Phylogeny and genetic history of the Siberian salamander (Salamandrella keyserlingii, Dybowski, 1870) inferred from complete mitochondrial genomes

Boris Malyarchuk *, Miroslava Derenko, Galina Denisova

Institute of Biological Problems of the North, Magadan 685000, Russia

Article info

Article history:
Received 31 October 2012
Revised 28 January 2013
Accepted 3 February 2013
Available online 12 February 2013

Keywords:
Mitochondrial DNA
Salamandrella keyserlingii
Phylogeny
Molecular phylogeography
Adaptation

Abstract

We assessed phylogeny of the Siberian salamander (Salamandrella keyserlingii, Dybowski, 1870), the most northern ectothermic, terrestrial vertebrate in Eurasia, by sequence analysis of complete mitochondrial genomes in 26 specimens from different localities (China, Khabarovsk region, Sakhalin, Yaktutia, Magadan region, Chukotka, Kamchatka, Ural, European part of Russia). In addition, a complete mitochondrial genome of the Schrenck salamander, Salamandrella schrenckii, was determined for the first time. Bayesian phylogenetic analysis of the entire mtDNA genomes of S. keyserlingii demonstrates that two haplogroups – clades A and C – radiated about 1.4 million years ago (Mya). Bayesian skyline plots of population size change through time show an expansion around 250 thousand years ago (kya) and then a decline around the Last Glacial Maximum (25 kya) with subsequent restoration of population size. Climatic changes during the Quaternary period have dramatically affected the population genetic structure of the Siberian salamanders. In addition, complete mtDNA sequence analysis allowed us to recognize that the vast area of Northern Eurasia was colonized only by the Siberian salamander clade C1b during the last 150 kya. Meanwhile, we were unable to find evidence of molecular adaptation in this clade by analyzing the whole mitochondrial genomes of the Siberian salamanders.

* Corresponding author. Address: Institute of Biological Problems of the North, Far East Branch of the Russian Academy of Sciences, Portovaya Str., 18, Magadan 685000, Russia. Fax: +7 4132 63 44 63.
E-mail address: malyarchuk@ibpn.ru (B. Malyarchuk).

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

The Siberian salamander (Salamandrella keyserlingii, Dybowski, 1870) occupies the widest range compared to other Palearctic salamanders, extending from Hokkaido in the east to the North-East of Europe in the west, and from Chukotkan peninsula in the north to northern China in the south (Borkin et al., 1984). Apparently, this wide distribution of the Siberian salamander was promoted by the fact that this species can tolerate extremely low winter temperatures (up to −40 °C) (Berman et al., 1984). Morphologists reported homogeneity of the Siberian salamander populations across most of the species range, albeit they suggested the presence of some variety at the southeast of the range (Borkin et al., 1984). Molecular genetic studies, based on both mitochondrial and nuclear genomic data, revealed a cryptic species, the Schrenck salamander (S. schrenckii), in the southeast of Russia (Primorye and Khabarovsk regions) (Berman et al., 2005; Matsui et al., 2008; Poyarkov and Kuzmin, 2008; Malyarchuk et al., 2009).

More detailed phylogeographic investigation of mitochondrial DNA (mtDNA) cytochrome b gene variation in populations of the Siberian salamander allowed us to establish that the middle Amur River Basin and the northern Sikhote-Alin Mountains region can be considered the site of origin of S. keyserlingii (Malyarchuk et al., 2010). This conclusion is based on the fact that only this part of the range shows maximum diversity of the mtDNA clades. All mtDNA monophyletic haplogroups – clades A, B, C1, C2, and C3 – were detected there, while the whole range westward and northward from the Khabarovsk region is populated by the Siberian salamanders belonging only to clade C1 (Malyarchuk et al., 2010). These data suggest that such selective dispersal pattern of the Siberian salamander could be mediated by selection of the animals most adapted to habitation in low temperature conditions. Interestingly, a close relative of the Siberian Salamander, the Schrenck salamander, did not extend its range as far as the Siberian salamander, even though it is thought to be a more ancient species based on the ages of the cytochrome b gene haplotype divergence (Berman et al., 2005; Matsui et al., 2008; Malyarchuk et al., 2009). Thus, as only the Siberian salamander occupies the most northern areas, it is possible that this species had an earlier opportunity to expand into formerly inaccessible territory and subsequently blocked the Schrenck salamander from entering the same territory.

To obtain more information concerning the phylogeny and genetic history of the Siberian salamander, we analyzed here the whole mitochondrial genome sequences in the Siberian salamander, as well as in the related Schrenck salamander, reported in this study for the first time.
2. Materials and methods

2.1. Samples collection, DNA extraction, PCR amplification and sequencing

We examined a total of 26 Siberian salamanders from different regions of Eurasia (Table 1, Fig. 1) and one Schrenck salamander (S. schrencki) from the Primorye region (Russia). Total DNA was purified from various tissues of adults, larvae, and embryos, which were frozen or fixed with 70% ethanol. DNA was extracted by the standard method involving tissue lysis in a solution containing 100 mM Tris–HCl (pH 8.0), 10 mM EDTA, 100 mM NaCl, 1% sodium dodecyl sulfate, and 0.2 mg/ml proteinase K (Sigma, USA) at 56 °C for 12–16 h, with subsequent phenol–chloroform deproteinization.

The sequencing strategy proposed by Zhang and Wake (2009) was used for the Siberian salamander mtDNA sequencing. A suite of 22 PCR primers was used to amplify overlapping fragments that covered the entire mitochondrial genome of the Siberian salamander (Table S1). Initially, primers were designed on the basis of complete mtDNA sequence of S. keyserlingii presented in GenBank under accession number DQ333814 (Zhang et al., 2006). For PCR fragments of the Schrenck salamander mtDNA, specific primers were designed according to newly obtained sequences to facilitate primer walking.

Sequencing was performed directly with the corresponding PCR primers using the BigDye Terminator version 3.1 cycle sequencing kit (Applied Biosystems) in an ABI 3130 or ABI 3500xL genetic analyzers following the manufacturer’s instructions. Multiple alignments of nucleotide sequences were prepared for complete mitochondrial genomes by using ClustalW at default settings as implemented in MEGA 5.05 (Tamura et al., 2011).

2.2. Data analysis

To reconstruct the tree topology for complete mtDNAs and the 16S rRNA–CO1 gene haplotypes, neighbor-joining (NJ) and maximum likelihood (ML) analyses were used as implemented in MEGA 5.05. For ML analysis, we chose the HKY (Hasegawa–Kishino–Yano) model of evolution with the gamma shape parameter and proportion of invariable sites (HKY+G+I) as the best model of evolution that fits our data, by MEGA 5.05.

The confidence of branches in NJ and ML trees was assessed using non-parametric bootstrapping searches of 1000 replicates. Topologies of NJ and ML trees with bootstrap values 70% or greater were regarded as sufficiently resolved (Huelsenbeck and Hillis, 1993). Maximum parsimony (MP) networks of complete mtDNA genomes were constructed by the median-joining (MJ) algorithm (Bandelt et al., 1999), using the Network 4.6 and DNA Alignment 1.3 packages (www.fluxus-engineering.com). Homoplasic sites were revealed by use of the phylogenetic network analysis as implemented in the Network 4.6 program.

Bayesian inference (BI) of phylogeny using the Markov Chain Monte Carlo technique (MCMC) was produced from whole mtDNA sequences and the 16S rRNA–CO1 gene haplotypes by means of the program BEAST v.1.6.1 (Drummond and Rambaut, 2007) with the HKY+G+I model of DNA substitution. We initiated three independent analyses with a random starting tree that ran for 30 million generations sampled every 3000 steps, with the first 3 million generations regarded as burn-in. Posterior probabilities 95% or greater were considered significant support (Leaché and Reeder, 2002). We used the mtDNA sequence of the Schrenck salamander as an outgroup. Changes in effective population size across time were inferred using Bayesian skyline analyses, which enable estimation of past population dynamics through time from a sample of modern DNA sequences (Drummond et al., 2005). The Bayesian skyline plots (BSPs) produced from the whole mtDNA sequences were obtained with a piecewise linear model using BEAST v.1.6.1.

In Bayesian analysis, to test the hypothesis of clock-like evolution, we have used the uncorrelated relaxed lognormal molecular clock and measured the ucld.stdev parameter (the standard deviation of the uncorrelated lognormal relaxed clock). Parameter estimate close to 0.0 (ucld.stdev = 0.129–0.135) indicates that genomic evolution is nearly clock-like, with no variation among lineages in a tree. To estimate times to the most recent common ancestors (TMRCA) of the Siberian salamander mtDNA clades, a relaxed molecular clock with the evolutionary rate for salamanders (0.64% per million years (My) per lineage), as proposed by Weisrock et al. (2001), was used. This rate was calculated based on variation of the segment of the mitochondrial genome encompassing three protein genes (ND1, ND2, and CO1) and eight tRNA genes, and derived from four geologic events in salamandrid salamanders (Weisrock et al., 2001).

Haplotype and nucleotide diversity indices and their variances as well as neutrality tests (Tajima’s D, Fu and Li’s D’, Fu and Li’s
were calculated using DnaSP version 5.0 software package (Librado and Rozas, 2009). Gene conversion was searched using an algorithm (Betran et al., 1997) as implemented in the DnaSP program. Calculations of different estimators of the population parameter \( \theta = \theta(S) \) and \( \theta(F) \) were performed using Arlequin 3.5 software package (Excoffier and Lischer, 2010).

Fig. 1. Map showing sampled localities of *S. keyserlingii* and geographical distribution of the main mtDNA clades.
Sliding-window analysis was performed on the aligned complete mitochondrial genome nucleotide sequences of Salamandrella species, as well as some representatives of Hynobiidae family, such as Onychodactylus fischeri (NC_008089; Zhang et al., 2006), Hynobius amjiensis (NC_008076; Zhang et al., 2006), Ranodon sibiricus (NC_004021; Zhang et al., 2003) and Batrachuperus tibetanus (NC_008085; Zhang et al., 2006). Pairwise comparisons were performed to estimate interspecific divergence (Dxy; average number of nucleotide substitutions per site between species) and intraspecific nucleotide diversity (Pi) using sliding windows of 300 nucleotide positions (np) and steps of 10 np as implemented in the DnaSP program.

The ratio of the number of nonsynonymous substitutions per nonsynonymous sites ($K_a$) to the number of synonymous substitutions per synonymous sites ($K_s$) indicates the level of selection against nonsynonymous substitutions relative to synonymous ones. $K_a/K_s$ distributions in pairwise comparisons between DNA sequences were calculated (Nei and Gojobori, 1986) using DnaSP. Ka/Ks > 1 indicates positive selection, Ka/Ks < 1 negative (purifying) selection, and Ka/Ks = 1 neutrality. A codon-based Z-test was used to test for a significant difference between two estimates, with the variance of the difference computed using the bootstrap method (1000 replicates) in MEGA 5.05.

Using the DnaSP program, the McDonald–Kreitman test was performed for comparison of between-species divergence ($D_1$) and within-species polymorphism ($P_1$) at nonsynonymous ($D_n$, $P_n$) and synonymous ($D_s$, $P_s$) sites to infer adaptive protein evolution (McDonald and Kreitman, 1991). For between-species divergence estimates, the mtDNA sequence of S. schrencki was used. $N_I$ is a neutrality index calculated as a ratio of $P_{s}\!/P_{n}$ to $D_{s}\!/D_{n}$, and its value should be 1.0 under selective neutrality. In the presence of negative selection $N_I$ should be >1, while in the presence of positive selection $N_I$ should be <1.

Significant physicochemical amino acid changes among residues in mitochondrial protein-coding genes were identified by the algorithm implemented in TreeSAAP 3.2, which compares the observed distribution of physicochemical changes inferred from a phylogenetic tree with an expected distribution based on the assumption of completely random amino acid replacement expected under the condition of selective neutrality (Woolley et al., 2003). Positive selection is detected when the number of inferred amino acid replacements significantly exceeds the number expected by chance alone, yielding positive Z-scores. Z-scores from 1 to 3 indicate small variation in the amino acid characteristics, while scores from 6 to 8 represent the most radical substitutions, those that cause local directional shifts in biochemical function, structure, or both (McClellan et al., 2005). Therefore, in the present study, to detect strong directional selective pressures as a consequence of adaptive evolution, only changes corresponding to categories 6–8 at the $P < 0.001$ level were considered, following McClellan et al. (2005).

3. Results

3.1. Sequence data analysis

We sequenced entire mitochondrial genomes of 26 Siberian salamander samples (GenBank accession numbers JX508739–JX508764) and one Schrenck salamander (JX508765) (Table 1). The final alignment used for phylogenetic analysis consisted of 16356 bp, of which 1692 were variable and 438 were parsimony-informative. Each of the sequences represents a unique haplotype. Mitochondrial genomes of two Salamandrella species differ in length by 3 point deletions and 6 point insertions of nucleotides and are characterized by 1050 fixed nucleotide differences, with interspecies divergence equal to 7.8% (Table S2). The length of the Schrenck salamander mitochondrial genome is 16,342 np, and the size of the Siberian salamander mtDNA varied from 16,334 to 16,340 np. In the Siberian salamander mtDNAs, 670 positions were variable and 369 were parsimony-informative.

Complete mitochondrial genomes of Salamandrella species were used to estimate interspecific nucleotide divergence (Dxy) and intraspecific nucleotide diversity (Pi) across the genome as revealed by sliding-window analysis using DnaSP (Librado and Rozas, 2009). Lowest divergence and diversity values were observed in the comparison between D-loop regions: Dxy values varied from 0.2% to 1%, and Pi values varied from 0% to 0.1%. Gene by gene Dxy and Pi values were highly variable (Fig. S1). The highest interspecific differences were observed in ND2, ND4L and ND4 genes (up to 15%), and in ND5 and CYTB (up to 12%). Nucleotide diversity in S. keyserlingii varies almost similarly with Dxy, with highest values of Pi (about 1.5%) detected in ND2, CO1, CO3, ND4L, ND4 and CYTB. Interspecific divergence and intraspecific diversity values are not high in 12S and 16S rRNA genes, being estimated at interval of 1.5–6% for Dxy and 0.08–0.6% for Pi.

Calculations of the nucleotide diversity in the D-loop region, rRNA genes and in the structural genes suggested that the variation among the 12S and 16S rRNA genes is quite low in comparison with the protein-coding genes (Table S2). The variation in the D-loop region was also lower than that in structural genes (with the exception of the CO2 gene), being ~0.6%.

Nucleotide sequences of 16S rRNA and CO1 genes appear to be the commonly used mitochondrial “barcodes” for amphibians (Vences et al., 2005; Maya-Soriano et al., 2012; Xia et al., 2012). Results of hynobiid barcoding have shown that the standard barcoding marker CO1 (with its divergence threshold value equal in amphibians to 10%), in contrast to 16S rRNA, can identify species of Asiatic salamanders (Xia et al., 2012). However, interspecific comparisons at the complete mtDNA-based level indicate that highest values of Dxy (>20%) are characteristic for ND2 and CYTB genes, so in hynobids, both CO1 and 16S rRNA regions used in barcoding studies are of relatively low divergence. To reach this conclusion, we compared the mitochondrial genomes of S. keyserlingii with S. schrencki, as well as other hinobiid species in genera Hyobius, Batrachuperus, Ranodon and Onychodactylus (data not shown).

3.2. The phylogeny of mtDNA clades in the Siberian salamander

Mitochondrial genomes were employed to build a detailed phylogeny (Fig. 2). Bayesian phylogenetic tree revealed two nonoverlapping clades, AB and C, with maximal nodal support (100% posterior probabilities). In turn, clade AB consists of clades A and B. This result was further supported by the ML and NJ trees (data not shown). Median-network analysis designed for the reconstruction of all possible MP trees from a given data set also allowed the identification of clade AB (with interior clades A and B) and clade C, each characterized by substantial substructure (Fig. S2).

Among the 26 complete sequences, 5 belonged to clades A and B (Fig. 1 and Fig. S2). Thus, most mtDNAs formed a separate large clade C, differing by 156 mutations from the AB branch. Mitochondrial genomes belonging to clade C clustered within two clades, C1 and C2. Clade C2 includes four haplotypes, two of which belong to clade C2a1, which is specific for the Sakhalin population of the Siberian salamander. Continental C2a–sequences found in the Khabarovsk region (SK8250) formed a separate branch, differing by 41 mutations from the Sakhalin C2a1–clade (Fig. S2).

Clade C1 is the most represented in our phylogeny with 17 distinct haplotypes. This clade harbors both the Far Eastern and North Eurasian samples. Clade C1a comprises three samples from Far East (Khabarovsk region), whereas C1b connects mtDNA sequences from a wide range of geographical areas, including North-East.
China, the whole of Siberia and North-East Europe. C1b is the most structured clade, with two major clades, C1b1 and C1b2. Five sequences from North-East Asia (Magadan region, Chukotka and Kamchatka) formed clade C1b1. C1b2 comprises three further clades: Chinese clade C1b2b (from Heilongjiang Province) marked by only two mutations, separate branch C1b2 found in the Far East (Jewish autonomous region) and clade C1b2a represented by salamanders from several geographical areas: South, Central and West Siberia, the Urals, and North-East Europe (Fig. S2).

Sample DQ333814 sequenced by Zhang et al. (2006) originates from China (Heilongjiang Province) and according to our results can be classified as belonging to clade C1b2a. However, we do not use this sample in our further analyses because some sequence ambiguities were revealed in the DQ333814 sample [i.e., unidentified nucleotides and multiple amino acid substitutions in ND2 (Y59H, H92L, L93F, S183F, V193G) and CYTB (K312N, Q313K) genes in comparison with our data set].

The nucleotide divergence (Dxy) between clades AB and C is 1.53%, much higher than divergence values between clades C1 and C2 (0.47%), C1a and C1b (0.042%), and C1b1 and C1b2 (0.028%). Analysis of the median network (Fig. S2) indicates that among 670 variable sites, 79 sites had recurrent nucleotide substitutions. Homoplasic events [i.e. independent (parallel) mutations occurring at particular sites during the evolution DNA lineages] were identified in 52 sites, and reversions (or back mutations) were observed in 32 sites.

Simultaneous occurrence of several homoplasies in relatively short nucleotide tracts provides a means for evaluating the possibility of recombination between maternal and paternal mtDNA, which has been suggested to take place in animals (Tsaousis et al., 2005; Sun et al., 2011b; Malyarchuk, 2012). Application of the gene conversion algorithm (as implemented in DnaSP program) allowed us to identify two tracts characterized by homoplasy at two and three sites in nucleotide sequences of 80 and 323 np in length, respectively (Fig. S2). Both nucleotide tracts detected in clade B individuals (SK363 and SK8391) were, in fact, clade C-specific. Therefore, this may indicate that these homoplasic events had arisen as either parallel mutations or a consequence of

**Fig. 2.** Complete mtDNA based Bayesian inference (BI) of the intraspecific phylogeny of the Siberian salamanders (*S. keyserlingii*) under the HKY+I+G model of nucleotide substitutions. The Schrenck salamander (*S. schrenckii*) was used as an outgroup. The analysis was run for 30 million iterations, with the first 10% discarded as burn-in. The statistical supports (in %) are listed in the order NJ/ML/BI.
the recombination. Future studies may help resolve above the problem (Chen et al., 2009; Sun et al., 2011b).

3.3. Age estimates of the Siberian salamander mtDNA clades

The Bayesian estimate of the age of mtDNA diversity in the Siberian salamanders based on the 26 complete mtDNA sequences is about 1.36 My according to the mutation rate proposed by Weisrock et al. (2001) (Table 2). Evolutionary ages of clades C, C1, C1a, C1b and C2 are very close to each other, with values ranging from 310 to 420 ky, while the age of clade AB is higher, corresponding to 780 ky. Much lower age estimates were obtained for clades C1b1 (80 ky) and C1b2a (150 ky), specific for North Eurasian regions.

We utilized our samples of complete mtDNA genome sequences to estimate population size changes over time, using Bayesian skyline analysis. The BSPs for species and major clades of the Siberian salamanders based on the 26 complete mtDNA sequences (Fig. 3 and Fig. S3). Thus, the BSPs suggest that the Siberian salamanders of all clades (the Far Eastern A, B, C1a and C2, as well as North Eurasian C1b) went through a bottleneck in the cold period around the Last Glacial Maximum about 25 ky.

3.4. Testing for the signatures of selection in the Siberian salamander mitochondrial genome

Haplotypes in clade AB are much more divergent from each other than are those in clade C taking into consideration such parameters as the average number of pairwise nucleotide differences and nucleotide diversity (Table 3). Only clade C displayed significantly negative values of sequence-based neutrality tests (Tajima’s D and Fu and Li’s F*), suggesting recent demographic expansion or the influence of positive or weak negative (purifying) selection (Tajima, 1989; Rogers and Harpending, 1992; Aris-Brosou and Excoffier, 1996). Results of the MacDonald–Kreitman test (with the use of S. schrenckii as an outgroup for interspecies analyses) also demonstrated that all but one clade (AB) underwent statistically significant purifying selection (NI > 1) (Table 4). To identify which of the mitochondrial proteins were most influenced by purifying selection, we applied the MacDonald–Kreitman test to each of the 13 mitochondrial protein-coding genes (Table S3). Our results suggest that six genes (ND1, NDS, CO1, CO2, CO3 and ATP6), mainly in clade C, have undergone stronger selective constraints than others.

We estimated the ratio of the number of nonsynonymous substitutions per nonsynonymous sites (Ka) to the number of synonymous substitutions per synonymous sites (Ks) and found low Ka/Ks values in both clades (Ka/Ks = 0.04 and 0.1 in clades AB and C, respectively), indicating the influence of negative selection (Table 3). Meanwhile, in the Siberian salamanders, relatively high Ka/Ks value (0.258) has been detected in North-East Asian clade C1b1. This is due to very low synonymous substitution rate (Ks = 0.003), on the order less than in other clades (Table 3). Additional analysis of purifying selection using Z-test has shown that the possibility of relaxation of purifying selection for clade C1b1 cannot be excluded, because values of P were of two kinds: P > 0.027 using the Li-Wu-Luo method (Li et al., 1985), but P = 0.113 using the Kumar method (Nei and Kumar, 2000). Consequently, Z-test did not unambiguously support positive selection, although it is possible that the branch leading to the common ancestor of C1b1-Siberian salamanders has undergone relaxation of purifying selective pressure, connected probably with adaptation to extreme environments of the North-East Asia. However, additional studies are required to confirm this suggestion due to the small sample size of C1b1-Siberian salamanders analyzed.

The program TreeSAAP indicated that 20 amino acid sites in different genes of the Siberian salamanders have been affected by positive selection during mtDNA evolution (Table S4, Fig. S2). The Z-scores for these substitutions were positive and significant (z > 3.09, P < 0.001), indicating radical changes in the physicochemical amino acid properties. Singular and clade-specific amino acid

### Table 2

<table>
<thead>
<tr>
<th>Species/clades</th>
<th>Sample size</th>
<th>Time (and 95% HPD of the posterior probability distribution), in Mya</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. keyserlingii</td>
<td>26</td>
<td>1.389 (1.153–1.636)</td>
</tr>
<tr>
<td>AB</td>
<td>5</td>
<td>0.776 (0.618–0.944)</td>
</tr>
<tr>
<td>C</td>
<td>21</td>
<td>0.424 (0.362–0.495)</td>
</tr>
<tr>
<td>C1</td>
<td>17</td>
<td>0.372 (0.315–0.431)</td>
</tr>
<tr>
<td>C1a</td>
<td>3</td>
<td>0.34 (0.281–0.4)</td>
</tr>
<tr>
<td>C1b</td>
<td>14</td>
<td>0.307 (0.253–0.361)</td>
</tr>
<tr>
<td>C1b1</td>
<td>5</td>
<td>0.079 (0.047–0.114)</td>
</tr>
<tr>
<td>C1b2a</td>
<td>6</td>
<td>0.147 (0.106–0.192)</td>
</tr>
<tr>
<td>C2</td>
<td>4</td>
<td>0.404 (0.333–0.481)</td>
</tr>
<tr>
<td>S. keyserlingii</td>
<td>26</td>
<td>1.389 (1.153–1.636)</td>
</tr>
<tr>
<td>AB</td>
<td>5</td>
<td>0.776 (0.618–0.944)</td>
</tr>
<tr>
<td>C</td>
<td>21</td>
<td>0.424 (0.362–0.495)</td>
</tr>
<tr>
<td>C1</td>
<td>17</td>
<td>0.372 (0.315–0.431)</td>
</tr>
<tr>
<td>C1a</td>
<td>3</td>
<td>0.34 (0.281–0.4)</td>
</tr>
<tr>
<td>C1b</td>
<td>14</td>
<td>0.307 (0.253–0.361)</td>
</tr>
<tr>
<td>C1b1</td>
<td>5</td>
<td>0.079 (0.047–0.114)</td>
</tr>
<tr>
<td>C1b2a</td>
<td>6</td>
<td>0.147 (0.106–0.192)</td>
</tr>
<tr>
<td>C2</td>
<td>4</td>
<td>0.404 (0.333–0.481)</td>
</tr>
<tr>
<td>S. keyserlingii</td>
<td>26</td>
<td>1.389 (1.153–1.636)</td>
</tr>
<tr>
<td>AB</td>
<td>5</td>
<td>0.776 (0.618–0.944)</td>
</tr>
<tr>
<td>C</td>
<td>21</td>
<td>0.424 (0.362–0.495)</td>
</tr>
<tr>
<td>C1</td>
<td>17</td>
<td>0.372 (0.315–0.431)</td>
</tr>
<tr>
<td>C1a</td>
<td>3</td>
<td>0.34 (0.281–0.4)</td>
</tr>
<tr>
<td>C1b</td>
<td>14</td>
<td>0.307 (0.253–0.361)</td>
</tr>
<tr>
<td>C1b1</td>
<td>5</td>
<td>0.079 (0.047–0.114)</td>
</tr>
<tr>
<td>C1b2a</td>
<td>6</td>
<td>0.147 (0.106–0.192)</td>
</tr>
<tr>
<td>C2</td>
<td>4</td>
<td>0.404 (0.333–0.481)</td>
</tr>
</tbody>
</table>

Fig. 3. Bayesian skyline plot derived from 26 complete mitochondrial genomes of the Siberian salamanders. The x-axis is the time from present in units of million years, and the y-axis is equal to Neu (the product of the effective population size and mutation rate). The thick solid line is the median estimate and the thin lines show the 95% highest posterior density limits estimated with 30 million chains with the first 3 million generations regarded as burn-in.
replacements were distributed equally (Table S4). Radical amino acid changes defining the whole clades appear to be most important. For instance, five changes at once – in genes ND5 (I48V), ND1 (I270V), ND2 (A321T), CO1 (Y206H in S. keyserlingii), and synonymous substitutions that were fixed between species (except the ND6). Statistically significant values (P < 0.05) are shown in bold.

<table>
<thead>
<tr>
<th>Species/clades</th>
<th>n</th>
<th>k</th>
<th>Ps</th>
<th>Neutrality tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. keyserlingii</td>
<td>26</td>
<td>121.7</td>
<td>0.008</td>
<td>-1.267</td>
</tr>
<tr>
<td>AB</td>
<td>5</td>
<td>105.2</td>
<td>0.006</td>
<td>0.333</td>
</tr>
<tr>
<td>C</td>
<td>21</td>
<td>58.7</td>
<td>0.004</td>
<td>-1.825</td>
</tr>
<tr>
<td>C1</td>
<td>17</td>
<td>49.3</td>
<td>0.003</td>
<td>-1.599</td>
</tr>
<tr>
<td>C1b</td>
<td>14</td>
<td>39.5</td>
<td>0.002</td>
<td>-1.252</td>
</tr>
<tr>
<td>C1b1</td>
<td>5</td>
<td>10.3</td>
<td>0.001</td>
<td>-0.183</td>
</tr>
<tr>
<td>C1b2</td>
<td>9</td>
<td>38.9</td>
<td>0.002</td>
<td>-1.264</td>
</tr>
<tr>
<td>C2</td>
<td>4</td>
<td>61.8</td>
<td>0.004</td>
<td>-0.579</td>
</tr>
</tbody>
</table>

Note: n, number of sequences used; k, average number of nucleotide differences; Ps, nucleotide diversity; Ks/Ka, ratio of nonsynonymous and synonymous substitutions at nonsynonymous and synonymous sites in mtDNA protein-coding genes (except the ND6).

Table 3
Parameters estimated using the neutrality tests and the Ks/Ka approach along the complete mtDNA sequence of S. keyserlingii.

Note: Protein-coding genes (except the ND6) were analyzed. Mitochondrial genome of S. schrenckii was used for interspecific comparison. Number of nonsynonymous and synonymous substitutions that were fixed between species (Dk, Ds) and that were polymorphic within one of those species (Ps) is shown. NI is neutrality index calculated as described in Section 2. P-values were determined with Fisher’s exact test. Significant values are shown in bold.

4. Discussion

The present study demonstrates that the highest diversity of the Siberian salamander mtDNA clades is observed in the Russian Far East (the northern Sikhote-Alin Mountains and the middle Amur River Basin) and Manchuria. This suggests that these territories can be considered the center of genetic diversification of S. keyserlingii. This is generally consistent with the results of previous studies of cytochrome b gene variation in the Siberian salamander populations (Poyarkov and Kuzmin, 2008; Mal'yarchuk et al., 2010). Salamandrella is one of multiple taxa for which the Russian Far East, the northern part of the Korean Peninsula and North-East China demonstrate distinct genetic subdivisions explained by ecological factors accompanying the changing climatic conditions during the Pleistocene (Serizawa et al., 2002; Oshida et al., 2005; Haring et al., 2007; Fedorov et al., 2008; Li et al., in press). For instance, the Primorye region is considered a refugium, well-known for its highly stable natural conditions, at least, in the Late Quaternary; unlike other areas of the boreal Palearctic, this region did not suffer widespread extinction of populations during glaciations (Nazarenko, 1990).

Our study suggests that genetic history of the Siberian salamander was affected by Pleistocene glacial cycles. Using Bayesian skyline plot analysis, an expansion of population size of S. keyserlingii is placed in the Middle Pleistocene during the Mindel-Riss interglacial (≈ Tobolsk, 200–375 kya). The Siberian salamander later experienced a severe bottleneck around 25 kya, coincident with the coldest period around the Last Glacial Maximum. During this period, Siberia was covered by many local, especially highland, glaciations (Velichko, 1984; Schirrmeister et al., 2002), which prevented amphibian migrations. Results of molecular dating suggest that the Mindel-Riss interglacial was the most probable time of formation of major Siberian salamander clades (such as C, C1 and C2) in the Far East and neighboring regions, while some clades (C1b2a) may have started to colonize the western part of North Eurasia during the warm Kazantsyev interglacial (130–100 kya), and another clade (C1b1) may have reached North-East Asia (up to Chukotka) in the Early Zyrarian stage (100–50 kya).

However, results of molecular dating should be considered tentative because of uncertainties concerning the mutation rates. Unfortunately, hynobiid salamanders lack calibration points that could be used in the BEAST analyses of complete mtDNA genomes. Another problem is that the Siberian salamander data analyzed here do not represent all branches of mtDNA phylogeny revealed in previous studies of mtDNA segments (Matsui et al., 2008; Xia et al., 2012). For instance, there is a divergent Hokkaido-specific haplotype belonging to clade AB, but its position within this clade were found in Heilongjiang, this indicates that first split could occur near the Manchurian region.
is still unclear (Matsui et al., 2008; Malyarchuk et al., 2010). In addition, Manchurian clade D requires further analysis with complete mtDNA genomes. Thus, further sequencing of the entire mitochondrial genomes of the Siberian salamanders from these clades may clarify the estimation of divergence time between mtDNA branches and the origin of S. keyserlingii.

Unique cold tolerance of the Siberian salamander, as well as the ability of larvae to survive at very low concentrations of oxygen (Ishchenko et al., 1995), allowed the Siberian salamander to colonize various water pools and wetlands formed during the Early Holocene warming and the Pleistocene glaciers melting (Malyarchuk et al., 2010; Malyarchuk et al., 2011). In the present study, we have found that the vast area of Northern Eurasia was colonized, probably during the last 150 ky, by the Siberian salamander haplotype clade C1b alone. We have tested for selection in the Siberian salamander mitochondrial genome and found that the mitochondrial proteins were most influenced by purifying selection. This pattern appears to be characteristic of mitochondrial genomes of many species (Popadin et al., 2007; Shen et al., 2009; Castellana et al., 2011; Sun et al., 2011a). Meanwhile, relaxation of purifying selection has been detected in the North-East Asian clade C1b1. However, this was an artifact of decreasing the synonymous substitution rate (Table 3). In addition, clade C1b1 is characterized by the lowest values of molecular diversity indexes, associated with the effective population size and the mutation rate (Ewens, 1972) (Table S5). As $K_A/K_S$ values should negatively scale with effective population size, according to the nearly neutral theory (Ohta, 1992), it is possible that the higher ratio of $K_A/K_S$ found in clade C1b1–Siberian salamanders is due to their low effective population size. In this case, however, it is unclear whether this was caused by extreme environments in the North.

It is probable that adaptation to the North environments might have led to appearance of amino acid changes in mtDNA-encoded proteins that were advantageous in those climatic conditions and have allowed the Siberian salamanders to expand their range. Previously, we speculated that radical amino acid change at site 160 of cytochrome $b$, which was found in the Siberian salamanders belonging to clade C, may have some influence on bioenergetic functions.

Fig. 4. Neighbour-joining tree illustrating the phylogenetic relationships among the 16S rRNA-COI gene haplotypes in Salamandrella genus. The statistical supports (>70% obtained in ML and NJ analysis and >95% in Bayesian inference) are listed in the order NJ/ML/BI. Our samples correspond to those shown in the Table 1, while the others (designated by the use of "XM") have been taken from Xia et al. (2012). SS77 is S. schrenckii specimen from the Primorye region (Russia).
such as production of ATP and cellular respiration (Malyarchuk et al., 2010; see also McClellan and McCracken, 2001; Fink et al., 2004; McClellan et al., 2005; da Fonseca et al., 2008). In the present study, we have found five radical amino acid changes (z > 3.09, P < 0.001) in the clade C, but only one such change has occurred in the clade C1b1 (Table S4). Thus, possible evidence for positive selection is limited to only a few sites, while the signal for positive selection usually is masked by the continuous negative selection that occurs on most sites in a gene sequence (Ruiz-Pesini et al., 2004; Zhang et al., 2005; da Fonseca et al., 2008; Shen et al., 2009). Further studies of nuclear genes are needed to identify the signals of adaptive evolution of the Siberian salamander.

Acknowledgments

This research was supported by a grant from the Far Eastern Branch of the Russian Academy of Sciences (10-III-B-06-136). We are grateful to Dr. D.I. Berman (IBPN, Magadan) for samples of the Siberian salamanders.

Appendix A. Supplementary material

Supplemental data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2013.02.004.

References


