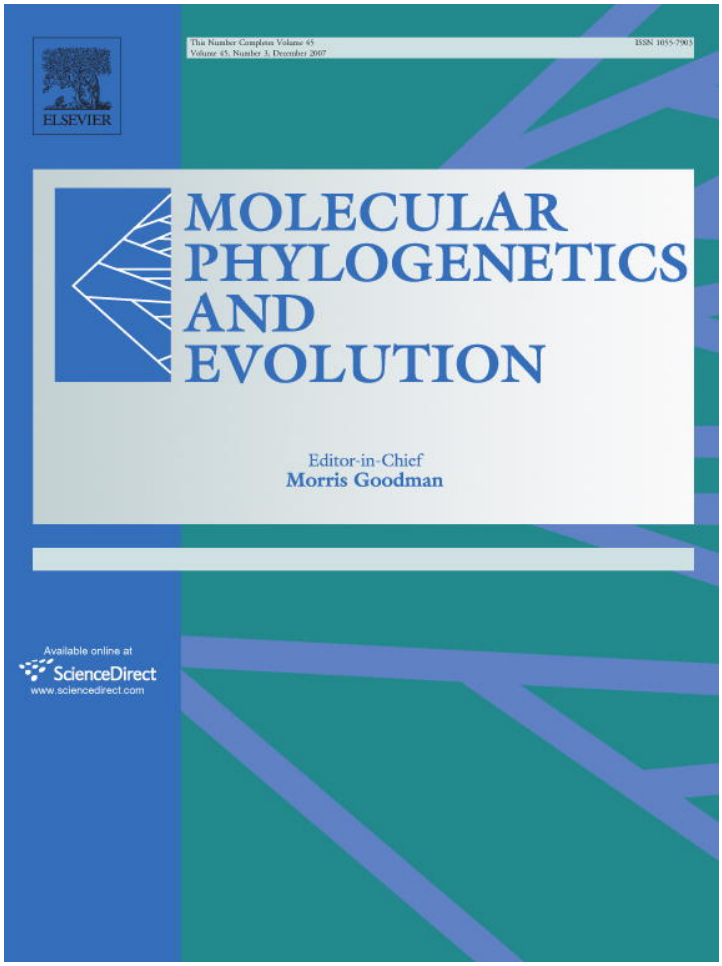

Available online at www.sciencedirect.com
ScienceDirect
Molecular Phylogenetics and Evolution 45 (2007) 981–996
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Reading the history of a hybrid fish complex from its molecular record

C. Sousa-Santos [a,*], M.J. Collares-Pereira [b], V. Almada [a]

[a] Instituto Superior de Psicologia Aplicada, Unidade de Investigação em Eco-Etologia, Rua Jardim do Tabaco 34, 1149-041 Lisboa, Portugal
[b] Universidade de Lisboa, Faculdade de Ciências, Centro de Biologia Ambiental, Campo Grande, 1749-016 Lisboa, Portugal

Received 15 March 2007; revised 7 May 2007; accepted 8 May 2007
Available online 29 May 2007

Abstract

Squalius alburnoides is a widely distributed intergeneric hybrid complex with fish of both sexes, varying ploidy levels and proportions of the parental genomes. Its dispersal routes were here delineated and framed by the reconstruction of the phylogeny and phylogeography of other *Squalius* with which it hybridizes, based on the available data on the paleohydrographical history of the Iberian Peninsula. Results based on sequences of cytochrome *b* and beta-actin genes showed that: proto-*Squalius pyrenaicus* originated at least five species as it dispersed throughout the Iberian Peninsula in the Mio-Pliocene; the *S. alburnoides* complex likely had a single origin in the bulk of Iberia, in the Upper Tagus/Guadiana area, when hydrographical rearrangements allowed the contact between its ancestors (around 700,000 years ago); interspecific crosses allowed the introgression of mitochondrial and nuclear genes of *S. alburnoides* in allopatric species/populations of other *Squalius* and vice-versa; and reconstituted *S. alburnoides* non-hybrid males may contribute to the replacement of the typical mtDNA of the complex (in the populations where they occur, crosses with females of other *Squalius* seem to have been especially frequent). A number of dispersal events and colonization routes are proposed.

Keywords: *Squalius alburnoides*; Hybrid complex; Introgression; Paleobiogeography; Iberia

1. Introduction

Squalius alburnoides is a curious example of a very successful intergeneric hybrid complex of cyprinid fishes endemic to the Iberian Peninsula. Allozymes (Alves et al., 1997a; Carmona et al., 1997), microsatellites (Crespo-López et al., 2006), beta-actin gene sequencing (Robalo et al., 2006) and cytogenetic studies (Gromicho et al., 2006) revealed that the crosses that originated *S. alburnoides* involved *S. pyrenaicus* females (P-haplotype) and a presumably extinct *Anaecypris*-like paternal ancestor (A-haplotype). As a result of this hybridization process, the bisexual reproductive mode was disturbed, originating new patterns of gamete production which generated a diverse array of ploidy levels and genomic constitutions (see Table S1—electronic supplementary material).

As regards primary freshwater fish, the Iberian Peninsula is almost an island which had very limited historical connections with the rest of Europe (Banarescu, 1973; Almaça, 1978). The first known Iberian cyprinid fossil, *Rutilus antiquus*, was detected in the eastern margin of the Peninsula, in the region of the Ebro river basin, and was dated from the Upper Oligocene (Cabrera and Gaudant, 1985; De la Peña, 1995). At this time, the uplift of the Pyrenean Mountains was still an active process that culminated only in the Late Pliocene (Andeweg, 2002), thus, the colonization of the Peninsula in the Oligocene by cyprinids (including the ancestors of *S. alburnoides*) likely occurred through a freshwater passage from south France to the Ebro river basin. For an unlikely alternative (Doadrio and Carmona, 2003) involving wide migrations across the Mediterranean during the Messinian salinity crisis see Bianco (1990). Also, although Iberia and North Africa were in contact for a short period in the Upper Miocene (Andeweg, 2002), the extreme scarcity of Leuciscinae in Africa (Froese

* Corresponding author. Fax: +351 21 8860954.
E-mail address: csousasantos@gmail.com (C. Sousa-Santos).

1055-7903/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2007.05.011

and Pauly, 2007) also make this African alternative very unlikely. For studies on the phylogeny of Iberian cyprinids see Zardoya and Doadrio (1998), Zardoya and Doadrio (1999) and Cunha et al. (2002).

Although the identity of the ancestors of the complex appears to be a clarified issue, the number of original hybridization events has been a matter of controversy. Indeed, the analysis of the mitochondrial DNA (mtDNA) diversity of *S. alburnoides* populations led some authors to postulate multiple independent origins for the complex: Alves et al. (1997b) postulated two independent origins in the Tagus–Guadiana and in the Sado basins and Cunha et al. (2004) postulated three additional ones: in Guadiana–Guadalquivir basins, in Douro and in the Quarteira river (see Fig. 1 for river locations).

Along its wide distribution range, *S. alburnoides* is currently sympatric with at least three other *Squalius* species, *S. pyrenaicus*, *S. carolitertii* and *S. aradensis* (Fig. 1), whose mitochondrial and nuclear genes are found in the complex (Alves et al., 1997b; Cunha et al., 2004; Sousa-Santos et al., 2006a). To discuss whether the origin of the *S. alburnoides* complex was a unique event or if it occurred independently in more than one river basin, it is crucial to evaluate the existence of past and present interspecific gene flow between *S. alburnoides* and the other three sympatric *Squalius* species. Moreover, the interpretation of the phylogeographic patterns of *S. alburnoides* has to be framed by the patterns exhibited by the other *Squalius* species with which the complex hybridizes and corroborated by the paleohydrographical history of Iberia.

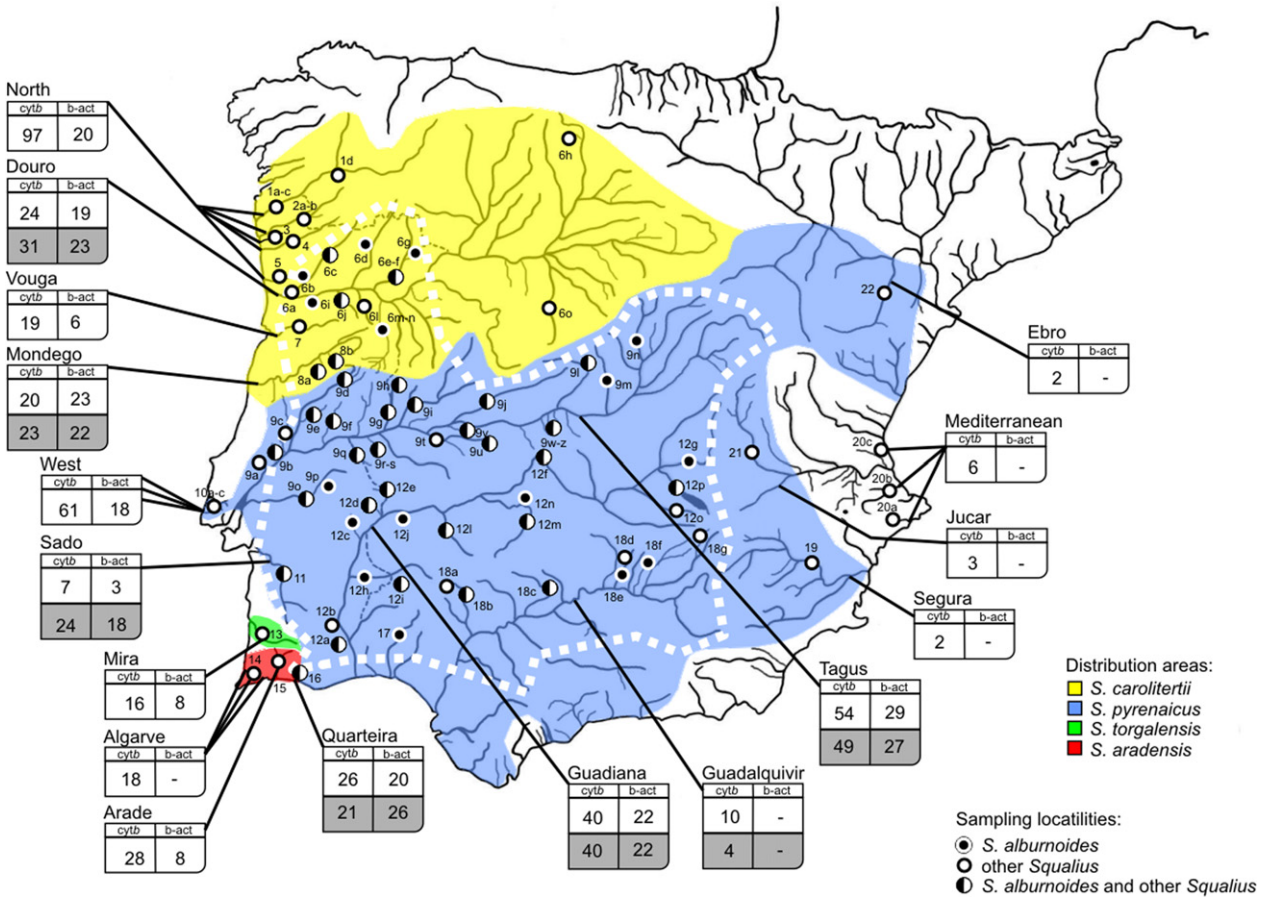

Fig. 1. Distribution areas of the *Squalius* species studied, sampling locations and number of individuals, from each river basin, sequenced for cyt*b* and beta-actin genes. The distribution area of *S. alburnoides* is represented by a white broken line, overlapping the distribution areas of *S. carolitertii* (in yellow), *S. pyrenaicus* (in blue) and *S. aradensis* (in red). Sampling locations are represented by different signs (legended in the figure) according to the species captured in each location. Each river basin has a box associated in which the number of *S. alburnoides* (grey half of the box) and the number of other *Squalius* (white half of the box) sequenced for cyt*b* and beta-actin genes are summarized. Legend: 1—Minho (1a—Coura, 1b—Mouro, 1c—Tea, 1d—Vivey); 2—Lima (2a—Salas, 2b—Vez); 3—Neiva; 4—Cávado; 5—Ave; 6—Douro (6a—Minas, 6b—Sousa, 6c—Tâmega; 6d-Calvo, 6e—Sabor, 6f—Azibo, 6g—Maçãs, 6h—Boedo, 6i—Arda, 6j—Paiva, 6l—Távora, 6m—Coa, 6n—Águeda, 6o—Adaja); 7—Vouga; 8—Mondego (8a—Alva, 8b—Ceira); 9—Tagus [9a—Maior, 9b—Alviela, 9c—Nabão, 9d—Zêzere, 9e—Sertã, 9f—Ocreza, 9g—Erges, 9h—Trevejana, 9i—Alagon tributaries (Acebo, Arrago, Jerte, Caparro), 9j—Tiétar, 9l—Cofio, 9m—Guadarrama, 9n—Jarama, 9o—Sorraia, 9p—Seda, 9q—Sever, 9r—Alburrel, 9s—Vid, 9t—Pesquero, 9u—Almonte, 9v—Aurela, 9w—Huso, 9x—Gebalo, 9z—Cedena]; 10—Western rivers (10a—Lizandro, 10b—Samarra, 10c—Colares); 11—Sado; 12—Guadiana (12a—Vascão, 12b—Oeiras, 12c—Degebe, 12d—Caia, 12e—Xévora, 12f—Estena, 12g—Zancara, 12h—Ardila, 12i—Sillo, 12j—Albuera, 12l—Matachel, 12m—Zujar, 12n—Quejigares, 12o—Azuer, 12p—Ruidera); 13—Mira; 14—Algarve rivers (Aljezur, Seixe and Alvor); 15—Arade; 16—Quarteira; 17—Odiel; 18—Guadalquivir (18a—Mollinos, 18b—Montemayor, 18c—Manzano, 18d—Robledillo, 18e—Jandula, 18f—Guadiel, 18g—Guadalmena); 19—Segura; 20—Mediterranean rivers (20a—Algar, 20b—Serpis, 20c—Valencia Lagoon); 21—Jucar; 22—Ebro. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

During the Miocene most river systems, instead of flowing to the sea, drained to a large number of inland lakes, some of which persisted well into the Pliocene (Friend and Dabrio, 1996; Andeweg, 2002). The current exorheic river network is very recent, being of Plio–Pleistocene origin (Andeweg, 2002). Thus, hypotheses drawn to explain the dispersal and foundation of new *Squalius* populations have to be concordant with the available geological data concerning the definition of the respective fluvial networks.

The aims of this study were: (1) to draw plausible dispersal routes for the Iberian *Squalius*; (2) to discuss possible locations for the origin of the *S. alburnoides* complex; (3) to postulate the putative dispersal routes of the *S. alburnoides* complex that are compatible with geological data and with the distribution of other *Squalius* species; and (4) to evaluate the relationships between *S. alburnoides* and other *Squalius* species and their implications to the genetic dynamics of the complex.

The opportunity to reconstruct the history of this complex was made possible because although recombination has been demonstrated to occur among similar genetic complements (e.g., among two A or two P chromosome sets—Alves et al., 2004; Crespo-López et al., 2006), no evidence of recombination among dissimilar genetic complements (A and P) was yet found. This absence of recombination between the A and P genomes means that by the combined use of a mitochondrial and a nuclear marker it is possible to infer the parentage of each form of the complex. In this paper, fragments of the cytochrome *b* and the beta-actin genes were used to achieve this goal.

As it was possible to obtain samples covering almost all the known distribution area of *S. alburnoides*, this study allowed a deep insight into the evolutionary dynamics of a intergeneric hybrid complex. Indeed, the reconstruction of paleogeographic scenarios may be helpful for the evaluation of its evolutionary potential and capacity to adapt to distinct environments and interact with sympatric species.

2. Materials and methods

2.1. Brief presentation of the S. alburnoides complex

The *S. alburnoides* complex is known to occur presently in nine Iberian river basins (Douro, Vouga, Mondego, Tagus, Sado, Guadiana, Quarteira, Odiel and Guadalquivir) (Cabral et al., 2005; Ribeiro et al., 2007) (Fig. 1). This complex comprises distinct forms whose frequency may differ significantly according to the river basin: the southern populations include diploids (PA), triploids (PAA and PPA), tetraploids (PPAA, PAAA and PPPA), and also a non-hybrid form (AA), reconstituted from the hybrids and almost constituted by males (females appear to be extremely rare—Sousa-Santos et al., 2006c), that is absent from the northern populations where all the other hybrid forms may be found (reviewed in Alves et al., 2001). The designation of the nuclear genomes of the *S. alburnoides* hybrid forms includes a capital "A" (refering to the paternal ancestor of the complex), while the letter "P" used above to denote the genome of the still existing maternal ancestor *S. pyrenaicus*, must be replaced by "C" or "Q" when the sympatric *Squalius* are, respectively, *S. carolitertii* or *S. aradensis*. A synthesis on the distribution and frequency of the *S. alburnoides* forms, including the northern ones, may be found in Table S2 (electronic supplementary material).

The mtDNA found in *S. alburnoides* is usually *S. pyrenaicus*-like (the maternal ancestor of the complex) but some introgressions were already reported: one individual with *S. carolitertii*-like mtDNA in the Douro drainage (Alves et al., 1997b), and several with *S. aradensis*-like mtDNA in the Quarteira River (Sousa-Santos et al., 2006a).

2.2. Field work and laboratorial procedures

Sequences of specimens from the *S. alburnoides* complex and from five other *Squalius* species (*S.aradensis, S. carolitertii, S. pyrenaicus, S. torgalensis* and *S. valentinus*), covering a total of 27 river basins, were analysed for the cytochrome *b* (cyt*b*) and beta-actin genes. Sampling locations and the respective number of individuals sequenced, in a total of 149 *S. alburnoides*, 143 *S. pyrenaicus*, 164 *S. carolitertii*, 42 *S. aradensis* and 18 *S. torgalensis* are depicted in Fig. 1. In order to get the most complete coverage of the populations, some additional sequences of the cyt*b* gene were retrieved from GenBank which, added to the sequences that resulted from this study, amounted to a total of 217 *S. alburnoides*, 214 *S. pyrenaicus*, 175 *S. carolitertii*, 75 *S. aradensis*, 18 *S. torgalensis* and 6 *S. valentinus*. GenBank Accession numbers may be found in Table S3 (electronic supplementary material). Samples of *S. alburnoides* from River Vouga should have been included, however, despite some attempts in the main river and in the River Sul, no *S. alburnoides* specimens were captured.

Fish were electrofished, morphologically identified, and in general returned to the river after the removal of small fin clips. Total genomic DNA was extracted from fin clips preserved in ethanol by an SDS/proteinase-k based protocol, precipitated with isopropanol and washed with ethanol before re-suspension in water (adapted from Sambrook et al., 1989). A total of 799 bp of the cyt*b* gene and of 935 bp of the beta-actin gene were amplified. The primers and PCR conditions may be found in Sousa-Santos et al. (2005).

2.3. Data analysis

Sequences of the cyt*b* gene were aligned with BioEdit v.5.0.6 and their phylogenetic relationships reconstructed with a maximum parsimony method using PAUP 4.0 (Swofford, 1998). The resultant phylogenetic tree was analysed to assess the existence of present and ancient crosses between *S. alburnoides* and other sympatric *Squalius* species in a given river basin. It was assumed that if *S. alburno-*

ides crossed only with conspecifics and/or with males of other *Squalius* species no haplotypes will be shared between *S. alburnoides* and other *Squalius* species. In contrast, the existence of derived haplotypes shared between *S. alburnoides* and other *Squalius* species would reflect the existence of interspecific crosses. Thus, when analysing the phylogenetic tree, (a) terminal haplotypes shared between *S. alburnoides* and other *Squalius* species were assumed to be representatives of recent crosses; and (b) missing common ancestors between *S. alburnoides* and other *Squalius* species located in the internal nodes of the tree were interpreted as indicative of ancient crosses. The mean number of mutational steps linking each pair of haplotypes to their common ancestor was used to construct a histogram reflecting the contacts through time between *S. alburnoides* and other *Squalius* species.

Concerning the beta-actin gene, the sequences of homozygous individuals were also aligned with BioEdit v.5.0.6. However, for heterozygous diploid and polyploid individuals the genome complements had to be recovered following the procedures described in Sousa-Santos et al. (2005) before the alignment process. Since the low differentiation of the beta-actin *Squalius* haplotypes does not allow the distinction of all the species validated by the mtDNA analysis, the nuclear genomes of *S. pyrenaicus, S. carolitertii*, *S. torgalensis* and *S. aradensis* were herein designated as P-haplotypes for simplicity reasons. The nuclear haplotypes derived from the paternal ancestor of the *S. alburnoides* complex were designated as A-haplotypes. In the presence of distinct P- or A-haplotypes in the same individual the designation P′ or A′ was used.

In previous studies with this marker in cyprinids, involving more than ten species and some hundreds of individuals (Robalo et al., 2006, 2007; Sousa-Santos et al., 2006a,c, 2007), only a single locus was detected in diploids for the beta-actin gene, assuring that it is a single copy gene, thus excluding the risk of using paralogous sequences when analysing polyploids of hybrid origin.

Networks of mitochondrial and nuclear haplotypes were performed with Network 4.201 (www.fluxus-engineering.com), using a median-joining algorithm (Bandelt et al., 1999). Mean number of pairwise differences, diversity indices, AMOVA and pairwise comparisons of haplotype frequencies among populations were calculated with Arlequin 3.01 (Excoffier et al., 2005). When computing mean divergence among populations, the values were corrected to remove within population variation. To perform this correction, the mean divergence between all possible pairs of sequences having a member in each population was first computed and then the average divergence of all pairs of sequences involving members of the same population was subtracted, as implemented in Arlequin.

Estimations of the divergence time based on the cyt*b* gene were calculated using an evolutionary rate of 1.05% sequence divergence per million years (MY), as suggested by Dowling et al. (2002) for North American cyprinids.

3. Results

3.1. MtDNA variation

From a total of 625 cyt*b* gene sequences, 211 distinct haplotypes were identified: 39.81% belonging to *S. alburnoides*, 53.08% belonging to other *Squalius* species and 7.11% shared haplotypes between *S. alburnoides* and other *Squalius* species. A summary of the genetic structure of each population is summarized in Table 1.

The haplotype network showed a clear distinction of the populations belonging to the five *Squalius* species studied (Fig. 2). The *S. pyrenaicus* sub-network was the most diverse, with 155 different haplotypes that ranged in their level of divergence from one to 22 mutations. The haplotypes from the River Tagus occupied a central position within this sub-network, from where other *S. pyrenaicus* populations from Guadiana, Guadalquivir and Sado river basins, and *S. valentinus* branched. Within the *S. carolitertii* sub-network, three groups of haplotypes were clearly identified, differing from one to 31 mutations: haplotypes that were found only in River Zêzere; haplotypes that were found only in River Mondego; and a third group of haplotypes belonging to northern rivers from Minho to Mondego. *S. aradensis* and *S. torgalensis* haplotypes also formed well defined sub-networks. Globally, eight phylogroups of *Squalius* were identified: *S. pyrenaicus*-Tagus/Guadiana, *S. pyrenaicus*-Sado, *S. valentinus*, *S. carolitertii*-North, *S. carolitertii*-Mondego, *S. carolitertii*-Zêzere, *S. aradensis* and *S. torgalensis*. The mean percentage of divergence between the described phylogroups is presented in Fig. 2.

In general, the *Squalius* individuals that were not *S. alburnoides* generally presented the expected mtDNA considering their geographical provenience, with only few exceptions: four *S. pyrenaicus* from Guadiana with Tagus-like mtDNA; five *S. pyrenaicus* in Mondego (where the expected species should be *S. carolitertii*); and one *S. pyrenaicus* from Lizandro with Guadiana-like mtDNA.

To allow the analysis of the gene flow between distinct river basins, only the populations of *S. alburnoides* bearing *S. pyrenaicus*-mtDNA were considered. Other *Squalius* harbouring mtDNA of other sympatric *Squalius* were excluded because they represent introgressions that would introduce artefacts in the estimation of divergence times. The calculated divergence values between pairs of populations ranged between 0 and 1.15% (Table 2). Also shown in Table 2 are the divergence times between populations of other *Squalius* species analysed in this study.

3.2. Nuclear DNA variation

The genomic constitution of 138 *S. alburnoides* and of 176 individuals belonging to the other *Squalius* species is presented in Table S3 (electronic supplementary material). Concerning the later, although the majority of the individuals were homozygous for the beta-actin gene, considerably

Table 1
Sample size (*N*), number of haplotypes (*N* hap), haplotype diversity (HD), gene diversity (GD), nucleotide diversity (ND) and mean number of pairwise differences (MNPD) for *S. alburnoides* and other *Squalius* species from distinct drainages sequenced for the cyt*b* gene

River basins	*N*	*N* hap	HD (%)	GD ± SD	MNPD ± SD	ND ± SD
S. alburnoides						
Douro	31	10	32.26	0.626 ± 0.100	11.755 ± 5.470	0.015 ± 0.008
Mondego	24	9	37.50	0.775 ± 0.079	16.670 ± 7.690	0.021 ± 0.011
Tagus	49	40	81.63	0.990 ± 0.007	7.051 ± 3.368	0.009 ± 0.005
Sado	24	10	41.67	0.620 ± 0.117	3.967 ± 2.057	0.005 ± 0.003
Guadiana	40	27	67.50	0.951 ± 0.022	5.362 ± 2.641	0.007 ± 0.004
Guadalquivir	4	4	100.00	1.000 ± 0.177	7.667 ± 4.533	0.010 ± 0.007
Quarteira	21	6	28.57	0.552 ± 0.122	14.219 ± 6.641	0.018 ± 0.009
Other Squalius						
Minho	22	1	4.55	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Lima	22	4	18.18	0.398 ± 0.122	0.429 ± 0.402	0.001 ± 0.001
Neiva	13	1	7.69	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Cávado	20	2	10.00	0.521 ± 0.042	0.521 ± 0.456	0.001 ± 0.001
Ave	20	2	10.00	0.100 ± 0.088	0.100 ± 0.178	0.000 ± 0.000
Douro	24	7	29.17	0.779 ± 0.057	1.667 ± 1.014	0.002 ± 0.001
Vouga	19	2	10.53	0.105 ± 0.092	0.105 ± 0.183	0.000 ± 0.000
Mondego	20	8	40.00	0.847 ± 0.051	17.584 ± 8.156	0.022 ± 0.011
Lizandro	21	8	38.10	0.791 ± 0.076	2.238 ± 1.284	0.003 ± 0.002
Samarra	22	2	9.09	0.091 ± 0.081	0.182 ± 0.245	0.000 ± 0.000
Colares	18	4	22.22	0.529 ± 0.117	0.699 ± 0.553	0.001 ± 0.001
Tagus	54	36	66.67	0.980 ± 0.008	15.881 ± 7.195	0.020 ± 0.010
Guadiana	40	24	60.00	0.942 ± 0.027	4.165 ± 2.115	0.005 ± 0.003
Mira	16	3	18.75	0.242 ± 0.135	0.250 ± 0.297	0.000 ± 0.000
Arade	28	9	32.14	0.833 ± 0.050	2.611 ± 1.439	0.003 ± 0.002
Quarteira	26	3	11.54	0.151 ± 0.093	0.2310 ± 7.779	0.000 ± 0.011
Sado	7	4	57.14	0.810 ± 0.130	1.238 ± 0.885	0.002 ± 0.001
Guadalquivir	10	9	90.00	0.978 ± 0.054	7.089 ± 3.636	0.009 ± 0.005
Alvor	6	1	16.67	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Aljezur	6	2	33.33	0.533 ± 0.172	0.533 ± 0.508	0.001 ± 0.001
Seixe	6	2	33.33	0.333 ± 0.215	0.333 ± 0.380	0.000 ± 0.001
Segura	2	2	100.00	1.000 ± 0.500	2.000 ± 1.732	0.003 ± 0.003
Algar	2	2	100.00	1.000 ± 0.500	2.000 ± 1.732	0.003 ± 0.003
Serpis	2	1	50.00	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000
Valencia Lagoon	2	2	100.00	1.000 ± 0.500	1.000 ± 1.000	0.001 ± 0.002
Júcar	3	2	66.67	0.667 ± 0.314	0.667 ± 0.667	0.001 ± 0.001
Ebro	2	1	50.00	0.000 ± 0.000	0.000 ± 0.000	0.000 ± 0.000

high numbers of heterozygous individuals (PP′) were found (40.34%).

From all the *S. alburnoides* individuals sampled, 23.19% were nuclear non-hybrids (AA or AA′), proceeding from the Tagus, Guadiana and Quarteira rivers. The remaining 76.81% were diploid or polyploid hybrids (13.04% PA, 50.00% PAA, 11.59% PPA, 1.45% PPAA and 0.72 PAAA)—Table S2 (electronic supplementary material).

From a total of 314 specimens of all the *Squalius* sampled, 37 P-haplotypes and 14 A-haplotypes were identified. The network of the A-haplotypes (Fig. 3) showed that the ancestral haplotype (A4) was dispersed throughout the sampling area with the exception of the Douro river basin. In this drainage only one haplotype (A9) was found, shared with the Tagus and Mondego Rivers. Mondego and Tagus specimens also shared three other haplotypes (A1, A15 and A5) and showed, respectively, one and three unique haplotypes. In the south, the Guadiana population showed one exclusive haplotype (A2) and four haplotypes shared with other river basins: Sado (A4 and A11), Quarteira (A4 and A6), Tagus (A4 and A3) and Mondego (A6). The Quarteira population, besides the two haplotypes shared with Guadiana, also presented two exclusive haplotypes (A13 and A14).

Concerning the P-haplotypes (Fig. 4), the ancestral haplotype (P3) was found in all the northern rivers from Minho to Mondego, in the Tagus and in the Guadiana. These last three river basins were the most diverse, with nine different haplotypes being found in Mondego and Guadiana, and ten in the Tagus. The network depicted in Fig. 4 showed that starting from the ancestral haplotype, three distinct southern lineages seem to have differentiated: one that includes haplotypes found in the southwestern rivers of Mira and Arade (P9, P10, P14, P15, P26 and P27); and two lineages linked by a missing common ancestor, one from the Sado (P8, P18, P36 and P37) and the other from the Guadiana (P7, P12, P13, P28, P29 and P30). In the Sado, in addition to the above mentioned four exclusive haplotypes, one haplotype from the Arade lineage was detected in one individual (P14). The Quarteira population also presented a mixture of haplotypes: from the Arade (P10, P14 and P19) and from the Guadiana (P7 and P32) lineages. In the

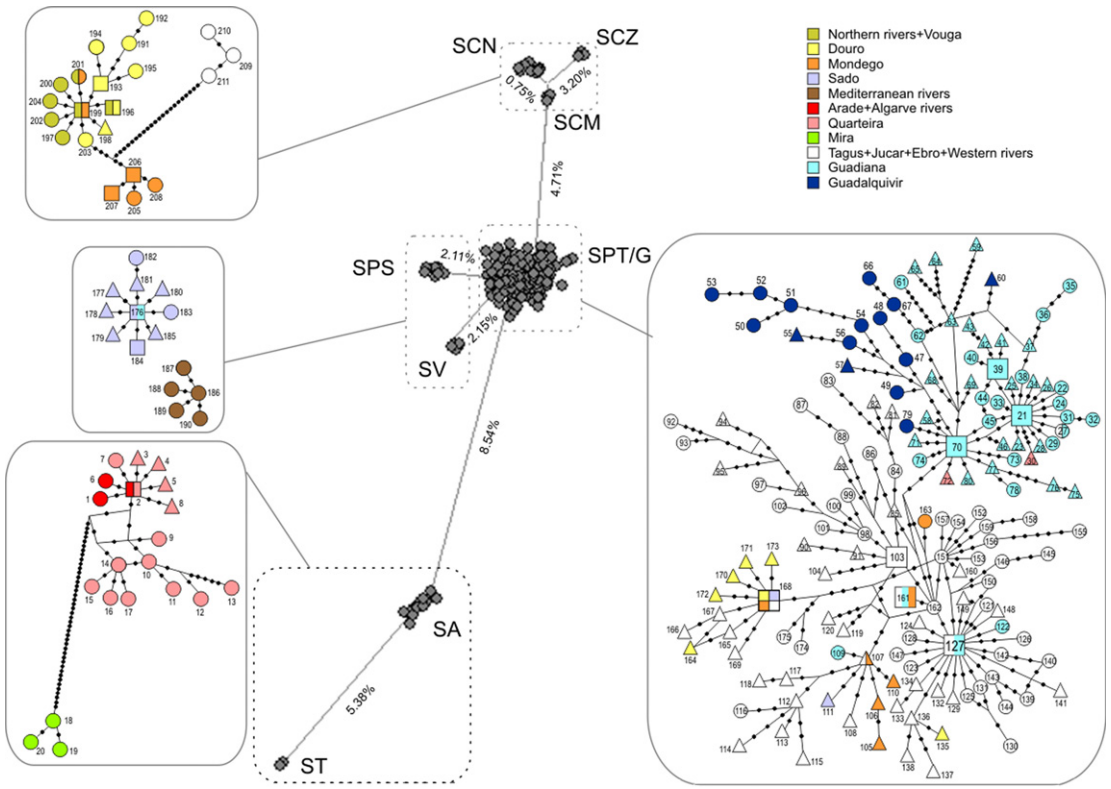

Fig. 2. Network of mtDNA haplotypes. The global network representing eight phylogroups (SCN—*S.carolitertii*-North, SCZ—*S.carolitertii*-Zêzere, SCM—*S.carolitertii*-Mondego, SPS—*S. pyrenaicus*-Sado, SPT/G-*S.pyrenaicus*-Tagus/Guadiana, SV—*S.valentinus*, SA—*S.aradensis* and ST—*S.torgalensis*), with the respective mean percentage of pairwise differences to the closest phylogroups, is depicted in a central position. On both sides of this central diagram the network of haplotypes of each phylogroup is depicted, with information concerning the species from where the haplotypes were retrieved and their respective river basin: haplotypes that were exclusive of *S. alburnoides* are represented by triangles; haplotypes that were exclusive of other *Squalius* are represented by circles; and haplotypes shared between *S. alburnoides* and other *Squalius* species are represented by squares. Each haplotype is identified with a numerical code (the same that was used in Figure S1 and Table S3—both in electronic supplementary material) and its geographic provenience is represented by different colour patterns (see legend in the Figure). The number of mutations between haplotypes is represented by the number of small black dots placed on the branches linking haplotypes.

north, the majority of the populations exhibited the ancient haplotypes P3 and P6 (only found in this region). The population of Vouga also presented haplotype P5 that was shared with Mondego and with the Erges tributary of Tagus. In the Mondego, four out of nine haplotypes were only found in this population (P20, P21, P22 and P34), the remaining being shared with the neighbour drainages of Vouga (P3 and P5) and Tagus (P1, P2, P3, P5 and P11). The Tagus populations presented four exclusive haplotypes (P17, P31, P33 and P35) and haplotypes shared with northern rivers (P3 and P5), with Mondego (P1, P2, P3, P5 and P11), with the western rivers (P1 and P2) and with Guadiana (P2, P3 and P11). The haplotypes found in the western rivers of Lizandro, Samarra and Colares were either shared with the Tagus populations (P1 and P2) or derived by one mutational step from the ancestral haplotype P3 also present in the Tagus (P24 and P25).

3.3. Interspecific gene flow

The nuclear P-haplotypes found in *S. alburnoides* were the ones found in the sympatric *Squalius* of the considered river basin, with a few exceptions (Fig. 4). These exceptions may represent new mutations, *Squalius* genomes brought by *S. alburnoides* that dispersed from other rivers, or simply insufficient sampling causing a failure to detect less common haplotypes in all *Squalius* of a given river basin.

In the case of Tagus, Guadiana and Guadalquivir drainages, the mtDNA haplotypes presented by *S. alburnoides* merged into the *S. pyrenaicus* sub-network. Some *S. alburnoides* haplotypes from Sado, Mondego, Douro and Quarteira river basins were included into the sub-network of the *Squalius* species that is sympatric with the complex in the same river basin (Fig. 2). The extent of the introgressions of mtDNA of different phylogroups in the *S. alburnoides* complex is presented in Table 3.

The parsimony phylogenetic tree exhibited a topology that was similar to trees already published by other authors (Sanjur et al., 2003; Cunha et al., 2004; Doadrio and Carmona, 2006) and, due to its extension it was herein presented as electronic supplementary material (Figure S1). However, in contrast to previous papers, the inclusion in this study of *S. alburnoides* from almost all its distribution area made possible the identification of crosses with mem-

Table 2
Mean percentage of divergence (below diagonal) between the *S. alburnoides* and other *Squalius* populations, intrapopulation divergence percentage values (in the diagonal) and estimated divergence times between populations in MY (above diagonal)

S. alburnoides	Tagus ALAG	Tagus ERM	Tagus ELM	Tagus WRM	Tagus WLM	Guadiana WRM	Guadiana WLM	Guadiana ERM	Guadiana ELM	Guadalquivir	Quarteira	Douro	Mondego	Sado
Tagus ALAG	0,28	0,77	0,77	0,72	0,76	0,85	0,89	0,80	0,99	1,09	0,85	0,01	0,69	0,01
Tagus ERM	0,80	0,60	< 0	0,16	0,21	0,78	0,81	0,75	0,93	1,01	0,74	0,70	0,30	0,17
Tagus ELM	0,81	-0,01	0,81	0,07	0,15	0,69	0,73	0,68	0,83	0,92	0,64	0,71	0,22	0,13
Tagus WRM	0,75	0,17	0,07	0,81	0,09	0,74	0,78	0,75	0,94	1,01	0,70	0,66	0,05	< 0
Tagus WLM	0,80	0,22	0,16	0,09	0,62	0,71	0,75	0,69	0,88	0,97	0,67	0,70	0,23	0,11
Guadiana WRM	0,89	0,82	0,73	0,78	0,75	0,43	0,00	0,02	0,23	0,28	< 0	0,81	0,84	0,50
Guadiana WLM	0,94	0,86	0,76	0,82	0,78	0,00	0,43	< 0	0,26	0,30	< 0	0,86	0,88	0,55
Guadiana ERM	0,84	0,79	0,71	0,78	0,73	0,02	-0,01	0,92	0,10	0,18	< 0	0,78	0,83	0,49
Guadiana ELM	1,04	0,98	0,87	0,99	0,93	0,24	0,27	0,10	1,75	0,05	0,24	1,03	1,01	0,72
Guadalquivir	1,15	1,06	0,96	1,06	1,01	0,29	0,32	0,19	0,05	0,96	0,29	1,11	1,11	0,79
Quarteira	0,89	0,78	0,67	0,74	0,70	-0,06	-0,06	-0,02	0,25	0,30	0,75	0,80	0,79	0,48
Douro	0,01	0,73	0,74	0,69	0,73	0,85	0,90	0,82	1,09	1,17	0,84	0,18	0,63	< 0
Mondego	0,72	0,31	0,23	0,05	0,24	0,88	0,92	0,87	1,06	1,17	0,83	0,66	0,30	< 0
Sado	0,01	0,18	0,13	-0,03	0,12	0,53	0,57	0,51	0,75	0,83	0,50	-0,02	-0,08	1,13

	Tagus ALAG	Tagus ERM	Tagus ELM	Tagus WRM	Tagus WLM	Guadiana WRM	Guadiana WLM	Guadiana ERM	Guadiana ELM	Guadalquivir	Samarra	Lizandro	Colares	Mondego a	Sado	Jucar	Ebro	Segura	Algar	Serpis	Valencia lagoon	Minho	Lima	Neiva	Cávado	Ave	Douro	Vouga	Mondego b	Mondego c	Tagus Zêzere	Quarteira	Arade	Alvor	Aljezur	Seixe	Mira
S. pyrenaicus																																					
Tagus ALAG	0,68	0,43	0,16	0,11	0,15	0,66	1,10	0,81	0,85	1,15	0,35	0,20	0,22	0,20	2,19	0,39	0,39	1,22	2,42	2,53	2,42	5,24	5,23	5,24	5,15	5,23	5,22	5,24	5,28	4,65	5,41	8,42	8,10	8,57	8,58	8,06	10,45
Tagus ERM	0,45	0,67	0,16	0,34	0,18	0,81	1,23	0,95	1,02	1,15	0,52	0,35	0,39	0,33	2,39	0,20	0,20	1,19	2,54	2,66	2,54	5,32	5,32	5,32	5,24	5,32	5,31	5,32	5,44	4,73	5,74	8,35	8,19	8,50	8,51	8,22	10,39
Tagus ELM	0,17	0,17	0,87	0,12	0,03	0,57	0,98	0,66	0,74	0,98	0,33	0,16	0,20	0,13	2,20	0,14	0,14	1,08	2,36	2,48	2,36	5,00	5,00	5,00	4,92	5,00	4,99	5,00	5,12	4,41	5,36	8,16	7,95	8,31	8,31	7,96	10,14
Tagus WRM	0,12	0,36	0,87	0,61	0,03	0,53	0,95	0,66	0,73	0,99	0,19	0,04	0,06	0,06	2,11	0,30	0,30	1,07	2,18	2,30	2,18	5,15	5,14	5,15	5,06	5,15	5,13	5,15	5,27	4,60	5,33	8,30	7,97	8,45	8,46	7,93	10,31
Tagus WLM	0,16	0,19	0,03	0,03	0,63	0,53	0,97	0,68	0,75	0,99	0,20	0,05	0,08	0,02	2,26	0,11	0,11	1,08	2,28	2,40	2,28	5,19	5,18	5,19	5,10	5,19	5,17	5,19	5,30	4,60	5,46	8,26	7,99	8,40	8,41	7,98	10,30
Guadiana WRM	0,70	0,85	0,60	0,55	0,56	0,53	0,49	0,07	0,02	0,42	0,73	0,49	0,62	0,45	2,05	0,77	0,77	0,63	2,21	2,23	2,21	5,30	5,29	5,30	5,21	5,30	5,23	5,30	5,42	4,71	5,29	8,73	8,41	8,90	8,90	8,38	10,39
Guadiana WLM	1,16	1,29	1,03	1,00	1,01	0,51	0,00	0,43	0,46	0,76	1,18	0,95	1,06	0,86	2,51	1,19	1,19	0,83	2,50	2,50	2,50	5,60	5,60	5,60	5,52	5,60	5,59	5,60	5,72	5,01	5,54	8,65	8,33	8,82	8,83	8,30	10,47
Guadiana ERM	0,85	1,00	0,70	0,69	0,72	0,07	0,64	0,38	0,05	0,39	0,89	0,65	0,78	0,56	2,13	0,91	0,91	0,55	2,19	2,19	2,19	5,32	5,31	5,32	5,23	5,32	5,27	5,32	5,44	4,73	5,39	8,76	8,44	8,93	8,94	8,41	10,36
Guadiana ELM	0,90	1,07	0,78	0,77	0,79	0,02	0,70	0,06	0,45	0,41	0,96	0,70	0,85	0,64	2,11	0,98	0,98	0,62	2,24	2,24	2,24	5,32	5,31	5,32	5,23	5,32	5,25	5,32	5,44	4,73	5,30	8,75	8,43	8,92	8,93	8,40	10,38
Guadalquivir	1,20	1,20	1,03	1,04	1,04	0,44	0,80	0,41	0,43	0,89	1,22	0,97	1,11	0,90	2,32	1,23	1,23	0,30	2,33	2,33	2,33	5,38	5,38	5,38	5,30	5,38	5,31	5,38	5,50	4,84	5,15	8,83	8,52	9,01	9,01	8,49	10,44
Samarra	0,72	0,89	0,80	0,51	0,54	1,05	1,25	1,14	1,24	1,73	0,02	0,13	0,12	0,29	2,49	0,48	0,48	1,30	2,49	2,61	2,49	5,60	5,60	5,60	5,52	5,60	5,58	5,60	5,72	5,01	5,77	8,79	8,46	8,94	8,95	8,42	10,83
Lizandro	0,69	0,85	0,75	0,49	0,51	0,92	1,14	1,01	1,10	1,60	0,14	0,28	0,02	0,12	2,22	0,32	0,32	1,07	2,24	2,35	2,24	5,33	5,32	5,33	5,24	5,33	5,30	5,33	5,45	4,74	5,47	8,56	8,24	8,72	8,72	8,20	10,55
Colares	0,62	0,79	0,69	0,42	0,44	0,96	1,16	1,05	1,16	1,65	0,13	0,02	0,09	0,16	2,39	0,35	0,35	1,18	2,38	2,49	2,38	5,47	5,47	5,47	5,39	5,47	5,46	5,47	5,59	4,88	5,65	8,61	8,28	8,76	8,77	8,24	10,70
Mondego a	0,78	0,90	0,80	0,59	0,56	0,97	1,13	1,00	1,12	1,61	0,54	0,49	0,44	0,45	2,18	0,29	0,29	0,98	2,12	2,24	2,12	5,08	5,07	5,08	4,99	5,08	5,06	5,08	5,20	4,49	5,30	8,36	8,04	8,51	8,52	7,99	10,24
Sado	2,30	2,51	2,31	2,22	2,37	2,15	2,64	2,23	2,22	2,96	2,62	2,33	2,51	2,29	0,15	2,51	2,51	2,63	3,11	3,23	3,11	5,94	5,93	5,94	5,96	5,94	5,84	5,94	6,06	5,38	6,62	9,35	9,25	9,53	9,54	9,25	11,17
Jucar	0,41	0,21	0,14	0,32	0,11	0,81	1,25	0,95	1,03	1,30	0,50	0,33	0,37	0,30	2,64	0,08	0,00	1,31	2,50	2,62	2,50	5,13	5,12	5,13	5,04	5,13	5,11	5,13	5,24	4,54	5,54	8,32	8,15	8,46	8,47	8,18	10,36
Ebro	0,41	0,21	0,14	0,32	0,11	0,81	1,25	0,95	1,03	1,30	0,50	0,33	0,37	0,30	2,64	0,00	0,00	1,31	2,50	2,62	2,50	5,13	5,12	5,13	5,04	5,13	5,11	5,13	5,24	4,54	5,54	8,32	8,15	8,46	8,47	8,18	10,36
Segura	1,28	1,25	1,13	1,13	1,14	0,66	0,88	0,58	0,65	0,32	1,37	1,12	1,24	1,03	2,77	1,38	1,38	0,25	2,26	2,38	2,38	5,48	5,48	5,48	5,40	5,48	5,47	5,48	5,60	4,89	5,66	8,76	8,45	8,94	8,95	8,42	10,36
S. valentinus																																					
Algar	2,54	2,67	2,48	2,29	2,39	2,32	2,63	2,29	2,36	2,45	2,62	2,35	2,49	2,23	3,27	2,63	2,63	2,38	0,25	0,12	0,00	5,72	5,72	5,72	5,74	5,72	5,71	5,72	5,84	5,13	6,37	9,60	9,29	9,77	9,78	9,26	11,31
Serpis	2,66	2,80	2,60	2,41	2,52	2,34	2,63	2,29	2,36	2,45	2,74	2,47	2,62	2,35	3,39	2,75	2,75	2,50	0,13	0,00	0,12	5,84	5,84	5,84	5,86	5,84	5,83	5,84	5,96	5,25	6,49	9,72	9,41	9,89	9,90	9,38	11,31
Valencia lagoon	2,54	2,67	2,48	2,29	2,39	2,32	2,63	2,29	2,36	2,45	2,62	2,35	2,49	2,23	3,27	2,63	2,63	2,50	0,00	0,13	0,13	5,72	5,71	5,72	5,74	5,72	5,71	5,72	5,84	5,13	6,37	9,49	9,17	9,65	9,66	9,14	11,19
S. carolitertii																																					
Minho	5,84	5,92	5,69	5,71	5,76	5,83	5,88	5,78	5,81	6,10	5,89	5,73	5,79	5,56	6,31	5,38	5,38	5,76	6,01	6,13	6,01	0,00	0,00	0,00	0,02	0,00	0,03	0,00	0,12	0,73	3,28	8,22	8,25	8,22	8,23	8,30	9,40
Lima	5,86	5,95	5,71	5,73	5,78	5,85	5,91	5,80	5,83	6,12	5,92	5,76	5,81	5,58	6,33	5,38	5,38	5,75	6,00	6,13	6,00	0,00	0,05	0,00	0,02	0,00	0,03	0,00	0,12	0,73	3,24	8,20	8,23	8,20	8,21	8,28	9,38
Neiva	5,84	5,92	5,69	5,71	5,76	5,83	5,88	5,78	5,81	6,10	5,89	5,73	5,79	5,56	6,31	5,38	5,38	5,76	6,01	6,13	6,01	0,00	0,00	0,00	0,02	0,00	0,03	0,00	0,12	0,73	3,28	8,22	8,25	8,22	8,23	8,30	9,40
Cávado	5,78	5,87	5,63	5,66	5,70	5,77	5,83	5,72	5,75	6,04	5,84	5,68	5,74	5,50	6,37	5,29	5,29	5,67	6,03	6,16	6,03	0,02	0,03	0,02	0,07	0,02	0,05	0,02	0,14	0,75	3,19	8,13	8,16	8,14	8,15	8,22	9,32
Ave	5,84	5,93	5,70	5,72	5,76	5,84	5,89	5,78	5,81	6,10	5,90	5,74	5,80	5,56	6,32	5,38	5,38	5,76	6,01	6,13	6,01	0,00	0,00	0,00	0,02	0,01	0,03	0,00	0,11	0,73	3,28	8,21	8,25	8,22	8,23	8,30	9,40
Douro	5,93	6,01	5,78	5,80	5,85	5,86	5,97	5,83	5,84	6,12	5,98	5,81	5,88	5,65	6,32	5,37	5,37	5,74	5,99	6,12	5,99	0,03	0,03	0,03	0,06	0,03	0,21	0,03	0,15	0,67	3,14	8,24	8,28	8,25	8,26	8,33	9,34
Vouga	5,85	5,93	5,70	5,72	5,76	5,84	5,89	5,78	5,81	6,10	5,90	5,74	5,80	5,56	6,32	5,38	5,38	5,76	6,01	6,13	6,01	0,00	0,00	0,00	0,02	0,00	0,03	0,01	0,12	0,73	3,28	8,22	8,25	8,22	8,23	8,30	9,40
Mondego b	5,88	6,05	5,81	5,84	5,88	5,96	6,01	5,90	5,93	6,22	6,02	5,86	5,92	5,68	6,44	5,51	5,51	5,88	6,13	6,26	6,13	0,13	0,13	0,13	0,15	0,11	0,16	0,13	0,00	0,85	3,39	8,32	8,37	8,34	8,35	8,42	9,52
Mondego c	5,27	5,36	5,12	5,19	5,19	5,26	5,31	5,21	5,24	5,58	5,33	5,17	5,22	4,71	5,78	4,76	4,76	5,14	5,39	5,51	5,39	0,77	0,77	0,77	0,79	0,77	0,70	0,77	0,89	0,10	3,05	8,21	8,24	8,22	8,23	8,30	8,90
Tagus Zêzere	6,07	6,41	6,11	5,96	6,09	5,87	5,86	5,90	5,84	5,90	6,06	5,74	5,93	5,57	7,08	5,82	5,82	5,94	6,69	6,82	6,69	3,44	3,40	3,44	3,35	3,44	3,29	3,44	3,56	3,20	0,10	9,36	9,00	9,47	9,48	8,96	10,05
S. aradensis																																					
Quarteira	8,84	8,77	8,57	8,72	8,67	9,16	9,08	9,19	9,18	10,77	9,23	8,99	9,04	8,77	10,95	8,74	8,74	9,20	10,08	10,21	9,96	8,63	8,61	8,63	8,54	8,63	8,66	8,63	8,74	8,62	9,83	2,12	0,27	0,12	0,13	0,48	5,18
Arade	8,50	8,60	8,34	8,37	8,39	8,83	8,75	8,86	8,85	9,55	8,88	8,65	8,69	8,44	9,95	8,56	8,56	8,87	9,75	9,88	9,63	8,66	8,64	8,66	8,57	8,66	8,69	8,66	8,79	8,66	9,45	1,50	0,33	0,39	0,39	0,04	5,17
Alvor	9,00	8,93	8,72	8,87	8,82	9,34	9,26	9,37	9,36	9,46	9,39	9,15	9,20	8,94	10,01	8,89	8,89	9,39	10,26	10,39	10,14	8,64	8,61	8,64	8,55	8,64	8,67	8,64	8,76	8,63	9,95	0,13	0,40	0,00	0,01	0,60	5,35
Aljezur	9,01	8,94	8,73	8,88	8,83	9,35	9,27	9,38	9,37	9,46	9,40	9,16	9,21	8,94	10,01	8,89	8,89	9,40	10,27	10,40	10,15	8,64	8,62	8,64	8,56	8,64	8,68	8,64	8,77	8,64	9,96	0,13	0,41	0,01	0,07	0,60	5,36
Seixe	8,46	8,64	8,36	8,33	8,38	8,80	8,72	8,83	8,82	8,91	8,84	8,61	8,65	8,39	9,72	8,59	8,59	8,84	9,72	9,85	9,60	8,72	8,70	8,72	8,63	8,72	8,75	8,72	8,84	8,72	9,40	0,50	0,05	0,63	0,63	0,04	5,23
S. torgalensis																																					
Mira	10,97	10,91	10,65	10,82	10,81	10,91	11,00	10,88	10,90	11,42	11,37	11,08	11,24	10,75	11,82	10,87	10,87	10,87	11,87	11,87	11,75	9,87	9,85	9,87	9,78	9,87	9,81	9,87	10,00	9,35	10,56	5,44	5,61	5,62	5,62	5,49	0,03

To infer more accurately the colonization routes followed by the complex, samples of the major drainages (Tagus and Guadiana) were grouped according to their geographic provenience: tributaries of the western left margin (WLM), eastern left margin (ELM), western right margin (WRM) and eastern right margin (ERM). Additionally, a fifth sub-group of the Tagus River containing the samples of the Alagon River (ALAG) was created since special haplotypes were previously reported to this tributary (Cunha et al., 2004). Concerning the river Mondego, besides the population of *S. alburnoides*, calculations were made independently for the groups of samples belonging to three distinct phylogroups found in this drainage: *S. pyrenaicus* (Mondego a), *S. carolitertii*-North (Mondego b) and *S. carolitertii*-Mondego (Mondego c).

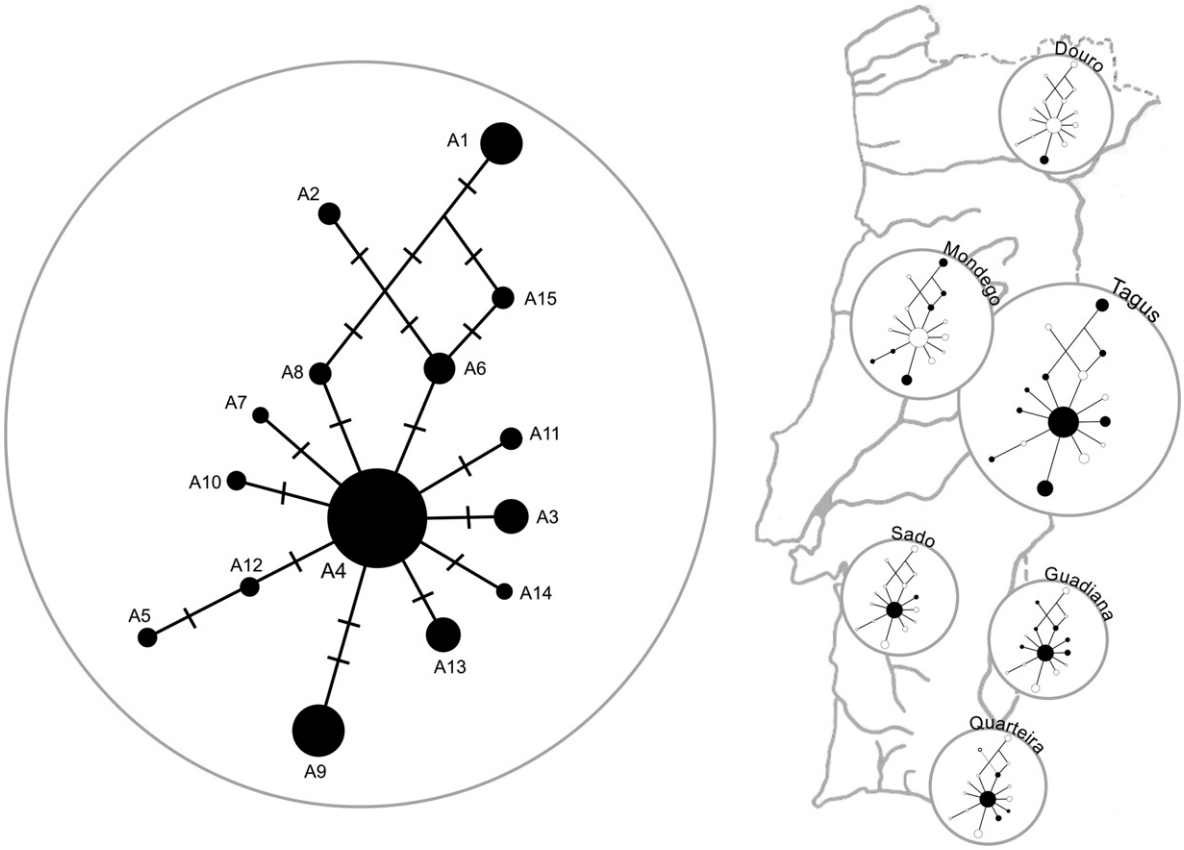

Fig. 3. Network of nuclear A-haplotypes and their respective occurrences in each of the populations/drainages. Haplotypes are represented by circles with diameters that are proportional to the number of individuals that shared each haplotype. Mutations between haplotypes are represented by the small lines perpendicular to the branch linking haplotypes. The haplotypes found in each population (map on the right) are depicted in black.

bers of other *Squalius* species that occurred at different time scales. As shown in Fig. 5, although the majority of the interspecific crosses were recent (15 shared haplotypes), there was also a high frequency of past interspecific crosses. Indeed, the average branch lengths measured from the missing common ancestor to the terminal nodes of each branch, ranged from 0.62 to 10.39 mutational steps. Note that, as stated above, only the branches of the tree that contained *S. alburnoides* and other *Squalius* species as terminal nodes were included in this analysis.

These findings clearly showed that *S. alburnoides* did not exchange mtDNA with other *Squalius* species only one to five times, as suggested by previous authors. Indeed, accepting a different origin of the complex for each major basin like Cunha et al. (2004), we would expect one or a few old common ancestors at the basal nodes of the subtree corresponding to that basin, with distinct evolutionary lines in *S. alburnoides* and other sympatric *Squalius* without shared haplotypes along the branches of the tree. On the contrary, the pattern presented in Figs. 5 and S1 clearly showed that sharing of haplotypes took place at multiple occasions within each basin, although the crosses were not so massive as to blur the distinctiveness of *S. alburnoides* and the other *Squalius*.

For the river basins Tagus and Guadiana, there were a sufficient number of samples of both *S. alburnoides* and *S. pyrenaicus* to perform a statistical comparison of the populations. In Table 4 we present the results of an AMOVA in which two groups were considered (Tagus and Guadiana) with two populations per group (*S. alburnoides* and *S. pyrenaicus* of the same drainage). Inspection of Table 4 shows that the variation among groups is much greater than that among populations, supporting the view that much of the history of *S. pyrenaicus* and *S. alburnoides* in each basin was shared. The populations of *S. alburnoides* and *S. pyrenaicus* from Tagus are significantly distinct ($p = 0.00$) and those from Guadiana approach significance ($p = 0.07$). At the same time, the corrected mean number of pairwise differences between *S. alburnoides* and *S. pyrenaicus* of the same basin were very low (Table 4). This could be explained in one of two ways: either there was a single origin of *S. alburnoides* in each basin and the populations were so recent that they had little time to diverge; or many instances of haplotype transfers between *S. pyrenaicus* and *S. alburnoides* took place on a much longer history. The within population variation is higher than the variation among populations and even greater than the variation among groups, indicating that each population had a considerably long history of mutation accumulation and haplotype diversification (Table 4)—the very low mean numbers of pairwise differences contrasted with the much higher levels of intrapopulation

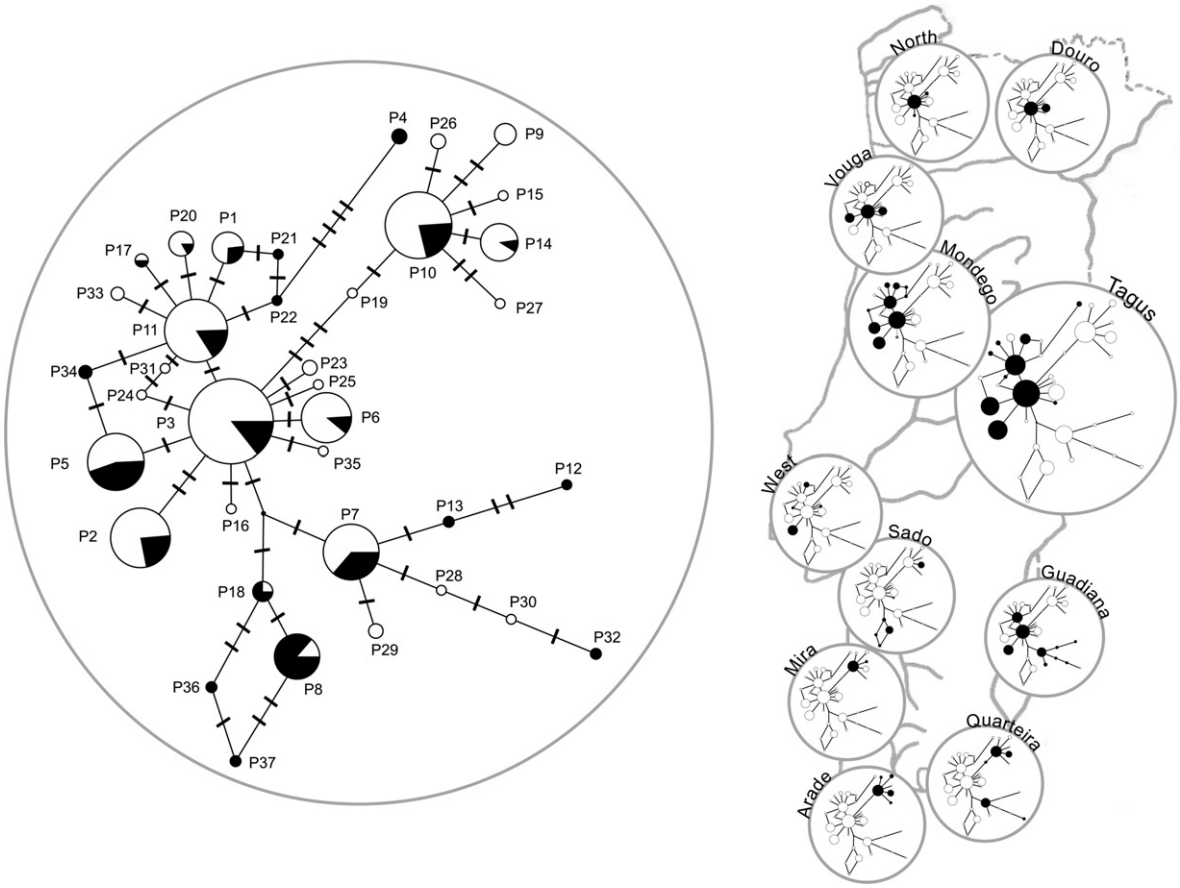

Fig. 4. Network of nuclear P-haplotypes, represented by circles with diameters that are proportional to the number of individuals that shared each haplotype. Each mutation between haplotypes is represented by a small line perpendicular to the branch linking haplotypes. Black and white circles were used to distinguish the haplotypes found, respectively, in *S. alburnoides* and in other *Squalius* species. When the haplotypes were shared between *S. alburnoides* and other *Squalius* species, black and white slices proportional to the respective number of individuals were depicted. The distribution of the haplotypes in the distinct populations is also depicted (haplotypes found in each drainage are black coloured).

Table 3
Distribution of the mtDNA of the different phylogroups of *S. alburnoide*s, as an indicator of the distinct levels of introgression by other species in the complex

	Douro	Mondego	Tagus	Sado	Guadiana	Quarteira	Guadalquivir
Phylogroup	$N = 31$	$N = 23$	$N = 49$	$N = 24$	$N = 40$	$N = 21$	$N = 4$
S. pyrenaicus Tagus/Guadiana	27 (87.1%)	18 (78.3%)	49 (100%)	2 (8.3%)	40 (100%)	2 (9.5%)	4 (100%)
S. pyrenaicus Sado	—	—	—	22 (91.7%)	—	—	—
S. carolitertii North	4 (12.9%)	1 (4.3%)	—	—	—	—	—
S. carolitertii Mondego	—	4 (17.4%)	—	—	—	—	—
S. aradensis	—	—	—	—	—	19 (90.5%)	—

For each river basin, the number (and percentage) of individuals with a given mtDNA type is indicated.

mean number of pairwise differences, as shown in Table 1. These results are in agreement with the structure of the tree (Figure S1) and of the networks presented in Fig. 2.

4. Discussion

The information retrieved from the molecular analyses led to the postulation of a single origin for the *S. alburnoides* complex (outlined below) and allowed the reconstitution of a hypothetical dispersal scenario that aimed to explain the foundation of populations in distinct river basins. However, since the dispersal of primary freshwater fish is only possible through fluvial connections, the routes followed by *S. alburnoides* had to be corroborated by the phylogeography of other *Squalius*.

4.1. Phylogeographical patterns of the Squalius species

The general pattern that emerged from the phylogenetic and phylogeographic analyses was that *S. pyrenaicus* is a highly diversified species with a wide distribution range whose ancestral population originated at least five new spe-

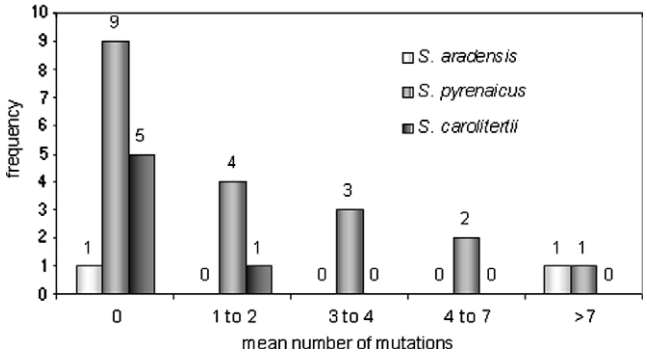

Fig. 5. Frequency of interspecific crosses, represented in the phylogenetic tree depicted in Figure S1 by shared haplotypes or missing common ancestors between *S. alburnoides* and other *Squalius* species, against a temporal line expressed by the mean number of mutations leading from the common ancestor to its terminal nodes.

cies as it dispersed towards the peripheral areas of the Iberian Peninsula: *S. carolitertii* in the North, *S. aradensis* and *S. torgalensis* in the Southwest, *S. valentinus* in the Southeast and *S. malacitanus* in the South (not studied; see Doadrio and Carmona, 2006). This picture involving a central, very widely distributed clade which originated a star-like pattern of peripheral derived clades resembles the peripatric speciation model proposed by Mayr (1982) and is congruent with our understanding of the Miocenic hydrography of Iberia, with a number of large endorheic lakes which subsequently branched, underwent fragmentation and became connected with rivers.

A detailed chronological description of those phylogeographical events, with estimated dates based on the estimated mtDNA divergence times (Table 2) and supported by geological data, is presented below—for a synthesis see Fig. 6.

4.2. Miocenic pathways (Fig. 6a)

Before it became an exorheic river in the Pliocene (Cunha et al., 1993; Andeweg, 2002), the River Tagus was a system of at least five endorheic lakes, likely connected at some time since Middle Miocene fossils morphologically similar to extant *Squalius* species (Doadrio and Carmona, 2006), were found in the Lower Tagus basin (Gaudant, 1977).

The Lower Tagus basin (Fig. 6a) may have acted as a littoral corridor that allowed the colonization of the southern Rivers Mira and Arade by a *Squalius* ancestor. Indeed, since *S. aradensis* and *S. torgalensis* are sister-species (Brito et al., 1997; Coelho et al., 1998), their differentiation depended on the arrival of a common ancestor to southwest Portugal at least in the Upper Miocene, the estimated age of the common ancestor of these species with the common ancestor of *S. pyrenaicus*/*S. carolitertii* (Doadrio and Carmona, 2003; Sanjur et al., 2003). Hypothetically, two routes could bring primary freshwater fish to southwest Portugal at that time: one from the East (from the Guadiana) or one from the North (involving the Lower Tagus and the primitive basin of the Sado, which was intermittently connected with the Tagus—T. Azevedo pers. com.). This last scenario seems more likely since (1) the southwestern endemism *Iberochondrostoma almacai* diverged from a common ancestor with *I. lusitanicum* (Robalo et al., 2007) that occurs in Sado but is absent from Guadiana; (2) *Squalius* from Arade and Sado shared a beta-actin haplotype; (3) the Guadiana River only drained to the south in the Pleistocene (Rodríguez-Vidal et al., 1991, 1993), which is posterior to the estimated age of these species; and (4) cyprinid species that inhabit the Guadiana are absent from the southwestern area (*Anaecypris hispanica*, *Pseudochondrostoma willkommii* and *Barbus microcephalus*).

The elevation of the Caldeirão Mountain, between the Guadiana and the southwestern rivers of Arade and Mira, 5.3–3.4 MY ago (Dias, 2001), must have isolated the recently arrived *Squalius* ancestor. Moreover, the geomorphological changes that took place may have isolated a subpopulation in the River Mira and another in the River Arade, allowing the differentiation of *S. torgalensis* and *S. aradensis*, respectively. The estimated age of about 5.13 MY for the divergence between these two species is congruent with the timing of the elevation of the Caldeirão Mountain and with the divergence values obtained by Doadrio and Carmona (2003), Sanjur et al. (2003) and Mesquita et al. (2005). The haplotype P10 of the beta-actin gene found in specimens from the Mira and Arade Rivers may be one of the last vestiges of the connection between the two populations.

4.3. Pliocenic pathways (Fig. 6b)

The northward migration to the Mondego of a common ancestor to the *S. pyrenaicus*-Tagus/Guadiana phylogroup

Table 4
AMOVA using pairwise differences with two groups (Tagus and Guadiana), each with two populations: *S. alburnoides* and *S. pyrenaicus* from Tagus (albT and pyrT); and *S. alburnoides* and *S. pyrenaicus* from Guadiana (albG and pyrG)

	pyrT	albT	pyrG	albG		DF	Variance components	% of variation
pyrT	15.881	0.000	0.000	0.000	Among groups	1	2.234	32.03
albT	1.119	7.051	0.000	0.000	Among populations within groups	2	0.358	4.90
pyrG	4.873	4.904	4.165	0.090	Within populations	179	4.450	63.06
albG	5.218	5.268	0.139	5.362	Total	182	7.042	

The corrected mean number of pairwise differences within (diagonal) and between (below diagonal) populations is also presented. Significance *p* values for the exact test of population differentiation are indicated above diagonal.

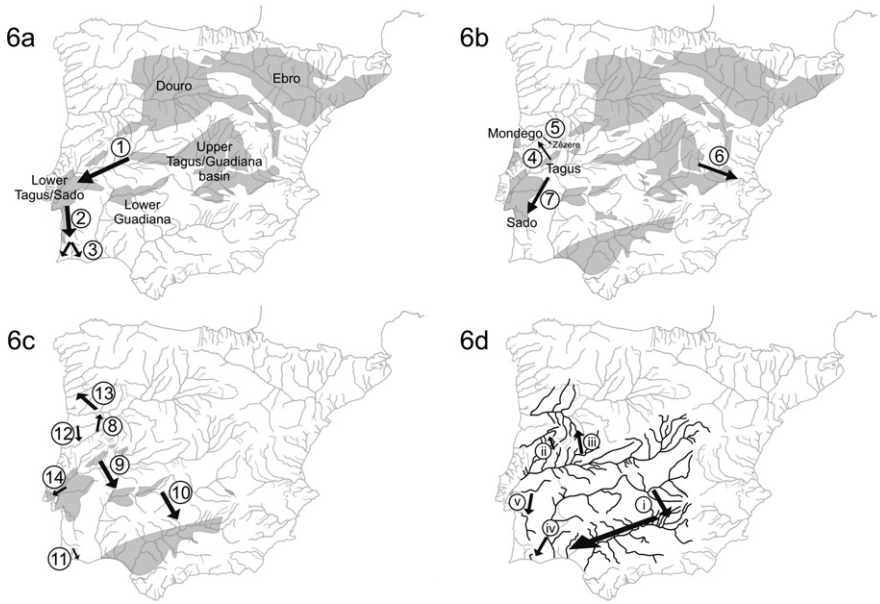

Fig. 6. Hypothesized pathways for *Squalius* species in the colonization of Iberian drainages. The numbered arrows represent the hypothetical pathways occurring in the Miocene (6a), Pliocene (6b), Pleistocene (6c) and Pleisto–Holocene (6d). Dates for each colonization event were estimated from the mtDNA divergence times (Table 2). Representations of the endorheic lakes were adapted from Andeweg (2002). Legend: 1—Arrival of a proto-*S. pyrenaicus* to western Iberia (Middle Miocene); 2—Colonization of the Southwest (Upper Miocene); 3—Differentiation of *S. aradensis* and *S. torgalensis* (5.13 MY); 4—Differentiation of the *S. carolitertii*-Mondego phylogroup (4.49 MY); 5—Differentiation of the *S. carolitertii*-Zêzere phylogroup (3.05 MY); 6—Differentiation of the *S. valentinus* phylogroup (2.05 MY); 7—Differentiation of the *S. pyrenaicus*-Sado phylogroup (2.01 MY); 8—Differentiation of the *S. carolitertii*-North phylogroup (0.72 MY); 9—Renewed contacts between Tagus and Guadiana (0.53 MY); 10—Dispersal Upper Guadiana–Upper Guadalquivir (0.39 MY); 11—Dispersal Arade–Quarteira (0.27 MY); 12—Dispersal Vouga–Mondego (0.12 MY); 13—Northward radiation of *S. carolitertii* (0.05-0.00 MY); 14—Dispersal Tagus–West (0.04 MY); (i to v)—hypothesized colonization pathways followed by *S. alburnoides* (see also text).

was corroborated by the positioning of the haplotypes of *S. carolitertii* from Mondego in both mtDNA and beta-actin networks. From the Middle Miocene to the Early Pliocene, at least three small Tagus endorheic lagoons were present in the vicinity of the spring of the Mondego River (Andeweg, 2002). Between 3.6 and 2.6 MY ago the Tagus River was acquiring its longitudinal profile, as the Upper Tagus basin drained towards west and the formerly isolated endorheic lakes were being united (Cunha et al., 1993). Thus, a connection with the adjacent Mondego basin was plausible, allowing the passage of the *S. carolitertii* ancestor.

According to the mtDNA analysis, *S. carolitertii* from the Zêzere River, a tributary of the right bank of the River Tagus located in the southeast side of the Estrela Mountain, diverged from that of Mondego in the Middle Pliocene. This suggests that after an initial dispersal of its ancestor from Tagus to Mondego, the derived phylogroup *S. carolitertii*-Mondego reinvaded a Tagus tributary (Zêzere) at a later date, leading to the foundation of the *S. carolitertii*-Zêzere phylogroup. Our hypothesis is that a tributary of Mondego (whose headwaters are on the northwestern side of the Estrela Mountain but only about 15 km away from the sampling location of the Zêzere) must have drained to one of the endorheic lakes that existed in the area (Cunha et al., 1993).

The foundation of the Zêzere population by migrants proceeding from the Douro seems unlikely since their divergence was higher (3.39%) than with the Mondego population (3.20%); the most common nuclear haplotype in Douro (P6) is absent from Zêzere; and there is no geological support for contacts between Zêzere and Douro basins.

In the Pliocene, *S. pyrenaicus* from the Tagus also seem to have dispersed southwards, reinvading the Sado (2.11 MY ago) and allowing the foundation of the *S. pyrenaicus*-Sado phylogroup (represented by the nuclear haplotypes P8, P18, P36 and P37). This dispersal path is supported by the mtDNA networks (Fig. 2) and by the nuclear P-haplotypes (Fig. 4). To corroborate the Tagus-Sado pathway in detriment of a hypothetical Guadiana–Sado pathway is the fact that the nase and barbel species present in Sado are *Pseudochondrostoma polylepis* and *Barbus bocagei* that are present in Tagus but not in Guadiana. The Albufeira Lagoon, located between the mouths of the Tagus and Sado rivers, and inhabited by *I. lusitanicum* (Collares-Pereira, 1983; Robalo et al., 2007), may be considered a last vestige of the connection between the two rivers. The differentiation of the *S. pyrenaicus*-Sado phylogroup could have been favoured by the isolation of subpopulations in more elevated areas were freshwaters persisted during the frequent transgression episodes that caused an intermittent regime of contact with the adjacent Tagus drainage (Pimentel, 1997; Andeweg, 2002). Additionally, in the Upper Pliocene, a climatic crisis impeded the persistence of fluvial canals and was responsible for

the disorganization of the Sado drainage network (Pimentel, 1997). Consequently, the original *Squalius* populations may have suffered declines, accelerating the process of lineage sorting and causing the loss of the proto-*S. aradensis* mtDNA (the presence of the haplotype P14 in a fish from Sado is probably a last vestige of the Miocenic spread of *Squalius* to Arade through the Tagus-Sado corridor).

As it presents divergence values from *S. pyrenaicus* that are similar to the ones that allowed the description of *S. valentinus* as a distinct species (Doadrio and Carmona, 2006), we suggest that the *S. pyrenaicus* population from Sado River basin should be eventually considered a new species.

In the Pliocene, the Upper Guadiana (which was isolated from its Lower section) was a tributary of the Upper Tagus (Moya-Palomares, 2002), so that both rivers must have shared the same *S. pyrenaicus* populations at that period. The existence of two distinct sets of species distributed in the Upper Tagus (*Iberochondrostoma lemmingii* and *Achondrostoma arcasii*) and in the Lower Tagus (*I. lusitanicum*), with similarities with the species present, respectively, in the Guadiana and Sado basins, led us to postulate that the Tagus may have established independent connections with these two basins, at a time when the communication between its Lower and Upper sections was interrupted. The fact that *B. comizo* is shared by Tagus and Guadiana but not Sado and *P. polylepis* is shared by Tagus and Sado but not Guadiana, also support this view.

The geographic proximity between the Upper Tagus and tributaries of some Mediterranean rivers may have also allowed the migration of the *S. valentinus* ancestors in the Pliocene.

4.4. Pleistocenic pathways (Fig. 6c)

In the Pleistocene the Iberian hydrographical network acquired its current profile but some connections between drainages were still possible. The following Pleistocenic colonizations were postulated according to the topology of the network of mtDNA haplotypes and to the respective values of divergence between haplotypes:

(1) Tagus–Guadiana (0.53 MY ago)—a contact between the lower sections of both drainages would explain the occurrence of four *S. pyrenaicus* from Guadiana carrying Tagus-like mtDNA and seems plausible according to geological data: one or more Portuguese tributaries of the Guadiana, in the region of Mora-Pavia, were tributaries of the Tagus (T. Azevedo, pers. com.); and contacts between the two basins occurred in the Badajoz area (Moya-Palomares, 2002);
(2) Tagus–Western rivers (0.04 MY ago)—migration of *S. pyrenaicus* to Rivers Lizandro, Samarra and Colares prior to the arrival of *S. alburnoides* to the lower section of the Tagus (since the complex is absent from the western rivers);
(3) Guadiana–Guadalquivir (0.39 MY ago)—migrants of *S. pyrenaicus* proceeding from the Upper Guadiana must have reached the adjacent Guadalquivir drainage;
(4) Arade–Quarteira (0.27 MY ago)—the geographic proximity between tributaries in the lowlands south of the Caldeirão Mountain might have allowed the migration of *S. aradensis* from Arade to Quarteira;
(5) Mondego–Douro (0.35 MY ago)—the low mitochondrial and nuclear diversity of the *S. carolitertii* populations of the Douro corroborates a very recent colonization and/or reflects major depletions of the original fauna caused by glaciations. A similar pattern of higher levels of genetic diversity in Mondego populations when compared to the Douro and other northern populations was also detected in the golden-striped salamanders *Chioglossa lusitanica* (Alexandrino et al., 2002), for which a recent colonization by a small number of founders was suggested to explain the almost genetically uniform populations located north of the Douro;
(6) Douro–Northern rivers (0.06 to 0.03 MY ago)—the very recent radiation of *S. carolitertii* to the other northern rivers is supported by the star-like mtDNA network (with a highly abundant root haplotype and many closely associated haplotypes) and may have been favoured by the major regression that took place 0.018 MY ago, during which almost all of the Portuguese continental shelf was above sea level, allowing the confluence of the mouths of the northern rivers (Dias et al., 2000);
(7) Vouga–Mondego (0.13 MY ago)—postulated to explain the sharing of the P5 nuclear haplotype and the existence of mtDNA haplotypes belonging to the *S. carolitertii*-North phylogroup in the Mondego drainage.

4.5. Origin of the S. alburnoides complex

In previous studies (Alves et al., 1997b; Cunha et al., 2004) the similarities between the mtDNA haplotypes of *S. alburnoides* and of other *Squalius* from the same river basin were interpreted as evidence of an independent origin of the complex in that particular river basin. In our view, however, they may reflect the occurrence of recent interspecific crosses involving females of the sympatric *Squalius* species. As hypothesized by Sousa-Santos et al. (2006a) and corroborated by the results from the present work, the available data are consistent with a single origin for the *S. alburnoides* complex, when both the maternal and paternal ancestors became sympatric, due to the historical rearrangements of the Iberian hydrographical network. This hypothesis seems more parsimonious than admitting the prolonged coexistence of the maternal and paternal ancestors in multiple river basins, the independent synthesis of the complex in each of those basins, and the subse-

quent extinction of one or both the ancestors depending on the river basin considered (Sousa-Santos et al., 2006a). The much higher levels of intrapopulation mean number of pairwise differences when compared to the very low mean numbers of pairwise differences between populations (Table 2) also contradict the hypothesis of a single origin per basin and supports the view that, from time to time, haplotypes of one population passed on to the other.

According to our findings, the origin of the complex must have occurred in the bulk of Iberia, in the Middle–Upper Pleistocene (less than 0.7 MY ago), more recently than the Upper Pliocene age proposed by Cunha et al. (2004). Our hypothesis is that the differentiation of the *Anaecypris*-like paternal ancestor of the complex occurred in a southern endorheic lake that remained isolated until it was captured by an ancient river carrying the *S. pyrenaicus* maternal ancestor. We suggest that this refuge was located in the area of what is now the River Guadiana since the distribution of the extant *A. hispanica* is restricted to this basin and the paternal ancestor, belonging to a derived clade, was also presumably favoured by the southern ecological conditions (namely, higher temperatures and intermitent conditions).

The capture of the endorheic lake must have been possible since with the tilting of the Peninsula towards the Atlantic in the Pleistocene, the Tagus and Guadiana rivers (which, at the time, were connected in the area where are now the headwaters of the Guadiana) began to drain towards west, acquiring their present longitudinal profiles and capturing the isolated endorheic lakes located on the way (Moya-Palomares, 2002). As a result, the paternal ancestor of the complex must have become sympatric with *S. pyrenaicus* (at the time already differentiated in Tagus and Guadiana) and interspecific crosses gave rise to the complex. Afterwards, with the ongoing basculation process, the Tagus and Guadiana became completely isolated from each other but continued their path towards west (Moya-Palomares, 2002), already carrying their respective *S. alburnoides* and *S. pyrenaicus* populations.

4.6. Dispersal of the S. alburnoides complex

Once originated, the complex must have dispersed throughout the connections between river basins that were still available in the Upper Pleistocene–Holocene, which explains why it has a wide distribution in the main drainages and is absent from the smaller and peripheral river basins of the Peninsula, already isolated at that time. These colonizations allowed the contact, not only with different *S. pyrenaicus* populations, but also with other *Squalius* species with which the complex interbreeds.

The dispersal route of the *S. alburnoides* complex, based on the estimated mtDNA divergence times (Table 2), likely included at least five colonization paths (represented by the same arabic numbers in the text below and in Fig. 6d):

(i) From Upper Guadiana to Upper Guadalquivir (0.05 MY ago)—Path corroborated by the lowest divergence values between the *S. alburnoides* populations from Guadalquivir and from the left bank tributaries of the Upper Guadiana. Stream captures may also explain the migration of *S. alburnoides* from Guadalquivir to the adjacent Odiel drainage;

(ii) From Tagus to Mondego (0.05 MY ago)—Through fluvial captures involving adjacent tributaries of the right bank of the Tagus basin, as corroborated by the lower divergence values involving *S. alburnoides* from the Zêzere-Erges area. These contacts may have also allowed the migration of *S. pyrenaicus* whose genes were probably diluted in the more abundant populations of its sister-species *S. carolitertii*. The five presumably *S. carolitertii* individuals from Mondego with *S. pyrenaicus* mtDNA may either be true *S. pyrenaicus* proceeding from Tagus or, alternatively, may be reconstituted from crosses between PPA females (carrying the *S. pyrenaicus* mtDNA) and *S. carolitertii* males. The later hypothesis would be discarded if the nuclear genomes of these five individuals showed P-haplotypes that were exclusive of the Tagus basin. However, two individuals were homozygous for a haplotype that was shared between Mondego and Tagus (P5) and the remaining three were heterozygous with one or both complements shared between the two basins. Brito et al. (1997) also found one *S. pyrenaicus* individual in the Mondego but interpreted it as a result of anthropogenic introductions;

(iii) From Tagus to Douro (0.01 MY ago)—The *S. alburnoides* population of Douro showed a very low divergence value from the population of the Alagon river (tributary of Tagus, in the vicinity of the Portuguese border), which suggested that this may have been the corridor used in the colonization of the Douro basin. This hypothesis is corroborated by the fact that, in contrast to the wide distribution in the Portuguese Douro basin, the distribution of the complex in the Spanish Douro basin is restricted to a few tributaries of the left bank that are located in the vicinity of the Alagon area. Thus, after the colonization of those tributaries, the *S. alburnoides* complex may have reached the main course of Douro and, from there, dispersed virtually to all Portuguese tributaries. An upstream migration may have been impeded by the existence of a geological barrier of about 400 meters near the Portuguese border (Ribeiro et al., 1987). As in the case of the colonization of Mondego, the contact between Tagus and Douro may also have allowed the passage of *S. pyrenaicus*. This introgression was not yet detected but increased sampling effort and the use of new nuclear markers (the beta-actin does not differ between *S. carolitertii* and *S. pyrenaicus*) will very likely solve this issue;

(iv) From lower Guadiana to Quarteira (0 MY ago)—After the colonization of Quarteira by *S. aradensis* from Arade River, a second colonization might have occurred: *S. alburnoides* proceeding from Guadiana seem to have colonized this river basin very recently, when the contact with the Arade had already ceased (since the complex is absent from Arade). The Guadiana acquired its present configuration and a southward draining pattern (that must have allowed the connections with Quarteira) very recently, in the Upper Pleistocene (Rodríguez-Vidal et al., 1993), which is in accordance with our results. The presence of *I. lemmingii*, which is present in Guadiana and Quarteira but not in Arade, also supports this route;

(v) From Tagus to Sado (0 MY ago)—The colonization of Sado probably occurred when the upper section of the Tagus, carrying the *S. alburnoides* complex, merged with its lower section, yet connected with the Sado River basin.The *S. alburnoides* from Sado, in addition to mtDNA that is typical of the *S. pyrenaicus* from this basin, also exhibited Tagus-like mtDNA, which corroborates the postulation of a third dispersal wave from Tagus towards Sado (see the other two postulated episodes above), that according to the geomorphological history of both drainages is not unlikely.

4.7. Relationships between *S. alburnoides* and other *Squalius* species

The mtDNA analysis showed a low number of haplotypes shared between *S. alburnoides* and other *Squalius* species, indicating that present crosses involving *S. alburnoides* males and females of other *Squalius* species are scarce. Scarcity is not, however, synonymous of absence and some proofs of the occurrence of interspecific crosses were found: (1) the complete replacement of the typical *S. pyrenaicus* mtDNA of the complex by *S. aradensis* mtDNA in Quarteira; and (2) the introgression of *S. carolitertii* mtDNA in some *S. alburnoides* individuals from Mondego and Douro. Thus, the reproduction of the *S. alburnoides* complex seems to involve mating with conspecifics, with males of other *Squalius* species, and, at least occasionally, with females of all the three sympatric *Squalius* species (*S. pyrenaicus*, *S. carolitertii* and *S. aradensis*), allowing the introgression of their mtDNA in the complex.

The presence of non-hybrid *S. alburnoides* males seems to be of extreme relevance to the process of replacement of the typical mtDNA of the complex. These males are probably more efficient in the diffusion of the mtDNA of other *Squalius* species, as corroborated by the finding that in the river basins where AA males are abundant (Sado, Guadiana, Tagus, Guadalquivir and Quarteira), the mtDNA of the *S. alburnoides* is identical to the mtDNA of the sympatric *Squalius* species. This is probably a result of higher attractiveness and fertilization success of these small males (Sousa-Santos et al., 2006b).

In contrast, in the river basins where non-hybrid males are absent (Mondego and Douro) the detected levels of introgression were much inferior. According to the proposed dispersal scenarios, the Mondego and Douro rivers were apparently colonised by *S. alburnoides* proceeding from tributaries of the right bank of the Tagus, where non-hybrid *S. alburnoides* males have not been found. This virtual absence of non-hybrids may be explained by unfavourable ecological conditions since they seem to prefer shallow waters with higher temperatures (Martins et al., 1998). Indeed, the tributaries of the right bank of the Tagus have higher discharges and lower water temperatures when compared to the tributaries of the left bank, whose ecological regimes resemble more the ones from the southern Mediterranean rivers. Moreover, the calculated age of the colonization of Mondego (0.05 MY) predated the last glacial maximum (0.018 MY), that may have been responsible for severe bottlenecks, as river discharges were extremely higher due to a longer pluvial season and to the effect of spring ice melting (Dias et al., 2000), combined with a cooling that was probably unfavourable to AA males. Moreover, the persistence of non-hybrid males in populations is self-dependent, as they can only be originated by crosses between males of their own type and PAA females producing A gametes (Alves et al., 2002; Sousa-Santos et al., 2006b). Thus, if a secondary loss of this kind of males occurred in the northern populations, it is unlikely that they could be originated *de novo*.

Conversely, while non-hybrids may contribute to the introgression of distinct mtDNA in the complex, triploid PPA females might play an extremely important role in the introgression of nuclear and mtDNA in other *Squalius* species. As these females discard the uneven genome and perform normal meiosis (Crespo-López et al., 2006), the generated eggs carry a single P-haplotype. Thus, populations with abundant PPA females reflect a certain degree of autonomy from the sympatric *Squalius* species as P-donors. Additionally, the PPA females that colonized the Douro and Mondego drainages and crossed with *S. carolitertii* males transmitted nuclear genes of *S. pyrenaicus* to the offspring. However, if only mtDNA sequencing was performed, this transference of nuclear genes would be undetectable since the resultant offspring would be classified as *S. pyrenaicus*, when they should instead be classified as hybrids between *S. pyrenaicus* and *S. carolitertii*. Thus, the mtDNA analysis, when considered alone, may underestimate the extent of gene introgression between *Squalius* species.

The *S. alburnoides* complex is, therefore, besides being introgressed with sympatric *Squalius* genes, also responsible for the transference of mitochondrial and nuclear genes to different *Squalius* species, contributing to a homogenization of the *Squalius* genomes. This situation is particularly relevant at the nuclear DNA level since recombination between nuclear genes belonging to distinct species may

occur, raising taxonomical problems related to the definition of the species.

To conclude, after reading the history of this hybrid complex from its molecular record, our results may be summarized as follows: (1) its origin may be traced back to the Pleistocene; (2) it is likely to have had a single origin, from hybridizations between an extinct *Anaecypris*-like species and *S. pyrenaicus* in the centre of Iberia, in the area of the Upper Tagus/Upper Guadiana; (3) it apparently dispersed afterwards along several routes, namely: Guadiana–Guadalquivir–Odiel; Guadiana–Quarteira; Tagus–Sado; Tagus–Douro and Tagus–Mondego; and (4) it may have played a major role in bidirectional nuclear and mtDNA gene transfer with allopatric species and populations of *Squalius*.

Many fish hybrid lineages are mere sinks for the genes of sexual species they parasitize for reproducing. However, *S. alburnoides*, in its history of about 700,000 years, interacted with several other *Squalius* species, promoting bidirectional gene transfers. In this respect, the peculiar modes of reproduction of this hybrid complex emphasized by Alves et al. (2001), place it in an unique position not only in terms of its own evolution but in the evolutionary dynamics of other fish species. If the scenarios here reconstructed are correct, the dispersion/colonization paths of many other primary fish species might have occurred using the same fluvial connections. Thus, it is essential to delineate a wider research program with a much more intense sampling of other sympatric species and molecular markers to test for the signature of the events now postulated.

In this study we combined conventional phylogenetic inference procedures with phylogeographical tools and geological information, and our results suggest that phylogeographical analysis of slowly evolving nuclear genes like the beta-actin gene, may help to get a better picture of the past of a clade because these genes will be also more slowly affected by the processes leading to lineage sorting. Thus, ancestral haplotypes and historic relationships that left no equivalent signature in the mtDNA may be recovered, providing ways to get a more accurate phylogenetic reconstruction.

This research also illustrated the advantage of analysing phylogroups of mtDNA haplotypes instead of simply taking each species as a single collection of samples, as a way of identifying the relevant clades. Thus, the concerted use of phylogenetic and phylogeographical methods designed for studies at various time scales may be considered a promising combination of tools in paleobiogeography. In the future, the use of nuclear genes varying in their rates of evolution, combined with mtDNA analysis, may provide the necessary tools for validating the history of groups of organisms at the multiple time scales now advocated.

Acknowledgments

We thank Sousa-Santos family and J. Robalo for help in sample collection and T. Azevedo for the helpful informations on the geological history of the Guadiana and Tagus drainages. Samples of *S. alburnoides* from River Almonte (Tagus) were kindly supplied by I. Doadrio. The study was funded by the FCT Pluriannual Program (UI & D 331/94 and UI & D 329/94) (FEDER participation). C. Sousa-Santos was supported by a Ph.D Grant from FCT (SFRH/BD/8320/2002).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2007.05.011.

References

Alexandrino, J., Arntzen, J.W., Ferrand, N., 2002. Nested clade analysis and the genetic evidence for population expansion in the phylogeography of the golden-striped salamander, *Chioglossa lusitanica* (Amphibia: Urodela). Heredity 88, 66–74.

Almaça, C., 1978. Répartition géographique des Cyprinidae Ibériques et secteurs ichthyogéographiques de la Péninsule Ibérique. Vestn. Cesk. Spol. Zool. 42, 241–248.

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

Alves, M.J., Coelho, M.M., Collares-Pereira, M.J., Dowling, T.E., 1997b. Maternal ancestry of the *Rutilus alburnoides* complex (Teleostei, Cyprinidae) as determined by analysis of cytochrome *b* sequences. Evolution 51, 1584–1592.

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

Alves, M.J., Gromicho, M., Collares-Pereira, M.J., Crespo-López, E., Coelho, M.M., 2004. Simultaneous production of triploid and haploid eggs by triploid *Squalius akburnoides* (Teleostei: Cyprinidae). J. Exp. Zool. 301A, 552–558.

Andeweg, B., 2002. Cenozoic tectonic evolution of the Iberian Peninsula, causes and effects of changing stress fields. PhD Thesis. Vrije Universiteit Amsterdam, Amsterdam, 178pp.

Banarescu, P., 1973. Some reconsiderations on the zoogeography of the Euro-Mediterranean freshwater fish fauna. Revue Romaine de Biologie (Zoologie) 18, 257–264.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol 1681, 37–48.

Bianco, P.G., 1990. Potential role of the paleohistory of the Mediterranean and Paratethys basins on the early dispersal of Euro-Mediterranean freshwater fishes. Ichthyol. Explor. Freshwaters 1, 167–184.

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

Cabral, M.J., Almeida, J., Almeida, P.R., Dellinger, T., Ferrand de Almeida, N., Oliveira, M.E., Palmeirim, J.M., Queiroz, A.I., Rogado, L., Santos-Reis, M. (Eds.), 2005. Livro Vermelho dos Vertebrados de Portugal. Instituto de Conservação da Natureza, Lisboa, p. 660pp.

Cabrera, L., Gaudant, J., 1985. Los Ciprínidos (Pisces) del sistema lacustre Oligocénico-Miocénico de los Monegros (sector SE de la Cuenca del Ebro, provincias de Lleida, Tarragona, Hoesca y Zaragoza). Acta Geol. Hisp. 20, 219–226.

Carmona, J.A., Sanjur, O.I., Doadrio, I., Machordom, A., Vrijenhoek, V.C., 1997. Hybridogenetic reproduction and maternal ancestry of

polyploid Iberian fish: the *Tropidophoxinellus alburnoides* complex. Genetics 146, 983–993.

Coelho, M.M., Bogutskaya, N., Rodrigues, J.A., Collares-Pereira, M.J., 1998. *Leuciscus torgalensis* and *Laradensis*, two new cyprinids for Portuguese fresh waters. J. Fish Biol. 52, 937–950.

Collares-Pereira, M.J., 1983. Estudo sistemático e citogenético dos pequenos ciprinídeos ibéricos pertencentes aos géneros *Chondrostoma* Agassiz, 1835, *Rutilus* Rafinesque, 1820 and *Anaecypris* Collares-Pereira, 1983. Ph.D. Thesis. University of Lisbon, Lisbon, Portugal.

Crespo-López, M.E., Duarte, T., Dowling, T., Coelho, M.M., 2006. Modes of reproduction of the hybridogenetic fish *Squalius alburnoides* in the Tejo and Guadiana rivers: an approach using microsatellites. Zoology 109, 277–286.

Cunha, P., Barbosa, B., Reis, R., 1993. Synthesis of the Piacenzian onshore record between the Aveiro and Setúbal parallels (Western Portuguese margin). Ciências da Terra (UNL) 12, 35–43.

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.

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

De la Peña, A., 1995. Tertiary fishes from the Iberia continental basins: History and fossil record. Coloquios de Paleontologia 47. Editorial Complutense, Madrid.

Dias, J.M.A., Boski, T., Rodrigues, A., Magalhães, F., 2000. Coast line evolution in Portugal since the Last Glacial Maximum until present—a synthesis. Mar. Geol. 170, 177–186.

Dias, R.P., 2001. Neotectónica da região do Algarve. Ph.D Thesis. University of Lisbon, Lisbon, 369pp.

Doadrio, I., Carmona, J.A., 2003. Testing freshwater Lago Mare dispersal theory on the phylogeny relationships of Iberian cyprinid genera *Chondrostoma* and *Squalius* (Cypriniformes, Cyprinidae). Graellsia 59, 457–473.

Doadrio, I., Carmona, J.A., 2006. Phylogenetic overview of the genus *Squalius* (Actinopterygii, Cyprinidae) in the Iberian Peninsula, with description of two new species. Cybium 30, 199–214.

Dowling, T., Tibbets, C.A., Minckley, W.L., Smith, G.R., 2002. Evolutionary eelationships of the Plagopterins (Teleostei: Cyprinidae) from cytochrome *b* sequences. Copeia 2002 (3), 665–678.

Excoffier, L., Laval, G., Schneider, S., 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1, 47–50.

Friend, P.F., Dabrio, C.J. (Eds.), 1996. Tertiary Basins of Spain: The Stratigraphic Record of Crustal Kinematics. Cambridge University Press, Cambridge, p. 400.

Froese, R., Pauly, D. (Eds.), 2007. FishBase. World Wide Web electronic publication, www.fishbase.org, version 01/2007.

Gaudant, J., 1977. Contributions à la paléontologie du Miocène moyen continental du bassin du Tage. II. Observations sur les dents pharyngiennes de poissons cyprinidés-Póvoa de Santarém. Ciências da Terra (UNL) 3, 129–141.

Gromicho, M., Coelho, M.M., Alves, M.J., Collares-Pereira, M.J., 2006. Cytogenetic analysis of *Anaecypris hispanica* and its relationship with the paternal ancestor of the diploid-polyploid *Squalius alburnoides* complex. Genome 49, 1621–1628.

Gromicho, M., Collares-Pereira, M.J., 2004. Polymorphism of major ribosomal gene chromosomal sites (NOR-phenotypes) in the hybridogenetic fish *Squalius alburnoides* complex (Cyprinidae) assessed through crossing experiments. Genetica 122, 291–302.

Martins, M.J., Collares-Pereira, M.J., Cowx, I.G., Coelho, M.M., 1998. Diploids *v*triploids of *Rutilus alburnoides*: spatial segregation and morphological differences. J. Fish Biol. 52, 817–828.

Mayr, E., 1982. Processes of speciation in animals. In: Barigozzi, C. (Ed.), Mechanisms of Speciation. Alan R. Liss, New York, pp. 1–19.

Mesquita, N., Hänfling, B., Carvalho, G.R., Coelho, M.M., 2005. Phylogeography of the cyprinid *Squalius aradensis* and implications for conservation of the endemic freshwater fauna of southern Portugal. Mol. Ecol. 14, 1939–1954.

Moya-Palomares, M.E., 2002. Evolución sedimentológica y geomorfológica de las Vegas Bajas del Guadiana entre Mérida y Badajoz (España). Ph.D Thesis. Universidad Complutense de Madrid, Madrid, 297pp.

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

Pimentel, N.L., 1997. O terciário da bacia do Sado. Sedimentologia e análise tectono-sedimentar. Ph.D Thesis. University of Lisbon, Lisbon, 381pp.

Ribeiro, F., Beldade, R., Dix, M., Bochechas, J., 2007. Carta Piscícola Nacional Direcção Geral dos Recursos Florestais—Fluviatilis, Lda. World Wide Web electronic publication, www.fluviatilis.com, version 01/2007.

Ribeiro, O., Lautensach, H., Daveau, S., 1987. Geografia de Portugal—I. A posicão geográfica e o território. Edicões João Sá da Costa, Lisboa, 334pp.

Robalo, J.I., Sousa-Santos, C., Levy, A., Almada, V.C., 2006. Molecular insights on the taxonomic position of the paternal ancestor of the *Squalius alburnoides* hybridogenetic complex. Mol. Phylogenet. Evol. 39, 276–281.

Robalo, J.I., Almada, V.C., Levy, A., Doadrio, I., 2007. Re-examination and phylogeny of the genus Chondrostoma based on mitochondrial and nuclear data and the definition of 5 new genera. Mol. Phylogenet. Evol. 42, 362–372.

Rodríguez-Vidal, J., Cáceres, L., Ramirez, A., 1991. La red fluvial cuaternaria en el piedemonte de Sierra Morena occidental. Cuadernos de Investigación Geográfica 17, 37–46.

Rodríguez-Vidal, J., Cáceres, L., Ramirez, A., 1993. Modelo evolutivo de la red fluvial cuaternaria en el suroeste de la Península Ibérica. Actas da 3ª Reunião do Quaternário Ibérico, Coimbra, pp. 93–96.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, New York, 999 pp.

Sanjur, O.I., Carmona, J.A., Doadrio, I., 2003. Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. Mol. Phylogenet. Evol. 29, 20–30.

Sousa-Santos, C., Collares-Pereira, M.J., Almada, V.C., 2006a. Evidence of extensive mitochondrial introgression with nearly complete substitution of the typical *Squalius pyrenaicus*-like mtDNA of the *Squalius alburnoides* complex (Cyprinidae) in an independent Iberian drainage. J. Fish Biol. 68 (Supplement B), 292–301.

Sousa-Santos, C., Collares-Pereira, M.J., Almada, V.C., 2006b. Reproductive success of nuclear non-hybrid males of *Squalius alburnoides* hybridogenetic complex (Teleostei, Cyprinidae): an example of interplay between female choice and ecological pressures? Acta Ethologica. 9, 31–36.

Sousa-Santos, C., Collares-Pereira, M.J., Almada, V.C., 2006c. May a hybridogenetic complex regenerate the nuclear genome of both sexes of a missing ancestor? First evidence on the occurrence of a nuclear non-hybrid *Squalius alburnoides* (Cyprinidae) female based on DNA sequencing. J. Nat. Hist. 40, 1443–1448.

Sousa-Santos, C., Collares-Pereira, M.J., Almada, V.C., 2007. Fertile triploid males—an uncommon case among hybrid vertebrates. J. Exp. Zool. 307A, 220–225.

Sousa-Santos, C., Robalo, J., Collares-Pereira, M.J., Almada, V.C., 2005. Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic composition of hybrid organisms. DNA Seq. 16, 462–467.

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

Zardoya, R., Doadrio, I., 1998. Phylogenetic relationships of Iberian cyprinids: systematic and biogeographical implications. Proc. R. Soc. 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.