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

Espregueria Themudo, G.

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

Unpublished manuscript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.

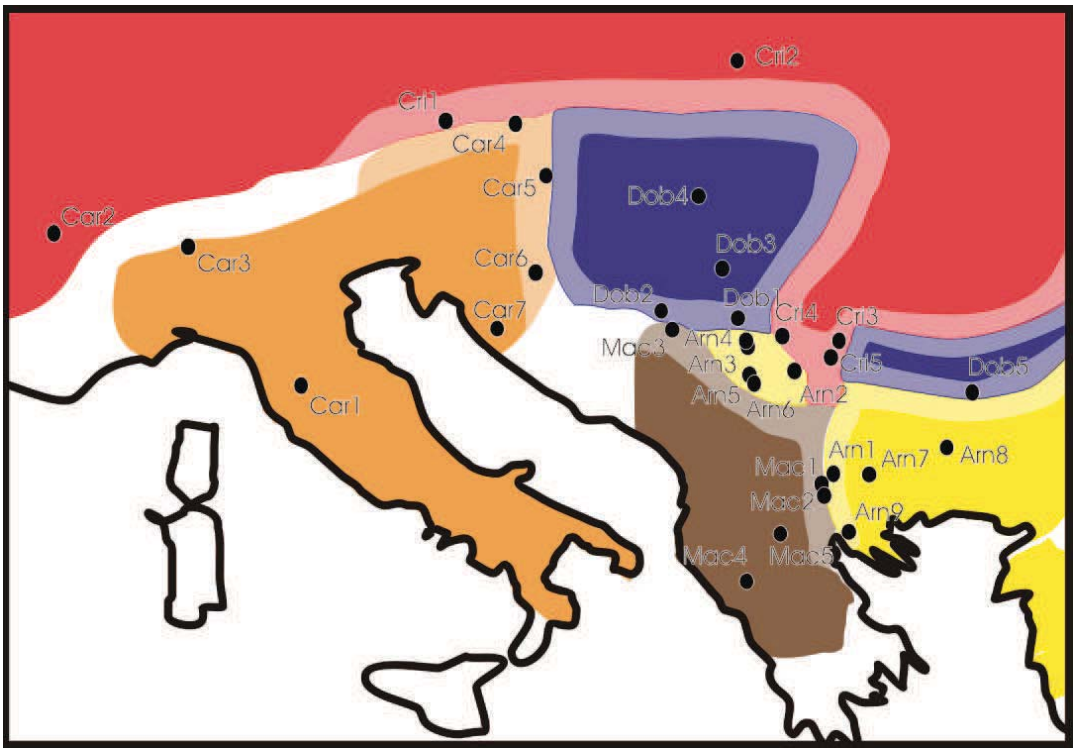

Figure 1 Distribution of *Triturus* species in the Balkans, represented by the following colours: orange - *T. carnifex*, red - *T. cristatus*, blue - *T. dobrogicus*, yellow - *T. arntzenii*, brown – *T. macedonicus*. Lighter colours represent the areas within 100 km of neighbouring species. Sample 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)	Coordinates 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	Monodentri, 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).

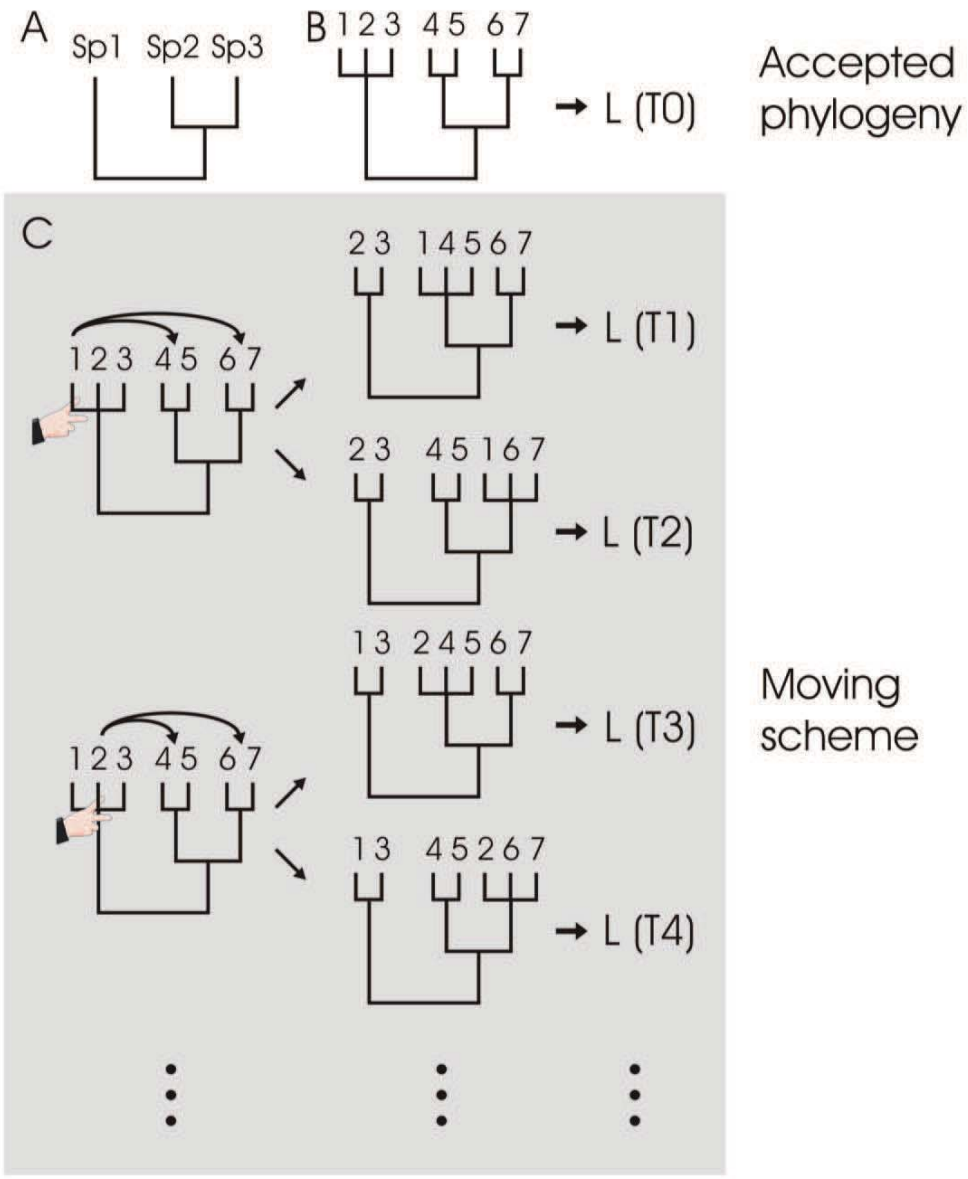

Figure 2 Schematic representation of the allele swapping procedure for a simple case with three species, with three, two and two individuals, respectively, starting with the 'true' or source phylogeny of Figure A, with an allelic species identity based on independent data such as morphology. Genetic 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.