Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 39 (2006) 276–281
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Short communication

Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex

J.I. Robalo *, C. Sousa Santos, A. Levy, V.C. Almada

Instituto Superior de Psicologia Aplicada, Unidade de Investigação em Eco-Etologia, Rua Jardim do Tabaco 34, 1149-041 Lisbon, Portugal

Received 23 June 2005; revised 9 August 2005; accepted 22 August 2005
Available online 4 October 2005

1. Introduction

Many hybrid fish lineages result from intrageneric crosses and are often asexual, being considered as "evolutionary dead-ends" (Vrijenhoek, 1998). This seems not to be true in the *Squalius alburnoides* (Iberian minnow) Steindachner 1866 complex. This complex, that seems to have originated in a cross between distant species, has a remarkable variability in ploidy level and genomic composition, includes fertile fish of both sexes and some of its forms retain meiosis and recombination (Alves et al., 2001).

The *S. alburnoides* complex probably originated from crosses between *S. pyrenaicus* females and males from an unknown species. The complex includes $2n = 50$, $3n = 75$, and $4n = 100$ hybrid forms (reviewed in Alves et al., 2001) with varying proportions of the two parental genomes (denoted by P and A, corresponding to the *S. pyrenaicus* and paternal ancestor genomes, respectively). Reconstituted diploid non-hybrids (AA genome) are also produced and are morphologically distinct from the diploid hybrid form of the complex (PA). These non-hybrid fish, normally males (females seem to be extremely rare), exhibit the nuclear genome of the paternal ancestor (AA) and *S. pyrenaicus*-like mtDNA. In the absence of an identified paternal species this suggests that they are reconstituted from the hybrids (Alves et al., 2002). The oogenesis of triploid females with PAA genomes frequently involves discarding the *Squalius* (P genome), followed by normal meiosis and recombination, generating A gametes. When these gametes fuse with the sperm of AA males, which also undergo normal meioses, new AA male progeny is generated. In the absence of *S. pyrenaicus*, such as in the northern basins of Portugal, this complex seems to be maintained by crosses with males of *S. carolitertii* (CC) and by diploid hybrid males (CA), although the mtDNA found in *S. alburnoides* fish is *S. pyrenaicus*-like (Cunha et al., 2004; Pala and Coelho, 2005).

The Leuciscini presently found in the Iberian Peninsula include mainly species of the genera *Squalius*, *Chondrostoma*, and *Anaecypris*. Studies based on allozymes showed that *S. alburnoides* did not originate from intrageneric crosses between *Squalius* species and also ruled out members of the genera *Chondrostoma* and *Anaecypris* as paternal ancestors (Alves et al., 1997; Carmona et al., 1997). Recent studies still unpublished (Gilles, Dowling, Alves, Coelho, and Collares-Pereira) using introns from two nuclear genes seem to suggest that perhaps *A. hispanica*-like individuals may represent the paternal ancestor of this complex.

The aim of the present work was to investigate which of the genera considered is phylogenetically closest to the paternal species that originated this complex. This approach is based on the amplification of a segment of the β-actin nuclear gene from a number of genera closely related to *S. pyrenaicus* and from the non-hybrid males of *S. alburnoides*. The topology obtained with β-actin was compared to relatively complete phylogenies of the European cyprinids based on cyt *b* (e.g., Briolay et al., 1998; Gilles et al., 1998; Zardoya and Doadrio, 1999) and with our own reconstruction using a set of species comparable to that used with β-actin.

2. Methods

Total genomic DNA was extracted from fin clips preserved in ethanol by an SDS/proteinase-k-based protocol (adapted from Sambrook et al., 1989).

For the β-actin gene a total of 1062 bp was amplified using the primers For–5′-ATGGATGATGAAATTGCC

* Corresponding author. Fax: +351 218 860 954.
E-mail address: jrobalo@ispa.pt (J.I. Robalo).

1055-7903/$ - see front matter
doi:10.1016/j.ympev.2005.08.009

GC-3′ and Rev–5′AGGATCTTCATGAGGTAGTC-3′ (J. Robalo, unpublished). The amplification process was conducted as follows: 35 cycles of 94 °C (30 s), 55 °C (40 s), and 72 °C (1 min 30 s). Further details may be requested from the corresponding author. The amplified fragment is homologous to a region of the β-actin gene of *Cyprinus carpio* (GenBank Accession No. M24113), including introns B and C and three exons.

For the cyt *b* gene a total of 1044 bp was amplified using the primers LCB1–5′ AATGACTTGAAGAACCACC GT-3′ (Brito et al., 1997) and HA–5′- CAACGATCTCCG GTTTACAAGAC-3′ (Schmidt and Gold, 1993). PCR conditions followed those in Cunha et al. (2004).

For both genes, each sample was sequenced in both directions with the same primers used for PCR. Sequencing reactions were performed by Macrogen Inc. in a MJ Research PTC-225 Peltier Thermal Cycler using a ABI PRISM BigDyeTM Terminator Cycle Sequencing Kits with AmpliTaq DNA polymerase (FS enzyme) (Applied Biosystems), following the protocols supplied by the manufacturer.

All accession numbers from the present and previous studies are listed in Table 1.

Sequences were aligned with BioEdit v.5.0.6.

The aligned sequences were analysed using distance (minimum evolution, ME), maximum likelihood (ML), maximum parsimony (MP), and Bayesian methods. For ME and ML analyses, we performed a hierarchical likelihood ratio test (LRT), using the program ModelTest 3.6 (Posada and Crandall, 1998) to find the model of evolution that best fitted our data. These phylogenetic analyses were performed using PAUP 4.0 (Swofford, 1998) and MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Bootstrap analysis were used to assess the relative robustness of branches of the MP (1000 replicates), the ME (1000 replicates), and the ML (100 replicates) trees. For the Bayesian analysis, cyt *b* data were partitioned by codon base position, and a GTR+I+Γ model was used for third base positions and a HKY model for first and second base positions. For β-actin, gaps were coded as separate characters (Simmons and Ochoterena, 2000), using GapCoder (Young and Healy, 2003). Data were partitioned and separate models used for each region, thus: GTR+I+Γ model used for third base positions of exon and intron region, HKY model for first and second base positions of exons, and single rate for gap characters. For both genes, four separate analyses were performed (one with four million generations and three with one million generations) with four chains per analysis. The first 50,000 generations were discarded as "burn-in," the remaining generations were sampled every 100 generations, and majority-rule consensus trees were calculated from samples at stationarity.

To compare topologies recovered with both genes, the same species were used in both analyses, except for the genera *Phoxinus* and *Misgurnus* where, for lack of corresponding sequences for both fragments, we used different congeneric species.

We decided to include a phylogeny for the cyt *b* gene because we could not get β-actin sequences for all the species used in previous cyt *b* studies. Thus, there was a risk that our reduced data set for β-actin was not directly comparable to cyt *b* studies with more taxa and different outgroups.

Species from the genus *Misgurnus* (Cobitidae) were used as outgroups.

The sequence of *S. alburnoides* chosen to integrate this paper is the most common haplotype found through out the species area of distribution, both in hybrids and in reconstituted non-hybrid males, and was recovered from 51 fishes out of 103 (unpublished data). All other haplotypes differ from this most common one by few mutational steps that rarely exceed one and are unlikely to affect the results presented. This sequence was found in homozygous condition in many reconstituted non-hybrid males which allowed us to reconstruct it from the chromatograms without ambiguities. As we also obtained homozygous sequences for the maternal ancestor of *S. alburnoides*, *S. pyrenaicus*, we were in a position to identify the overlapping patterns of the two species when they were present in hybrids. In this respect, the β-actin gene proved a very useful marker because it possesses a combination of three very convenient features: (a) well-preserved exons that provide good landmarks to align the sequences and identify homologous regions; (b) introns that are sufficiently

Table 1
Species considered in this study and their GenBank accession numbers (cyt *b* and β-actin gene)

Species name	GenBank accession number	
	Cyt *b* gene	β-actin gene
Phoxinus lagowskii steindachneri	AB162650	
Misgurnus anguillicaudatus	AF051868	
Misgurnus mizolepis		AF270649
Leuciscus idus	AY026397	DQ061947[a]
Scardinius erythrophthalmus	AY509848	DQ061949[a]
Phoxinus oxycephalus		AF200957
Anaecypris hispanica	AJ427814	DQ061936[a]
Chondrostoma genei	AF533766	DQ061938[a]
Chondrostoma lemmingii	DQ089654[a]	DQ061940[a]
Chondrostoma lusitanicum	AY254584	DQ061941[a]
Chondrostoma oligolepis	DQ061932[a]	DQ061942[a]
Chondrostoma polylepis	AF045982	DQ061945[a]
Chondrostoma prespensis	AF090747	DQ061944[a]
Chondrostoma soetta	AY568623	DQ061939[a]
Chondrostoma turiensis	AY568619	DQ061946[a]
Telestes souffia	AY509862	DQ061950[a]
Rutilus rutilus	DQ061933[a]	DQ061948[a]
Squalius aradensis	AF421825	DQ061937[a]
Squalius carolitertii	AF045994	AY943882[a]
Squalius pyrenaicus	AF421826	AY943882[a]
Squalius torgalensis	DQ061934[a]	DQ061937[a]
Squalius alburnoides		AY943863[a]
Cyprinus carpio	NC001606	M24113
Gobio gobio	AY426589	DQ061935[a]

Note. DQ061937 and AY943882 are haplotypes found for the β-actin gene that are shared by *S. aradensis–S. torgalensis* and *S. pyrenaicus–S. carolitertii*, respectively.

[a] Result from this work.

variable to accumulate information useful in phylogenetic reconstructions; and (c) the presence of frequent indels. These indels, when in heterozygous condition, make possible the reconstruction of the two sequences that overlap in the chromatogram of a hybrid fish. Indeed, was recently published a method that permits the recognition and reconstruction of different nuclear DNA sequences in the same chromatogram through the detection of indels (Bhangale et al., 2005) that with some alterations (Sousa Santos et al., in press) makes possible to distinguish copies of the gene of maternal origin (similar to those found in *S. pyrenaicus*) and reconstruct the remaining ones, of paternal origin.

For both genes, the saturation of transitions and transversions was checked by plotting the absolute number of changes of each codon position against uncorrected sequence divergence values (*p*). There was no evidence of saturation in the ingroup (graph not shown).

3. Results

3.1. β-Actin gene

Among all the sequences studied, 283 sites were variable and 77 were parsimony informative.

The general time-reversible model with among-site rate heterogeneity HKY+G (HKY, Hasegawa et al., 1985) was selected by ModelTest as the best fit to the data. Base frequencies were A = 0.22, C = 0.24, G = 0.20, and T = 0.34. Among-site rate variation was approximated with the gamma distribution shape parameter $\alpha = 0.27$. The proportion of invariable sites (I) was 0.7.

MP analysis resulted in a consensus tree of 392 steps (Consistency Index, CI = 0.87; Homoplasy Index, HI = 0.12; Retention Index, RI = 0.70). The results of the four phylogenetic inference methods are summarized in Fig. 1.

3.2. Cyt b gene

Among all the sequences studied, 445 sites were variable and 363 were parsimony informative.

The general time-reversible model with among-site rate heterogeneity GTR+I+G (Lanave et al., 1984; Yang, 1994) was selected by ModelTest as the best fit to the data. The rate matrix parameters estimated were: $R(a) = 1.10$, $R(b) = 38.6$, $R(c) = 0.63$, $R(d) = 1.81$, and $R(e) = 1.0$. Base frequencies were A = 0.29, C = 0.31, G = 0.11, and $T = 0.28$. Among-site rate variation was approximated with the gamma distribution shape parameter $\alpha = 1.02$. The proportion of invariable sites (I) was 0.52.

MP analysis resulted in a consensus tree of 1596 steps (CI = 0.43; Homoplasy Index, HI = 0.57; Retention Index, RI = 0.41). The results of the four phylogenetic inference methods are summarized in Fig. 2.

Phylogenetic reconstructions with all inference methods show some important congruent features for both genes (Figs. 1 and 2) and for the most part agreed with the findings of Briolay et al. (1998), Gilles et al. (1998), and Zardoya and Doadrio (1998, 1999).

The division between the subfamilies Cyprininae and Leuciscinae was recovered with Cyprininae in a basal position.

Gobio gobio is more closely related to Leuciscinae than to Cyprininae.

Phoxinins and leuciscins emerge in a single clade, although in the present work the position of the *Phoxinus* species varied according with the reconstruction method.

The β-actin gene also confirms the polyphyly of the old genus *Leuciscus*, *Telestes* (= *Leuciscus*) *souffia* being undoubtedly related with the genus *Chondrostoma*. *Leuciscus idus* remains in an unresolved situation, at least with the taxa included in this analysis. The Iberian *Squalius* remain divided in the two groups: *S. aradensis–S. torgalensis* and *S. pyrenaicus–S. carolitertii*.

In the monophyletic genus *Chondrostoma* the β-actin gene did not resolve the polytomy found in the previous cyt *b* studies of Doadrio and Carmona (2004), Durand et al. (2003), and Zardoya and Doadrio (1999).

Trees derived from the two genes disagree in the placement of *Scardinius erythrophtalmus*.

In the β-actin tree, *S. alburnoides* is, according with all methods, strongly associated with *Anaecypris hispanica* (in the cyt *b* tree it is not present because its mitochondrial DNA is *Squalius*-like).

4. Discussion

The results of the present study clearly point to two main conclusions: (1) the paternal ancestor of *S. alburnoides* did not belong to the genera *Squalius* or *Chondrostoma* and (2) of the species included in this study, *A. hispanica* is the closest to the paternal ancestor (as already suggested by Gilles, Dowling, Alves, Coelho, and Collares-Pereira, unpublished).

During the history of *S. alburnoides* some recombination may have taken place between the *Squalius* and *alburnoides* genomes and thus the β-actin sequences of the reconstituted non-hybrid males may not be identical to the sequence of the paternal ancestor. The patristic distance between *S. alburnoides* and *Squalius pyrenaicus*/*S. carolitertii* β-actin sequences is 1.48% while that between *A. hispanica* and *S. pyrenaicus*/*S. carolitertii* is 1.80%. The *p* distance between *S. alburnoides* and *A. hispanica* is 1.48%. Although the hypothesis of some recombination cannot be excluded, the fact that all inference methods recovered the *S. alburnoides* β-actin sequence with that of *A. hispanica*, with very high bootstrap support, indicates that the signal still present in the β-actin gene is sufficiently strong to trace its phylogenetic relationships.

Our suggestion is that, in the past, one species of the same clade of *A. hispanica* may have hybridized with *S. pyrenaicus* and originated *S. alburnoides*. The morphology of the reconstituted AA males supports this conclusion: they are similar to *A. hispanica* in size, general body shape, and coloration, but they also differ from it in several important characters

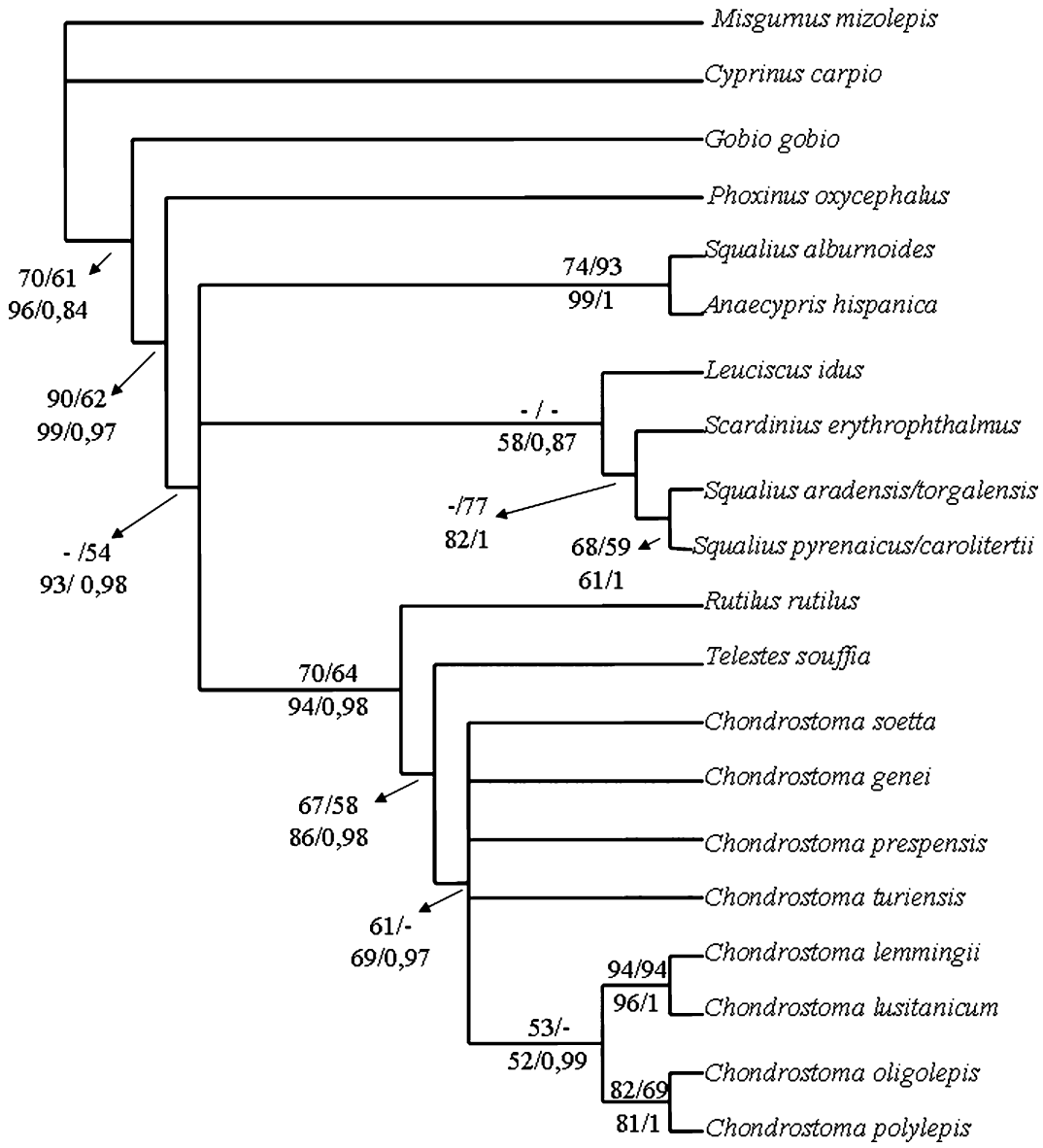

Fig. 1. This figure is based on the MP tree. Phylogenetic relationships among the species based on cyt *b* sequences. For each branch, numbers above represent the bootstrap values obtained for ME and ML; numbers below indicate those for MP and the posterior probabilities for Bayesian inference.

(e.g., structure of the lateral line, lateral scale counts, and number of gill rakers) (Collares-Pereira, 1983). The recent discovery of frequent natural intergeneric hybrids between *Squalius cephalus* and a species of *Chalcalburbus* (Ünver and Erk'Akan, 2005), a member of the same clade as *A. hispanica* (Zardoya and Doadrio, 1999), supports our hypothesis of an intergeneric hybrid origin.

In spite of its hybrid origin *S. alburnoides* seems to be, from an evolutionary perspective, a very successful fish. It is often much more abundant than other sympatric *Squalius*. Alves et al. (2001) argue that the continuous shifting between forms, with P and A nuclear genomes being cyclically lost, gained or replaced by new genomes, allows the introduction of new genetic material. The evolutionary potential of this species, in terms of recombination and maintenance of genetic variability, may be even enhanced by the presence of tetraploid fishes of both sexes in some natural populations. Indeed, PPAA tetraploids seem to undergo normal meiosis. Crosses between them mean that sexual reproduction is restored, while their crosses with other forms of the complex are a way to introduce recombination in the whole system (Alves et al., 2001).

The β-actin gene evolves at a much slower rate than the cyt *b* gene. For example, the difference between *S. pyrenaicus* and *S. carolitertii* was 6.13% for the cyt *b* gene, while the majority of fish from both species share the same haplotype for the β-actin gene. The distance between species of *Squalius* and *Chondrostoma* averaged 14.05% for the cyt *b* and 1.77% for β-actin gene. Because of its slow rate of evolution, the β-actin may prove potentially very useful in studies of cyprinid phylogeny, particularly in resolving intrageneric and intertribe relationships.

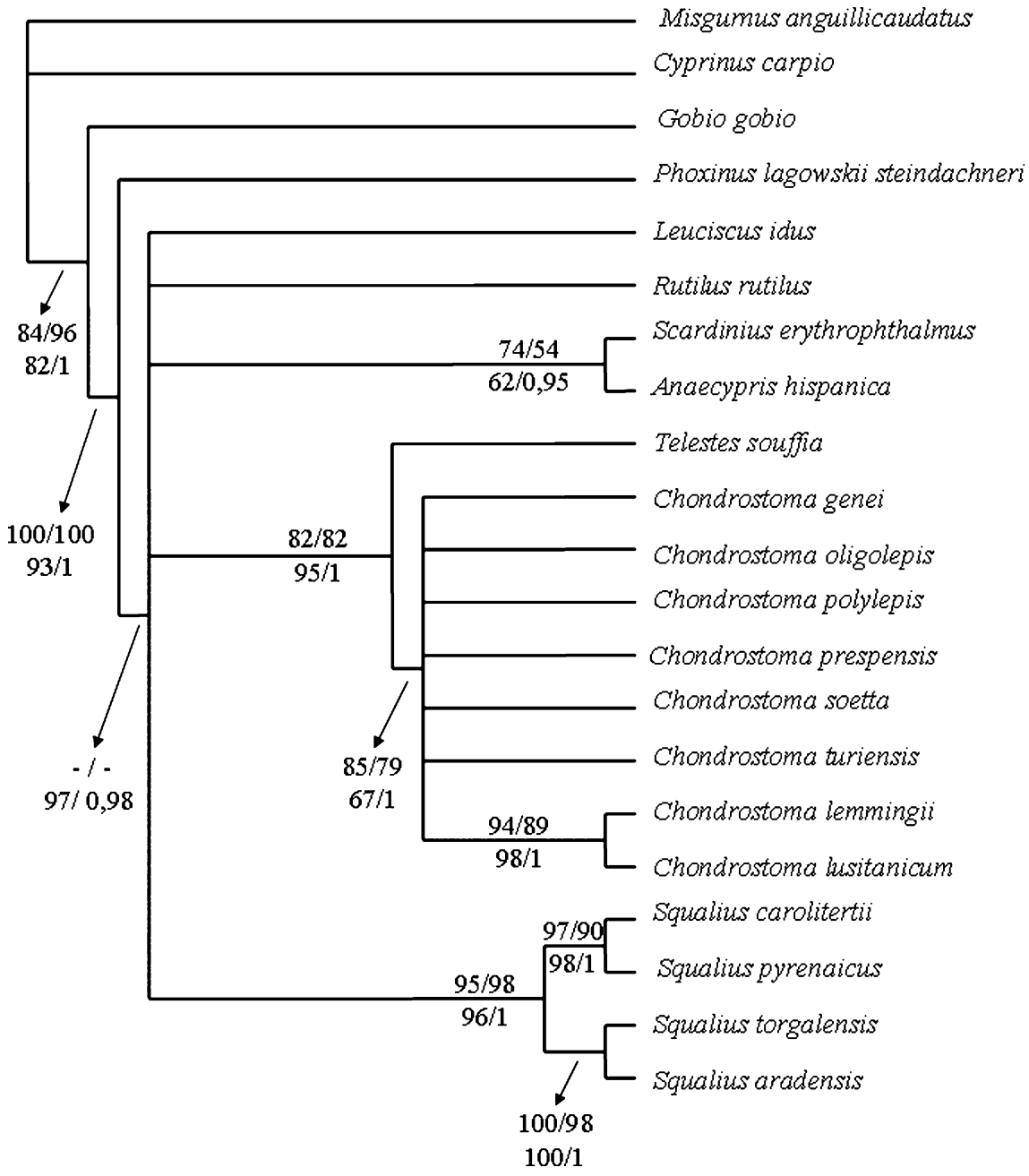

Fig. 2. This figure is based on the MP tree. Phylogenetic relationships among the species based on β-actin sequences. For each branch, numbers above represent the bootstrap values obtained for ME and ML; numbers below indicate those for MP and the posterior probabilities for Bayesian inference.

Acknowledgments

We thank G. Lemos, Sousa-Santos family, and T.Bento for all the help in field work and fish maintenance. We also thank I. Doadrio for the samples of European cyprinids. DGF provided all authorizations for field work. This study was funded by the Pluriannual Program (FCT, UI&D 331/94 with FEDER participation). C. Sousa-Santos and A. Levy were supported by FCT grants (SFRH/BD/8320/2002 and SFRH/BPD/18067/2004, respectively).

References

Alves, M.J., Coelho, M.M., Collares-Pereira, M.J., 1997. The *Rutilus alburnoides* complex (Cyprinidae): evidence for a hybrid origin. J. Zool. Syst. Evol. Res. 35, 1–10.

Alves, M.J., Coelho, M.M., Collares-Pereira, M.J., 2001. Evolution in action through hybridisation and polyploidy in an Iberian freshwater fish: a genetic review. Genetica 111, 375–385.

Alves, M.J, Collares-Pereira, M.J., Dowling, T.E., Coelho, M.M., 2002. The genetics of maintenance of an all-male lineage in the *Squalius alburnoides* complex. J. Fish Biol. 60, 649–662.

Bhangale, T.R., Rieder, M.J., Livingston, R.J., Nickerson, D.A., 2005. Comprehensive identification and characterization of diallelic inser-

tion-deletion polymorphisms in 330 human candidate genes. Hum. Mol. Genet. 14, 59–69.

Briolay, J., Galtier, N., Brito, M.R., Bouvet, Y., 1998. Molecular phylogeny of Cyprinidae inferred from cytochrome *b* DNA sequences. Mol. Phylogenet. Evol. 9, 100–108.

Brito, R.M., Briolay, J., Galtier, N., Bouvet, Y., Coelho, M.M., 1997. Phylogenetic relationships within the genus *Leuciscus* (Pisces, Cyprinidae) in Portuguese Fresh waters, based on mitochondrial DNA cytochrome *b* sequences. Mol. Phylogenet. Evol. 8, 435–442.

Carmona, J.A., Sanjur, O.I., Doadrio, I., Machordom, A., Vrijenhoek, R.C., 1997. Hybridogenetic reproduction and maternal ancestry of polyploid Iberian fish: the *Tropidophoxinellus alburnoides* complex. Genetics 146, 983–993.

Collares-Pereira, M.J., 1983. Estudo sistemático e citogenético dos pequenos ciprinídeos ibéricos pertencentes aos géneros *Chondrostoma* Agassiz 1835, *Rutilus* Rafinesque, 1820 e *Anaecypris* Collares-Pereira, 1893. PhD Thesis, Universidade de Lisboa.

Cunha, C., Coelho, M.M., Carmona, J.A., Doadrio, I., 2004. Phylogeographical insights into the origins of the *Squalius alburnoides* complex via multiple hybridization events. Mol. Ecol. 13, 2807–2817.

Doadrio, I., Carmona, J.A., 2004. Phylogenetic relationships and biogeography of the genus *Chondrostoma* inferred from mitochondrial DNA sequences. Mol. Phylogenet. Evol. 33, 802–815.

Durand, J.D., Bianco, P.G., Laroche, J., Gilles, A., 2003. Insight into the origin of endemic Mediterranean ichthyofauna—phylogeography of *Chondrostoma* genus (Teleostean, Cyprinidae). J. Hered. 94, 315–328.

Gilles, A., Lecointre, G., Faure, E., Chappaz, R., Brun, G., 1998. Mitochondrial phylogeny of the European cyprinids: implications for their systematics, reticulate evolution, and colonization time. Mol. Phylogenet. Evol. 10, 132–143.

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 21, 160–174.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17 (8), 754–755.

Lanave, C., Preparate, G., Saccone, C., Serio, G., 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. 20, 86–93.

Pala, I., Coelho, M.M., 2005. Contrasting views over a hybrid complex: between speciation and evolutionary "dead-end". Gene 347, 283–294.

Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14 (9), 817–818.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19 (12), 1572–1574.

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

Schmidt, T.R., Gold, J.R., 1993. Complete sequence of the mitochondrial cytochrome *b* gene in the Cherryfin Shinner, *Liturus roseipinnis* (Teleostei: Cyprinidae). Copeia 3, 880–883.

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

Sousa-Santos, C., Robalo, J.I., Collaers-Pereira, M.J., Almada, V.C. in press. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic constitution of hybrid organisms. DNA Seq.

Swofford, D.L., 1998. PAUP *—phylogenetic analysis using parsimony (* and other methods) version 4.0. Sinauer Associates, Sunderland, MA.

Ünver, B., Erk'Akan, F., 2005. A natural hybrid of *Leuciscus cephalus* (L.) and *Chalcalburnus chalcoides* (Güldenstädt) (Osteichthyes-Cyprinidae) from Lake Tödürge (Sivas, Turkey). J. Fish Biol. 66, 899–910.

Vrijenhoek, R.C., 1998. Clonal organisms and the benefits of sex. In: Carvalho, G.R. (Ed.), Advances in Molecular Ecology. IOS Press, Amsterdam, pp. 151–172.

Yang, Z., 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. J. Mol. Evol 39, 306–314.

Young, N.D., Healy, J., 2003. GapCoder automates the use of indel characters in phylogenetic analysis. BMC Bioinformatics 4, 6.

Zardoya, R., Doadrio, I., 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. Proc. R. Soc. Lond. B 265, 1365–1372.

Zardoya, R., Doadrio, I., 1999. Molecular evidence on the evolutionary and biogeographical patterns of European cyprinids. J. Mol. Evol. 49, 227–237.