

SHORT COMMUNICATIONS

Phylogeny of Salmonids (Salmoniformes: Salmonidae) and its Molecular Dating: Analysis of Nuclear *RAG1* Gene

S. V. Shedko, I. L. Miroshnichenko, and G. A. Nemkova

Institute of Biology and Soil Science, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, 690022, Russia

e-mail: shedko@biosoil.ru

Received November 3, 2011

Abstract—The phylogenetic relationships among 26 species of salmonid fishes (family Salmonidae) were studied using the *RAG1* gene as phylogenetic marker. No unambiguous relationships between thymallins, coregonins, and salmonins it was possible to establish. It seems likely, that divergence of these lineages took place during rather short time interval (about 3 to 4 million years). The thymallins are thought to be the first separated lineage. The genera of the subfamily Salmoninae form two distinct monophyletic groups, represented by (1) *Brachymystax* and *Hucho* and (2) *Salmo*, *Parahucho*, *Salvelinus*, *Parasalmo* and *Oncorhynchus*. Ancestral forms of these two evolutionary lineages could diverge at the Oligocene–Miocene boundary (about 24 million years ago). It is suggested that diversification of the main lineages within the second group was rather rapid, and took place in middle Miocene (about 19–16 million years ago). Moreover, the lineages of *Salvelinus*, *Parasalmo* and *Oncorhynchus* were the latest to diverge. It seems likely that divergence of the *Prosopium* and *Coregonus* ancestral lineages occurred during the same time interval.

DOI: 10.1134/S1022795412050201

Salmonid fishes (family Salmonidae) are used as a model group in studies in evolutionary biology, genomics, and other fields of modern biology [1, 2]. It is evident that the existence of well-grounded phylogenetic scheme is the requirement for proper design of comparative studies and interpretation of their results. Despite of the recent progress in understanding of the phylogenetic relationships within salmonids, some basic issues still remain unresolved. First, the divergence order of three main lineages of Salmonidae, thymallins (subfamily Thymallinae), coregonins (Coregoninae), and salmonins (Salmoninae) is unclear. Second, there is no consensus on phylogenetic relationships of the main lineages within Salmoninae: *Hucho* and *Brachymystax*; *Parahucho*; *Salmo*; *Salvelinus*; *Parasalmo* and *Oncorhynchus*.

In the present study, these issues are examined using a proven phylogenetic marker, single-copy nuclear gene *RAG1* (recombination activating gene 1) [3, 4], which was new for this fish group. As a result, quite acceptable phylogenetic scheme of Salmonidae was constructed. In addition, with some suppositions, molecular dating of this phylogenetic scheme was performed.

A total of 62 DNA samples were examined. The samples were extracted with the help of standard phenol–chloroform technique from muscle tissues of 26 salmonid species, including *Thymallus arcticus* (Katun River, $n = 1$), *Th. tugarinae* (Kievka River, $n = 2$), *Coregonus migratoris* (Baikal Lake, $n = 2$), *Hucho taimen* (Selenga River, $n = 2$), *Brachymystax lenok* (Markakol Lake, $n = 1$; Ingoda River, $n = 1$: Armu

River, $n = 1$), *B. tumensis* (Onon River, $n = 1$; Komissarovka River, $n = 1$; Anan'evka River, $n = 1$), *Salmo trutta* (Luga River, $n = 2$), *Parahucho perryi* (Kievka River, $n = 7$), *Salvelinus alpinus* (Luomus Lake, Finland; Kanes Laddu, Finland, $n = 3$; goggle-eyed char from Lama Lake, $n = 3$; davatchan from Frolikha Lake, $n = 3$; Kolyma River, $n = 2$), *S. leucomaensis* (Maksimovka River, $n = 2$), *S. levanidovi* (Yama River, $n = 1$), *S. elgyticus* (El'gygytgyn Lake, $n = 2$), *S. boganidae* (El'gygytgyn Lake, $n = 2$), *S. taranetzi* (Vykvynaivaam River, $n = 1$), *S. malma* (Kamchatka River, $n = 1$; Vilyuchinsky Brook, $n = 1$), *S. curilus* (Severny Bay, $n = 1$; Tym' River, $n = 1$; Rasshua Island, $n = 2$; Maksimovka River, $n = 1$), *Oncorhynchus masou* (Lyutoga River, $n = 1$; Serebryanka River, $n = 1$), *O. keta* (Narva River, $n = 1$; Kievka River, $n = 2$), *O. gorbusha* ($n = 1$), *O. nerka* ($n = 2$), *O. kisutch* ($n = 1$), *O. tshawytscha* ($n = 1$), and *Parasalmo mykiss* ($n = 1$) from Kamchatka River.

The DNA fragment containing the larger part of the *RAG1* exon 2 was amplified with the help of polymerase chain reaction. The product obtained was purified and sequenced using the Big Dye Terminator 3.1 kit (Applied Biosystems, United States). The reaction products were fractionated using an ABI Prizm 3130 automated analyzer (Applied Biosystems, United States/Hitachi, Japan). All manipulations were carried out at the Institute of Biology and Soil Science, Russian Academy of Sciences, Russian Academy of Sciences, Vladivostok. Amplification and sequencing was performed using a set of primers described in [3], as well as personally designed primers

F2817 (GAAGCACAGCCGTCTCATCCTG) and F3269 (TGGACAARCAGCTGAGGAAGAAGA).

As a result, in each of 62 specimens examined, 1524 bp of the *RAG1* gene fragment were sequenced. For most of the specimens, no mixed signals in the ABI chromatograms were detected. However, in 13 specimens peak overlapping at 1 to 2 (1.6, on average) different nucleotide positions was identified. Such individuals were considered to be heterozygous for the *RAG1* fragment examined. Taken together, in 62 specimens a total of 23 *RAG1* variants, differing in unambiguous nucleotide substitutions, were detected. Typically, each of the 23 Salmonidae species was characterized by one to two unique variants. The exclusion were the specimens of *B. lenok* and *B. tumensis*, which were indifferent with respect to the *RAG1* fragment examined. The same was true for the specimens of *S. alpinus* and *S. taranetzi*, and one of two *S. elgyticus* individuals, which had one and the same *RAG1* variant. In addition to the sequences determined (deposited in the GenBank/NCBI under accession numbers GQ871466 to GQ871494), the analysis included the GenBank/NCBI *RAG1* sequences for: *Thymallus thymallus* (AY430214), *Prosopium williamsoni* (AY430213), *Salmo salar* (DW577969, EG768285), as well as *Salvelinus leucomaenis* (AY380535) from Kuril Islands, and *P. mykiss* (NM_001124737) from North America. The Esociformes species taken as outgroup were *Esox lucius* (AY380542) and *Umbra krameri* (AY380547). Thus, phylogeny of salmonid fishes was reconstructed using 30 sequence variants of the *RAG1* gene, which represented 26 species of Salmonidae and two outgroups.

The phylogeny of Salmonidae was reconstructed using the following approaches: (1) Bayesian (BA), based on the GTR + G nucleotide substitution model, selected with the help of the Modeltest 3.7 software program [5]; (2) maximum likelihood (ML) with the use of the same evolutionary model; (3) maximum parsimony (MP); (4) distance method, using neighbor joining (NJ) algorithm based on LogDet genetic distances.

The Markov chain Monte Carlo analysis (Bayesian phylogenetic inference) was performed using the MrBayes 3.2 software program [6] via simultaneous running of six chains (five hot and one cold) during 3×10^6 cycles sampling each 500th tree generated and discarding the first 1001 trees out of 6001 generated (burn-in = 1001).

Heuristic search of maximum likelihood or maximum parsimony trees (ML and MP trees) was conducted in the PAUP 4.0b10 software program in 40 rounds of random sequence addition and subsequent TBR swapping. The same program was used for the construction of NJ tree. Robustness of the ML, MP, and NJ tree branching patterns was evaluated using 1000 bootstrap pseudosamples.

Irrespective of the methods used, final phylogenetic trees appeared to be virtually the same. The BA,

ML, and MP tree topologies were identical (Fig. 1). The NJ tree differed from the first ones in two points, including sister position of *Thymallus* and *Coregonus* and clustering of *Parahucho*, *Salvelinus* and *Oncorhynchus*. However, bootstrap support of these alternative clusters appeared to be low (< 50%). Basic similarity of the results received using different phylogenetic reconstruction methods, along with highly statistically significant support of most of the clades identified enables considering phylogenetic tree in Fig. 1 as valid.

Noteworthy, the testing of molecular clock based on likelihood ratio test [8] showed that evolutionary rate of *RAG1* sequences in different parts of the tree was relatively homogenous. The difference between the likelihood estimates obtained for additive and ultrametric ML trees was not statistically significant. Specifically, $2\Delta = 2(-4963.02 - (-4976.32)) = 26.6$; $P = 0.54$ at $d.f. = 28$ (without outgroups, $2\Delta = 2(-3740.08 - (-3753.01)) = 25.9$, $P = 0.47$ at $d.f. = 26$). This property enabled building of ultrametric tree of Salmonidae, which was realized in Bayesian approach as implemented in the BEAST 1.4.8 software package [9]. The conditions used were as follows: molecular clock model, strict clock; nucleotide substitution model, GTR + R; the extent of Bayesian analysis, 3×10^6 cycles sampling each 1000th tree; burn-in, 601; outgroups were excluded from the analysis. To associate the ultrametric tree with absolute time scale, two calibration intervals were used. The first interval was based on numerous salmonid fossils, which displayed advanced characters of Pacific salmon from the genus *Oncorhynchus* (see [10] and references therein). This interval confined the divergence time of recent *Oncorhynchus* lineages to late Miocene (5.3 to 11.6 million years ago). The second interval confined isolation of the lineage of Salmonidae to the limits of Eocene (33.9 to 55.8 million years ago). This interval is associated with the discovery of well-preserved fossils of the *Eosalmo* salmon in Middle Eocene. These fossils possessed some advanced characteristics, which among modern Salmonidae, are found exclusively in the representatives of Salmoninae [11].

Ultrametric Bayesian tree built following the requirements mentioned (Fig. 2) was virtually similar to the topology of the tree in Fig. 1, and demonstrated the posterior probability estimates similar to ordinary Bayesian tree. Summation of the data obtained (Figs. 1, 2) led to the following conclusions:

(1) No unambiguous relationships between thymallins, coregonins, and salmonins were established. It seems likely that the divergence of these lineages took place during a relatively short time interval (about 3 to 4 million years). The thymallins are thought to be the first separated lineage.

(2) The genera of the subfamily Salmoninae form two distinct monophyletic groups, represented by (1) *Brachymystax* and *Hucho* and (2) *Salmo*, *Parahucho*, *Salvelinus*, and *Oncorhynchus*. The ancestral forms of

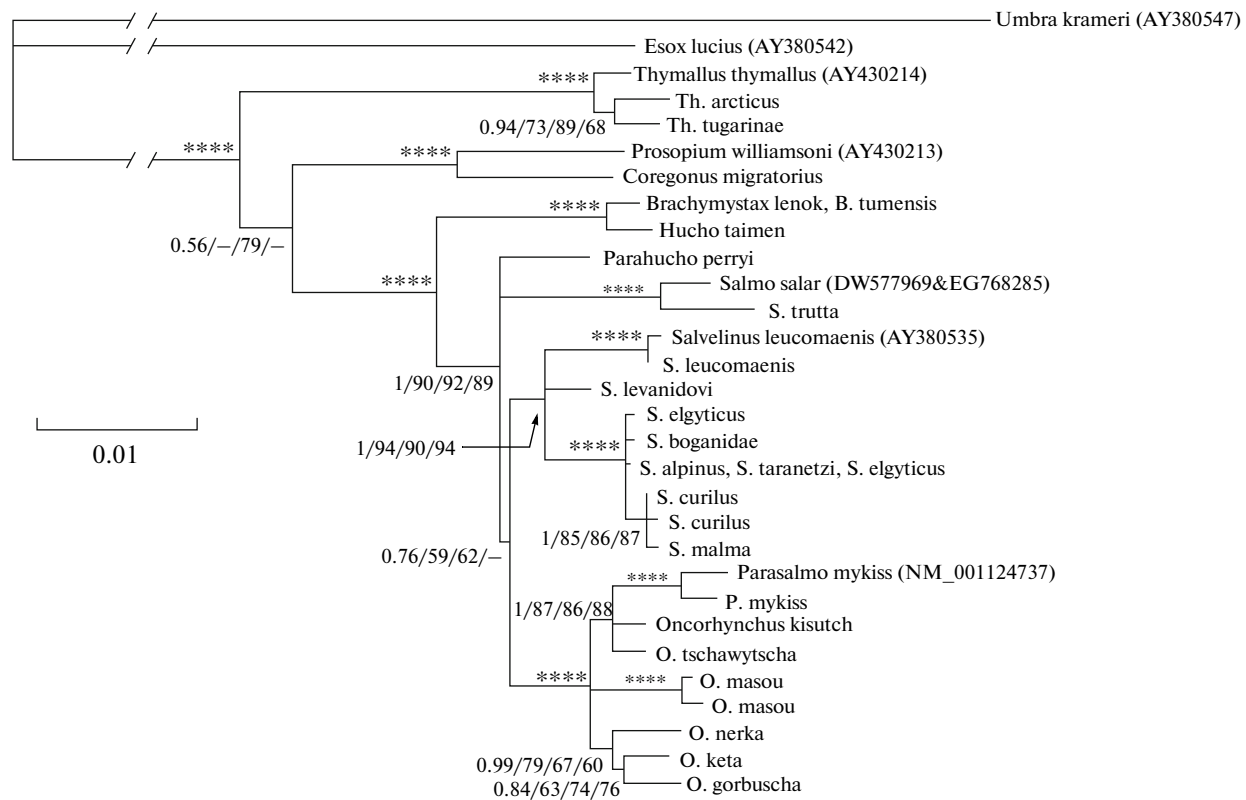


Fig. 1. Phylogram of salmonid fishes (ML tree), inferred from the data for *RAG1* gene fragment. Figures at the clade bases are the posterior probabilities for Bayesian tree, as well as the bootstrap values (% from 1000 bootstrap replicates) for ML, MP, and NJ trees, respectively. The reproducibility of a certain clade in less than 50% of bootstrap replicates is designated by a dash; *****, the highest support value (for this clade the values range constituted from 0.95 to 1.0, for Bayesian tree and from 95 to 100 for ML, MP, and NJ trees). The scale reflects the number of substitutions per nucleotide position.

these two evolutionary lineages could diverge at the Oligocene–Miocene boundary (about 24 million years ago).

(3) It is suggested that diversification of the main lineages within the second group was rather rapid, and took place in middle Miocene (about 19–16 million years ago). Moreover, the lineages of *Salvelinus*, *Parasalmo* and *Oncorhynchus* were the latest to diverge. It seems likely, that divergence of the *Prosopium* and *Coregonus* ancestral lineages occurred during the same time interval.

In general, the *RAG1* fragment sequenced proved its applicability for resolution of the tasks of the present study. This marker, like many other nuclear genes characterized by low evolutionary rate (according to our estimates, 0.57×10^{-9} substitutions per site per year), displayed distinct, with minimum homoplasy, phylogenetic signal (the consistency index for MP tree without outgroups, 0.79). The phylogenetic trees (Figs. 1, 2) agrees with the data of recent molecular genetic studies, based on the analysis of nuclear DNA, in a part of relationships between thymallins, coregonins, and salmonins [12], the divergence order of the main evolutionary lineages within

Salmoninae [13–15], phylogenetic relationships of *O. keta*, *O. gorbuscha*, and *O. nerka* [14, 16], and in some other details. The results of molecular dating of the Salmonidae cladogeny (Fig. 2) are to a large extent similar to the estimates obtained earlier using other approaches and based on other data [17, 18].

However, some other molecular genetic studies described another pattern of phylogenetic relationships between thymallins, coregonins, and salmonins [15, 19, 20]. Moreover, the position of *Brachymystax* and *Hucho* among other evolutionary lineages of Salmoninae, inferred from the analysis of mtDNA [14, 21] and introns of the GH1 gene [16], was different. In addition, position of the *Parasalmo* lineage relative to the species of *Oncorhynchus*, reported in different studies, was also different (Figs. 1, 2 [13, 14, 16, 18, 20–22]). Because of this, it seems reasonable that further clarification and detailed elaboration of the salmonidae phylogeny is performed using more phylogenetic markers and simultaneous analysis of many unlinked loci of nuclear genome and mtDNA.

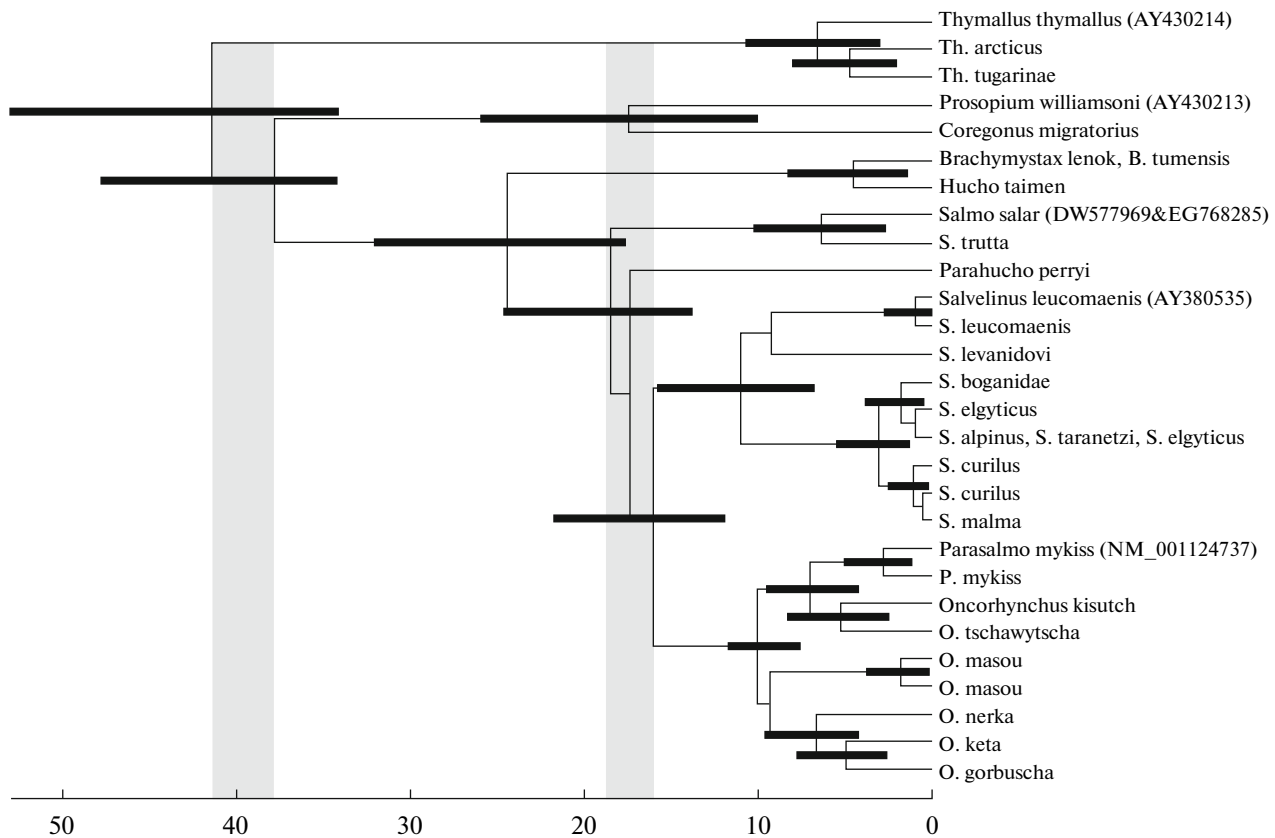


Fig. 2. Bayesian chronogram of salmonid cladogenesis, expressed in absolute time scale (million years). Boxes designate 95% intervals of the highest posterior densities of the clade lifetime estimates, characterized by posterior probabilities higher than 0.5. The periods that produced the main morphoecological diversity of Salmonidae are shown in dark.

ACKNOWLEDGEMENTS

The authors are thankful to O. A. Radchenko, D. S. Sendek, and M. B. Shedko for kindly providing the samples of some salmonid species.

This work was supported by the Russian Foundation for Basic Research (grant no. 09-04-00370-a) and the Far Eastern Branch of the Russian Academy of Sciences (grant no. 09-III-A-06-185).

REFERENCES

1. *Evolution illuminated: Salmon and Their Relatives*, Hendry, A.P. and Stearns, S.C., Eds., New York: Oxford Univ. Press, 2004.
2. Davidson, W.S., Koop, B.F., Jones, S.J.M., et al., Sequencing the Genome of the Atlantic Salmon (*Salmo salar*), *Genome Biol.*, 2010, vol. 11, no. 9, p. 403.
3. López, J.A., Chen, W.-J., and Ort, G., Esociform Phylogeny, *Copeia*, 2004, no. 3, pp. 449–464.
4. Chiari, Y., Meijden, A., Madsen, O., et al., Base Composition, Selection, and Phylogenetic Significance of Indels in the Recombination Activating Gene-1 in Vertebrates, *Front. Zool.*, 2009, vol. 6, p. 32.
5. Posada, D. and Crandall, K.A., MODELTEST: Testing the Model of DNA Substitution, *Bioinformatics*, 1998, vol. 14, no. 9, pp. 817–818.
6. Ronquist, F. and Huelsenbeck, J.P., MRBAYES 3: Bayesian Phylogenetic Inference under Mixed Models, *Bioinformatics*, 2003, vol. 19, no. 12, pp. 1572–1574.
7. Swofford, D.L., *PAUP* Phylogenetic Analysis Using Parsimony (and Other Methods): Beta Version 10*, Sunderland: Sinauer, 2002.
8. Felsenstein, J., Evolutionary Trees from DNA Sequences: A Maximum Likelihood Approach, *J. Mol. Evol.*, 1981, vol. 17, no. 6, pp. 368–376.
9. Drummond, A.J. and Rambaut, A., BEAST: Bayesian Evolutionary Analysis by Sampling Trees, *BMC Evol. Biol.*, 2007, vol. 7, no. 1, p. 214.
10. Eiting, T.P. and Smith, G.R., Miocene Salmon (*Oncorhynchus*) from Western North America: Gill Raker Evolution Correlated with Plankton Productivity in the Eastern Pacific, *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 2007, vol. 249, nos. 3–4, pp. 412–424.
11. Wilson, M.V.H. and Li, G.-Q., Osteology and Systematic Position of the Eocene Salmonid †*Eosalmo driftwoodensis* Wilson from Western North America, *Zool. J. Linn. Soc.*, 1999, vol. 125, no. 3, pp. 279–311.
12. Koop, B.F., Schalburg, K.R., Leong, J., et al., A Salmonid EST Genomic Study: Genes, Duplications, Phylogeny and Microarrays, *BMC Genomics*, 2008, vol. 9, p. 545.
13. Oakley, T.H. and Phillips, R.B., Phylogeny of Salmonine Fishes Based on Growth Hormone Introns: Atlan-

- tic (*Salmo*) and Pacific (*Oncorhynchus*) Salmon Are not Sister Taxa, *Mol. Phylogenet. Evol.*, 1999, vol. 11, no. 3, pp. 381–393.
14. Crespi, B.J. and Fulton, M.J., Molecular Systematics of Salmonidae: Combined Nuclear Data Yields a Robust Phylogeny, *Mol. Phylogenet. Evol.*, 2004, vol. 31, no. 2, pp. 658–679.
 15. Matveev, V. and Okada, N., Retroposons of Salmonoid Fishes (Actinopterygii: Salmonoidei) and Their Evolution, *Gene*, 2009, vol. 434, nos. 1–2, pp. 16–28.
 16. Phillips, R.B., Matsuoka, M.P., Konkol, N.R., and McKay, S., Molecular Systematics and Evolution of the Growth Hormone Introns in the Salmoninae, *Environ. Biol. Fishes*, 2004, vol. 69, nos. 1–4, pp. 433–440.
 17. Osinov, A.G. and Lebedev, V.S., Salmonid Fishes (Salmonidae, Salmoniformes): The Systematic Position in the Superorder Protacanthopterygii, the Main Stages of Evolution, and Molecular Dating, *Vopr. Ikhtiol.*, 2004, vol. 44, no. 6, pp. 738–765.
 18. Wilson, W.D. and Turner, T.F., Phylogenetic Analysis of the Pacific Cutthroat Trout (*Oncorhynchus clarki* ssp.: Salmonidae) Based on Partial mtDNA ND4 Sequences: A Closer Look at the Highly Fragmented Inland Species, *Mol. Phylogenet. Evol.*, 2009, vol. 52, no. 2, pp. 406–415.
 19. Li, J., Xia, R., McDowall, R.M., et al., Phylogenetic Position of the Enigmatic *Lepidogalaxias salamandroides* with Comment on the Orders of Lower Euteleostean Fishes, *Mol. Phylogenet. Evol.*, 2010, vol. 57, no. 2, pp. 932–936.
 20. Yasuike, M., Jantzen, S., Cooper, G.A., et al., Grayling (Thymallinae) Phylogeny within Salmonids: Complete Mitochondrial DNA Sequences of *Thymallus arcticus* and *Thymallus thymallus*, *J. Fish. Biol.*, 2010, vol. 76, pp. 395–400.
 21. Shedko, S.V., Phylogeny of Mitochondrial DNA in Salmonids of the Subfamily Salmoninae: Analysis of the Cytochrome b Gene Sequences, *Russ. J. Genet.*, 2002, vol. 38, no. 3, pp. 277–285.
 22. Osinov, A. and Lebedev, V.S., Genetic Divergence and Phylogeny of the Salmoninae Based on Allozyme Data, *J. Fish. Biol.*, 2000, vol. 57, no. 2, pp. 354–381.

УДК 575.86:597.553.2

ФИЛОГЕНИЯ ЛОСОСЕВЫХ РЫБ (Salmoniformes: Salmonidae) И ЕЕ МОЛЕКУЛЯРНАЯ ДАТИРОВКА: АНАЛИЗ ЯДЕРНОГО ГЕНА *RAG1*

© 2012 г. С. В. Шедько, И. Л. Мирошниченко, Г. А. Немкова

Учреждение Российской академии наук Биолого-почвенный институт
Дальневосточного отделения, Владивосток 690022

e-mail: shedko@biosoil.ru

Поступила в редакцию 03.11.2011 г.

Исследованы филогенетические взаимоотношения 26 видов лососевых рыб (сем. Salmonidae) с использованием гена *RAG1* в качестве филогенетического маркера. Взаимоотношения хариусов, сига и лососей однозначно установить не удалось; дивергенция этих линий, по-видимому, произошла в сравнительно короткий (≈ 3 –4 млн. лет) промежуток времени. Первой, вероятно, обособилась линия хариусовых рыб. Роды подсемейства Salmoninae слагают две хорошо выраженные монофилетические группировки – (1) *Brachymystax* и *Hucho*, а также (2) *Salmo*, *Parahucho*, *Salvelinus*, *Parasalmo* и *Oncorhynchus*. Предковые формы этих двух эволюционных линий могли дивергировать на границе олигоцена и миоцена (≈ 24 млн. лет назад). Диверсификация основных линий внутри второй группировки происходила, по всей видимости, сравнительно быстро в среднем миоцене (≈ 19 –16 млн. лет назад). При этом линии *Salvelinus* и *Parasalmo* вместе с *Oncorhynchus* разделились в последнюю очередь. Вероятно, в этот же период времени разошлись предковые линии родов *Prosopium* и *Coregonus*.

Лососевые рыбы (сем. Salmonidae) – модельная группа для изучения ряда актуальных проблем эволюционной биологии, геномики и других разделов современной биологии [1, 2]. Очевидно, что наличие обоснованной филогенетической схемы является необходимым условием для корректного планирования сравнительных исследований и интерпретации их результатов. Несмотря на достигнутый в последнее время прогресс в понимании филогенетических связей внутри Salmonidae, некоторые принципиальные вопросы все еще остаются нерешенными. Во-первых, не ясен порядок дивергенции трех главных линий Salmonidae – хариусовых (подсем. Thymallinae), сиговых (Coregoninae) и лососевых (Salmoninae) рыб. Во-вторых, нет единого мнения относительно филогенетических взаимоотношений основных линий внутри Salmoninae – *Hucho* и *Brachymystax*; *Parahucho*; *Salmo*; *Salvelinus*; *Parasalmo* и *Oncorhynchus*.

Настоящее исследование направлено на решение указанных вопросов при использовании проверенного и хорошо себя зарекомендовавшего [3, 4] филогенетического маркера, нового для данной группы рыб – однокопийного ядерного гена *RAG1* (гена активирующего рекомбинацию 1). В результате получена вполне приемлемая схема филогении Salmonidae, а также, с некоторыми допущениями, проведена ее молекулярная датировка.

В работе использовали 62 образца ДНК, полученных стандартной фенольно-хлороформной экстракцией из мышечных тканей 23 видов лососевых рыб: *Thymallus arcticus* (р. Катунь, $n = 1$), *Th. tugarinae* (р. Киевка, $n = 2$), *Coregonus migratorius* (оз. Байкал, $n = 2$), *Hucho taimen* (р. Селенга, $n = 2$), *Brachymystax lenok* (оз. Маркаколь, $n = 1$; р. Ингода, $n = 1$; р. Арму, $n = 1$), *B. tumensis* (р. Онон, $n = 1$; р. Комиссаровка, $n = 1$; р. Ачаньевка, $n = 1$), *Salmo trutta* (р. Луга, $n = 2$), *Parahucho perryi* (р. Киевка, $n = 7$), *Salvelinus alpinus* (оз. Luomus, Финляндия, $n = 3$; Kanes Laddu, Финляндия, $n = 3$; “пучеглазка” из оз. Лама, $n = 3$; даватчан из оз. Фролиха, $n = 3$; р. Колыма, $n = 2$), *S. leucomaenis* (р. Максимовка, $n = 2$), *S. levanidovi* (р. Яма, $n = 1$), *S. elgyticus* (оз. Эльгыгытгын, $n = 2$), *S. boganidae* (оз. Эльгыгытгын, $n = 2$), *S. taranetzi* (р. Выквынайваам, $n = 1$), *S. malma* (р. Камчатка, $n = 1$; руч. Вилючинский, $n = 1$), *S. curilus* (зал. Северный, $n = 1$; р. Тымь, $n = 1$; о-в Расшуа, $n = 2$; р. Максимовка, $n = 1$), *Oncorhynchus masou* (р. Лютога, $n = 1$; р. Серебрянка, $n = 1$), *O. keta* (р. Нарва, $n = 1$; р. Киевка, $n = 2$), а также *O. gorbuscha* ($n = 1$), *O. nerka* ($n = 2$), *O. kisutch* ($n = 1$), *O. tschawytscha* ($n = 1$) и *Parasalmo mykiss* ($n = 1$) из р. Камчатка.

Фрагмент, включающий большую часть второго (3') экзона гена *RAG1*, амплифицировали методом полимеразной цепной реакции; полученный продукт очищали, секвенировали с помо-

щью набора Big Dye Terminator 3.1 (“Applied Biosystems”, США) и разгоняли продукты реакции на автоматическом анализаторе ABI Prizm 3130 (“Applied Biosystems”, США/“Hitachi”, Япония) на базе Биолого-почвенного института ДВО РАН (г. Владивосток). Для амплификации и секвенирования использовали набор праймеров из работы [3], а также праймеры, разработанные нами: F2817 – GAAGCACAGCCGTCTCATCCTG и F3269 – TGGACAARCAGCTGAGGAAGAAGA.

В результате у каждого из 62 исследованных образцов было секвенировано по 1524 пн фрагмента гена *RAG1*. Для большинства образцов смешанный сигнал в ABI-хроматограммах выявлен не был. Однако у 13 образцов обнаружено наложение пиков по 1–2 (в среднем – 1.6) различным нуклеотидным позициям. Такие особи рассматривались как гетерозиготные по исследованному участку гена *RAG1*. Всего у 62 образцов было выявлено 23 варианта гена *RAG1*, различавшихся недвусмысленными нуклеотидными замещениями. Как правило, каждый из 23 видов Salmonidae характеризовался одним–двумя уникальными вариантами. Исключение составили представители *B. lenok* и *B. tumensis*, неотличимые по исследованному участку гена *RAG1*, а также образцы *S. alpinus*, *S. taranetzi* и одна из двух особей *S. elgyticus*, имевших один и тот же вариант гена *RAG1*. Кроме последовательностей, полученных нами (депонированы в Genbank/NCBI под номерами GQ871466–GQ871494), в анализ были включены данные по гену *RAG1* из базы Genbank/NCBI для следующих видов рыб: *Thymallus thymallus* (AY430214), *Prosopium williamsoni* (AY430213), *Salmo salar* (DW577969, EG768285), а также *Salvelinus leucomaenis* (AY380535) с Курил и *P. mykiss* (NM_001124737) из Северной Америки и видов Esociformes, взятых в качестве внешних групп – *Esox lucius* (AY380542) и *Umbra krameri* (AY380547). Таким образом, для построения филогении лососевых рыб было использовано 30 вариантов последовательностей гена *RAG1*, представлявших 26 видов Salmonidae и 2 внешние группы.

Реконструкцию филогении Salmonidae осуществляли различными методами: 1) байесовским (BA) на основе модели нуклеотидных замещений GTR + G, отобранной программой Modeltest 3.7 [5]; 2) наибольшего правдоподобия (ML) при использовании той же эволюционной модели; 3) максимальной экономии (MP); 4) дистанционным – по алгоритму ближайшего соседства (NJ) на основе LogDet генетических дистанций.

Анализ Монте-Карло с использованием цепей Маркова (байесовский анализ филогении) выполняли с помощью программы MrBayes 3.2 [6] путем одновременного запуска шести цепей (пяти “горячих” и одной “холодной”) в течение $3 \times$

$\times 10^6$ циклов с отбором каждого пятисотого из генерированных деревьев и исключением первых 1001 из 6001 полученных деревьев (burnin = 1001).

Эвристический поиск наиболее правдоподобных или максимально экономных деревьев (ML- и MP-деревьев) проводили с помощью программы RAUP 4.0b10 [7] в 40 повторностях со случайным характером включения последовательностей в анализ и перестановками по TBR-алгоритму. Также программа использовалась и для построения NJ-дерева. Устойчивость порядка ветвлений ML-, MP- и NJ-деревьев оценивалась методом бутстрэпа в 1000 повторных псевдослучайных выборках.

Вне зависимости от использованного метода итоговые филогенетические деревья оказались практически одними и теми же. Топология BA-, ML- и MP-деревьев была идентичной (рис. 1), а NJ-дерево отличалось от них в двух моментах – сестринским положением *Thymallus* и *Coregonus*, а также объединением в одну группировку *Parahucho*, *Salvelinus* и *Oncorhynchus*, но бутстрэп-поддержка этих альтернативных вариантов кластеризации оказалась низкой (<50%). Принципиальное сходство результатов, полученных с помощью различных методов филогенетической реконструкции, а также высокая статистическая поддержка большинства из выявленных клад позволяют рассматривать филогенетическое дерево на рис. 1 как обоснованное.

Важно, что тест молекулярных часов, основанный на отношении оценок правдоподобия [8], показал, что скорость эволюции нуклеотидных последовательностей гена *RAG1* в разных частях этого дерева относительно однородна. Различие между оценками правдоподобия, полученными для аддитивного и ультраметрического ML-деревьев, оказалось статистически не значимо – $2\Delta = 2(-4963.02 - (-4976.32)) = 26.6$, $P = 0.54$ при $d.f. = 28$ (без внешних групп – $2\Delta = 2(-3740.08 - (-3753.01)) = 25.9$, $P = 0.47$ при $d.f. = 26$). Это свойство позволило построить ультраметрическое дерево Salmonidae, что было реализовано в рамках байесовского подхода с помощью пакета программ BEAST 1.4.8 [9] при следующих условиях: модель молекулярных часов – строгие часы; модель нуклеотидных замещений – GTR + G; протяженность байесовского анализа – 3×10^6 циклов с отбором каждого тысячного из генерированных деревьев; burnin – 601; внешние группы из рассмотрения были исключены. Для привязки ультраметрического дерева к абсолютной шкале времени было использовано два калибровочных интервала. Первый был основан на многочисленных ископаемых остатках лососевых рыб, проявляющих продвинутое черты тихоокеанских лососей рода *Oncorhynchus* (см.: [10], а также литературу из этой работы), и ограничивал время разделения рецентных линий *Oncorhynchus* ин-

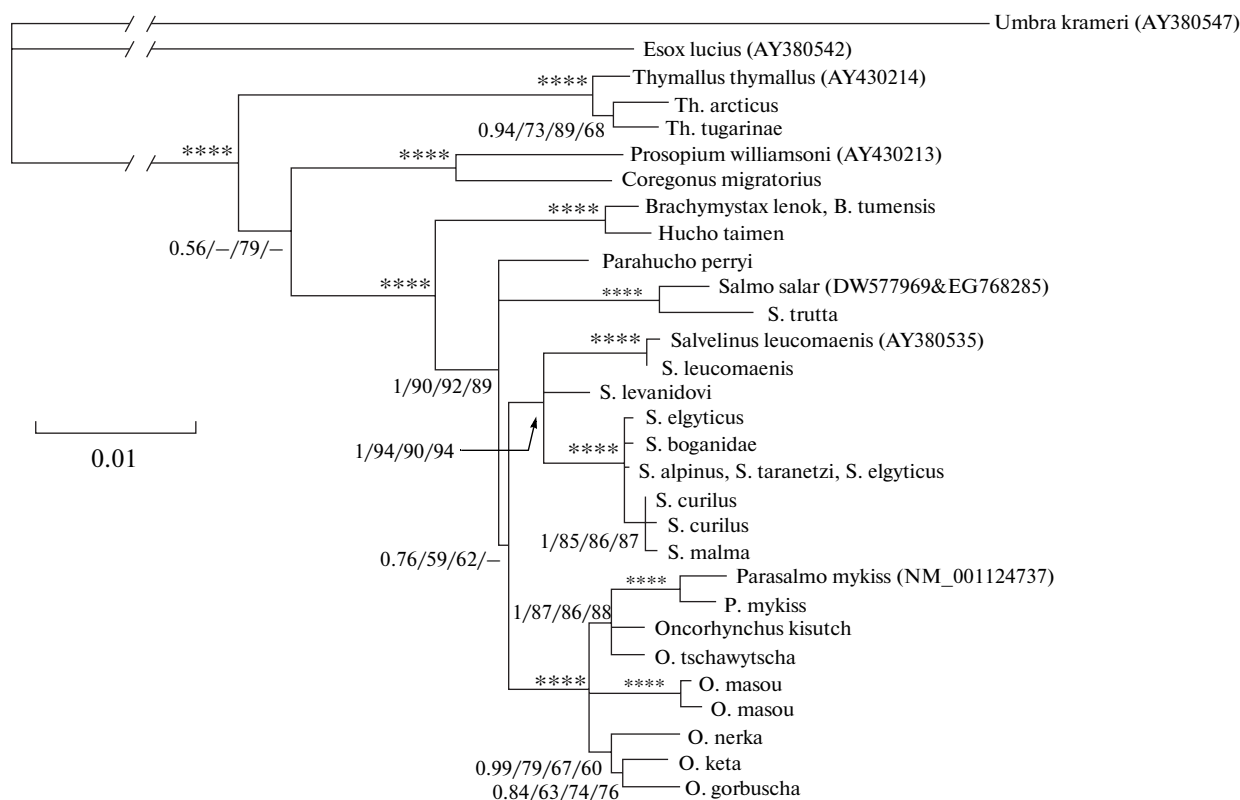


Рис. 1. Филогенетическое древо лососевых рыб (ML-дерево), полученная по результатам анализа фрагмента гена *RAG1*. Цифры в основаниях клад — значения апостериорной вероятности для байесовского дерева, а также бутстрэп-оценки (% от 1000 реплик бутстрэпа) для ML-, MP- и NJ-деревьев, соответственно. Прочерк означает воспроизводимость данной клады в менее чем 50% реплик бутстрэпа, **** — наивысший уровень поддержки (интервал значений для данной клады составлял 0.95–1.0 в случае байесовского дерева и 95–100 для ML-, MP- и NJ-деревьев). Шкала отражает число замен на нуклеотидную позицию.

тервалом позднего миоцена (5.3–11.6 млн. лет назад). Второй интервал ограничивал обособление линии Salmoninae пределами эоцена (33.9–55.8 млн. лет назад) и связан с обнаружением в среднем эоцене хорошо сохранившихся ископаемых лососей *Eosalmo*, имеющих несколько продвинутых характеристик, которые среди современных Salmonidae присущи исключительно представителям Salmoninae [11].

Полученное при указанных условиях ультраметрическое байесовское дерево (рис. 2) практически полностью повторило топологию дерева на рис. 1 и имело сходные с ординарным байесовским деревом оценки апостериорных вероятностей. Таким образом, суммируя полученные результаты (рис. 1, 2), можно прийти к следующим основным выводам:

1. Взаимоотношения линий хариусов, сигов и лососей однозначно установить не удалось. Их дивергенция, по-видимому, произошла в сравнительно короткий (≈ 3 –4 млн. лет) промежуток времени. Вероятно, что первой обособилась линия хариусовых рыб.

2. Роды подсемейства Salmoninae слагают две хорошо выраженные монофилетические группировки — (1) *Brachymystax* и *Hucho*, а также (2) *Salmo*, *Parahucho*, *Salvelinus*, *Parasalmo* и *Oncorhynchus*. Предковые формы этих двух эволюционных линий могли дивергировать на границе олигоцена и миоцена (≈ 24 млн. лет назад).

3. Диверсификация основных линий внутри второй группировки происходила, по всей видимости, сравнительно быстро в среднем миоцене (≈ 19 –16 млн. лет назад). При этом линии *Salvelinus* и *Parasalmo* вместе с *Oncorhynchus* разделились в последнюю очередь. Вероятно, в этот же период времени разошлись предковые линии родов *Prosopium* и *Coregonus*.

В целом секвенированный фрагмент гена *RAG1* показал свою пригодность для решения задач настоящего исследования. Данный маркер, имеющий, как и многие другие ядерные гены, низкую скорость эволюции (по нашим расчетам — 0.57×10^{-9} замен на сайт в год), проявил ясный, с минимальным количеством гомоплазии, филогенетический сигнал (индекс соответствия для MP-дерева без внешних групп — 0.79). Полученные

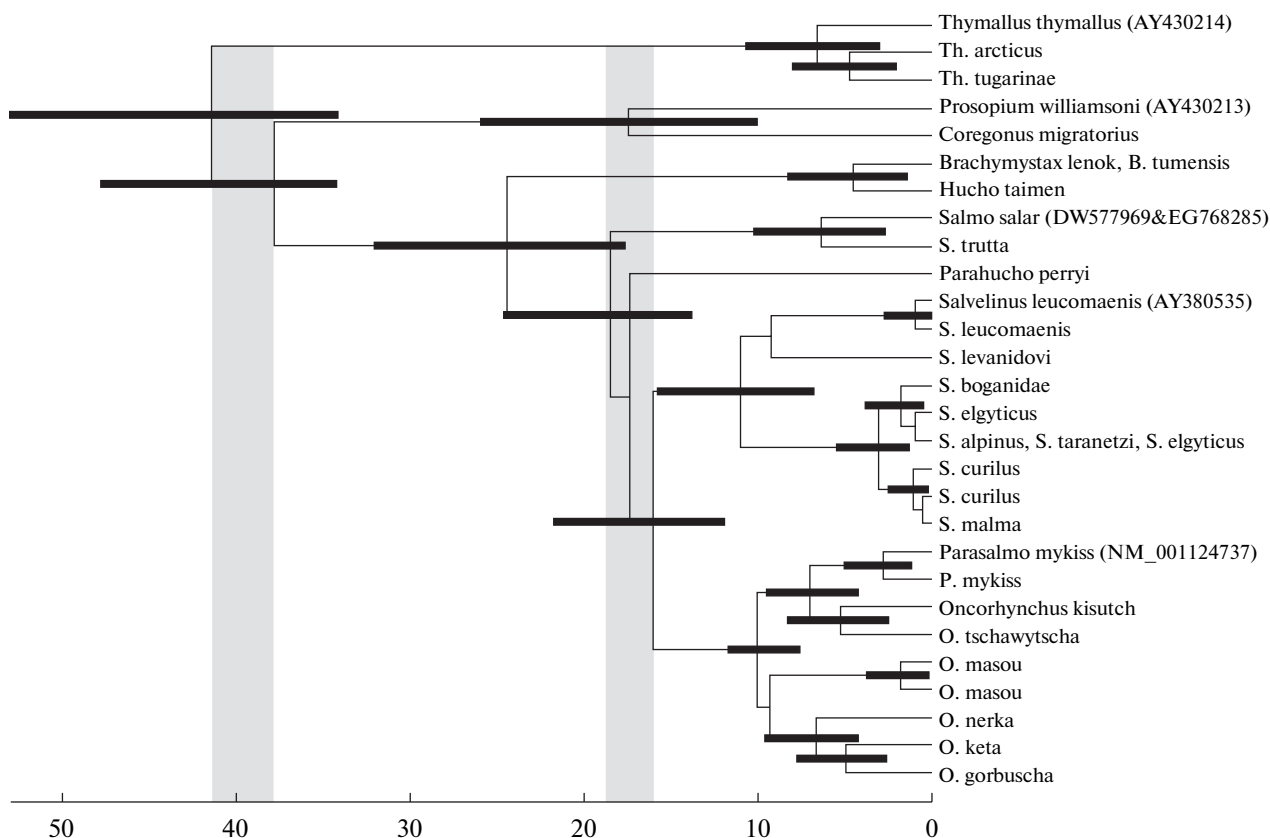


Рис. 2. Байесовская хронограмма кладогенеза лососевых рыб, выраженная в абсолютной временной шкале (млн. лет). Прямоугольники – 95%-ные интервалы наивысшей апостериорной плотности оценок времени существования клад, имевших апостериорные вероятности выше 0.5. Затемнены периоды, давшие основное морфо-экологическое разнообразие Salmonidae.

филогенетические деревья (рис. 1–2) находятся в соответствии с результатами недавних молекулярно-генетических исследований, основанных на анализе ядерной ДНК, в части взаимоотношений линий хариусов, сига и лососей [12], порядка дивергенции основных эволюционных линий внутри Salmoninae [13–15], филогенетических взаимоотношений *O. keta*, *O. gorbuscha* и *O. nerka* [14, 16] и других деталей. Результаты молекулярной датировки кладогенеза Salmonidae (рис. 2) также во многом сходны с оценками, полученными ранее, исходя из других данных и при использовании иных методических подходов [17, 18].

Однако филогенетические связи хариусов, сига и лососей по результатам других молекулярно-генетических работ выглядят иначе [15, 19, 20]. Положение *Brachymystax* и *Hucho* среди остальных эволюционных линий Salmoninae в работах, основанных на анализе мтДНК [14, 21] или интронов ядерного гена GN1 [16], также выглядит по-другому. Кроме того, от работы к работе варьирует положение линии *Parasalmo* относительно видов *Oncorhynchus* (рис. 1, 2, [13, 14, 16, 18, 20–22]). В связи с этим дальнейшее уточнение и детализация схемы филогении Salmonidae видятся в уве-

личении числа используемых филогенетических маркеров и одновременном анализе многих не сцепленных локусов ядерного генома, а также мтДНК.

Авторы глубоко признательны О.А. Радченко, Д.С. Сендек и М.Б. Шедько за предоставленные образцы ряда видов лососевых рыб.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (проект № 09-04-00370-а) и ДВО РАН (проект № 09-III-A-06-185).

СПИСОК ЛИТЕРАТУРЫ

1. Evolution illuminated. Salmon and their relatives / Hendry A.P., Stearns S.C. N.Y.: Oxford Univ. Press, 2004. 510 p.
2. Davidson W.S., Koop B.F., Jones S.J.M. et al. Sequencing the genome of the Atlantic salmon (*Salmo salar*) // Genome Biology. 2010. V. 11. № 9. P. 403.
3. Lo'pez J.A., Chen W.-J., Ort G. Esociform phylogeny // Copeia. 2004. № 3. P. 449–464.
4. Chiari Y., Meijden A. van der, Madsen O. et al. Base composition, selection, and phylogenetic significance

- of indels in the recombination activating gene-1 in vertebrates // *Front. Zoology*. 2009. V. 6. P. 32.
5. *Posada D., Crandall K.A.* MODELTEST: testing the model of DNA substitution // *Bioinformatics*. 1998. V. 14. № 9. P. 817–818.
 6. *Ronquist F., Huelsenbeck J.P.* MRBAYES 3: Bayesian phylogenetic inference under mixed models // *Bioinformatics*. 2003. V. 19. № 12. P. 1572–1574.
 7. *Swofford D.L.* PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. MA: Sinauer Associates, 2002.
 8. *Felsenstein J.* Evolutionary trees from DNA sequences: a maximum likelihood approach // *J. Mol. Evol.* 1981. V. 17. № 6. P. 368–376.
 9. *Drummond A.J., Rambaut A.* BEAST: Bayesian evolutionary analysis by sampling trees // *BMC Evol. Biol.* 2007. V. 7. № 1. P. 214.
 10. *Eiting T.P., Smith G.R.* Miocene salmon (*Oncorhynchus*) from Western North America: gill raker evolution correlated with plankton productivity in the Eastern Pacific // *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 2007. V. 249. № 3–4. P. 412–424.
 11. *Wilson M.V.H., Li G.-Q.* Osteology and systematic position of the Eocene salmonid †*Eosalmo driftwoodensis* Wilson from western North America // *Zool. J. Linn. Soc.* 1999. V. 125. № 3. P. 279–311.
 12. *Koop B.F., Schalburg K.R. von, Leong J. et al.* A salmonid EST genomic study: genes, duplications, phylogeny and microarrays // *BMC genomics*. 2008. V. 9. P. 545.
 13. *Oakley T.H., Phillips R.B.* Phylogeny of salmonine fishes based on growth hormone introns: Atlantic (*Salmo*) and Pacific (*Oncorhynchus*) salmon are not sister taxa // *Mol. Phylog. Evol.* 1999. V. 11. № 3. P. 381–393.
 14. *Crespi B.J., Fulton M.J.* Molecular systematics of Salmonidae: combined nuclear data yields a robust phylogeny // *Mol. Phylog. Evol.* 2004. V. 31. № 2. P. 658–679.
 15. *Matveev V., Okada N.* Retroposons of salmonoid fishes (Actinopterygii: Salmonoidei) and their evolution // *Gene*. 2009. V. 434. № 1–2. P. 16–28.
 16. *Phillips R.B., Matsuoka M.P., Konkol N.R., McKay S.* Molecular systematics and evolution of the Growth Hormone introns in the Salmoninae // *Env. Biol. Fishes*. 2004. V. 69. № 1–4. P. 433–440.
 17. *Осинов А.Г., Лебедев В.С.* Лососевые рыбы (Salmonidae, Salmoniformes): положение в надотряде Protacanthopterygii, основные этапы эволюционной истории, молекулярные датировки // *Вопр. ихтиологии*. 2004. Т. 44. № 6. С. 738–765.
 18. *Wilson W.D., Turner T.F.* Phylogenetic analysis of the Pacific cutthroat trout (*Oncorhynchus clarki* ssp.: Salmonidae) based on partial mtDNA ND4 sequences: a closer look at the highly fragmented inland species // *Mol. Phylog. Evol.* 2009. V. 52. № 2. P. 406–415.
 19. *Li J., Xia R., McDowall R.M. et al.* Phylogenetic position of the enigmatic *Lepidogalaxias salamandroides* with comment on the orders of lower euteleostean fishes // *Mol. Phylog. Evol.* 2010. V. 57. № 2. P. 932–936.
 20. *Yasuike M., Jantzen S., Cooper G.A. et al.* Grayling (Thymallinae) phylogeny within salmonids: complete mitochondrial DNA sequences of *Thymallus arcticus* and *Thymallus thymallus* // *J. Fish Biol.* 2010. V. 76. P. 395–400.
 21. *Шедько С.В.* Филогения митохондриальной ДНК лососевых рыб подсемейства Salmoninae: анализ последовательностей гена цитохрома b // *Генетика*. 2002. Т. 38. № 3. С. 357–367.
 22. *Osinov A., Lebedev V.S.* Genetic divergence and phylogeny of the Salmoninae based on allozyme data // *J. Fish Biol.* 2000. V. 57. № 2. P. 354–381.

ФИЛОГЕНИЯ ЛОСОСЕВЫХ РЫБ (Salmoniformes: Salmonidae) И ЕЕ МОЛЕКУЛЯРНАЯ ДАТИРОВКА: АНАЛИЗ ЯДЕРНОГО ГЕНА *RAG1*

© 2012 г. С. В. Шедько, И. Л. Мирошниченко, Г. А. Немкова

*Учреждение Российской академии наук Биолого-почвенный институт
Дальневосточного отделения, Владивосток 690022*

e-mail: shedko@biosoil.ru

Поступила в редакцию 03.11.2011 г.

Исследованы филогенетические взаимоотношения 26 видов лососевых рыб (сем. Salmonidae) с использованием гена *RAG1* в качестве филогенетического маркера. Взаимоотношения хариусов, сига и лососей однозначно установить не удалось; дивергенция этих линий, по-видимому, произошла в сравнительно короткий (~3–4 млн. лет) промежуток времени. Первой, вероятно, обособилась линия хариусовых рыб. Роды подсемейства Salmoninae слагают две хорошо выраженные монофилетические группировки – (1) *Brachymystax* и *Hucho*, а также (2) *Salmo*, *Parahucho*, *Salvelinus*, *Parasalmo* и *Oncorhynchus*. Предковые формы этих двух эволюционных линий могли дивергировать на границе олигоцена и миоцена (~24 млн. лет назад). Диверсификация основных линий внутри второй группировки происходила, по всей видимости, сравнительно быстро в среднем миоцене (~19–16 млн. лет назад). При этом линии *Salvelinus* и *Parasalmo* вместе с *Oncorhynchus* разделились в последнюю очередь. Вероятно, в этот же период времени разошлись предковые линии родов *Prosopium* и *Coregonus*.