range expansion of *Triturus pygmaeus* isolates populations of its sister species. *Diversity and Distributions* **13**, 580-586.

ESTRELA, T., MARCUELLO C. and IGLESIAS A. (1996) *Water resources problem in Southern Europe - An overview report.* European Environment Agency, Copenhagen.

FIELDING, A.H. and BELL J.F. (2002) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* **24**, 38-49.

FRANCILLON-VIEILLOT, H., ARNTZEN J.W. and GERAUDIE J. (1990) Age, growth and longevity of sympatric *Triturus cristatus*, *T. marmoratus* and their hybrids(Amphibia, Urodela): A skeletochronological comparison. *Journal of Herpetology* **24**, 13-22.

GARCÍA-PARÍS, M., MONTORI A. and HERRERO P. (2004) Amphibia: Lissamphibia. In: *Fauna Ibérica*. Consejo Superior de Investigaciones Científicas.

GODINHO, R., TEIXEIRA J., REBELO R*., et al.* (1999) Atlas of the continental Portuguese herpetofauna: an assemblage of published and new data. *Revista Española de Herpetología* **13**, 61-82.

GROSS, S.J. and PRICE T.D. (2000) Determinants of the northern and southern range limits of a warbler. *Journal of Biogeography* **27**, 869-878.

HOLM, S. (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* **6**, 1979.

ITC (2001) Ilwis 3.0 for Windows. International Institute for Geo-Information Science and Earth Observation (ITC), Enschede, The Netherlands.

JAKOB, C., MIAUD C., CRIVELLI A.J. and VEITH M. (2003) How to cope with periods of drought? Age at maturity, longevity, and growth of marbled newts

(*Triturus marmoratus*) in Mediterranean temporary ponds. *Canadian Journal of Zoology/Revue Canadienne de Zoologie* **81**, 1905-1911.

JEHLE, R. and SINSCH U. (2007) Wanderleistung und Orientierung von Amphibien: eine Übersicht. *Zeitschrift fur Feldherpetologie* **14**, 137-152.

KEY, K.H.L. (1981) Species, Parapatry, and the Morabine Grasshoppers. *Systematic Zoology* **30**, 425-458.

LOUREIRO, A., FERRAND DE ALMEIDA N., CARRETERO M.A. and PAULO O.S. (2008) *Atlas dos Anfíbios e Répteis de Portugal* Instituto da Conservação da Natureza e da Biodiversidade, Lisboa.

MALKMUS, R. and LOUREIRO A. (2007) Die Herpetofauna der östlichen Beira Baixa (Portugal). *Zeitschrift fur Feldherpetologie* **14**, 191-224.

PEARCE, J. and FERRIER S. (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* **133**, 225-245.

PLEGUEZUELOS, J.M., MÁRQUEZ R. and LIZANA M. (2004) *Atlas y libro rojo de los anfibios y reptiles de España* Organismo Autónomo de Parques Nacionales.

REAL, R., BARBOSA A.M., MARTINEZ-SOLANO I. and GARCIA-PARIS M. (2005) Distinguishing the distributions of two cryptic frogs(Anura: Discoglossidae) using molecular data and environmental modeling. *Canadian Journal of Zoology/Revue Canadienne de Zoologie* **83**, 536-545.

RUNNING, S.W., JUSTICE C.O., SALOMONSON V*., et al.* (1994) Terrestrial remote sensing science and algorithms planned for EOS/MODIS. *International Journal of Remote Sensing* **15**, 3587-3620.

SPSS (2005) SPSS for Windows. SPSS Inc., Chicago.

TEIXEIRA, J. and ARNTZEN J.W. (2002) Potential impact of climate warming on the distribution of the Golden-striped salamander, *Chioglossa lusitanica*, on the Iberian Peninsula. *Biodiversity and Conservation* **11**, 2167-2176.

CHAPTER 9

NEWTS UNDER SIEGE: RANGE EXPANSION OF *TRITURUS PYGMAEUS* ISOLATES POPULATIONS OF ITS SISTER SPECIES

Espregueira Themudo, G. and J.W. Arntzen

[1] National Museum of Natural History – Naturalis, PO Box 9517, 2300 RA Leiden, the Netherlands.

[2] CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Campus Agrário de Vairão, 4485-661 Vairão, Portugal.

Contents

Abstract 158
Introduction 159
Material and Methods 161
  Sampling 161
  Genetic Data 161
  Ecological Data 162
  Model selection 163
  Results 164
  Discussion 167
Acknowledgments 169
References 170
Appendix S1 174
Appendix S2 175

Published in *Diversity and Distributions* **13** (5) 580-586 (2007).

Abstract

The newt species *Triturus marmoratus* and *T. pygmaeus* are both present in central Portugal where they have parapatric distributions. We used four genetic markers to determine which species was present in 31 populations. In the centre of the study area we found a *T. marmoratus* enclave. Despite small inter-population distances, hybridization is locally rare. We built several models to try to explain this distribution using environmental data. The best model, chosen by Akaike's Information Criterion, relates the presence of *T. marmoratus* with the temperature in July, the relief of the landscape and a higher use of the land for orchards. The current distribution can best be explained by *T. pygmaeus* expanding north and replacing *T. marmoratus*, the latter only persisting where ecological conditions are relatively favourable.

Keywords: Akaike's Information Criterion, Allozymes, Amphibia, Enclaves, Mosaic Hybrid Zones, Parapatry

Introduction

Secondary contact between closely related species often generates spectacular events and provides insight into the evolutionary process. Hybrid zones formed this way are considered 'natural laboratories' or 'windows on evolutionary processes' (e.g. HEWITT, 1988 in Anartia butterflies; HARRISON, 1990). Evolution is not a static process and to understand it, we must look into dynamic systems like these.

The dynamics of species ranges are not independent of ecological conditions, competition with sister species, or dispersal capabilities. Present day distributions are echoes from past events. One particular event that provides clues to the relative movement of species is the formation of enclaves. Enclaves are here defined as populations of one species *completely* surrounded by populations of closely related species and genetically isolated from other populations of the same species (ARNTZEN, 1978; in geographical terms, they are simultaneously exclaves and enclaves). This is reminiscent of the 'internal parapatry' concept of Key (1981) that, however, does not deal with disjunct distributions.

Mosaic hybrid zones are bimodal hybrid zones with few hybrids and predominantly parental genotypes present. The contact between the species is more strongly shaped by ecological constraints than by genetic interactions. In a recent review, Jiggins & Mallet (2000) go one step further and suggest that ecology contributes more to speciation than genetic incompatibility. Well known mosaic hybrid zones are for example those in *Gryllus* crickets (RAND and HARRISON, 1989), *Chorthippus* grasshoppers (BRIDLE *et al.*, 2001), and *Mytilus* mussels (BIERNE *et al.*, 2003). Examples of mosaic hybrid zones in salamanders are *Triturus cristatus* and *T. marmoratus* in western France (ARNTZEN and WALLIS, 1991) and *Plethodon cinereus* and *P. shenandoah* in the Appalachian Mountains of North America (JAEGER, 1970; JAEGER, 1971; see also SITES *et al.*, 2004). Although they present patches of populations of one species distributed among patches of the other, these are not all necessarily enclaves because dispersal among patches may be frequent.

To the best of our knowledge the only enclaves recorded in the literature are those for *Bombina* toads in central Europe (ARNTZEN, 1978) and *Triturus* newts in western France and the northern Balkans (ARNTZEN and WALLIS, 1991; ARNTZEN and WALLIS, 1999). Perhaps enclaves are more likely to arise in

organisms with structured populations and low dispersal capability than in organisms that disperse well. Amphibian populations in particular are well-delimited because of their dependence on water for reproduction and they have low individual mobility. Once formed, enclaves will take some time to dissolve, or be stable or disappear by reconnection to the main distribution.

The two species of marbled newts living in the Iberian Peninsula, *Triturus marmoratus* (Latreille, 1800) and *T. pygmaeus* (Wolterstorff, 1908) have a parapatric distribution. Some reports, however, indicated the presence of *T. marmoratus* where only *T. pygmaeus* was expected, near Caldas da Rainha and a spatial-environmental model for the two species suggests that the local conditions may indeed be favourable to *T. marmoratus* (see the southernmost record in Fig. 1c in ARNTZEN, 2006)). This would indicate an area of sympatry or a mosaic distribution. A mosaic distribution would point to differential ecological requirements with patches where the environmental conditions are more suitable for one species than for the other.

An intuitive explanation from fieldwork across the Iberian Peninsula is that *T. pygmaeus* thrives in ephemeral ponds with a fluctuating reproductive output and that *T. marmoratus* thrives in smaller, more permanent water bodies (e.g., springs) with regular but low annual recruitment reaching metamorphosis (J. W. Arntzen, unpublished). To confirm the record at Caldas da Rainha and to learn more about the events that lead to this occurrence and its extent, we conducted a detailed study. Adult marbled newts have clearly distinguishable morphologies while embryos and larvae are difficult to identify. Conversely, embryos and larvae are easy to find at the aftermath of the reproductive season while adults may be elusive. To facilitate a fast, detailed and reliable surveying we employed genetic markers for species identification (CHAPTER 7).

In the present paper, we look into the spatial structure of the distribution of the two species of marbled newts and determine if and what ecological constraints are shaping it. We also analyze the presence of hybrids in our sample.

Material and Methods

Sampling

In April 2003, we searched around Caldas da Rainha, in central Portugal, for water bodies containing marbled newts. The region is characterized by flat dunes on the coast and an undulating agricultural land, with an abundance of orchards in the centre. We located and sampled 31 populations of *Triturus marmoratus* and *T. pygmaeus*, over 1600 km^2. Breeding female newts obtain one or several spermatophores from which eggs are internally fertilized and then deposited one by one, mostly on aquatic vegetation. The process takes place over a ca. two month period.

To increase random collection and to reduce pseudoreplication we screened the entire accessible area of each pond and collected no more than one egg per leaf or two eggs per plant. In marbled newts adult population size can be very small, especially in small water bodies such as springs (SCHOORL and ZUIDERWIJK, 1980; JEHLE *et al.*, 2001; 2005) and this explains why sample size is small on some occasions (N<5 in four populations). Adults and larvae were captured by dip-netting. Tail tips were collected from adults and larvae were sacrificed. All samples were immediately stored in liquid nitrogen and later transferred to -80º C until the day they were analysed.

Genetic Data

All the tissue samples (from n≥10 individuals for most locales) were analyzed for four allozyme loci: peptidase A (*Pep-A*), peptidase B (*Pep-B*), peptidase D (*Pep-D*) and lactate dehydrogenase (*Ldh-2*; this locus is not yet expressed in embryos), that yield a species specific enzyme profile using standard starch gel electrophoresis and isoelectric focusing. The genetic signature is consistent with morphological identification of adult marbled newts (CHAPTER 7).

With the program ARLEQUIN (Version 3.1; EXCOFFIER *et al.*, 2005), we tested for departures from Hardy-Weinberg expectations and linkage disequilibrium. We used the program FSTAT (v. 2.9.3.2; GOUDET, 1995) to calculate F-statistics and, to detect population sub-structuring, we analysed the results with STRUCTURE 2.1 (PRITCHARD *et al.*, 2000). Using a Markov chain Monte Carlo (MCMC) algorithm, STRUCTURE assigns individuals to a population, or jointly to two or more populations, if their genotypes indicate that they are admixed. This is done assuming a model with K populations (where K may be

unknown), where each is characterized by a set of alleles for each locus. Individuals are assigned to populations as to maximize Hardy-Weinberg and linkage equilibrium. We choose for the 'admixture model' because neighbouring populations may interbreed and varied K from 1 to 5 with 10 000 generations as the length of burn-in period and 100 000 MCMC generations after burn-in. Convergence occurred before 10 000 in test runs. The value of λ was inferred by the program. We accepted K as indicating the true number of genetic partitions when the difference in likelihood between two consecutive values of K was maximal.

We investigated the existence of hybrids in our sample by performing a Bayesian model-based clustering as implemented in the software NEWHYBRIDS (ANDERSON and THOMPSON, 2002). This computes the posterior probability that each individual belongs to each of six predetermined classes (pure species A, pure species B, F_1-hybrid, F_2-hybrid, backcross to species A and backcross to species B).

Ecological Data

For the spatial environmental analysis in a Geographical Information System (GIS), we selected 21 ecological parameters following Teixeira *et al.* (2001) and an additional explanatory variable (land surface occupied by orchards, arcsin transformed percentages – ORCH) that appeared locally informative. We used ORCH as a proxy to one or more unidentified variables that - possibly more directly than ORCH - would help to explain the newt distribution. The advantage of the parameter ORCH is that it manifests itself from field observations and that blanket data are readily available (INE, 1999).

For all variables, information was available in digital format for Portugal (DGA, 1995). A vegetation map (normalised difference vegetation index or NDVI) was obtained courtesy of the Royal Dutch Meteorological Institute (KNMI). An altitude map was taken from the internet (http://edcwww.cr.usgs.gov/doc/edchome/ datasets/edcdata.html) and used to produce a relief map by a set of filter operations (ITC, 1997). Maps on the mean January and July temperature were digitalised from the Portuguese climate atlas (SCN, 1974).

A hierarchical clustering based on Spearman’s correlation coefficient was used to evaluate the level at which ecological information appeared redundant.

One variable was selected arbitrarily out of a set of variables that correlated at Spearman's r_s>0.8. Sixteen variables were retained: acidity of the soil (ACID), altitude , chlorates content of subterranean water (CHLO), frost days (FROD), frost months , hardness of subterranean water (HARD), humidity of the air (HUMI), insolation (INSO), lithology (LITH), vegetation index (NDVI), orchard land coverage , mean annual precipitation (PRET), relief (RELI), the sulphate content of subterranean water (SULP), mean annual temperature (TEMP) and mean July temperature (TJUL). To increase the comparability of their effects, all continuous variables were standardized to an average of zero and a standard deviation of one. The variables were introduced into the GIS analytical software as raster layers with 1 km spatial resolution. Values for 10*10 km UTM grids were obtained by averaging the data (modal values for the categorical variable LITH).

Model selection

We used an information-theoretic model selection approach for the statistical analysis of the data (BURNHAM and ANDERSON, 2002). Using our field knowledge, we built several *a priori* models that would explain the current distribution of the two species. By contrasting the presence of one species against the other, we circumvented the inclusion of false absence data. To understand the effect of the variable ORCH, we used our available data on fruit growing in the region as dependent variable, and applied a Stepwise Multiple Regression (SMR) with the same environmental data (all but ORCH) as explanatory variables, using the software SPSS v14 (SPSS, Inc., 2005). We then substituted ORCH by the model derived from this analysis in our *a priori* models, and added them to the list as *a posteriori* models. In a second step, we used the small-sample Akaike's Information Criterion (AICc; BURNHAM and ANDERSON, 2002) to rank the models and chose the best one. AICc is defined as:

$$AIC_c = -2\ln L + 2K + \frac{2K(K+1)}{n-K-1}$$

where ln L is the natural logarithm of the likelihood function, K is the number of parameters from the model, and n is the sample size. Akaike's Information Criterion prevents overfitting the model, by taking the number of parameters into

consideration. To fit the models to the data and obtain the log-likelihoods of the models, we used the logistic regression procedure available in SPSS.

Results

We analyzed 398 individuals (101 adults, 50 larvae and 247 embryos; see allele frequencies in supplementary material – Table S1).

F-statistics show a high degree of population substructuring (Fst) in *Pep-A* (0.84) and *Pep-D* (0.66), intermediate in *Ldh-2* (0.41) and low in *Pep-B* (0.14). The exact test on Hardy-Weinberg expectations showed that population 4 presents a significant departure at one locus (*Pep-D*) under Bonferroni correction. Linkage disequilibrium test are significant at $P<0.05$ in two *T. marmoratus* populations that are neighbouring *T. pygmaeus* populations (population 23 and 31; see Figure 1 and Table S2 in supplementary information).

The model-based clustering method implemented in STRUCTURE 2.1 showed that the most likely number of partitions for the present data is K=2 (Table 1; for details see table S2 in supplementary material). The partitions correspond to *T. marmoratus* and *T. pygmaeus*. Since the likelihood continues to increase with K, we looked into the population structure when K=3, to examine the possibility that additional sub-structuring existed. The only result was that one of the clusters split into two equal parts. Fst values are consistent with the relative diagnostic power of the *loci* (high for *Pep-A* and *Pep-D*, intermediate for *Ldh-2* and low for *Pep-B*). Accurate species identification is independent of the number of loci studied (Figure 1).

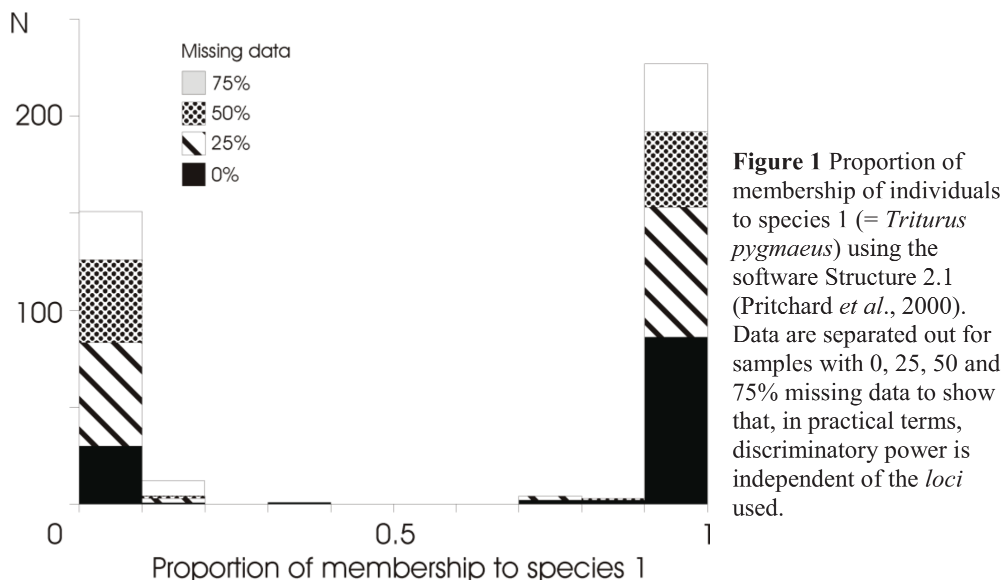


Figure 1 Proportion of membership of individuals to species 1 (= *Triturus pygmaeus*) using the software Structure 2.1 (Pritchard *et al*., 2000). Data are separated out for samples with 0, 25, 50 and 75% missing data to show that, in practical terms, discriminatory power is independent of the *loci* used.

K	logL	SD
1	-1737,7	5,3
2	-961,9	0,6
3	-887,5	1,4
4	-849,4	30,7
5	-802,2	6,9

Table 1 – Log-likelihood (logL) and standard deviation (SD) of the number of partitions (K) in our data set. K varied from 1 to 5. The increase in likelihood is not significant other than at K=2.

The spatial distribution of the two groups indicates the existence of a set of *T. marmoratus* populations surrounded by *T. pygmaeus* populations, forming an enclave in the centre of our study area (Figure 2; note that the isolated pocket is - in proper terms - an exclave of *T. marmoratus* and an enclave of *T. pygmaeus*). Only population 18 had individuals of both species.

The Bayesian-based assignment of individuals to hybrid classes using NEWHYBRIDS indicates that hybridization, backcrossing and introgression are locally rare Thirty out of 31 populations were classified as either *T. marmoratus* or *T. pygmaeus* and not both. Population 18 contained both species as well as one individual with about equal probability of being a pure *T. marmoratus* (p=0.53) as having mixed species parentage (p=0.47, being either a F2-hybrid (P=0.35) or a backcross hybrid in direction of *T. marmoratus* (p=0.12)). This mixed population (population in grey in figure 2) was excluded from the spatial-environmental analysis.

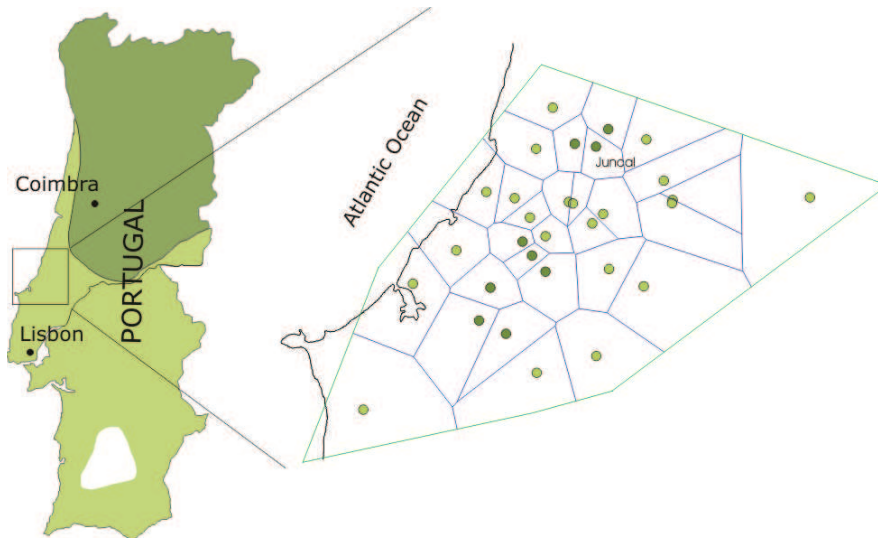


Figure 2 Distribution of marbled newts in Portugal with *Triturus marmoratus* (hatched) and *Triturus pygmaeus* (shaded). The box shows the research area around Caldas da Rainha (a). The detailed map (b) shows the localities with *T. marmoratus* (solid dots) and *T. pygmaeus* (open dots). See Table 1 for population details. Voronoi polygons are used to estimate the contiguous species distribution over the area. Note that the scale of extrapolation at the exterior of the study area is set at *c*. 6 km as to match the level of interpolation.

Table 2 List of *a priori* and *a posteriori* models used to determine the relationship between the distribution of *Triturus marmoratus* and *Triturus pygmaeus* in the Caldas da Rainha area, Portugal. The parameter values were obtained by logistic regression analysis of presence data in the area (by contrasting the presence of one species with the presence of the other). The log likelihood of the models ($-2 \ln L$), number of parameters in each model (K), and the number of data points (n) were used to calculate the small-sample Akaike Information Criterion (AICc). The hypothesis behind each model is also listed. Lowest AICc value is highlighted in bold.

Model	-2 ln L	K	n	AICc	Hypothesis	Model equation												
						ORCH	RELI	TJUL	FROD	ALTI	PRET	HARD	HUMI	NDVI	TEMP	SULP	ACID	Constant
a priori																		
1	25.50	1	30	53.14	Extent of orchard plantations	2.85												-1.23
2	36.87	1	30	75.88	Relief		-1.04											-0.12
3	40.24	1	30	82.62	July temperature			-0.42										0.05
4	41.23	1	30	84.60	Number of frost days				-0.04									0.08
5	24.85	2	30	54.15	Combination of model 1 and model 2	2.57	-0.61											-1.24
6	24.02	3	30	54.96	Model 5 plus altitude	2.78	-0.33			-0.83								-1.49
7	24.74	3	30	56.40	Model 5 plus precipitation	2.50	-0.56				-0.22							-1.21
8	24.77	3	30	56.47	Combination of model 4 and model 5	2.62	-0.57		-0.18									-1.25
9	20.72	3	30	**48.36**	Combination of model 3 and model 5	5.10	-0.40	-2.10										-3.30
10	28.52	6	30	72.70	from Arntzen (unpublished)		-0.98	-0.44				0.16	-1.81	-0.87	-1.88			-0.68
a posteriori																		
11	28.73	5	30	69.97	SMR model					-0.92			-3.07		-2.88	0.79	-0.38	-0.72
12	27.06	6	30	69.76	Model 5 with ORCH replaced by SMR model		-1.06			-0.70			-3.18		-2.22	0.90	-0.75	-0.77
13	26.12	7	30	71.32	Model 7 with ORCH replaced by SMR model		-0.74			0.00	-1.63		-3.69		-2.70	0.24	-0.48	-0.90
14	26.80	7	30	72.70	Model 8 with ORCH replaced by SMR model		-1.14		0.42	-0.92			-3.19		-2.14	1.02	-0.84	-0.83
15	22.82	7	30	64.75	Model 9 with ORCH replaced by SMR model		-2.17	-2.82					-4.53		-2.07	2.24	-0.31	-2.01

SMR, stepwise multiple linear regression; ORCH, orchard land coverage; RELI, relief; TJUL, mean July temperature; FROD, frost days; ALTI, altitude; PRET, mean annual precipitation; HARD, hardness of subterranean water; HUMI, humidity of the air; NDVI, vegetation index; TEMP, mean annual temperature; SULP, sulphate content of subterranean water; ACID, acidity of the soil (details in Teixeira *et al*. 2001 and Arntzen, 2006).

GIS-model 9 (ORCH + TJUL + RELI; see table 3) showed the lowest AICc value. This model explains 50% of the total variance observed. The presence of *T. marmoratus* relates to the higher abundance of orchards, lower July temperature and smoother relief than for *T. pygmaeus*. In the subsequent analysis, five environmental variables: ALTI, ACID, HUMI, SULP and TEMP explained the variable ORCH. This model explained 69% of the variance of ORCH. When, however, ORCH was replaced by this set of variables, the models did not perform as well (Table 2).

Discussion

One of us (JWA) first described the existence of *T. marmoratus* in the research area through a sporadic observation in March 1998. Our results confirm the presence of *T. marmoratus* in six populations outside its documented range in a pocket near Caldas da Rainha. This pocket of *T. marmoratus* is fully surrounded by populations of *T. pygmaeus*. Considering that i) the distance of ca. 10 km that separates it from the main *T. marmoratus* distribution exceeds the dispersal capability of large bodied newts (ARNTZEN and WALLIS, 1991; THIESMEIER and KUPFER, 2000) and ii) that the species are locally strongly parapatric, we conclude that the pocket equals to an enclave.

The variable ORCH has a stronger effect than its fellow explanatory variables in models 5-9 (Table 3). It would be inappropriate though to extrapolate any model with ORCH over wider areas because in Portugal extensive fruit growing is particular to the Caldas da Rainha region.

In the habitat preference model, we assume that strongly preferred habitat is of high quality for the species (RAILSBACK *et al.*, 2003). So, if a model has a good fit, it is usually assumed that the species/habitat system is in equilibrium and that the species distribution will only change if the environment around also changes. This ignores, however, effects of life history and dispersal. Areas may be suitable but out-of-reach, due for example to unsuitable habitat in between realized and prospective ranges.

The current distribution of marbled newts in central Portugal is best explained by *T. pygmaeus* moving north from its previous range and superseding *T. marmoratus* that only persisted in areas with ecological conditions more suitable for the species. Because *T. marmoratus* and *T. pygmaeus* occasionally

hybridize, this scenario can in principle be tested by looking for *T. marmoratus* 'genetic footprints' in areas that are hypothesized to have been inhabited by *T. marmoratus* in the past (ARNTZEN and WALLIS, 1991). Similarly, we predict the absence of *T. pygmaeus* 'genetic footprints' inside the enclave.

There are several accounts of moving hybrid zones in, for example, butterflies (MALLET *et al.*, 1990; BLUM, 2002; DASMAHAPATRA *et al.*, 2002), birds (PEARSON, 2000; ROHWER *et al.*, 2001), crayfish (PERRY *et al.*, 2001), lizards (HILLIS and SIMMONS, 1986) and amphibians (ARNTZEN, 1978; ARNTZEN and WALLIS, 1991); see also Barton & Hewitt (1985; 116-119). Most studies provide direct evidence of hybrid zone movement through the tracing of genetically interacting species distributions over time. Our study utilizes a single temporal window and provides compelling evidence for spatial change in a mosaic hybrid zone nevertheless, through the demonstration of an enclave.

The distance between *T. marmoratus* in the enclave and the main distribution is minimally 6 km (the distance between populations 13 and 15) and maximally 15 km (the distance between populations 22 and 23). Although we cannot pertinently exclude the presence of long distance dispersal, such a scenario is unlikely given the absence of *T. marmoratus* or genetically mixed individuals in populations 13 and 15. Similarly, we cannot exclude the possibility that a human introduction is responsible for the enclave. There is, however, no tradition of newt husbandry in Portugal, and a deliberate or accidental release is improbable.

To infer the direction of the movement in a hybrid zone, it is equally possible to follow a direct or an indirect strategy. Direct demonstrations employ two or more temporally separated observations, either on position (e.g. HILLIS and SIMMONS, 1986 in *Pholidobolus* lizards) or shape of the cline that separates in this case connects the hybridizing species (e.g. DASMAHAPATRA *et al.*, 2002 in *Anartia* butterflies). An indirect way is to look at disequilibrium measures. Cruzan (2005) showed in the wide *Piriqueta coaroliniana/viridis* (flowering plants from the family Turneracea) hybrid zone that *P. viridis* alleles were moving north. This was done by showing that the southern border of the hybrid zone presented relatively high levels of disequilibria, indicating recent gene flow from parental populations south of the hybrid zone. Interestingly, we observed significant levels of linkage disequilibrium in two *T. marmoratus* populations (populations 23 and 31) that are both within the dispersal range of *T. pygmaeus* populations. This strengthens the

argument that *T. pygmaeus* advances into *T. marmoratus* territory, even though we were unable to demonstrate current *T. pygmaeus* presence.

As a further test to our explanation, we predict that the presence of alien *T. marmoratus* genes in *T. pygmaeus* exceeds that of the reverse condition, in a spatial pattern consistent with enclave formation.

Despite the range expansion of *T. pygmaeus* in its northern border, the situation in the south of Spain and Portugal is different. Due to desertification and an intensification of agricultural practices, *T. pygmaeus* is losing many breeding sites and has been classified as 'near threatened' (ARNTZEN *et al.*, 2006). This is consistent with recent evidence that suggests that climate warming will not only increase the northern range of species but also decrease the southern one (THOMAS *et al.*, 2006).

If *T. pygmaeus* continues its competitive advance, the *T. marmoratus* enclave would eventually disappear. As yet, we have no indication on the speed of the process. Monitoring the area would provide an additional test to the hypothesis of *T. pygmaeus* expansion and document the speed of the process. It is remarkable that *T. marmoratus* is not only losing out to its sister-species at the southern edge of its range, but also to the related species *T. cristatus* at its northern edge. The advance of *T. cristatus* at the expense of *T. marmoratus* was estimated to occur at a speed of ca. one km per year (ARNTZEN and WALLIS, 1991). If range replacement would proceed at this speed at either side of its range *T. marmoratus* would be squeezed out in ca. 500 years.

Acknowledgments.
We are grateful to Armando da Costa Pais from Direcção Regional da Agricultura do Alto Oeste for discussion on the relevance of agricultural data for our work. The study was carried under license from the ICN (Instituto de Conservação da Natureza) in accordance to National law for capturing wild fauna and was financed by FCT (Fundação para a Ciência e Tecnologia) research project POCTI/34110/99.

References

ANDERSON, E. C. and E. A. THOMPSON (2002) A Model-Based Method for Identifying Species Hybrids Using Multilocus Genetic Data. *Genetics* **160**, 1217-1229.

ARNTZEN, J. W. (1978) Some Hypotheses on Postglacial Migrations of the Fire-Bellied Toad, Bombina bombina (Linnaeus) and the Yellow-Bellied Toad, Bombina variegata (Linnaeus). *Journal of Biogeography* **5**, 339-345.

ARNTZEN, J. W. (2006) From descriptive to predictive distribution models: a working example with Iberian amphibians and reptiles. *Frontiers in Zoology* **3**, 8.

ARNTZEN, J. W., P. BEJA, J. BOSCH, M. TEJEDO, M. LIZANA, R. JEHLE, I. MARTÍNEZ-SOLANO, A. SALVADOR, M. GARCÍA-PARÍS, E. R. GIL, R. MARQUEZ, P. SÁ-SOUSA and C. D. PANIAGUA (2006) *Triturus pygmaeus*. *In*: IUCN 2008. 2008 IUCN Red List of Threatened Species. www.iucnredlist.org. Acedido em 12 de Novembro de 2008.

ARNTZEN, J. W. and G. P. WALLIS (1991) Restricted gene flow in a moving hybrid zone of the newts *Triturus cristatus* and *T. marmoratus* in western France. *Evolution* **45**, 805-826.

ARNTZEN, J. W. and G. P. WALLIS (1999) Geographic variation and taxonomy of crested newts (*Triturus cristatus* supersecies): morphological and mitochondrial data. *Contributions to Zoology* **68**, 181-203.

BARTON, N. H. and G. M. HEWITT (1985) Analysis of Hybrid Zones. *Annual Review of Ecology and Systematics* **16**, 113-148.

BIERNE, N., P. BORSA, C. DAGUIN, D. JOLLIVET, F. VIARD, F. BONHOMME and P. DAVID (2003) Introgression patterns in the mosaic hybrid zone between Mytilus edulis and M. galloprovincialis. *Molecular Ecology* **12**, 447-461.

BLUM, M. J. (2002) Rapid movement of a *Heliconius* hybrid zone: Evidence for Phase III of Wright's shifting balance theory? *Evolution* **56**, 1992-1998.

BRIDLE, J. R., S. J. E. BAIRD and R. K. BUTLIN (2001) Spatial structure and habitat variation in a grasshopper hybrid zone. *Evolution* **55**, 1832-1843.

BURNHAM, K. P. and D. R. ANDERSON (2002) *Model Selection and Multimodel Inference: A Practical-Theoretic Approach*, 2nd Edition edn. Springer, New York.

CRUZAN, M. B. (2005) Patterns of introgression across an expanding hybrid zone: analysing historical patterns of gene flow using nonequilibrium approaches. *New Phytologist* **167**, 267-278.

DASMAHAPATRA, K. K., M. J. BLUM, A. AIELLO, S. HACKWELL, N. DAVIES, E. P. BERMINGHAM and J. MALLET (2002) Inferences from a rapidly moving hybrid zone. *Evolution* **56**, 741-753.

DGA (1995) *Atlas do Ambiente. Versão 1.1.* Direcção Geral do Ambiente. Ministério do Ambiente e dos Recursos Naturais., Lisboa, Portugal.

EXCOFFIER, L., G. LAVAL and S. SCHNEIDER (2005) Arlequin version 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1**, 47-50.

GOUDET, J. (1995) FSTAT (Version 1.2): A computer program to calculate F-statistics. *Journal of Heredity* **86**, 485.

HARRISON, R. G. (1990) *Hybrid zones: Windows on evolutionary processes.* Oxford University Press, New York.

HEWITT, G. M. (1988) Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution* **3**, 158-167.

HILLIS, D. M. and J. E. SIMMONS (1986) Dynamic change of a zone of parapatry between two species of *Pholidobolus* (Sauria: Gymnophtalmidae). *Journal of Herpetology* **20**, 85-87.

INE (1999) Recenseamento Geral da Agricultura (Available from http://www.ine.Part/prodserv/Rga/rga99_2.asp).

JAEGER, R. G. (1970) Potential Extinction Through Competition Between Two Species of Terrestrial Salamanders. *Evolution* **24**, 632-642.

JAEGER, R. G. (1971) Competitive Exclusion as a Factor Influencing the Distributions of Two Species of Terrestrial Salamanders. *Ecology* **52**, 632-637.

JEHLE, R., J. W. ARNTZEN, T. BURKE, P. KRUPA and W. HÖDL (2001) The annual number of breeding adults and the effective population size of syntopic newts (*Triturus cristatus*, *T. marmoratus*). *Molecular Ecology* **10**, 839-850.

JEHLE, R., G. A. WILSON, J. W. ARNTZEN and T. BURKE (2005) Contemporary gene flow and the spatio-temporal genetic structure of subdivided newt

populations (*Triturus cristatus*, *T. marmoratus*). *Journal of Evolutionary Biology* **18**, 619-628.

JIGGINS, C. D. and J. MALLET (2000) Bimodal hybrid zones and speciation. *Trends in Ecology and Evolution* **15**, 250-255.

KEY, K. H. L. (1981) Species, Parapatry, and the Morabine Grasshoppers. *Systematic Zoology* **30**, 425-458.

MALLET, J., N. BARTON, L. M. GERARDO, S. C. JOSE, M. M. MANUEL and H. EELEY (1990) Estimates of Selection and Gene Flow From Measures of Cline Width and Linkage Disequilibrium in Heliconius Hybrid Zones. *Genetics* **124**, 921-936.

OBSERVATION, I. I. F. G.-I. S. A. E. (1997) Ilwis 2.1 for Windows. ITC, Enschede, the Netherlands.

PEARSON, S. F. (2000) Behavioral asymmetries in a moving hybrid zone. *Behav. Ecol.* **11**, 84-92.

PERRY, W. L., J. L. FEDER, G. DWYER and D. M. LODGE (2001) Hybrid zone dynamics and species replacement between orconectes crayfishes in a northern Wisconsin lake. *Evolution* **55**, 1153-1166.

PRITCHARD, J. K., M. STEPHENS and P. DONNELLY (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics* **155**, 945-959.

RAILSBACK, S. F., H. B. STAUFFER and B. C. HARVEY (2003) What can habitat preference models tell us? Tests using a virtual trout population. *Ecological Applications* **13**, 1580-1594.

RAND, D. M. and R. G. HARRISON (1989) Ecological genetics of a mosaic hybrid zone: Mitochondrial, nuclear, and reproductive differentiation of crickets by soil type. *Evolution* **43**, 432-449.

ROHWER, S., E. BERMINGHAM and C. WOOD (2001) Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. *Evolution* **55**, 405-422.

SCHOORL, J. and A. ZUIDERWIJK (1980) Ecological Isolation in Triturus cristatus and Triturus marmoratus (Amphibia: Salamandridae). *Amphibia-Reptilia* **1**, 235-252.

SCN (1974) *Atlas climatológico de Portugal* Serviço Climático Nacional, Lisboa, Portugal.

SITES, J. W. J., M. MORANDO, R. HIGHTON, F. HUBER and R. E. JUNG (2004) Phylogenetic Relationships of the Endangered Shenandoah Salamander (*Plethodon shenandoah*) and Other Salamanders of the *Plethodon cinereus* Group (Caudata: Plethodontidae). *Journal of Herpetology* **38**, 96-105.

SPSS, I. (2005) SPSS for Windows. SPSS Inc., Chicago.

TEIXEIRA, J., N. FERRAND and J. W. ARNTZEN (2001) Biogeography of the golden-striped salamander *Chioglossa lusitanica*: a field survey and spatial modelling approach. *Ecography* **24**, 618-624.

THIESMEIER, B. and A. KUPFER (2000) *Der kammolch — ein wasserdrache in gefahr.* Beiheft 1 der Zeitschrift für Feldherpetologie. Laurenti-Verlag, Berlin.

THOMAS, C. D., A. M. A. FRANCO and J. K. HILL (2006) Range retractions and extinction in the face of climate warming. *Trends in Ecology & Evolution* **21**, 415-416.

Appendix S1. Allele frequencies in XLS format. Allele frequencies over four loci in marbled newts from the Caldas da Rainha area, Portugal. Analysis of the results indicates the existence of two separate genetic units that correspond to *T. pygmaeus* (22 populations) and *T. marmoratus* (nine populations). Pond numbering and alleles indication are as in Espregueira Themudo & Arntzen (2007). N is sample number; zero sample size refers to missing data.

	Triturus pygmaeus																						*Triturus marmoratus*								
Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	26	27	28	29	30	18	19	20	21	22	23	24	25	31
Adults	0	20	20	22	0	6	0	1	1	1	8	10	1	0	0	0	0	0	0	1	0	0	7	0	1	1	0	2	0	0	0
Larvae	5	0	0	0	6	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	0	0	9	0	1	23	0	2	0	0	0
Eggs	10	0	5	0	0	8	11	10	0	13	0	0	8	2	10	0	11	5	10	10	10	2	0	21	14	0	14	13	18	20	21
Total	15	20	25	22	6	14	11	11	1	14	8	10	10	2	10	3	11	5	10	11	10	2	16	21	16	24	14	17	18	20	21
Locus and alleles																															
Pep-A																															
N	*13*	*19*	*12*	*22*	*6*	*10*	*10*	*10*	*1*	*10*	*7*	*10*	*9*	*2*	*10*	*3*	*10*	*0*	*10*	*10*	*10*	*2*	*9*	*10*	*10*	*22*	*10*	*10*	*11*	*10*	*10*
a																							0,56	0,85	0,85	0,91	0,95	0,95	1,00	1,00	0,60
b																											0,05	0,05			
d	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00		1,00	1,00	1,00	1,00	0,44	0,15	0,15	0,05					0,40
e																										0,05					
Pep-B																															
N	*15*	*20*	*25*	*22*	*6*	*13*	*10*	*10*	*1*	*11*	*8*	*10*	*10*	*2*	*8*	*3*	*4*	*5*	*0*	*2*	*0*	*1*	*16*	*20*	*14*	*24*	*14*	*14*	*11*	*11*	*9*
b		0,13	0,24	0,07	0,08	0,23					0,06	0,15				0,17							0,09	0,03		0,06					0,06
d	0,37	0,78	0,54	0,66	0,92	0,69	0,80	1,00	1,00	0,95	0,88	0,60	0,95	1,00	0,81	0,83	1,00	0,70		1,00		1,00	0,84	0,95	0,96	0,90	1,00	0,96	0,91	0,77	0,89
e	0,63	0,10	0,22	0,27		0,08	0,20	0,00	0,00	0,05	0,06	0,25	0,05		0,19			0,30					0,06	0,03	0,04	0,04		0,04	0,09	0,23	0,06
Pep-D																															
N	*15*	*20*	*24*	*22*	*6*	*12*	*9*	*10*	*1*	*12*	*7*	*10*	*8*	*2*	*8*	*3*	*11*	*5*	*0*	*1*	*2*	*1*	*16*	*20*	*15*	*23*	*14*	*15*	*11*	*20*	*11*
f	0,97	0,85	0,69	0,70	0,92	0,71	0,94	0,50	1,00	0,33	1,00	1,00	0,94	1,00	0,81	0,83	1,00	0,90		1,00	1,00	1,00	0,13					0,03			
d			0,02		0,08										0,06	0,17							0,63	1,00	1,00	0,83	1,00	0,87	1,00	1,00	1,00
e		0,10	0,02	0,18		0,17				0,67																					
g	0,03	0,05		0,11			0,06						0,06		0,13			0,10													
c			0,02					0,50															0,03					0,03			
a																							0,09			0,17		0,07			
b																							0,13								
h			0,25			0,13																									
Ldh-2																															
N	*5*	*20*	*20*	*22*	*6*	*6*	*0*	*0*	*0*	*1*	*8*	*10*	*2*	*0*	*0*	*3*	*0*	*0*	*0*	*1*	*0*	*0*	*14*	*0*	*1*	*23*	*0*	*3*	*0*	*0*	*0*
c			0,03																												
b	1,00	1,00	0,98	1,00	1,00	1,00				1,00	1,00	1,00	1,00			0,83				1,00			0,46		0,50	0,43		0,67			
f																0,17							0,54		0,50	0,57					
e																												0,33			

Appendix S2

Clustering assignment of 31 sample sites in XLS format. Assignment of 31 populations of marbled newts (*Tritrus marmoratus* and *T. pygmaeus*) to two clusters in STRUCTURE 2.1 . Columns represent sample sites localities, Universal Transverse Mercator coordinates (UTM_X and UTM_Y), sample number (N) and the probability to belong to cluster 1 (that corresponds to *T. marmoratus*).

Code	Sample site	UTM_X (km)	UTM_Y (km)	N	Inferred cluster 1
1	Porto de Mós	514.0	4384.3	15	0.009
2	Rio Maior	507.3	4355.2	20	0.008
3	Alqueidão	536.0	4376.6	25	0.020
4	Valado dos Frades	499.2	4383.1	22	0.007
5	São Bartolomeu dos Galegos	476.0	4348.0	6	0.044
6	Mosteiro de Alcanene	513.7	4364.5	14	0.014
7	Foz do Arelho	482.6	4365.0	11	0.016
8	Casais dos Morgados	488.6	4369.5	11	0.015
9	Carrascal	503.7	4375.9	1	0.010
10	Molianos	506.8	4373.0	14	0.016
11	Covas	517.5	4376.1	8	0.020
12	Casais Monizes	509.0	4367.0	10	0.007
13	Carrascal II	503.8	4375.8	10	0.020
14	Cela	496.3	4376.6	2	0.011
15	Genrinhas	498.4	4373.9	10	0.037
16	Sta Susana	499.3	4353.0	3	0.119
17	Pataias Gare	501.5	4388.6	11	0.011
18	Famalicão da Nazaré	492.5	4377.3	5	0.018
19	Chão	516.4	4378.8	10	0.032
20	Covão da Fonte	517.6	4376.3	11	0.031
21	Molianos II	508.2	4374.4	10	0.029
22	Ribeira da Maceira	500.4	4371.4	2	0.027
23	Juncal	507.3	4383.4	16	0.778
24	Casal da Charneca	491.5	4360.0	21	0.984
25	Andam	508.9	4385.8	16	0.974
26	Salir de Matos	493.0	4364.4	24	0.987
27	Cós	504.4	4383.8	14	0.990
28	Fonte da Pena da Couvinha	497.5	4370.6	17	0.971
29	Vidais	495.2	4358.3	18	0.984
30	Casal da Coita	500.5	4366.5	20	0.986
31	Vimeiro	498.6	4368.8	21	0.934

Sample	Probability Inferred cluster 1
J1	0.035
J2	0.036
J3	0.226
J4	0.631
J5	0.870
J6	0.938
J7	0.941
J8	0.942
J9	0.955
J10	0.974
J11	0.974
J12	0.977
J13	0.987
J14	0.987
J15	0.987
J16	0.992

ACKNOWLEDGMENTS

Although the name on the cover is mine, this thesis is the product of collaboration. This is reflected in the co-authorships of the chapters. Others have affected the ultimate shape of this thesis.

Of course, this thesis would not be possible without the many samples collected by numerous people. I am especially indebted to the colleagues at Aydin University, Turkey, for their hospitality and guidance in finding newts in Anatolia, during our field trip in 2006, and to the Walen family who generously provided accommodation in Nisa for many of the field trips in Portugal.

I could not forget all the people from CIBIO, in Portugal, and Naturalis, The Netherlands that have, one way or another, helped me, either in the lab, or elsewhere, analysing data, discussing results and sharing their knowledge.

Early on, I was *adopted* by the Evolutionary Biology group (Patrícia's group, as I sometimes call it), here at Leiden University. Not only did I make many friends within this group, but they also proved that there is life beyond trees.

More importantly, I would have lost my mind a long time ago, if it weren't for Paula constantly reminding me of my goal, putting up with my irritations and frustrations, and reading and commenting everything I wrote.

CURRICULUM VITAE

Gonçalo Espregueira Cruz Themudo was born on November 25, 1979, in Vila Nova de Gaia, Portugal. Although he was always a city-dweller, he decided to become a biologist early on, because he wanted to understand Nature (because he didn't).

In 2001, he finished his *Licenciatura* (licentiate degree) in Biology from the University of Porto, Portugal, and immediately started working in CIBIO (Centro de Investigação em Biodiversidade e Recursos Genéticos) in Vairão, Portugal, with a research stipendium on a project with Dr. J. W. [Pim] Arntzen and Prof. Dr. Nuno Ferrand. This three years' work culminated with the completion of his MSc in Applied Ecology in 2005 from the University of Porto. During this period he also visited Leiden, The Netherlands, for the first time, thanks to a grant of the Jan Joost ter Pelkwijk Fund.

In April 2005, he moved to Leiden, to start working on his PhD in Biology in Naturalis, continuing his work with Dr. Arntzen. For two months in 2005, he visited Dr. Wieslaw Babik at the UFZ laboratory in Halle, Germany, within the scope of an Alexander von Humboldt Foundation grant.

He is now living in Copenhagen, Denmark, where he works as a post-doctoral researcher at the Faculty of Life Sciences, University of Copenhagen.

Publications in peer-reviewed journals

Espregueira Themudo, G., B. Wielstra, and J. W. Arntzen. 2009. Multiple nuclear and mitochondrial genes resolve the branching order of a rapid radiation of crested newts (*Triturus*, Salamandridae). *Molecular Phylogenetics and Evolutions* **52**, 321-328.

Espregueira Themudo, G., W. Babik, and J. W. Arntzen. 2009. A combination of techniques proves useful in the development of nuclear markers in the genus *Triturus*. *Molecular Ecology Resources* **9**:1160-1162.

Arntzen, J. W., and G. Espregueira Themudo. 2008. Environmental parameters that determine species geographical range limits as a matter of time and space. *Journal of Biogeography* **35**:1177-1186.

Arntzen, J. W., G. Espregueira Themudo, and B. Wielstra. 2007. The phylogeny of crested newts (*Triturus cristatus* superspecies): nuclear and mitochondrial genetic characters suggest a hard polytomy, in line with the paleogeography of the centre of origin. *Contributions to Zoology* **76**:261-278.

Espregueira Themudo, G., and J. W. Arntzen. 2007b. Newts under siege: range expansion of *Triturus pygmaeus* isolates populations of its sister species. *Diversity and Distributions* **13**:580-586.

Espregueira Themudo, G., and J. W. Arntzen. 2007a. Molecular identification of marbled newts and a justification of the *Triturus marmoratus* and *T. pygmaeus* species status. *Herpetological Journal* **17**:24-30.

Book chapters

Espregueira Themudo, G., and J. W. Arntzen (2008) *Triturus marmoratus*. pp. 100-101, in Loureiro, A., N. Ferrand de Almeida, M.A. Carretero and O.S. Paulo (eds.) *Atlas dos Anfíbios e Répteis de Portugal.* Instituto da Conservação da Natureza e Biodiversidade, Lisboa.