

Comparative analysis of inverted sequences of noncoding region in the mitochondrial DNA of the Baikal sponges (Fam. Lubomirskiidae)

Maikova O.O.^{1,*}, Balakshin D.D.^{1,2}, Belikov S.I.¹ 

¹ Limnological Institute, Siberian Branch of the Russian Academy of Sciences, Ulan-Batorskaya Str., 3, Irkutsk, 664033, Russia

² Irkutsk State University, Sukhe-Batora Str., 5, 664011, Russia

ABSTRACT. In recent years, mitochondrial DNA (mtDNA) has been used to solve many phylogenetic problems at different taxonomic levels. Previously, the nucleotide sequences of the mitochondrial genomes of the Baikal endemic sponge species of the Lubomirskiidae family have been identified. Their phylogenetic links are still being actively studied, and their systematics is considered to be not definitive. In order to study the mechanisms of speciation and evolution of mtDNA, special attention is paid to the investigations of noncoding DNA. We have studied the characteristics of the organization of the intergenic mtDNA regions in the Baikal sponges on the example of the longest region between the tRNATyr and tRNAIle genes of the species *L. baicalensis*, *S. papyracea*, *R. echinata*, and *B. intermedia profundalis*. A comparative analysis of the sequences has shown the presence of secondary structures represented by individual hairpins and complex multilevel structures. Based on a comparative analysis of secondary structures, we have suggested their role as both regulative elements and potential mobile elements. We have determined the continuity of these structures among representatives of different genera of the Lubomirskiidae family. The highest similarity in their distribution and localization was found in phylogenetically related species of the Baikal sponges.

Keywords: sponges, Lake Baikal, mitochondrial DNA, intergenic regions, secondary-structure elements

1. Introduction

The Baikal endemic sponges (the Lubomirskiidae family) are a bouquet of closely related species that diverged from cosmopolitan sponges approximately 10 million years ago (Maikova et al., 2015). To date, 14 species of the Lubomirskiidae family have been described, but so far their systematics is not definitive and requires further research (Efremova, 2001; 2004; Itskovich et al., 2017). In the past few years, sequences of the mitochondrial genome have been used to resolve the phylogenetic relationships of the Baikal sponges at various taxonomic levels.

Previously, the nucleotide sequences of the mitochondrial genomes of the Baikal endemic species *Lubomirskia baicalensis*, *Swartchewskia papyracea*, *Rezinkovia echinata*, and *Baikalospongia intermedia profundalis* were identified (Lavrov, 2010; Lavrov et al., 2012). Based on the nucleotide sequences of 14 protein-coding genes, a phylogenetic analysis of the Baikal sponges was performed. The analysis showed

that the mitochondrial DNA (mtDNA) of the Baikal sponges evolves at different rates. The nucleotide substitution rate in the protein-coding genes of the species *S. Papyracea* was twice higher relative to other species of the Lubomirskiidae family (Maikova et al., 2015). There were also different nucleotide substitution rates in different regions of the mitochondrial DNA. The accumulation of single substitutions in intergenic regions was four times higher than in protein-coding genes (Maikova et al., 2012). The causes of the different nucleotide substitution rates in different species and different mtDNA regions of the Baikal sponges are still unknown.

The evolution of mitochondrial genome of sponges has been well studied at high taxonomic levels. The presence of several common features characterize mitochondrial genomes in sponges of the Demospongiae class: minimally modified genetic code, the presence of several additional genes, conservatism of transport RNA (tRNA), and many noncoding regions. However, there were also significant genomic changes, mainly, in gene

*Corresponding author.

E-mail address: maikova@lin.irk.ru (O.O. Maikova)

rearrangement, number of tRNA genes and noncoding DNA. The length of mitochondrial DNA in most sponges of the class Demospongiae varied within 16-25 kbp (Wang and Lavrov, 2008; Gazave et al., 2010). The longest mitochondrial genome of Demospongiae was found in the Baikal endemic family Lubomirskiidae dominated by the species *L. Baicalensis*, approximately 30 kbp (Lavrov, 2010). The variability in the length of the mitochondrial DNA in the Baikal sponges is mainly an elongation of noncoding sequences, which comprise from 28.6 to 33.6 % of the total size of genome compared to 2-24 % in cosmopolitan and marine sponges (Wang and Lavrov, 2008). Direct and inverted repeats comprised the intergenic mtDNA regions in the Baikal sponges. For example, *L. Baicalensis* contained approximately 160 secondary structures, hairpins (Lavrov, 2010), *S. papyracea* – 97 such structures, *B. intermedia profundalis* – 96, and *R. Echinata* – 144 (Lavrov et al., 2012). Other Demospongiae representatives also contained such secondary structures, but the nature of their occurrence and distributions, as well as their functions remain unknown (Erpenbeck et al., 2009). Stable secondary structures were found in some sponges of the class Demospongiae, which repeatedly occurred in different phylogenetic lines during evolution. For example, *S. domuncula* had approximately 700 direct repeats with a predominant length of 12-15 bp, 100 inverted repeats (the longest repeat was 44 bp) and 100 palindromes. In *Vaceletia* sp., intergenic regions contained inverted repeats of considerable length (up to 339 bp). Such repeats were found in different regions. A+T-pairs comprised more than 90 % in these long repeats. In *Igernella notabilis* (Demospongiae, Keratosa), intergenic regions also contained AT-rich inverted repeats, which can form double and triple secondary elements. However, the mitochondrial DNA of most representatives of the class Demospongiae, which have intergenic regions with a length varying from 340 to 2134 bp, either cannot contain inverted repeats at all, or have short GC-rich hairpins. A comparative analysis of intergenic mtDNA regions in most Demospongiae sponges showed that the introduction and distribution of hairpin elements among sponges is uneven, i.e. some sponge families are more susceptible to their introduction (Keratosa and Myxospongiae), and others – less (Hexactinellida and Homoscleromorpha) (Lukic-Bilela et al., 2008; Erpenbeck et al., 2009; Lavrov, 2010).

Noncoding DNA plays a great role in the functioning of the entire genome. The inverted repeats in the noncoding mitochondrial DNA can contribute to intramolecular recombination and are involved in various regulatory processes, including replication and transcription (Ling et al., 2011; Kolesnikov and Gerasimov, 2012). Unlike the bilaterian animals, most non-bilaterian animals do not have a common control region for regulating replication and transcription of the mitochondrial DNA, and regulatory elements are distributed over intergenic regions (Lavrov and Pett, 2016). Probably, the accumulation of inverted repeats in the future may result in gene rearrangements and other significant structural transformations of the genome.

In order to study the distribution mechanisms of inverted repeats in noncoding mitochondrial DNA of Baikal sponges, we have carried out a comparative analysis of sequences of one of the longest regions located between the tRNA^{Tyr} and tRNA^{Leu} genes in representatives of all four genera of the Lubomirskiidae family. The length of the studied intergenic region varied from 300 bp in *E. fluviatilis* (Spongillidae) to 600 bp in *L. fusifera* (Lubomirskiidae), and these changes were associated with the presence of indels. Previously, based on the nucleotide sequences of this region, we performed a phylogenetic analysis. It showed that some of these indels were species-specific, and the sequences of the intergenic region appeared to be suitable for studying phylogenetic relationships of closely related sponges of the Lubomirskiidae family (Maikova et al., 2012).

2. Materials and methods

For a comparative analysis of noncoding mtDNA regions located between the tRNA^{Tyr} and tRNA^{Leu} genes, we used previously published sequences of the mitochondrial genomes for the species *L. baicalensis*, *S. papyracea*, *R. Echinata*, and *B. intermedia profundalis* (Lavrov et al., 2012). The sequences were aligned using the MAFFT v.6.240 and BioEdit programs. Pairwise distances based on indels of the noncoding regions using Jukes-Cantor (Jukes and Cantor, 1969) method, were determined using MEGA 4.0.

Inverted repeats were found using Unipro UGENE software (<http://ugene.unipro.ru/index.html>). The secondary elements were constructed at 4° C and visualized using Mfold webserver (version 3.2) (Zuker, 2003) and Vienna RNA package (RNAfold, RNAalifold) (Hofacker, 2003).

3. Results and discussion

We have carried out a detailed study of the characteristics of organization and distribution of secondary structures in the Baikal sponges on the example of the mtDNA region between the tRNA^{Tyr} and tRNA^{Leu} of four species of the Baikal sponges, *Lubomirskia baicalensis*, *Swartchewskaia papyracea*, *Rezinkovia echinata*, and *Baikalospongia intermedia profundalis*, which mitochondrial genomes were published previously (Lavrov, 2010; Lavrov et al., 2012). To name the hairpins, we used their previous classification, in which they are divided into 9 families, H1-H9 (Lavrov, 2010).

Modelling of the potential secondary structure of the nucleotide sequence has shown that all studied species form complex elements: individual hairpins (Fig. 1) and complex multilevel elements (Fig. 2). We have indicated a specific occurrence of some hairpins that form stable clusters. For instance, the structures H3 and H4 are mainly observed together and resemble double hairpin elements (Fig. 2) found by Paquin with co-authors (Paquin et al., 2000) in the mitochondrial DNA of *Allomyces macrogynus* fungi, which were described as mobile elements.

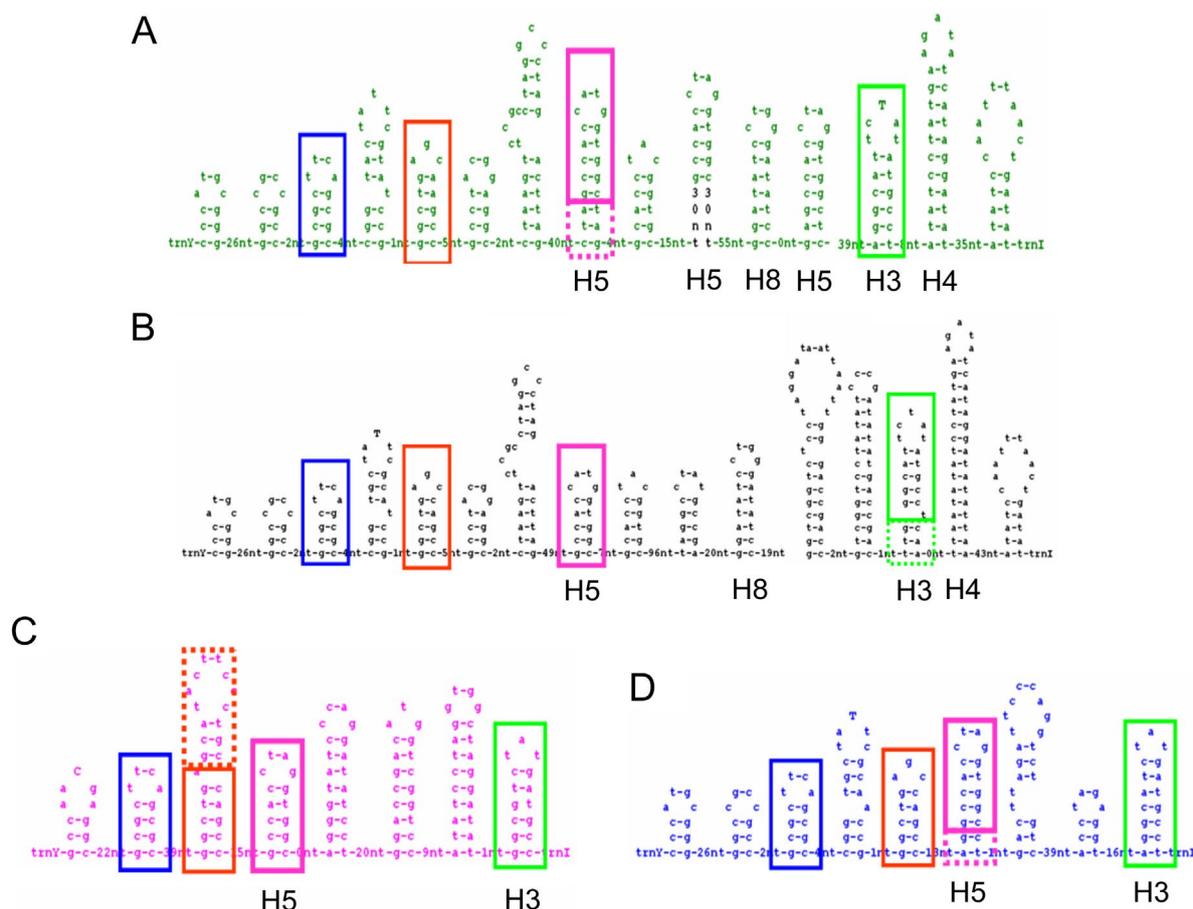


Fig. 1. Potential secondary elements located in the region of the mitochondrial DNA between the tRNA^{Tyr} and tRNA^{Ile} genes: A – *L. baicalensis*, B – *R. echinata*, C – *S. papyracea*, and D – *B. intermedia profundalis*.

In the studied intergenic region, we have found one conservative (marked in blue) and three homologous hairpins (Nos. 1-3, marked in red, pink and green, respectively). Notably, the conservative hairpin and homologous hairpin No.1 have been found for the first time. *L. baicalensis* and *R. echinata* have another common AT-rich H4-element. Homologous hairpins differ in insertion or deletion of complementary base pairs. The sequences of *L. baicalensis* and *R. echinata* have the highest similarity in the number and distribution of hairpins in the region between tRNA^{Tyr} and tRNA^{Ile} genes. At the same time, *R. echinata* and *S. Papyracea* show the highest similarity in the region between tRNA^{Ile} и tRNA^{Met} genes.

Three homologous and one conservative hairpins in all four species have GC-clusters increasing their GC-content. Table 1 shows the GC-content of these hairpins (excluding the nucleotide composition of the loops).

Using the data on the presence or absence of hairpins, we have determined the genetic distances of two intergenic regions between four studied species of the Baikal sponges. We observed minimal differences between *L. baicalensis* and *S. papyracea* (90%). Therefore, the example of two intergenic regions shows the general tendency of the maximum similarity of *L. baicalensis* and *R. echinata* in distribution and localization of secondary elements in the noncoding mito-

chondrial DNA. This can be explained by the presence of the common ancestor for these two species, which phylogenetic reconstructions confirm based on the nucleotide sequences of 14 protein-coding genes of the mitochondrial DNA (Maikova et al., 2015).

The presence of both, identical hairpins in all sponges of the Lubomirskiidae family and unique hairpins in certain species, may indicate their repeated introduction into the mitochondrial genome of the Baikal sponges during evolution. Interestingly, in the Baikal sponges, like in different Demospongiae representatives, the bulk of secondary elements found in the intergenic mtDNA regions have increased GC-content, which can reach 100 % (Gazave et al., 2010). This can also indicate the continuity of these structures at high taxonomic levels.

The presence of hairpins conservative in the nucleotide sequence and localization among representatives of different genera of closely related Baikal sponges suggests that they can have functional load.

Similar hairpins with GC-clusters were characterized as recombination sites in fungi (Paquin et al., 2000). This assumption is also supported by the fact that homologous hairpins vary significantly in loops, hence, they cannot be involved in various regulatory processes, such as the regulation of transcription. At the same time, the existence of conservative and homologous secondary elements in closely related sponge

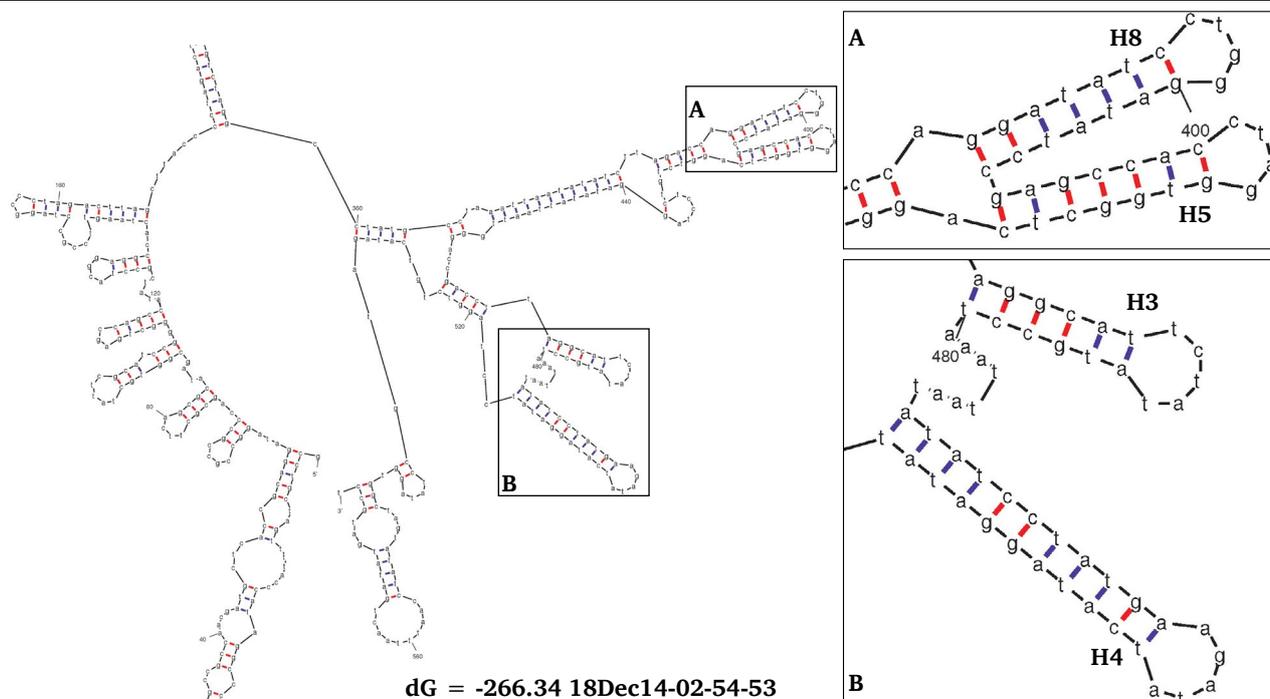


Fig. 2. Fragment of the sequence secondary structure from the mitochondrial DNA region of *L. baicalensis* located between the tRNA^{Tyr} and tRNA^{Leu} genes: A – possible combination of H8 and H5 hairpins, but not conservative in the mitochondrial DNA of the Baikal sponges, B – secondary structure conservative in the mitochondrial DNA of the Baikal sponges consisting of H3 and H4 hairpins.

species indicates a possible exchange of genetic material between them (Paquin et al., 2000; Erpenbeck et al., 2009).

4. Conclusions

Search and comparative analysis of inverted repeats has identified interesting features. Firstly, the mitochondrial DNA of the Baikal sponges has a large number of inverted repeats capable of forming hairpins, which is unusual for sponges. Secondly, we have found the continuity of these structures among representatives of different genera of the Lubomirskiidae family. The genetically closest species have shown the highest similarity in their distribution and localization. We have indicated that secondary elements are single and double hairpins similar to those of some other organisms in the mitochondrial DNA, where they are regarded as mobile elements. The results obtained within this study allow us to have a deeper insight into

not only mechanisms of explosive speciation of closely related species of the Baikal sponges, but also the evolution of the mitochondrial genome in general.

Acknowledgments

The study was performed within the framework of the LIN SB RAS State task No. 0345-2016-0002 (AAAA-A16-116122110066-1) “Molecular ecology and evolution of living systems of the Central Asia under global ecological changes”.

References

Efremova S.M. 2001. Sponges (Porifera). In: Timoshkin O.A. (Ed.), Index of animal species of Lake Baikal fauna and its catchment area. Novosibirsk, pp. 179–192. (in Russian)
 Efremova S.M. 2004. Novel genus and novel species of the Lubomirskiidae Rezvoj, 1936 sponge family. In: Timoshkin O.A. (Ed.), Index of animal species of Lake Baikal fauna and

Table 1. GC-content of secondary structures shown in four species of the Baikal sponges in the region between the tRNA^{Tyr} and tRNA^{Leu} genes. The colours correspond to those shown in Fig. 1.

Species	Conservative hairpin (blue)	Homologous hairpin No. 1 (red)	Homologous hairpin No. 2 (pink)	Homologous hairpin No. 3 (green)
<i>L. baicalensis</i>	100%	80%	62.5%	50%
<i>R. echinata</i>	100%	80%	80%	50%
<i>S. papyracea</i>	100%	75%	80%	66.6%
<i>B. int. prof.</i>	100%	80%	71.4%	57%

its catchment area. Novosibirsk, pp. 1261–1278. (in Russian)

Erpenbeck D., Voigt O., Wörheide G. et al. 2009. The mitochondrial genomes of sponges provide evidence for multiple invasions by Repetitive Hairpin-forming Elements (RHE). *BMC Genomics* 10: 1–14. DOI: 10.1186/1471-2164-10-591

Gazave E., Lapebie P., Renard E. et al. 2010. Molecular phylogeny restores the supra-generic subdivision of homoscleromorph sponges (Porifera, Homoscleromorpha). *PLoS ONE* 5: 12. e14290. DOI: 10.1371/journal.pone.001429

Hofacker I.L. 2003. Vienna RNA secondary structure server. *Nucleotide Acids Research* 31: 3429–3431. DOI: 10.1093/nar/gkg599

Itskovich V., Kaluzhnaya O., Veynberg Y. et al. 2017. Endemic Lake Baikal sponges from deep water. 2: Study of the taxonomy and distribution of deep-water sponges of Lake Baikal. *Zootaxa* 4236: 335–342. DOI: 10.11646/zootaxa.4236.2.8

Jukes T.H., Cantor C.R. 1969. Evolution of protein molecules. In: Munro H.N. (Ed.), *Mammalian Protein Metabolism*. New York, pp. 21–123.

Kolesnikov A.A., Gerasimov E.S. 2012. Diversity of mitochondrial genome organization. *Biochemistry* 77: 1424–1435. DOI: 10.1134/S0006297912130020

Lavrov D.V. 2010. Rapid proliferation of repetitive palindromic elements in mtDNA of the endemic Baikalian sponge *Lubomirskia baicalensis*. *Molecular Biology and Evolution* 27: 757–760. DOI: 10.1093/molbev/msp317

Lavrov D.V., Maikova O.O., Pett W. et al. 2012. Small inverted repeats drive mitochondrial genome evolution in Lake Baikal sponges. *Gene* 505: 91–99. DOI: 10.1016/j.gene.2012.05.039

Lavrov D.V., Pett W. 2016. Animal mitochondrial DNA as we do not know it: mt-genome organization and evolution in nonbilaterian lineages. *Genome Biology and Evolution* 8: 2896–2913. DOI: 10.1093/gbe/evw195

Ling W., Xuming Z., Liuwang N. 2011. Organization and variation of mitochondrial DNA control region in pleurodiran turtles. *ZOOLOGIA* 28: 495–504. DOI: 10.1590/S1984-46702011000400011

Lukic-Bilela L., Brandt D., Pojskic N. et al. 2008. Mitochondrial genome of *Suberites domuncula*: Palindromes and inverted repeats are abundant in non-coding regions. *Gene* 412: 1–11. DOI: 10.1016/j.gene.2008.01.001

Maikova O.O., Stepanova G.N., Belikov S.I. 2012. Variations in noncoding sequences of the mitochondrial DNA in sponges from family Lubomirskiidae. *Doklady Biochemistry and Biophysics* 442: 46–48. DOI: 10.1134/S1607672912010140

Maikova O., Khanaev I., Belikov S. et al. 2015. Two hypotheses of the evolution of endemic sponges in Lake Baikal (Lubomirskiidae). *Journal of Zoological Systematics and Evolutionary Research* 53: 175–179. DOI: 10.1111/jzs.12086

Paquin B., Laforest M.-J., Lang B.F. 2000. Double-hairpin elements in the mitochondrial DNA of Allomyces: Evidence for Mobility. *Molecular Biology and Evolution* 17: 1760–1768. DOI: 10.1093/oxfordjournals.molbev.a026274

Wang X., Lavrov D.V. 2008. Seventeen new complete mtDNA sequences reveal extensive mitochondrial genome evolution within the Demospongiae. *PLoS ONE* 3: 1–11. DOI: 10.1371/journal.pone.0002723

Zuker M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research* 31: 3406–3415. DOI: 10.1093/nar/gkg595