Molecular Phylogenetics and Evolution 47 (2008) 1061–1075

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Molecular phylogeny of the genus *Gobio* Cuvier, 1816 (Teleostei: Cyprinidae) and its contribution to taxonomy

Jan Mendel [a,*], Stanislav Lusk [a], Ekaterina D. Vasil'eva [b], Victor P. Vasil'ev [c], Věra Lusková [a], F. Guler Ekmekci [d], Füsun Erk'akan [d], Alexander Ruchin [e], Ján Koščo [f], Lukáš Vetešník [a], Karel Halačka [a], Radek Šanda [g], Andrei N. Pashkov [h], Sergei I. Reshetnikov [h]

[a] *Institute of Vertebrate Biology of the Academy of Sciences of the Czech Republic, v.v.i., Department of Ichthyology, Květná 8, 603 65 Brno, Czech Republic*
[b] *Zoological Museum, Bolshaya Nikitskaya 6, 125009 Moscow, Russia*
[c] *Institute of Problems of Ecology and Evolution RAS, Leninsky Prospect 33, 117071 Moscow, Russia*
[d] *Department of Biology, Hacettepe University, Beytepe Campus, 06532 Ankara, Turkey*
[e] *Mordovian State University, Bolshevitskaya 68, 430000 Saransk, Russia*
[f] *Department of Ecology, University of Prešov, 17. Novembra 1, 081 16 Prešov, Slovakia*
[g] *National Museum, Department of Zoology, Václavské náměstí 68, 115 79 Praha, Czech Republic*
[h] *Kuban State University, Stavropolskaya 149, 350040 Krasnodar, Russia*

ARTICLE INFO

Article history:
Received 7 September 2007
Revised 29 February 2008
Accepted 4 March 2008
Available online 14 March 2008

Keywords:
Control region
S7 ribosomal protein gene
Intron
Cytonuclear disequilibrium
S7indel diagnostics
Taxonomy
New species
Phylogeography
Gobio

ABSTRACT

The phylogenetic relationships among gudgeons that represent most nominal taxa within *Gobio gobio* sensu lato were examined by mitochondrial and nuclear genome sequencing. The molecular analyses confirmed the separate generic status of *Gobio* as a monophyletic group and revealed 15 Eurasian lineages divided into two main clades, the Northern European and the Ponto-Caspian. The validity of eleven nominal taxa as distinct species was confirmed, gudgeons from the Volga River basin were described as a new species *G. volgensis*, and three revealed phylogenetic lineages were submitted for a comprehensive revision as "species-in-waiting". The species *G. gobio* showed a wide range extending from the British Isles to the Black Sea coast and overlapped the areas of several other species. Four pure lineages were detected in the middle Danube River basin. The Crimean Peninsula was found to be a region with the occurrence of individuals of hybrid origin. This region will require special investigation to define species participating in hybridization events, and to establish further steps for the conservation of endemic native gudgeon species. A simple diagnostic method, based on different lengths of the PCR products, called "S7indel diagnostics" is presented for further taxonomic investigations in the genus *Gobio*.

1. Introduction

The genus *Gobio* Cuvier, 1816 belongs to the subfamily Gobioninae, which is part of the large family Cyprinidae. Its distribution reaches from Spain and the British Isles to the far East and Northern China, and its representatives live in all types of waters, i.e., in standing and flowing waters, in freshwaters and in some cases found in brackish waters. The Palaearctic gudgeon *G. gobio* (Linnaeus, 1758) sensu lato is a complicated species especially in terms of taxonomy, due to its exceptional phenotypical diversity, and is therefore considered one of the most variable fish species in Europe (Bănărescu et al., 1999). *G. gobio* sensu lato includes many subspecies and local forms described in the past, whose validity is under extensive discussion now. For example, Naseka et al. (2006) argue that most of the designations attached to these fish do not have a real basis, as they apparently arose as artifacts due to the combination of inadequate material, discrepancies in the use of different species concepts, language barriers and an insufficient attention to detailed morphological studies (analogous to Kottelat, 1997; Kottelat and Persat, 2005). Moreover, we have noticed several descriptions of newer species of the genus *Gobio* from different geograghical areas which were derived from incomplete information, being based solely on morphological data (Freyhof and Naseka, 2005; Kottelat and Persat, 2005; Naseka et al., 2006). On the other hand, several genetic investigations show signs of a more comprehensive approach to gudgeon phylogeny and taxonomy (Doadrio and Madeira, 2004 along with Madeira et al., 2005; Bianco and Ketmaier, 2005).

This study is an attempt to adopt a comprehensive approach as well, as it is based on morphologically defined specimens of several nominal *Gobio* taxa from type localities or their close surroundings,

* Corresponding author. Fax: +420 543 211 346.
E-mail address: jmendel@seznam.cz (J. Mendel).

1055-7903/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2008.03.005

which were subjected to analysis by molecular markers of both mitochondrial and nuclear genomes.

The main aim of this study was to evaluate the validity of particular species and subspecies of the genus *Gobio* and to estimate phylogenetic relations between them, and thus gain a clear view of the taxonomy of gudgeons, to identify "species-in-waiting," and to propose that several previously described species/subspecies should be subject to a comprehensive revision. Furthermore, we wanted to introduce new information in the genetic field of this genus, including diagnostic markers.

2. Materials and methods

2.1. Sample collection

In the period from 2000 to 2006, 139 gudgeon specimens were collected from 43 localities (see Table 1), which represented areas of the most nominal taxa included in *G. gobio* s. lato (Fig. 1). The species *Rhodeus ocellatus* and *Sarcocheilichthys microoculus* were selected as outgroups (Table 1) based on recent knowledge of phylogenetic relationships among cyprinid fishes (Cunha et al., 2002; Yang et al., 2006). Two species of the genus *Romanogobio* (*R. albipinnatus* and *R. frici*) were also used for comparison. Voucher specimens are deposited in the collections of the Department of Ichthyology of the Institute of Vertebrate Biology, v.v.i. (Brno, Czech Republic) and Zoological Museum of the Moscow State University (ZMMU).

2.2. DNA extraction, PCR amplification and sequencing

Total genomic DNA was extracted from a small piece of the pectoral fin by proteinase K digestion followed by phenol–chloroform–isoamylalcohol purification and ethanol precipitation (Sambrook et al., 1989). Sequences of the control region (CR), and the first intron of the S7 r-protein (S7) were amplified by polymerase chain reaction (PCR) with primers specified in Table 2. PCRs were performed in 50 μl volume containing 10 mM Tris–HCl, 50 mM KCl, 0.1% Triton X-100, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 μM each primer, 2.5 U *Taq* DNA polymerase (TopBio) and approximately 100–500 ng of genomic DNA. Reactions were performed in TGRADIENT Thermocycler (Whatman Biometra) under the following conditions: CR: 95 °C for 1 min, followed by 37 cycles of 94 °C for 45 s, annealing at 52.6 °C (the first fragment) and 54.8 °C (the second fragment) for 30 s, and an extension temperature of 72 °C for 45 s, followed by a final extension at 72 °C for 5 min. In some cases, other pairs of primers (STIR_CR) were used under the following conditions: 95 °C for 3 min, followed by 34 cycles of 95 °C for 30 s, annealing at 55.0 °C for 30 s, with an extension temperature of 72 °C for 1 min, and a final extension at 72 °C for 5 min. S7: 95 °C for 1 min, followed by 30 cycles of 94 °C for 45 s, annealing at 52.4 °C for 30 s, and an extension temperature of 72 °C for 25 s, with a final extension at 72 °C for 5 min. The PCR products were visualized by mini-gel electrophoresis using ethidium bromide staining and 1.7% agarose gels. The PCR products were purified by means of precipitation PEG/Mg/NaAc (26% Polyethylene glycol, 6.5 mM $MgCl_2{\cdot}6H_2O$, 0.6 M·NaAc·$3H_2O$). Direct sequencing of purified PCR products was performed with the BigDye™ Terminator Cycle Sequencing Ready Reaction Kit version 1.1 (Applied Biosystems) according to the manufacturer's instructions, and purified with EtOH/EDTA precipitation. The sequencing was performed on an ABI PRISM 310 Genetic Analyzer (Applied Biosystems). All PCR amplicons were multiple sequenced from both directions to ensure high quality reads. The DNA sequences were edited and aligned using the Seqman module within Lasergene v. 6.0 (DNASTAR Inc.) and also checked manually. The accuracy of the sequence was confirmed by comparison with the NCBI database.

2.3. Phylogenetic analyses

The web-based ModelTest 3.8 program was used to ascertain the best-fit model of nucleotide substitution for separate nuclear and mitochondrial regions (Posada, 2006). Phylogenetic relationships among the two gene sequences were examined using the neighbour-joining (NJ) algorithm, the criteria of optimality: maximum parsimony (MP) and maximum-likelihood (ML), as well as using Bayesian inference (BI). The sequences were imported into PAUP* 4.0B.10 (Swofford, 2002) and MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2005) for phylogenetic analysis. For NJ analysis, DNA distances were calculated. Non-parametric bootstrap analyses with 1000 pseudo-replicates were performed to obtain supporting estimates for each node of the NJ trees. For MP tree construction, unweighted parsimony analysis using a branch-and-bound search was used. The confidence levels in the resulting relationship were assessed using the bootstrap procedure with 1000 replications. ML search was performed under the best-fit model with the branch-and-bound algorithm on 100 bootstrap replicates. Bayesian analysis was performed using MrBayes 3.1.2. Starting from a random tree, four Markov chains were run for 1×10^6 generations with a sampling frequency of 100. The best-fit models were then specified. The combined data set was treated as two partitions with different models accounting for their heterogeneity. We utilized the "unlink" command in MrBayes 3.1.2 to unlink the following parameters: "unlink shape = (all) statefreq = (all) revmat = (all)". The application Tracer 1.2 (Rambaut and Drummond, 2003) was used to view the output of the *sump* file generated by MrBayes. The trees generated prior to reaching stationarity were discarded as burn-in. We then took the resulting 50% majority rule consensus tree. Congruence among tree topologies generated for the combined data (CR and S7 sequences) was tested with the incongruence length difference test (ILD) as implemented in the partition homogeneity test in PAUP* (Farris et al., 1995).

As for treating the gaps as phylogenetic characters, three types of analyses were compared during the process of phylogenetic inference from the sequence of nuclear marker (S7): (1) gaps as missing data, (2) gaps as the fifth state character (Barriel, 1994) and (3) gaps as a separate binary character (Simmons and Ochoterena, 2000). The best elaboration for dealing with indels (insertion/deletion) together with their incorporation to phylogenetic analyses was described by Simmons and Ochoterena (2000). They determined that there were two ways of coding gaps: (a) the procedure "simple indel coding" (SIC), and (b) the procedure "complex indel coding" (CIC). Müller (2006) described a third approach of dealing with indels: "modified complex indel coding" (MCIC). The coding of indels was provided by the SeqState program (Müller, 2005) containing the implemented program IndelCoder. All of these approaches and methods were applied in this study. Instead of the CIC procedure, we used its newer, modified version, MCIC.

Only known nuclear haplotypes were included in the S7 analyses; heterozygous genotypes were excluded.

Haplotype and nucleotype networks were constructed to estimate the genealogical intraspecific relationships employing the statistical parsimony (Templeton et al., 1992) implemented into the TCS 1.21 program (Clement et al., 2000). Indels were coded as the fifth state characters. A 95% connection limit was then calculated, meaning that the haplotypes were disconnected when more than ten mutational steps divided them.

Table 1
Species examined in this study, source of tissue samples used, number of haplotype/nucleotype and GenBank Accession numbers

Lineage/species	[Locality Nos.] River, drainage, country (CRH/N, S7N/N)	Accession Nos.
Outgroups		
R. albipinnatus	R. Moksha in Mordovia, Volga, RUS	EF427390, EF427408
R. frici	R. Hornád at Košice, Tisza, SK	EF427392, EF427414
R. ocellatus[a]	Unknown	AY017149, AY325789
S. microoculus[a]	Unknown	NC_004694
Ingroup taxa		
Lineage_I *G. gobio* s. stricto	[1] R. Bečva at Rybáře, Danube, CZ (H1/1, H2/2); [2] R. Blanice at Vlašim and Vodňany, Elbe, CZ (H1/5); [3] R. D. Orlice at Kostelec n. Orlicí, Elbe, CZ (H1/3, H4/1, H5/1); [4] R. Elbe at Neratovice and Srnojedy, Elbe, CZ (H1/3, N1/1, N2/2, N3/1, N4/1); [5] R. Haná at Vyškov, Danube, CZ (H1/2); [6] R. Lahn, Rhine, D (H1/3, H6/2); [7] R. Lark near Isleham, Great Ouse, UK (H1/2, N2/1); [8] R. Odra at Odry, Odra, CZ (H1/1,H2/1, H3/2); [9] Reviština channel, Tisza, SK (H1/2, N2/3); [10] Rhone basin, FR[a] (H1/1); [11] R. S. Bug, UA (H1/3, H7/1, H8/1); [12] R. Stěnava at Broumov, Odra, CZ (H1/3)	EU131542, EU131543, EU131544, EU131545, EU131546, EU131547, EU131548, EU131550, EU131589, EU131590, EU131591, EU131592
Lineage_II *G. obtusirostris*	[1] R. Bečva at Rybáře, Danube, CZ (H13/1, H14/1); [13] R. Bystrička at Martin, Danube, SK (H10/2, H11/1, H13/2); [14] R. Dyje at Soutok, Danube, CZ (H10/2, N5/2); [5] R. Haná at Vyškov, Danube, CZ (H10/1, H13/2); [15] R. Ipel' at st. border, Danube, SK (H10/4, H12/1); [16] R. Jevišovka at Božice, Danube, CZ (H10/1); [8] R. Odra at Odry, Odra, CZ (H13/1); [17] R. Váh at Hlohovec, Danube, SK (H13/5, N5/2)	EU131554, EU131555, EU131556, EU131557, EU131558, EU131607
Lineage_III *Gobio* sp. 1	[18] Belžan Stream, Tisza, SK (H15/3, H18/2); [19] R. Laborec at Kochánovce, Tisza, SK (H16/1); [9] Reviština channel, Tisza, SK (H17/1); [20] R. Topl'a at Poliakovce, Tisza, SK (H16/5, N6/1)	EU131563, EU131564, EU131565, EU131562, EU131603
Lineage_IV *G. carpathicus*	[14] R. Dyje at Soutok, Danube, CZ (H19/1); [19] R. Laborec at Kochánovce, Tisza, SK (H19/3, N8/1); [9] Reviština channel, Tisza, SK (H19/1, H20/1); [21] R. Tereshva at Krive, Tisza, UA (H21/1, N7/1); [22] R. Ublianka at Ubl'a, Tisza, SK (H22/1, N7/1, N9/1)	EU131552, EU131559, EU131560, EU131561, EU131604, EU131605, EU131606
Lineage_V *Gobio* sp. 2	[24] Bakacak deresi - Biga, TR (H23/1, N10/1, N11/2)	EU131551, EU131593, EU131594
Lineage_VI *G. volgensis*	[25] R. Bol'shaya Lašva, Perm District, Volga, RUS (N12/1, N13/1); [26] R.Chardym at Aryash, Volga, RUS (N12/2); [27] R. Malaya Tsivil' at Shichazany, Volga, RUS (N12/1); [28] R. Moskva at Zvenigorod, Volga, RUS (H24/1, N12/1); [29] R. Sura at Nikolaevka, Volga, RUS (N12/3)	EU131566, EU131613, EU131614
Lineage_VII *G. caucasicus*	[30] R. Ashe, Sochi region, Krasnodar district, RUS (H27/1, N14/1, N17/1); [31] R. Mzymta at Adler, Krasnodar district, RUS (H25/1, H26/1, N15/1); [32] R. Shakhe, Sochi region, Krasnodar district, RUS (H28/1, N14/2, N16/1)	EU131584, EU131585, EU131586, EU131587, EU131615, EU131616, EU131617, EU131618
Lineage_VIII *G. insuyanus*	[33] Ayranci Dam Lake at Karaman, TR (H33/1, H34/1, H35/1, N20/2, N21/1); [34] R. Insuyu at Cihanbeyli, Tuz Lake, TR (H29/2, H30/1, H31/1, H32/1, N21/5); [35] R. Sugla at Seydisheir, TR (N20/1, N22/1)	EU131574, EU131575, EU131576, EU131577, EU131578, EU131579, EU131580, EU131621, EU131622, EU131623
Lineage_IX *Gobio* sp. 3	[36] Bilecik, TR (H36/1, N23/1, N24/1)	EU131581, EU131624, EU131625
Lineage X *G. ohridanus*	[37] R. Mat at Milot, AL (H39/1); [38] Lake Ohrid, Drin, AL (H37/1, H38/1, H40/1, N25/3)	EU131570, EU131571, EU131572, EU131573, EU131626
Lineage_XI *G. cynocephalus*	[39] R. Zeya at Blagoveshchensk, Amur, RUS (H41/1, N26/1)	EU131582, EU131608
Lineage_XIV *G. skadarensis*	[37] R. Mat at Milot, AL (H45/1, H47/1); [42] R. Zeta at Danilovgrad, Morača, MN (H45/4, H46/1, N30/2, N31/2)	EU131567, EU131568, EU131569, EU131601, EU131602
Ingroup taxa—lineages discriminated by the nuclear marker S7		
Lineage_XII *G. tauricus*	[40] R. Chernaya at Sevastopol, Crimean Peninsula, UA, 4549 (N27/1)	EU131609
Lineage_XIII *G. brevicirris*	[41] R. Sosna at Elets, Don, RUS, 4568 (N28/1), 4570 (N29/1)	EU131611, EU131612
Lineage_XV *G. krymensis*	[43] R. Bel'bek at Lyubimovka village, Crimean Peninsula, UA, 4607 (N32/1)	EU131620
Ingroup taxon—unambiguous hybrid (mixed lineages; mtDNA × nDNA)		
Hybrid L_I × L_XV	[43] R. Bel'bek at Lyubimovka village, Crimean Peninsula, UA, 4607 (H9/1, N32/1)	EU131549, EU131620
Ingroup taxa—submitted specimens for further investigation[b]		
Pure L_XII or Hybrid L_XIII × L_XII	[40] R. Chernaya at Sevastopol, Crimean Peninsula, UA, 4549 (H42/1, N27/1)	EU131583, EU131609
Pure L_? or Hybrid_?	[40] R. Chernaya at Sevastopol, Crimean Peninsula, UA; 4550 (H43/1, N19/1)	EU131553, EU131610
Hybrid L_I × L_?	[43] R. Bel'bek at Lyubimovka village, Crimean Peninsula, UA; 4605 (H1/1, N18/1)	EU131542, EU131619
Pure L_XIII or Hybrid L_XII × L_XIII	[41] R. Sosna at Elets, Don, RUS, 4570 (H44/1, N29/1)	EU131588, EU131612

[a] Indicates sequences from the GenBank; CRH/N = control region haplotype/number of analyzed specimens, S7N/N = S7 nucleotype/number of analyzed specimens.
[b] Explained in the text; e.g. 4549 = identification number; AL, Albania; CZ, Czech Republic; D, Germany; MN, Montenegro; RUS, Russia; SK, Slovakia; TR, Turkey; UA, Ukraine; UK, United Kingdom.

3. Results and discussion

With regard to the extent of the submitted study, which gives the taxonomic and systematic overview of more than two thirds of the valid representatives of the genus *Gobio*, and in an effort to provide a clearer perspective while saving space as well, we have reorganised this part of the paper by connecting our results with the discussion.

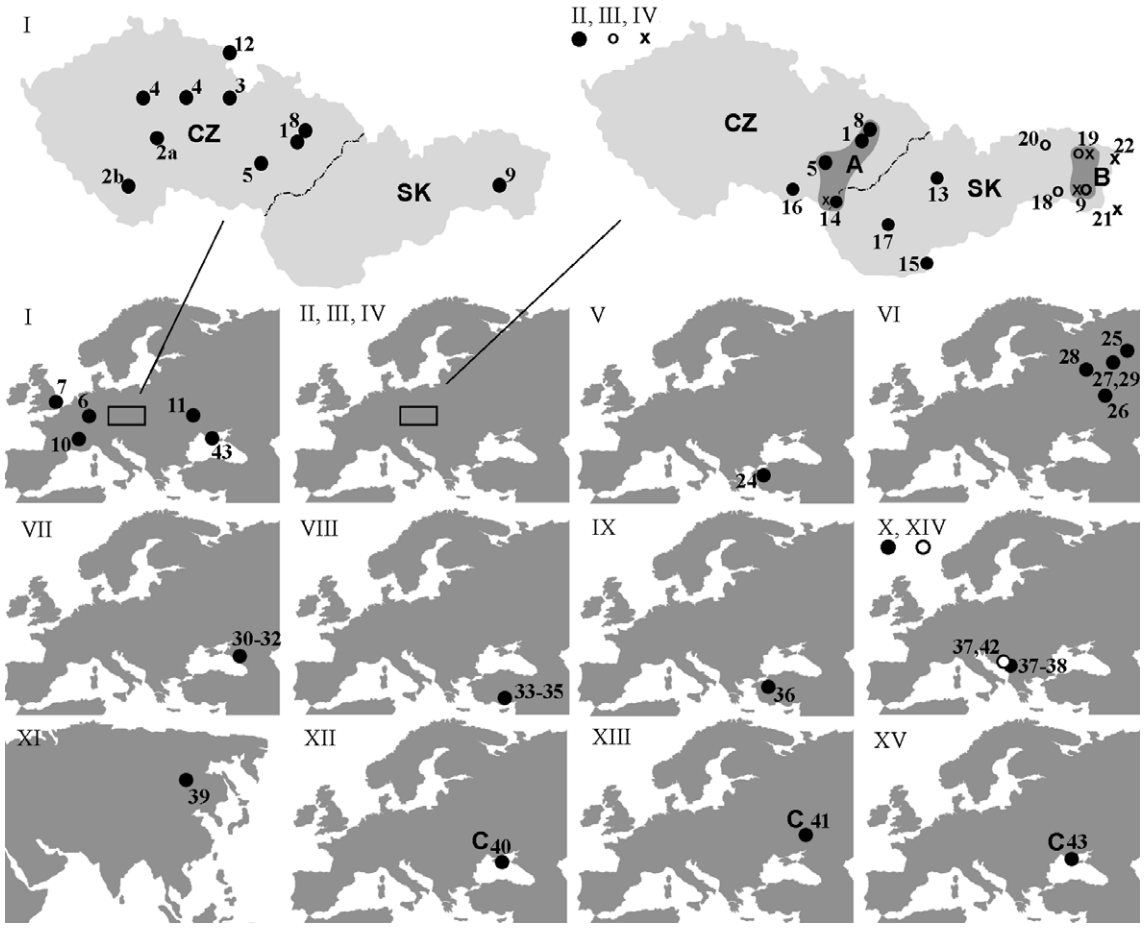

Fig. 1. Geographical origins of the fifteen lineages of the genus *Gobio*. Locality numbers correspond to the locality numbers in Table 1. In cases of the existence of more lineages, they are designated by various symbols. The rectangle demarcates the areas of large concentrations of collecting localities and these are displayed in the larger map. CZ = Czech Republic, SK = Slovakia. The dotted line indicates the border between countries. A–C = localities with occurrence of more species or possibly hybrids.

Table 2
List of primers used in this study

Gene	Primer	Sequences (5′–3′)	Reference
CR	CR159	CCC AAA GCA AGT ACT AAC GTC	This study
	CR439	AAC TGT TTT TCC CAC ACT TA	This study
	CR493	TTG GGT AAC GAG GAG TAT GTA	This study
	CR851	TGC GAT GGC TAA CTC ATA C	This study
STIR_CR	Carp-Pro	AAC TCT CAC CCC TGG CTA CCA AAG	Thai et al. (2004)
STIR_CR	Carp-Phe	CTA GGA CTC ATC TTA GCA TCT TCA GTG	Thai et al. (2004)
S7	S7univL	ACA ATT GTA AGT CGG AGA TG	This study
	S7univP	CCC ACA AAA TAA GAT ATT AGG	This study

3.1. Sequence characteristics

Sequence data was deposited in the GenBank database under Accession Nos. (CR: EU131542–EU131588; S7: EU131589–EU131594 and EU131601–EU131626; Table 1). Up to 1097 bp of analyzable sequence data was obtained from the nuclear and mitochondrial genome fragments for each specimen. Fragments of the control region (713 bp alignment) and the first intron of the S7 r-protein (384 bp alignment) were analyzed both separately and in combination. As some sequences from the same taxa were identical, these taxa were reduced to one representative per taxon in all subsequent analyses, excluding construction of the mtDNA and nDNA networks. Nucleotide base composition showed high values of AT pairs (A = 33.4% and T = 31.6% across all sites/all taxa) in the control region sequences. A similar composition has been noted for Cypriniformes (Liu et al., 2002). In the case of the first intron sequences of the S7 gene, low values of GC pairs were found (16.7% and 19.3%, respectively) which is typical for noncoding regions of the genome.

3.2. Haplotype and nucleotype richness—haplotype and nucleotype networks

Within the haplotype network, a total of 47 CR haplotypes representing 116 sequences were detected. Sixty-one sequences were used for the construction of the nucleotype network and altogether 32 S7 nucleotypes were identified (Table 3). A detailed list of all studied taxa, their haplotype and nucleotype assignment, sampling localities, haplotype and nucleotype frequency, and the GenBank accession numbers are given in Table 1. The schematic diagram of the statistical parsimony network shows a complex pattern of relationships within the genus *Gobio*. CR haplotype network revealed five disconnected groups consisting of 13 lineages (Fig. 2), and the S7 nucleotype network revealed eleven disconnected groups consisting of 15 lineages, see the supplementary material (Supplementary Fig. SM_1).

3.3. Phylogenetic analyses

For both NJ and Bayesian analyses, the best-fit model under Akaike information criterion, (AIC; Akaike, 1974) was determined using the software ModelTest 3.8, see Table 4. The levels of diver-

Table 3
Summary statistics of the gudgeon lineages within the CR and S7 networks and S7indel diagnostics

Lineage	$N_{cr/s7}$	$NHN_{cr/s7}$	PCR product (bp)
Lineage_I	42/9	9/4	338
Lineage_II	24/4	5/1	364
Lineage_III	12/1	4/1	366
Lineage_IV	8/4	4/3	366
Lineage_V	1/2	1/2	364
Lineage_VI	1/9	1/2	360
Lineage_VII	4/6	4/4	359
Lineage_VIII	8/10	7/3	367
Lineage_IX	1/2	1/2	369
Lineage_X	4/3	4/1	366
Lineage_XI	1/1	1/1	371
Lineage_XII	2[a]/1	2[a]/1	366
Lineage_XIII	2[a]/2	2[a]/2	337
Lineage_XIV	7/4	3/2	366
Lineage_XV	[a]/1	[a]/1	359
Lineage_?	1/2	1/2	
Overall	116/61	47/32	

N, the number of specimens; NHN, the number of CRhaplotypes/S7nucleotypes.
[a] Explained in the text.

Table 4
Analyzed fragments of both genomes, their characteristics resulting from the MP analysis and the appropriate models selected by Modeltest

Partition	No. characters (pars. inf.)	TL	CI	RI	Model
CR	713 (80)	375	0.5551	0.7183	HKY + Γ
S7	384 (62) + 15 gaps	276	0.8840	0.9516	K81uf
All combined data	1097 (142) + 15 gaps	819	0.7082	0.8446	Mixed model

CI, consistency index (excluding uninformative characters); Γ, gamma; pars. inf., number of parsimony informative characters; RI, retention index; TL, tree length.

gence within and among *Gobio* lineages are shown in Table 5. The summary statistics of MP analyses for the separate and combined data sets for each gene are shown in Table 4. For ML analyses, the likelihood settings of the best-fit model for CR based on the hierarchical likelihood ratio tests (hLRTs) were as follows: base frequencies (A = 0.3337, C = 0.2132, G = 0.1372 and T = 0.3158); ti/tv ratio = 1.6559; and the shape parameter of the gamma distribution 0.3362. Likelihood settings of the best-fit model for S7 were: base frequencies (A = 0.2968, C = 0.1931, G = 0.1667 and T = 0.3433).

The three most common approaches of treating gaps in multiple sequence alignment S7 within a parsimony framework were used in this study. In the case of gap coding as separate characters, the results of two indel coding methods have been compared. The SIC and MCIC methods provided the same results in terms of topological accuracy and similar values of bootstrap support (data not shown).

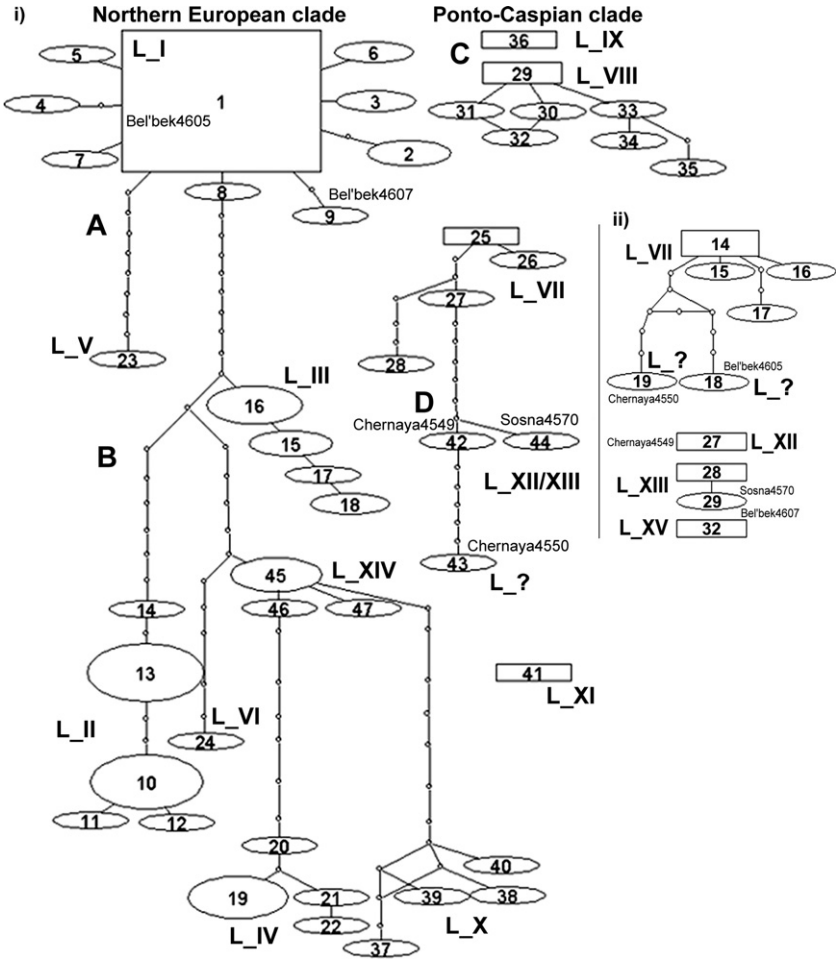

Fig. 2. The unrooted haplotype network based on sequences of the control region of certain representatives of the genus *Gobio* (i). The haplotype numbers refer to the numbers in Table 1. The node sizes are proportional to the haplotype frequency (see Table 1). The pertinence of individual haplotypes to the lineages (e.g. L_I) and subclades (A–D) is marked. Marked individuals of certain lineages (e.g. Bel'bek4607) show the hybrid origin. The names in the network indicate the collection locality followed by the identification number of the individual. The complete S7 nucleotype network is shown as supplementary material. Its partial version (ii) shows the differences between both networks.

Table 5
Mutual comparison of the representatives of the genus *Gobio* and their sequence divergences obtained by analysis of both mitochondrial (CR) and nuclear (S7) genomes

	Lineage I	Lineage II	Lineage III	Lineage IV	Lineage V	Lineage VI	Lineage VII	Lineage VIII	Lineage IX	Lineage X	Lineage XI	Lineage XII	Lineage XIII	Lineage XIV	Lineage XV
Lineage_I	**0.14 ± 0.00**	2.94 ± 0.60	1.62 ± 0.41	2.83 ± 0.58	1.34 ± 0.38	1.93 ± 0.47	6.06 ± 0.75	4.88 ± 0.70	8.30 ± 0.91	2.98 ± 0.60	4.93 ± 0.74	5.79 ± 0.75	5.33 ± 0.71	1.85 ± 0.46	–
Lineage_II	5.08 ± 0.85	**0.34 ± 0.09**	1.87 ± 0.45	2.09 ± 0.55	2.47 ± 0.54	2.51 ± 0.55	6.69 ± 0.79	5.33 ± 0.75	9.53 ± 0.89	2.74 ± 0.58	5.57 ± 0.77	6.50 ± 0.79	6.01 ± 0.78	2.26 ± 0.55	–
Lineage_III	5.91 ± 0.95	2.72 ± 0.65	**0.20 ± 0.07**	1.69 ± 0.47	2.08 ± 0.52	1.78 ± 0.45	6.22 ± 0.77	4.17 ± 0.66	8.12 ± 0.90	2.33 ± 0.55	4.76 ± 0.70	5.63 ± 0.75	5.17 ± 0.69	1.24 ± 0.38	–
Lineage_IV	5.05 ± 0.85	1.07 ± 0.39	2.70 ± 0.65	**0.23 ± 0.08**	2.08 ± 0.47	1.71 ± 0.45	6.10 ± 0.76	4.95 ± 0.71	9.52 ± 0.89	2.57 ± 0.57	4.84 ± 0.75	5.91 ± 0.78	5.44 ± 0.75	1.47 ± 0.43	–
Lineage_V	1.70 ± 0.45	5.37 ± 0.91	5.91 ± 0.95	5.34 ± 0.91	–	2.56 ± 0.58	4.70 ± 0.68	5.04 ± 0.69	9.62 ± 0.89	2.77 ± 0.58	5.28 ± 0.76	4.87 ± 0.75	4.42 ± 0.68	2.00 ± 0.47	–
Lineage_VI	3.72 ± 0.71	5.98 ± 0.95	6.77 ± 1.01	6.22 ± 0.97	3.00 ± 0.69	–	6.61 ± 0.80	4.70 ± 0.68	9.00 ± 0.90	2.18 ± 0.53	4.22 ± 0.66	6.01 ± 0.75	5.53 ± 0.75	0.94 ± 0.35	–
Lineage_VII	1.42 ± 0.45	5.22 ± 0.90	6.34 ± 0.98	5.19 ± 0.86	1.48 ± 0.46	3.41 ± 0.70	**0.55 ± 0.17**	6.47 ± 0.78	9.70 ± 0.92	5.90 ± 0.78	8.38 ± 0.88	1.63 ± 0.47	1.25 ± 0.37	6.11 ± 0.76	–
Lineage_VIII	7.18 ± 1.12	2.96 ± 0.68	4.08 ± 0.80	2.95 ± 0.68	6.76 ± 1.01	7.68 ± 1.19	7.21 ± 1.12	**0.28 ± 0.08**	4.72 ± 0.68	3.57 ± 0.59	6.33 ± 0.78	5.90 ± 0.78	5.42 ± 0.75	3.89 ± 0.66	–
Lineage_IX	7.00 ± 1.10	2.81 ± 0.66	3.91 ± 0.79	2.80 ± 0.66	6.57 ± 1.00	7.49 ± 1.16	7.03 ± 1.10	1.99 ± 0.50	–	8.07 ± 0.89	10.27 ± 0.93	9.86 ± 0.95	9.39 ± 0.88	8.13 ± 0.89	–
Lineage_X	5.33 ± 0.91	1.34 ± 0.43	2.43 ± 0.60	1.34 ± 0.43	5.07 ± 0.85	5.95 ± 0.95	5.50 ± 0.92	2.68 ± 0.65	2.54 ± 0.62	**0.48 ± 0.18**	5.19 ± 0.75	5.25 ± 0.75	4.87 ± 0.71	1.31 ± 0.38	–
Lineage_XI	3.75 ± 0.71	4.34 ± 0.82	5.17 ± 0.87	4.31 ± 0.82	3.83 ± 0.79	4.68 ± 0.83	3.70 ± 0.71	6.12 ± 0.96	6.08 ± 0.96	4.60 ± 0.83	–	7.95 ± 0.85	7.44 ± 0.84	4.13 ± 0.66	–
Lineage_XII	1.79 ± 0.46	5.51 ± 0.92	6.63 ± 1.00	5.48 ± 0.92	1.76 ± 0.46	3.70 ± 0.71	1.62 ± 0.45	7.52 ± 1.16	7.32 ± 1.14	5.80 ± 0.94	3.98 ± 0.80	**0.81 ± 0.28**	0.73 ± 0.18	5.53 ± 0.76	–
Lineage_XIII	1.68 ± 0.45	5.56 ± 0.92	6.46 ± 0.99	5.52 ± 0.92	1.56 ± 0.45	3.61 ± 0.71	1.42 ± 0.45	7.69 ± 1.19	7.48 ± 1.16	5.85 ± 0.94	3.63 ± 0.71	1.42 ± 0.45	–	5.06 ± 0.74	–
Lineage_XIV	5.23 ± 0.91	1.08 ± 0.39	2.84 ± 0.66	1.21 ± 0.41	5.50 ± 0.92	6.40 ± 0.98	5.35 ± 0.91	3.09 ± 0.69	2.94 ± 0.68	0.93 ± 0.38	4.46 ± 0.82	5.65 ± 0.93	5.69 ± 0.93	**0.27 ± 0.08**	–
Lineage_XV	2.57 ± 0.62	6.26 ± 0.98	7.38 ± 1.14	6.22 ± 0.98	2.44 ± 0.58	4.40 ± 0.68	2.31 ± 0.56	8.26 ± 1.20	8.07 ± 1.19	6.53 ± 0.98	4.69 ± 0.83	2.31 ± 0.56	2.42 ± 0.58	6.39 ± 0.98	**0.14 ± 0.00**

The mean DNA distance in percents ±SD above the diagonal is for control region, under the diagonal there are the values for the intron S7 r-protein; values on the diagonal (in bold) indicate within-CR lineage divergences.

Comparing different approaches of treating gaps:

(i) In terms of the topological accuracy view: Significant differences between topology trees obtained by coding gaps as "missing" (GM) characters and by the other two approaches—coding gaps as the fifth state character (G5) and coding gaps as separate present/absent characters (GS)—were found. Certain taxa were clustered without apparent logic when indels were excluded from the phylogenetic analysis. No significant differences were found between the results obtained by G5 and GS analyses (data not shown).

(ii) In terms of the bootstrap support view: The results of branch support based on GM coding differed in comparison with the results obtained by G5 and GS methods (data not shown). In the case of comparing nodal supports by G5 and GS analyses, there was a significant increase in bootstrap values obtained by G5 method, which was apparent in the inner and terminal nodes (Fig. 3ii). Similar findings were reached by Simmons et al. (2001), and by Ogden and Rosenberg (2007). Hence, while building phylograms, as well as nucleotype and haplotype networks, we used indels as an additional phylogenetic signal besides substitutions, and in addition, gaps were treated as the fifth state character.

The data obtained by analysing the mitochondrial and nuclear markers was first analyzed singularly, and then in combination. In terms of the combined dataset, the ILD test revealed significant incongruencies between the two analyzed loci. The *P* values were computed from 1000 replicates, and when S7 gaps were treated as the "additional state", the resultant value was $P = 0.001$. When gaps were treated as "missing", the resultant value was $P = 0.004$. Both values are very similar, which points out the fact that indels are not the reason for the reduction of phylogenetic accuracy. We then made a visual comparison of particular trees obtained from both markers (S7 and CR). The aim was to determine that these results did not represent the case described by various authors (Barker and Lutzoni, 2002; Darlu and Lecointre, 2002) when the incongruence length difference test failed. The comparison revealed certain differences, but only in terms of the topology of trees. In contrast, the case of the bootstrap values or Bayesian posterior probabilities of terminal taxa, an increase in resolution and support was found. The topological diversity was apparent when certain lineages were separated differently into two main clades, whereas the results obtained by mtDNA analyses appeared to be more logical (Figs. 3 and 4). On the basis of the above mentioned facts, we concluded that we would use information from both markers, including combined analysis for the evaluation of the taxonomic state of the studied representatives of the genus *Gobio* (the terminal nodes of the tree) similar to that reported by Gatesy et al. (1999) and Lavoué et al. (2003). However, in the systematics of the genus *Gobio* we only used the phylogenetic signal from the mitochondrial marker.

From the taxonomic point of view, we identified 13 separate pure monophyletic lineages of the genus *Gobio* (Figs. 2–4). Two gudgeon lineages, Lineage_XII or XIII (see below) and Lineage_XV were designated only by the S7 marker and the mitochondrial marker revealed a hybrid origin for the analyzed individuals (Fig. 2 and Supplementary Fig. SM_1). The phylogenetic analyses based on both markers, CR and S7, showed *G. cynocephalus* to be the most divergent species (Lineage_XI). The systematics of the genus *Gobio* based on the analysis of the mitochondrial control region revealed a clustering of the mentioned lineages into two main clades with a strong bootstrap support (BP) and a significant Bayesian posterior probability (PP), exluding Lineages_XII and XIII (Fig. 4). The first major clade we designated Northern European, and according to the BI analysis with PP = 1.00, we subdivided it into two subclades:

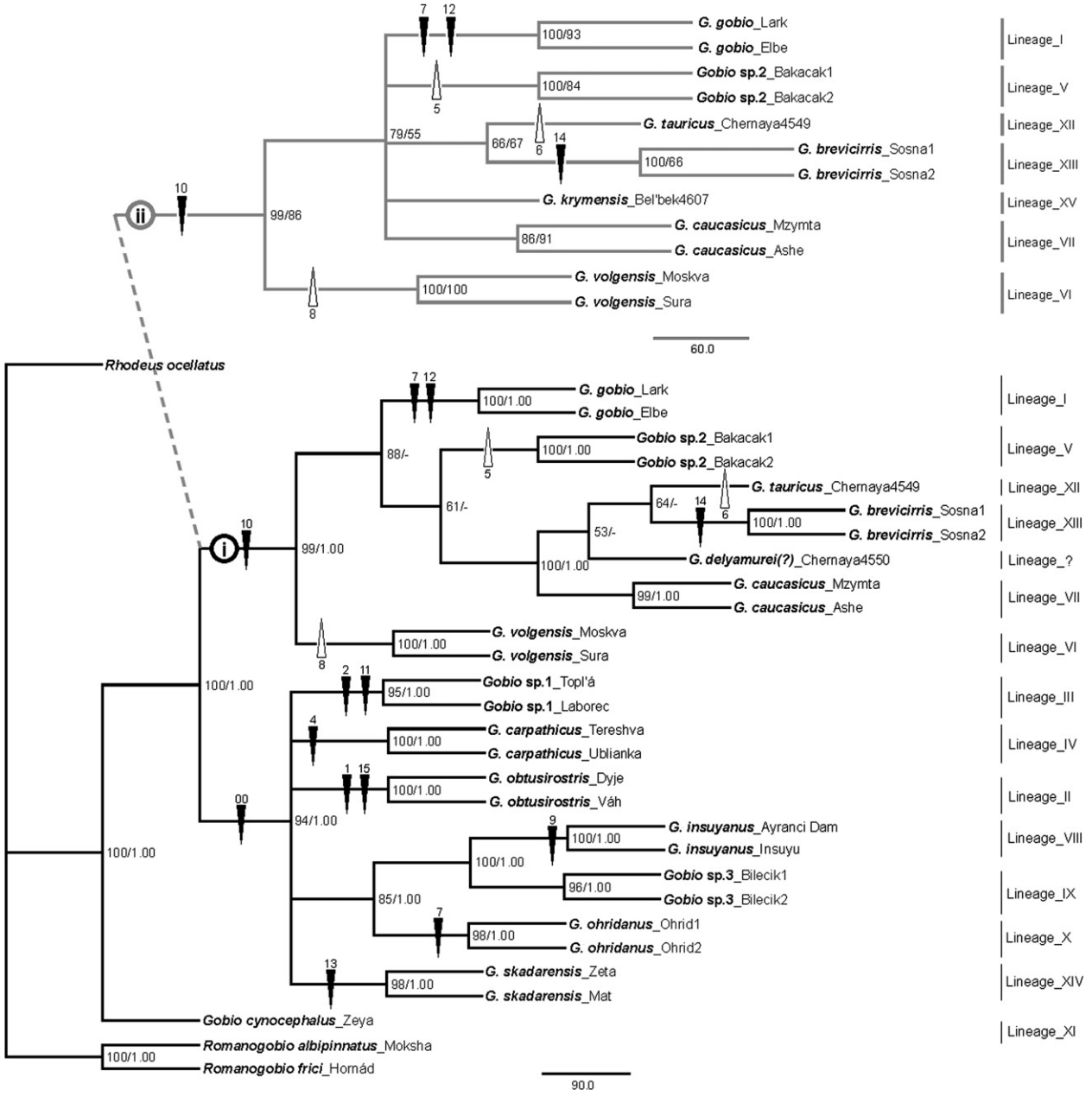

Fig. 3. Maximum parsimony tree inferred from combined data (CR and S7; i). Bootstrap values for MP and Bayesian posterior probabilities are listed near the nodes. Only values ⩾50% are shown. The fifteen lineages are highlighted in the phylogram. Up and down arrows represent insertions and deletions, respectively. Numbers on the arrows correspond to the gap codes in Fig. 5. The nominal species name is followed by the name of the locality, as well as in the subsequent phylogram. The partial MP tree inferred from the S7 sequences and showing Lineage_XV (ii). The number preceding the slash describes the value of support based on the G5 method; the number after the slash describes the value of support based on the GS method.

the Northwestern European (A) and the Northeastern European (B). The second major clade is designated Ponto-Caspian. We also subdivided this into two subclades: the Southern Ponto-Caspian (C) and the Northern Ponto-Caspian (D), being supported by a significant Bayesian posterior probability (PP = 1.00). However, the subdivision of the Northern European clade into subclades A and B is not supported by the other three statistical methods (MP, ML, NJ) and the subclades C and D received only moderate support (Fig. 4). In conclusion, without the division into subclades, the Northern European major clade was formed by the nominotypical Lineage_I, Lineage_II, Lineage_III, Lineage_IV, Lineage_V, Lineage_VI, Lineage_X, and Lineage_XIV. The Ponto-Caspian major clade was composed of the Transcaucasian Lineage_VII, two Turkey Lineages_VIII and IX, and two Northern Pontic Lineages_XII and XIII. Localities of common appearances of more lineages were revealed (Fig. 1 and Table 1): the Odra River and the Morava River tributaries – the Bečva R., the Dyje R., the Haná R. (region A, the Czech Republic) and the upper Tisza R. tributaries—the Laborec R., the Revištia channel (region B, Slovakia). Region C is restricted to the Northern part of the Black Sea (the Southwestern part of the Crimean Peninsula and the Sosna River—the Don River basin). The situation in areas A and B concerning hybridization of the gudgeons is very complicated and will therefore be specifically addressed in a subsequent article.

3.4. The first intron of the gene coding S7 r-protein as a diagnostic marker—S7indel diagnostics

The above presented results reveal the usefulness of the intron sequence for the evaluation of gudgeon taxonomy of the genus

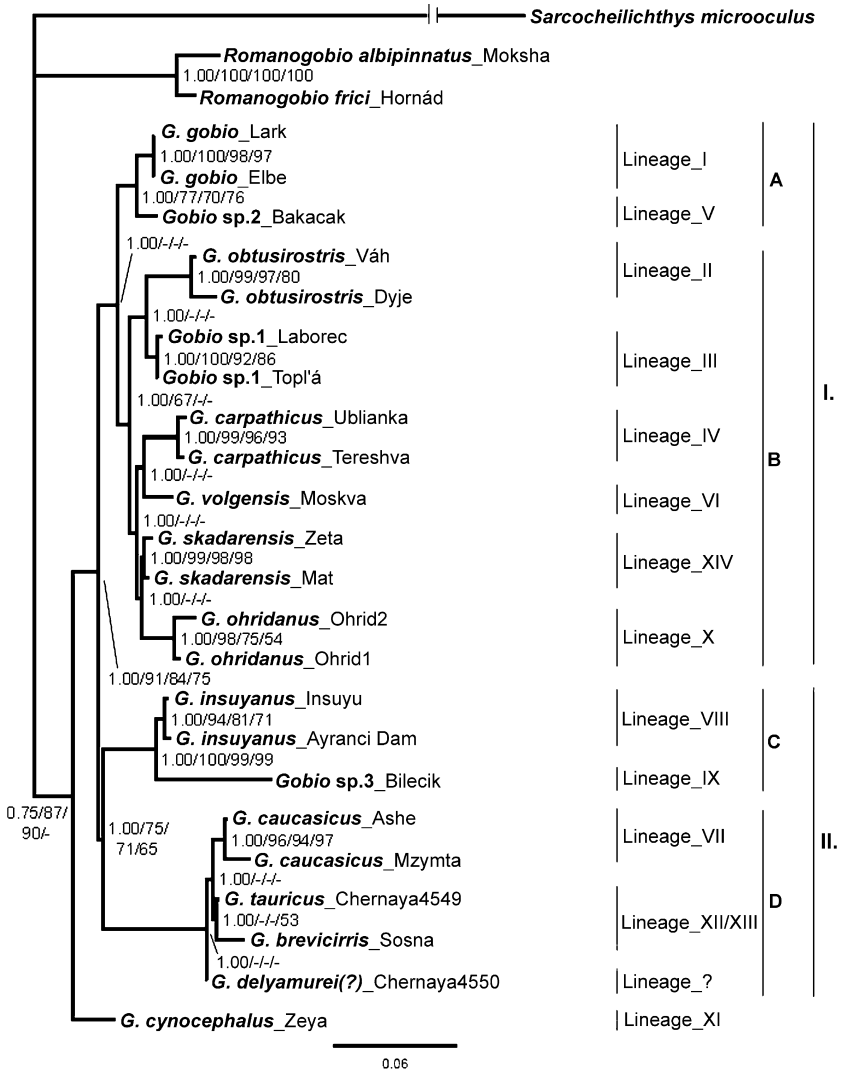

Fig. 4. Bayesian consensus tree resulting from the analysis of the control region data in studied gudgeon taxa, with Bayesian posterior probabilities/NJ bootstrap/MP bootstrap/ML bootstrap values listed near the nodes. Only values ⩾50% are shown. The species *G. cynocephalus* was associated with the three outgroups based on the ML method. The fourteen highlighted lineages are categorized into four subclades and two major clades.

Gobio. In a more detailed comparison of nuclear sequences in the individual species we found the first intron S7 to be a suitable diagnostic marker. We found that this unencoding nuclear region contains numerous deletions and insertions (indels), which are responsible for the different lengths of amplified PCR products. Partial S7 alignment of regions related to gaps only is shown in Fig. 5. It documents the finding of the region 00 and 15 indels, which always appear in at least two specimens of each lineage, except for the specimens from the Crimean Peninsula (the Chernaya River—4549 and 4550). In this lineage only one specimen containing 7nt (seven nucleotide) insertion (gap-6) was detected. The combination of the presence or absence of gaps led to the differentiation of certain species and to the discovery of a fast, simple and cheap PCR diagnostic method, the "S7indel diagnostics". Different lengths of the PCR products in unambiguous lineages are listed in Table 3. Individual diagnostic indels and their significance are displayed in the phylogenetic tree generated by the MP method (Fig. 3). The general benefit of this diagnostic method could be evaluated and verified only after the completion of the not yet examined species of the genus *Gobio*. If necessary, this technique can be further developed on the basis of primer position, or RFLP method, etc.

3.5. Phylogeny of Gobio, overview

The taxonomy of gudgeons in the Palaearctic zone still classified as the common gudgeon *G. gobio* by most authors (Bănărescu et al., 1999; Tsepkin, 2002; Golubtsov and Malkov, 2007) is an issue of great importance which is currently undergoing constant development. This is documented by numerous new species described over the last four years (Doadrio and Madeira, 2004; Vasil'eva et al., 2004, 2005; Kottelat and Persat, 2005; Freyhof and Naseka, 2005; Naseka et al., 2006). In the past, Bănărescu and Nalbant (1973) recognized 19 subspecies in polytypical species *G. gobio* s. lato, and Bănărescu (1992a) documented only 17. In total, 149 nominal names were proposed for these fishes in scientific literature; the majority of them were later reevaluated as synonyms. Thus, approximately 20 are now considered to be valid species at the present time (Froese and Pauly, 2007). In an effort to maintain transparency we will treat the individual analyzed gudgeon taxa separately.

3.5.1. Gobio cynocephalus Dybowski, 1869

According to our molecular analyses, specimens from the Zeya River (the Middle Amur River drainage) represent the most diver-

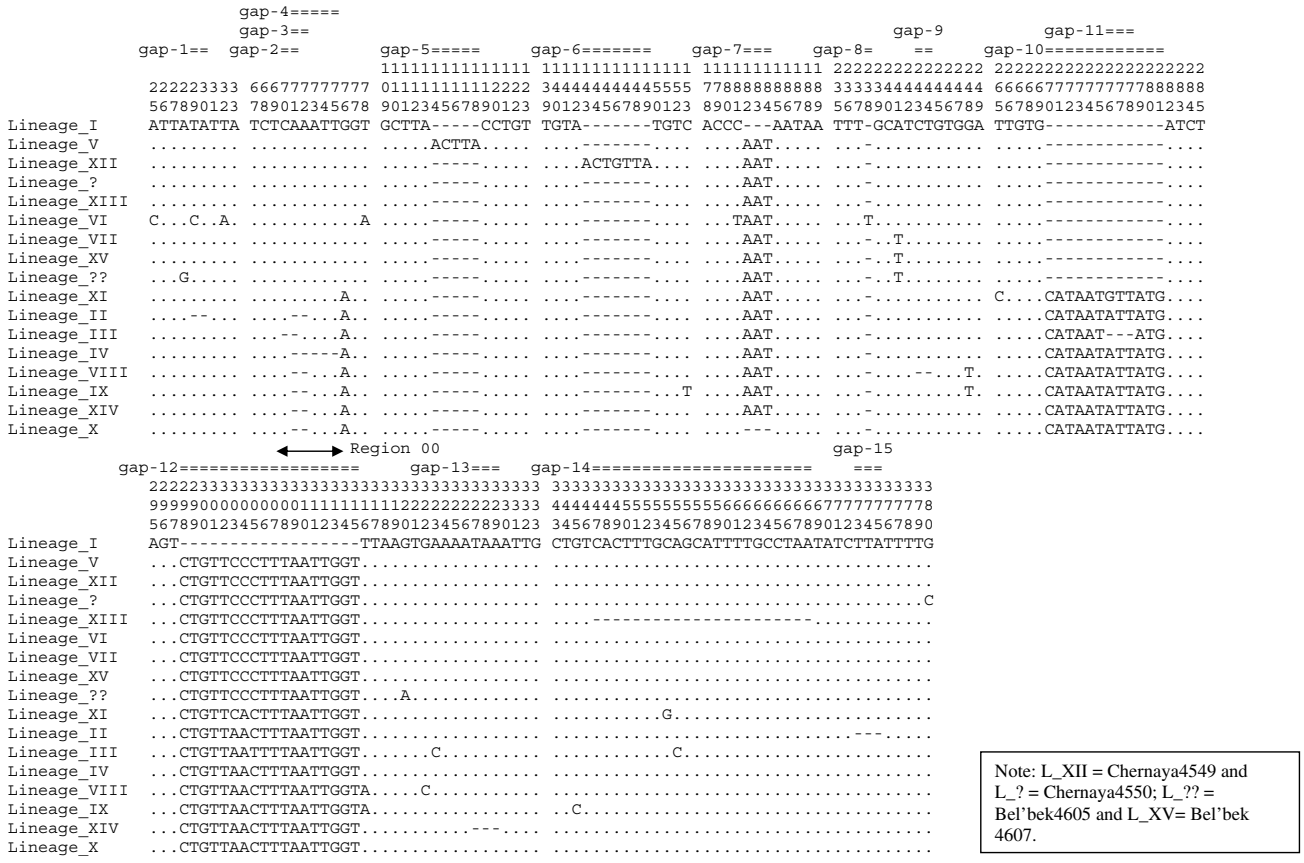

Fig. 5. Partial S7 alignment of the region containing gaps in selected representatives of fifteen lineages of the genus *Gobio*. Parsimony-informative gaps, treated as single indel mutations, are indicated by = = =.

gent group among the studied gudgeons. The Amur River basin is inhabited by the Amur gudgeon *G. cynocephalus* Dybowski, 1869 s. stricto, described as a variety of the common gudgeon from the Onon and Ingoda rivers (the Upper Amur). Some authors consider the Amur gudgeon to be a subspecies of *G. gobio* distributed from Siberia to the Far East (Berg, 1949; Naseka, 1998; Tsepkin, 2002). In the phylogenetic tree (Figs. 3 and 4) the Amur gudgeon represents Lineage_XI in which one CR haplotype H_41 and one S7 nucleotype N_26 were detected (Fig. 2 and Supplementary Fig. SM_1). The sequence divergencies towards the other analyzed gudgeons are listed in Table 5. In terms of the "S7indel diagnostics" Lineage_XI has the longest PCR product (371 bp) and is therefore distinguishable from many other species (Table 3). Thus, our data supports the status of the Amur gudgeon as an independent species, and its relationship to gudgeons from Siberia needs further investigation.

Gudgeons of the major clade I—Northern European

3.5.2. G. gobio (Linnaeus, 1758) sensu stricto

The common gudgeon, *G. gobio* s. lato was claimed to have a large distribution area from the Iberian Peninsula and the British Islands to the Far East rivers (Berg, 1949; Bănărescu et al., 1999). The type specimen of *Cyprinus gobio* Linnaeus, 1758 is still unknown, and the type locality of the nominotypical subspecies of the common gudgeon, *G. gobio gobio*, was supposedly located in the southeastern part of England (Bănărescu et al., 1999). Kottelat and Persat (2005) redescribed *G. gobio* and designated its neotype from the stream Sieg at Eitorf (Rhine River drainage in Germany). On the basis of our molecular analyses and comparison with the representatives from the Lark River in southeastern England we can prove that *G. gobio* s. stricto is widely distributed in the following river basins (Table 1): Rhone (the Mediterranean Sea drainage area), Rhine (Lahn R.) and Elbe (Blanice R., D. Orlice R., Elbe R.) (the North Sea drainage area), Odra (Odra R., Stěnava R.; the Baltic Sea drainage area), Danube (Bečva R., Haná R.), Tisza (Revištia channel), and Southern Bug (the Black Sea drainage area). Nine mitochondrial haplotypes H_1–9 and four nucleotypes N_1–4 were distinguished (Fig. 2 and Supplementary Fig. SM_1). Haplotype H_1 is widespread. The representatives of this species form Lineage_I in the phylogram and they are part of subclade A—Northwestern European (Fig. 4). The important diagnostic feature of this species is the 18nt deletion (gap-12), which is present in this species only (Figs. 3 and 5). The interspecific divergencies are shown in Table 5. The intraspecific divergence of sequences in both markers did not exceed 0.37%.

3.5.3. Gudgeons from the Danube River basin

The situation with the gudgeons from the Danube River basin is complicated due to the presence of a large number of lineages. We found four pure lineages (L_I–IV) in this area. The first of them is L_I, representing *G. gobio* s. stricto and is discussed above. Lineages III and IV are formed by Danubian samples only, whereas L_II is also represented in the Baltic Sea drainage. Three gudgeon taxa were described from the Danube River basin, and each of them is discussed further as available names for Lineages II–IV.

3.5.3.1. Gobio obtusirostris Valenciennes, 1842. This species was described from the Isar R. at Munich (the upper Danube R. basin). Vladykov (1925, 1931) considered it to be a valid subspecies of the common gudgeon and extended its distribution to the middle and lower part of the Danube River basin, excluding the Tisza R. populated by another subspecies. Bănărescu (1961) presumed *G.*

gobio obtusirostris to be the only subspecies of the common gudgeon, but Bănărescu et al. (1999) began to doubt the validity of this argument. Freyhof and Naseka (2005) considered it a valid species. Our phylogenetic analyses confirmed the validity of the Danubian gudgeon being widely distributed in the Danube River basin and also occurring in the Odra River (Fig. 1 and Table 1). One nucleotype N_5 and five haplotypes H_10 – 14 were found for these populations, of which H_10 and H_13 occured most frequently (Fig. 2 and Supplementary Fig. SM_1). In the phylogram based on mtDNA analysis it is differentiated as monophyletic Lineage_II with strong support belonging to subclade B—Northeastern European (Fig. 4). The intraspecific sequence differences on both markers did not exceed 0.31%. At the same time this lineage demonstrated high genetic differences in comparison to the common gudgeon Lineage_I (Table 5). Moreover, Lineage_I and Lineage_II differed significantly in the nuclear marker, namely by numerous substitutions and especially by six indels (Fig. 5). This means that they are simply distinguishable by "S7indel diagnostics" (Table 3). Finally, the areas of both lineages demonstrate significant overlapping (Fig. 1), but no hybridization between them was noted. The species status based on morphologic data, suggested by Kottelat and Freyhof (2007) is also supported by our results.

3.5.3.2. G. gobio carpathicus Vladykov, 1925, G. gobio muresia Jaszfalusi, 1951 and/or Gobio sp. 1. G. gobio carpathicus was described as a subspecies of the common gudgeon occurring in the Upper Tisza basin (middle Danube R. drainage). Berg (1949) presumed that this subspecies might also occur in the Lower Danube. The type locality of another subspecies *G. gobio muresia* was designated at the Mures River near Gödemesterháza and at the confluence with the creeks Zebrak and Göde in Romania (Kottelat, 1997; the lower part of the Tisza River basin). Bănărescu (1961) concluded that both were synonyms of *G. g. obtusirostris*, whereas Freyhof and Naseka (2005) classified gudgeons from the Tisza and Mures rivers as *G. carpathicus.* Our phylogenetic analyses revealed two sympatric lineages in the Upper Tisza drainage: Lineage_III and Lineage_IV (Figs. 3 and 4 and Table 1), both included in the B—Northeastern European subclade in the CR phylogram (Fig. 4). Their genetic differences (Table 5) reach 1.69% (CR) and 2.70% (S7) concerning substitutions, and in terms of nDNA they differ in three indels (gap-2, gap-4, gap-11; Figs. 3 and 5). We identified four haplotypes H_15–18 and one nucleotype N_6 of Lineage_III in the rivers Laborec and Top'lá, the Belžanský stream and four haplotypes H_19–22 and three nucleotypes N_7–9 of Lineage_IV in the rivers Dyje, Laborec, Tereshva, Ublianka and the Revištia channel (Fig. 2 and Supplementary Fig. SM_1). The intraspecific diversity in both lineages did not exceed 0.30%. Their ranges overlap significantly (Fig. 1), however, no hybrid individual of either Lineages_III × IV was observed up to the present time.

Our data can lead us to assume the presence of two independent species in the Tisza R. basin, but their relationship to the aforementioned nominal names needs further investigation. Because the specimens of *G. gobio muresia* from the type localities were not investigated, its validity or conspecificity with *G. carpathicus* will remain problematic. Our study should prove the availability of the name *carpathicus* for one of the two lineages. In this work we will initially designate it for L_IV in concordance with the study of Freyhof and Naseka (2005), who describe the occurence of this species in the same localities. Lineage_III is designated as *Gobio* sp. 1—"species-in-waiting".

3.5.4. Gudgeons from the Volga River basin

Previously, gudgeons from the Volga R. basin were classified as the common gudgeon, *G. gobio* or its nominotypical subspecies (Berg, 1949; Bănărescu, 1961; Naseka, 1998; Bănărescu et al., 1999; Tsepkin, 2002; Ruchin and Naseka, 2003; Vasil'eva et al., 2004; Vasil'eva and Kuga, 2005; Freyhof and Naseka, 2005). In the phylogenetic tree (Fig. 4), the specimens from the Volga River basin form the monophyletic Lineage_VI, belonging to subclade B—Northeastern European. We identified one haplotype H_24 and two nucleotypes N_12–13 in specimens from the rivers Bol'shaya Lašva, Chardym, Malaya Tsivil', Moskva and Sura (Table 1 and Fig. 2 and Supplementary Fig. SM_1). The interpopulation divergency did not exceed 0.60% and the values of the interspecific divergency are shown in Table 5. The "S7indel diagnostics" is able to distinguish it from several gudgeon lineages, including Lineage_I (Table 3). On the basis of the above mentioned findings we propose that gudgeons from the Volga R. basin should be classified as a separate species. According to a previous study (Vasil'eva et al., 2004), this species is very similar to the common gudgeon in external morphology, and thus should be considered as a cryptic species. We did not find any available name for it in previous publications and therefore we consequently describe it as a new species in this paper (the description of the species is attached).

3.5.5. Gudgeons from the Ohrid-Drim-Skadar hydrologic system (the Adriatic Sea drainage)

Two local gudgeon forms and one subspecies were described from this largest hydrologic system in the western Balkan zoogeographic region (Bănărescu, 1992b): *G. gobio* var. *ohridana* Karaman 1924 from Ohrid Lake, *G. gobio lepidolaemus* form *skadarensis* Karaman, 1936 from Skadar Lake, and *G. gobio albanicus* Oliva, 1961 from the Kiri R. (the Drin R. system) in Albania. Further comparative morphological and meristic analyses showed that all gudgeons within the Ohrid-Drim-Skadar system are conspecific (Grupče and Dimovski, 1975; Šorić and Ilic, 1988). Most authors consider *G. gobio ohridanus* to be the only valid subspecies of the common gudgeon (Grupče and Dimovski, 1975; Šorić and Ilic, 1988; Šorić, 1990; Bănărescu, 1992a; Bănărescu et al., 1999). In constrast to this opinion, Šanda et al. (2005) consider *G. g. ohridanus* to be a junior synonym of *G. gobio gobio* based on the results from allozyme analysis. We studied gudgeon populations related to both the *ohridanus* and *skadarensis* nominal names.

3.5.5.1. Gobio ohridanus Karaman, 1924. The Ohrid Lake is the type locality of this taxon. Our phylogenetic analyses extended its distribution range: besides the Ohrid Lake we also noted its occurrence in the Albanian Mat River (Table 1 and Fig. 1). In the phylogram, this species is marked as the monophyletic Lineage_X (Figs. 3 and 4), which is part of subclade B—Northeastern European, and contains four haplotypes H_37–40 and one nucleotype N_25 (Fig. 2 and Supplementary Fig. SM_1). The Lineage_X demonstrates the high level of sequence divergence from L_I and can be clearly differentiated by "S7indel diagnostics" (Table 3). The intraspecific divergence did not exceed 0.48%. Proposed species status (Kottelat and Freyhof, 2007) is supported also by our results.

3.5.5.2. Gobio skadarensis Karaman, 1936. The local gudgeon form *skadarensis* was described from the Skutari or Skadar Lake (the Drim River basin) in present Montenegro and Albania. Bănărescu et al. (1999) noted that this form should be considered a synonym of *G. g. ohridanus*. We studied specimens from the Zeta River, the largest tributary of the Morača River, the main tributary of the Skadar Lake. These specimens are included into subclade B—Northeastern European—and form the monophyletic Lineage_XIV (Fig. 4), in which four haplotypes H_45–47 and two nucleotypes N_30–31 were found (Fig. 2 and Supplementary Fig. SM_1). The identical sequence pattern was also found in the Albanian Mat River where gudgeons from the Ohrid Lake (Lineage_X) were noted (Table 1 and Fig. 1). Nevertheless no hybridization between L_XIV and L_X was observed in the zone of sympatry. The intraspecific variability in both markers did not exceed 0.27%. The 3nt dele-

tion is characteristic (gap-13; Figs. 3 and 5) for these gudgeons and they can be easily discriminated from the nominotypical Lineage_I by "S7indel diagnostics" (Table 3). The molecular results support the species status proposed on the basis of the morphologic data by Kottelat and Freyhof (2007).

Gudgeons of the major clade II—Ponto-Caspian, and Turkish gudgeons with European relations

3.5.6. Gudgeons from Turkey

The Turkish area is of great interest since it belongs to one of the most important glacial refuges, the Ponto-Caspian (Bănărescu, 1992b), and represents a region with a high variability of gudgeon populations resulting in the identification of eight species/subspecies of the genus *Gobio* (Erk'akan et al., 2005; Naseka et al., 2006). Our analyses revealed three different monophyletic lineages from this area with strong support. On the basis of mitochondrial analysis, one of the lineages (Lineage_V) was assigned to subclade A—Northwestern European, and two lineages (Lineage_VIII and Lineage_IX) were assigned to subclade C—Southern Ponto-Caspian (Fig. 4).

3.5.6.1. Gobio sp. 2. Specimens of Lineage_V (Figs. 3 and 4) come from Northwest Anatolia (Bakacak deresi, Biga; Fig. 1). Erk'akan et al. (2005) presumed this area to be populated by the nominotypical species *G. gobio*. One haplotype H_23 and two nucleotypes N_10, 11 were identified within this lineage (Fig. 2 and Supplementary Fig. SM_1). The intraspecific sequence divergence did not exceed 0.45%. This lineage shows certain similarity with the nominotypical Lineage_I (Table 5) and belongs to the same subclade A (Fig. 4) as well, but two deletions (gap-7 and gap-12), which are typical for the common gudgeon were not found in the Turkish lineage (Figs. 3 and 5). Therefore Lineages_I and V are easy distinguishable by "S7 diagnostics" (Table 3) and it is evident that they are not conspecific. Freyhof and Naseka (2005) classified gudgeons from the Meria R. in the European part of Turkey as *G. bulgaricus* Drensky, 1926 that were originally described from the Maritza River (Southern Bulgaria). Since gudgeon populations from the Maritza R. basin were not studied, we will refrain from any conclusion on the availability of the name *bulgaricus* for gudgeons from our Anatolian Lineage_V and designate it as *Gobio* sp. 2, a species which needs a comprehensive revision.

3.5.6.2. Gobio insuyanus Ladiges, 1960. The specimens of the second distinct Turkish Lineage_VIII originate from Central Anatolia, specifically from three localities, the Ayranci Dam Lake, the Sugla River, and the Insuyu Stream (Fig. 1 and Table 1). The last of these is the type locality of *G. gobio insuyanus* (Erk'akan et al., 2005; Naseka et al., 2006). Seven haplotypes H_29–35 and three nucleotypes N_20–22 were found in this lineage (Fig. 2 and Supplementary Fig. SM_1). The intraspecific diversity did not exceed 0.30%, the values of genetic divergence from L_V were high for both markers (Table 5). Figs. 3 and 5 show the existence of the typical gap-9. These results support the specific status of *G. insuyanus* proposed by Naseka et al. (2006) and further extends its occurrence outside the type locality.

3.5.6.3. Gobio sp. 3. Gudgeons from the locality Bilecik in northwestern Anatolia, form the third Turkish monophyletic Lineage_IX with haplotype H_36 and nucleotypes N_23–24 (Figs. 1–4 and Supplementary Fig. SM_1) with an intraspecific diversity about 0.30%. The aforementioned locality belongs to the area of the Sakarya River basin, where the occurrence of the subspecies *G. g. obtusirostris* was previously recorded (Erk'akan et al., 2005). However, Lineage_IX significantly differs from Lineage_II in both molecular markers: by 9.53% (CR) and 2.81% (S7). Moreover, due to the "S7 diagnostics," L_IX is easily distinguishable (Table 3). Thus Lineages_II and IX obviously represent different species, and we have designated the latter as *Gobio* sp. 3, "species-in-waiting" until a comprehensive revision can be undertaken.

3.5.7. Gudgeons from the northeastern coast of the Black Sea and the Crimean Peninsula

Phylogenetic analyses revealed four separate lineages of different origin among samples collected in this area. Only one of them, Lineage_VII, represents an unquestionably pure lineage with its own mitochondrial and nuclear patterns. In contrast, lineages L_XII and L_XIII demonstrated an ambiguous situation with an indefinite origin of the CR pattern. Therefore, we used the label L_XII/XIII in Figs. 2 and 4. The fourth lineage L_XV includes specimens that were demonstrated to be of hybrid origin.

3.5.7.1. Gobio caucasicus Kamensky, 1901. The specimens examined from the northeastern coast of the Black Sea (Fig. 1) form the monophyletic Transcaucasian Lineage_VII in the phylogram (Fig. 3), which belongs to subclade D—Northern Ponto-Caspian (Fig. 4). Four haplotypes H_25–28 and four nucleotypes N_14–17 were determined for this group (Fig. 2). The intraspecific diversity on both markers did not exceed 0.55%. The "S7indel diagnostics" differentiates this lineage within subclade D (excluding Lineage_XV; Table 3). Freyhof and Naseka (2005) stated that rivers from the Black Sea basin in Krasnodar province (including rivers presented in this study) are populated by the Caucasian gudgeon *G. caucasicus* which was described as a variety *caucasica* of the Central Asian gudgeon *G. lepidolaemus* Kessler, 1872 from both the Caspian Sea (Podkumok and Sulak rivers) and the Black Sea (Rioni R. system) basins (see Kottelat, 1997). We are not sure of the conspecificity of the Caspian and Black Sea populations, but we will presume to use *G. caucasicus* as an available name for gudgeon species represented by the phylogenetic Lineage _VII.

3.5.7.2. Gobio brevicirris Fowler, 1976. Freyhof and Naseka (2005) classified gudgeons from the Don River drainage as a distinct species *G. brevicirris.* This name became available after Fowler (1976) concluded *G. gobio* morpha *brevicirris* Berg, 1932 to be a valid subspecies *G. gobio brevicirris* distributed in Ukraine and Russia (see Kottelat, 1997). The specimens from the Don River basin (Sosna R.; Fig. 1) represent the second Northern Pontic lineage (subclade D; Fig. 4) designated as Lineage_XIII in the phylogram (Fig. 3). This lineage is characterised by the haplotype H_44 and unique nucleotypes N_28–29 (Fig. 2) which differ from each other by 0.26%. The significant long 22nd deletion was found (gap-14; Figs. 3 and 5). "S7indel diagnostics" enables us to distinguish this lineage and Lineage XII specifically, (Table 3) despite their ambiguous species status of CR patterns. Molecular results support the species status of the Don gudgeon proposed on the basis of morphologic data by Kottelat and Freyhof (2007). Further analyses will be necessary to resolve the lineage classification of the haplotype H_44 (L_XII/XIII; Fig. 2 and Table 1).

3.5.7.3. Gobio tauricus Vasil'eva, 2005 and Gobio delyamurei Freyhof and Naseka, 2005. The molecular analysis of specimens from the Chernaya River raises complicated questions about the taxonomic status of gudgeon populations in the western part of the Crimean Peninsula. The specimen labeled as Chernaya4549 had haplotype H_42 and a unique nucleotype N_27, whereas the specimen labeled as Chernaya4550 had a similar haplotype H_43 but a different nucleotype N_19, which is related to nucleotypes of Lineage_VII (Fig. 2). The sequence divergences between these specimens reached 0.81% and 1.74% on mtDNA and nDNA markers,

respectively (Table 5). This increased "interindividual" difference results from substitutions and the presence of 7nt insertion (gap-6; Fig. 5), not found in other gudgeon lineages. In accordance to the aforementioned differences, the specimen Chernaya4549 represents a distinct Lineage XII in the phylogram (Fig. 3). We will refrain from classifying of the specimen Chernaya4550 and leave this question open for the present (Table 1 and Figs. 2–4). The above mentioned data indicates the genetic and taxonomic heterogeneity of gudgeons from the Chernaya River. This conclusion agrees with obvious morphological heterogeneity observed in samples used by Freyhof and Naseka (2005) in their description of *G. delyamurei*, and also with our new materials collected from the same river. In accordance with the molecular data we surmise that the present gudgeon population in the Chernaya River results from hybridization between native species (Lineage XII) and from species having penetrated into the river during recent years as a result of acclimatization and irrigation activity in the Crimean Peninsula. The native species was subjected to karyological and craniological analyses based on specimens collected in the Chernaya River in 1981. The karyological and craniological peculiarities of these gudgeons are the main grounds for the description of the species *G. tauricus* and were presented as its main diagnostic characters (Vasil'eva et al., 2005). Thus, we classify the specimen Chernaya4549 as *G. tauricus*, despite the hybrid origin of several type specimens identified by molecular analyses (ICZN, 1999, art. 17). At the same time, the morphological characters of the holotype of *G. delyamurei* presented by Freyhof and Naseka (2005) allow consideration of this specimen to not be conspecific to native gudgeons that have been distributed in the Chernaya R. in the past. This situation with gudgeons in the Chernaya R. is very complicated and also needs further investigation.

3.5.7.4. Gobio krymensis Bănărescu and Nalbant, 1973. In contrast to previous Ponto-Caspian lineages, the molecular analyses of gudgeons from the Bel'bek River in the Steppe Crimea (Fig. 1 and Table 1) labeled as Bel'bek4605 and Bel'bek4607 revealed a specifically mixed origin. Bel'bek4607 had haplotype H_9 closely related to haplotypes from Lineage_I, and unique nucleotype N_32 (Fig. 2). Therefore this specimen forms (based on S7 marker analysis) a separate Lineage_XV in phylogram (Fig. 3ii). According to obtained data, we concluded that this specimen most probably represents a hybrid between a female from L_I and a male of another species distributed in the Bel'bek River. The other specimen from this sample (Bel'bek4605) had haplotype H_1 and nucleotype N_18 (Fig. 2) and should be considered as a hybrid between a female from Lineage_I and a male from another phylogenetic lineage with CR haplotypes most related to the ones of Lineage_VII. The S7 sequence variability of both the representatives from the Bel'bek River was 2.24%. We will refrain from the classification of males having participated in the aforementioned hybridization as members of Lineage_VII representing *G. caucasicus* since the hybridization between this species and Crimean gudgeons seems impossible due to their geographic isolation. It is possible that, arrangements of acclimatization and irrigation in the Crimean Peninsula indicate there may have been an accidental introduction (and further distribution) of gudgeons from the Dnieper River basin, which have not been subject to molecular studies. Therefore it is quite possible that at least one of the nucleotypes N_18 or N_19 belong to *G. sarmaticus* Berg, 1949 distributed in the Dnieper R. basin. As to an available name for the Lineage_XV distributed in the Steppe Crimea, we are inclined towards Freyhof and Naseka (2005), who classify these gudgeons as independent species *G. krymensis*. The above stated data indicates that the situation with gudgeons from the Crimean Peninsula is extremely complicated and requires a more exhaustive analysis, both morphological and genetic.

3.6. Phylogeography, genetic aspects and taxonomic implications

The aim of this phylogenetic study was to bring up new findings in the field of genetics and the phylogeography of the genus *Gobio* and to throw light on the current and rather complicated taxonomy and systematics of this genus. We attempted a more comprehensive molecular approach based on combinations of both mitochondrial and nuclear genomic markers. On the basis of sequence analyses of the material collected at type localities or in their close surroundings, and on the background of the data from literature, we have arrived at the following findings and conclusions. The gudgeons of the genus *Gobio* show a large scale of haplo- and nucleotype patterns, which also exhibit a large distribution spectrum ranging from small areas (*G. ohridanus*, *G. skadarensis*, some Turkish gudgeons, etc.) to vast territories covering thousands of kilometres, e.g. extending from the British Isles to the Black Sea, as in the case of the nominotypical species *G. gobio* s. stricto. Localities (Bečva R., Dyje R., Haná R., Laborec R., Odra R., Revištia ch.; Table 1 and Fig. 1) demonstrate the sympatry of several different species (*G. gobio* and *G. obtusirostris*, gudgeons from the Tisza River basin), which can lead to problems in their identification, as a result of possible hybridization events. Our analysis reveals that the cytonuclear disequilibrium is an actual phenomenon among gudgeons of the genus *Gobio* (rivers Bel'bek, Chernaya/Sosna; Table 1). The phylogeography of different species, including the zones of their sympatry (A–C), are shown in Fig. 1. We surmise that the existence of many closely related species living together in the same proximity presents the likelihood of frequent hybridization or introgresive hybridization. We consider this phenomenon to be one of the principal reasons for the wide variability of gudgeons, as mentioned by most previous authors. The question as to whether or not specimens of hybrid origin form a numerous, viable and spawning lineage can be answered only after further investigation.

The above mentioned results also leads us to the conclusion that data resulting from phylogenetic studies based on mtDNA analyses only are not sufficient for taxonomical reconstructions. The application of a suitable nDNA locus in combination with the mtDNA marker provides more useful tools to answer systematic questions. In addition to the employed molecular methods we discovered a new and promising method called "S7indel diagnostics" which is based on different lengths of the PCR products in most studied gudgeon lineages and therefore allows for a more simple identification of species of the genus *Gobio* undistinguishable by traditional morphological characters; for example, *G. gobio* and the new species *G. volgensis*. We presume that further investigations of gudgeons from different parts of the generic distribution will illustrate the convenience of this method with regard to the taxonomy.

The phylogentic analyses based on both mtDNA and nDNA markers confirm the validity of the genus *Gobio* as a monophyletic group with strong support, similarly mentioned by Yang et al. (2006). Altogether 15 gudgeon lineages are distinguishable in this genus, most of them identified as pure species. The phylogenetic relations obtained by control region analysis and applied statistical methods (NJ, MP, ML, BI) demonstrate these lineages are divided into two main clades, namely, the Northern European and the Ponto-Caspian. According to the BI analysis, the first clade was subdivided into two subclades—Northwestern European (A) and Northeastern European (B). The second main clade was subdivided into the Southern Ponto-Caspian (C) and the Northern Ponto-Caspian (D). These results agree with previous zoogeographic data.

The molecular analyses confirmed the validity of 11 taxa as independent species of the genus *Gobio*, namely *G. gobio*, *G. obtusirostris*, *G. carpathicus*, *G. caucasicus*, *G. insuyanus*, *G. ohridanus*, *G. skadarensis*, *G. cynocephalus*, *G. brevicirris*, *G. tauricus*, and *G. krym-*

ensis. Their genetic diagnostic characters were revealed as well. Based on these studies, gudgeons from the Volga River basin were separated from *G. gobio* s. stricto and described as a new species *G. volgensis*. Moreover, three phylogenetic lineages designated as *Gobio* sp. 1–3 were submitted for a comprehensive revision owing to their description/redescription as separate species. Thus, this study offers new views of the present taxonomy of the genus *Gobio*.

Acknowledgments

The study was carried out within the framework of the research project no. VaV-SM/6/3/05 supported by the Grant from the Ministry of Environment of the Czech Republic and by the Grant project no. IAA6045005 from the Grant Agency of the Academy of Sciences of the Czech Republic. Our investigations in Russia were partially supported by RAS Programs "Biodiversity" and "Biological resources of Russia". The investigations in Albania and Montenegro were supported by the Grant project no. MK00002327201 of the Czech Ministry of Culture. We are thankful to C.S. Özeren, J. Taggart, N.G. Bogutskaya for their help with the obtaining of samples. We thank G. Speas for English correction.

Appendix A. Description of new species

A.1. Gobio volgensis Vasil'eva, Mendel, Vasil'ev, Lusk, Lusková sp. nova (Supplementary Fig. SM_2)

Cyprinus gobio (not of Linnaeus, 1758): Pallas, 1814: 295 (part.: Volga R. basin).

Gobio fluviatilis (not of Fleming, 1828): Cuvier in Cuvier and Valenciennes, 1842: 300 (Europe-part.); Kessler, 1877: 251 (part.: Eastern Europe-part.).

Gobio gobio (not of Linnaeus, 1758): Berg, 1914: 428 (part.: Europe-part.); Berg, 1916: 218 (part.: Europe-part.); Lukash, 1923: 174, 176 (Vychegda); Lukash, 1933: 56 (Rivers Vyatka, Voya, Iryuk); Lukash, 1940: 26 (Vyatka and Kama basins); Berg, 1949: 640 (part.: Europe-part.); Bănărescu, 1992a: 317 (part.: Caspian Sea basin); Naseka, 1998: 82 (part.: European part of Russia-part.); Bănărescu et al., 1999: 81 (part.: Europe-part.); Tsepkin, 2002: 249 (part.: Europe-part.); Ruchin and Naseka, 2003: 334–335 (Sura R.); Vasil'eva et al., 2004: 772 (part.: Volga R. basin). Freyhof and Naseka, 2005: 336 (part.: Sura, Volga).

Gobio gobio gobio (not of Linnaeus, 1758): Bănărescu et al., 1999: 109 (part.); Ruchin and Naseka, 2003: 334 (Volga).

A.1.1. Holotype

ZMMU P-21861, SL 91.5 mm, TL 109.0 mm, the Moskva River at Staraya Ruza City, Moskovskaya District; collector V.P. Vasil'ev, 2004, August 21.

A.1.2. Paratypes

ZMMU P-21865, 4 spec., SL 64.2–86.0 mm, TL 76.5–103.0 mm the Moskva River at Zvenigorod City, Moskovskaya District; collector V.P. Vasil'ev, 2005, June 14; P-21910, 4 spec., SL 46.5–66.0 mm, TL 57.0–78.0 mm the Moskva R. at Zvenigorod City, collector V.P. Vasil'ev, 2004, June 05, voucher specimens for this molecular study.

A.1.3. Additional materials

The Moskva R. basin: P-442 (5 spec.), P-2705 (1 spec.), P-16229 (4 spec.), P-16819 (1 spec.), P-17966 (2 spec.), P-21235 (9 spec.), P-21413 (1 spec.), P-21422 (10 spec.), P-21426 (5 spec.). The Volga R. basin: P-3441 (Moksha R., 45 spec.), P-4164 (Oka R., 3 spec.), P-9561 (Ozerna R., 1 spec.), P-21040 (Sura R., 1 spec.), P-21234 (Kobra R., 7 spec.), P-21206 (Vytebet' R., 1 spec.), P-21236 (Mytets R., 5 spec.), P-21844 (Sura R., 3 spec.).

Comparative materials on *G. gobio* s. stricto. England: P-9423 (Thames R. at Reading, 1 spec.). The Baltic Sea basin: P-2762 (Neman R., Lithuania, 4 spec.), P-13033 (pound Vira at Třeboň, Southern Bohemia, 1 spec.), P-19034 (Ahja R., South Estonia, 1 spec.).

A.1.4. Diagnosis

D (II) III (7) 8; *A* II (III) 6 - 7; *V* I (II) 7(8); *P* I (14) 15-16; *l.l.* 40–43, usually 42–43; the body and the caudal peduncle are moderately compressed; the minimum body depth is somewhat smaller than the width of the caudal peduncle at the level of the last anal ray in larger specimens and somewhat greater in smaller fishes; the anus is closer to the insertion of the anal fin than to the origin of the pelvic fins; there are no epithelial crests on the dorsal scales and there are no barbellike prolongations at the corners of the mouth; barbels are moderately long: they usually extend beyond the anterior edge of the eye (only rarely do they not reach anterior eye edge), sometimes reaching up to the middle of the eye, but never reach to its posterior edge; the barbel length varies from 15% to 28% of the head length with modal values between 21% and 22%; paired fins are moderately long: pectoral fins never reach the pelvic fin insertion, and their average length varies from 74.7% to 84.8% of the distance between the base of paired fins; ventral fins never reach the anal fin insertion, and their average length varies from 72.8% to 75.7% of the distance between ventral and anal fin bases; there are large, more or less rounded, dark spots located along the lateral line and several rows of small dark spots on the dorsal and caudal fins; the eye is large with a diameter greater than ¾ of the interorbital distance; the breast in front of the level of the rear extent of the pelvic fin insertions usually lacks scales; the lateral branch of the supraorbital cephalic sensory canal (*CSO*) is connected with the infraorbital canal behind the eye; there are usually seven pores in the frontoparietal area of CSO and five pores in the pteroticum; both supra- and infraorbital bones are wide: the average width of the supraorbital bone exceeds 40% of its length, and the average width of the last infraorbitals comes to more than half of the bone length; $2n = 50$ (24 meta-, 24 submeta-, 2 subtelo-acrocentric chromosomes), NF = 98; the total number of vertebrae (according to Naseka, 2001)—40 (caudal, 19; preanal caudal, 2; abdominal, 21).

A.1.5. Other morphological features

Morphometric characters of gudgeons from the Sura River have been presented earlier by Ruchin and Naseka (2003). The variability of the relative length of paired fins among different populations from the Volga River basin, as well as the karyotype of specimens from the Yakot' River (Volga R. basin) were described by Vasil'eva et al. (2004). The craniological features and indices were demonstrated for gudgeons from the Yakot' R. (Vasil'eva et al., 2004; Vasil'eva and Kuga, 2005).

A.1.6. Distribution

According to our molecular data we have restricted the range of this species to the Volga River basin only. Its occurrence in neighbouring river systems needs further investigation.

A.1.7. Etymology

The name *volgensis* refers to the range of the species.

A.1.8. Comparative remarks

G. volgensis differs from most of the species previously included in *G. gobio* s. lato with the complex of features presented in the diagnosis, but, as mentioned previously, this species is very similar to *G. gobio* s. stricto in its external morphology. According to our preliminary study *G. gobio* (we examined several specimens of this species for comparison) differs due to the smaller average number of lateral line pored scales. The analysis of karyotypes presented by different authors for *G. gobio* s. lato reveals that the karyotype of *G.*

volgensis obviously differed from karyotypes of such species as *G. tauricus* Vasil'eva, 2005 and *G. kubanicus* Vasil'eva et Vasil'ev, 2004, but was quite similar to karyotypes obtained from gudgeons from the Odra R. basin, Lower Danube R. and Garonna R. (Hafez et al., 1978; Vujošević et al., 1983; Raicu et al., 1996; Kirtiklis et al., 2005). This result indicates karyological similarity between *G. gobio* and *G. volgensis*. Thus, the last taxon represents a cryptic species distinguishing from *G. gobio* only by molecular analysis and "S7indel diagnostics".

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2008.03.005.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Contr. 19, 716–723.

Bănărescu, P., 1961. Weitere systematische Studien über die Gattung *Gobio* (Pisces, Cyprinidae) insbesondere im Donaubecken. Vestn. Česko. Zool. Spol. 25 (4), 318–346.

Bănărescu, P., 1992a. A critical update checklist of Gobioninae (Pisces, Cyprinidae). Trav. Mus. Hist. Nat. Gr. Antipa. 32, 303–330.

Bănărescu, P., 1992b. Zoogeography of Freshwaters, vol. 2, Distribution and Dispersal of Freshwater Animals in North America and Eurasia. AULA-Verlag, Wiesbaden. pp. 517–1091.

Bănărescu, P., Nalbant, T.T., 1973. Pisces, Teleostei. Cyprinidae (Gobioninae). Das Tierreich 93, 1–304.

Bănărescu, P., Šorić, V.M., Economidis, P., 1999. *Gobio gobio*. In: Bănărescu, P. (Ed.), The Freshwater Fishes of Europe, 5/I. AULA-Verlag, Wiebelsheim, pp. 81–134.

Barker, F.K., Lutzoni, F.M., 2002. The utility of the incongruence length difference test. Systematic Biology 51 (4), 625–637.

Barriel, V., 1994. Molecular phylogenies and how to code insertion/deletion events. C. R. Acad. Sci. III 317, 693–701.

Berg, L.S., 1914. Fishes. Marsipobranchii and Pisces. V. III. Ostariophysi. 2. Petrograd (in Russian).

Berg, L.S., 1916. Fishes of Fresh Waters of Russian Impire. Moscow (in Russian).

Berg, L.S., 1949. Fishes of Fresh Waters of the USSR and Adjacent Countries. Nauka, Moscow-Leningrad (in Russian).

Bianco, P.G., Ketmaier, V., 2005. Will the Italian endemic gudgeon, *Gobio benacensis*, survive the interaction with the invasive introduced *Gobio gobio*? Folia Zool. 54 (Suppl. 1), 42–49.

Clement, M., Posada, D., Crandall, K., 2000. TCS: a computer program to estimate gene genealogies. Mol. Ecol. 9 (10), 1657–1660.

Cunha, C., Mesquita, N., Dowling, T.E., Gilles, A., Coelho, M.M., 2002. Phylogenetic relationships of Eurasian and American cyprinids using cytochrome *b* sequences. J. Fish Biol. 61, 929–944.

Cuvier, G., Valenciennes, A., 1842. Histoire naturelle des poissons. Tome seizième. Livre dix-huitième. Les Cyprinoïdes. 16.

Darlu, P., Lecointre, G., 2002. When does the incongruence length difference test fail? Mol. Biol. Evol. 19 (4), 432–437.

Doadrio, I., Madeira, M.J., 2004. A new species of the genus *Gobio* Cuvier, 1816 (Actinopterigii, Cyprinidae) from the Iberian Peninsula and southwestern France. Graellsia 60 (1), 107–116.

Erk'akan, F., Tatlidil, H., Özeren, S.C., 2005. Biometric investigation on *Gobio gobio* subspecies from Turkey. Folia Zool. 54 (Suppl. 1), 90–98.

Farris, J.S., Källerjö, M., Kluge, A.G., Bult, C., 1995. Testing significance of incongruence. Cladistics 10, 315–319.

Freyhof, J., Naseka, A.M., 2005. *Gobio delyamurei*, a new gudgeon from Crimean Peninsula (Cyprinidae, Gobioninae). Ichthyol. Explor. Freshwaters 16 (4), 331–338.

Froese, R., Pauly, D., (Eds.), 2007. FishBase. World Wide Web electronic publication. Available from: <www.fishbase.org>.

Gatesy, J., Milinkovitch, M., Waddell, V., Stanhope, M., 1999. Stability of cladistic relationships between Cetacea and higher-level Artiodactyl taxa. Syst. Biol. 48, 6–20.

Golubtsov, A.S., Malkov, N.P., 2007. Essay on the Fish Fauna of the Altai Republic: Systematic Diversity, Distribution and Conservation. KMK Scientific Press, Moscow (in Russian).

Grupče, R., Dimovski, A., 1975. Morfološko-sistematski proučuvanja vrz rodot *Gobio* (Pisces, Cyprinidae) vo Makedonija II. *Gobio gobio ohridanus* Kar. od Ohridskoto i Skadarskoto ezera [Morphological and systematic investigation of the genus Gobio (Pisces, Cyprinidae) in Macedonia II. *Gobio gobio ohridanus* Kar. from lakes Ohrid and Skadar]. Acta Musei Macedonici Scientarium Naturalium 14 (6), 145–169 (in Macedonian with French summary).

Hafez, R., Labat, R., Quillier, R., 1978. Etude cytogenetique chez quelques especes de cyprinides de la region Midi Pyrenees. Bull. Soc. Hist. Natur. Toulouse 144 (1/2), 122–159.

ICZN, 1999. International Code of Zoological Nomenclature, fourth ed. International Trust for Zoological Nomenclature, London.

Kessler, K., 1877. Fishes living and occurring in the Aral-Caspian-Pontic ichthyological region. Tr. Aralo-Caspian Exp. 2, 1–360.

Kirtiklis, L., Boroń, A., Porycka, K., 2005. Chromosome banding patterns of the gudgeon, *Gobio gobio* (Actinopterygii, Cyprinidae). Acta Ichthyol. Pisc. 35 (2), 119–123.

Kottelat, M., 1997. European freshwater fishes. Biologia, Bratislava 52 (Suppl. 5), 1–271.

Kottelat, M., Freyhof, J., 2007. Handbook of European Freshwater Fishes. Kottelat, Cornol, Switzerland and Freyhof, Berlin, Germany.

Kottelat, M., Persat, H., 2005. The genus *Gobio* in France, with redescription of *G. gobio* and description of two new species (Teleostei: Cyprinidae). Cybium 29 (3), 211–234.

Lavoué, S., Sullivan, J.P., Hopkins, C.D., 2003. Phylogenetic utility of the first two introns of the S7 ribosomal protein gene in African electric fishes (Mormyroidea: Teleostei) and congruence with other molecular markers. Biol. J. Linnean Soc. 78, 273–292.

Liu, H., Tzeng, C.-S., Teng, H.-Y., 2002. Sequence variations in the mtDNA control region and their implications for the phylogeny of the Cypriniformes. Can. J. Zool. 80, 569–581.

Lukash, B.S., 1923. Fishes of the Vychegda River (Zyryanskyi krai). Sever 3, 163–176.

Lukash, B.S., 1933. Fishes of the lower section of the Vyatka River. Tr. Vyatskogo nauchno-issledov. Inst. Kraeveden. 6, 5–110 (in Russian).

Lukash, B.S., 1940. Fishes of Kirov district. Tr. Kirovskogo Obl. Nauchno-issledov. Inst. Kraeveden. 14, 1–72 (in Russian).

Madeira, M.J., Gómez-Moliner, B.J., Doadrio, I., 2005. Genetic characterization of *Gobio gobio* populations of the Iberian Peninsula based on cytochrome *b* sequences. Folia Zool. 54 (Suppl. 1), 5–12.

Müller, K., 2005. SeqState-primer design and sequence statistics for phylogenetic DNA data sets. Appl. Bioinformatics 4, 65–69.

Müller, K., 2006. Incorporating information from length-mutational events into phylogenetic analysis. Mol. Phylogenet. Evol. 38, 667–676.

Naseka, A.M., 1998. Subfamily Gobioninae. In: Reshetnikov, Yu.S. (Ed.), An Annotated Catalog of Cyclostomata and Fishes of Russia. Nauka, Moscow, pp. 81–87 (in Russian).

Naseka, A.M., 2001. Redescription of the white-fin gudgeon *Romanogobio albipinnatus* (Cypriniformes, Gobioninae), with comments on its taxonomic position. Zoolog. Zh. 80 (11), 1372–1383 (in Russian with English summary).

Naseka, A.M., Erk'akan, F., Kücük, F., 2006. A description of two new species of the genus *Gobio* from Central Anatolia (Turkey) (Teleostei: Cyprinidae). Zoosyst. Rossica 15, 185–194.

Ogden, T.H., Rosenberg, M.S., 2007. How should gaps be treated in parsimony? A comparison of approaches using simulation. Mol. Phylogenet. Evol. 42, 817–826.

Pallas, P.S., 1814. Zoographia Rosso-Asiatica, sistens omnium animalium in extenso Imperio Rossico et adjacentibus maribus observatorum recensionem, domicilia, mores et descriptiones, anatomen atque icones plurimorum, vol. 3. Petropoli.

Posada, D., 2006. ModelTest Server: a web-based tool for the statistical selection of models of nucleotide substitution online. Nucleic Acids Res. 34, Web Server issue. Available from: <http://darwin.uvigo.es/software/modeltest_server.html>.

Raicu, P., Taicescu, E., Bănărescu, P., 1973. A comparative study of the karyotype in the genus *Gobio* (Pisces, Cyprinidae). Cytologia (Tokyo) 38, 731–736.

Rambaut, A., Drummond, A., 2003. Tracer v 1.2. Available from: <http://evolve.zoo.ox.ac.uk/>.

Ronquist, F., Huelsenbeck, J.P., 2005. Bayesian analysis of molecular evolution using MrBayes. In: Nielsen, R. (Ed.), Statistical Methods in Molecular Evolution. Springer, New York.

Ruchin, A., Naseka, A.M., 2003. Morphological characteristic of two gudgeons sympatric in the Sura River (Mordovia). Vopr. Ikhtiol. 43 (3), 334–337 (in Russian).

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning. A laboratory Manual, second ed. Cold Spring Harbor Laboratory Press.

Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequencebased phylogenetic analyses. Syst. Biol. 49, 369–381.

Simmons, M.P., Ochoterena, H., Carr, T.G., 2001. Incorporation, relative homoplasy, and effect of gap characters in sequence-based phylogenetic analyses. Syst. Biol. 50 (3), 454–462.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis Using Parsimony (* and other methods). Sinauer Associates, Sunderland, MA.

Šanda, R., Lusková, V., Vukic, J., 2005. Notes on the distribution and taxonomic status of *Gobio gobio* from the Morača River basin (Montenegro). Folia Zool. 54 (Suppl. 1), 73–80.

Šorić, V., 1990. Ichthyofauna of the Ohrid-Drim-Skadar system. Ichthyologia 22 (1), 31–43.

Šorić, V., Ilic, K.R., 1988. Vrste roda *Gobio* II—*Gobio gobio ohridanus* u jadranskom, Ohrid-Drim-Skadar sistemu i *Gobio gobio obtusirostris* u crnomorskom i egejskom sistemu [Species of the genus *Gobio* II—*Gobio gobio ohridanus* in the Adriatic, Ohrid-Drim-Skadar system and *Gobio gobio obtusirostris* in the Black Sea and Aegean systems]. Ichthyologia 20 (1), 47–60 (in Serbian with English summary).

Templeton, A.R., Crandall, K.A., Sing, C.F., 1992. A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation.. Genetics 132, 619–633.

Thai, B.T., Burridge, C.P., Pham, T.A., Austin, C.M., 2004. Using mitochondrial nucleotide sequences to investigate diversity and genealogical relationships within common carp (*Cyprinus carpio* L.). Anim. Genet. 36, 23–28.

Tsepkin, E.A., 2002. *Gobio gobio* (Linnaeus, 1758)—gudgeon. In: Reshetnikov, Yu.S. (Ed.), Atlas of Freshwater Fishes of Russia, 1. Nauka, Moscow, pp. 249–251 (in Russian).

Vasil'eva, E.D., Kuga, T.I., 2005. Craniological features as available diagnostic characters for the short-barb Crimean gudgeon *Gobio krymensis* (Gobioninae, Cyprinidae). Folia Zool. 54 (Suppl. 1), 33–41.

Vasil'eva, E.D., Vasil'ev, V.P., Boltachev, A.P., 2005. Taxonomic relationships among gudgeons (*Gobio*, Gobioninae, Cyprinidae) of Crimea. Vopr. Ikhtiol. 45 (6), 768–781 (in Russian).

Vasil'eva, E.D., Vasil'ev, V.P., Kuga, T.I., 2004. On taxonomy of gudgeons of the genus *Gobio* (Gobioninae, Cyprinidae) of Europe: a new gudgeon species *Gobio kubanicus* from the basin of the Kuban River. J. Ichthylol. 44 (6), 716–731.

Vladykov, V., 1925. Über einige neue Fische aus der Tschechoslowakei (Karpathorussland). Zool. Anz. (Jena) 64 (11/12), 248–252.

Vladykov, V., 1931. Les poissons de la Russie Sous-Carpathique (Tchécoslovaquie). Mém. Soc. Zool. (France) 29 (4), 217–374.

Vujošević, M., Žizković, S., Rimsa, D., Jurišić, S., Cakić, P., 1983. The chromosomes of 9 fish species from Dunav basin in Yugoslavia. Acta Biol. Jug. Ichthyologia 15 (2), 29–40.

Yang, J., He, S., Freyhof, J., Witte, K., Liu, H., 2006. The phylogenetic relationships of the Gobioninae (Teleostei: Cyprinidae) inferred from mitochondrial cytochrome *b* sequences. Hydrobiologia 553, 255–266.