

Phylogeny of Salmonids (Salmoniformes: Salmonidae) and its Molecular Dating: Analysis of mtDNA Data

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

Institute of Biology and Soil Science, Far Eastern Branch, Russian Academy of Sciences, Vladivostok, 690022 Russia
e-mail: shedko@biosoil.ru

Received August 7, 2012; in final form November 6, 2012

Abstract—Phylogenetic relationships among 41 species of salmonid fish and some aspects of their diversification-time history were studied using the GenBank and original mtDNA data. The position of the root of the Salmonidae phylogenetic tree was uncertain. Among the possible variants, the most reasonable seems to be that in which thymallins are grouped into the same clade as coregonins and the lineage of salmonins occupied a basal position relative to this clade. The genera of Salmoninae formed two distinct clades, i.e., (*Brachymystax*, *Hucho*) and (*Salmo*, *Parahucho*, (*Salvelinus*, (*Parasalmo*, *Oncorhynchus*))). Furthermore, the genera *Parasalmo* and *Oncorhynchus* were reciprocally monophyletic. The congruence of Salmonidae phylogenetic trees obtained using different types of phylogenetic markers is discussed. According to Bayesian dating, ancestral lineages of salmonids and their sister esocoids diverged about 106 million years ago. Sometime after, probably 100–70 million years ago, the salmonid-specific whole genome duplication took place. The divergence of salmonid lineages on the genus level occurred much later, within the time interval of 42–20 million years ago. The main wave of the diversification of salmonids at the species level occurred during the last 12 million years. The possible effect of genome duplication on the Salmonidae diversification pattern is discussed.

DOI: 10.1134/S1022795413060112

INTRODUCTION

Salmonid fish (family Salmonidae) are a remarkable model for use in evolutionary biology, genomics, and other fields of modern biology, largely because the lineage ancestor experienced an additional round of whole genome duplication [1–4]. The total number of publications on salmonid fish exceeds several tens of thousands [4], and Salmonidae are rightfully considered one of the most comprehensively studied group of lower vertebrates. However, in many areas, there are still blank spots and some basic issues still remain unresolved, including the phylogeny.

First, the divergence order of three main lineages of Salmonidae, thymallins (subfamily Thymallinae), coregonins (Coregoninae), and salmonins (Salmoninae) is unclear [5–9].

Second, the existing schemes of phylogenetic relationships among taxa within the above-mentioned subfamilies [8–15] require improvement and/or additional testing using different types of phylogenetic markers.

In addition, Salmonidae still require a more or less complete time-calibrated portrait of their diversification, which includes most of the contemporary representatives of the subfamilies.

Mitochondrial DNA (mtDNA) is a well-proven phylogenetic marker tested at different levels of taxonomic hierarchy in many different groups of organ-

isms with clear evolutionary dynamics, and is free from many of the problems associated with the phenomenon of paralogy [16]. The mtDNA data, starting from the pioneer study of Berg and Ferris [17], were often used to solve various problems of the phylogeny of salmonids. However, these studies usually covered only a few representatives of Salmonidae.

For these reasons, the main objective of this study was to elucidate the phylogenetic relationships among 41 species of salmonid fish, which accounts for about two-thirds of the existing diversity of Salmonidae [18]. In addition, the molecular dating of cladogenetic events in this group was performed based on the analysis of the mitogenome sequence data. Moreover, in terms of current discussion on the role of polyploidization in the formation of diversity in bony fish (Teleostei) [19–21], it seemed interesting to assess how close in time the genome duplication event and the main periods of their diversification were in salmonid fish.

MATERIALS AND METHODS

The analysis included mtDNA sequences retrieved from the GenBank/NCBI database, as well as the sequences of the *COI*, *Cytb*, and the D loop region of salmonid fish amplified in the present study using primers reported in [22–24]. Amplification products were sequenced using the Big Dye Terminator 3.1 kit

(Applied Biosystems, United States) and the ABI Prism 3130 automated analyzer (Applied Biosystems, United States/Hitachi, Japan), especially for this case (Table 1: JX261982 to JX262010). In total, 41 species of salmonid fish were examined and 11 species represented three outgroups, including Esociformes (5), Argentiformes (3), and Stomiiformes (3). For 24 species, the data on whole mitogenomes (16 species of Salmonidae and eight species from outgroups) were used. For another 28 species, only incomplete data were available, of which eight mitogenome fragments, which contain either parts or full sequences of the *AT6*, *COI*, *Cytb*, *ND1*, *ND6*, *12S* and *16S* rRNA genes, and the D-loop region were selected for further analysis (Table 1). A rational basis for including taxa with incomplete data into the analysis lies first in the need to minimize the number of long branches in the phylogenetic tree (fourth strategy of taxa selection [26]), which yields a positive effect, even in case of data gaps [27, 28]. Second, the goal of the study was to encompass as much of the Salmonidae morphological and ecological diversity as possible, including the taxa, which were either previously treated as independent genera or still have a claim on this position (*Acantholingua*, *Baione*, *Cristivomer*, *Phylogeophya*, *Salmothy-mus*, and some others; third strategy [26]).

Sequence data for each mtDNA functional region were aligned using the MAFFT v. 6 software program [29]. Then, individual alignments were concatenated. The matrix was reviewed and the alignment of ambiguous regions, as well as stop codons in the protein-coding genes were excluded from the analysis.

Phylogenetic reconstructions were performed using different approaches, including weighted maximum parsimony (MP), maximum likelihood (ML), and Bayesian analysis (BA). Moreover, in each case two variants of phylogenetic reconstructions were performed, i.e., using the data for all types of substitutions, and excluding the nucleotide substitutions of transition type.

A search for an MP tree taking into consideration all types of nucleotide substitutions was performed with the partition of the matrix into six presumably structurally homogeneous blocks of nucleotides, including (1) the first, (2) second, and (3) third codon positions of 13 protein-coding genes; the genes of (4) ribosomal and (5) transport RNA, and the (6) D-loop. For each block, the ratio of the rates of transition to transversion was determined using the PAUP 4.0b10 software program [30]. These estimates, rounded to the nearest whole number, were used as transversion weights in step matrices upon construction of MP trees.

The heuristic search for an MP tree was conducted in the PAUP software program in 30 rounds of random sequence addition and subsequent TBR swapping. The robustness of the tree branching pattern was evaluated using 1000 bootstrap pseudoreplicates. The same approach was used to construct the ML tree. The

tree robustness was assessed based on the data from which the information on nucleotide substitutions of transition type was excluded using RY coding.

The construction of an ML tree based on all types of nucleotide substitutions was performed using the data matrix, in which seven blocks of nucleotides were identified, including (1) the first, (2) second, and (3) third codon positions of 12 protein-coding genes; (4) the *ND6* gene, which differs substantially from the other protein-coding genes in its base composition; and the genes of (5) ribosomal and (6) transport RNA and the (7) D-loop. Using the Modeltest 3.7 software program [31] and based on the BIC criterion for each nucleotide block, the best-fit model of nucleotide substitutions was determined as follows: GTR+I+G, TVM+I+G, TrN+I+G, HKY+I+G, GTR+I+G, SYM+G, and HKY+G, respectively.

In the variant of ML analysis based on the RY-coded data matrix, the CF+I+G model of nucleotide substitutions was used.

A heuristic search of the ML tree was conducted in the Garli v. 2 software program [32] in 30 rounds. The robustness of the branching pattern of the tree was evaluated using 1000 bootstrap pseudosamples.

Bayesian phylogenetic inference was performed using the MrBayes 3.2 software program [33]. The partitioned analysis was performed using the above-mentioned ML scheme, except that the TVM+I+G and TrN+I+G models were replaced by the more general GTR+I+G model (due to limitations of the MrBayes program). Transversion analysis was also conducted using the CF+I+G model. In both variants, the analysis included 6×10^6 cycles with sampling each 1000th tree generated. The first 1001 trees out all generated were discarded and the remaining 5000 trees, which were characterized by the stabilized probability estimates (LnL), parameters of nucleotide substitution models, and the tree lengths, were used to construct consensus tree and to obtain the posterior probability estimates of its branching pattern.

Molecular dating of Salmonidae cladogenesis was performed within the framework of Bayesian inference using the BEAST 1.6.2 software program [34] with conditions including a fixed tree topology (ML tree); a relaxed (nonstrict) molecular clock model with uncorrelated log normal distribution of nucleotide substitution rates across the tree branches and speciation events following the Yule model of speciation; a GTR+I+G model of nucleotide substitutions; a Bayesian analysis run for 30×10^6 cycles with sampling of every 2000th trees generated; burnin, 1501. The ultrametric tree was calibrated to absolute time using four calibration intervals. The first interval allowed the split of salmonids and esocoids in the Cretaceous (145.5 to 65.5 million years ago) and was based on numerous Late Cretaceous fossil records of Esociformes [35]. The second interval associated the split of *Esox* and *Novumbra* with the Paleocene (65.5 to 55.8 million years ago) and was based on a number of well-preserved fish fos-

Table 1. List of fish species examined with GenBank accession numbers of fragments of their mtDNA sequences (prefix NC marks mitogenomes with full nucleotide sequences)

Specie	AT6	COI	Cytb	NDI	ND6	12S	16S	D-loop
Stomiiformes								
<i>Diplphos taenia</i>	NC_002647	NC_002647	NC_002647	NC_002647	NC_002647	NC_002647	NC_002647	
<i>Gonostoma gracile</i>	NC_002574	NC_002574	NC_002574	NC_002574	NC_002574	NC_002574	NC_002574	
<i>Chauliodus sloani</i>	NC_003159	NC_003159	NC_003159	NC_003159	NC_003159	NC_003159	NC_003159	
Argentiformes								
<i>Bathylagus ochotensis</i>	NC_004591	NC_004591	NC_004591	NC_004591	NC_004591	NC_004591	NC_004591	
<i>Opisthoproctus soleatus</i>	NC_004600	NC_004600	NC_004600	NC_004600	NC_004600	NC_004600	NC_004600	
<i>Glossanodon semifasciatus</i>	NC_004595	NC_004595	NC_004595	NC_004595	NC_004595	NC_004595	NC_004595	
Esociformes								
<i>Esox lucius</i>	NC_004593	NC_004593	NC_004593	NC_004593	NC_004593	NC_004593	NC_004593	NC_004593
<i>Dallia pectoralis</i>	NC_004592	NC_004592	NC_004592	NC_004592	NC_004592	NC_004592	NC_004592	NC_004592
<i>Umbra limi</i>	FJ918928		AY497458			AY430268	AF060443	
<i>U. krameri</i>	JX261985		JX261998			AY430269	AF060444	
<i>Novumbra hubbsi</i>			AY497457			AY430272	AF060447	
Salmoniformes								
Salmonidae								
Coregoninae								
<i>Coregonus lavaretus</i>	NC_002646	NC_002646	NC_002646	NC_002646	NC_002646	NC_002646	NC_002646	NC_002646
<i>C. peled</i>			DQ185402	DQ399871			DQ399871	AF195647
<i>Stenodus leucichthys</i>	AY778968	EU202658	EU003527					HM448022
<i>Prosopium williamsoni</i>	AY778970	FJ918926						
<i>P. cylindraceum</i>	HQ616435	EU202657	EU003531					
<i>P. coulterii</i>	HQ616436	EU202654	AY382564			AY430254	AY430236	AY382565
Thymallinae								
<i>Thymallus thymallus</i>	NC_012928	NC_012928	NC_012928	NC_012928	NC_012928	NC_012928	NC_012928	NC_012928
<i>Th. arcticus</i>	NC_012929	NC_012929	NC_012929	NC_012929	NC_012929	NC_012929	NC_012929	NC_012929
<i>Th. burejensis</i>	AY778976							AY779007
<i>Th. brevitrostris</i>	AY778990							AY168387
<i>Th. flavomaculatus</i>	AY778983	JX261987	JX262000					AY246410
<i>Th. grubii</i>	AY778972	JX261988	JX262001					AY168390
<i>Th. tugariniae</i>	AY778998	JX261986	JX261999					AY246389

Table 1. (Contd.)

Species	AT6	COI	Cyrb	NDI	ND6	I2S	I6S	D-loop
Salmoninae								
<i>Brachymystax tumensis</i> (Korea)			AF125213		AF125052	AF125513	AF125513	AF125519
<i>B. tumensis</i> (Amur River)	AY862340	JX261989	JX262002	EU760486				EU760491
<i>B. lenok</i> (Amur River)	AY862339	JX261990	JX262003	EU760488				EU760490
<i>Hucho taimen</i>	AY862334	JX261991	JX262004	EU760487				EU760489
<i>H. hucho</i>	AY862338		AF172397	EU729362				EU729360
<i>H. bleekeri</i>	NC_015995	NC_015995	NC_015995	NC_015995	NC_015995	NC_015995	NC_015995	NC_015995
<i>Parahucho perryi</i>	AY862342	JX261992	JX262005	AY862386				AY862364
<i>Salmo salar</i>	NC_001960	NC_001960	NC_001960	NC_001960	NC_001960	NC_001960	NC_001960	NC_001960
<i>S. trutta</i>	NC_010007	NC_010007	NC_010007	NC_010007	NC_010007	NC_010007	NC_010007	NC_010007
<i>S. (Salmothymus) ohridanus</i>	EU282329		AF202033					AY926564
<i>S. (Salmothymus) obtusirostris</i>	EF469834		AF488534					EF469832
<i>Salvelinus fontinalis</i>	NC_000860	NC_000860	NC_000860	NC_000860	NC_000860	NC_000860	NC_000860	NC_000860
<i>S. alpinus</i>	NC_000861	NC_000861	NC_000861	NC_000861	NC_000861	NC_000861	NC_000861	NC_000861
<i>S. leucomaenis</i>		JX261993	JX262006				AF060445	DQ403177
<i>S. namaycush</i>		EU522425	DQ451386					AF297989
<i>S. taranetzi</i>		JX261995	JX262008					JX261983
<i>S. curilus</i>		JX261997	JX262010					DQ403180
<i>S. malma</i>		JX261996	JX262009					JX261984
<i>S. levanidovi</i>		JX261994	JX262007					JX261982
<i>S. confluentus</i>	EU524365		Table 1 in [25]					
<i>S. (Salvethymus) svetovidovi</i>								AF297990
<i>Parasalmo clarkii</i>	NC_006897	NC_006897	NC_006897	NC_006897	NC_006897	NC_006897	NC_006897	NC_006897
<i>P. mykiss</i>	NC_001717	NC_001717	NC_001717	NC_001717	NC_001717	NC_001717	NC_001717	NC_001717
<i>Oncorhynchus tshawytscha</i>	NC_002980	NC_002980	NC_002980	NC_002980	NC_002980	NC_002980	NC_002980	NC_002980
<i>O. kisutch</i>	NC_009263	NC_009263	NC_009263	NC_009263	NC_009263	NC_009263	NC_009263	NC_009263
<i>O. nerka</i>	NC_008615	NC_008615	NC_008615	NC_008615	NC_008615	NC_008615	NC_008615	NC_008615
<i>O. gorbusha</i>	NC_010959	NC_010959	NC_010959	NC_010959	NC_010959	NC_010959	NC_010959	NC_010959
<i>O. keta</i>	NC_017838	NC_017838	NC_017838	NC_017838	NC_017838	NC_017838	NC_017838	NC_017838
<i>O. masou</i>	NC_008747	NC_008747	NC_008747	NC_008747	NC_008747	NC_008747	NC_008747	NC_008747

Note: Numbers in bold type are sequences determined in the present study.

sils, which were quite validly attributed to the genus *Esox* [36, 37]. The third interval limited the isolation of the Salmoninae lineage to the Eocene (55.8–33.9 million years ago). This calibration interval was associated with the discovery in the Middle Eocene of well-preserved fossil salmonids *Eosalmo*, which possessed some derived features that, among present-day Salmonidae, are exclusively typical of Salmoninae [38]. The fourth calibration interval was based on many fossil records of salmonids, which express the derived features of Pacific salmon of the genus *Oncorhynchus* (see [39] and the references from this article) and limited the start of diversification of recent lineages of *Oncorhynchus* to Late Miocene (11.6–5.3 million years ago).

RESULTS

After the removal of the alignment ambiguous regions and stop codons of the protein-coding genes from concatenated matrix, its length was 53×16388 bp. The resulting data matrix included sequences of 18 full mtDNA genomes, 6 mitogenomes without control regions (in the genomes of the representatives of Stomiiformes and Argentiformes, these regions were excluded from the analysis due to inability to perform adequate alignment with the same regions of other sequences), and 29 concatenated nucleotide sequences. The total length of the latter sequences varied from 543 to 5064 bp (with an average length of 2660 bp).

As a result of the phylogenetic analysis of the partitioned data matrix performed using weighted MP, the only MP tree was found (Fig. 1). Most of the internal branches (37 out of 49) were supported with high bootstrap values, which were either equal to or greater than 70%. The topology of the MP tree obtained using transversion MP was almost the same as that of the tree in Fig. 1. However, in this case, the overall level of bootstrap support was somewhat lower; only 30 out of 49 internal branches were supported with the values equal to or greater than 70%.

The cladogram obtained as a result of partitioned ML analysis is demonstrated in Fig. 2; its topology has only one substantial difference from the MP tree topology. Specifically, in the ML tree, coregonins formed one clade with thymallins, while in the MP tree, they grouped together with salmonins. The ML tree topology inferred from the analysis of the RY-coded data was almost the same as that of the tree in Fig. 2. In total, 38 out of 50 internal branches of the tree in Fig. 2 were highly bootstrap supported (70% or higher) in a given variant of ML analysis (in both cases, 31 internal branches).

The cladograms obtained as a result of Bayesian analysis considering either all types of nucleotide substitutions or ignoring the transitions did not differ greatly from one another, and their topology (Fig. 3) was almost the same as that of ML cladogram. With respect to the total number of branches with high

bootstrap support values (38 out of 49 internal branches had the probability of 0.95 and higher) and their distribution across the tree, BA and ML cladograms were very similar.

As can be seen, Figs. 1–3, which were constructed using three different methods, were generally congruent. The exclusion was the remarkably different collocation of the lineages of thymallins, coregonins, and salmonins in MP (Fig. 1) and the cladogram constructed by means of either maximum likelihood or Bayesian methods (Figs. 2, 3). To assess whether this difference is statistically significant, we compared three alternative variants of the phylogenetic relationships of thymallins, coregonins, and salmonins using a number of statistical tests based on an analysis of the probability estimates and implemented in the CONSEL software package [40]. Most of these tests showed that the difference in the most preferable variant of the Salmonidae phylogeny (Figs. 2, 3) from the alternative variants (including the variant demonstrated in Fig. 1) was not statistically significant.

Molecular dating was performed based on the fixed topology of phylogenetic tree constructed using the maximum likelihood approach (Fig. 2). The resulting chronogram of the cladogenesis events of salmonid fish and representatives of the closest outgroup (*Esociformes*) is demonstrated in Fig. 4.

According to our estimates, for the given set of Salmonidae taxa, the mean evolutionary species age (terminal branch length of the tree in Fig. 4) was 5.6 million years (range was 0.9–21.4; median was 4.8).

Among the species of Salmonidae, a number of relatively compact groups that unite the representatives of one genus (*Thymallus*, *Prospium*), or more frequently, the representatives of closely related genera (*Stenodus* and *Coregonus*; *Hucho* and *Brachymystax*; *Salmo* and *Salmothymus*; *Parasalmo* and *Oncorhynchus*; *Salvelinus* and *Salvethymus*) can be identified. In these groups, the evolutionary age of the most recent common ancestor varied in the range of 6.1–17.0 million years with an average of 10.4 million years (median was 10.3).

Correspondingly, the ancestral lineages of these groups of species, as well as the lineage of the monotypic species *Parahucho* diverged from each other much earlier, about 42.3–19.9 million years ago. Moreover, five out of seven of these divergence events occurred 27.8–19.9 (23.4, on average) million years ago.

Dating of the Salmonidae lineage-specific whole genome duplication was carried out using the data for nuclear genome regions of the four salmonid species (*Coregonus clupeaformis*, *Salmo salar*, *Salvelinus alpinus*, and *Parasalmo mykiss*), flanked by conservative noncoding elements (CNE) (Table 1 in [41]). The divergence time for nine pairs of these paralogous regions was determined using estimates of the divergence time for the coregonins–salmonins and (42 million years) and the *Salmo*–*Salvelinus/Parasalmo*

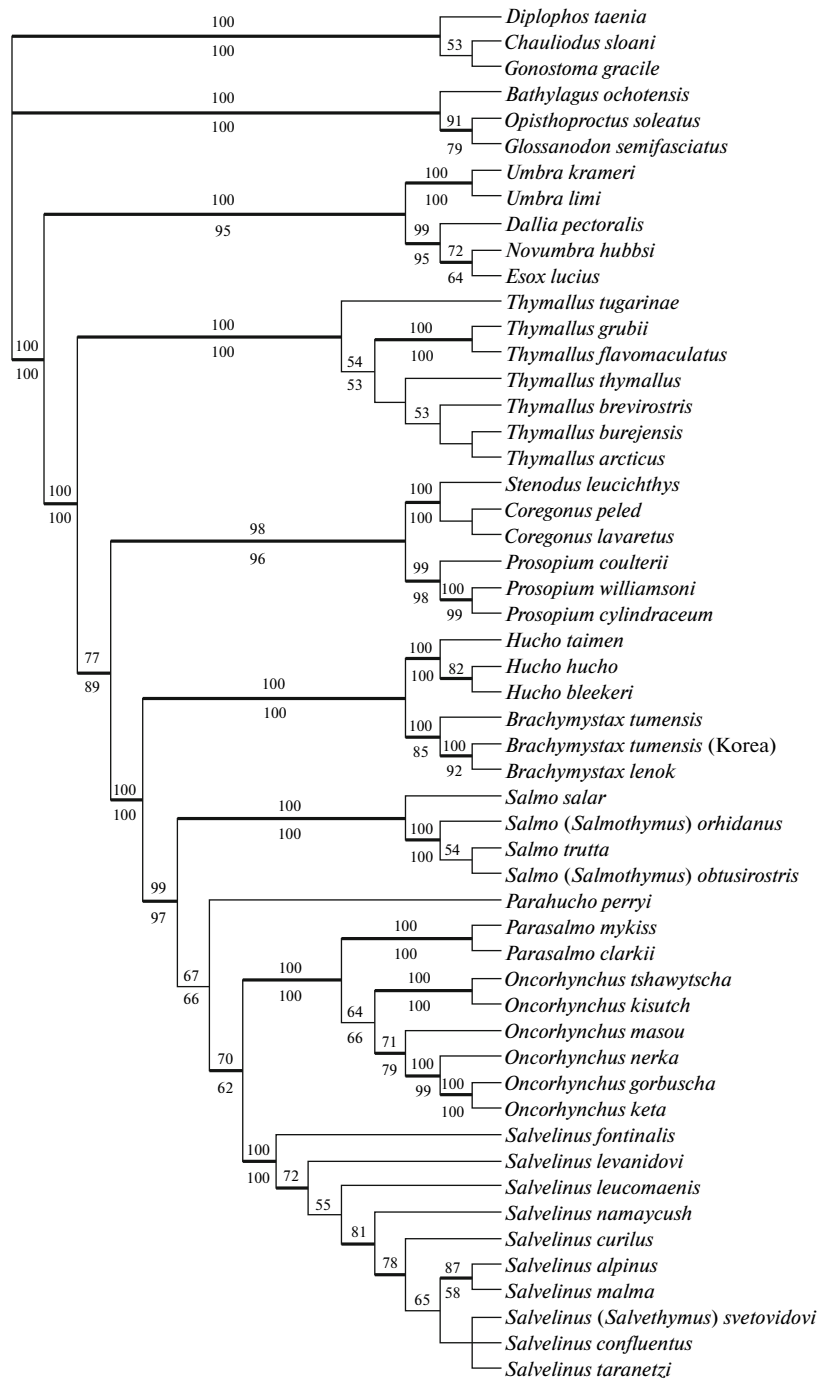


Fig. 1. MP cladogram (76011 steps) obtained as a result of analysis of mtDNA data matrix using method of weighted MP. Figures above and below the branches are bootstrap support values (% of 1000 bootstrap replicates) obtained for MP trees generated using data on all types of substitutions (6713 phylogenetically informative positions) or excluding the substitutions of transition type (3391 phylogenetically informative positions), respectively. Branches with support values of 70% and higher are marked.

(23 million years) dichotomies obtained in the present study were used as calibration points. Similar to the previous estimation, molecular dating was performed using the BEAST software package [34] and additional conditions that consist of the HKY+G model of nucleotide substitutions; the strict molecular clock

model; Bayesian analysis run for 2×10^6 cycles with sampling of each 1000th of the trees generated; and burnin, 501.

As a result, it was demonstrated (Fig. 4) that medians of the Bayesian time estimates from the date of paralog divergence (the date of disomy establishment)

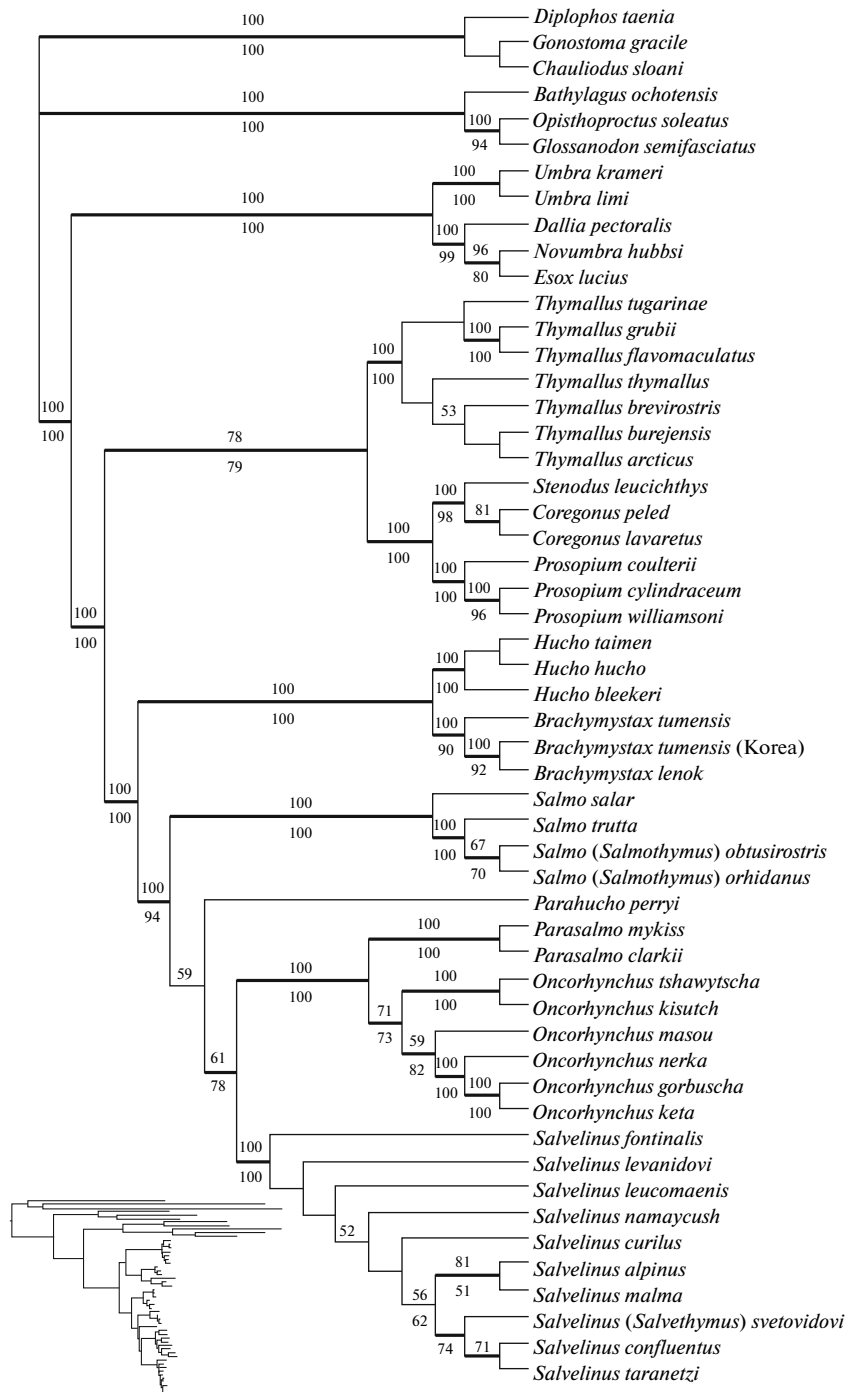


Fig. 2. ML cladogram generated based on partitioned analysis of mtDNA data. Figures show bootstrap support values (% of 1000 bootstrap replicates) obtained for ML trees, generated using data on all types of substitutions (above the branches) or excluding transition-type substitutions (below the branches). Branches with support values of 70% and higher are marked.

fell into the time interval of 102–71 million years ago. Accepting the point of view of the authors of the study [41] that these regions were duplicated as a result of whole genome duplication, it can be hypothesized that the latter could occur just after the appearance of the Salmonidae lineage, or even at the moment of its formation.

DISCUSSION

Phylogenetic Relationships of Thymallins, Coregonins, and Salmonins

In recent studies of the mitogenomes of the two species of *Thymallus* [7] and taimen *Hucho bleekeri* [8], it was demonstrated that thymallins were grouped

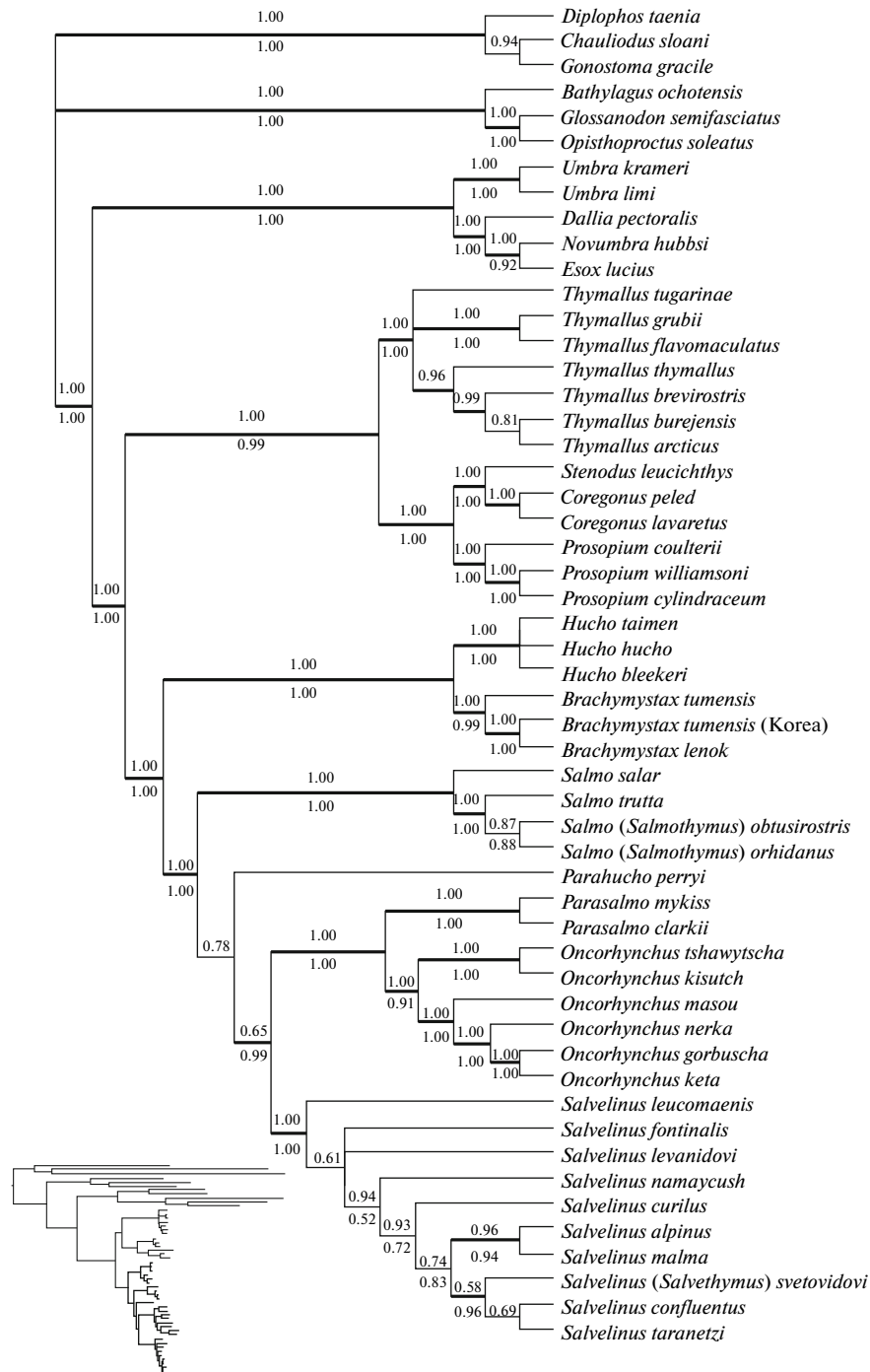


Fig. 3. Bayesian cladogram (50% majority rule consensus of 5000 trees) obtained based on partitioned analysis of mtDNA data. Figures are posterior probabilities for Bayesian trees constructed using data on all substitution types (above the branches) or excluding the substitutions of transition type (below the branches). Branches with posterior probability values of 0.95 and higher are marked.

together with salmonines, while the position of coregonines was basal relative to this group, variant (OUT,((C,(T,S))), according to the designations in Table 2. In each of these studies, about 15 species of Salmonidae and from two to three representatives of

the two outgroups were examined. At the same time, in another study [6] focused on establishing the phylogenetic relationships of the genus *Lepidogalaxias* based on the analysis of a large set (about 40) of the mitogenomes of lower Euteleostei, including five genomes of

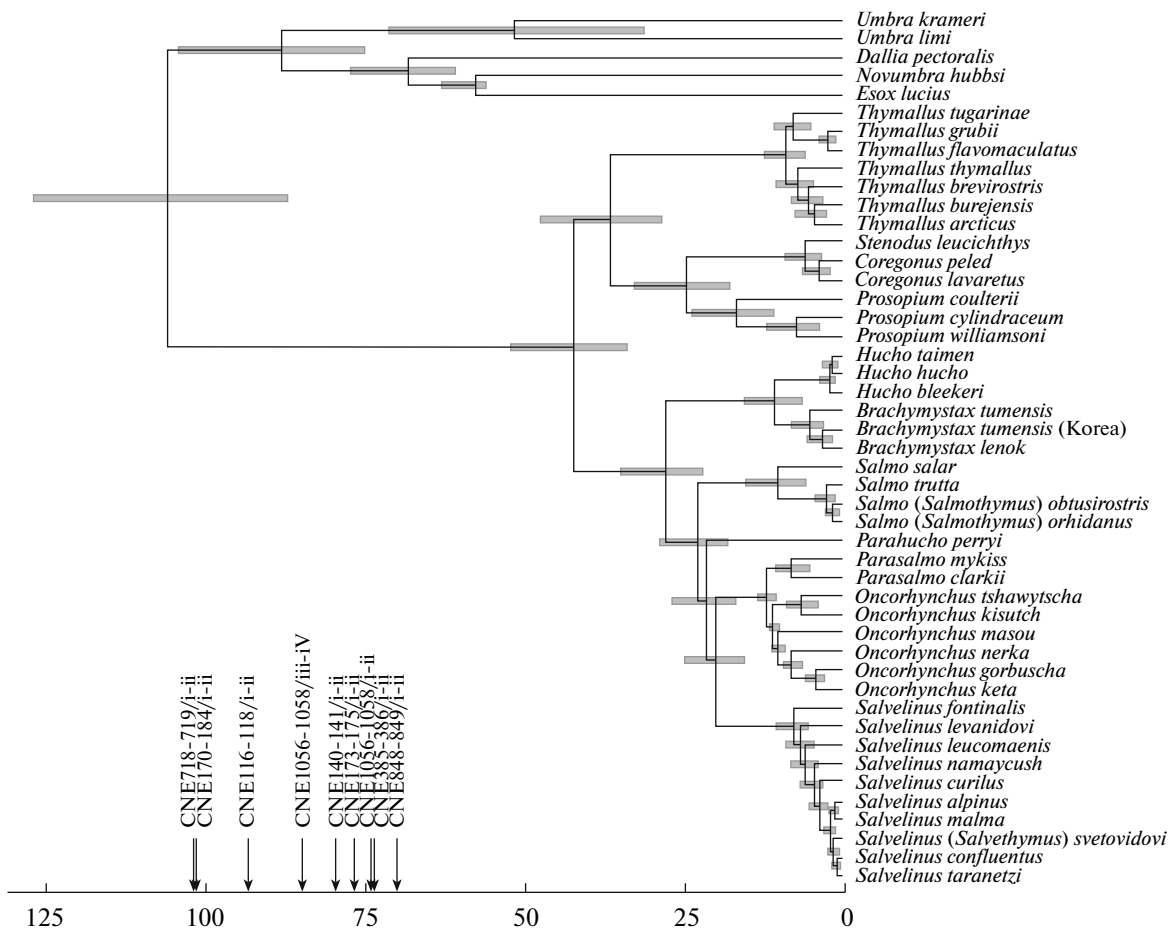


Fig. 4. Bayesian chronogram of cladogenesis of Salmonidae and close taxa translated into absolute time scale (millions of years). Boxes represent the 95% highest posterior density intervals of clade ages. Arrows indicate Bayesian time estimates of disomy establishment calculated for nine duplicated nuclear genome regions of salmonid fish flanked by the pairs of corresponding CNE.

the representatives of Salmonidae, another variant of phylogenetic relationships was found. According to this variant, thymallins and coregonins were sister taxa, (OUT,(T,C),S)).

In the present study, two variants of the phylogenetic relationships of the three main lineages of Salmonidae

were described, including (OUT,((T,C),S)) for ML and Bayesian trees and (OUT,(T,(C,S))) for MP trees. The first variant is considered to be better supported, since it is well known [42] that, under the conditions of a relatively high level of homoplasy (in our case the consistency index for MP trees (CI) constituted 0.39 and

Table 2. Statistical comparison of alternative phylogenies of Salmonidae (T, C, and S, thymallins, coregonins, and salmonins, respectively; OUT, outgroup) carried out for two data sets (PART, partitioned data with inclusion of all types of nucleotide substitutions; RY, only transversions) using the tests based on an analysis of the maximum likelihood estimates. AU is approximately unbiased test; BP is bootstrap; KH is Kishino–Hasegawa test; SH is Shimodaira–Hasegawa test

Data type	Topology	(Ln L)/ΔLn L	Test			
			AU	BP	KH	SH
PART	(OUT,((T,C),S))	(-150281.2)	0.867	0.784	0.831	0.911
	(OUT,(T,(C,S)))	11.0	0.213	0.147	0.169	0.244
	(OUT,(C,(T,S)))	13.7	0.123	0.069	0.109	0.165
RY	(OUT,((T,C),S))	(-56195.7)	0.821	0.784	0.796	0.878
	(OUT,(T,(C,S)))	8.8	0.054	0.015	0.066	0.081
	(OUT,(C,(T,S)))	5.5	0.232	0.201	0.204	0.243

0.37, for full and RY-recoded data, respectively). Maximum likelihood and Bayesian approaches were more efficient than the MP method of phylogenetic reconstruction.

It seems likely that the differences in the results of [7, 8] and [6], as well as the present study, with respect to the relationships of thymallins, coregonins, and salmonins lineages are at most associated with the difference in the number and composition of the taxa taken as outgroups relative to Salmonidae. It was expected that the inclusion of a large number of representatives of Salmonidae in the analysis, along with a fairly revealing set of outgroups, would increase the reliability of our reconstruction and minimize the number of long branches in the phylogenetic tree of Salmonidae, including its basal part. However, the closest outgroup to Salmonidae (esocoids) is not close to it in the true sense. According to our estimates, the events of the lineage splitting of the Salmonidae and Esociformes and the diversification of the main lineages within the Salmonidae are separated by a considerable time interval that constitutes about 64 million years (Fig. 4). In addition, the time interval from the divergence of the lineage of Salmoninae to the split of the lineages of Thymallinae and Coregoninae was rather small, about six million years, in terms of mtDNA dating (Fig. 4) (analogous interval for the (OUT,(T,(C,S))) variant, about three to four million years, according to the *RAG1* dating [9]). These factors, plus the inevitable degradation of the phylogenetic signal due to homoplasy, which is typical of rapidly evolving mitogenomes (in our case the rate of evolution constituted 0.44×10^{-8} substitutions per site per year per lineage, which is, e.g., ten times faster than the rate of evolution for the nuclear genome *RAG1* sequences [9]), seem to be the main causes of the uncertainty manifested in the attempts to establish phylogenetic relationships of the lineages of thymallins, coregonins, and salmonins using the mtDNA data.

The studies based on an analysis of nuclear DNA do not help in resolving the situation. An analysis of the expressed sequence tags (ESTs) of Salmonidae indicated the variant with basal position of thymallins relative to the sister taxa of coregonins and salmonins (OUT,(T,(C,S))) as the most plausible [5]. Among the 25 informative gene trees obtained in [5], 14 showed in favor of this variant, while 8 trees supported the (OUT,((T,C),S)) variant, and 3 trees supported the (OUT,(C,(T,S))) variant. These results were obtained based on an analysis of distant matrices using the NJ algorithm of the phylogenetic tree construction. We repeated the data analysis from the study [5], but using the ML method with selection of the most adequate model of nucleotide substitutions for each of the 78 potentially informative EST sets. As a result, it was demonstrated that, at a bootstrap support level of 70% and higher, eight out of nineteen informative gene trees supported the (OUT,(T,(C,S))) variant, six trees

supported the (OUT,((T,C),S)) variant, and five trees supported the (OUT,(C,(T,S))) (personal unpublished results). At a bootstrap level higher than 50% but lower than 70%, the number of these gene trees was six, eight, and four, respectively. As a result, according to our data, none of the three alternative variants had a convincing advantage, and all of them can be considered equally probable.

Based on an analysis of different copies of duplicated nuclear *Vig* gene, two different phylogeny variants, (OUT,(C,(T,S))) and (OUT,(T,(C,S))), were suggested [43]. The second variant was also inferred from an analysis of the nuclear *RAG1* gene; however, its Bayesian and bootstrap support values were low [9].

According to the opinions of some of researchers, the distribution of the copies of *MT* (type III–IV) and *MS* (type V) subfamilies of the *HpaI* family of short interspersed repetitive elements (SINEs) among the salmonid fish supports the (OUT,(C,(T,S))) variant. However, these data are fragmentary and preliminary [15].

In a number of studies aimed at resolving the phylogeny of Salmonidae based on a cladistics analysis on the morphological characteristics, the (OUT,(C,(T,S))) variant was obtained [38, 44–47]. However, the relationships within the Salmonidae found in these studies at a number of key points differ from those reported in many other phylogenetic studies performed using molecular genetic data, and which are considered to be more or less established (e.g., the phylogenetic position of *Acantholingua*, *Hucho*, *Parahucho*, *Salmo*, *Salmothymus*, *Salvelinus*, and others). Therefore, as in the case of the development of morphological phylogeny for Osmeridae fish [48], which are relatively close to Salmonidae, it seems likely that the accuracy of the Salmonidae phylogeny reconstructions listed could suffer from the inability to adequately neutralize the negative effect of homoplasy that is typical of the evolutionary transformation of different morphological structures.

Thus, the overall situation seems to be ambiguous, and the question on the true relationships of thymallins, coregonins, and salmonins remains uncertain.

Phylogenetic Relationships of the Main Evolutionary Lineages of Salmoninae

According to the data of the mitogenome analyses ([8], present study), the salmonid fish genera form two clades. One of these clades includes *Hucho* and *Brachymystax*, and another clade consists of all other genera of Salmoninae. The same pattern was obtained previously based on the sequence data for introns *C* and *D* of the growth hormone 2 gene (*GH2C* and *GH2D*) [13], the *RAG1* gene [9], and the distribution pattern of the members of *SlimII* family of SINE retroposons among the selected representatives of Salmoninae [15]. An analysis of introns *C* and *D* of the growth hormone 1 gene (*GH1C* and *GH1D*) revealed a somewhat

different pattern. Specifically, *Hucho* formed a common clade with *Parahucho*, and not with *Brachymystax* [13]. Since sister relationships of *Hucho* and *Brachymystax* are considered to be reliably established based on the analysis of a variety of phylogenetic markers, including allozymes [11], and morphological characters [1, 49], the surprising data inferred from the analysis of the *GHIC* and *GH1D* introns seem to be an artifact.

The evolutionary age of the recent species of *Hucho* and *Brachymystax* is different. The mtDNA lineages identified in species of *Brachymystax* diverged much earlier than the radiation of the mtDNA lineages in the genus *Hucho* (Fig. 4). The differences within the blunt-scouted lenok *Brachymystax* are high for the interspecific level and suggest that this species can actually be represented by two independent species. On the other hand, it cannot be excluded that this pattern was formed as a result of introgressive hybridization.

The pattern of the phylogenetic relationships in the clade of *Salmo*, *Parahucho*, *Salvelinus*, *Parasalmo*, and *Oncorhynchus* inferred from the mitogenome data was similar to that revealed previously based on an analysis of nuclear genes [9, 10, 13, 14]. In this clade, *Salmo* occupies basal position, and the lineage of *Salvelinus* (including *Salvethymus*) and that of Pacific salmon *Oncorhynchus* and trout *Parasalmo* demonstrate sister relationships. The lineage of *Parahucho* joins the later grouping (Figs. 1–3). However, the position of *Parahucho* is uncertain and, in considerable proportion to either bootstrap pseudoreplicates or Bayesian cycles, the lineage of *Parahucho* formed monophyletic group with the lineage of *Salvelinus*, thus reducing the level of support for preferable grouping of *Salvelinus*, *Parasalmo*, and *Oncorhynchus* into one clade. It should be noted in this respect that analyses of either introns *C* and *D* of the growth hormone genes 1 and 2 [13], or the *RAG1* gene [9] failed to determine the unambiguous position of the lineage of *Parahucho* relative to *Salmo*, *Salvelinus*, *Parasalmo*, and *Oncorhynchus*. One of the causes of these difficulties can be the relatively narrow time period during which the divergence of the given lineages took place (about three million years, according to mtDNA (Fig. 4), as well as the *RAG1* calibration [9]). This period was probably too short to accumulate the required number of congruent phylogenetically informative characters.

With regard to the long-discussed issue on the relationships of Pacific salmon and trout, the following should be noted. The morphological genera of *Oncorhynchus* and *Parasalmo* are well distinguished, and there are at least six unique derived characters (synapomorphies) that unite the species of Pacific salmon into monophyletic group [1]. These data are supported by the results of recent detailed allozyme analysis, which demonstrated monophyly of the genus *Oncorhynchus* [11]. In addition, based on the data on introns *C* and *D* of the growth hormone gene 1 [13], as

well as those on mitogenomes (Figs. 1–3), it follows that the genera *Oncorhynchus* and *Parasalmo* are reciprocally monophyletic. According to our estimates (Fig. 4), the time between the divergence of the ancestors of these genera and the lineage diversification within the *Oncorhynchus* is short, i.e., only about 0.8 million years. Because of this, the cases of paraphyly of the genus *Oncorhynchus* identified during the analysis of other regions of nuclear DNA, including introns *C* and *D* of the growth hormone 2 gene [13], internal transcribed spacers *ITS1* and *ITS2* of the rRNA genes [14], and *RAG1* gene [9], can be interpreted in terms of incomplete lineage sorting.

The questions on the number of species in the *Salvelinus* genus and their phylogenetic relationships have been discussed for a long time. The species of this genus can be conditionally divided into two groups. The first group will include the species with unambiguous taxonomic status, *S. fontinalis*, *S. levanidovi*, *S. leucomaenis*, *S. confluentus*, and *S. namaycush*. The second group will be comprised of *S. alpinus* and *S. malma*, and the charr species grouping around them (so-called alpinoid and malmoid charr, respectively). The species status of many of these charr is disputable.

The data of the present study once again demonstrated the following. First, the species from the second group are evolutionarily young (Fig. 4). Second, the mitogenome data alone are not sufficient to elucidate the phylogenetic relationships in the *Salvelinus* genus due to usual to this group phenomenon of introgressive hybridization accompanied by the transfer of mtDNA from one species to another ([50] and others). For instance, it follows from the mtDNA data that alpinoid (*S. alpinus*, *S. taranetzi*) and malmoid (*S. malma*, *S. curilus*) charr are reciprocally paraphyletic ([51–53], Figs. 1–3). However, based on allozyme [54, 55] and the *RAG1* gene data [9], these species form monophyletic groups. Similarly, the situation with *S. confluentus* can be considered as an evidence of introgressive hybridization [50, 53, 56]. According to the mitogenome data, this species groups in one clade with *S. taranetzi* ([50, 57], Figs. 1–3). However, the phylogenetic tree position of *S. confluentus* as revealed by allozyme data [54], the data on the *ITS1* of nuclear rRNA genes [14], and those on the introns *GHIC* and *GH2C* [58] indicates that this species is sister to *S. leucomaenis*. For these reasons, phylogenetic position of monotypic *Salvethymus* within the group of alpinoid and malmoid charr (Figs. 1–3), inferred from the mtDNA data, cannot be considered as strongly valid. At the same time, it should be noted that the results of RAPD PCR analysis also place for *Salvethymus* in the group mentioned [59]. In addition, the chromosome arm number (NF) in the karyotype of *Salvethymus* constitutes 98 [60]. This means that it is the same as in the other alpinoid and malmoid charr, including those discussed in the present study. In *S. fontinalis*, *S. levanidovi*, *S. leucomaenis*, *S. confluentus* and *S. namaycush*, NF = 100, and this state is considered

to be ancestral for charrs [60]. For these reasons, the position of *Salvelinus* in the obtained phylogenetic trees (Figs. 1–3) seems to be close to true.

Polyploidization and Diversification of Salmonidae

It is known that, after its development, but prior to its split into lineages of thymallins, coregonins, and salmonins, the Salmonidae lineage experienced whole-genome duplication, probably via autopolyploidization [61]. The first attempt to time calibrate this Salmonidae-specific whole-genome duplication was based on the evolutionary age estimates of the *LdhB* paralogs (80–100 million years), and assumed that this event could coincide in time with the formation of the salmonid lineage [62, 63]. More recently, this conclusion was reached by Osinov and Lebedev [64] based on the evolutionary age estimate of the *Vtg* paralogs (75–85 million years), which appeared to be about 10 million years younger than dichotomy of salmonids and esocoids. Our molecular calibrations, which are based on the analysis of a number of paralogous regions of nuclear genome (Fig. 4), are congruent with the estimates mentioned above and also assume that polyploidization could occur just after the isolation of the salmonid lineage.

The duration of the diversification period of the genus-level lineages in salmonids (22 million years, taking place 42–20 million years ago) is similar to that in esocoids (30 million years, taking place 88–58 million years ago). However, in the lineages of Esociformes, if this period began about 18 million years after the emergence of the lineage, then it began much later in the lineages of Salmonidae, with a large about 64 million years time lag (Fig. 4). Interestingly, the date of recent Salmonidae diversification coincided with the period when the transition from tetra- to disomic inheritance was probably the most intensive. The median of the close-to-normal distribution of estimates of synonymous distances between the 408 pairs of salmonid-specific genome-duplication ohnologs from the genome of *S. salar* constituted 0.19, which was about two times lower than that obtained upon their comparison with the orthologs from the genome of *Esox lucius* (0.43) [65]. This implies that the median of time estimates from the date when disomy is established for these genes lies somewhere around 47 million years ago ($0.19/0.43 \times 106$). To date, about 76% out of 33 isozyme loci examined in *P. mykiss* have passed through diploidization [61]. Taking this assessment as a guideline, it can be calculated that, by the time of diversification of modern Salmonidae (about 42 million years ago), the proportion of these loci in their common ancestor could be about 40%. It can be hypothesized that the achievement of this level of diploidization served as a trigger of the process of diversification of recent Salmonidae.

In plants, it was demonstrated that the rate of speciation in recently formed polyploid lineages was con-

siderably lower than in diploids [66]. In other words, in most of the cases, polyploidization is an evolutionary dead end. However, the lineages that are able to pass through it can then achieve long-term evolutionary success [66]. Despite the aforementioned lag period, salmonid fish can be attributed to the last category because, at the root of the Salmonidae phylogenetic tree, the net rate of diversification is rather high (in total, 0.078 lineages per million years; calculations were made using the LASER software package [67] based on the Bayesian chronogram obtained in the present study) and comparable to that (0.067–0.081) in rapidly evolving groups of Teleostei [21]. It should be noted in this respect that, during the past 12 million years, the diversification rate (0.108) was two times higher than in the preceding period (42–13 million years ago) (0.054). Finally, the present-day diversity of Salmonidae is much higher than that of sister to them Esociformes (66 species of 11 genera versus ten species of four genera, respectively [18]).

An analysis of the whole-genome duplication (WGD) phenomenon in the yeast *Saccharomyces cerevisiae* showed that the loss (physically, by deletion or pseudogenization) of one of the ohnologs is a common outcome for the overwhelming majority (up to 90%) of duplicated genes [68, 69]. Moreover, these processes are most intensive during the first several ten million years after the WGD event [70, 71]. Similar evolutionary dynamics of WGD ohnologs is also typical of teleost fish (Teleostei) [72].

It is assumed that divergent resolution, i.e., losses of different members of WGD duplicates in different populations, can play an important role in speciation of polyploids [71, 73]. It was believed that divergent resolution promoted the diversification of salmonid fish [73]. In our opinion, in the case of Salmonidae, the role of divergent resolution was inessential.

Compared to other whole genome duplication events happened in teleost fish, the rate of the ohnolog losses in the genomes of salmonid fish seems extremely low. In the 70–100 million years since the Salmonidae-specific WGD in lineages that lead to *P. mykiss* and *S. salar*, one duplicate was lost in 36% (12 out of 33) of isozyme loci [61] and 30% (23 out of 74) of other nuclear genes examined [5], as well as in 22% (13 out of 59) of the *Hox* cluster genes [74, 75].

In the lineage of Clupeocephala, only during the first 50–75 million years from the date of the Teleostei-specific whole-genome duplication (TSWGD) up to the radiation of the ancestors of Ostarioclupeomorpha and Euteleostei, one of the ohnologs was lost in 61% (78 out of 128) of the pigmentation genes examined, 79% (149 out of 187) of liver-specific genes [76], 61% (in 71 out of 116) of the genes involved in signal transduction and metabolic pathways [77], and 59–72% (26–31 out of 43–44) of the *Hox* cluster genes [75, 78, 79]. Taken together, the losses during this time period constituted 65–90% of all losses of duplicate genes that formed as a result of TSGD.

In the lineage of *Cyprinus*, during the short period of time from the moment of the *Cyprinus*- and *Carasius*-specific WGD (11–16 million years) [80–82], the expression of one of duplicate genes was lost in 48% of the 23 isozyme loci [83], duplicated state was lost in 40–52% of the examined microsatellite loci [81–84], and one out four (25%) *Hox* ohnologs was lost [82].

It is suggested that the main reason for the extremely low rate of the loss of WGD duplicates in the lineage of Salmonidae is the low rate of restoration of disomic inheritance in these fish [85], which, in principle, is expected in autopolyploids. In our opinion, however, this feature makes it unlikely that divergent resolution plays an important role in the diversification of recent Salmonidae.

ACKNOWLEDGMENTS

The authors are thankful to O.A. Radchenko and J. Bohlen 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. Glubokovskii, M.K., *Evolutsionnaya biologiya lososyvykh ryb* (Evolutionary Biology of Salmonids), Moscow: Nauka, 1995.
2. *Evolution Illuminated: Salmon and Their Relatives*, Hendry, A.P. and Stearns, S.C., Eds., New York: Oxford Univ. Press, 2004.
3. Volff, J.-N., Genome evolution and biodiversity in teleost fish, *Heredity*, 2005, vol. 94, no. 3, pp. 280–294.
4. 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.
5. 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.
6. 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.
7. 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.
8. Wang, Y., Guo, R., Li, H., et al., The complete mitochondrial genome of the Sichuan taimen (*Hucho bleekeri*): repetitive sequences in the control region and phylogenetic implications for Salmonidae, *Mar. Genomics*, 2011, vol. 4, no. 3, pp. 221–228.
9. Shedko S.V., Miroshnichenko, I.L., and Nemkova, G.A., Phylogeny of salmonids (Salmoniformes: Salmonidae) and its molecular dating: analysis of nuclear *RAG1* gene, *Russ. J. Genet.*, vol. 48, no. 5, pp. 575–579.
10. Oakley, T.H. and Phillips, R.B., Phylogeny of salmonine fishes based on growth hormone introns: atlantic (*Salmo*) and Pacific (*Oncorhynchus*) salmon are not sister taxa, *Mol. Phylogenet. Evol.*, 1999, vol. 11, no. 3, pp. 381–393.
11. Osinov, A.G. 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.
12. 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.
13. 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.
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. Avise, J.C., *Molecular Markers, Natural History, and Evolution*, Sunderland: Sinauer, 2004, 2nd ed.
17. Berg, W.J. and Ferris, S.D., Restriction endonuclease analysis of salmonid mitochondrial DNA, *Can. J. Fish. Aquat. Sci.*, 1984, vol. 41, pp. 1041–1047.
18. Nelson, J.S., *Fishes of the World*, New Jersey: Wiley, 2006, 4th ed.
19. Crow, K.D. and Wagner, G.P., What is the role of genome duplication in the evolution of complexity and diversity?, *Mol. Biol. Evol.*, 2006, vol. 23, no. 5, pp. 725–732.
20. Van de Peer, Y., Maere, S., and Meyer, A., The evolutionary significance of ancient genome duplications, *Nat. Rev. Genet.*, 2009, vol. 10, no. 10, pp. 725–732.
21. Santini, F., Harmon, L.J., Carnevale, G., et al., Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes, *BMC Evol. Biol.*, 2009, vol. 9, p. 194.
22. Ward, R.D., Zemplak, T.S., Innes, B.H., et al., DNA barcoding Australia's fish species, *Philos. Trans. Soc., B*, 2005, vol. 360, no. 1462, pp. 1847–1857.
23. Sevilla, R.G., Diez, A., Norén, M., et al., Primers and polymerase chain reaction conditions for DNA barcoding teleost fish based on the mitochondrial cytochrome b and nuclear rhodopsin genes, *Mol. Ecol. Notes*, 2007, vol. 7, no. 5, pp. 730–734.
24. Brunner, P.C., Douglas, M.S., Osinov, A.G., et al., Holarctic phylogeography of Arctic charr (*Salvelinus alpinus* complex) inferred from mitochondrial DNA sequences, *Evolution*, 2001, vol. 55, no. 3, pp. 573–586.
25. Yamamoto, S., Morita, K., Kitano, S., et al., Phylogeography of white-spotted charr (*Salvelinus leucomaenis*) inferred from mitochondrial DNA sequences, *Zool. Sci.*, 2004, vol. 21, no. 2, pp. 229–240.

26. Hillis, D.M., Taxonomic sampling, phylogenetic accuracy, and investigator bias, *Syst. Biol.*, 1998, vol. 47, no. 1, pp. 3–8.
27. Wiens, J.J. and Tiu, J., Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling, *PLoS One*, 2012, vol. 7, no. 8, p. e42925.
28. Roure, B., Baurain, D., and Philippe, H., Impact of missing data on phylogenies inferred from empirical phylogenomic data sets, *Mol. Biol. Evol.*, 2013, vol. 30, no. 1, pp. 197–214.
29. Katoh, K., Misawa, K., Kuma, K., and Miyata, T., MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform, *Nucleic Acids Res.*, 2002, vol. 30, pp. 3059–3066.
30. Swofford, D.L., *PAUP* Phylogenetic Analysis Using Parsimony (and Other Methods): Beta Version 10*, Sunderland: Sinauer Associates, 2002.
31. Posada, D. and Crandall, K.A., MODELTEST: testing the model of DNA substitution, *Bioinformatics*, 1998, vol. 14, no. 9, pp. 817–818.
32. Zwickl, D.J., Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion, *PhD Dissertation*, Austin: Univ. of Texas, 2006.
33. Ronquist, F. and Huelsenbeck, J.P., MRBAYES 3: Bayesian phylogenetic inference under mixed models, *Bioinformatics*, 2003, vol. 19, no. 12, pp. 1572–1574.
34. Drummond, A.J. and Rambaut, A., BEAST: Bayesian evolutionary analysis by sampling trees, *BMC Evol. Biol.*, 2007, vol. 7, no. 1, p. 214.
35. Wilson, M.V.H., Brinkman, D.B., and Neuman, A.G., Cretaceous Esocidae (Teleostei): early radiation of the pikes in North American fresh waters, *J. Paleontol.*, 1992, vol. 66, no. 5, pp. 839–846.
36. Wilson, M.V.H., Osteology of the Palaeocene teleost *Esox tiemani*, *Palaeontol.*, 1984, vol. 27, no. 3, pp. 597–608.
37. Grande, L., The first *Esox* (Esocidae: Teleostei) from the Eocene Green River Formation, and a brief review of esocid fishes, *J. Vert. Paleontol.*, 1999, vol. 19, no. 2, pp. 271–292.
38. 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.
39. 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.
40. Shimodaira, H. and Hasegawa, M., CONSEL: for assessing the confidence of phylogenetic tree selection, *Bioinformatics*, 2001, vol. 17, no. 12, pp. 1246–1247.
41. Moghadam, H.K., Ferguson, M.M., and Danzmann, R.G., Whole genome duplication: challenges and considerations associated with sequence orthology assignment in Salmoninae, *J. Fish. Biol.*, 2011, vol. 79, pp. 561–574.
42. Felsenstein, J., *Inferring Phylogenies*, Sanderland: Sinauer Associates, 2004.
43. Finn, R.N. and Kristoffersen, B.A., Vertebrate vitellogenin gene duplication in relation to the “3R hypothesis”: correlation to the pelagic egg and the oceanic radiation of teleosts, *PLoS One*, 2007, vol. 2, no. 1, p. e169.
44. Sanford, C.P.J., The phylogenetic relationships of salmonoid fishes, *Bull. Br. Mus. Nat. Hist. (Zool.)*, 1990, vol. 56, no. 2, pp. 145–153.
45. Sanford, C.P.J., Salmonoid fish osteology and phylogeny (Teleostei: Salmonoidei), vol. 33 of *Theses Zoologicae*, Vaduz: ARG Gantner, 2000.
46. Stearley, R.F. and Smith, G.R., Phylogeny of the pacific trouts and salmon (*Oncorhynchus*) and genera of the family Salmonidae, *Trans. Am. Fish. Soc.*, 1993, vol. 122, no. 1, pp. 1–33.
47. Johnson, G.D. and Patterson, C., Relationships of lower euteleostean fishes, *Interrelationships of Fishes*, Stiassny, M.L.J., Parenti, L., and Johnson, G.D., Eds., New York: Academic, 1996, pp. 251–332.
48. Ilves, K.L. and Taylor, E.B., Molecular resolution of the systematics of a problematic group of fishes (Teleostei: Osmeridae) and evidence for morphological homoplasy, *Mol. Phylogenet. Evol.*, 2009, vol. 50, no. 1, pp. 163–178.
49. Shedko, S.V., Phylogenetic relations of lenoks of the Genus *Brachymystax* (Salmonidae, Salmoniformes) and characteristics of their speciation, *Cand. Sci. (Biol.) Dissertation*, Vladivostok: Institute of Biology and Soil Science, Far Eastern Branch of Russian Academy of Sciences, 2003.
50. Phillips, R.B., Sajdak, S.L., and Domanico, M.J., Relationships among charrs based on DNA sequences, *Nord. J. Freshwater Res.*, 1995, vol. 71, pp. 378–391.
51. Radchenko, O.A., Variability of nucleotide sequences of mitochondrial DNA cytochrome b gene in charrs of the genus *Salvelinus*, *Russ. J. Genet.*, 2004, vol. 40, no. 3, pp. 244–254.
52. Oleinik, A.G., Skurikhina, L.A., and Brykov, V.A., Divergence of *Salvelinus* species from northeastern Asia based on mitochondrial DNA, *Ecol. Freshwater Fish.*, 2007, vol. 16, no. 1, pp. 87–98.
53. Shedko, S.V., Ginatulina, L.K., Miroshnichenko, I.L., and Nemkova, G.A., Phylogeography of mitochondrial DNA in South Asian dolly varden char *Salvelinus curilus* Pallas, 1814 (Salmoniformes, Salmonidae): Mediated gene introgression?, *Russ. J. Genet.*, 2007, vol. 43, no. 2, pp. 165–176.
54. Crane, P.A., Seeb, L.W., and Seeb, J.E., Genetic relationships among *Salvelinus* species inferred from allozyme data, *Can. J. Fish. Aquat. Sci.*, 1994, vol. 51, pp. 182–197.
55. Salmenkova, E.A., Omelchenko, V.T., Kolesnikov, A.A., and Malinina, T.V., Genetic differentiation of charrs in the Russian North and Far East, *J. Fish. Biol.*, 2000, vol. 57, suppl. A, pp. 136–157.
56. Phillips, R.B. and Oakley, T.H., Phylogenetic relationships among the Salmoninae based on nuclear and mitochondrial DNA sequences, *Molecular Systematics of Fishes*, Kocher, T.D. and Stepien, C.A., Eds., San Diego: Academic, 1997, pp. 145–162.
57. Grewe, P.M., Billington, N., and Hebert, P.D.N., Phylogenetic relationships among members of *Salvelinus*

- inferred from mitochondrial DNA divergence, *Can. J. Fish. Aquat. Sci.*, 1990, vol. 47, pp. 984–991.
58. Westrich, K.M., Konkol, N.R., Matsuoka, M.P., and Phillips, R.B., Interspecific relationships among charrs based on phylogenetic analysis of nuclear growth hormone intron sequences, *Environ. Biol. Fishes*, 2002, vol. 64, pp. 217–222.
 59. Shubina, E.A., Ponomareva, E.V., and Gritsenko, O.F., Population genetic structure of the char species of the Northern Kuril Islands and the rank of the Dolly Varden Char in the system of the genus *Salvelinus* (Salmonidae: Teleostei), *Zh. Obshch. Biol.*, 2006, vol. 67, no. 4, pp. 280–297.
 60. Frolov, S.V., *Izmenchivost' i evolyutsiya kariotipov lososyvykh ryb* (Karyotype Variation and the Evolution of Salmonids), Vladivostok: Dal'nauka, 2000.
 61. Allendorf, F.W. and Thorgaard, G.H., Tetraploidy and the evolution of salmonid fishes, *Evolutionary Genetics of Fishes*, Turner, B.J., Ed., New York: Plenum, 1984.
 62. Bailey, G.S. and Lim, S.T., Gene duplication in salmonid fish: Evolution of a lactate dehydrogenase with an altered function, vol. 6 of *Isozymes*, Markert, C.L., Ed., San Francisco: Academic, 1975.
 63. Lim, S.T., Kay, R.M., and Bailey, G.S., Lactate dehydrogenase isozymes of salmonid fish: evidence for unique and rapid functional divergence of duplicate H4 lactate dehydrogenases, *J. Biol. Chem.*, 1975, vol. 250, no. 5, pp. 1790–1800.
 64. 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.
 65. Leong, J. and Jantzen, S., *Salmo salar* and *Esox lucius* full-length cDNA sequences reveal changes in evolutionary pressures on a post-tetraploidization genome, *BMC Genomics*, 2010, vol. 11, p. 279.
 66. Mayrose, I., Zhan, S.H., Rothfels, C.J., et al., Recently formed polyploid plants diversify at lower rates, *Science*, 2011, vol. 333, no. 6047, p. 1257.
 67. Rabosky, D.L., LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies, *Evol. Bioinform.*, 2006, vol. 2, pp. 273–276.
 68. Wolfe, K.H. and Shields, D.C., Molecular evidence for an ancient duplication of the entire yeast genome, *Nature*, 1997, vol. 387, no. 6634, pp. 708–713.
 69. Kellis, M., Birren, B.W., and Lander, E.S., Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*, *Nature*, 2004, vol. 428, no. 6983, pp. 617–624.
 70. Byrne, K.P. and Blanc, G., Computational analyses of ancient polyploidy, *Curr. Bioinform.*, 2006, vol. 1, no. 2, pp. 131–146.
 71. Scannell, D.R., Byrne, K.P., Gordon, J.L., et al., Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts, *Nature*, 2006, vol. 440, no. 7082, pp. 341–345.
 72. Jaillon, O., Aury, J.-M., Brunet, F., et al., Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype, *Nature*, 2004, vol. 431, no. 7011, pp. 946–957.
 73. Taylor, J.S., Van de Peer, Y., and Meyer, A., Genome duplication, divergent resolution and speciation, *Trends Genet.*, 2001, vol. 17, no. 6, pp. 299–301.
 74. Mungpakdee, S., Seo, H.-C., Angotzi, A.R., et al., Differential evolution of the 13 Atlantic salmon *Hox* clusters, *Mol. Biol. Evol.*, 2008, vol. 25, no. 7, pp. 1333–1343.
 75. Henkel, C.V., Burgerhout, E., de Wijze, D.L., et al., Primitive duplicate *Hox* clusters in the European eel's genome, *PLoS One*, 2012, vol. 7, no. 2, p. e32231.
 76. Braasch, I., Brunet, F., Volff, J.-N., and Scharlt, M., Pigmentation pathway evolution after whole-genome duplication in fish, *Genome Biol. Evol.*, 2009, vol. 1, pp. 479–493.
 77. Sato, Y., Hashiguchi, Y., and Nishida, M., Temporal pattern of loss/persistence of duplicate genes involved in signal transduction and metabolic pathways after teleost-specific genome duplication, *BMC Evol. Biol.*, 2009, vol. 9, p. 127.
 78. Hoegg, S., Boore, J.L., Kuehl, J.V., and Meyer, A., Comparative phylogenomic analyses of teleost fish *Hox* gene clusters: lessons from the cichlid fish *Astatotilapia burtoni*, *BMC Genomics*, 2007, vol. 8, p. 317.
 79. Kuraku, S. and Meyer, A., The evolution and maintenance of *Hox* gene clusters in vertebrates and the teleost-specific genome duplication, *Int. J. Dev. Biol.*, 2009, vol. 53, nos. 5–6, pp. 765–773.
 80. Larhammar, D. and Risinger, C., Molecular genetic aspects of tetraploidy in the common carp *Cyprinus carpio*, *Mol. Phylogenet. Evol.*, 1994, vol. 3, no. 1, pp. 59–68.
 81. David, L., Blum, S., Feldman, M.W., et al., Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci, *Mol. Biol. Evol.*, 2003, vol. 20, no. 9, pp. 1425–1434.
 82. Yuan, J., He, Z., Yuan, X., et al., Speciation of polyploid Cyprinidae fish of common carp, crucian carp, and silver crucian carp derived from duplicated *Hox* genes, *J. Exp. Zool., Part B*, 2010, vol. 314, no. 6, pp. 445–456.
 83. Ferris, S.D. and Whitt, G.S., The evolution of duplicate gene expression in the carp (*Cyprinus carpio*), *Experientia*, 1977, vol. 33, no. 10, pp. 1299–1301.
 84. Zhang, Y., Liang, L., Jiang, P., et al., Genome evolution trend of common carp (*Cyprinus carpio* L.) as revealed by the analysis of microsatellite loci in a gynogenetic family, *J. Genet. Genomics*, 2008, vol. 35, no. 2, pp. 97–103.
 85. Li, W.H., Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes, *Genetics*, 1980, vol. 95, no. 1, pp. 237–258.

Translated by N. Maleeva

УДК 575.86+575.13:597.553.2

ФИЛОГЕНИЯ ЛОСОСЕВЫХ РЫБ (Salmoniformes: Salmonidae) И ЕЕ МОЛЕКУЛЯРНАЯ ДАТИРОВКА: АНАЛИЗ мтДНК-ДАННЫХ

© 2013 г. С. В. Шедько, И. Л. Мирошниченко, Г. А. Немкова

Биолого-почвенный институт Дальневосточного отделения Российской академии наук, Владивосток 690022
e-mail: shedko@biosoil.ru

Поступила в редакцию 07.08.2012 г.

На основе данных по митохондриальной ДНК, взятых из Genbank, и оригинальных материалов проанализированы филогенетические взаимоотношения 41 вида лососевых рыб, а также некоторые аспекты их диверсификации во времени. Положение корня филогенетического дерева Salmonidae оказалось неустойчивым. Среди возможных вариантов наиболее обоснованным выглядит тот, в котором хариусовые объединены в одну кладу с сиговыми, а линия лососевых занимает базальное относительно нее положение. Роды подсемейства Salmoninae составили две хорошо выраженные клады: (*Brachymystax*, *Hucho*) и (*Salmo*, *Parahucho*, (*Salvelinus*, (*Parasalmo*, *Oncorhynchus*))). При этом роды *Parasalmo* и *Oncorhynchus* оказались взаимно монофилетическими. Обсуждены вопросы конгруэнтности филогенетических деревьев Salmonidae, построенных по разным типам маркеров. Согласно байесовскому датированию, предковые линии лососевых и сестринских к ним щукообразных разошлись около 106 млн. лет назад. Некоторое время спустя – вероятно 100–70 млн. лет назад – произошла специфическая для лососевых полногеномная дупликация. Разделение у лососевых рыб линий родового уровня происходило много позже – в интервале 42–20 млн. лет назад. Основная волна диверсификации лососевых на видовом уровне имела место в последние 12 млн. лет. Обсуждено возможное влияние геномной дупликации на характер диверсификации Salmonidae.

DOI: 10.7868/S0016675813060118

Лососевые рыбы (сем. Salmonidae) – одна из модельных групп рыб для решения ряда актуальных проблем эволюционной биологии, геномики и других разделов современной биологии, предок которой прошел через дополнительный раунд удвоения генома [1–4]. Общее число публикаций по лососевым составляет величину, превышающую несколько десятков тысяч источников [4], и Salmonidae по праву могут считаться одной из наиболее полно и разносторонне изученных групп низших позвоночных. Тем не менее “белые пятна” во многих областях все еще сохраняются, и некоторые принципиальные вопросы ожидают своего окончательного решения, в том числе касающиеся и филогении.

Во-первых, неясен порядок дивергенции трех главных линий Salmonidae – хариусовых (подсем. Thymallinae), сиговых (Coregoninae) и лососевых (Salmoninae) рыб [5–9].

Во-вторых, нуждаются в уточнении и/или дополнительной проверке с использованием разного типа филогенетических маркеров существующие схемы филогенетических связей таксонов внутри указанных подсемейств [8–15].

Кроме того, до сих пор для Salmonidae не создана более или менее цельная, включающая большинство современных представителей, картина диверсификации линий во времени.

Митохондриальная ДНК (мтДНК) – хорошо себя зарекомендовавший филогенетический маркер, проверенный на разных уровнях таксономической иерархии в самых разных группах организмов, с понятной эволюционной динамикой и свободный от большинства проблем, связанных с явлением паралогии [16]. Данные по мтДНК, начиная с пионерской работы Берга и Фэрриса [17], неоднократно привлекались для решения тех или иных вопросов филогении лососевых рыб. Однако такие работы, как правило, охватывали лишь немногих представителей Salmonidae.

В связи с этим целью настоящего исследования было выяснение филогенетических взаимоотношений 41 вида лососевых рыб (что составляет примерно 2/3 существующего разнообразия Salmonidae [18]), а также молекулярное датирование их кладогенеза на основе анализа последовательностей митохондриальных геномов. Кроме того, в плане текущего обсуждения значения полиплоидизации в формировании разнообразия костистых рыб (Teleostei) [19–21] представлялось также интересным оценить, насколько у лососевых рыб близки во времени акт удвоения генома и основные периоды их диверсификации.

МАТЕРИАЛЫ И МЕТОДЫ

Материалом послужили последовательности мтДНК, взятые из базы Genbank/NCBI, а также последовательности генов *COI* и *Cytb*, D-петли лососевых рыб, амплифицированные нами с помощью праймеров из работ [22–24] и секвенированные с использованием набора Big Dye Terminator 3.1 и автоматического анализатора ABI Prizm 3130 (“Applied Biosystems”, США/“Hitachi”, Япония) специально для этого случая (табл. 1: JX261982–JX262010). Всего в анализ был включен 41 вид лососевых рыб, а также 11 видов, представлявших три внешние группы – Esociformes (5), Argentiformes (3) и Stomiiformes (3). Для 24 видов использованы данные по полному митохондриальному геному (16 видов Salmonidae и 8 видов из внешних групп). Для других 28 видов были доступны лишь неполные данные, из которых нами были отобраны восемь участков митогеномов: фрагменты или полные последовательности генов *AT6*, *COI*, *Cytb*, *ND1*, *ND6*, *12S* и *16S* рРНК, а также D-петли (табл. 1). Рациональность включения в анализ таксонов с неполными данными была обусловлена, во-первых, необходимостью минимизации количества длинных ветвей в филогенетическом дереве (четвертая стратегия подбора таксонов [26]), что дает положительный эффект даже при наличии пробелов в данных [27, 28]. Во-вторых, преследовалась цель охватить, по возможности, все морфо-экологическое разнообразие Salmonidae, включая те таксоны, которые в прошлом рассматривались в статусе самостоятельных родов или претендуют на это до сих пор (*Acantholingua*, *Baione*, *Cristivomer*, *Phylogephyra*, *Salmothymus* и др.) (третья стратегия – [26]).

По каждому функциональному участку мтДНК произведено выравнивание с помощью программы MAFFT v. 6 [29], а затем – конкатенация этих индивидуальных выравниваний. Матрица была просмотрена, и участки с двусмысленным выравниванием, а также все стоп-кодоны в белоккодирующих генах исключены из анализа.

Реконструкцию филогении осуществляли различными методами: взвешенной максимальной экономии (MP), наибольшего правдоподобия (ML) и байесовским (BA). При этом в каждом случае произведено по два варианта филогенетической реконструкции – с использованием информации по всем типам замен, а также исключая нуклеотидные замещения, относящиеся к транзициям.

При поиске MP-дерева с учетом всех типов нуклеотидных замен матрица данных была подразделена на шесть, предположительно структурно-однородных, блоков нуклеотидов: (1) первая, (2) вторая и (3) третья позиции кодонов 13 белоккодирующих генов, гены (4) рибосомальной и (5) транспортной РНК, а также (6) D-петля. Для

каждого из этих блоков с помощью программы RAUP 4.0b10 [30] было определено отношение скорости транзиций к скорости трансверсий, и эти округленные до целого значения оценки были использованы в качестве веса трансверсий в матрицах пошаговых переходов (step matrix) при построении MP-дерева.

Эвристический поиск MP-дерева выполнен с помощью программы RAUP в 30 повторностях со случайным характером включения последовательностей в анализ и перестановками по TBR-алгоритму. Воспроизводимость ветвлений дерева оценена методом бутстрэпа в 1000 повторных псевдослучайных выборок. Аналогичными приемами найдено MP-дерево и оценена его устойчивость на основе данных, из которых посредством RY-кодирования была исключена информация о нуклеотидных замещениях, относящихся к транзициям.

Построение ML-дерева по всем типам нуклеотидных замен проведено на матрице данных, в которой было выделено семь блоков нуклеотидов: (1) первая, (2) вторая и (3) третья позиции кодонов 12 белоккодирующих генов, (4) ген *ND6* с существенно отличным от остальных белоккодирующих генов частотным составом нуклеотидов, гены (5) рибосомальной и (6) транспортной РНК, а также (7) D-петля. Для каждого из нуклеотидных блоков с помощью программы Modeltest 3.7 [31] на основе BIC-критерия была определена наиболее подходящая модель нуклеотидных замещений: GTR + I + G, TVM + I + G, TrN + I + G, HKY + I + G, GTR + I + G, SYM + G, HKY + G, соответственно.

В варианте ML-анализа на основе матрицы RY-кодированных данных была использована модель нуклеотидных замещений CF + I + G.

Эвристический поиск ML-деревьев осуществлен с помощью программы Garli v. 2 [32] в 30 повторностях. Устойчивость ветвлений дерева оценена методом бутстрэпа с использованием 1000 повторных псевдослучайных выборок.

Байесовский анализ филогении выполняли с помощью программы MrBayes 3.2 [33]. Подразделенный анализ выполнен по предыдущей ML-схеме, за исключением того, что модели TVM + I + G и TrN + I + G были заменены на более общую GTR + I + G (из-за ограничений программы MrBayes). Трансверсионный анализ был также проведен с использованием модели нуклеотидных замещений CF + I + G. В обоих вариантах анализ включал в себя 6×10^6 циклов с отбором каждого тысячного из генерированных деревьев. Из полученных деревьев первые 1001 были отброшены, а оставшиеся 5000 деревьев, характеризовавшихся стабилизированными оценками правдоподобия (LnL), параметрами моделей нуклеотидных замещений и длинами деревьев,

Таблица 1. Перечень рассмотренных видов рыб с Genbank-номерами участков их мтДНК (префикс "NC_" маркирует митогеномы с полной нуклеотидной последовательностью)

Виды	AT6	COI	Cytb	ND1	ND6	12S	16S	D-петля
Stomiiformes								
<i>Diplophos taenia</i>	NC_002647	NC_002647	NC_002647	NC_002647	NC_002647	NC_002647	NC_002647	
<i>Gonostoma gracile</i>	NC_002574	NC_002574	NC_002574	NC_002574	NC_002574	NC_002574	NC_002574	
<i>Chauliodus sloani</i>	NC_003159	NC_003159	NC_003159	NC_003159	NC_003159	NC_003159	NC_003159	
Argentiformes								
<i>Bathylagus ochotensis</i>	NC_004591	NC_004591	NC_004591	NC_004591	NC_004591	NC_004591	NC_004591	
<i>Opisthoproctus soleatus</i>	NC_004600	NC_004600	NC_004600	NC_004600	NC_004600	NC_004600	NC_004600	
<i>Glossanodon semifasciatus</i>	NC_004595	NC_004595	NC_004595	NC_004595	NC_004595	NC_004595	NC_004595	
Esociformes								
<i>Esox lucius</i>	NC_004593	NC_004593	NC_004593	NC_004593	NC_004593	NC_004593	NC_004593	NC_004593
<i>Dallia pectoralis</i>	NC_004592	NC_004592	NC_004592	NC_004592	NC_004592	NC_004592	NC_004592	NC_004592
<i>Umbra limi</i>		FJ918928	AY497458			AY430268	AF060443	
<i>U. krameri</i>		JX261985	JX261998			AY430269	AF060444	
<i>Novumbra hubbsi</i>			AY497457			AY430272	AF060447	
Salmoniformes								
Salmonidae								
Coregoninae								
<i>Coregonus lavaretus</i>	NC_002646	NC_002646	NC_002646	NC_002646	NC_002646	NC_002646	NC_002646	NC_002646
<i>C. peled</i>			DQ185402	DQ399871			DQ399871	AF195647
<i>Stenodus leucichthys</i>	AY778968	EU202658	EU003527					HM448022
<i>Prosopium williamsoni</i>	AY778970	FJ918926						
<i>P. cylindraceum</i>	HQ616435	EU202657	EU003531					
<i>P. coulterii</i>	HQ616436	EU202654	AY382564				AY430236	AY382565
Thymallinae								
<i>Thymallus thymallus</i>	NC_012928	NC_012928	NC_012928	NC_012928	NC_012928	NC_012928	NC_012928	NC_012928
<i>Th. arcticus</i>	NC_012929	NC_012929	NC_012929	NC_012929	NC_012929	NC_012929	NC_012929	NC_012929
<i>Th. burejensis</i>	AY778976							AY779007
<i>Th. brevirostris</i>	AY778990							AY168387
<i>Th. flavomaculatus</i>	AY778983	JX261987	JX262000					AY246410
<i>Th. grubii</i>	AY778972	JX261988	JX262001					AY168390
<i>Th. tugarinae</i>	AY778998	JX261986	JX261999					AY246389

Таблица 1. Окончание

Виды	AT6	COI	Cytb	NDI	ND6	I2S	I6S	D-петля
Salmoninae								
<i>Brachymystax tumensis</i> (Корея)			AF125213		AF125052	AF125513	AF125513	AF125519
<i>B. tumensis</i> (р. Амур)	AY862340	JX261989	JX262002	EU760486				EU760491
<i>B. lenok</i> (р. Амур)	AY862339	JX261990	JX262003	EU760488				EU760490
<i>Hucho taimen</i>	AY862334	JX261991	JX262004	EU760487				EU760489
<i>H. hucho</i>	AY862338		AF172397	EU729362				EU729360
<i>H. bleekeri</i>	NC_015995	NC_015995	NC_015995	NC_015995	NC_015995	NC_015995	NC_015995	NC_015995
<i>Parahucho perryi</i>	AY862342	JX261992	JX262005	AY862386				AY862364
<i>Salmo salar</i>	NC_001960	NC_001960	NC_001960	NC_001960	NC_001960	NC_001960	NC_001960	NC_001960
<i>S. trutta</i>	NC_010007	NC_010007	NC_010007	NC_010007	NC_010007	NC_010007	NC_010007	NC_010007
<i>S. (Salmothymus) ohridanus</i>	EU282329		AF202033					AY926564
<i>S. (Salmothymus) obtusirostris</i>	EF469834		AF488534					EF469832
<i>Salvelinus fontinalis</i>	NC_000860	NC_000860	NC_000860	NC_000860	NC_000860	NC_000860	NC_000860	NC_000860
<i>S. alpinus</i>	NC_000861	NC_000861	NC_000861	NC_000861	NC_000861	NC_000861	NC_000861	NC_000861
<i>S. leucomaenis</i>		JX261993	JX262006					DQ403177
<i>S. namaycush</i>		EU522425	DQ451386				AF060445	AF297989
<i>S. taranetzi</i>		JX261995	JX262008					JX261983
<i>S. curilus</i>		JX261997	JX262010					DQ403180
<i>S. malma</i>		JX261996	JX262009					JX261984
<i>S. levanidovi</i>		JX261994	JX262007					JX261982
<i>S. confluentus</i>		EU524365	табл. 1 в [25]					
<i>S. (Salvelinichthys) svetovidovi</i>								AF297990
<i>Parasalmo clarkii</i>	NC_006897	NC_006897	NC_006897	NC_006897	NC_006897	NC_006897	NC_006897	NC_006897
<i>P. mykiss</i>	NC_001717	NC_001717	NC_001717	NC_001717	NC_001717	NC_001717	NC_001717	NC_001717
<i>Oncorhynchus tshawytscha</i>	NC_002980	NC_002980	NC_002980	NC_002980	NC_002980	NC_002980	NC_002980	NC_002980
<i>O. kisutch</i>	NC_009263	NC_009263	NC_009263	NC_009263	NC_009263	NC_009263	NC_009263	NC_009263
<i>O. nerka</i>	NC_008615	NC_008615	NC_008615	NC_008615	NC_008615	NC_008615	NC_008615	NC_008615
<i>O. gorbuscha</i>	NC_010959	NC_010959	NC_010959	NC_010959	NC_010959	NC_010959	NC_010959	NC_010959
<i>O. keta</i>	NC_017838	NC_017838	NC_017838	NC_017838	NC_017838	NC_017838	NC_017838	NC_017838
<i>O. masou</i>	NC_008747	NC_008747	NC_008747	NC_008747	NC_008747	NC_008747	NC_008747	NC_008747

Примечание. Номера, выделенные жирным шрифтом, – последовательности, секвенированные в настоящей работе.

использованы для построения консенсусного филогенетического дерева и получения оценок апостериорной вероятности его ветвлений.

Молекулярное датирование кладогенеза Salmonidae проведено в рамках байесовского подхода с помощью пакета программ BEAST 1.6.2 [34] при следующих условиях: топология дерева фиксирована — ML-дерево; модель молекулярных часов — варьирующие (не строгие) часы с некоррелированным логнормальным распределением скоростей нуклеотидных замен по ветвям дерева и актами видообразования, подчиняющимися распределению Юла; модель нуклеотидных замещений — GTR + I + G; протяженность байесовского анализа — 30×10^6 циклов с отбором каждого двухтысячного из генерированных деревьев; burnin — 1501. Для привязки ультраметрического дерева к абсолютной шкале времени было использовано четыре калибровочных интервала. Первый из них допускал разделение лососеобразных и щукообразных в мелу (145.5–65.5 млн. лет назад) и базировался на множестве позднемиоценовых ископаемых остатков Esociformes [35]. Второй связывал разделение *Esox* и *Novumbra* с палеоценом (65.5–55.8 млн. лет назад) и основывался на ряде хорошо сохранившихся ископаемых остатков рыб, вполне обоснованно относимых к роду *Esox* [36, 37]. Третий интервал ограничивал обособление линии Salmoninae пределами эоцена (55.8–33.9 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 [38]. Четвертый был основан на многочисленных ископаемых остатках лососевых рыб, проявляющих продвинутые черты тихоокеанских лососей рода *Oncorhynchus* (см.: [39], а также литературу из этой работы), и ограничивал время начала диверсификации рецентных линий *Oncorhynchus* интервалом позднего миоцена (11.6–5.3 млн. лет назад).

РЕЗУЛЬТАТЫ

После удаления из составной матрицы участков с двусмысленным выравниванием, а также всех стоп-кодонов в белковых генах, ее размерность составила 53×16388 пн. Итоговая матрица данных включала в себя последовательности 18 полных мтДНК-геномов, 6 митогеномов без контролируемых участков (в митогеномах представителей Stomiiformes и Argentiformes эти участки были исключены из-за невозможности произвести их полноценное выравнивание с таковыми остальными последовательностями), а также 29 составных нуклеотидных последовательностей. Суммарная длина последних варьировала в пределах 543–5064 пн (в среднем 2660 пн).

В результате филогенетического анализа подразделенной матрицы данных с помощью метода взвешенной максимальной экономии было найдено единственное MP-дерево (рис. 1). Большая часть внутренних ветвей (37 из 49) этого дерева имела высокий уровень бутстрэп-поддержки, равный или превышающий 70%. Топология MP-дерева, полученного методом трансверсионной максимальной экономии, практически полностью повторяла таковую дерева на рис. 1, но общий уровень бутстрэп-поддержки оказался несколько ниже — только 30 из 49 внутренних ветвей имели уровень поддержки в 70 и более процентов.

Кладограмма, полученная в результате подразделенного ML-анализа, приведена на рис. 2. Ее топология имеет лишь одно существенное отличие от топологии MP-дерева — на ML-дереве сиговые образовали одну кладу с хариусовыми, в то время как на MP-дереве сиговые объединились с лососевыми. Топология ML-дерева, построенного на основе анализа RY-кодированных данных, ничем существенным не отличалась от таковой дерева на рис. 2. В целом 38 из 50 внутренних ветвей дерева на рис. 2 имели высокий (равный или превышающий 70%) уровень бутстрэп-поддержки в том или ином варианте ML-анализа (в обоих случаях — 31 внутренняя ветвь).

Кладограммы, полученные в результате байесовского анализа, с учетом всех типов нуклеотидных замен или без учета транзиций, мало чем отличались друг от друга, и их топология (рис. 3) практически полностью повторила таковую ML-кладограммы. По общему числу ветвлений, получивших высокую поддержку (38 из 49 внутренних ветвей имели вероятность 0.95 и выше), и их распределению по дереву BA- и ML-кладограммы также оказались чрезвычайно сходны.

Как это можно видеть (рис. 1–3), построенные тремя различными методами кладограммы оказались в основном конгруэнтны. Исключение составило резко отличное взаиморасположение линий хариусовых, сиговых и лососевых рыб на максимально экономной кладограмме (рис. 1) и кладограммах, построенных при использовании максимально-правдоподобного или байесовского методов (рис. 2, 3). Чтобы оценить, насколько это отличие существенно, было проведено сравнение трех альтернативных вариантов филогенетических взаимоотношений хариусовых, сиговых и лососевых рыб с помощью нескольких статистических тестов, основанных на анализе оценок правдоподобия и реализованных в пакете программ CONSEL [40] (табл. 2). Большинство из них показали, что отличия наиболее предпочтительного варианта филогении Salmonidae (рис. 2, 3) от альтернативных (включая вариант, приведенный на рис. 1) статистически не значимы.

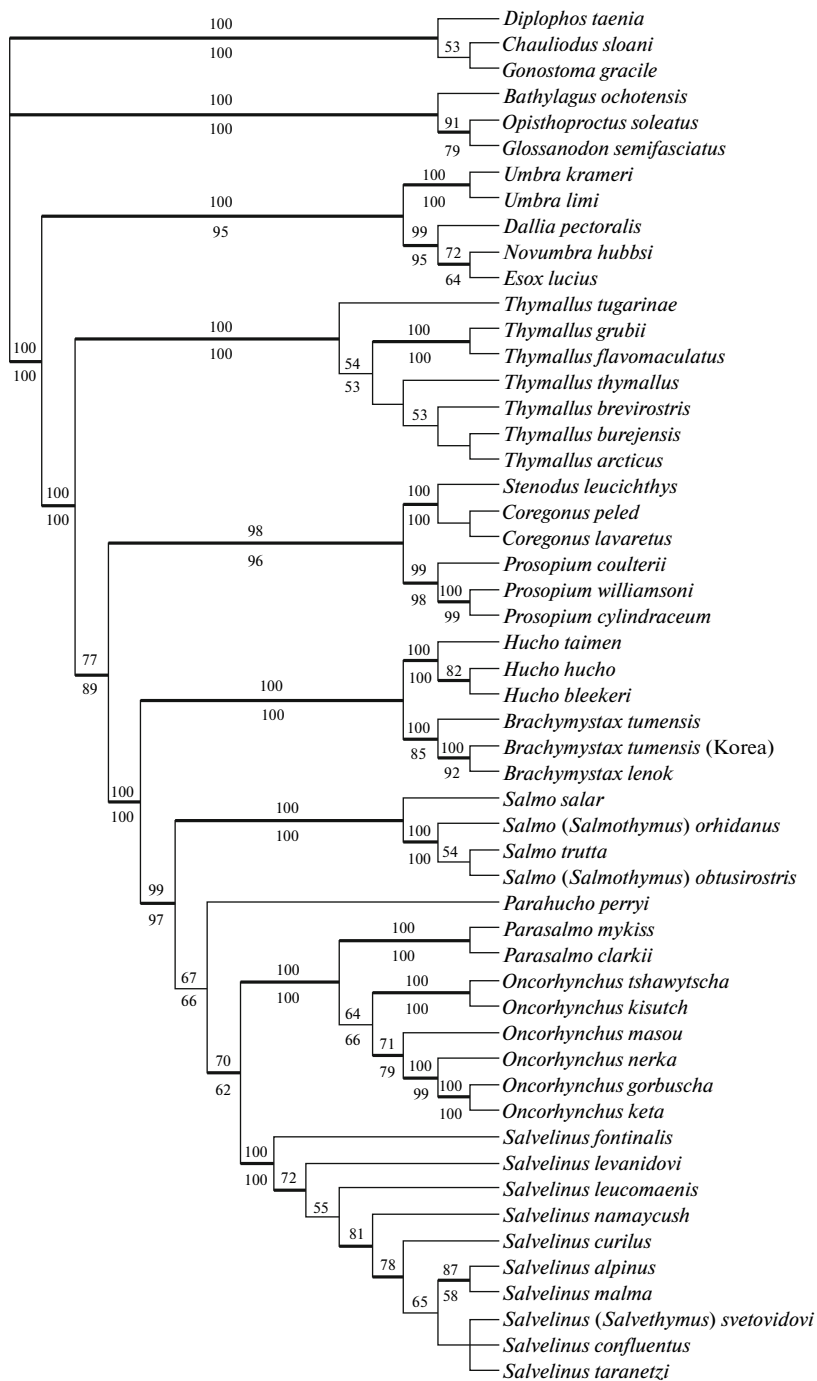


Рис. 1. МР-кладограмма (76011 шагов), полученная в результате анализа матрицы мтДНК-данных методом взвешенной максимальной экономии. Цифры над и под ветвями – бутстрэп-оценки (% от 1000 реплик бутстрэпа), полученные для МР-деревьев, построенных с использованием информации по всем типам замен (6713 филогенетически информативных позиций) или исключая нуклеотидные замещения, относящиеся к транзициям (3391 филогенетически информативная позиция), соответственно. Выделены ветви с оценками, превышающими или равными 70%.

Молекулярная датировка проведена на основе фиксированной топологии филогенетического дерева, полученного методом максимального правдоподобия (рис. 2). Итоговая хронограмма кладогенеза лососевых рыб и представителей

ближайшей к ним внешней группы (Esociformes) дана на рис. 4.

Согласно проведенным расчетам для данного набора таксонов Salmonidae средний эволюционный возраст видов (длина терминальных ветвей

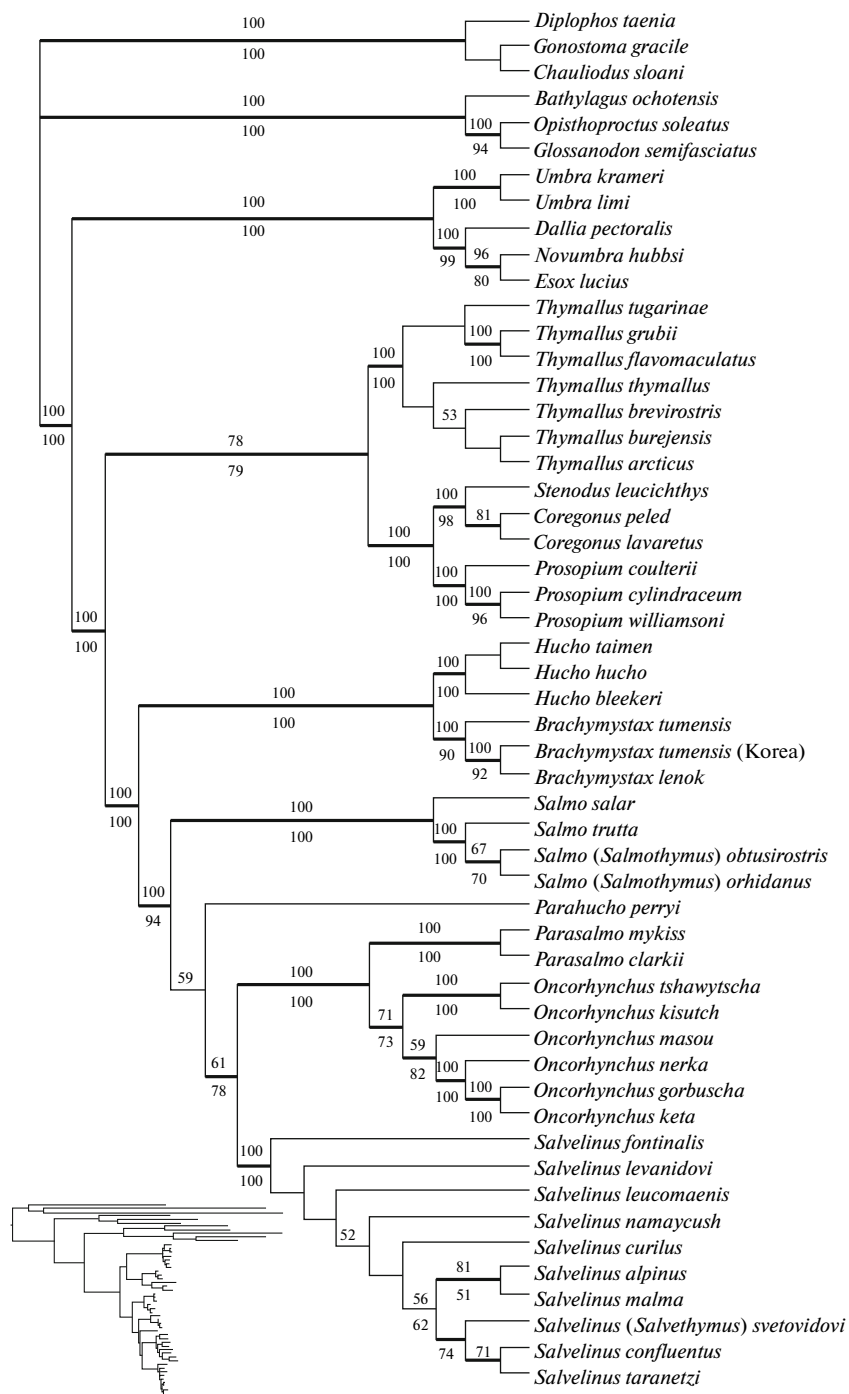


Рис. 2. ML-клатограмма, полученная на основе подразделенного анализа мтДНК-данных. Цифры – бутстрэп-оценки (% от 1000 реплик бутстрэпа) для ML-деревьев, построенных с использованием информации по всем типам замен (над ветвями) или исключая нуклеотидные замещения, относящиеся к транзициям (под ветвями). Выделены ветви с оценками, превышающими или равными 70%.

деревья на рис. 4) составил величину в 5.6 млн. лет (пределы варьирования – 0.9–21.4, медиана – 4.8).

Среди видов Salmonidae можно выделить несколько сравнительно компактных групп, объ-

единяющих представителей одного рода (*Thymallus*, *Prosopium*) или, что чаще, близких родов (*Stenodus* и *Coregonus*; *Hucho* и *Brachymystax*; *Salmo* и *Salmothymus*; *Parasalmo* и *Oncorhynchus*; *Salvelinus* и *Salvethymus*). Эволюционный возраст наиболее

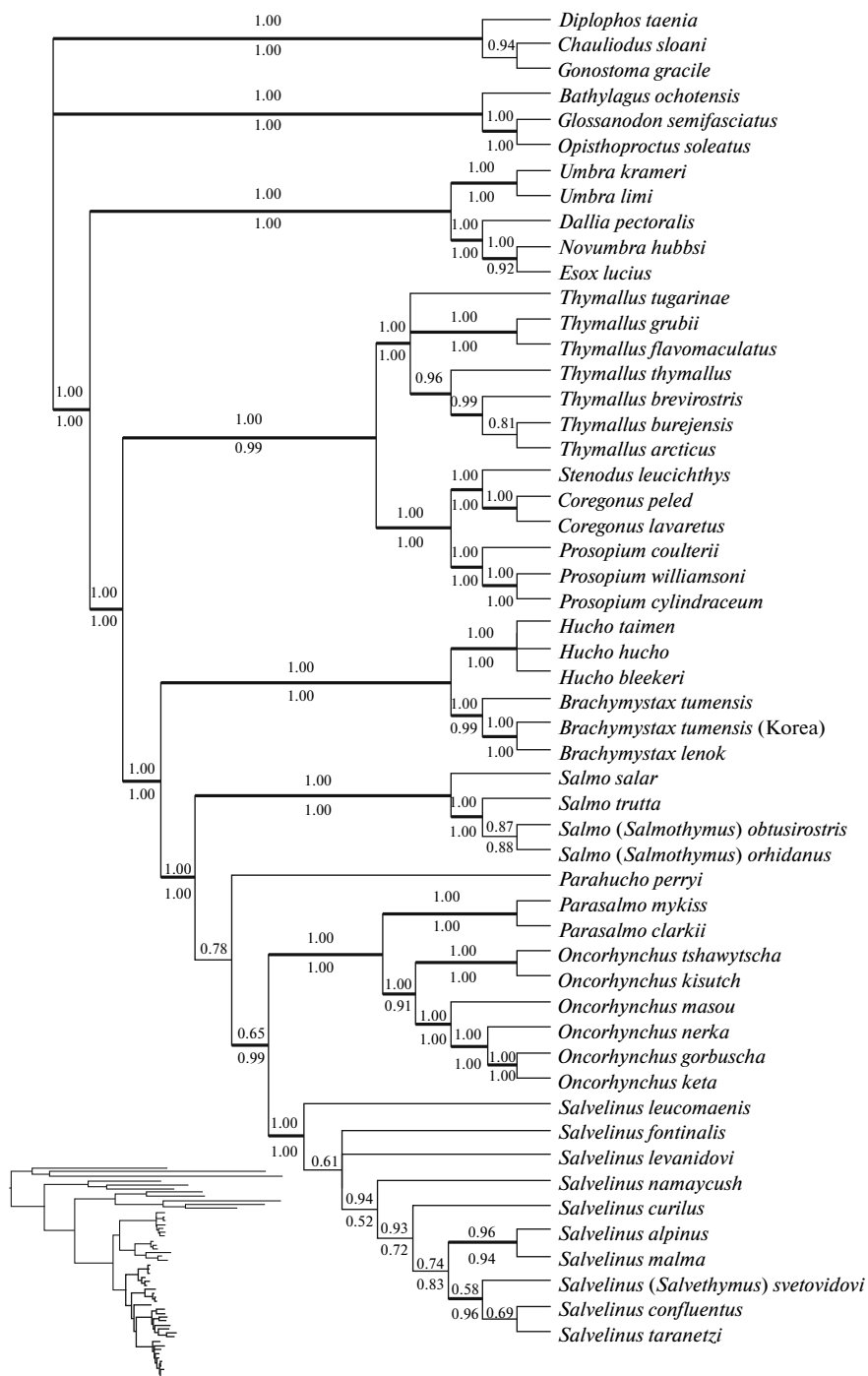


Рис. 3. Байесовская кладограмма (50%-ный консенсус 5000 деревьев), полученная на основе подразделенного анализа мтДНК-данных. Цифры – значения апостериорной вероятности для байесовских деревьев, построенных с использованием информации по всем типам замен (над ветвями) или исключая нуклеотидные замещения, относящиеся к транзициям (под ветвями). Выделены ветви с оценками, превышающими или равными 0.95.

близкого общего предка в таких группах варьировал в пределах 6.1–17.0 млн. лет и в среднем составил 10.4 млн. лет (медиана – 10.3).

В свою очередь, предковые линии этих групп видов, а также линия монотипического рода *Parahucho* дивергировали друг от друга много

раньше – от 42.3 до 19.9 млн. лет назад. Причем, пять из семи таких актов дивергенции произошли в интервале 27.8–19.9 (в среднем 23.4) млн. лет назад.

Датирование специфической для линии Salmonidae полногеномной дупликации проведе-

Таблица 2. Результаты статистического сравнения альтернативных вариантов филогении Salmonidae (Т, С и S – хариусовые, сиговые и лососевые рыбы, соответственно; OUT – внешняя группа), выполненного для двух наборов данных (PART – подразделенные данные с включением всех типов нуклеотидных замещений, RY – только трансверсии) с использованием тестов, основанных на анализе оценок правдоподобия: AU – несмещенного, BP – бутстрэп, KH – Кишино-Хэсегавы, SH – Шимодайра-Хэсегавы

Тип данных	Топология	$(\ln L)/\Delta \ln L$	Тест			
			AU	BP	KH	SH
PART	(OUT,((T,C),S))	(-150281.2)	0.867	0.784	0.831	0.911
	(OUT,(T,(C,S)))	11.0	0.213	0.147	0.169	0.244
	(OUT,(C,(T,S)))	13.7	0.123	0.069	0.109	0.165
RY	(OUT,((T,C),S))	(-56195.7)	0.821	0.784	0.796	0.878
	(OUT,(T,(C,S)))	8.8	0.054	0.015	0.066	0.081
	(OUT,(C,(T,S)))	5.5	0.232	0.201	0.204	0.243

но с использованием данных по участкам ядерных геномов четырех видов лососевых рыб (*Coregonus clupeaformis*, *Salmo salar*, *Salvelinus alpinus* и *Parasalmo mykiss*), фланкированных консерватив-

ными некодирующими элементами (CNE) (табл. 1 в: [41]). При расчете времени дивергенции девяти пар таких паралогических участков в качестве калибровочных выступили оценки эво-

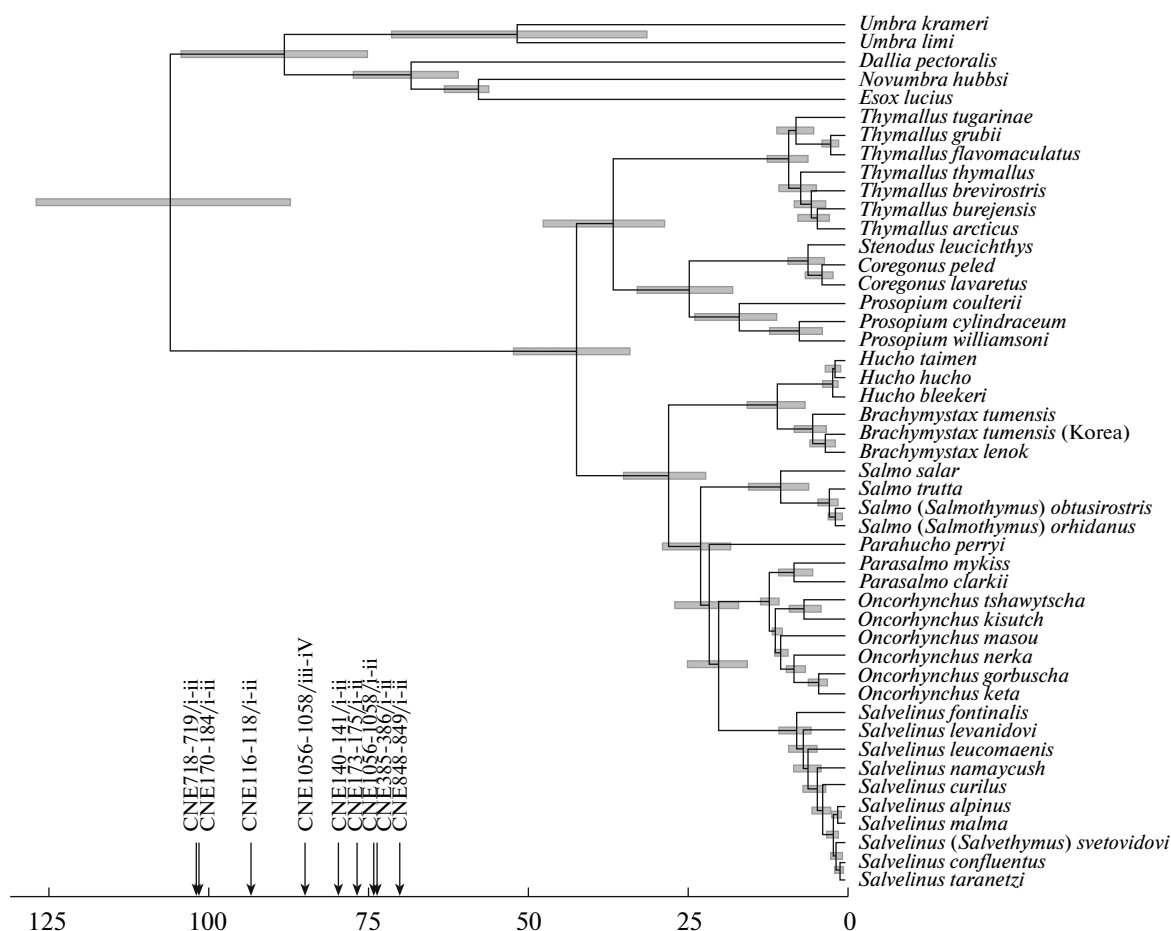


Рис. 4. Байесовская хронограмма кладогенеза Salmonidae и близких к ним таксонов, выраженная в абсолютной временной шкале (млн. лет). Прямоугольники – 95%-ные интервалы наивысшей апостериорной плотности оценок времени существования клад. Стрелки – байесовские оценки времени установления дисомии, рассчитанные для девяти дублированных участков ядерного генома лососевых рыб, фланкированных парами соответствующих CNE.

люционного возраста дихотомий сиговые–лососевые (42 млн. лет) и *Salmo–Salvelinus/Parasalmo* (23 млн. лет), полученные в настоящей работе. Датировка, как и ранее, была выполнена с помощью программ пакета BEAST [34] при следующих дополнительных условиях: модель нуклеотидных замещений – HKY + G; молекулярные часы – строгие; протяженность байесовского анализа – 2×10^6 циклов с отбором каждого тысячного; burnin – 501.

В результате было найдено (рис. 4), что медианы байесовских оценок времени с момента начала дивергенции паралогов (времени установления дисомии) уложились в промежуток 102–71 млн. лет назад. Принимая точку зрения авторов работы [41], что эти участки были удвоены в результате полногеномной дупликации, можно утверждать, что последняя могла произойти почти сразу же после появления линии Salmonidae или даже в момент ее образования.

ОБСУЖДЕНИЕ

Филогенетические взаимоотношения хариусовых, сиговых и лососевых

В недавних исследованиях митогеномов двух видов *Thymallus* [7] и тайменя *Hucho bleekeri* [8] было найдено, что хариусовые рыбы составляют с лососевыми одну группировку, а сиговые занимают базальное, относительно нее, положение – вариант (OUT,(C,(T,S))), согласно обозначениям из табл. 2. В каждой из этих работ было задействовано около полутора десятка видов Salmonidae и 2–3 представителя от пары внешних групп. Однако в другом исследовании [6], посвященном установлению филогенетических связей рода *Lepidogalaxias* на основе анализа большого набора (порядка 40) митогеномов низших Euteleostei, включая пять митогеномов представителей Salmonidae, был найден иной вариант, в котором хариусовые и сиговые рыбы находятся в сестринских отношениях – (OUT,((T,C),S)).

В настоящей работе получено два варианта филогенетических взаимоотношений трех главных линий Salmonidae – (OUT,((T,C),S)) в случае ML или байесовских деревьев, а также (OUT,(T,(C,S))) в случае MP деревьев. Первый из них можно считать более обоснованным, поскольку хорошо известно [42], что в условиях сравнительно высокого уровня гомоплазии (в нашем случае индекс соответствия для MP-деревьев CI – 0.39 и 0.37 для полных и RY-перекодированных данных, соответственно) максимально правдоподобный и байесовский подходы оказываются более эффективными, чем максимально экономный метод филогенетической реконструкции.

Различие результатов исследований [7, 8], с одной стороны, а также [6] и настоящей работы, с другой, в плане взаимоотношений линий хариусовых, сиговых и лососевых рыб, по-видимому, обусловлено в основном разницей в числе и составе таксонов, взятых в качестве внешних, по отношению к Salmonidae, групп. Большое число включенных в анализ представителей Salmonidae вкупе с достаточно показательным набором внешних групп должны были, по идее, повысить надежность нашей реконструкции, минимизируя количество длинных ветвей в филогенетическом дереве Salmonidae в целом и базальной его части в том числе. Однако ближайшая к Salmonidae внешняя группа (щукообразные) не является, к сожалению, близкой в прямом смысле. Момент расхождения линий Salmonidae и Esociformes и момент обособления главных линий внутри Salmonidae разделяет, по нашей оценке, значительный промежуток времени – около 64 млн. лет (рис. 4). Вдобавок, промежуток времени от момента отделения линии Salmoninae и до разделения линий Thymallinae и Coregoninae сравнительно мал – порядка 6 млн. лет по мтДНК-датированию (рис. 4) (аналогичный интервал для варианта (OUT,(T,(C,S))) – около 3–4 млн. лет, согласно RAG1-датированию [9]). Эти обстоятельства плюс неизбежная деградация филогенетического сигнала вследствие гомоплазии, свойственной быстро эволюционирующим митогеномам (а в нашем случае скорость составила в среднем 0.44×10^{-8} замен на сайт в год на линию, что, к примеру, в 10 раз быстрее, чем таковая для последовательностей ядерного гена RAG1 [9]), по-видимому и являются основными причинами той неопределенности, которая проявляется при попытках установления филогенетических взаимоотношений линий хариусовых, сиговых и лососевых рыб с помощью мтДНК-данных.

Работы, основанные на анализе ядерной ДНК, разрешить ситуацию не помогают. По результатам рассмотрения концевых меток экспрессирующихся последовательностей (EST) Salmonidae было, в частности, найдено [5], что наиболее вероятным является вариант с базальной позицией хариусовых рыб относительно сестринской сиговых и лососевых – (OUT,(T,(C,S))). Из 25 информативных генных деревьев, полученных в работе [5], 14 свидетельствовали в пользу этого варианта, 8 – в пользу (OUT,((T,C),S)) и 3 – в пользу варианта (OUT,(C,(T,S))). Эти результаты были получены на основе анализа матриц дистанций с помощью NJ-алгоритма построения филогенетических деревьев. Мы повторили анализ данных из работы [5], но при использовании ML-метода и с подбором наиболее адекватной модели нуклео-

тидных замещений для каждого из 78 потенциально информативных наборов EST. В результате нашли (наши неопубликованные данные), что на уровне бутстрэп-поддержки в 70 и более процентов из 19 информативных генных деревьев 8 свидетельствовали в пользу (OUT,(T,(C,S))), 6 – (OUT,((T,C),S)), 5 – (OUT,(C,(T,S))). На уровне бутстрэп-поддержки более 50%, но менее 70% таких генных деревьев было 6, 8 и 4, соответственно. В итоге, по нашим данным, ни один из трех альтернативных вариантов убедительного перевеса не получил, и все они могут считаться равновероятными.

Анализ разных копий дублированного ядерного гена *Vtg* приводит к двум разным вариантам – (OUT,(C,(T,S))) и (OUT,(T,(C,S))) [43]. Последний из них получен также при анализе ядерного гена *RAG1*, но его бутстрэп и байесовская поддержки оказались невелики [9].

По мнению некоторых исследователей, распределение среди лососевых рыб копий *MT* (тип III–IV) и *MS* (тип V) подсемейств семейства *HpaI* коротких рассеянных повторов (SINE) свидетельствует в пользу варианта (OUT,(C,(T,S))), но эти данные фрагментарны и носят предварительный характер [15].

В нескольких исследованиях, направленных на решение вопросов филогении Salmonidae на основе кладистического анализа морфологических признаков, был получен вариант (OUT,(C,(T,S))) [38, 44–47]. Однако установленные в этих работах взаимоотношения внутри Salmoninae в ряде принципиальных моментов отличались от тех, что были найдены в многочисленных филогенетических исследованиях, проведенных с использованием молекулярно-генетических данных, и которые сейчас считаются более или менее твердо установленными (это, к примеру, филогенетическое положение *Acantholingua*, *Hucho*, *Parahucho*, *Salmo*, *Salmothymus*, *Salvelinus* и т.п.). Поэтому, как и в случае разработки морфологической филогении относительно близких к лососевым корюшкообразных рыб [48], вполне вероятно, что точность вышеуказанных реконструкций филогении Salmonidae могла пострадать из-за невозможности адекватным образом нивелировать негативный эффект гомоплазии, обычной для эволюционной трансформации различного рода морфоструктур.

Таким образом, ситуация в целом складывается неоднозначная, и вопрос об истинных филогенетических взаимоотношениях хариусовых, сиговых и лососевых рыб все еще остается открытым.

Филогенетические взаимоотношения основных эволюционных линий Salmoninae

Согласно результатам анализа митогеномов ([8], настоящая работа), роды лососевых рыб образуют две клады, одну из которых составляют *Hucho* и *Brachymystax*, а другую – все остальные роды Salmoninae. Такая же картина была получена ранее по результатам анализа нуклеотидных последовательностей интронов *C* и *D* гена гормона роста 2 (*GH2C* и *GH2D*) [13], гена *RAG1* [9], а также паттерна распределения членов семейства *SlmII* SINE-ретропозонов среди отобранных представителей Salmoninae [15]. Анализ интронов *C* и *D* гена гормона роста 1 (*GH1C* и *GH1D*) выявил несколько иную картину – род *Hucho* образовал общую кладу не с *Brachymystax*, а с *Parahucho* [13]. Поскольку сестринские отношения *Hucho* и *Brachymystax* считаются надежно установленными по результатам анализа самых различных филогенетических маркеров, включая аллозимы [11] и морфологические признаки [1, 49], то необычный результат по интронам *GH1C* и *GH1D*, скорее всего, является неким артефактом.

Эволюционный возраст рецентных видов *Hucho* и *Brachymystax* не одинаков. Выявленные у видов *Brachymystax* линии мтДНК дивергировали намного раньше, чем разделились таковые в роде *Hucho* (рис. 4). Различия внутри тупорылового ленка *B. tumensis* велики для внутривидового уровня и сигнализируют о том, что этот вид может на самом деле состоять из двух независимых видов. С другой стороны, не исключено, что такая картина могла сложиться в результате интрогрессивной гибридизации.

Паттерн филогенетических связей в кладе *Salmo*, *Parahucho*, *Salvelinus*, *Parasalmo* и *Oncorhynchus* по митогеномным данным сходен с тем, что был выявлен ранее при анализе ядерных генов [9, 10, 13, 14]. *Salmo* в этой кладе занимает базальное положение, а линия *Salvelinus* (включая *Salvethymus*) и линия тихоокеанских лососей *Oncorhynchus* и форелей *Parasalmo* демонстрируют сестринские отношения. Линия *Parahucho* присоединяется к последней группировке (рис. 1–3). Положение *Parahucho*, однако, неустойчиво, и в заметной доле псевдореплик бутстрэпа или байесовских циклов линия *Parahucho* составляла монофилетическую группировку с линией *Salvelinus*, существенно снижая таким образом уровень поддержки преимущественного объединения в одну кладу *Salvelinus*, *Parasalmo* и *Oncorhynchus*. В связи с этим стоит указать на то, что при анализе интронов *C* и *D* генов гормона роста 1 и 2 [13] или гена *RAG1* [9] положение линии *Parahucho* относительно *Salmo*, *Salvelinus*, *Parasalmo* и *Oncorhynchus*

установить однозначно также не удалось. Одной из причин этих затруднений может быть сравнительно узкий временной промежуток, в течение которого произошла их дивергенция (примерно 3 млн. лет согласно результатам как мтДНК — рис. 4, так и *RAG1*-датирования [9]) и который, вероятно, недостаточен для накопления необходимого количества согласующихся филогенетически информативных признаков.

В отношении долгое время дискутируемого вопроса о взаимоотношениях тихоокеанских лососей и форелей отметим следующее. Морфологически роды *Oncorhynchus* и *Parasalmo* хорошо отличаются, и известно по меньшей мере шесть уникальных производных признаков (синапоморфий), объединяющих виды тихоокеанских лососей в монофилетическую группу [1]. В соответствии с этим находятся результаты недавнего подробного анализа аллозимных данных, продемонстрировавшие монофилию рода *Oncorhynchus* [11]. Кроме того, роды *Oncorhynchus* и *Parasalmo* оказываются взаимно монофилетическими по интронам *C* и *D* гена гормона роста 1 [13] и митогеномам (рис. 1–3). По нашим расчетам (рис. 4), промежуток времени от момента дивергенции предков этих родов и до момента диверсификации линий внутри *Oncorhynchus* мал — всего лишь около 0.8 млн. лет. Поэтому случаи парафилии рода *Oncorhynchus*, выявленные при анализе других участков ядерной ДНК — интронов *C* и *D* гена гормона роста 2 [13], внутренних транскрибируемых спейсеров *ITS1* и *ITS2* генов рРНК [14], гена *RAG1* [9] — можно трактовать как результат сохранения предкового полиморфизма.

Вопросы о числе видов в роде *Salvelinus* и их филогенетических взаимоотношениях также долгое время являются предметом дискуссий. Виды этого рода можно разделить условно на две группы. В первую войдут виды, чей таксономический статус не вызывает сомнений, — *S. fontinalis*, *S. levanidovi*, *S. leucomaenis*, *S. confluentus* и *S. namaycush*. Во вторую — *S. alpinus* и *S. malma* и группирующиеся вокруг них виды гольцов (т.н. альпиноидные и мальмоидные гольцы, соответственно), видовой статус многих из которых выглядит спорным.

Результаты настоящей работы еще раз продемонстрировали, во-первых, что виды в последней группе эволюционно молоды (рис. 4); во-вторых, что митогеномные данные сами по себе не являются достаточными для выяснения филогенетических связей в роде *Salvelinus* ввиду обычности в этой группе рыб явления интрогрессивной гибридизации, сопровождающейся переносом мтДНК от одного вида к другому ([50] и мн. др.). К примеру, альпиноидные (*S. alpinus*, *S. taranetzi*)

и мальмоидные (*S. malma*, *S. curilus*) гольцы по мтДНК-данным ([51–53], рис. 1–3) взаимно парафилетичны. Однако по аллозимам [54, 55] или по гену *RAG1* [9] они составляют монофилетические группировки. Аналогично, как свидетельство интрогрессивной гибридизации можно рассматривать ситуацию с *S. confluentus* [50, 53, 56]. По митогеномным данным этот вид объединяется в одну кладу с *S. taranetzi* ([50, 57], рис. 1–3). Однако на деревьях, построенных по аллозимным данным [54], *ITS1* генов ядерной рРНК [14] или интронам *GH1C* и *GH2C* [58], *S. confluentus* является видом, сестринским к *S. leucomaenis*. В связи с этим филогенетическое положение монотипического *Salvethymus* — внутри группировки видов альпиноидных и мальмоидных гольцов (рис. 1–3) — по мтДНК-данным нельзя признать веско обоснованным. Тем не менее отметим, что по результатам PCR-RAPD-анализа *Salvethymus* также оказывается внутри данной группировки [59]. Кроме того, число хромосомных плеч (NF) в кариотипе *Salvethymus* равно 98 [60]. То есть оно такое же, как у рассмотренных в настоящей работе и других видов альпиноидных и мальмоидных гольцов. У *S. fontinalis*, *S. levanidovi*, *S. leucomaenis*, *S. confluentus* и *S. namaycush* NF = 100, и это состояние для гольцовых рыб считается исходным [60]. Поэтому положение *Salvethymus* на полученных филогенетических деревьях (рис. 1–3), вероятно, близко к истинному.

Полиплоидизация и диверсификация Salmonidae

Как известно, имеются все основания считать, что в линии Salmonidae, после ее формирования, но до ее разделения на линии хариусов, сига и лососей, произошло удвоение генома, путем, вероятно, автополиплоидизации [61]. Первая попытка датирования этой, специфической для Salmonidae, полногеномной дупликации (ЛСГД — лососево-специфическая геномная дупликация) основывалась на оценке эволюционного возраста паралогов гена *LdhB* (80–100 млн. лет) и допускала, что это событие могло по времени совпасть с собственно самим формированием линии лососевых рыб [62, 63]. Позже к аналогичному заключению пришли Осинов и Лебедев [64], исходя из полученной ими оценки эволюционного возраста паралогов гена *Vig* (75–85 млн. лет), оказавшихся примерно на 10 млн. лет моложе возраста дихотомии лососевых и шукообразных. Наши молекулярные датировки, основанные на анализе нескольких паралогических участков ядерного генома (рис. 4), находятся в согласии с перечисленными оценками и также допускают, что полиплоидизация могла произойти почти сразу же после обособления линии лососевых рыб.

Протяженность периода диверсификации линейного родового уровня у лососевых рыб (22 млн. лет — с 42 по 20 млн. лет назад) сходна с таковой щукообразных (30 млн. лет — с 88 по 58 млн. лет назад). Однако, если в линии *Esociformes* этот период начался спустя примерно 18 млн. лет после ее появления, то в линии *Salmonidae* он наступил много позже, с большим (около 64 млн. лет) временным лагом (рис. 4). Любопытно, что начало диверсификации рецентных *Salmonidae* совпадает с периодом, когда переход от тетра- к дисомическому наследованию осуществлялся, по-видимому, наиболее интенсивно. Медиана близкого к нормальному распределению оценок синонимических дистанций между 408 парами ЛСГД-онологов из генома *S. salar* равна 0.19, что примерно в два раза меньше таковой при их сравнении с ортологами из генома *Esox lucius* — 0.43 [65]. Из чего можно ожидать, что медиана распределения оценок времени с момента установления дисомии для этих генов находится где-то в районе оценки в 47 млн. лет назад ($0.19/0.43 \times 106$). К настоящему моменту через диплоидизацию прошло примерно 76% из 33 проанализированных изозимных локусов *P. mykiss* [61]. Тогда, беря эту оценку за ориентир, можно подсчитать, что к моменту начала диверсификации современных *Salmonidae* (около 42 млн. лет назад) доля таких локусов у их общего предка могла составлять около 40%. Может достижение именно такого уровня диплоидизации и явилось “спусковым крючком” процесса диверсификации рецентных *Salmonidae*?

На примере растений показано, что скорость видообразования в недавно сформированных полиплоидных линиях существенно ниже, чем в диплоидных [66], т.е. полиплоидия чаще всего является эволюционным тупиком. Однако те из полиплоидных линий, которые оказываются способными его миновать, впоследствии могут иметь долговременный эволюционный успех [66]. Лососевые рыбы, несмотря на указанную задержку, относятся, очевидно, к последней категории, поскольку в кроне филогенетического дерева *Salmonidae* чистая скорость диверсификации (net diversification rate) довольно высока (в целом 0.078 линий на млн. лет; расчет проведен с помощью пакета LASER [67] на основе полученной в настоящей работе байесовской хронограммы) и вполне сопоставима с таковой (0.067–0.081) в быстро эволюционирующих группах *Teleostei* [21]. Отметим, что при этом в последние 12 млн. лет ее темпы (0.108) были в два раза выше, чем в предшествующий (42–13 млн. лет назад) отрезок времени (0.054). В конечном итоге в настоящее время разнообразие *Salmonidae* оказывается много выше разнообразия сестринских к ним *Esoci-*

formes (66 видов 11 родов против 10 видов 4 родов, соответственно [18]).

Как показано при изучении явления полногеномной дупликации (ПГД) у дрожжей *Saccharomyces*, потеря (физически, через делецию или путем псевдогенизации) одного из онологов — обычный исход для подавляющего большинства (вплоть до 90%) дублицированных генов [68, 69]. При этом наиболее интенсивно такие процессы идут на протяжении первых нескольких десятков млн. лет спустя события ПГД [70, 71]. Сходная эволюционная динамика ПГД-онологов характерна и для костистых рыб (*Teleostei*) [72].

Считается, что дивергентное разрешение — потери разных членов ПГД-дубликатов в разных популяциях — может играть важную роль в процессе видообразования у полиплоидов [71, 73]. Предполагалось, что дивергентное разрешение способствовало диверсификации лососевых рыб [73]. На наш взгляд, в случае с *Salmonidae* его роль, скорее всего, была несущественная.

В сравнении с другими, произошедшими у костистых рыб, случаями полногеномных дупликаций скорость потерь онологов геномами лососевых рыб выглядит крайне низкой. За 70–100 млн. лет, прошедших с момента специфической для *Salmonidae* ПГД, в линиях, ведущих к *P. mykiss* или *S. salar*, произошла потеря одного из дубликатов в 36% (в 12 из 33) изозимных локусов [61], в 30% (в 23 из 74) других исследованных ядерных генов [5], а также в 22% (в 13 из 59) генов *Hox* кластеров [74, 75].

В линии *Clupeocephala* только в течение первых 50–75 млн. лет, начиная с момента специфической для *Teleostei* полногеномной дупликации (ТСГД) и до расхождения предков *Ostarioclupeomorpha* и *Euteleostei*, был утерян один из онологов в 61% (в 78 из 128) исследованных генов пигментации, в 79% (в 149 из 187) печеночных генов [76], в 61% (в 71 из 116) генов, вовлеченных в процессы передачи сигналов и обмена веществ [77], а также в 59–72% (в 26–31 из 43–44) генов *Hox* кластеров [75, 78, 79]. В целом потери в этот промежуток времени составили примерно от 65 до 90% всех потерь генов-дубликатов, образовавшихся в результате ТСГД.

В линии *Cyprinus* за то небольшое время, прошедшее с момента специфической для *Cyprinus* и *Carassius* ПГД — 11–16 млн. лет [80–82], — утеряна экспрессия одного из дубликатов в 48% из 23 проанализированных изозимных локусов [83], потерян дублицированное состояние в 40–52% исследованных микросателлитных локусов [81–84], а также утрачен один из четырех (25%) *Hox*-онологов [82].

Основной причиной чрезвычайно низкой скорости утраты ПГД-дубликатов в линии Salmonidae являются, по-видимому, низкие темпы восстановления у этих рыб дисомического наследования [85], что, в принципе, ожидаемо для автополиплоидов. Так или иначе, но, на наш взгляд, эта особенность делает маловероятной какую-либо значимую роль дивергентного разрешения в диверсификации современных Salmonidae.

Авторы глубоко признательны О.А. Радченко и J. Vohlen за любезно предоставленные образцы некоторых видов 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. Глубоковский М.К. Эволюционная биология лососевых рыб. М.: Наука, 1995. 343 с.
2. Evolution Illuminated. Salmon and their Relatives / Eds. Hendry A.P., Stearns S.C. N.Y.: Oxford Univ. Press, 2004. 510 p.
3. Volf J.-N. Genome evolution and biodiversity in teleost fish // Heredity. 2005. V. 94. № 3. P. 280–294.
4. 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.
5. 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.
6. 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.
7. 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.
8. Wang Y., Guo R., Li H. et al. The complete mitochondrial genome of the Sichuan taimen (*Hucho bleekeri*): repetitive sequences in the control region and phylogenetic implications for Salmonidae // Mar. Genomics. 2011. V. 4. № 3. P. 221–228.
9. Шедько С.В., Мирошниченко И.Л., Немкова Г.А. Филогения лососевых рыб (Salmoniformes: Salmonidae) и ее молекулярная датировка: анализ ядерного гена *RAG1* // Генетика. 2012. Т. 48. № 5. С. 676–680.
10. 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.
11. Osinov A.G., Lebedev V.S. Genetic divergence and phylogeny of the Salmoninae based on allozyme data // J. Fish Biol. 2000. V. 57. № 2. P. 354–381.
12. Шедько С.В. Филогения митохондриальной ДНК лососевых рыб подсемейства Salmoninae: анализ последовательностей гена цитохрома b // Генетика. 2002. Т. 38. № 3. С. 357–367.
13. 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.
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. Avise J.C. Molecular Markers, Natural History, and Evolution. MA: Sinauer Assoc. Inc., 2004. 684 p.
17. Berg W.J., Ferris S.D. Restriction endonuclease analysis of salmonid mitochondrial DNA // Can. J. Fish. Aquat. Sci. 1984. V. 41. P. 1041–1047.
18. Nelson J.S. Fishes of the World. New Jersey: John Wiley & Sons, Inc., 2006. 601 p.
19. Crow K.D., Wagner G.P. What is the role of genome duplication in the evolution of complexity and diversity? // Mol. Biol. Evol. 2006. V. 23. № 5. P. 725–732.
20. Van de Peer Y., Maere S., Meyer A. The evolutionary significance of ancient genome duplications // Nat. Rev. Genet. 2009. V. 10. № 10. P. 725–732.
21. Santini F., Harmon L.J., Carnevale G. et al. Did genome duplication drive the origin of teleosts? A comparative study of diversification in ray-finned fishes // BMC Evol. Biol. 2009. V. 9. P. 194.
22. Ward R.D., Zemlak T.S., Innes B.H. et al. DNA barcoding Australia's fish species // Phil. Trans. R. Soc. B. 2005. V. 360. № 1462. P. 1847–1857.
23. Sevilla R.G., Diez A., Norén M. et al. Primers and polymerase chain reaction conditions for DNA barcoding teleost fish based on the mitochondrial cytochrome b and nuclear rhodopsin genes // Mol. Ecol. Notes. 2007. V. 7. № 5. P. 730–734.
24. Brunner P.C., Douglas M.S., Osinov A.G. et al. Holarctic phylogeography of Arctic charr (*Salvelinus alpinus* complex) inferred from mitochondrial DNA sequences // Evolution. 2001. V. 55. № 3. P. 573–586.
25. Yamamoto S., Morita K., Kitano S. et al. Phylogeography of white-spotted charr (*Salvelinus leucomaenis*) inferred from mitochondrial DNA sequences // Zool. Sci. 2004. V. 21. № 2. P. 229–240.
26. Hillis D.M. Taxonomic sampling, phylogenetic accuracy, and investigator bias // Syst. Biol. 1998. V. 47. № 1. P. 3–8.
27. Wiens J.J., Tiu J. Highly incomplete taxa can rescue phylogenetic analyses from the negative impacts of limited taxon sampling // PLoS ONE. 2012. V. 7. № 8. e42925.

28. Roure B., Baurain D., Philippe H. Impact of missing data on phylogenies inferred from empirical phylogenomic data sets // *Mol. Biol. Evol.* 2013. V. 30. № 1. P. 197–214.
29. Katoh K., Misawa K., Kuma K., Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform // *Nucl. Acids Res.* 2002. V. 30. P. 3059–3066.
30. Swofford D.L. PAUP*: Phylogenetic Analysis Using Parsimony (*and other methods). Version 4.0b10. MA: Sinauer Assoc. Inc., 2002.
31. Posada D., Crandall K.A. MODELTEST: testing the model of DNA substitution // *Bioinformatics.* 1998. V. 14. № 9. P. 817–818.
32. Zwickl D.J. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion: Ph.D. dissertation. Austin: Univ. Texas, 2006. 115 p.
33. Ronquist F., Huelsenbeck J.P. MRBAYES 3: Bayesian phylogenetic inference under mixed models // *Bioinformatics.* 2003. V. 19. № 12. P. 1572–1574.
34. Drummond A.J., Rambaut A. BEAST: Bayesian Evolutionary Analysis by Sampling Trees // *BMC Evol. Biol.* 2007. V. 7. № 1. P. 214.
35. Wilson M.V.H., Brinkman D.B., Neuman A.G. Cretaceous Esocidae (Teleostei): early radiation of the pikes in North American fresh waters // *J. Paleontol.* 1992. V. 66. № 5. P. 839–846.
36. Wilson M.V.H. Osteology of the Palaeocene teleost *Esox tiemani* // *Palaeontology.* 1984. V. 27. № 3. P. 597–608.
37. Grande L. The first *Esox* (Esocidae: Teleostei) from the Eocene Green River Formation, and a brief review of esocid fishes // *J. Vert. Paleontol.* 1999. V. 19. № 2. P. 271–292.
38. 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.
39. 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.
40. Shimodaira H., Hasegawa M. CONSEL: for assessing the confidence of phylogenetic tree selection // *Bioinformatics.* 2001. V. 17. № 12. P. 1246–1247.
41. Moghadam H.K., Ferguson M.M., Danzmann R.G. Whole genome duplication: challenges and considerations associated with sequence orthology assignment in Salmoninae // *J. Fish Biol.* 2011. V. 79. P. 561–574.
42. Felsenstein J. *Inferring Phylogenies*. MA: Sinauer Assoc. Inc., 2004. 664 p.
43. Finn R.N., Kristoffersen B.A. Vertebrate vitellogenin gene duplication in relation to the “3R hypothesis”: correlation to the pelagic egg and the oceanic radiation of teleosts // *PloS ONE.* 2007. V. 2. № 1. e169.
44. Sanford C.P.J. The phylogenetic relationships of salmonoid fishes // *Bull. Br. Mus. Nat. Hist. (Zool.)*. 1990. V. 56. № 2. P. 145–153.
45. Sanford C.P.J. *Salmonoid Fish Osteology and Phylogeny (Teleostei: Salmonoidei)*. Theses Zoologicae 33. Liechtenstein: ARG Gantner Verlag KG, 2000. 264 p.
46. Stearley R.F., Smith G.R. Phylogeny of the Pacific trouts and salmon (*Oncorhynchus*) and genera of the family Salmonidae // *Trans. Am. Fish. Soc.* 1993. V. 122. № 1. P. 1–33.
47. Johnson G.D., Patterson C. Relationships of lower Euteleostean fishes // *Interrelationships of Fishes* / Eds Stiassny M.L.J., Parenti L., Johnson G.D. N.Y.: Acad. Press, 1996. P. 251–332.
48. Ives K.L., Taylor E.B. Molecular resolution of the systematics of a problematic group of fishes (Teleostei: Osmeridae) and evidence for morphological homoplasy // *Mol. Phylog. Evol.* 2009. V. 50. № 1. P. 163–178.
49. Шедько С.В. Филогенетические связи ленков рода *Brachymystax* (Salmonidae, Salmoniformes) и особенности их видообразования: Дис. ... канд. биол. наук. Владивосток: Биолого-почвенный институт ДВО РАН, 2003. 207 с.
50. Phillips R.B., Sajdak S.L., Domanico M.J. Relationships among charrs based on DNA sequences // *Nordic J. Freshw. Res.* 1995. V. 71. P. 378–391.
51. Радченко О.А. Изменчивость нуклеотидных последовательностей гена цитохрома b митохондриальной ДНК голецов рода *Salvelinus* // *Генетика.* 2004. Т. 40. № 3. С. 322–333.
52. Oleinik A.G., Skurikhina L.A., Brykov V.A. Divergence of *Salvelinus* species from northeastern Asia based on mitochondrial DNA // *Ecol. Freshw. Fish.* 2007. V. 16. № 1. P. 87–98.
53. Шедько С.В., Гинатулина Л.К., Мирошниченко И.Л., Немкова Г.А. Филогеография митохондриальной ДНК южной азиатской мальмы *Salvelinus curilus* Pallas, 1814 (Salmoniformes, Salmonidae): опосредованная интрогрессия генов? // *Генетика.* 2007. Т. 43. № 2. С. 227–239.
54. Crane P.A., Seeb L.W., Seeb J.E. Genetic relationships among *Salvelinus* species inferred from allozyme data // *Can. J. Fish. Aquat. Sci.* 1994. V. 51. Suppl. 1. P. 182–197.
55. Salmenkova E.A., Omelchenko V.T., Kolesnikov A.A., Malinina T.V. Genetic differentiation of charrs in the Russian North and Far East // *J. Fish Biol.* 2000. V. 57. Suppl. A. P. 136–157.
56. Phillips R.B., Oakley T.H. Phylogenetic relationships among the Salmoninae based on nuclear and mitochondrial DNA sequences // *Molecular Systematics of Fishes* / Eds Kocher T.D., Stepien C.A. San Diego: Acad. Press, 1997. P. 145–162.
57. Grewe P.M., Billington N., Hebert P.D.N. Phylogenetic relationships among members of *Salvelinus* inferred from mitochondrial DNA divergence // *Can. J. Fish. Aquat. Sci.* 1990. V. 47. P. 984–991.
58. Westrich K.M., Konkol N.R., Matsuoka M.P., Phillips R.B. Interspecific relationships among charrs

- based on phylogenetic analysis of nuclear growth hormone intron sequences // *Env. Biol. Fishes.* 2002. V. 64. P. 217–222.
59. Шубина Е.А., Пономарева Е.В., Гриценко О.Ф. Популяционно-генетическая структура гольцов Северных Курильских островов и положение мальмы в системе рода *Salvelinus* (Salmonidae: Teleostei) // *Журн. общей биологии.* 2006. Т. 67. № 4. С. 280–297.
 60. Фролов С.В. Изменчивость и эволюция кариотипов лососевых рыб. Владивосток: Дальнаука, 2000. 229 с.
 61. Allendorf F.W., Thorgaard G.H. Tetraploidy and the evolution of salmonid fishes // *Evolutionary Genetics of Fishes* / Ed. Turner B.J. N.Y.: Plenum Press, 1984. P. 1–53.
 62. Bailey G.S., Lim S.T. Gene duplication in salmonid fish: Evolution of a lactate dehydrogenase with an altered function // *Isozymes* / Ed. Markert C.L. San Francisco: Acad. Press, 1975. V. IV. P. 29–44.
 63. Lim S.T., Kay R.M., Bailey G.S. Lactate dehydrogenase isozymes of salmonid fish: Evidence for unique and rapid functional divergence of duplicate H4 Lactate dehydrogenases // *J. Biol. Chem.* 1975. V. 250. № 5. P. 1790–1800.
 64. Осинев 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.
 65. Leong J., Jantzen S. *Salmo salar* and *Esox lucius* full-length cDNA sequences reveal changes in evolutionary pressures on a post-tetraploidization genome // *BMC Genomics.* 2010. V. 11. P. 279.
 66. Mayrose I., Zhan S.H., Rothfels C.J. et al. Recently formed polyploid plants diversify at lower rates // *Science.* 2011. V. 333. № 6047. P. 1257.
 67. Rabosky D.L. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates from molecular phylogenies // *Evol. Bioinf.* 2006. V. 2. P. 273–276.
 68. Wolfe K.H., Shields D.C. Molecular evidence for an ancient duplication of the entire yeast genome // *Nature.* 1997. V. 387. № 6634. P. 708–713.
 69. Kellis M., Birren B.W., Lander E.S. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae* // *Nature.* 2004. V. 428. № 6983. P. 617–624.
 70. Byrne K.P., Blanc G. Computational analyses of ancient polyploidy // *Curr. Bioinform.* 2006. V. 1. № 2. P. 131–146.
 71. Scannell D.R., Byrne K.P., Gordon J.L. et al. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts // *Nature.* 2006. V. 440. № 7082. P. 341–345.
 72. Jaillon O., Aury J.-M., Brunet F. et al. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype // *Nature.* 2004. V. 431. № 7011. P. 946–957.
 73. Taylor J.S., Van de Peer Y., Meyer A. Genome duplication, divergent resolution and speciation // *Trends in Genetics.* 2001. V. 17. № 6. P. 299–301.
 74. Mungpakdee S., Seo H.-C., Angotzi A.R. et al. Differential evolution of the 13 Atlantic salmon *Hox* clusters // *Mol. Biol. Evol.* 2008. V. 25. № 7. P. 1333–1343.
 75. Henkel C.V., Burgerhout E., de Wijze D.L. et al. Primitive duplicate *Hox* clusters in the European Eel's genome // *PLoS ONE.* 2012. V. 7. № 2. e32231.
 76. Braasch I., Brunet F., Voff J.-N., Scharlt M. Pigmentation pathway evolution after whole-genome duplication in fish // *Genome Biol. Evol.* 2009. V. 1. P. 479–493.
 77. Sato Y., Hashiguchi Y., Nishida M. Temporal pattern of loss/persistence of duplicate genes involved in signal transduction and metabolic pathways after teleost-specific genome duplication // *BMC Evol. Biol.* 2009. V. 9. P. 127.
 78. Hoegg S., Boore J.L., Kuehl J.V., Meyer A. Comparative phylogenomic analyses of teleost fish *Hox* gene clusters: lessons from the cichlid fish *Astatotilapia burtoni* // *BMC Genomics.* 2007. V. 8. P. 317.
 79. Kuraku S., Meyer A. The evolution and maintenance of *Hox* gene clusters in vertebrates and the teleost-specific genome duplication // *Int. J. Dev. Biol.* 2009. V. 53. № 5–6. P. 765–773.
 80. Larhammar D., Risinger C. Molecular genetic aspects of tetraploidy in the common carp *Cyprinus carpio* // *Mol. Phylog. Evol.* 1994. V. 3. № 1. P. 59–68.
 81. David L., Blum S., Feldman M.W. et al. Recent duplication of the common carp (*Cyprinus carpio* L.) genome as revealed by analyses of microsatellite loci // *Mol. Biol. Evol.* 2003. V. 20. № 9. P. 1425–1434.
 82. Yuan J., He Z., Yuan X. et al. Speciation of polyploid Cyprinidae fish of common carp, crucian carp, and silver crucian carp derived from duplicated *Hox* genes // *J. Exp. Zool. Part B, Mol. Dev. Evol.* 2010. V. 314. № 6. P. 445–456.
 83. Ferris S.D., Whitt G.S. The evolution of duplicate gene expression in the carp (*Cyprinus carpio*) // *Experientia.* 1977. V. 33. № 10. P. 1299–1301.
 84. Zhang Y., Liang L., Jiang P. et al. Genome evolution trend of common carp (*Cyprinus carpio* L.) as revealed by the analysis of microsatellite loci in a gynogenetic family // *J. Genet. Genomics.* 2008. V. 35. № 2. P. 97–103.
 85. Li W.H. Rate of gene silencing at duplicate loci: a theoretical study and interpretation of data from tetraploid fishes // *Genetics.* 1980. V. 95. № 1. P. 237–258.