Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 46 (2008) 738–759
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Phylogeny of *Paramysis* (Crustacea: Mysida) and the origin of Ponto-Caspian endemic diversity: Resolving power from nuclear protein-coding genes

Asta Audzijonyte [a,*], Mikhail E. Daneliya [a], Nikolai Mugue [b], Risto Väinölä [a]

[a] *Finnish Museum of Natural History, POB 17, FI-00014, University of Helsinki, Helsinki, Finland*
[b] *Institute of Developmental Biology, Vavilov Street 26, Moscow 119991, Russia*

Received 6 July 2007; revised 1 November 2007; accepted 15 November 2007
Available online 9 January 2008

Abstract

The Ponto-Caspian (Black and Caspian seas) brackish-water fauna represents a special case of the endemic diversification in world's ancient lakes; it also involves a hotspot of continental diversity in the predominantly marine mysid crustaceans. We explored the origins and history of the mysid diversification in a phylogenetic analysis of some 20 endemic Ponto-Caspian species mainly of the genus *Paramysis* and their marine congeners, using sequences of two nuclear protein-coding genes, two nuclear rRNA genes, the mitochondrial COI gene and morphological data. A nearly completely resolved phylogeny was recovered, with no indication of rapid diversification bursts. Deep divergences were found among the main endemic clades, attesting to a long independent faunal history in the continental Paratethys waters. The current marine *Paramysis* species make a monophyletic cluster secondarily derived from the continental Paratethyan (Ponto-Caspian) *Paramysis* ancestors. The good phylogenetic resolution was mainly due to the two nuclear protein-coding genes, opsin and EPRS, here for the first time applied to peracarid systematics. In contrast, 'conventional' mtDNA and nuclear rRNA genes provided poor topological resolution and weak congruence of divergence rates. The two nuclear protein-coding genes had more congruent rates of evolution, and were about 10–15 times slower than the mitochondrial COI gene.

Keywords: Caspian sea; COI; EPRS; Molecular clock; Opsin; RNA; Speciation; Vicariance

1. Introduction

The origin of endemic diversity in ancient lakes is a question related to several contentious topics in evolutionary biology, including the importance of sympatric speciation and the roles of long-term environmental stability versus instability in promoting diversification. Ancient lakes are usually defined as long-lived lakes (typically older than 100 KY) that harbour endemic species flocks (Brooks, 1950; Martens, 1997). While lakes Baikal and Tanganyika are the best known ancient lakes, the largest one by these criteria is the Caspian Sea (which in general is the largest lake on Earth), whose fauna is 50–80% endemic (Mordukhai-Boltovskoi, 1979; Dumont, 1998). The Caspian Sea currently has no outlet and is naturally brackish. Yet unlike other ancient lakes that remained isolated from oceans throughout most of their existence and have origins of their endemism in freshwater lineages (Schön and Martens, 2004), the Caspian has undergone complex paleogeographic and hydrographic evolution, inextricably linked with that of the adjacent water bodies—the Black and Mediterranean seas. In the course of recurrent isolations and connections to the world ocean, the endemic fauna of the Black and Caspian seas, which together constitute the Ponto-Caspian basin, has had multiple links with related marine lineages. The timing and direction of these marine–continental faunal connections are key questions to understanding the evolution of the Ponto-Caspian diversity.

* Corresponding author. Fax: +358 9 19128888.
E-mail address: audzi@mbari.org (A. Audzijonyte).

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.11.009

The history of the Ponto-Caspian has involved extensive environmental perturbations caused by both tectonic and climatic changes. The Black and Caspian seas are descendants of the Paratethys epicontinental sea basin which gradually became isolated from the Mediterranean in the Late Miocene times. Since then during the last several million years, the salinity in the component water bodies has varied from nearly oceanic to almost freshwater and water levels have fluctuated over several hundred metres (Reid and Orlova, 2002; Popov et al., 2006). Numerous extinctions and appearances of endemic lineages are correspondingly known from the fossil record (reviewed by Nevesskaya et al., 2006), but it remains controversial whether local faunas have been completely wiped out during adverse conditions or survived in local refugia (Popov et al., 2006). The idea of ancient origins and long-term survival is supported by Caspian Sea endemism that is manifested not only at the species but also the genus level, and by the pronounced morphological differentiation in some taxa (Derzhavin, 1939; Zenkevitch, 1963; Dumont, 1998, 2000). On the other hand, the hypothesis is challenged by the absence of brackish-water mollusc taxa from Caspian Pliocene strata (5–2 MY) (Mordukhai-Boltovskoi, 1979; Tarasov, 1997; Grigorovich et al., 2003). Similarly, extinction of brackish-water fauna has been claimed for the freshened Late Pleistocene Black Sea (15 KY), suggesting that the current diversity in that basin largely originates from a post-glacial reinvasion from the Caspian (Mordukhai-Boltovskoi, 1979). Yet the Black Sea also harbours a number of endemic brackish-water species that are not found in the Caspian.

In trying to understand the history of Ponto-Caspian endemism, two major challenges are (i) to determine the primary sources and numbers of founding endemic lineages (i.e., single versus multiple colonisations in a given group of organisms), and (ii) to reveal the age of these colonisation and radiation events. From well-resolved molecular trees we should be able to infer the relationships between the Ponto-Caspian endemics and their extra-basin relatives, and indicate the direction and minimum number of colonisation events in or out of the region. Despite the inherent inaccuracy of molecular clocks, the depth of molecular divergence should be able to discriminate competing hypotheses of Late Miocene (10–6 MY) versus Pleistocene (<2 MY) diversification. Phylogenies based on sufficient unlinked molecular characters should also demonstrate whether speciation events in the Ponto-Caspian region occurred concurrently in bursts, resulting in bush-like and largely unresolved trees, or whether the radiation took place at a more even pace, leaving clear traces of the branching order in character distributions. To date, molecular studies of the endemic Ponto-Caspian fauna are rare and have dealt with few species. Nonetheless these studies have suggested both the presence of ancient lineages, in bivalve molluscs and amphipods (Therriault et al., 2004; Cristescu and Hebert, 2005), and fast recent radiations in cladocerans and mysids (Väinölä, 1995; Cristescu and Hebert, 2002; Audzijonyte et al., 2005).

The Ponto-Caspian mysid crustaceans (Malacostraca: Mysida) are one of the diversified and highly endemic groups of the Ponto-Caspian basin. They also represent a unique concentration of continental (non-marine) species diversity in a predominantly marine order, whose history and evolution from the crustacean zoogeography point of view are intriguing in themselves (Ekman, 1953; Bănărescu, 1991). As an exemplary case for study, the total list of 25 autochthonous Ponto-Caspian brackish-water mysid species makes a suitable number for a complete taxonomic assessment of the group. The group also involves varying levels of taxonomic and morphological differentiation: apart from the speciose and dominant genus *Paramysis* Czerniavsky, 1882, three morphologically distinct but probably phylogenetically close monotypic endemic genera *Caspiomysis* G.O. Sars, 1907, *Katamysis* G.O. Sars, 1893 and *Limnomysis* Czerniavsky, 1882 are recognised. Further the Ponto-Caspian fauna includes species with different regional or ecological distributions, such as a number of strictly Caspian endemics confined to the brackish central and southern Caspian, Black Sea endemics, and a true Ponto-Caspian element shared by the Caspian and Black Sea basins and distributed in their rivers and estuaries (Fig. 1 and Table 1). Finally, the genus *Paramysis* also contains seven more species from the Mediterranean Sea and Northeast Atlantic, and the phylogenetic relationship of these truly marine species to the Ponto-Caspian taxa has been a matter of continued debate (Czerniavsky, 1882; Sars, 1907; Derzhavin, 1939).

In this study, we explore the zoogeographical and evolutionary history of Ponto-Caspian diversity in a phylogenetic analysis of the genus *Paramysis* and the three associated monotypic endemic genera *Caspiomysis*, *Katamysis* and *Limnomysis*. For this purpose we use a morphological data matrix and a multi-gene molecular data set comprising parts of two nuclear protein-coding genes, two rRNA genes and one mitochondrial gene. We use these data to assess the relationships among the species and to reconstruct the dynamics and geography of their colonisation history and their endemic radiation in the framework of the complex Ponto-Caspian geological history. In particular we test the following controversies and hypotheses: (i) multiple versus single origins of the Ponto-Caspian endemics and their relation to marine congeners—are the Ponto-Caspian endemics monophyletic, are the oceanic *Paramysis* species basal to the Ponto-Caspian taxa or secondarily derived from the Ponto-Caspian clade(s), and do the monotypic genera represent offshoots from endemic *Paramysis* or independent colonising lineages? (ii) The age of the Ponto-Caspian endemism—is there evidence for survival of ancient lineages from the early Paratethys (Miocene, see Fig. 1b) or did speciation mostly occur during the Pleistocene? (iii) The relative phylogenetic information contents and congru-

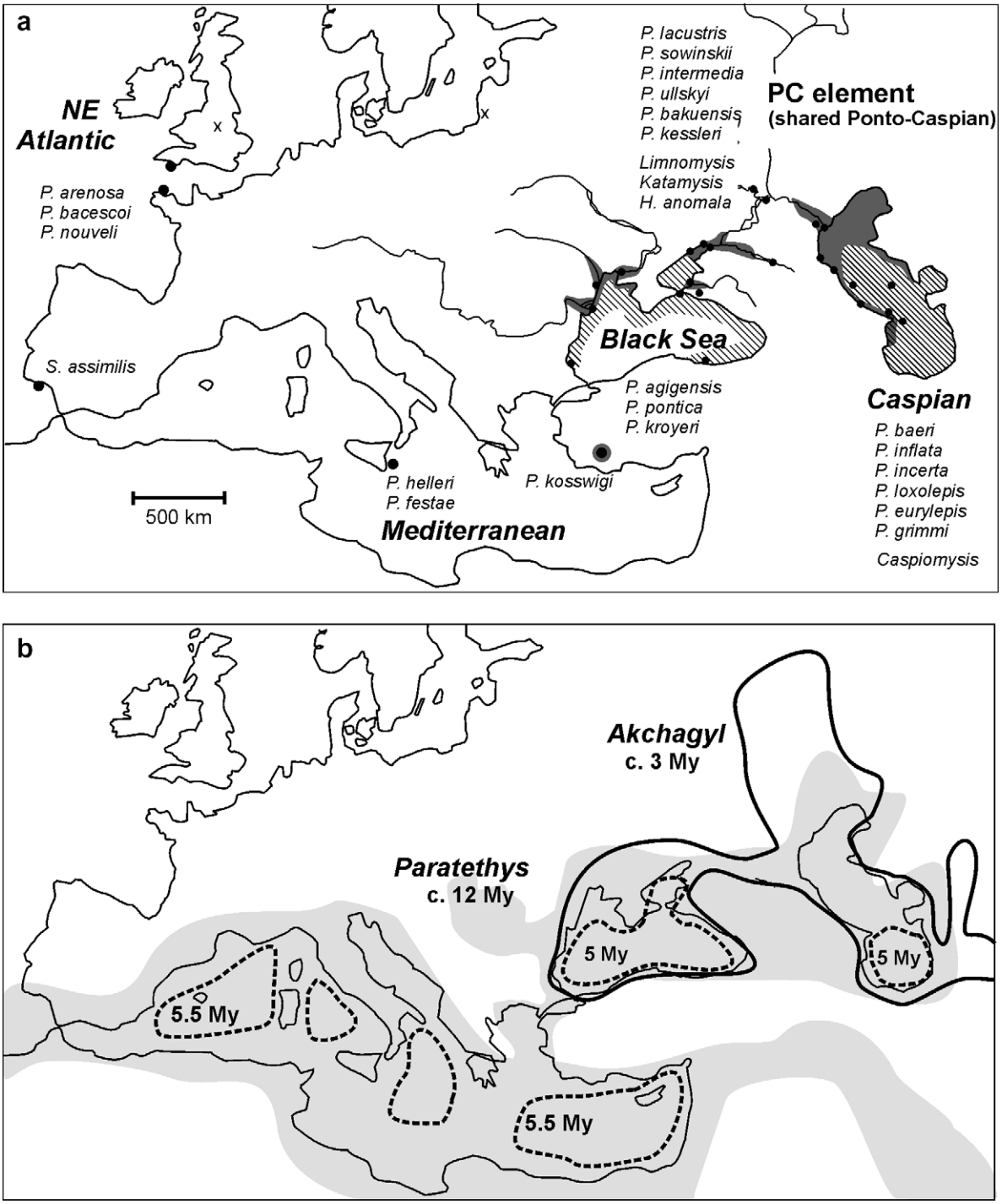

Fig. 1. (a) Map with sampling sites (black dots) and indication of the general distribution of the zoogeographical groups addressed and species studied of each of them: shared Ponto-Caspian element (dark shading), Caspian Sea endemics (fine hatching), and Black Sea endemics (coarse hatching). For exact sampling locations see Electronic Supplement I. (b) Schematic presentation of some major stages in the history of the Mediterranean, Black and Caspian seas (according to Starobogatov, 1970; Popov et al., 2006). The grey shading shows the Late Miocene situation with the semi-isolated Paratethys Sea (Sarmatian) north of the Mediterranean. Dashed lines depict limits of the water bodies during Pliocene regression stages. The thick line shows approximate extent of the Akchagyl transgression c. 3 MY ago.

ence of divergence time estimates from different genes, and correspondence between morphological and molecular data. In particular, we contrast the information contents of the conventionally used nuclear rRNA genes (18S, 28S) and mtDNA COI gene to that of the here introduced nuclear protein-coding genes.

2. Materials and methods

2.1. Taxonomic sampling

A total of 27 species were included in the phylogenetic analyses, of which six were studied for morpholog-

Table 1
Taxa studied (subgeneric names according to Daneliya, 2004) listed by zoogeographical groups, with GenBank accession numbers of sequences used in the phylogenetic analysis; for the complete list of localities and specimens examined see Electronic Supplements I and II

Zoogeographic group	GenBank numbers for sequences used in the phylogenetic analysis					
Species (subgenus) author	COI	OPS	EPRS	18S[a]	28S	MOR
Caspian Sea endemics (Central and Southern Caspian)						
P. (*Paramysis*) *baeri* Czerniavsky, 1882	DQ779858*	EU233595	EU233555	EU233479, -497	EU233516	4
P. (*Metamysis*) *inflata* (G.O. Sars, 1907)	EU233537	EU233601	EU233556	EU233480, -498	EU233517	3
P. (*Serrapalpisis*) *incerta* (G.O. Sars, 1895)	EU233538	EU233602	EU233557	EU233481, -499	EU233518	3
P. (*Nanoparamysis*) *loxolepis* (G.O. Sars, 1895)	EU233539	EU233610	EU233558	EU233482, -500	EU233519	3
P. (*Paramysis*) *eurylepis* G.O. Sars, 1907	—	—	—	—		2
P. (*Metamysis*) *grimmi* G.O. Sars, 1895	—	—	—	—	—	3
Caspiomysis knipowitschi G.O. Sars, 1907	EU233543	EU233622	EU233559	—	EU233520	2
Shared PC element (Ponto-Caspian rivers and estuaries)						
P. (*Paramysis*) *bakuensis* G.O. Sars, 1895	DQ779863*	EU233592	EU233560	EU233483, -501	EU233521	7
P. (*Paramysis*) *kessleri* G.O. Sars, 1895	DQ779848*	EU233597	EU233562	EU233484, -502	EU233522	5
P. (*Metamysis*) *ullskyi* (Czerniavsky, 1882)	DQ779871*	EU233599	EU233565	EU233485, -503	EU233523	4
P. (*Serrapalpisis*) *lacustris* (Czerniavsky, 1882) CB	DQ779823*	EU233603	EU233568	—	—	6
P. (*Serrapalpisis*) *lacustris* (Czerniavsky, 1882) AZ	DQ779841*	EU233604	EU233569	EU233486, -504	EU233524	5
P. (*Serrapalpisis*) *sowinskii* Daneliya, 2004	DQ779812*	EU233606	EU233570	EU233487, -505	EU233525	3
P. (*Mesomysis*) *intermedia* (Czerniavsky, 1882)	DQ779852*	EU233608	EU233572	EU233488, -506	EU233526	3
Hemimysis anomala G.O. Sars, 1907	EU029170*	EU233620	EU233575	AJ566104	EU233527	2
Katamysis warpachowskyi G.O. Sars, 1893	—	EU233617	EU233577	EU233489, -507	EU233528	2
Limnomysis benedeni Czerniavsky, 1882	DQ779792*	—	EU233579	EU233490, -508	EU233529	3
Black Sea endemics						
P. (*Longidentia*) *kroyeri* (Czerniavsky, 1882)	EU233544	EU233612	EU233582	EU233491, -509	EU233530	2
P. (*Occiparamysis*) *agigensis* Băcescu, 1940	EU233545	EU233611	EU233583	EU233492, -510	EU233531	2
P. (*Pseudoparamysis*) *pontica* (Băcescu, 1940)	—	—	—	—	—	1
Mediterranean, NE Atlantic						
P. (*Longidentia*) *nouveli* Labat, 1953 comb. nov.	EU233546	EU233613	EU233584	EU233493, -511	EU233532	3
P. (*Pseudoparamysis*) *bacescoi* Labat, 1953 comb. nov.	EU233550	EU233615	EU233586	EU233494, -512	EU233533	2
P. arenosa (G.O. Sars, 1877)	DQ779811	EU233616	EU233588	EU233495, -513	EU233534	2
P. (*Longidentia*) *helleri* (G.O. Sars, 1877) comb. nov.	—	—	—	—	—	1
P. festae Colosi, 1922	—	—	—	—	—	1[c]
Schistomysis assimilis (G.O. Sars, 1877)	EU233554	EU233621	EU233590	EU233496, -514	EU233535	2
Other						
P. (*Serrapalpisis*) *kosswigi* Băcescu, 1948 comb. nov.	—	—	—	—	—	2
Mysis segerstralei Audzijonyte and Väinölä, 2005	DQ524909	EU233619	EU233591	DQ189132 EU233515[b]	EU233536	1

MOR column indicates the number of completely dissected specimens, asterisk (*) denotes that a larger number of specimens has been examined for the COI gene in previous studies (Audzijonyte et al., 2006, 2008). All specimens with GenBank numbers were also inspected for morphological characters; species with no GenBank numbers were studied only for morphology.

[a] The 18S gene was sequenced in two non-overlapping parts, each given one GenBank number.

[b] The second half of the 18S sequence of *M. segerstralei* was substituted by the closely related *Mysis oculata* (EU233515) (see Section 2).

[c] Scored from an intact specimen.

ical characters only (Table 1 and Fig. 1). The 21 species analysed with both molecular and morphological characters included 12 Ponto-Caspian *Paramysis* species representing all the eight subgenera defined by Daneliya (2004) (Table 1), three further Mediterranean–Atlantic *Paramysis* species, and three Ponto-Caspian endemic monotypic genera *Caspiomysis*, *Katamysis* and *Limnomysis*. In addition we included a representative from each of two morphologically related genera *Hemimysis* and *Schistomysis*, and an outgroup species *Mysis segerstralei*. The three latter genera are all primarily marine, but each of them also contains species restricted to the Ponto-Caspian (of which only *H. anomala* is studied here). Note that throughout the paper, the species ranges are discussed in terms of natural distributions only, excluding range expansions due to recent invasions and intentional introductions.

Most samples were collected in 2000–2006, and stored in 80–96% ethanol, or deep-frozen and stored at −80 °C. For species with no fresh samples available, museum specimens were used for morphology (Table 1). In all, the data included specimens from 27 localities (Electronic Supplements I and II). Based on the previous finding of two relatively deep mtDNA lineages within *P. lacustris* (5% COI divergence, Audzijonyte et al., 2006), we included representatives from both clades as different OTUs, and refer to them as *P. lacustris* CB (from the Caspian and Black Seas) and *P. lacustris* AZ (from the Sea of Azov).

For the presentation and discussion, the taxa studied here are divided into four zoogeographical groups (Table

1 and Fig. 1): (i) species endemic to the Caspian Sea; (ii) truly Ponto-Caspian species—or the PC element—present in the estuaries and rivers of both the Caspian and Black Sea basins; (iii) species endemic to the Black Sea; (iv) taxa from the Mediterranean and NE Atlantic waters (i.e., non-Ponto-Caspian).

2.2. Morphological data

The morphological data set included 35 external characters which were deemed stable within species and amenable to coding into discrete states (Appendix A). The characters were assessed from 2 to 9 specimens of each species, dissected, mounted on slides and examined using light microscopy. Characters of *P. festae* were scored only from intact specimens. All individuals studied here with molecular characters were also inspected morphologically for some of the characters. The morphological data matrix included both characters traditionally used in systematics of *Paramysis* and related genera, e.g., characters of carapace, antennal scales, maxillae, pereiopods, telson and uropods (Sars, 1907; Băcescu, 1954; Komarova, 1991), and others not used in taxonomy before, i.e., those of antennulae, mandibular palps, maxillipedes, penis and pleopods. Twenty-one characters were scored with multiple states, and in parsimony analyses were treated as unordered to minimise *a priori* assumptions on character state changes. All morphological data were produced by one person (MED).

2.3. DNA sequence data

We analysed partial sequences of two protein-coding nuclear genes, the long-wave opsin (OPS) and glutamyl- and prolyl-tRNA synthetase (EPRS); two nuclear ribosomal RNA genes (18S, 28S); and the mitochondrial cytochrome oxidase subunit I gene (COI). One to four specimens per species were sequenced for each gene; for the mitochondrial COI gene a larger number of individuals (5–60) from seven of the species have been analysed in a previous study (Audzijonyte et al., 2006). Information for molecular data used in this study and GenBank accession numbers are given in Table 1 and in Electronic Supplement II. The DNA extraction, amplification and sequencing procedures were as described in Audzijonyte et al. (2005). Primers and PCR conditions are listed in the Electronic Supplement III. Generally, different genes were sequenced from the same specimen, or from specimens from the same locality. We were unable to obtain COI data from *K. warpachowskyi*, 18S from *C. knipowitschi*, and OPS from *L. benedeni*, and therefore coded these genes as missing data in the respective species in the simultaneous phylogenetic analyses. 18S data were not obtained for the outgroup *M. segerstralei* and was substituted by data from the closely related *M. oculata*. The close relationship of these two species and the coherence of the *Mysis* genus are strongly supported (Audzijonyte et al., 2005) and besides, the first half of the gene, sequenced for both *Mysis* species, had only two nucleotide differences between them. Only one sequence of each of 18S and 28S was obtained for *P. lacustris*, and was used for both CB and AZ OTUs.

2.4. Data analyses

2.4.1. Molecular diversity

The translated protein-coding gene data did not include any stop codons to indicate presence of pseudogenes. Mutational saturation of substitutions and relative amount of divergence among different genes was assessed visually by plotting model-corrected or uncorrected distances between genes. Because sequences of the OPS gene were longer than of the other two protein-coding genes and the gene was less affected by mutation saturation, it was used as a basis for comparisons. Given the deep molecular divergences among the analysed genera, only *Paramysis* species were included in these comparisons (strongly diverged *P. arenosa* was also excluded, see Section 3). The model parameters for correcting the mutation saturation were inferred in a Bayesian phylogenetic analysis of *Paramysis* species and one outgroup only (same as used in the penalized likelihood analysis, see below) and are listed in Table 2. To assess functional (selective) constraints on the protein-coding genes we calculated ratio of non-synonymous to synonymous substitutions per non-synonymous and synonymous sites (dN/dS) in Mega3 (Kumar et al., 2004).

2.4.2. Alignment and phylogenetic analyses

Alignment of the protein-coding sequences (COI, OPS, EPRS) was straightforward, as sequences contained no length variation, except for a 3 bp deletion in EPRS of *L. benedeni*. Ribosomal RNA genes (18S and 28S) were more variable in length and were aligned using two approaches. First, we inferred the secondary structure of the studied 18S rRNA region. For that purpose, sequences were aligned to *Saccharomyces cerevisiae* (GenBank Accession No. J01353) for which secondary rRNA structure has been defined. Stem and loop structure of mysid sequences was then inferred according to that in yeast, and the 2D model of the molecule was drawn with RnaViz software (De Rijk and De Watcher, 1997) using the predefined *Saccharomyces* skeleton file. The obtained 2D structure was further modified manually, if additional pairing could be inferred from complementary substitutions. As a reference we also used RNA structures from the European ribosomal RNA database (http://www.psb.ugent.be/rRNA/index.html) and 18S secondary structure predictions published in Meland and Willassen (2007). The inferred secondary structure for *Hemimysis anomala* is available in Electronic Supplement IV, and approximate annotation in other taxa is shown in the Electronic Supplement V. The inferred secondary structure however did not aid alignment of loop regions, where most of the variation was observed. Therefore, we also applied automatic alignment procedures, using both ClustalW software (Thomson et al., 1994) with default parameters, and an iterative alignment procedure

Table 2
Substitution model parameters used in the maximum likelihood (ML) and estimated through Bayesian inference (BI) for the different molecular data partitions

	ML			BI			BI*	
	PRO	PRN	RNA	COI	PRN	RNA	COI	PRN
r(AC)	1.0	1.9	1.0	0.60 ± 0.31	1.84 ± 0.18	1.04 ± 0.12	0.33	1.85
r(AG)	3.6	5.0	2.1	6.25 ± 1.17	3.89 ± 0.28	2.14 ± 0.16	7.78	4.80
r(AT)	1.0	1.7	2.1	0.51 ± 0.19	1.52 ± 0.14	2.18 ± 0.16	1.00	1.72
r(CG)	1.0	1.1	0.5	0.89 ± 0.32	1.17 ± 0.18	0.53 ± 0.09	0.16	1.12
r(CT)	5.3	5.0	4.1	23.01 ± 1.37	4.90 ± 0.30	3.90 ± 0.21	11.96	6.18
r(GT)	1.0	1.0	1.0	1.00 ± 0.23	1.00 ± 0.13	1.00 ± 0.11	1.00	1.00
I	0.4	0.6	0	0.38 ± 0.03	0.42 ± 0.03	0.41 ± 0.12	0.58	0.60
α	0.7	6.0	0.4	0.25 ± 0.02	1.70 ± 0.36	15.21 ± 36.2	10.78	2.00

For BI, means and standard deviations are given, with first 100,000 generation discarded. BI*, model parameters estimated in the BI of protein-coding data of *Paramysis* species only, and applied to correct for mutational saturation in pairwise molecular distance comparisons (Fig. 2). PRO, all protein-coding data (COI, EPRS, OPS); PRN, nuclear protein-coding data (EPRS, OPS); RNA, rRNA data (18S, 28S). Inferred nucleotide frequencies were very similar to the empirical frequencies and are not presented. The parameter values used in ML analyses correspond to models: TVM + I + Γ for PRN data (nearly identical results in terms of topology and bootstrap support were obtained with a simpler but less fitting model HKY + Γ); TRN + I + Γ for PRO data, and GTR + Γ for RNA data.

implemented in program Muscle (Edgar, 2004). The two software packages gave nearly identical results in the relatively conservative parts of the gene, but results of either program were not judged reliable in several areas that included large insertions. The most length-variable parts of 18S and 28S were therefore excluded from the phylogenetic analyses. When the alignment was ambiguous only in a few taxa, these were coded as missing for that particular region. It is noted that the species that deviated greatly in rRNA were also strongly diverged in other genes, and their phylogenetic position had good support from non-rRNA data. Therefore, excluding parts of rRNA genes would not affect the inference of topology. The complete alignment and the data set used in the analysis are given in the Electronic Supplements V and VI.

Phylogenetic analyses were performed on the total data set, on the morphological data and on the combined molecular data set separately, on each gene separately, and on three combinations of the genes (or partitions of the molecular data set): "RNA", including the two rRNA genes; "PRO", including all three protein-coding genes (OPS, EPRS, COI), and "PRN", including the two nuclear protein-coding genes only (OPS, EPRS). As COI showed extensive saturation at 3rd codon positions, analyses were also conducted with these positions excluded. All data sets were analysed under maximum parsimony (MP), the RNA, PRO, PRN data sets were analysed using maximum likelihood (ML), and the combined morphological and molecular data using Bayesian inference (BI).

The MP analyses were conducted using PAUP* 4.0b10 (Swofford, 2003), NONA (Goloboff, 1993) and Winclada (Nixon, 2002). In PAUP heuristic searches involved 10–100 replications with random taxon addition order followed by tree bisection reconnection (TBR) branch swapping; nodes were collapsed if minimum branch lengths were zero (by default in PAUP* they are collapsed if maximum possible branch lengths are zero). More exhaustive searches conducted in NONA resulted in the same topology. Gaps were treated as a 5th character state, all characters were unordered. Branch support was assessed by non-parametric bootstrapping (1000 pseudoreplicates, 10 random addition replicates for each bootstrap replicate). For ML analyses, the best-fit model for each of the data sets was selected from nested models using likelihood ratio-tests (significance level $\alpha = 0.01$) in Modeltest (Posada and Crandall, 1998) (Table 2). To minimise a selection bias towards parameter rich models caused by the order of model input (Posada and Crandall, 2001) we manually calculated likelihood ratio-test statistics for all pairs of nested models. Heuristic searches were performed using 10 random taxa additions followed by TBR branch swapping; bootstrap support was assessed from 250 to 500 bootstrap replicates (three random addition replicates for each bootstrap replicate).

Simultaneous model-based phylogenetic analysis of the entire data set was conducted using a Bayesian approach as implemented in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). Data were subdivided into four partitions: the mtDNA COI gene, PRN, RNA and morphology. Molecular data were analysed with the GTR + I + Γ model, but the substitution matrix and rate variability across sites (proportion of invariable sites I and the α-parameter of the gamma distribution) were allowed to differ for each data partition (were unlinked). We used GTR + I + Γ model, because likelihood ratio tests of hierarchical model comparisons showed that six nucleotide substitution types (GTR) and rate variation across sites described the data significantly better than the alternative two substitution type (HKY) model available in MrBayes or uniform rate models. However, prior to the analysis we checked that all possible nucleotide substitutions did exist in the data, as inference of model parameters may be biased if some of the substitutions are absent (Ronquist et al., 2005). For morphological data we used the default model implemented in MrBayes, which for each character assumes equal (symmetrical, unordered) change probabilities among the character states observed. Analyses in MrBayes were conducted with two parallel runs of two million generations each, using four chains, as set by default.

The standard deviation of split frequencies (SDSF) after 2 million generations was 0.012. Longer runs did not reduce the SDSF value, yet the model parameters from the two parallel runs converged with <4% (in most cases <2%) difference. Likelihoods of Bayesian analyses converged after the initial 100,000 generations; we therefore discarded initial 1000 trees (representing 100,000 generations) from each of the two parallel runs and computed posterior probabilities from the remaining 38,002 trees. Estimated parameters are shown in Table 2.

Given that results of phylogenetic analysis suggested paraphyly of *Paramysis* genus, we compared the obtained optimum MP and ML trees with those constrained for the monophyly of *Paramysis*. We compared MP trees from simultaneous analysis of all data using the non-parametric (Templeton and winning-sites) and Kishino and Hasegawa (1989) tests implemented in PAUP; the constrained and unconstrained ML trees obtained by PAUP from the PRO, PRN and RNA data were compared using the AU (approximately unbiased) test implemented in the Consel package (Shimodaira and Hasegawa, 2001; Shimodaira, 2002).

The contributions of individual data partitions to the simultaneous analysis topology in the MP analysis and possible incongruence among the data sets were assessed by comparing results from separate and combined data analyses, and through partitioned Bremer support (PBS) and partitioned hidden Bremer support values (PHBS) (Baker and DeSalle, 1997; Gatesy et al., 1999). Negative PBS values indicate conflict of the data partition with the topology of the simultaneous analysis, whereas PHBS indicates how much additional support a particular data partition provides for a particular node in a simultaneous analysis as compared to a separate analysis of that data partition (see further explanation in Audzijonyte et al., 2005).

2.4.3. Inferring divergence times

The molecular clock assumption among *Paramysis* taxa was tested using a likelihood ratio test for the protein-coding and rRNA genes separately; this was done by comparing likelihood scores of trees with identical topology but with a clock relaxed and enforced ($\alpha = 0.01$, df = number of taxa—2). Given that non-*Paramysis* species were strongly diverged, the same analysis was repeated with *Paramysis* taxa and one outgroup species only. Since the molecular clock assumption was readily rejected, the divergence times were estimated using a penalised likelihood (PL) approach that allows for lineage-specific rate variation (Sanderson, 2002), as implemented in r8s v1.7 (Sanderson, 2006). Divergence times were estimated for *Paramysis* species only, and were based on the PRO data set only, as RNA data showed extreme and irregular rate variation. For the PL analyses, we used a Bayesian tree obtained for the PRO data (parameters listed in Table 2); the tree had identical topology with that from the simultaneous analyses, except for the poorly supported position of *P. agigensis* and *P. bacescoi*. The tree with branch lengths was supplied to the r8s program and the optimal smoothing level explored using a cross-validation procedure and TN algorithm and additive penalty functions. Because no fossil calibration points are available for *Paramysis* (and mysids in general) we assigned the basal *Paramysis* node an arbitrary value of 100 and considered the age of other nodes relative to this basal split. We also tested the robustness of the estimates by calculating divergence times with different smoothing factors (log values of 0.5, 1.0, 1.5. 2.0 and 2.5), but the effects of this on node ages were minor, no larger than 2%-units.

To explore potential systematic differences in divergence estimates from separate genes, we also conducted PL analysis for each protein-coding gene separately. We enforced the topology from the Bayesian analysis of PRO data described above and estimated branch lengths by ML for each gene separately using the GTR + I + Γ model (all parameters estimated during the analysis). From these phylograms divergence times were then estimated by r8s, as described above. The obtained ranges of divergence values in this approach are not directly comparable to confidence intervals from Bayesian inference (BI) (e.g., Thorne and Kishino, 2002) or from multiple simulated data sets. At any rate, in the absence of a fossil-based calibration the time estimates will inevitably be tentative; because of this and of recent reservations regarding the power of BI to simultaneously estimate divergence times and evolutionary rates without adequate parameterisation (e.g., Emerson, 2007), we do not focus on the precision of the estimates themselves, but on potential systematic differences in estimates from nuclear and mtDNA data.

3. Results

3.1. Patterns of molecular diversity and relative rates of molecular evolution

No departures from homogeneity of base composition among species were recorded in the molecular data partitions when tested using χ^2 test, with an exception in the COI gene's 3rd codon positions. The difference in base composition was between *Paramysis* and the distantly related genera *Schistomysis* and *Mysis*; when the latter two taxa were excluded, or when 3rd codon positions were excluded for the entire data set, the hypothesis of homogeneity was not rejected. The patterns of intraspecific diversity in OPS and EPRS agreed with those expected for single-copy genes in diploid genomes, with both homo- and heterozygote genotypes present.

Based on the data obtained in this study and from the previous COI assessments (Audzijonyte et al., 2006), the maximum intraspecific differences p in the protein-coding genes were 3.4% for COI (except for *P. lacustris*, see Section 2), 0.5% for OPS, and 0.5% for EPRS. Minimum interspecific differences were 3.8%, 1.0% and 0.7%, respectively, and the maxima 28%, 32% and 24%. With a few exceptions, the protein-coding data strongly supported monophyly of all species for which more than one

specimen were analysed (100% bootstrap support). One specimen was therefore used in further phylogenetic analyses to represent each species (except for *P. lacustris* CB and AZ taxa). A single exception to the species monophyly was found in the *P. baeri–P. bakuensis* species pair (Daneliya et al., 2007) characterised by ca. 7% mtDNA COI divergence; *P. baeri* was paraphyletic and basal to *P. bakuensis* in the OPS analysis, whereas in EPRS *P. bakuensis* was basal to *P. baeri*. Even in these cases intraspecific differences were always smaller than interspecific ones, and the choice of individuals was deemed not to affect the results essentially. We also found that the deep intraspecific COI divergence previously suggested for *P. kroyeri* in Cristescu and Hebert (2005) more plausibly represented misidentification; their "marine" sequence from the Romanian Black Sea coast closely matched our Sea of Azov *P. kroyeri* (<1% sequence difference), whereas the two "brackish" Dniester liman sequences (AY529035, -36) were similarly close to our *P. agigensis* and probably came from this species.

In terms of total nucleotide and amino acid variation the three protein-coding genes reflected different molecular rates and constraints (Table 3). The COI gene was most conserved at the amino acid level, with only seven (3.5%) amino acids variable within *Paramysis* (excluding *P. arenosa*, see below). For instance, *P. baeri*, *P. bakuensis* and *P. kessleri* which were characterised by 8–17% overall COI divergence still were all identical in amino acid sequence. The OPS gene in contrast had ca. 20% of the amino acids variable within *Paramysis*. Correspondingly the d*N*/d*S* ratio varied from 0.03 in COI to 0.12 in OPS genes (Table 3).

A clear saturation of changes in the COI gene when compared to OPS divergence was seen in the plots of pairwise distances between genes, whereas relationship between the two nuclear genes EPRS and OPS was more linear (Fig. 2a, b and e). Plots of OPS (and EPRS) versus rRNA divergence showed poor linear correspondence (Fig. 2c), suggesting that divergence of rRNA genes is strongly irregular and not useful for time estimation. Nevertheless there was a good correspondence of pairwise distances between 18S and 28S (Fig. 2d), suggesting correlated evolution in the two rRNA genes. In terms of the model-corrected distances, the divergence in OPS among *Paramysis* taxa (*P. arenosa* excluded) was on average 1.9 (±0.6 SD) times larger than in EPRS gene. The ratio between COI and OPS distances was 10.0 (±5.1 SD), and between COI and EPRS 16.7 (±6.0 SD). Neither the mean nor the dispersion of the ratio was much different if we excluded comparisons where COI uncorrected distance exceeded 15% or 20% (indicating saturation).

The sequences of 18S spanned almost the entire SSU molecule, i.e., from stem 3 to 50, ca. 1700 bp. Generally the most variable parts in the sequences in this study corresponded well with those identified as variable in the broader Mysida analysis by Meland and Willassen (2007). Three species of the *Paramysis* subgenus (*P. baeri*, *P. bakuensis* and *P. kessleri*) had particularly long insertions in loops 10 and E10_1, but also variation in rather conservative areas, such as stems 8 or 12 (see Electronic Supplement V, and Meland and Willassen, 2007). The sequences of 28S used here corresponded to the positions 824–1423 of the 25S subsequence of the yeast rRNA (GenBank Accession No. U53879), spanning from the positive strand of stem D2 to stem D19.

3.2. Phylogeny: total data set

A simultaneous MP analysis of all the data sets combined (22 taxa, a total of 4288 nucleotides and 35 morphological characters, comprising 1230 informative characters in all; Table 4) gave two equally parsimonious trees that differed only in the positions of two non-*Paramysis* species, the Atlantic and Mediterranean *Schistomysis assimilis* and the shared Ponto-Caspian (PC) *Katamysis warpachowskyi*. These MP trees strongly supported the monophyly of the *Paramysis* genus, excepting a single Atlantic–Mediterranean species *P. arenosa*, whose placement nevertheless was not well resolved (Fig. 3). The Bayesian analysis of the total data set returned an identical topology of the ingroup *Paramysis* (excluding *P. arenosa*), and likewise could not resolve the positions of *S. assimilis* and *K. warpachowskyi*. The best tree constrained for the monophyly of *Paramysis* including *P. arenosa* was 54 steps longer than the overall MP tree, and significantly worse in terms of the non-parametric and Kishino–Hasegawa tests ($p < 0.01$) (Table 4). In simple comparisons of the best similarly constrained ML tree and the overall ML tree from the PRO, PRN and RNA data sets, the AU tests also suggested significant ($\alpha = 0.05$) difference in two data sets (p values 0.043, 0.092 and 0.041, respectively). Given this evidence of exclusion of *P. arenosa* from a *Paramysis* clade and its strong molecular and phenotypic divergence, we

Table 3
Molecular diversity of the analysed gene segments in *Paramysis* species ("*P*". *arenosa* excluded) and in all studied taxa (separated by a slash)

Gene	Length	Variable	PI	AT%	AA	d*N*/d*S*
COI	603	229/276	201/242	53/54	7/39	0.031
EPRS	435	84/191	46/141	59/55	15/47	0.087
OPS	783 (780–783)	215/398	131/287	62/56	56/111	0.123
18S	1749 (1690–1723)	184/473	152/352	53/53		
28S	698 (617–671)	692/212	73/175	52/51		

Unaligned lengths of genes are given in parentheses, Variable, variable sites; PI, parsimony informative characters; AT%, percent A + T nucleotide content; AA, number of variable amino acids. d*N*/d*S*, ratio of non-synonymous versus synonymous substitutions per non-synonymous and synonymous sites in *Paramysis* species.

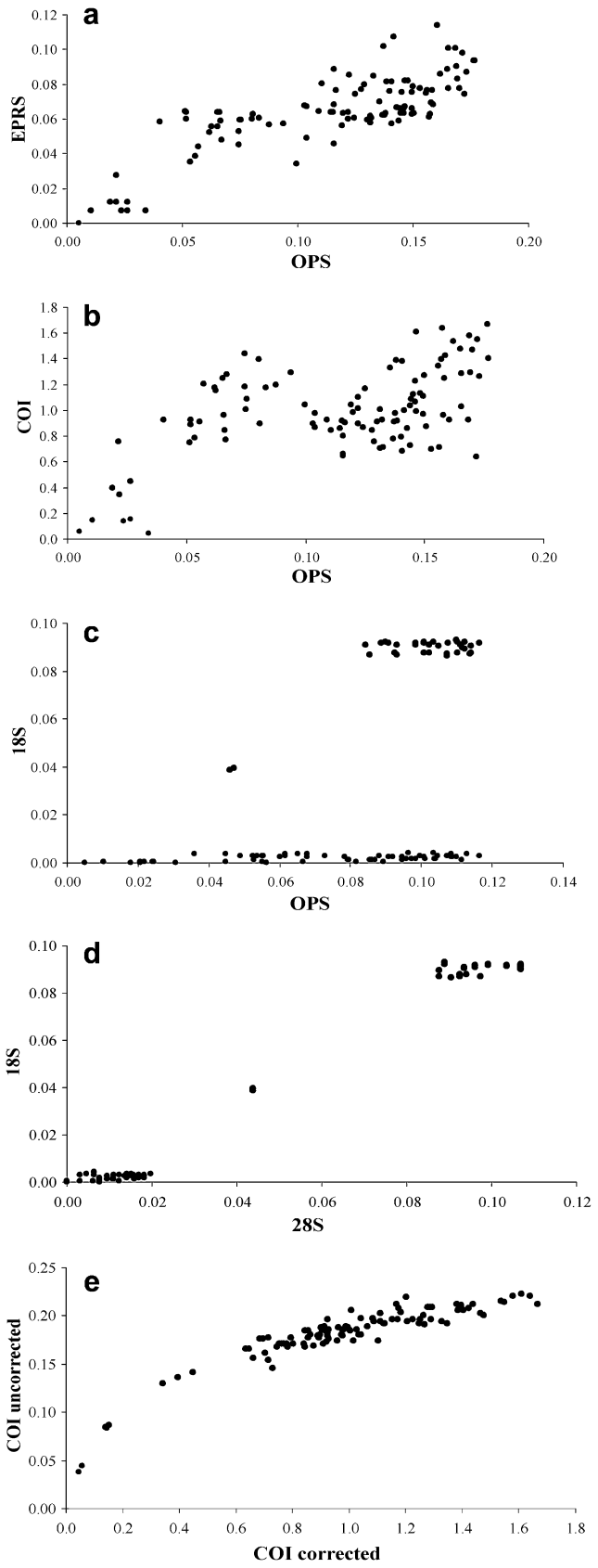

Fig. 2. Scatter-plots of pair-wise molecular distances among *Paramysis* taxa ("*P*". *arenosa* excluded), between different genes and distance measures. (a and b) Model-corrected distances; (c and d) uncorrected *p* distance; (e) model-corrected versus uncorrected distance for the COI gene.

shall henceforth treat "*P*". *arenosa* as a member of the outgroup, and use the name *Paramysis* to refer to the remaining monophyletic part of the genus only. The two further endemic monotypic genera *Caspiomysis* and *Limnomysis* were both distant from *Paramysis* in the analyses.

Relationships within *Paramysis* were fully resolved and most branches supported by >95% bootstrap support in overall MP and in protein data based ML analyses, and 100% posterior probability in BI tree (Fig. 3 and Table 4). Five main clades of *Paramysis* were identified. Four of these included a Caspian endemic and a more widespread shared PC species each. With reference to their previous subgeneric assignments, "type" species, or biogeographical characteristics, these clades will here be termed the BAERI, METAMYSIS, INTERMEDIA, SERRAPALPISIS and BA (=Black Sea—Atlantic) groups.

The deepest node within *Paramysis* separated two well-supported groups: the BAERI group of three species in the subgenus *Paramysis* s. str. (*P. baeri*, *P. kessleri* and *P. bakuensis*) versus the remaining *Paramysis*. In the BAERI group, the Caspian endemic *P. baeri* and the PC *P. bakuensis* grouped together to the exclusion of the PC *P. kessleri*.

The next node separated a clade of two morphologically close species of the *Metamysis* subgenus, the Caspian *P. inflata* and the PC *P. ullskyi*. The INTERMEDIA group comprised another biogeographically similar species pair, but with a considerably deeper divergence from each other—the Caspian *P. loxolepis* and the PC *P. intermedia*. The remaining Ponto-Caspian taxa were in the SERRAPALPISIS group. The two intraspecific *P. lacustris* lineages were evidently the closest pair of taxa in the data set. They grouped together with the Azov-Caspian *P. sowinskii* to the exclusion of the Caspian endemic *P. incerta*.

The final, well-supported BA group included four taxa which do not occur in the Caspian. Within the group, all the data strongly supported a sister relationship between the NE Atlantic *P. nouveli* and the Black Sea *P. kroyeri*, to the exclusion of the NE Atlantic *P. bacescoi*. The relative positions of *P. bacescoi* and the Black Sea *P. agigensis* were weakly supported. The BA group was placed as a sister of the SERRAPALPISIS group, but intra-group divergences were considerably deeper within the BA group. The BA–SERRAPALPISIS clade was sister to the INTERMEDIA clade both in the MP and Bayesian analyses, but the support in MP was weak (47%) and subject to conflicts between PRO and RNA data (see below). MP analysis of the molecular data alone gave the same two topologies as all the data combined, but with still weaker support for the three less supported nodes (Table 4).

3.3. Phylogeny: morphological data with complete taxon sampling

For the 35 morphological characters selected for the analysis, *P. kosswigi* was scored identical to *P. lacustris*, and *P. helleri* and *P. festae* identical to *P. kroyeri* (excluding missing characters); only *P. lacustris* and *P. kroyeri* of

Table 4
Tree statistics and bootstrap support (MP, ML) and posterior probability (BI) values for the nodes of the trees obtained with different combinations of data sets

Data set:	All		Mol	Mol*	PRO		PRN		RNA		COI	CO*	CO3	EPR	OPS	18S	28S	MOR
Method:	MP	BI	MP	MP	MP	ML	MP	ML	MP	ML	MP	MP	MP	MP	MP	MP	MP	MP
No MP trees	2	—	2	1	1		2		7		2	12	1	3	1	2	10	32
Tree length/$-\ln L$	4398	—	4275	3068	3021	14787.1	1605	8527.1	1174	8190.2	1379	203	1164	509	1086	792	429	111
PI	230	—	1198	1005	671	—	429	—	527	—	242	49	193	141	287	352	175	32
RC	0.35	—	0.34	0.47	0.22	—	0.32	—	0.75	—	0.14	0.22	0.13	0.32	0.33	0.8	0.67	0.54
HI	0.48	—	0.48	0.39	0.57	—	0.47	—	0.20	—	0.66	0.64	0.66	0.46	0.47	0.16	0.25	0.31
BAERI	100	*100*	100	100	100	*100*	100	*99*	100	*100*	—	—	—	95	98	100	100	93
(*bakuensis*, *baeri*)	100	*100*	100	100	100	*100*	100	*100*	100	*100*	100	94	98	99	100	100	100	83
METAMYSIS	100	*100*	100	100	100	*100*	100	*100*	90	—	100	87	100	100	100	—	—	99
SERRAPALPISIS	100	*100*	100	100	100	*100*	100	*98*	—	—	79	60	54	93	91	60	—	70
(*lacustris*, *sowinskii*)	100	*100*	100	96	100	*100*	100	*96*	—	—	90	31	93	96	93	—	—	—
BA	100	*100*	100	93	97	*100*	78	*96*	92	*81*	<5	—	18	—	77	40	84	—
(*kroyeri*, *nouveli*)	100	*100*	99	99	100	*100*	100	*100*	—	—	32	—	20	93	100	—	—	—
(*kroyeri*, *nouveli*, *bacescoi*)	73	*61*	66	58	53	—	—	—	—	—	19	—	16	—	—	—	—	—
INTERMEDIA	92	*100*	88	99	72	*72*	88	*86*	92	—	—	—	—	—	86	—	—	—
(SERRAPALPISIS, BA)	97	*100*	98	100	98	*100*	100	*98*	—	—	—	—	—	15	100	—	—	—
(SERRAPALPISIS, BA, INTERMEDIA)	49	*100*	45	75	40	*74*	86	*88*	—	—	—	—	—	—	78	—	—	—
BAERI versus other *Paramysis*	100	*100*	99	100	72	*97*	98	*97*	97	*69*	—	—	—	—	89	100	—	—
PARAMYSIS (exc. *arenosa*)	100	*100*	100	100	97	*98*	100	*96*	100	*100*	—	—	—	62	99	100	99	38
(*Paramysis*, *Katamysis*, *Schistomysis*)	90	*65*	93	98	89	*66*	99	*89*	90	*100*	—	—	—	82	96	99	—	—
(PARAMYSIS, incl. *arenosa*)	4452	—	—	—	3057	14799.3	1638	8536.0	1195	8210.6	—	—	—	—	—	—	—	—
p value	*<0.01*				*<0.01*	*0.043*	*<0.01*	*0.092*	*<0.01*	*0.041*								

Support from model-based approaches is shown in *italics*. — denotes nodes not supported by strict consensus in MP, or with <50% bootstrap support in ML. Mol*, molecular data without COI 3rd codon positions; CO*, COI without 3rd positions; CO3, only 3rd positions of COI; PI, number of parsimony informative characters; RC, rescaled consistency index; HI, homoplasy index (excluding uninformative characters). Taxa with missing data were deleted from corresponding analyses; the statistics for MOR are from analyses including only the taxa for which molecular characters were available. Results (MP and ML tree scores) of the analyses constrained for the monophyly of all *Paramysis* taxa (including *P. arenosa*) are also shown, together with *p* values of the tests comparing them with unconstrained searches (K–H test for MP analyses and approximately unbiased test for ML, see Section 2.4.2).

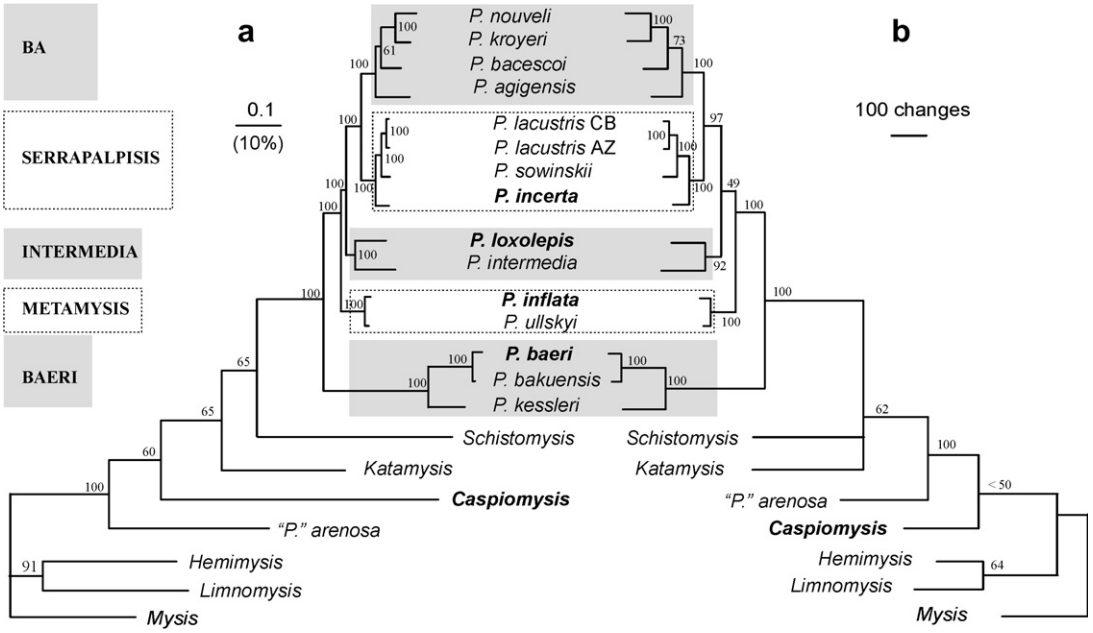

Fig. 3. Phylogeny of Ponto-Caspian mysids based on the simultaneous analysis of five genes and morphological data. (a) Bayesian inference and (b) maximum parsimony. Posterior probability (a) and bootstrap support values (b) are indicated. Branch lengths reflect the number of changes. Caspian endemics are shown in bold. Five phylogenetic clades of *Paramysis* are indicated with white/grey shading (see Section 3.2).

these groups were therefore used in the analysis. MP analysis of the morphological data set gave 152 equally parsimonious trees of 115 steps, supporting three of the main groups identified from the molecular data (Fig. 4): (i) The BAERI group, which here also included a Caspian endemic species *P. eurylepis*; its position next to the *P. bakuensis*–*P. baeri* species pair, leaving *P. kessleri* as basal, was supported by two unambiguously optimised characters. (ii) METAMYSIS group, with an additional Caspian endemic species *P. grimmi*; relationships within the group were however unresolved. (iii) SERRAPALPISIS group, where relationships among species were conflicting with the simultaneous analysis topology.

The main differences between molecular and morphological analyses were in the position of the BA group (which here included the additional Mediterranean taxa *P. helleri* and *P. festae*, and the Black Sea *P. pontica*), where morphological characters did not support (but not conflict either) the monophyly of the group and implied their basal placement on the *Paramysis* tree. Further, the morphology-based grouping of BAERI and METAMYSIS groups contrasted with the main molecular split of the of BAERI clade versus all other *Paramysis* species. The morphological distinction of the BAERI group was nevertheless evidenced by several synapomorphies. Excluding the six taxa with only morphological data available from the analysis did not affect the relationships among the remaining species.

3.4. Phylogenetic information and conflicts in separate data sets

Considering the contribution of the various subsets (partitions) of the data to the overall tree resolution in terms of partition Bremer support values (PBS), the OPS and EPRS genes had highest information content and they were most congruent with the simultaneous analysis topology. These genes had positive PBS values for all nodes and high ratio of total PBS to number of parsimony informative (PI) characters (Table 5). Of the individual gene data sets, a separate MP analysis of OPS gave the best resolved and supported trees, and a nearly complete resolution was obtained with the two PRN genes (OPS + EPRS) combined, both with the MP and ML approaches (Table 4).

Separate analysis of the mtDNA COI gene resolved only nodes at the tips of the tree and the data set had highest homoplasy index among all the data partitions. Excluding 3rd codon positions did not substantially alter MP results, or decrease the homoplasy index (Table 4). A number of outgroup taxa were placed within *Paramysis*, rendering the above defined *Paramysis* clades non-monophyletic. COI contributed positively to the simultaneous analysis topology, but most (ca. 80%) of the support for the final topology was obtained only in the simultaneous analysis of this gene with other data sets, as seen from its high PHBS value. In contrast PHBS value of PRN genes was only about 25% of the total PBS. Morphological data were congruent with the simultaneous analysis topology, and had mostly positive PBS values. Although none of the well-supported nodes obtained in the separate analyses of RNA data conflicted with the simultaneous analysis topology, the negative PHBS values nevertheless suggested incongruence of RNA with other data partitions. The main feature of both 18S and 28S based trees was the strong distinction of BAERI group from the rest; when these three species were excluded, 18S data contained only nine variable

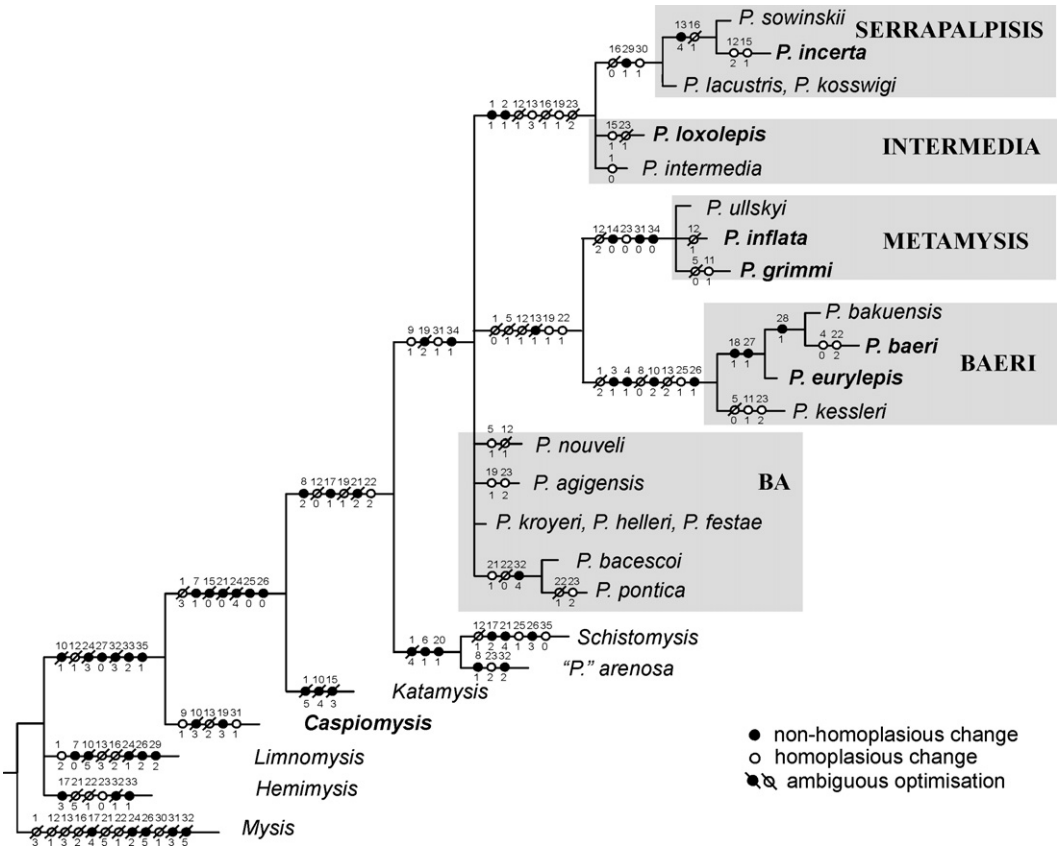

Fig. 4. Phylogeny of Ponto-Caspian mysids based on morphological data, strict consensus of 152 equally parsimonious trees. Numbers above the circles indicate character codes and those below the circles show the inferred states (Appendix A). Ambiguous character optimisations were resolved using ACCTRAN optimisation.

Table 5
Partitioned Bremer support and partitioned hidden Bremer support (PBS/PHBS) values for the different data sets, for clades within *Paramysis* ("*P.*" *arenosa* excluded), from the MP analysis of all taxa included; the statistics for MOR are from analyses including only the taxa for which molecular characters were available

Clade	MOR	COI	OPS	EPRS	18S	28S	Total
No. PI characters:	32	242	287	141	352	175	1230
BAERI	4/1	−4/−1	14.5/5.5	4.5/0.5	138/−3	23/−6	180/−3
(*bakuensis*, *baeri*)	2/ 0	16/5	19.5/6.5	5.5/0.5	34/−3	12/−1	89/8
METAMYSIS	5/1	22/8	14/0	9/1	0/0	2/2	52/12
SERRAPALPISIS	4/3	3/−1	10.5/6.5	4.5/0.5	−2/−3	0/0	20/6
(*lacustris*, *sowinskii*)	0.5/0.5	9.5/9.5	4/0	3/0	−0.5/−0.5	−0.5/−0.5	16/9
(*lacustris* AZ, *lacustris* CB)	1/1	7/2	4/0	2/0	0/0	0/0	14/3
BA	1/1	5/7	8.5/6.5	0.5/0.5	−1/−2	3/1	17/14
(*kroyeri*, *nouveli*)	0/0	4/4	12/0	4/1	0/0	0/0	20/5
(*kroyeri*, *nouveli*, *bacescoi*)	1/1	1.4/3.4	1.1/1.1	0.7/0.7	0.4/0.4	0.4/0.4	5/7
INTERMEDIA	2/2	−4/−1	10.5/7.5	0.5/0.5	−4/−4	2/2	7/7
(SERRAPALPISIS, BA)	0/0	−6/−3	16.5/2.5	2.5/2.5	−4/−4	0/0	9/−2
(SERRAPALPISIS, BA, INTERMEDIA)	0/0	−3/0	6.5/2.5	0.5/0.5	−3/−3	0/0	1/0
BAERI versus other *Paramysis*	−1/−1	−2/1	11.5/5.5	3.5/3.5	9/−3	−1/−1	20/5
PARAMYSIS (exc. *arenosa*)	2/0	4/7	26.5/9.5	2.5/0.5	19/0	4/−8	58/9
Total: PBS/PHBS	21.5/9.5	53/41	160/53.5	43.2/12.2	186/−25	45/−11	508/80
PBS per PI characters	0.65	0.22	0.56	0.30	0.53	0.26	0.41

and six PI characters within *Paramysis*. As a consequence, most of the total PBS values in the two RNA genes came just from the two nodes in the BAERI group (Table 5). In all, separate analyses of 18S and 28S data provided support only for four nodes each, and when the two genes were combined, MP tree supported two additional (METAMYSIS and INTERMEDIA) groups; the ML topology was similar to that from MP, but had lower node support.

3.5. Molecular dating

Molecular clock was rejected by the likelihood ratio test for each of the PRN, PRO and RNA data sets, both when all outgroups or only *Schistomysis* was included in the analysis ($-2\ln\Lambda$=70–720, df = 15–20, $p < 0.001$); the clock was also rejected in separate analyses of the COI, OPS and EPRS genes. Rate heterogeneities appeared to be in different parts of the tree in the different data sets. The ML phylogram from PRO showed disproportionately long branches in the BA group, whereas in the RNA tree it was branches of the BAERI group that were 10 times longer than in the other *Paramysis* species.

Using the divergence time estimates from the penalised likelihood (PL) analysis of PRO data the separation of the METAMYSIS group was at ca. 85% of the *Paramysis* root age, divergence of the INTERMEDIA group at ca. 80%, and divergence of the SERRAPALPISIS from the BA group at ca. 50%. Three sister-species pairs, *P. ullskyi–P. inflata*, *P. lacustris–P. sowinskii* and *P. baeri–P. bakuensis*, all appeared to have diverged at approximately similar time, at about 15% of the *Paramysis* root age (Fig. 5 and Table 6).

The overall rate variation among branches of the COI and OPS trees, as inferred by PL analysis, was 5- to 6-fold and in EPRS about 3-fold. There was an overall agreement in lineage-specific rate variability among the three genes, yet divergence time estimates from the individual genes varied (Table 6). Particularly node ages inferred from the mtDNA COI gene in most cases were older than those from the nuclear genes, suggesting that its substitution model could not adequately correct for both the shallow and deeper divergences. From the OPS gene data the *P. ullskyi–P. inflata* divergence appeared exceptionally old,

Table 6
Ages of nodes inferred by the penalised likelihood analysis, shown in percentage of the *Paramysis* root age (100) (see Fig. 5)

Node	PRO	COI	EPRS	OPS
BAERI	62–63	*86*	46–47	52–53
(*bakuensis*, *baeri*)	15–16	*24*	10	10
METAMYSIS	16	12–13	12	*31–32*
SERRAPALPISIS	28–29	44	30	29–30
(*lacustris*, *sowinskii*)	16	29–30	10	21
(*lacustris* AZ, *lacustris* CB)	6	12	—	7
BA	41–42	67	63	36
(*kroyeri*, *nouveli*)	20	39–40	36	*12*
(*kroyeri*, *nouveli*, *agigensis*)	36–37	56–57	58	*34*
INTERMEDIA	66	89	74	59
(SERRAPALPISIS, BA)	52–53	*86*	68	*43*
(SERRAPALPISIS, BA, INTERMEDIA)	77	−79	71	
BAERI versus other *Paramysis*	85	—	84	83

Dates from the combined analyses of all protein-coding genes (PRO) and from individual genes are given; variation due to application of different smoothing values ($\alpha = 0.5$–2.5) is also shown for each data set. The most discordant dates are shown in *italics*; dashes indicate nodes that were not supported by the data.

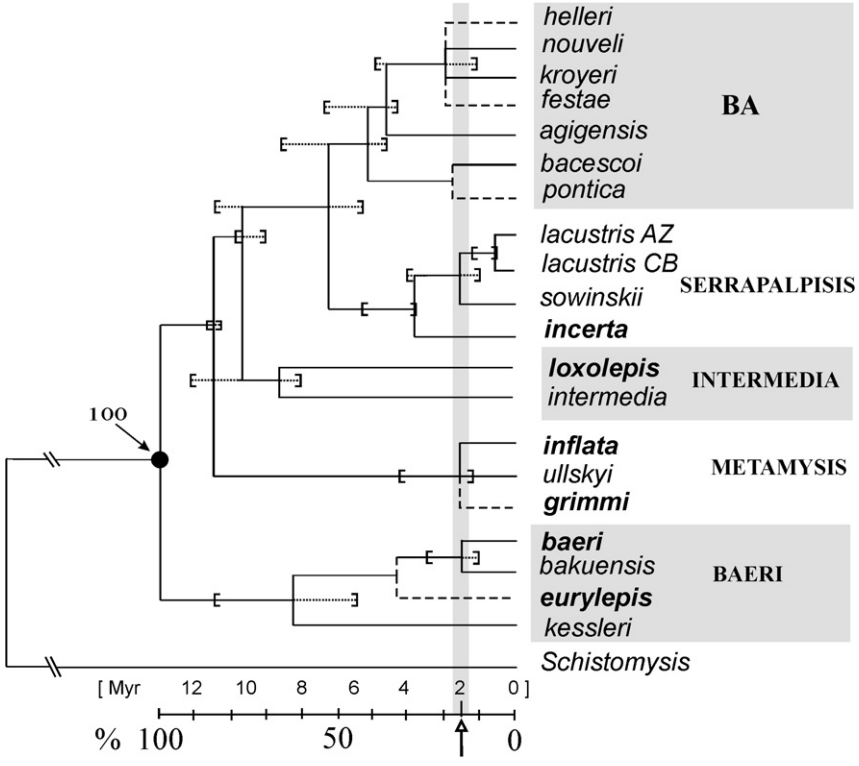

Fig. 5. Chronogram of *Paramysis* species obtained from the penalised likelihood analysis of a tree based on the three protein-coding genes OPS, EPRS and COI. The brackets around nodes indicate the range of relative ages from analyses of the three genes separately (Table 6). The primary divergence scale indicates the relative age (%) with respect to the *Paramysis* root age. The tentative absolute time scale is based on calibrating the most recent divergences at a putative vicariance following the end of Akchagyl transgression about 2 MY ago (shading; see text for details). The species with branches shown by dashed line were not included in the molecular analyses, but their position is tentatively shown based on the morphological tree (Fig. 4).

whereas nodes associated with BA group had younger ages than inferred from other genes.

No fossils are available to calibrate molecular clocks of mysid crustaceans, and no calibrations are available for OPS and EPRS genes in invertebrates in general. With the inferred several-fold difference in rates of molecular evolution among lineages within *Paramysis* itself, it is hard to justify application of a "universal" mtDNA clock either (e.g., 2% per 1 MY). Instead, to convert relative ages to the absolute time scales we used a biogeography assumption of vicariant subdivision. Three Ponto-Caspian species pairs, *P. ullskyi–P. inflata*, *P. baeri–P. bakuensis* and *P. sowinskii–P. lacustris* seem to have diverged at approximately simultaneously at ca. 15% of the *Paramysis* root age, and the former two species pairs remain largely vicariant. From paleogeography the most plausible event to cause simultaneous vicariant splits would be the termination of Akchagyl transgression about 1.8 MY ago, leading to separation of the Caspian from the Black Sea basin (see Section 4). Assuming that 15% of *Paramysis* root divergence corresponds to roughly 2 MY, the separation of older *Paramysis* clades (METAMYSIS, INTERMEDIA) would have occurred 10–11 MY ago, divergence between SERRAPALPISIS and BA about 6 MY ago, and diversification within the clades within the last 5 MY (Fig. 5).

Viewing the rates of individual genes in the light of these age estimates again points to conspicuous rate variations among the lineages, particularly in mtDNA. Considering divergence in each of the sister pairs *P. ullskyi–P. inflata*, *P. baeri–P. bakuensis*, *P. sowinskii–P. lacustris* and *P. kroyeri–P. nouveli* (the last pair appearing slightly older than the others, 20% of *Paramysis* root or 2.5 MY), the inferred COI rates would vary over 10-fold, i.e., 2.5%, 7%, 7% and 30% per million years, respectively, in the four pairs. The corresponding rates for OPS would be 1.7%, 0.6%, 1.2% and 0.9%, and those for EPRS 0.35%, 0.35%, 0.35% and 1.1% per MY.

4. Discussion

Ancient lakes are by definition characterised by endemic diversification, but which of their features are responsible for generating and maintaining the diversity remains contentious. Is it the long and continuous history *per se* that allowed accumulation of lineages in the course of steady rates of speciation? Or do these lakes provide unique ecological conditions and protection from competition that in turn allow species to radiate into a variety of niches? More generally, are ancient lakes more appropriately characterised as evolutionary reservoirs or hot spots of diversification (Wilson et al., 2004)? The history of endemism in the Caspian Sea and the Ponto-Caspian region can be, on the one hand, viewed in the light of these general questions. On the other hand, the area makes a special case in that the origin and diversity of its fauna have largely been shaped by the recurrent connections to the marine environments. The main challenge in deciphering its faunal history lies in understanding the consequences of these connections—the number of lineages dispersing in and out of the basin to true seas, the timing of these events and subsequent local radiations.

The main results from our study of Ponto-Caspian mysid evolution include (i) good phylogenetic resolution, suggesting gradual rather than burst-like speciation; (ii) evidence of an early (probably Late Miocene) start to an endemic *Paramysis* diversification in continental seas, with no recurrent inputs from the ocean; (iii) evidence of a secondary, single (re)colonisation of the marine realm from the originally continental *Paramysis* lineage; (iv) the overwhelming contribution of nuclear protein-coding genes to the phylogenetic resolution, as contrasted with poor performance of the widely used mtDNA and rRNA characters. These points will be discussed in the light of the paleogeographic history and information on diversification in other animal groups.

4.1. The Ponto-Caspian as an ancient cradle of continental mysid diversity

The origin of the Ponto-Caspian inland basins goes back to the final phases of the Tethys Sea that spanned between the African and Eurasian continents. With the approach of the Arabian and African plates towards Europe an epicontinental Paratethys inland sea gradually became separated from the Mediterranean and the world ocean. These connections were severed in the Late Miocene 12 MY ago (Sarmatian time), and closed by about 6 MY ago (Popov et al., 2006) (Fig. 1). After the separation, 5.5–5.3 MY ago, the Mediterranean experienced the Messinian desiccation and salinity crisis (a salinity increase then decrease), which is believed to have had a great impact on faunal subdivisions and vicariant isolations (Krijgsman et al., 1999; Vasiliev et al., 2005). At this time the Ponto-Caspian basin enjoyed favourable brackish-water conditions known as the Pontian Sea (6–5 MY). Judging by the molluscan fossil record, both the Sarmatian and Pontian stages harboured rich endemic faunas, thought to have resulted from local radiations initiated by independent marine lineages (Nevesskaya et al., 2006). At the beginning of Pliocene 5 MY ago, the rise of the Caucasus Mountains subdivided the Pontian into the Black and Caspian Sea basins; in the Mediterranean, the start of Pliocene is marked by the opening of the Gibraltar and reconstitution of marine conditions. Since then the Black and Caspian seas have largely had independent histories and at times distinctly different faunas but, on the other hand, they have also been linked by recurrent waterway connections that likely facilitated faunal exchange. After the initial subdivision ca. 5 MY ago, the sagging of the southern Caspian and Black Sea beds caused a prominent drop in the water level. It has been suggested that between ca. 5 and 3.5 MY ago the Caspian Sea level dropped by >700 m (Zubakov, 2001), but there is disagreement as to whether that was associated with a decrease of salinity to nearly freshwater or with an increase

to hypersaline levels (Popov et al., 2006). In any case, the Caspian environment then seems to have been unsuited for the autochthonous brackish-water species, as indicated by their absence in the mollusc fossil record. The Black Sea in contrast remained brackish throughout the Pliocene, and seems to have supported endemic radiations, with about 20 new genera of bivalve molluscs described from the period (Nevesskaya et al., 2006). A major faunal exchange likely occurred in the Late Pliocene (3.4–2 MY ago) when the Akchagyl transgression engulfed the Black and Caspian seas into a large brackish-water body. There was also an important marine input to the Akchagyl, seen in the presence of primarily marine mollusc taxa in its sediments (Starobogatov, 1970). Several marine intrusions to the Caspian Sea during the Pliocene (5–2 MY) are also suggested by the microfossil record, but their geography and timing remain unclear (Zubakov, 2001; Popov et al., 2006). Finally, the inter-basin sharing of the Ponto-Caspian element in the present time also implies continued Pleistocene connections (e.g., Audzijonyte et al., 2006).

One important feature of the Ponto-Caspian history is that despite its long-term hydrological isolation from the ocean, temporary connections have still existed allowing immigration. This is exemplified by the repeated contributions of marine lineages to Miocene to Pleistocene speciation bursts in molluscs (Nevesskaya et al., 2006), and also by molecular phylogeny of onychopod cladocerans, where at least two marine lineages seem to have contributed to the Caspian endemic diversity during the Pliocene, perhaps as late as ca. 2 MY ago (Cristescu and Hebert, 2002). These data indicate that the roots of the endemic Ponto-Caspian diversity are not necessarily as old as the long history of the basins might imply, contradicting the hypothesis that the old Sarmatian and Pontian lineages were important in the current endemism.

Our data on autochthonous mysid diversity suggest a different story, inconsistent with previously depicted scenarios. These species show deep molecular subdivisions of endemic lineages, conspicuous even at the intra-generic level in *Paramysis* and more so when including the other endemic genera *Caspiomysis*, *Katamysis* and *Limnomysis*. In this respect, the Ponto-Caspian region well qualifies as an ancient cradle of brackish-water mysid diversification, originating from early ancestral marine lineages isolated from the ocean together with the Paratethys itself. If our dating exercises are even approximately correct, the inception of the main *Paramysis* clades would be of Late Miocene age (6–10 MY; Pontian to Sarmatian) and differences among the genera considerably older. A number of other Ponto-Caspian taxa, such as amphipod crustaceans and *Dreissena* bivalve molluscs, are also characterised by deep intra-generic molecular subdivisions (Therriault et al., 2004; Cristescu and Hebert, 2005), but no broader phylogenies are available for these groups. Also from the fossil record, two extant bivalve species *D. polymorpha* and *D. rostriformis* are described already from Pontian deposits of 5–6 MY age (Kasymov, 1987). While mysids themselves have no appreciable fossil record, their statoliths (a mineral body situated in the uropod) are common in Late Miocene strata of Paratethys, and indicate a close relationship to the extant members of *Paramysis* (Wittmann et al., 1993).

A long independent history of the Ponto-Caspian lineages is also supported by the apparent absence of any immediate marine sister groups in the molecular phylogenies or morphological features. The closest connection is with *Schistomysis*, a genus itself comprised of strongly molecularly and morphologically divergent taxa, and thus of questionable monophyly (Meland and Willassen, 2007; K.J. Wittmann, personal communication). In fact, there is a species *Schistomysis elegans* that is endemic to the eastern coasts of the Caspian Sea, and a similar Caspian endemic is known for another mainly Mediterranean genus *Diamysis*. These two taxa, along with the "arctic immigrant" *Mysis* spp. (see below) could indeed represent more recent marine contributions to the Ponto-Caspian mysid diversity; alas both species are very rare or even extinct (Bondarenko, 1991) and were not available for our analyses. Yet, at least *S. elegans* appears morphologically strongly diverged from its marine congeners, leaving its taxonomic position vague. It might as well represent a monotypic genus, such as *Caspiomysis*, *Katamysis* or *Limnomysis*, and thus unlikely to be a recent addition to the Caspian fauna (Daneliya, unpublished).

Despite the lack of any links between marine and true Caspian or Ponto-Caspian (PC) taxa in our data, it is however notable that within *Paramysis* the Black Sea endemic species do make sister groups to the Atlantic–Mediterranean ones in the BA group (Figs. 3 and 5), and thus exemplify a probably more recent re-invasion of the Ponto-Caspian basin sensu lato (cf. next section). The time-scale of these invasions and the history of the endemic Black Sea lineages during the Pleistocene environmental disturbances remain open questions.

4.2. Secondary colonisation of the marine realm by a Paratethys lineage

The inference that one lineage of the initially Paratethyan *Paramysis* secondarily colonised the true marine waters and diversified thereafter is another exceptional feature in the evolutionary history of autochthonous Ponto-Caspian mysids. Neither the direction of invasion nor the inferred monophyly of the 'emigrating' BA clade were anticipated from previous hypotheses or data. If, as judged from the molecular divergence, the separation of the BA clade took place ca. 7 MY ago, it would roughly fit the ceased connection between the Pontian and Mediterranean basins. The subsequent Messinian low water stand could have facilitated further allopatric divisions in the Mediterranean and vicariance events with lineages in the NE Atlantic (*P. nouveli* and *P. bacescoi*). The inferred continental-brackish ancestry of marine *Paramysis* makes one of very few supported examples of secondary dispersal to

the marine realm. Although hypotheses about such Ponto-Caspian to marine invasions have also been proposed for other groups including onychopods (Cladocera) (Dumont, 1998) and seals (Ray, 1976; Koretsky, 2001), none of them so far was supported by molecular phylogenetic analyses (Cristescu and Hebert, 2002; Palo and Väinölä, 2006). Notably in *Paramysis* the evidence was obtained from the (multi-gene) molecular data. The morphological data, while not strongly contradicting the inference (as shown by the PBS values, Table 5), still rather suggested a basal (ancestral) position of the BA taxa (Fig. 4). Strictly, we do not have direct molecular data for three of the marine species *P. festae*, *P. helleri* and *P. pontica*, but their overall morphological similarity to *P. kroyeri* and *P. bacescoi* strongly point to their position within the BA clade.

Another character in support of a Ponto-Caspian ancestry of the Mediterranean–Atlantic *Paramysis* is the chemical composition of their statoliths. While most members of the primarily marine Mysinae subfamily have statoliths composed of fluorite, the Ponto-Caspian taxa, both extant and fossil, have statoliths of vaterite (calcium carbonate). Ariani et al. (1993) found vaterite statoliths in five species from the BA cluster (*P. bacescoi*, *P. helleri*, *P. kroyeri*, *P. nouveli* and *P. pontica*), and in six Ponto-Caspian species (*P. intermedia*, *P. kessleri*, *P. lacustris* and in *H. anomala*, *K. warpachowskyi*, *L. benedeni*). Notably, the Atlantic "*Paramysis*" *arenosa* excluded from *Paramysis* by our data, had fluorite statoliths, further supporting its distinction.

4.3. Vicariance-driven diversification?

The inference of long and gradual evolution of *Paramysis* diversity is not only based on the molecular dating but also on the good resolution of molecular phylogenies and congruence among different loci. Under the alternative hypotheses of recent diversification during the Pleistocene, or in radiation bursts even earlier, considerable incongruence among genes is expected from incomplete lineage sorting (Maddison, 1997), particularly for taxa with large population sizes such as mysids. Also, where speciation bursts in ancient lakes have been inferred empirically there indeed seems to be strong incongruence between characters because of deep coalescences or hybridisation, e.g., the endemic *Mysis* species flock (Väinölä, 1995; Audzijonyte et al., 2005) and *Podonevadne* cladocerans (Cristescu and Hebert, 2002) in the Caspian, and cichlids in East African Rift Lakes (Seehausen, 2004; Koblmüller et al., 2007). Another feature in the endemic *Paramysis* diversification is the absence of evident major eco-morphological radiation or niche shifts. Naturally, some differentiation in the depth and habitat preference exists, but most species still occupy the upper sublittoral zone and only *P. loxolepis* is truly pelagic and distributed down to 900 m in the Caspian Sea (Derzhavin, 1939; Bondarenko, 1991). Also interesting, the high endemic Ponto-Caspian diversity of amphipod crustaceans involves relatively little ecological and morphological specialisation (Tarasov, 1996), if contrasted with the extreme morphological and ecological variety of endemic amphipods of Lake Baikal (Kamaltynov, 1999). This contrast has been presented as an argument for a recent Pleistocene origin of the extant Ponto-Caspian amphipod fauna (Mordukhai-Boltovskoi, 1979; Tarasov, 1997), which however does not seem to agree with molecular data (Cristescu and Hebert, 2005).

A plausible reason for the rarity of specialised deep-water forms in autochthonous crustaceans is in the history of anoxic periods of the deeper waters, caused by changes in vertical mixing with varying freshwater river input and inter-basin connections. The current oxygenated status of the deep water Caspian Sea is probably not older than the Pleistocene, when salinity conditions in the basin became more stable (Dumont, 1998). Notably however, it is not the autochthonous mysids but an "arctic immigrant" element that has exploited this new environment. The primarily arctic marine genus *Mysis* has diversified into a flock of four distinct Caspian species specialised for deep pelagic waters. From molecular data, the radiation of *Mysis* seems to have occurred only during the Middle-to-Late Pleistocene, and is still characterised by substantial nuclear-mitochondrial incongruence (Väinölä, 1995; Audzijonyte et al., 2005). Moreover, the rates of morphological differentiation in *Mysis* indeed seem to have been fast, unlike in *Paramysis*.

The diversification of the autochthonous mysids can rather be explained in the context of allopatric divergence in the course of vicariant isolation between (sub)basins. Notably, even at present most sister species pairs are largely allopatric or represent different zoogeographical elements (Fig. 1 and Table 1). In practice, such contrasts are mostly between a more widespread estuarine (shared PC) and a central/southern Caspian endemic species, e.g., *P. bakuensis–P. baeri*, *P. intermedia–P. loxolepis* and *P. ullskyi–P. inflata*. For marine taxa this applies to the *P. kroyeri–P. helleri–P. nouveli* clade, with species confined to the Black Sea, Mediterranean and Atlantic, respectively, and to the Black Sea versus Atlantic pair *P. pontica–P. bacescoi*.

Basing our dating on an ad hoc inferred simultaneous vicariance of the youngest sister-taxon pairs, the attribution of other nodes dated to previous vicariance events does not completely avoid circularity, but may still give valuable insight. Several events since the Late Miocene stand out as probable triggers of vicariance and allopatric divergence (notwithstanding the poor understanding of the actual paleogeographic complexity). These include (i) the separation of the Paratethys (Ponto-Caspian) from the Mediterranean at ca. 8–6 MY ago; (ii) the Messinian desiccation in the Mediterranean 5.5 MY ago and subsequent reconstitution of connections to the ocean and Black Sea; (iii) subdivision of the Pontian into Black and Caspian basins ca. 5 MY ago, and (iv) the end of Akchagyl transgression ca. 2 MY ago and re-isolation of the basins. The closing of the Paratethys could account for the BA clade—SERRAPALPISIS split, the Messinian crisis for the

deeper divergences seen within the BA clade, the Pontian subdivision for early vicariant speciations within each of BAERI, SERRAPALPISIS, METAMYSIS and INTERMEDIA clades, whereas the post-Akchagyl event would have caused the final sister-species splits.

While the series of vicariance events would provide a plausible scenario for the origin of the extant lineage diversity, problems remain in understanding the geography of (vicariant) lineage survival trough the turbulent Ponto-Caspian paleoenvironmental history and its suggested effects, e.g., the repeated regional extinctions of brackish-water fauna. For instance, the persistence of the strictly Caspian endemics *C. knipowitschi*, *P. loxolepis* or *P. incerta* in the basin since the putative ancient vicariance event(s) does not fit the presumed absence of brackish-water fauna in the Caspian during the Balakhanian regression (5–3.4 MY). A general problem also concerns the history of the endemic Black Sea element now only found in salinities >10‰ (e.g., *P. pontica* and *P. agigensis*, *Hemimysis serrata* Bacescu, 1938), since the basin is believed to have been much more dilute (ca. 5‰ or even less) as recently as 15 KY ago (Mudie et al., 2002; Ryan et al., 2003). Clearly, generalising from a current vicariance of species to a continued vicariant history will not always be warranted. The true history probably has been more complex, involving dispersal, recolonisation and extirpation events, and/or the persistence of more favourable local refugia. Yet the palaeogegraphic-genetic vicariance scenario presented above makes a good basis to guide further work.

4.4. Relative performance of different genes: new nuclear power for resolution

So far, mitochondrial genes and those of nuclear rRNA have been the most popular markers and remain the backbone of crustacean molecular phylogenetics. Nuclear protein-coding genes have already become commonplace in insect and decapod systematics (Williams et al., 2001; Damgaard and Cognato, 2003; Moulton and Wiegman, 2007), but our work seems to represent their first proper application in peracarid crustaceans. Comparing the relative power of the different characters, e.g., in terms of PBS values and nodes supported in separate analyses of individual genes, shows that the nuclear loci make an overwhelmingly strong contribution to phylogenetic inferences in this study (Tables 4 and 5). Using the OPS and EPRS genes only, all but the two conflicting nodes were resolved with >85% bootstrap support. In contrast, use of the conventional mtDNA and rRNA data alone would leave the relationships among the main *Paramysis* clades unresolved, suggesting a history of an early simultaneous radiation. This realisation provides new prospects for fast progress in tracing peracarid evolutionary histories in species-to-genus-level studies.

The poor phylogenetic performance of rRNA genes in our analysis is traced back to at least four issues. First, there was remarkably little, if any, divergence to provide phylogenetic information at the intra-generic level between most *Paramysis* species characterised by deep divergence in protein genes. Second, just at the next divergence level, i.e., for most inter-generic comparisons, there was extensive length variation which made the alignment unreliable. While potential alignment errors are not likely to have biased conclusions in this study, where the coding genes provided sufficient information to resolve the branching order, this is likely to be a problem in phylogenies based on rRNA alone (e.g., Remerie et al., 2004; Meland and Willassen, 2007). Third, the rRNA rates of evolution in *Paramysis* seem to vary up to an order of magnitude between lineages even at the intra-generic level, as demonstrated by the extreme divergence of the BAERI group from the others. Similar intra-generic 18S rate variation in the mysid genus *Schistomysis* is evident from data of Meland and Willassen (2007). In particular, there was no indication of a regular (e.g., linear) relationship between divergence in protein and rRNA genes, suggesting that rRNA are of little value in relative divergence comparisons or for molecular dating, a practice still widespread in molecular systematics (Jansen et al., 2006; Hendrixson and Bond, 2007). Fourth, although there was little correlation in protein versus rRNA divergence, the rates between the two rRNA genes 18S and 28S appeared strongly linked across lineages. Both genes showed acceleration in the BAERI group. This may be expected in view of the constraints from the close structural interaction of the two molecules in the ribosome, and it further stresses the case of the two genes not representing independent characters in phylogenetic analysis (Dixon and Hillis, 1993).

A potential strength of the conventional mtDNA protein-coding genes, and particularly of COI, has been seen in their combination of fast-evolving silent substitutions and the slower non-synonymous changes, to simultaneously trace both recent and deeper branching events (Simon et al., 2006). For the COI gene and the intra- and inter-generic levels treated here, neither the fast nor slow sites seem very informative. Rather, both are too extreme to provide useful signal. The fast silent rate entails saturation at levels representing most interspecies comparisons, and basically can only separate the youngest divergences (5–10%) from the deeper ones. On the other hand, extreme selective constraints on amino acid evolution of COI (generally considered the most conservative of mtDNA-coded proteins e.g., Mueller, 2006) allows only a few variable characters. The poor COI performance is seen in the bootstrap and PBS values in Tables 4 and 5; it is also evident from comparisons with previous COI-only based assessments of *Paramysis* relationships, which showed little congruence with the present ones (Cristescu and Hebert, 2005; Audzijonyte et al., 2006). Likewise in the analysis of another mysid genus *Mysis*, mtDNA gave little resolution of most nodes, except for the most recent ones where the signal evidently had been strengthened by post-speciation sister-species introgression (Audzijonyte et al., 2005). Such nuclear-mitochon-

drial discordances are frequently found in animal mtDNA (Funk and Omland, 2003); they have the additional effect of distorting the clock-like behaviour of mtDNA, even if the substitution rates would not vary, because species that experienced post-speciation mtDNA introgression will appear unreasonably close in the mtDNA phylogenies. In summary, the strengths of COI in peracarid crustaceans seem to remain at the intraspecific phylogeographic levels, representing the Pleistocene time scale (e.g., Audzijonyte et al., 2006). Our conclusion is not in line with recent overviews that see the COI gene as a generally valuable marker for Cainozoic divergences and for example for nodes up to 40–60 MY old in salamanders (Simon et al., 2006; Mueller, 2006). Notably in our data the use of the model-based analysis did not help to recover the signal from the saturated COI data. Rather, adding the COI gene to the PRO data in the ML analysis decreased the bootstrap support for several nodes.

The rate of nucleotide substitution in nuclear protein-coding genes in *Paramysis* seems to be an order of magnitude slower than in the mtDNA COI (Fig. 2), which is consistent with data from other taxa (Simon et al., 2006). The relationship is similar when considering the 3rd position data only in the three closest sister-species pairs; the average difference is about 8-fold, although different pairs show widely different ratios, ranging from 2- to 15-fold. Between the OPS and EPRS genes there however seems to be a relatively good linear correlation at different levels of divergence; the OPS gene evolves roughly twice faster than the EPRS. From our calibrations these PRN divergence rates would be approximately 0.5% and 1.0% per MY. These markers therefore appear the most useful both for the resolution on Cainozoic phylogeny and in divergence comparisons; a conclusion that contradicts the widely held view that most nuclear genes are of little use at time scales <15 MY (Simon et al., 2006).

The strong performance of the OPS gene was not entirely expected, given the presumably strong differential selection pressures on the encoded opsin protein, responsible for visual adaptations and spectral tuning (Yokoyama, 2000; Porter et al., 2007). While our data do not rule out adaptive change, the good match between the two PRN genes nevertheless suggest that much of the change is driven by neutral divergence. From the best-fit-substitution models in *Paramysis* (Table 2) it indeed appears that while a large part (ca. 50%) of nucleotides are constrained (judged invariable), the probabilities of change at the remaining sites are more even for PRN genes than for COI. Still in several instances one may suspect that selection has been important in shaping OPS variation. Thus, no OPS amino acid changes were seen among the four BA group species (otherwise strongly diverged), while the other, less diverged sister-species pairs had 2–14 differences. Consequently, dates of BA clade diversification inferred from the OPS gene are younger than from other genes. To control such biases, a larger set of independent nuclear loci should be used for a reliable divergence dating.

4.5. Implications for taxonomy

Our results on *Paramysis* phylogeny are generally in a good agreement with the morphology-based subdivision of its Ponto-Caspian members into eight subgenera (Daneliya, 2004), and complement the classification with non Ponto-Caspian taxa. Both molecular and morphological data indicate the position of the Atlantic *P. nouveli* and its morphological Mediterranean sister taxon *P. helleri* in *Longidentia*, together with *P.* (*L.*) *kroyeri*. All three share a similar thin, sharp and prolonged serration of the paradactylar setae of pereiopods and long sharp denticles in an acute-angled telson cleft. The Atlantic *P. bacescoi* is now placed in the previously monotypic *Pseudoparamysis* with *P.* (*P.*) *pontica*, on the grounds of the similarly reduced 5th and 6th carpopropodus of pereiopods and long, thin masculine process of antennules. Finally the freshwater *P. kosswigi* from Turkey shows close affinity to *P. lacustris*, with typical characteristics of the *Serrapalpisis* subgenus: serrated setae of mandibular palp and wide, obtuse-angled telson cleft. While the two species did not differ in the characters used for the present analyses, they do differ in the shape of carapace and paradactylar setae. The single remaining Atlantic–Mediterranean species "*Paramysis*" *arenosa* fell outside other *Paramysis* taxa in our trees, rendering the *Paramysis* as paraphyletic. The species actually appears morphologically closer to *Schistomysis* (Fig. 4); this was already noted by Zimmer (1909), who nevertheless later moved the species to *Paramysis* on the grounds of similarity of the swollen 1st segment of the carpopropodus of pereiopods (Zimmer, 1915). The molecular relationships of "*P.*" *arenosa* to other genera remain ambiguous and for a proper generic assignment further study is needed.

Acknowledgments

We thank all those that contributed to sampling, including T. Aleksenko, A.G. Gorbunov, L. Lévêque, V.A. Minoranskii, and V.A. Ponomarenko, and the collection managers J. Olesen (Natural History Museum of Denmark), V.V. Petryashev (Zoological Institute of Russian Academy of Sciences), and Å. Wilhelmsen (Natural History Museum of University of Oslo), who arranged loans and access to the museum material. We also thank K.J. Wittmann and C. Aygen for donating important samples, M. Ahjos for help with the molecular labwork, and two anonymous referees for valuable comments to the manuscript. The study was supported by grants from the University of Helsinki research funds and the Walter and Andrée de Nottbeck Foundation, and additional travel support from the CORONA network and the Academy of Finland.

Appendix A

Morphological matrix of the studied mysid taxa.

	1									10									20									30				
P.baeri	2	0	1	0	1	0	1	0	1	2	0	1	2	1	0	0	1	1	1	0	2	2	1	4	1	1	1	1	0	0	1	3	2	1	1
P.bakuensis	2	0	1	1	1	0	1	0	1	2	0	1	2	1	0	0	1	1	1	0	2	1	1	4	1	1	1	1	0	0	1	3	2	1	1
P.eurylepis	2	0	1	1	1	0	1	0	1	2	0	1	2	1	0	0	1	1	1	0	2	1	1	4	1	1	1	1	0	0	0	1	3	2	1	1
P.kessleri	2	0	1	1	0	0	1	N	1	2	1	1	2	1	0	0	1	0	1	0	2	1	2	4	1	1	0	0	0	0	1	3	2	1	1
P.grimmi	0	0	0	0	0	0	0	1	2	1	1	1	2	1	0	0	0	1	0	1	0	2	1	0	4	0	0	0	0	0	0	0	3	2	0	1
P.inflata	0	0	0	0	1	0	1	2	1	1	0	1	1	0	0	0	1	0	1	0	2	1	0	4	0	0	0	0	0	0	0	0	3	2	0	1
P.ullskyi	0	0	0	0	1	0	1	2	1	1	0	2	1	0	0	0	1	0	1	0	2	1	0	4	0	0	0	0	0	0	0	0	3	2	0	1
P.incerta	1	1	0	0	0	0	1	2	1	1	0	2	4	1	1	1	1	0	1	0	2	2	2	4	0	0	0	0	1	1	1	3	2	1	1
P.kosswigi	1	1	0	0	0	0	1	2	1	1	0	1	3	1	0	0	1	0	1	0	2	2	2	4	0	0	0	0	1	1	1	3	2	1	1
P.lacustris	1	1	0	0	0	0	1	2	1	1	0	1	3	1	0	0	1	0	1	0	2	2	2	4	0	0	0	0	1	1	1	3	2	1	1
P.sowinskii	1	1	0	0	0	0	1	2	1	1	0	1	4	1	0	1	1	0	1	0	2	2	2	4	0	0	0	0	1	1	1	3	2	1	1
P.intermedia	0	1	0	0	0	0	1	2	1	1	0	1	3	1	0	1	1	0	1	0	2	2	2	4	0	0	0	0	0	0	1	3	2	1	1
P.loxolepis	1	1	0	0	0	0	1	2	1	1	0	1	3	1	1	1	1	0	1	0	2	2	1	4	0	0	0	0	0	0	1	3	2	1	1
P.agigensis	3	0	0	0	0	0	1	2	1	1	0	0	0	1	0	0	1	0	1	0	2	2	2	4	0	0	0	0	0	0	1	3	2	1	1
P.festae	3	0	0	0	0	0	N	N	N	N	N	N	N	1	0	0	1	0	2	0	2	2	N	N	0	0	0	0	0	0	1	N	2	1	1
P.kroyeri	3	0	0	0	0	0	1	2	1	1	0	0	0	1	0	0	1	0	2	0	2	2	1	4	0	0	0	0	0	0	1	3	2	1	1
P.helleri	3	0	0	0	0	0	1	2	1	1	0	0	0	1	0	0	1	0	2	0	2	2	1	4	0	0	0	0	0	0	1	3	2	1	1
P.nouveli	3	0	0	0	1	0	1	2	1	1	0	1	0	1	0	0	1	0	2	0	2	2	1	4	0	0	0	0	0	0	1	3	2	1	1
P.bacesoi	3	0	0	0	0	0	1	2	1	1	0	0	0	1	0	0	1	0	2	0	1	0	1	4	0	0	0	0	0	0	1	4	2	1	1
P.pontica	3	0	0	0	0	0	1	2	1	1	0	0	0	1	0	0	1	0	2	0	1	1	2	4	0	0	0	0	0	0	1	4	2	1	1
P.arenosa	4	0	0	0	0	1	1	1	0	1	0	0	0	1	0	0	1	0	1	1	2	2	2	4	0	0	0	0	0	0	2	2	2	2	1
C.knipowitschi	0	0	0	0	0	0	2	0	1	3	0	1	2	1	2	0	0	0	3	0	1	0	1	3	1	4	0	0	0	0	1	3	2	2	1
K.warpachowskyi	5	0	0	0	0	0	1	0	0	4	0	1	0	1	3	0	0	0	0	0	0	0	1	4	0	0	0	0	0	0	2	3	2	2	1
S.assimilis	4	0	0	0	0	1	1	2	0	1	0	1	0	1	0	0	2	0	1	1	4	2	1	4	1	3	0	0	0	0	2	3	2	2	0
H.anomala	0	0	0	0	0	0	2	0	0	0	0	0	0	1	2	0	3	0	0	0	5	1	0	0	1	4	2	0	0	0	2	1	1	2	0
L.benedeni	2	0	0	0	0	0	0	0	0	5	0	0	3	1	2	2	0	0	0	0	1	0	1	1	1	2	2	0	2	0	2	0	0	2	0
M.segerstralei	3	0	0	0	0	0	N	N	0	N	0	1	3	1	2	2	4	0	0	0	5	1	1	2	1	5	2	0	0	1	3	5	0	2	0

Description of the morphological characters and states

1. Telson, shape of distal margin: 0, truncated; 1, with broad; shallow cleft; 2, with narrow, small cleft; 3, with broad, acute-angled cleft; 4, with deep cleft; 5, rounded.
2. Telson, distal margin, wing-like expansions on the denticles: 0, absent; 1, present.
3. Telson, fine setae on margins: 0, absent; 1, present.
4. Telson, fine setae in the cleft: 0, absent; 1, present.
5. Uropod, endopod, number of spine-setae lines in proximal part: 0, one; 1, two-three.
6. Uropod, endopod, position of spine-setae along the lateral margin: 0, broadly set; 1, closely set;
7. Male pleopod IV, 4th segment: 0, absent; 1, equal or slightly longer than 5th segment; 2, twice longer than 5th segment.
8. Male pleopod IV, size of the distal outer seta: 0, shorter than endopod; 1, equal to endopod; 2, longer than endopod.
9. Penis, the outer blade: 0, weakly developed; 1, large.
10. Penis, the outer blade setae: 0, absent; 1, thin and short in distal part; 2, thin and short together with 1–3 long in distal part; 3, rare short along all the blade; 4, rare short along all the blade and 1–2 long in distal part; 5, many long in distal part.
11. Thoracopods, setae on the exopod basis: 0, absent; 1, present.
12. Pereiopod I, setae on preischium: 0, absent; 1, 1–4; 2, 7.
13. Pereiopod I, setae on the dorsal side of the ischium: 0, few (1–3) distally; 1, many along all side, especially distally; 2, two small groups, central and distal; 3, rare along all side; 4, many proximally.
14. Pereiopods, size of merus: 0, wider or equal width to ischium and twice shorter than ischium; 1, narrower and less than 1.5 times shorter than ischium.
15. Pereiopods, merus, setae on the ventral side: 0, in few bunches (4–8); 1, in many bunches (13–16); 2, in few groups of major and small bunches; 3, many single setae.
16. Pereiopods, carpopropodus, setae on ventral side of the subsegments: 0, absent; 1, present on all subsegments; 2, present only on 1st subsegment.
17. Pereiopod I, carpopropodus, number of subsegments: 0, three; 1, four; 2, four–five; 3, five; 4, seven.
18. Pereiopod I, carpopropodus, the last subsegment: 0, slightly shorter than previous subsegment; 1, less than a half of previous subsegment's length.
19. Pereiopod I, denticles on paradactylar claw-setae: 0, thin, short; 1, strong, straight; 2, thin, long, curved forward; 3, strong, sabre-like, curved backward.
20. Pereiopods, dactylar setae: 0, smooth; 1, serrated.
21. Pereiopod VI, carpopropodus, number of subsegments: 0, one; 1, three; 2, four; 3, five; 4, five–six; 5, six.
22. Pereiopod VI, paradactylar claw-setae: 0, rudimentary; 1, with fine serration; 2, with strong serration.
23. Maxillipede I, merus: 0, length 1.4–1.6 of width; 1, length 1.8–2.0 of width; 2, length 2.1–2.4 of width.

24. Maxillipede I, endites: 0, almost undeveloped; 1, small, about the width of segments; 2, moderately larger, wider than segments; 3, very big, as long as merus; 4, endite of ischium big and endite of preischium small.
25. Maxilla II, terminal endopod spine-setae: 0, absent; 1, present.
26. Maxilla II, terminal endopod spine-setae: 0, absent; 1, two-three short distal; 2, many short along outer margin; 3, many long and short; 4, long; 5, long, serrated.
27. Maxilla II, exopod: 0, narrow; 1, wide; 2, large; pyriform.
28. Maxilla II, length of exopod setae: 0, all equal; 1, proximal long, distal short.
29. Mandibular palp, 2nd segment, serration on setae: 0, absent; 1, along anterior margin of setae; 2, along posterior margin of setae.
30. Mandibular palp, 2nd segment, disto-lateral setae: 0, absent; 1, present.
31. Antennular peduncle: 0, half of antennal peduncle; 1, slightly shorter than antennal peduncle; 2, longer than antennal peduncle; 3, as long as antennal peduncle.
32. Antennular male process: 0, short, thin, makes claw with antennular prominence; 1, shorter than 3rd segment of peduncle, wide; 2, as long as 2nd and 3rd segments together; 3, longer than 2nd and 3rd segments; 4, as long as peduncle; 5, as long as peduncle, conical without setae.
33. Antennal scale, outer margin: 0, with setae all along; 1, with setae along half of the length; 2, bare with terminal spine.
34. Carapace, frontal margin: 0, straight; 1, convex; smoothly rounded; 2, convex, angled.
35. Carapace, subrostral plate: 0, absent; 1, present.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2007.11.009.

References

Ariani, A.P., Wittmann, K.J., Franco, E., 1993. A comparative study of static bodies in mysid crustaceans: evolutionary implications of crystallographic characteristics. Biological Bulletin 185, 393–404.

Audzijonyte, A., Damgaard, J., Varvio, S.-L., Vainio, J.K., Väinölä, R., 2005. Phylogeny of *Mysis* (Crustacea, Mysida): history of continental invasions inferred from molecular and morphological data. Cladistics 21, 575–596.

Audzijonyte, A., Daneliya, M.E., Väinölä, R., 2006. Comparative phylogeography of Ponto-Caspian mysid crustaceans: isolation and exchange among dynamic inland sea basins. Molecular Ecology 15, 2969–2984.

Audzijonyte, A., Wittmann, K.J., Väinölä, R., 2008. Tracing recent invasions of the Ponto-Caspian mysid shrimp *Hemimysis anomala* across Europe and to North America with mitochondrial DNA. Diversity and Distributions, doi:10.1111/j.1472-4642.2007.00434.x.

Baker, R.H., DeSalle, R., 1997. Multiple sources of character information and the phylogeny of Hawaiian drosophilids. Systematic Biology 46, 654–673.

Băcescu, M., 1954. Mysidacea. In: Fauna Republicii Populare Romine, Crustacea, 4(3), 1–126.

Bănărescu, P., 1991. Zoogeography of fresh waters. Distribution and dispersal of freshwater animals in North America and Eurasia, vol. 2. Aula Verlag, Wiesbaden.

Bondarenko, M.V., 1991. Mizidy Kaspiya i ikh rol v ekosistemakh morya. Dissertation Summary. Institute of Oceanology, Moscow.

Brooks, J.L., 1950. Speciation in ancient lakes. The Quarterly Reviews of Biology 25, 131–176.

Cristescu, M.E.A., Hebert, P.D.N., 2002. Phylogeny and adaptive radiation in the Onychopoda (Crustacea, Cladocera): evidence from multiple gene sequences. Journal of Evolutionary Biology 15, 838–849.

Cristescu, M.E.A., Hebert, P.D.N., 2005. The 'Crustacean Seas'—an evolutionary perspective on the Ponto-Caspian peracarids. Canadian Journal of Fisheries and Aquatic Sciences 62, 505–517.

Czerniavsky, V.I. 1882. Monographia Mysidarum inprimis Imperii Rossici. Fasc. 1, 2. Trudy Sankt-Peterburgskago Obshchestva Estestvoispytatelei 12, 1–170; 13, 1–85.

Damgaard, J., Cognato, A.I., 2003. Sources of character conflict in a clade of water striders (Heteroptera: Gerridae). Cladistics 19, 512–526.

Daneliya, M.E., 2004. K sistematike mizid roda *Paramysis* (Crustacea, Mysidacea) iz basseina Ponto-Kaspiya. Zoologicheskii Zhurnal 83, 408–416.

Daneliya, M.E., Audzijonyte, A., Väinölä, R., 2007. Diversity within the Ponto-Caspian *Paramysis baeri* Czerniavsky sensu lato revisited: *P. bakuensis* G.O. Sars restored (Crustacea: Mysida: Mysidae) Zootaxa 1632, 21–36.

De Rijk, P., De Watcher, R., 1997. RnaViz, a program for the visualization of RNA secondary structure. Nucleic Acids Research 25, 4679–4684.

Derzhavin, A.N., 1939. Mizidy Kaspiya. Izdatelstvo AzFAN, Baku, Azerbaijan.

Dixon, M.T., Hillis, D.M., 1993. Ribosomal RNA secondary structure: compensatory mutations and implications for phylogenetic inference. Molecular Biology and Evolution 10, 256–267.

Dumont, H.J., 1998. The Caspian Lake: History, biota, structure, and function. Limnology and Oceanography 43, 44–52.

Dumont, H.J., 2000. Endemism in the Ponto-Caspian fauna, with special emphasis on the Onychopoda (Crustacea). Advances in Ecological Research 31, 181–196.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32, 1792–1797.

Ekman, S., 1953. Zoogeography of the Sea. Sidgwick and Jackson, London.

Emerson, B.C., 2007. Alarm bells for the molecular clock? No support for Ho et al.'s model of time-dependent molecular rate estimates. Systematic Biology 56, 337–345.

Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes and consequences, with insights from animal mitochondrial DNA. Annual Review of Ecology and Systematics 34, 397–423.

Gatesy, J., O'Grady, P., Baker, R.H., 1999. Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. Cladistics 15, 271–313.

Goloboff, P.A., 1993. NONA. Program and documentation distributed via www.cladistics.com.

Grigorovich, I.A., Therriault, T.W., MacIsaac, H.J., 2003. History of aquatic invertebrate invasions in the Caspian Sea. Biological Invasions 5, 103–115.

Hendrixson, B.E., Bond, J.E., 2007. Molecular phylogeny and biogeography of an ancient Holarctic lineage of mygalomorph spiders (Araneae: Antrodiaetidae: Antrodiaetus). Molecular Phylogenetics and Evolution 42, 738–755.

Jansen, G., Devaere, S., Weekers, P.H.H., Adriaens, D., 2006. Phylogenetic relationships and divergence time estimate of African anguilliform catfish (Siluriformes: Clariidae) inferred from ribosomal gene and spacer sequences. Molecular Phylogenetics and Evolution 38, 65–78.

Kamaltynov, R.M., 1999. On the higher classification of Lake Baikal amphipods. Crustaceana 72, 933–944.

Kasymov, A.G., 1987. Zhivotnyi mir Kaspiiskogo Morya. Izdatelstvo Elm, Baku, Azerbaijan.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. Journal of Molecular Evolution 29, 170–179.

Koblmüller, S., Duftner, N., Sefc, K.M., Aibara, M., Stipacek, M., Blanc, M., Egger, B., Sturmbauer, C., 2007. Reticulate phylogeny of gastropod-shell-breeding cichlids from Lake Tanganyika—the result of repeated introgressive hybridization. BMC Evolutionary Biology 7, 7.

Komarova, T.I., 1991. Mizidy. In: Fauna Ukrainy, Crustacea, 1–101.

Koretsky, I.A., 2001. Morphology and systematics of Miocene Phocinae (Mammalia: Carnivora) from Paratethys and the North Atlantic region. Geologica Hungarica, Seria Paleontologica 54, 1–109.

Krijgsman, W., Hilgen, F.J., Raffi, I., Sierro, F.J., Wilson, D.S., 1999. Chronology, causes and progression of the Messinian salinity crisis. Nature 400, 652–655.

Kumar, S., Tamura, K., Nei, M., 2004. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5, 150–163.

Maddison, W.A., 1997. Gene trees in species trees. Systematic Biology 46, 523–536.

Martens, K., 1997. Speciation in ancient lakes. Trends in Ecology and Evolution 12, 177–182.

Meland, K., Willassen, E., 2007. The disunity of "Mysidacea" (Crustacea). Molecular Phylogenetics and Evolution 44, 1083–1104.

Mordukhai-Boltovskoi, F.D., 1979. Composition and distribution Caspian fauna in the light of modern data. Internationale Revue der Gesamten Hydrobiologie und Hydrographie 64, 1–38.

Moulton, J.K., Wiegman, B.M., 2007. The phylogenetic relationships of flies in the superfamily Empidoidea (Insecta: Diptera). Molecular Phylogenetics and Evolution 43, 701–713.

Mueller, R.L., 2006. Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. Systematic Biology 55, 289–300.

Mudie, P.J., Rochon, A., Aksu, A.E., Gillespie, H., 2002. Dinoflagellate cysts, freshwater algae and fungal spores as salinity indicators in Late Quaternary cores from Marmara and Black seas. Marine Geology 190, 203–231.

Nevesskaya, L.A., Popov, S.V., Goncharova, I.A., Ilyina, L.B., Paramonova, N.P., 2006. Uskarionaya evolyutsiya mollyuskov Vostochnogo Paratetisa v uslovyakh ponizhenoi konkurentsii. In: Rozhnov, S.V. (Ed.), Evolyutsiya biosfery i bioraznoobraziya. Tovarishchestvo nauchnykh izdanii KMK, Moscow, pp. 334–358.

Nixon, K.C., 2002. Winclada, Version 1.00.08. Published by the author, Ithaca, NY.

Palo, J.U., Väinölä, R., 2006. The enigma of the landlocked Baikal and Caspian seals addressed through phylogeny of phocine mitochondrial sequences. Biological Journal of the Linnean Society 88, 61–72.

Popov, S.V., Shcherba, I.G., Ilyina, L.B., Nevesskaya, L.A., Paramonova, N.P., Khondkarian, S.O., Magyar, I., 2006. Late Miocene to Pliocene palaeogeography of the Paratethys and its relation to the Mediterranean. Paleogeography Paleoclimatology Paleoecology 238, 91–106.

Porter, M.L., Cronin, T.W., McClellan, D.A., Crandall, K.A., 2007. Molecular characterization of crustacean visual pigments and the evolution of pancrustacean opsins. Molecular Biology and Evolution 24, 253–268.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Posada, D., Crandall, K.A., 2001. Selecting the best-fit model of nucleotide substitution. Systematic Biology 50, 580–601.

Ray, C.E., 1976. Geography of phocid evolution. Systematic Zoology 25, 391–406.

Reid, D.F., Orlova, M.I., 2002. Geological and evolutionary underpinnings for the success of Ponto-Caspian species invasions in the Black Sea and North American Great Lakes. Canadian Journal of Fisheries and Aquatic Sciences 59, 1144–1158.

Remerie, T., Bulckaen, B., Calderon, J., Deprez, T., Mees, J., Vanfleteren, J., Vanreusel, A., Vierstraete, A., Vincx, M., Wittmann, K.J., Wooldridge, T., 2004. Phylogenetic relationships within the Mysidae (Crustacea, Peracarida, Mysida) based on nuclear 18S ribosomal RNA sequences. Molecular Phylogenetics and Evolution 32, 770–777.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Ronquist, F., Huelsenbeck, J.P., van der Mark, P., 2005. MrBayes 3.1. Available from: <mrbayes.csit.fsu.edu/mb3.1_manual.pdf/>.

Ryan, W.B.F., Major, C.O., Lericolais, G., Goldstein, S.L., 2003. Catastrophic flooding of the Black Sea. The Annual Review of Earth and Planetary Science 31, 525–554.

Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Molecular Biology and Evolution 19, 101–109.

Sanderson, M., 2006. r8s. Version 1.71. Available from: <http://ginger.ucdavis.edu/r8s/>.

Sars, G.O., 1907. Mysidae. Trudy Kaspiiskoi expeditsii 1904 goda 1, 242–313.

Schön, I., Martens, K., 2004. Adaptive, pre-adaptive and non-adaptive components of radiations in ancient lakes: a review. Organisms Diversity and Evolution 4, 137–156.

Seehausen, O., 2004. Hybridization and adaptive radiation. Trends in Ecology and Evolution 19, 198–207.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Systematic Biology 51, 492–508.

Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.

Simon, C., Buckley, T.R., Frati, F., Stewart, J.B., Beckenbach, A.T., 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. Annual Review of Ecology, Evolution and Systematics 37, 545–579.

Starobogatov, Y.I., 1970. Fauna molluskov i zoogeograficheskoe raionirovanie kontinentalnykh vodoemov zemnogo shara. Nauka, Leningrad.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tarasov, A.G., 1996. Donnaya fauna glubokovodnykh vpadin Kaspiya. 1. Biologicheskoe raznoobrazie. Zoologicheskii Zhurnal 75, 1763–1775.

Tarasov, A.G., 1997. Donnaya fauna glubokovodnykh vpadin Kaspiya. 2. Genezis i vertikalnaya zonalnost. Zoologicheskii Zhurnal 76, 5–15.

Therriault, T.W., Docker, M.F., Orlova, M.I., Heath, D.D., MacIsaac, H.J., 2004. Molecular resolution of the family Dreissenidae (Mollusca: Bivalvia) with emphasis on Ponto-Caspian species, including first report of *Mytilopsis leucophaeata* in the Black Sea basin. Molecular Phylogenetics and Evolution 30, 479–489.

Thomson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research 22, 4673–4680.

Thorne, J.L., Kishino, H., 2002. Divergence time and evolutionary rate estimation with multilocus data. Systematic Biology 51, 689–702.

Väinölä, R., 1995. Origin and recent endemic divergence of a Caspian *Mysis* species flock with affinities to the "glacial relict" crustaceans in boreal lakes. Evolution 49, 1215–1223.

Vasiliev, I., Krijgsman, W., Stoica, M., Langereis, C.G., 2005. Mio-Pliocene magnetostratigraphy in the southern Carpathian foredeep and Mediterranean–Paratethys correlations. Terra Nova 17, 376–384.

Williams, S.T., Knowlton, N., Weigt, L.A., Jara, J.A., 2001. Evidence for three major clades within the snapping shrimp genus *Alpheus* inferred from nuclear and mitochondrial gene sequence data. Molecular Phylogenetics and Evolution 20, 375–389.

Wilson, A.B., Glaubrecht, M., Meyer, A., 2004. Ancient lakes as evolutionary reservoirs: evidence from the thalassoid gastropods of Lake Tanganyika. Proceedings of the Royal Society of London B Biological Sciences 271, 529–536.

Wittmann, K.J., Schlacher, T.A., Ariani, A.P., 1993. Structure of recent and fossil mysid statoliths (Crustacea, Mysidacea). Journal of Morphology 215, 31–49.

Yokoyama, S., 2000. Molecular evolution of vertebrate visual pigments. Progress in Retinal and Eye Research 19, 385–419.

Zenkevitch, L., 1963. Biology of the Seas of the USSR. Allen and Unwin, London, UK.

Zimmer, C., 1909. Die nordischen Schizopoden. VI. Nordisches Plankton 3, 1–179.

Zimmer, C., 1915. Die Systematik der Tribus Mysini H.J. Hansen. Zoologischer Anzeiger 46, 202–216.

Zubakov, V.A., 2001. History and causes of variations in the Caspian Sea level: the Miopliocene, 7.1–1.95 million years ago. Water Research 28, 249–256.