Available online at www.sciencedirect.com
SCIENCE DIRECT
Gene 378 (2006) 31–41
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Structure and evolution of the mitochondrial control region and flanking sequences in the European cave salamander *Proteus anguinus*

Špela Gorički, Peter Trontelj *

University of Ljubljana, Biotechnical Faculty, Department of Biology, Večna pot 111, SI-1000 Ljubljana, Slovenia

Received 3 February 2006; received in revised form 18 April 2006; accepted 20 April 2006
Available online 4 May 2006
Received by C. Saccone

Abstract

The European cave salamander *Proteus anguinus Laurenti 1768* is one of the best-known subterranean animals, yet its evolutionary history and systematic relationships remain enigmatic. This is the first comprehensive study on molecular evolution within the taxon, using an mtDNA segment containing the control region (CR) and adjacent sequences. Two to seven tandem repeats of 24–32 bp were found in the intergenic spacer region (VNTR1), and three, four or six repeats, 59–77 bp each, in the 3′ end of the CR (VNTR2). Different molecular mechanisms account for VNTR2 formation in different lineages of *Proteus*. The overall CR variation was lower than that of the spacer region, the 3′ end of the *cytb* gene, or the tRNA genes. Individual genes and the concatenated non-repetitive sequences produced similar, well resolved maximum likelihood, Bayesian inference and parsimony trees. The numbers of repeat elements as well as the genealogy of the VNTR2 repeat units were mostly inconsistent with the groupings of the non-repetitive sequences. Different degrees of repeat array homogenization were detected in all major groups. Orthology was established for the first and the second VNTR2 elements of some populations. These two copies may therefore be used for analyses at the population level. The pattern of CR sequence variation points to strong genetic isolation of hydrographically separated populations. Genetic separation of the major groups of populations is incongruent with the current division into subspecies.

Keywords: Mitochondrial DNA; Tandem repeats; Proteidae; Amphibia; Phylogeny

1. Introduction

The European cave salamander *Proteus anguinus Laurenti 1768* inhabits subterranean karst waters of the Dinaric karst in the western Balkan Peninsula. Two subspecies are presently recognized: the white, troglomorphic *P. anguinus anguinus* and the non-troglomorphic, dark-pigmented *P. anguinus parkelj Sket and Arntzen 1994*. The former is distributed throughout the species' range in several hydrographically isolated areas (Sket, 1997), whereas the black subspecies occurs only in the vicinity of Črnomelj in south-eastern Slovenia (Sket and Arntzen, 1994). High morphological resemblance between hydrographically isolated populations of the white subspecies is in stark contrast to the findings of the allozyme study by Sket and Arntzen (1994), that the degree of differentiation within the white subspecies exceeds that between both subspecies. Due, however, to the limited number of sampled populations and small sample size, the above study failed to provide a full insight into the relationships among populations of *Proteus*.

A common approach to resolving intraspecific relationships is the phylogenetic analysis of mitochondrial control region (CR) sequences. Owing primarily to the lack of selective constraints on the majority of nucleotide positions, it is typically considered, on average, the fastest-evolving part of vertebrate mitochondrial genome (e.g. Lunt et al., 1998). However, in a recent analysis of complete mitochondrial genomes of the salamanders *Ambystoma* spp., low variation was detected in the CR, comparable to the levels of variation in the mitochondrial rRNA genes, and exceeded by variation in protein-coding

Abbreviations: A, adenine; AIC, Akaike Information Criterion; C, cytosine; CR, control region; CSB, conserved sequence block; Cy, cyanine; *cytb*, cytochrome *b*; ETAS, extended termination associated sequence; G, guanine; hLRT, hierarchical likelihood ratio tests; ML, maximum likelihood; mtDNA, mitochondrial DNA; O_H, origin of heavy strand replication; PCR, polymerase chain reaction; rRNA, ribosomal RNA; T, thymine; tRNA, transfer RNA; Ts/Tv, transition/transversion ratio; VNTR1, variable number of tandem repeats in the intergenic spacer region; VNTR2, variable number of tandem repeats in the control region; Y, cytosine or thymine.

* Corresponding author. Tel.: +386 1 423 33 88; fax: +386 1 257 33 90.
E-mail address: peter.trontelj@bf.uni-lj.si (P. Trontelj).

0378-1119/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2006.04.016

regions (Samuels et al., 2005). Low relative rate of CR evolution is known also in birds (e.g. Randi and Lucchini, 1998; Crochet and Desmarais, 2000; Pereira et al., 2004).

Three domains have been described in the vertebrate CR: the extended termination associated sequence (ETAS) domain, the central domain, and the conserved sequence block (CSB) domain (Sbisà et al., 1997; Saccone et al., 1999). Both marginal domains accumulate base substitutions, insertions and deletions at a substantially higher rate than the central domain, and often contain tandemly repeated sequences of varying length and copy number. Tandem repeats have been reported in a range of vertebrate taxa (e.g. Lunt et al., 1998 — review; Zardoya and Meyer, 2000; San Mauro et al., 2004), including plethodontid salamanders (Mueller and Boore, 2005). However, a detailed description of any salamander CR repeats has not yet been given. In salamander mtDNA, a hypervariable non-coding region close to the CR, between the $tRNA^{Thr}$ and $tRNA^{Pro}$ genes, has been identified and found to contain tandemly repeated sequences (McKnight and Shaffer, 1997; Zardoya et al., 2003; Zhang et al., 2003a,b; Mueller and Boore, 2005). Often difficult to align due to a variable copy number, and not sharing the evolutionary and mutational processes with the rest of the sequence, tandem repeats may bias phylogenetic analyses. Examples of recognition of specific repeats as useful markers for resolving relationships at different levels include shrews (Stewart and Baker, 1994), cyprinid fishes (Broughton and Dowling, 1997), percid fishes (Nesbø et al., 1998; Faber and Stepien, 1998), sturgeons (Ludwig et al., 2000), hornbills (Delport et al., 2002) and leaf beetles (Mardulyn et al., 2003). It is therefore necessary to investigate the information contained in the tandem repeats for each case individually and at different hierarchical levels.

In the present study we describe the structure and organization of the CR and flanking genes in *Proteus*. We address its usefulness in resolving genetic variability and potential cryptic taxonomic diversity by exploring the level of nucleotide variation and the strength of phylogenetic signal. We report on the presence of a putatively heteroplasmic repeat region in the spacer between the $tRNA^{Thr}$ and $tRNA^{Pro}$ genes (VNTR1), as well as a repeat region at the 3′ end of the CR (VNTR2). By analyzing the repeats in terms of their intraspecific phylogenetic distribution, we investigate the possibility of their use in further analyses of relationships among and within populations of *Proteus*.

2. Materials and methods

2.1. Specimens

Within the framework of a wider systematic study, sequences from 84 individuals from 26 localities were obtained. Of these, 31 individuals from 21 localities belonging to 10 hydrographically separated regions were selected for the present analysis. The individuals used are listed in Supplem. 1, and the localities are shown in Fig. 1. Selection criteria were: (1) each hydrographic region is represented by at least one individual, (2) identical haplotypes in the same hydrographic region are represented by one individual, (3) of the haplotypes pertaining to the same hydrographic region, the most common is included; if there are two haplotypes equally represented in a sample, then both are included, and (4) less common haplotypes are included only if they differ from the most common in either of the repeat regions. A reduced dataset was chosen to enable numerous time consuming maximum-likelihood (ML) searches. By following the above selection criteria, all variation in the repeat regions as well as all major sequence variation in the rest of the fragment was taken into account. Nineteen specimens were obtained from the zoological collection of the Department of Biology, University of Ljubljana, and the remainder were sampled and released at the site of capture.

2.2. DNA extraction, amplification and sequencing

In the individuals that were released, a sample 3–4 mm^2 in size was cut off the tip of the tail fin. Genomic DNA was extracted using the NucleoSpin Tissue kit (Macherey-Nagel) or the GenElute Mammalian Genomic DNA Miniprep Kit (Sigma), following the manufacturers' instructions. Genomic DNA of one individual was obtained from swabs of the skin surface according to the FTA Starter Pack (Whatman) protocol.

In four individuals, each from a different locality, the fragment was amplified using the light strand primer cytb17062a (5′-GAATGATACTTCCTATTYGC-3′), and the heavy strand primer 12Sb (5′-GAGAGTGACGGGCGATGTGT-3′; modified from Kocher et al. (1989)). The fragment was amplified in a hot-start PCR using Expand Long Template PCR System (Roche). Cycling conditions were 30 cycles of 60 s at 94 °C, 60 s at 51 °C, 3 min 30 s at 72 °C, and 6 min at 72 °C following the last cycle. The amplification product of about 2400 bp was gel-purified and sequenced on an ALFexpress II automated sequencer (Amersham Biosciences) with primers cytb17062a and 12Sb-Cy5. Because of low variation in the 12S rRNA gene, the CR and flanking sequences of the remaining individuals were amplified using genus-specific primers Pcytb (5′-CATTCCGCCCAATCAGCCAA-3′) and PFr (5′-AGTAAGGCTAGGACCAAACC-3′) for conserved regions in the cytochrome *b* and $tRNA^{Phe}$ genes, respectively (Fig. 2). An Expand High Fidelity PCR System (Roche) mediated PCR included 30 cycles of 45 s at 94 °C, 45 s at 57 °C and 2 min at 72 °C, followed by a 5-min step at 72 °C. The amplicons were sequenced with internal sequencing primers PAcytb-Cy5 (5′-CGCCCAATCAGCCAACTTAT-3′), Pactr-Cy5 (5′-CACCCAAGGCCAAAATTCTGA-3′) and PAFr-Cy5 (5′-CGGACTATATTGGCATTTTCAG-3′). Complete control region and flanking sequences have been deposited in GenBank under accession numbers DQ494754–DQ494786.

2.3. Structural and sequence analysis

To check whether the fragment amplified was indeed derived from the mitochondrial genome, its nucleotide frequencies in the individual PL98 from SW Slovenia were compared to the frequencies in the putatively homologous fragment of *Lyciasalamandra atifi* mtDNA (NC_002756; Zardoya et al., 2003) by a χ^2 test. The sequence of a distantly related salamander was used because the corresponding fragment of *Necturus*, the sister taxon to *Proteus* (Trontelj and Gorički, 2003; Wiens et al.,

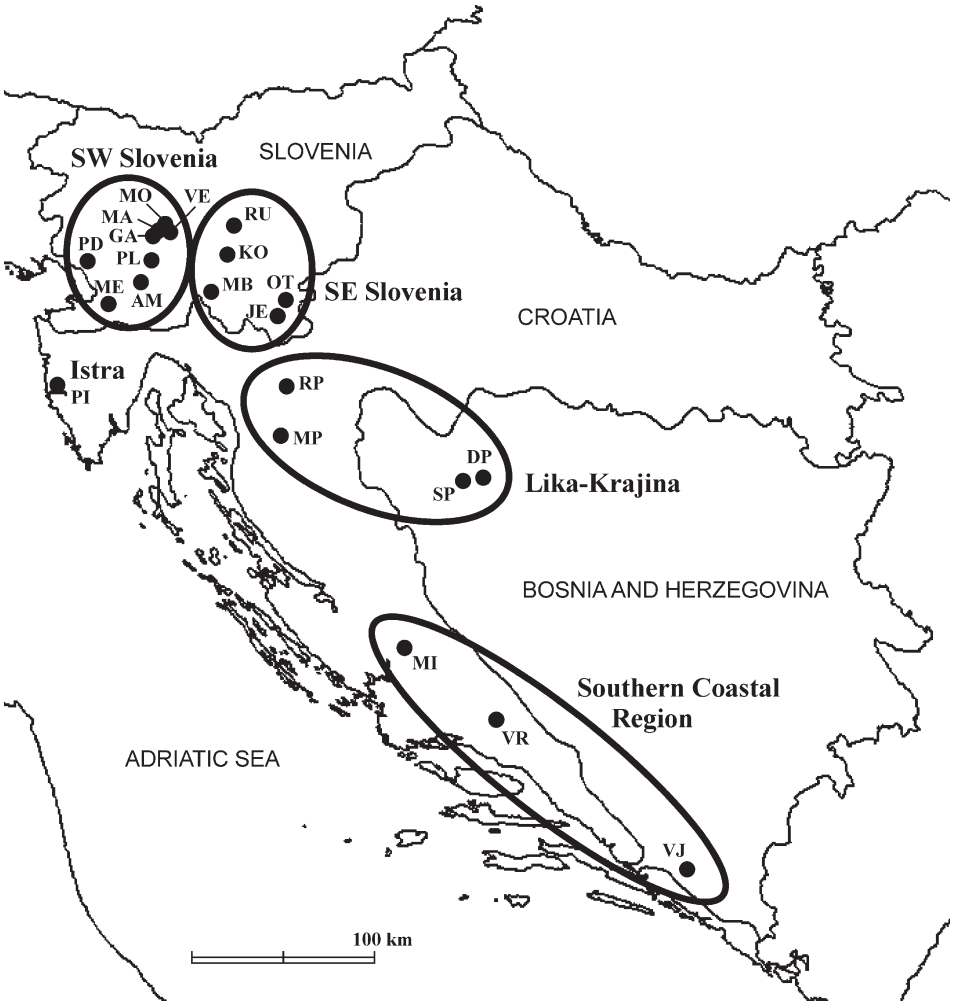


Fig. 1. Approximate locations of sampling sites. Locality abbreviations correspond to those shown in Supplem. 1 and in Figs. 4, 5 and 6. For a clearer presentation, the localities are assembled into five major geographic groups. The abbreviation "JE" (in SE Slovenia) stands for the locality of the black subspecies.

2005; but see Weisrock et al., 2005), could not be amplified with the primers employed. The comparison was performed on the entire region between primers cytb17062a and 12Sb, and for each partition (*cytb* gene, tRNA genes, spacer region, CR, and 12S rRNA gene) separately. In subsequent structural and phylogenetic analyses the fragment between primers PAcytb-Cy5 and PAFr-Cy5 was used.

Sequences were aligned using ClustalX v1.8 (Thompson et al., 1997) at default setting and refined by eye in BioEdit v5.0.9 (Hall, 1999). Functional regions were characterized by comparison with corresponding sequences of the following salamander taxa: *Andrias davidianus* (NC_004926; Zhang et al., 2003a), *Ranodon sibiricus* (NC_004021; Zhang et al., 2003b), *L. atifi* (NC_002756; Zardoya et al., 2003), *Salamandra corsica* (AF448816; Steinfartz et al., 2000) and *Ambystoma tigrinum* (AY659992; Samuels et al., 2005). Base composition and nucleotide variation were calculated using MEGA v3.0 (Kumar et al., 2004). Control region domains and elements were identified on the basis of data for mammals (Doda et al., 1981; Chang and Clayton, 1985; Hixson and Clayton, 1985; Sbisà et al., 1997), fish (Nesbø et al., 1998), amphibians (Bogenhagen et al., 1986; McKnight and Shaffer, 1997; Sumida et al., 2000; Zhang et al., 2003a,b), and reptiles (Brehm et al., 2003). Potential secondary structures of the tRNAs and the tandem repeats were identified using RNA *mfold* v2.3 and DNA *mfold* v3.1 (Zuker, 2003), respectively. All foldings were calculated at 8 °C, the annual average water temperature in habitats of *Proteus* (Sket, 1997). All sequence descriptions and relative position statements refer to the mtDNA light strand. Tandem repeats are numbered in the 5′ to 3′ direction.

2.4. Phylogenetic analyses

The dataset was analyzed separately for the four regions sequenced (*cytb*, tRNA genes, spacer region and CR) and combined, excluding all tandem repeats. All analyses were unrooted. Homogeneity of phylogenetic signal between regions was assessed using the partition-homogeneity test implemented in PAUP* v40b10 (Swofford, 2002). For the 1098-bp unpartitioned dataset, the GTR+Γ model of evolution was selected by hierarchical likelihood ratio tests (hLRT) and Akaike Information Criterion (AIC) implemented in Modeltest v3.7.c (Posada and Crandall, 1998). Parsimony and ML trees were constructed under the heuristic search option with 100 random-

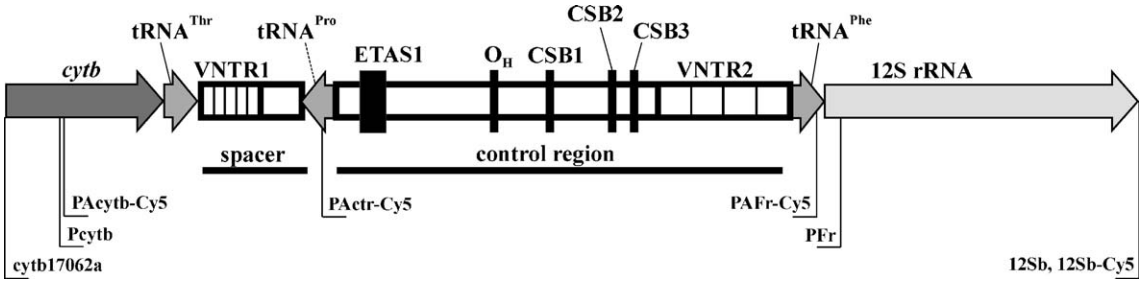


Fig. 2. Schematic representation of mtDNA fragment between primers cytb17062a and 12Sb in *Proteus*. The right and left arrows indicate genes encoded by the heavy and light strand, respectively. Non-coding regions are in white. Conserved elements in the CR are presented by dark boxes: ETAS1 — extended termination sequence 1, O_H — origin of heavy strand replication, CSB — putative conserved sequence block. VNTR1 and VNTR2 denote regions containing tandem repeats. The annealing sites of primers for amplification (Pcytb and PFr) and sequencing (PAcytb-Cy5, PActr-Cy5 and PAFr-Cy5) of the fragment used in structural and phylogenetic analyses are marked.

taxon-addition replicates and tree bisection–reconnection branch swapping, using PAUP*. Node support was assessed on the basis of 1000 bootstrap replicates. For Bayesian phylogenetic inference MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003) was used with the above mentioned model of evolution. A Markov chain Monte Carlo search was run with four chains for 10^6 generations, taking samples every 100 generations. The first 700 trees were discarded as burn-in after exploring the convergence of likelihood values graphically. From the remaining 9300 trees a posteriori probabilities for individual nodes were assessed based on their observed frequencies.

The evolutionary history of the VNTR2 region was analyzed treating each repeat unit in the tandem array as a separate operational taxonomic unit. Three alternative alignments (available at http://www.bf.uni-lj.si/bi/zoologija/peter_trontelj/data.htm) were used in the tree search under the ML optimality criterion using Phyml v2.4.4 (Guindon and Gascuel, 2003). The HKY+Γ model of evolution was selected using Modeltest. Trees were inferred also by parsimony under the heuristic search option with 100 random-taxon-addition replicates and tree bisection–reconnection branch swapping, using PAUP*. Node support was assessed on the basis of 1000 bootstrap replicates. Bayesian searches were performed as described above. The topologies of the trees obtained from repeat copies were compared to the genealogy obtained from the 1098-bp alignment with repeat regions excluded.

3. Results and discussion

3.1. Structure and variability of the sequence

The sequenced fragment of mtDNA between primers cytb17062a and 12Sb in *Proteus* has the same gene arrangement as found in other salamander taxa (Fig. 2). No significant differences in nucleotide composition were found between *Proteus* and a representative salamander sequence (Zardoya et al., 2003), except in the spacer region ($P=0.0006$). The $tRNA^{Thr}$ and $tRNA^{Pro}$ genes are presumed to be functional as they encode apparently stable secondary structures (Supplem. 2). We therefore believe that the fragment amplified is not a nuclear insertion. Sequences between primers PAcytb-Cy5 and PAFr-Cy5 ranged from 1359 (in DP206 from Lika–Krajina) to 1551 bp (in PL3 from SW Slovenia). Table 1 lists the nucleotide composition of each segment and other comparative sequence parameters. A high number, 15 sites out of 56 (26.8%) of nonsynonymous substitutions were found in the last 157-bp segment of the *cytb* gene. As many as 16.1% of the substitutions occurred at the second codon position, while only slightly more than one half (58.9%) at the third.

The $tRNA^{Thr}$ and $tRNA^{Pro}$ genes are separated by an intergenic spacer region (Fig. 2), identified in representatives of all salamander families sequenced to date. In *Proteus* the spacer is 137–275 bp long. Its 5′ end is composed of two (in PI121 from the Istra Peninsula) to seven (in PL3 from SW Slovenia) direct tandem repeats, followed by a non-repetitive, highly variable sequence of 66 (in the haplotype from Istra), 87 (in the haplotypes from Lika–Krajina and the Southern Coastal Region) or 88 (in the Slovenian haplotypes) nucleotides. The tandemly repeated sequence starts immediately after the $tRNA^{Thr}$ gene in the haplotype from Lika, where it involves the last adenine of the gene, or, in other individuals eight to 28 nucleotides from the 5′ end of the spacer. The number of repeats differs among geographical regions, among localities from the same region, among individuals from the same locality and possibly within individuals. Length heteroplasmy has been documented or proposed in individuals of several amphibian species (Bermingham et al., 1986; Wallis, 1987; Szymura et al., 2000; Mueller and Boore, 2005; Spolsky et al., 2006). In *Proteus*, sequence chromatograms suggested copy-number heteroplasmy in the individual MI158 from the Southern Coastal Region and the black individual JE143 from SE Slovenia. Haplotypes with three and four repeat copies were inferred in both of them. In four other black individuals sampled and in two further individuals from Miljacka pećina, these haplotypes were present singly. In the absence of paternal transmission, putative heteroplasmy in *Proteus* implies a rapid change in copy number, exceeding the rate of haplotype segregation.

The repeats in *Proteus* differ in length and sequence from the spacer repeats in other salamanders (McKnight and Shaffer, 1997; Zardoya et al., 2003; Zhang et al., 2003a, b). Altogether 20 types of repeats were found (Fig. 3). Repeats of 24 bp (types A to O in Fig. 3) were present in all haplotypes. In the haplotypes from SW Slovenia they were accompanied by longer variations containing extra cytosine residues (types P to T in

Table 1
Composition and variability of the analyzed mtDNA of *Proteus anguinus*

Region	Sites	NS [a]	% A	% C	% G	% T	NPS [b] (%)	NIS [c] (%)	Model [d]	Ts/Tv [e]	α [f]
Cytochrome *b*	Total	157	25.6	33.4	18.4	22.6	56 (35.7)	32 (20.4)	K80+Γ	3.43	0.71
	1st	53	20.1	30.7	31.6	17.6	14 (26.4)	6 (11.3)			
	2nd	52	20.2	26.4	11.2	42.2	9 (17.3)	5 (9.6)			
	3rd	52	36.7	43.1	12.2	8.0	33 (63.5)	18 (34.6)			
	Polymorphic	56	30.8	31.1	19.5	18.6					
$tRNA^{Thr}$		70	31.3	28.0	21.9	18.8	17 (24.3)	8 (11.4)	K80	2.97	
$tRNA^{Pro}$		69	30.9	30.2	17.2	21.7	13 (18.8)	5 (7.2)	K80	14.63	
Spacer [g]		88	27.9	34.9	15.6	21.6	51 (58.0)	36 (40.9)	HKY+Γ	2.44	1.32
	Polymorphic	51	27.2	28.5	23.1	21.1					
VNTR1		48–180	13.2	46.7	10.1	30.0					
CR [g]		693	30.9	24.1	14.8	30.2	115 (16.6)	77 (11.1)	HKY+Γ	2.32	0.09
	Polymorphic	115	27.6	25.5	18.6	28.3					
VNTR2		200–409	42.4	19.0	10.8	27.8					
Non-repetitive [h]		1098	30.1	26.4	15.9	27.6	253 (23.0)	163 (14.9)	GTR+Γ	3.53	0.21
	Polymorphic	253	28.6	27.3	20.1	24.1					

[a] Number of sites of aligned sequences, except in the VNTRs, where the alignments were ambiguous.
[b] Number and percent of polymorphic sites.
[c] Number and percent of parsimony informative sites.
[d] hLRT-estimated model of sequence evolution.
[e] Estimated transition/transversion (Ts/Tv) ratio.
[f] Shape parameter of the discrete Γ distribution.
[g] VNTR exluded.
[h] Cytochrome *b* (partial)+$tRNA^{Thr}$+spacer+$tRNA^{Pro}$+CR+$tRNA^{Phe}$ (partial).

Fig. 3). The CCCGTTAACC sequence was conserved across all types. Its core GTTAAC was involved in stem formation among adjacent and more distant repeat elements.

In the putatively heteroplasmic individuals all copies were of identical sequence (type A in MI158 and type B in the black individual JE143), a condition also found in the haplotype from Lika (all type F). In seven haplotypes (approximately one quarter) all copies were different. The greatest diversity of repeat types was observed among the haplotypes from SW Slovenia (Fig. 3). The haplotypes from the Southern Coastal Region possessed only three different types, and none was unique to this region. All individuals ($n=10$) sampled at the locality in Istra possessed the same haplotype containing two distinct complete repeats and an incomplete copy at the 5′ end of the array. The sequence of the latter matched the last nine nucleotides of the repeat at the opposite end (type N). The distribution of repeat types among and within groups of haplotypes of the non-repetitive sequence (see Section 3.3) supported only the groupings of very recently diverged haplotypes: SW Slovenian and parts of SE Slovenian and Southern Coastal groups. Hence, homoplasies appear to be frequent in VNTR1 elements, due to a low number of sites that are free to vary, within an already short repeat sequence.

The sequence of 880–1094 bp between the $tRNA^{Pro}$ and $tRNA^{Phe}$ genes was identified as the control region comprised of 680–692 bp of non-repetitive DNA followed by a tandemly repeated region (Fig. 2). The putative heavy strand origin of replication (O_H) was sited at 335 bp from the beginning of the aligned CR sequences, in the conserved segment. Downstream from the O_H, a putative conserved sequence block apparently homologous with CSB1 in mammals (Sbisà et al., 1997) was identified at position 450. Putative CSBs apparently homologous with CSB2 and CSB3 in *A. davidianus* (Zhang et al., 2003a), *A. tigrinum* (Samuels et al., 2005), *Lacerta dugesii* (Brehm et al., 2003) and mammals (Sbisà et al., 1997) were identified at positions 588 and 635, respectively. Upstream from the O_H at 34 bp, a sequence with 58.7–65% identity to the mammalian consensus ETAS1 sequence was located.

The level of nucleotide variation within the unrepeated section was lower than in any of the coding regions analyzed (Table 1). This observation is consistent with reports on low relative rates of CR evolution in the tiger salamander complex (Shaffer and McKnight, 1996; Samuels et al., 2005), the salamandrid salamanders (Steinfartz et al., 2000) and in birds (Randi and Lucchini, 1998; Crochet and Desmarais, 2000; Pereira et al., 2004). Steinfartz et al. (2000) attributed the slow evolutionary rate to the shortness of the 5′ (ETAS) domain in salamanders. Since the observed structural organization of the non-repetitive fragment of CR in *Proteus* matches that of other salamanders, their explanation may account for the low sequence variability observed in the CR of *Proteus*.

The last 200–409 bp of the CR are organized into imperfect direct repeats of 59–77 bp each. No copy number heteroplasmy was detected. The Istrian haplotype possessed an array of six copies, the haplotypes from SW Slovenia, Stična (RU70) and the black subspecies (JE143) – both from SE Slovenia – four, and the haplotypes from other regions three copies. A limited repeat number has been attributed to the stabilizing selection on regulatory elements within the repeats (Buroker et al., 1990; Brown et al., 1996; Wilkinson et al., 1997). In the VNTR2 of *Proteus* we did not identify any known functional sequence. Selection on repeat copy number may also act at a different level, possibly by a higher mtDNA replication rate of shorter molecules (Rand, 1993). Finally, migration could reduce variation in copy number. However, no gene flow between both three-copy groups of haplotypes, postdating the occurrence

Type	123456789111111111122222222223330123456789012345678901234567890	Total no. of copies	Occurence
A	ATTTT-CGCCC-------GTTAACCGCCTCTC	34	All major geographic groups, except Istra
B	.C...-.A...-------.....	22	SE Slovenia, Southern Coastal Region
C	.C...-.....-------.....	12	SE Slovenia, Lika-Krajina, Southern Coastal Region
D	....C-.T...-------.....	9	SW Slovenia
E	....C-.T...-------.....A..	7	SW Slovenia
F	....C-.....-------.....	3	Lika-Krajina
G	.....-.....-------.....T..	4	SE Slovenia
H	.CC..-.A...-------.....	2	SE Slovenia
I	.....-.....-------.....T	2	Lika-Krajina
J	....C-.T...-------.....C...	1	SW Slovenia
K	....C-.C...-------.....C...	1	SW Slovenia
L	.C...-.A...-------.....T..	1	SE Slovenia
M	.C...-.A...-------.....C...A..	1	SE Slovenia
N	.AA..-.....-------.....T..	1	Istra
O	..AC.-.....-------.....	1	Istra
P	...CCC.T...CCC----.....C...	13	SW Slovenia
Q	G..CCC.T...CCCCCCC.....	2	SW Slovenia
R	...CCC.T...CCCC---.....A..	2	SW Slovenia
S	G..CCC.T...CCCC---.....	2	SW Slovenia
T	G..CCA.T...CCC----.....	1	SW Slovenia

Fig. 3. Alignment of VNTR1 types A through T, their frequencies and distribution among major geographic groups. Dots indicate identity with repeat type A and dashes indicate gaps.

of repeats, was detected in the non-repetitive segment (see Section 3.3).

Altogether 45 different repeat sequence types were found. No identical repeats were found in any series. Repetitive sequences have been detected in the CR of a few species of the salamander family Plethodontidae (Mueller et al., 2004; Mueller and Boore, 2005), but representatives of other salamander families sequenced to date do not possess repeats in the CR. Together with the lack of substantial sequence similarities between the repeats in the two families, this strongly suggests their independent origin.

3.2. Mechanisms of VNTR2 repeat array formation

The occurrence of tandem repeats at the 3′ end of the CR is primarily attributed to either heavy strand (Broughton and Dowling, 1994; Mundy et al., 1996) or light strand (Ray and Densmore, 2003) slippage and mispairing during replication, facilitated by within-strand base pairing. Free binding energies of the most stable secondary structures of individual repeat elements in *Proteus* were high, ranging from −18.0 kJ/mol (the first repeat in VR36 from the Southern Coastal region) to −109.3 kJ/mol (the last repeat in the SE Slovenian haplotypes with three repeats). The free binding energies increased linearly with increasing number of repeats in both directions (the lowest linear regression R^2 was 0.960). When only one repeat was folded, the stem(s) included a poly-T sequence in the beginning of all but the first element, a poly-A block in the middle of each repeat and a dinucleotide microsatellite region at the 3′ end of each repeat. Such microsatellite regions within repeat elements were proposed to be implicated in the original formation of longer repeats (Wright, 1994; Zardoya and Meyer, 1998).

VNTR2 arrays in *Proteus* are the first documented case where duplications in different lineages of a single species followed seemingly different mechanisms of tandem repeat formation. In the heavy strand model, addition and deletion occur at the 5′ end of the array (Fumagalli et al., 1996; Mundy and Helbig, 2004). Providing the slippage occurs in neighboring repeats, the first copies in the array should be more similar, while the repeat(s) at the 3′ end would diverge. Such divergence pattern was indicated in the VNTR2 array of the Istrian haplotype, with the last copy being clearly divergent from the remaining copies. Although the relationships among the latter copies were less clear, the indication of homogenization by recurring additions and deletions of copies allows the specific repeat used as template to vary among subsequent duplications.

Conversely, the largest divergence within the VNTR2 arrays of the remaining haplotypes was observed between the first and the middle copies, with the first copy tentatively showing the greatest conservation across arrays (see Section 3.3 and Fig. 6). The latter was interpreted as the 5′ copy being the template of the original duplication with subsequent additions and deletions occurring at the 3′ end of the array (Broughton and Dowling, 1997). The observed pattern may thus be indicative of the light strand propagation of the tandem series. Low divergence between the middle and the last repeats was observed in the arrays of the Slovenian haplotypes, which is also consistent with the light strand model of tandem repeat expansion. In the individuals from Lika–Krajina and the Southern Coastal Region, either the middle and the last, or the first and the last copy were the least divergent. Although these results are less reliable due to the uncertainty of the alignment, the similarity between the terminal copies in the haplotypes from Lika–Krajina and the Southern Coastal Region might be a consequence of a hairpin consisting of the first and the second copy during the addition of the third element.

3.3. Evolutionary dynamics and phylogenetic applicability of the sequence

A very low value of the shape parameter of the discrete Γ distribution (Table 1) indicates high rate heterogeneity among sites in the 1098-bp alignment excluding both tandem repeat regions. This suggests that the few sites that vary are prone to

rapid substitution saturation. However, even at these sites mutational rate is relatively low, and plotting of the number of Ts and Tv against the GTR+Γ distances between haplotypes did not detect any substitution saturation (plot not shown). The combined *cytb*, tRNA, spacer and CR sequences, excluding the tandemly repeated regions, produced a well resolved phylogenetic tree of populations of *Proteus* (Fig. 4). Our attempts to root the tree using CR and adjacent sequences of *Ambystoma mexicanum* (AJ584639; Arnason et al., 2004), *A. tigrinum* (U36419; McKnight and Shaffer, 1997), *Lyciasalamndra atifi* (NC_002756; Zardoya et al., 2003) and *S. corsica* (AF448816; Steinfartz et al., 2000) resulted in inconsistent placements of the root. Depending on whether the entire fragment or only CR was used, either the Istrian or the Southern Coastal+Lika–Krajina clade appeared basal. We are therefore presenting the unrooted tree. The conclusions, however, are consistent with both possible rootings.

Groups of individuals from major geographic areas were strongly supported by all tree construction methods used, the only exception being the lineages from Lika–Krajina, which were not supported as monophyletic. No haplotypes were shared among major groups, indicating the absence of recent gene flow among the groups. The mtDNA sequences of *Proteus* support the relationships suggested in the previous allozyme analysis by Sket and Arntzen (1994): the black subspecies is part of the SE Slovenian group, and does not constitute an independent lineage as might be expected from its distinctive morphology. In separate analyses of the *cytb*, tRNA, CR, and spacer sequences, the same major groups were obtained, though groupings of the haplotypes from SE Slovenia and Lika–Krajina+Southern Coastal Region were less well supported using the CR and spacer sequences, respectively (not shown). The observed genetic separation of the major groups is in agreement with their mutual hydrographic

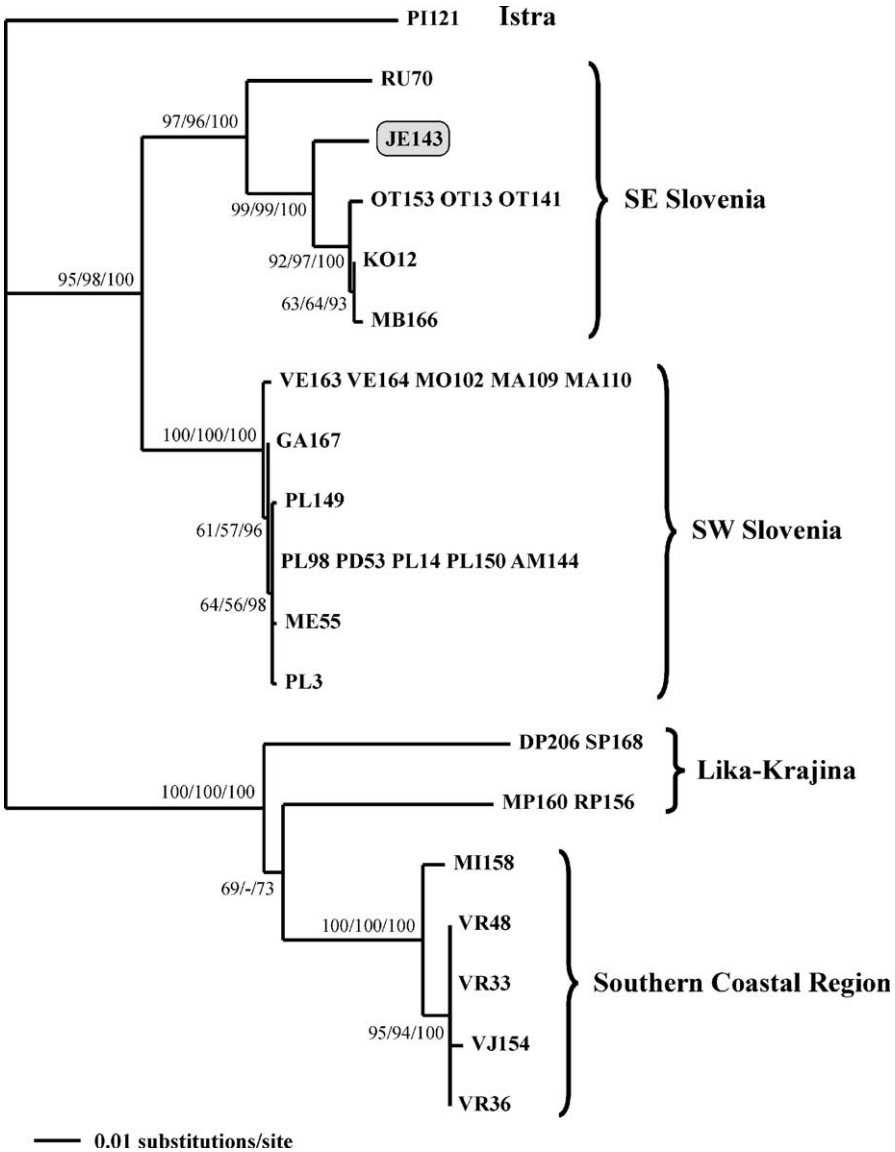


Fig. 4. Unrooted maximum likelihood tree of non-repetitive mtDNA sequences of *Proteus*. Bootstrap support values above 50% for each node for the ML/parsimony analyses are given, followed by Bayesian posterior probability values above 50%. On all three most parsimonious trees the branching order of the haplotypes from Lika–Krajina was reversed. The shaded box indicates the haplotype of the black subspecies.

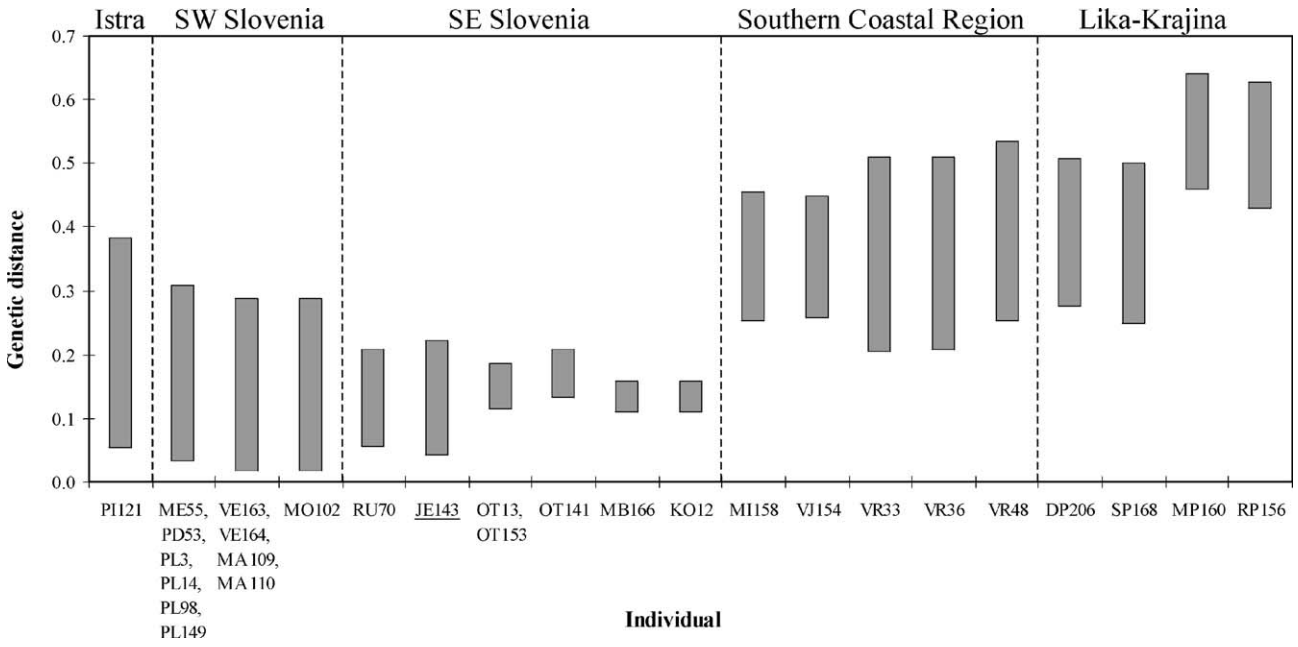


Fig. 5. Extent of sequence variation within individual VNTR2 arrays. Genetic distances were estimated by ML using the model selected for each haplotype (F81 or HKY85) by hLRTs. The black individual is underlined.

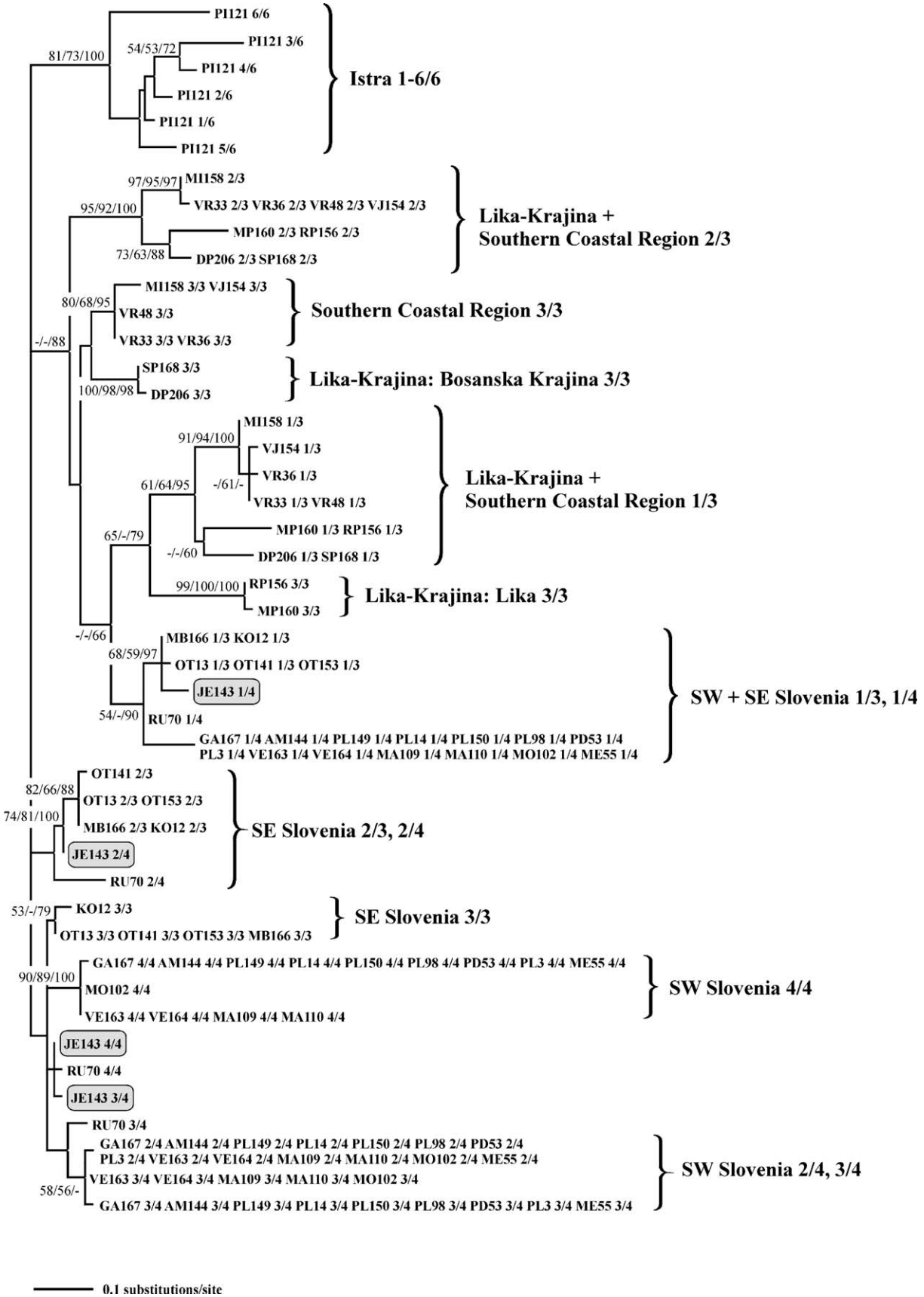
isolation effected by impermeable rock (Baučić, 1965; Milanović, 1979; Urumović et al., 1997; Gams, 2004). It coincides with the genetic structure of cave shrimp *Troglocaris anophthalmus* (V. Zakšek, B. Sket and P. Trontelj, unpubl. data) and, particularly in the Slovenian region, of subterranean populations of water lice *Asellus aquaticus* (Verovnik et al., 2004, 2005).

Intra-individual genetic distances between VNTR2 elements differed greatly among major geographic groups (Fig. 5). Tandemly repeated elements in the individuals from the Southern Coastal Region and Lika–Krajina were more variable than those within the Slovenian and Istrian individuals. In the haplotypes from SE Slovenia with three copies, the estimated genetic distance between the most divergent pair of copies was not much higher than the distance between the least divergent pair. Although recombination cannot be ruled out to influence variability of repeats through concerted evolution, there is at present little undisputed evidence for its occurrence in animal mtDNA (for reviews see Rokas et al., 2003; Barr et al., 2005). Alternatively, the degree of sequence divergence in an array may be explained by the balance between the copy insertion/deletion rate and the rate of point mutations in individual copies (Fumagalli et al., 1996). Based on the magnitudes of both absolute distances as well as the differences in the distances among the most and the least divergent copies, we deduce that in the individuals from SE Slovenia the ratio of copy turnover rate to the rate of point mutations is higher than in the individuals from Lika–Krajina and the Southern Coastal Region, having the same number of repeats.

Consistent with the intra-individual divergence patterns were the phylogenetic reconstructions performed on 84-bp alignments of the VNTR2 region, with each repeat unit treated separately (Fig. 6). Most relationships between repeat units were poorly resolved, but certain groups were nevertheless stable on all alignments tested and with all tree construction methods used. The topologies of the trees based on the VNTR2 units and the non-repetitive sequence were largely incongruent. All six repeats of the Istrian individual formed a separate group. The high conservation of repeat sequences within the Istrian individual indicates homogenization of its repeat array. Furthermore, it suggests a possible independent, lineage-specific origin of the Istrian group of repeats. Although less parsimonious, this hypothesis is based on the prediction that in closely related populations a similar sequence (in this case repeats) can be derived in parallel from a similar ancestral sequence (that has a tendency to form a stem–loop structure). While the outcome is essentially the same, its possible parallel origin may be suspected from the presence of minor differences. VNTR2 elements in the Istrian individual possess the following unique features: a unique pattern of sequence divergence within the array, a deviation from repeat number in other populations and a 4-bp insertion at the beginning of all except the first element.

In the remaining individuals, incomplete homogenization of the arrays was observed: Homogenization of the second and the third copy may have occurred in the repeat arrays of the SW Slovenian individuals and possibly in the RU70 individual from SE Slovenia, as suggested by the parsimony and Bayesian inference trees (trees

Fig. 6. Unrooted maximum likelihood tree of the VNTR2 repeats with bootstrap support values above 50% (ML and parsimony, respectively) and Bayesian posterior probability values above 50%. Each group is marked according to the major geographic region (Table 1) and the position of the elements within the array (e.g., 1/3 means the first element out of three). Shaded boxes indicate repeats of the black subspecies. On the ML and majority-rule consensus of 3185 most parsimonious trees, the first repeat in the Slovenian individuals was joined to the group containing the first repeat in the individuals from Lika–Krajina and Southern Coastal Region and the third repeat in Lika individuals, but the grouping received bootstrap support of less than 50%. Using alternative alignments, its relationship to other repeats was either unresolved (Bayesian tree of one possible alignment) or it grouped with the Istrian individual's repeats, again with low node support values.


PI121 6/6
PI121 3/6
PI121 4/6
PI121 2/6
PI121 1/6
PI121 5/6
81/73/100
54/53/72
Istra 1-6/6
MI158 2/3
VR33 2/3 VR36 2/3 VR48 2/3 VJ154 2/3
MP160 2/3 RP156 2/3
DP206 2/3 SP168 2/3
97/95/97
95/92/100
73/63/88
Lika-Krajina +
Southern Coastal Region 2/3
MI158 3/3 VJ154 3/3
VR48 3/3
VR33 3/3 VR36 3/3
80/68/95
Southern Coastal Region 3/3
-/-/88
SP168 3/3
DP206 3/3
100/98/98
Lika-Krajina: Bosanska Krajina 3/3
MI158 1/3
VJ154 1/3
VR36 1/3
VR33 1/3 VR48 1/3
91/94/100
-/61/-
61/64/95
MP160 1/3 RP156 1/3
DP206 1/3 SP168 1/3
65/-/79
-/-/60
Lika-Krajina +
Southern Coastal Region 1/3
RP156 3/3
MP160 3/3
99/100/100
Lika-Krajina: Lika 3/3
-/-/66
MB166 1/3 KO12 1/3
OT13 1/3 OT141 1/3 OT153 1/3
JE143 1/4
RU70 1/4
GA167 1/4 AM144 1/4 PL149 1/4 PL14 1/4 PL150 1/4 PL98 1/4 PD53 1/4
PL3 1/4 VE163 1/4 VE164 1/4 MA109 1/4 MA110 1/4 MO102 1/4 ME55 1/4
68/59/97
54/-/90
SW + SE Slovenia 1/3, 1/4
OT141 2/3
OT13 2/3 OT153 2/3
MB166 2/3 KO12 2/3
JE143 2/4
RU70 2/4
82/66/88
74/81/100
SE Slovenia 2/3, 2/4
KO12 3/3
OT13 3/3 OT141 3/3 OT153 3/3 MB166 3/3
53/-/79
SE Slovenia 3/3
GA167 4/4 AM144 4/4 PL149 4/4 PL14 4/4 PL150 4/4 PL98 4/4 PD53 4/4 PL3 4/4 ME55 4/4
MO102 4/4
VE163 4/4 VE164 4/4 MA109 4/4 MA110 4/4
90/89/100
SW Slovenia 4/4
JE143 4/4
RU70 4/4
JE143 3/4
RU70 3/4
GA167 2/4 AM144 2/4 PL149 2/4 PL14 2/4 PL150 2/4 PL98 2/4 PD53 2/4
PL3 2/4 VE163 2/4 VE164 2/4 MA109 2/4 MA110 2/4 MO102 2/4 ME55 2/4
VE163 3/4 VE164 3/4 MA109 3/4 MA110 3/4 MO102 3/4
GA167 3/4 AM144 3/4 PL149 3/4 PL14 3/4 PL150 3/4 PL98 3/4 PD53 3/4 PL3 3/4 ME55 3/4
58/56/-
SW Slovenia 2/4, 3/4
0.1 substitutions/site

not shown). It is likely to have occurred in the black individual JE143, involving the third and the fourth copy. Also, the third copy appears to have formed independently in the three groups from Lika–Krajina and the Southern Coastal Region. On the other hand, the sequence of the first copy in the Slovenian and Lika–Krajina+Southern Coastal populations matches the evolution of the non-repetitive segment in these groups. It is therefore likely to reflect their histories.

3.4. Conclusions

We have determined the sequence of a fragment of mtDNA in *Proteus* that is informative for resolving relationships within this taxon. The observed pattern of sequence variation points to a strong isolation between different hydrographic areas, discordant with the apparent morphological homogeneity. At the same time, the sequence contains regions that are conservative enough to allow for incorporation of more distantly related outgroups, including species from other genera and families. This is particularly important, since in its current perception the genus *Proteus* is monotypic.

Two sets of tandem repeats were detected in the analyzed mtDNA fragment. Several characteristics of tandem repeats in *Proteus* advocate against their use in phylogenetic studies of this species: the variable, but low repeat number and repeat length, the putatively different directions of repeat array formation, and the seemingly different relative repeat turnover rates in different groups of populations. On the other hand, when orthology of individual copies from different individuals is postulated, and the rate of point mutations exceeds the rate of gain or loss of copies, then the repetitive sequence or parts thereof can be used for revealing population genetic and phylogeographic patterns within lineages. Orthology could be established for the first and the second VNTR2 elements in the Lika–Krajina and Southern Coastal populations. These two copies may therefore be used in, while the third should be excluded from, subsequent analyses at the population level. Similarly, surveys of the Slovenian populations may include the first and the second elements, but the third and the fourth copies, which were presumably formed later, should be omitted.

Acknowledgments

We are grateful to Boris Sket for his guidance in biospeleology and comprehensive help with the project. We thank Janko Boštjančič, Boris Bulog, Arne Hodalič, Eduard Kletečki, Franc Kljun, Ljerka Lah, Slavko Polak, Tonči Rađa and Maja Zagmajster for their help with field work and for providing samples. Two anonymous reviewers provided helpful comments on an earlier version of the manuscript. This study was funded by the Slovenian Research Agency and from the contract no. EVK2-CT-2001-00121-PASCALIS of the Fifth Research and Technological Development Framework Program supported by the European Community.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.gene.2006.04.016.

References

Arnason, U., Gullberg, A., Janke, A., Joss, J., Elmerot, C., 2004. Mitogenomic analyses of deep gnathostome divergences: a fish is a fish. Gene 333, 61–70.

Barr, C.M., Neiman, M., Taylor, D.R., 2005. Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. New Phytol. 168, 39–50.

Baučić, I., 1965. Hydrological characteristics of the Dinaric karst in Croatia with a special regard to the underground water connections. Naše jame 7, 61–72.

Bermingham, E., Lamb, T., Avise, J.C., 1986. Size polymorphism and heteroplasmy in the mitochondrial DNA of lower vertebrates. J. Heredity 77, 249–252.

Bogenhagen, D.F., Yoza, B.K., Cairns, S.S., 1986. Identification of initiation sites for transcription of *Xenopus laevis* mitochondrial DNA. J. Biol. Chem. 261, 8488–8494.

Brehm, A., Harris, D.J., Alves, C., Jesus, J., Thomarat, F., Vicente, L., 2003. Structure and evolution of the mitochondrial DNA complete control region in the lizard *Lacerta dugesii* (Lacertidae, Sauria). J. Mol. Evol. 56, 46–53.

Broughton, R.E., Dowling, T.E., 1994. Length variation in mitochondrial DNA of the minnow *Cyprinella spiloptera*. Genetics 138, 179–190.

Broughton, R.E., Dowling, T.E., 1997. Evolutionary dynamics of tandem repeats in the mitochondrial DNA control region of the minnow *Cyprinella spiloptera*. Mol. Biol. Evol. 14, 1187–1196.

Brown, J.R., Beckenbach, K., Beckenbach, A.T., Smith, M.J., 1996. Length variation, heteroplasmy and sequence divergence in the mitochondrial DNA of four species of sturgeon (*Acipenser*). Genetics 142, 525–535.

Buroker, N.E., et al., 1990. Length heteroplasmy of sturgeon mitochondrial DNA: an illegitimate elongation model. Genetics 124, 157–163.

Chang, D.D., Clayton, D.A., 1985. Priming of human mitochondrial DNA replication occurs at the light-strand promoter. Proc. Natl. Acad. Sci. U. S. A. 82, 351–355.

Crochet, P.-A., Desmarais, E., 2000. Slow rate of evolution in the mitochondrial control region of gulls (Aves: Laridae). Mol. Biol. Evol. 17, 1797–1806.

Delport, W., Ferguson, J.W.H., Bloomer, P., 2002. Characterization and evolution of the mitochondrial DNA control region in hornbills (Bucerotiformes). J. Mol. Evol. 54, 794–806.

Doda, J.N., Wright, C.T., Clayton, D.A., 1981. Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc. Natl. Acad. Sci. U. S. A. 78, 6116–6120.

Faber, J.E., Stepien, C.A., 1998. Tandemly repeated sequences in the mitochondrial DNA control region and phylogeography of the pike-perches *Stizostedion*. Mol. Phylogenet. Evol. 10, 310–322.

Fumagalli, L., Taberlet, P., Favre, L., Hausser, J., 1996. Origin and evolution of homologous repeated sequences in the mitochondrial DNA control region of shrews. Mol. Biol. Evol. 13, 31–46.

Gams, I., 2004. Kras v Sloveniji v prostoru in času (Karst in Slovenia in Space and Time). ZRC SAZU, Ljubljana.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Hixson, J.E., Clayton, D.A., 1985. Initiation of transcription from each of the two human mitochondrial promoters requires unique nucleotides at the transcriptional start sites. Proc. Natl. Acad. Sci. U. S. A. 82, 2660–2664.

Kocher, T.D., et al., 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. U. S. A. 86, 6196–6200.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. Brief. Bioinform. 5, 150–163.

Ludwig, A., May, B., Debus, L., Jenneckens, I., 2000. Heteroplasmy in the mtDNA control region of sturgeon (*Acipenser*, *Huso* and *Scaphirhynchus*). Genetics 156, 1933–1947.

Lunt, D.H., Whipple, L.E., Hyman, B.C., 1998. Mitochondrial DNA variable number tandem repeats (VNTRs): utility and problems in molecular ecology. Mol. Ecol. 7, 1441–1455.

Mardulyn, P., Termonia, A., Milinkovitch, M.C., 2003. Structure and evolution of the mitochondrial control region of leaf beetles (Coleoptera: Chrysomelidae): a hierarchical analysis of nucleotide sequence variation. J. Mol. Evol. 56, 38–45.

McKnight, M.L., Shaffer, H.B., 1997. Large, rapidly evolving intergenic spacers in the mitochondrial DNA of the salamander family Ambystomatidae (Amphibia: Caudata). Mol. Biol. Evol. 14, 1167–1176.

Milanović, P.T., 1979. Hidrogeologija karsta i metode istražvanja [Karst hydrogeology and research methods]. Hidroelektrarne na Trebišnjici, Inštitut za korištenje i zaštitu voda na kršu, Trebinje.

Mueller, R.L., Boore, J.L., 2005. Molecular mechanisms of extensive mitochondrial gene rearrangement in plethodontid salamanders. Mol. Biol. Evol. 22, 2104–2112.

Mueller, R.L., Macey, J.R., Jaekel, M., Wake, D.B., Boore, J.L., 2004. Morphological homoplasy, life history evolution, and historical biogeography of plethodontid salamanders inferred from complete mitochondrial genomes. Proc. Natl. Acad. Sci. U. S. A. 101, 13820–13825.

Mundy, N.I., Helbig, A.J., 2004. Origin and evolution of tandem repeats in the mitochondrial DNA control region of shrikes (*Lanius* spp.). J. Mol. Evol. 59, 250–257.

Mundy, N.I., Winchell, C.S., Woodruff, D.S., 1996. Tandem repeats and heteroplasmy in the mitochondrial DNA control region of the loggerhead shrike (*Lanius ludovicianus*). J. Heredity 87, 21–26.

Nesbø, C.L., Arab, M.O., Jakobsen, K.S., 1998. Heteroplasmy, length and sequence variation in the mtDNA control regions of three percid fish species (*Perca fluviatilis*, *Acerina cernua*, *Stizostedion lucioperca*). Genetics 148, 1907–1919.

Pereira, S.L., Grau, E.T., Wajntal, A., 2004. Molecular architecture and rates of DNA substitutions of the mitochondrial control region of cracid birds. Genome 47, 535–545.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rand, D.M., 1993. Endotherms, ectotherms, and mitochondrial genome size variation. J. Mol. Evol. 37, 281–295.

Randi, E., Lucchini, V., 1998. Organization and evolution of the mitochondrial DNA control region in the avian genus *Alectoris*. J. Mol. Evol. 47, 449–462.

Ray, D.A., Densmore, L.D., 2003. Repetitive sequences in the crocodilian mitochondrial control region: poly-A sequences and heteroplasmic tandem repeats. Mol. Biol. Evol. 20, 1006–1013.

Rokas, A., Ladoukakis, E., Zouros, E., 2003. Animal mitochondrial DNA recombination revisited. Trends Ecol. Evol. 18, 411–417.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Saccone, C., De Giorgi, C., Gissi, C., Pesole, G., Reyes, A., 1999. Evolutionary genomics in Metazoa: the mitochondrial DNA as a model system. Gene 238, 95–209.

Samuels, A.K., et al., 2005. Transcriptional and phylogenetic analysis of five complete ambystomatid salamander mitochondrial genomes. Gene 349, 43–53.

San Mauro, D., García-París, M., Zardoya, R., 2004. Phylogenetic relationships of discoglossid frogs (Amphibia:Anura:Discoglossidae) based on complete mitochondrial genomes and nuclear genes. Gene 343, 357–366.

Sbisà, E., Tanzariello, F., Reyes, A., Pesole, G., Saccone, C., 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. Gene 205, 125–140.

Shaffer, H.B., McKnight, M.L., 1996. The polytypic species revisited: genetic differentiation and molecular phylogenetics of the tiger salamander *Ambystoma tigrinum* (Amphibia: Caudata) complex. Evolution 50, 417–433.

Sket, B., 1997. Distribution of *Proteus* (Amphibia: Urodela: Proteidae) and its possible explanation. J. Biogeogr. 24, 263–280.

Sket, B., Arntzen, J.W., 1994. A black, non-troglomorphic amphibian from the karst of Slovenia: *Proteus anguinus parkelj* n. ssp. (Urodela: Proteidae). Contrib. Zool. 64, 33–53.

Spolsky, C.M., Szymura, J.M., Uzzell, T., 2006. Mapping *Bombina* mitochondrial genomes: the conundrum of Carpathian *Bombina variegata* (Anura: Discoglossidae). J. Zoolog. Syst. Evol. Res. 44, 100–104.

Steinfartz, S., Veith, M., Tautz, D., 2000. Mitochondrial sequence analysis of *Salamandra* taxa suggests old splits of major lineages and postglacial recolonizations of central Europe from distinct source populations of *Salamandra salamandra*. Mol. Ecol. 9, 397–410.

Stewart, D.T., Baker, A.J., 1994. Patterns of sequence variation in the mitochondrial D-loop region of shrews. Mol. Biol. Evol. 11, 9–21.

Sumida, M., Kaneda, H., Kato, Y., Kanamori, Y., Yonekawa, H., Nishioka, M., 2000. Sequence variation and structural conservation in the D-loop region and flanking genes of mitochondrial DNA from Japanese pond frogs. Genes Genet. Syst. 75, 79–92.

Swofford, D.L., 2002. PAUP*: Phylogenetic analysis using parsimony (* and other methods), Version 40b10. Sinauer, Sunderland, Massachusetts.

Szymura, J.M., Uzzell, T., Spolsky, C., 2000. Mitochondrial DNA variation in the hybridizing fire-bellied toads, *Bombina bombina* and *B. variegata*. Mol. Ecol. 9, 891–899.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Trontelj, P., Gorički, Š., 2003. Monophyly of the family Proteidae (Amphibia: Caudata) tested by phylogenetic analysis of mitochondial 12S rDNA sequences. Nat. Croat. 12, 113–120.

Urumović, K., Vazdar, T., Dragičević, I., Tomljenović, B., 1997. Environmental impact on karstic aquifers in Istria in Western Croatia. In: Günay, G., Johnson, A.I. (Eds.), Karst waters and Environmental Impacts. Proceedings 5th International Symposium and Field Seminar on Karst Waters and Environmental Impacts, Antalya, Turkey, pp. 45–51.

Verovnik, R., Sket, B., Trontelj, P., 2004. Phylogeography of subterranean and surface populations of water lice *Asellus aquaticus* (Crustacea: Isopoda). Mol. Ecol. 13, 1519–1532.

Verovnik, R., Sket, B., Trontelj, P., 2005. The colonization of Europe by the freshwater crustacean *Asellus aquaticus* (Crustacea: Isopoda) proceeded from ancient refugia and was directed by habitat connectivity. Mol. Ecol. 14, 4355–4369.

Wallis, G.P., 1987. Mitochondrial DNA insertion polymorphism and germ line heteroplasmy in the *Triturus cristatus* complex. Heredity 58, 229–238.

Weisrock, D.W., Harmon, L.J., Larson, A., 2005. Resolving deep phylogenetic relationships in salamanders: analyses of mitochondrial and nuclear genomic data. Syst. Biol. 54, 758–777.

Wiens, J., Bonett, R., Chippindale, P., 2005. Ontogeny discombobulates phylogeny: paedomorphosis and higher-level salamander relationships. Syst. Biol. 54, 91–110.

Wilkinson, G.S., Mayer, F., Kerth, G., Petri, B., 1997. Evolution of repeated sequence arrays in the D-loop region of bat mitochondrial DNA. Genetics 146, 1035–1048.

Wright, J.M., 1994. Mutation at VNTRs: are minisatellites the evolutionary progeny of microsatellites? Genome 37, 345–347.

Zardoya, R., Meyer, A., 1998. Cloning and characterization of a microsatellite in the mitochondrial control region of the African side-necked turtle, *Pelomedusa subrufa*. Gene 216, 149–153.

Zardoya, R., Meyer, A., 2000. Mitochondrial evidence on the phylogenetic position of caecilians (Amphibia: Gymnophiona). Genetics 155, 765–775.

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.

Zhang, P., Chen, Y.Q., Liu, Y.F., Zhou, H., Qu, L.H., 2003a. The complete mitochondrial genome of the Chinese giant salamander, *Andrias davidianus* (Amphibia: Caudata). Gene 311, 93–98.

Zhang, P., Chen, Y.Q., Zhou, H., Wang, X.L., Qu, L.H., 2003b. The complete mitochondrial genome of a relic salamander, *Ranodon sibiricus* (Amphibia: Caudata) and implications for amphibian phylogeny. Mol. Phylogenet. Evol. 28, 620–626.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 31, 3406–3415.