

# MtDNA markers reveal the existence of allopatric evolutionary lineages in the threatened lampreys *Lampetra fluviatilis* (L.) and *Lampetra planeri* (Bloch) in the Iberian glacial refugium

C. S. Mateus · P. R. Almeida ·  
B. R. Quintella · M. J. Alves

Received: 1 August 2010 / Accepted: 1 April 2011 / Published online: 30 April 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** The Iberian Peninsula has been identified as an important glacial refugium during the Pliocene and Pleistocene epochs for the genus *Lampetra*, providing intermittent refuge and postglacial opportunities for colonization and expansion. We used mitochondrial DNA markers to investigate the processes that have shaped present-day genetic constitution of the genus *Lampetra* within the Iberian Peninsula. We surveyed 1,173 bp of the cytochrome *b* gene and 829 bp of the genes ATPase subunits 6 and 8 in 233 individuals of *Lampetra fluviatilis* (L.) and *Lampetra planeri* (Bloch) from 21 localities

along their distribution range in the Iberian Peninsula. We identified four highly divergent allopatric evolutionary lineages that evolved by fragmentation during the Pliocene and Pleistocene likely driven by environmental factors, namely regional geomorphic events. The high level of genetic divergence between the four lineages suggests that sufficient time has apparently passed by to originate a complex of incipient or cryptic resident species and allows the definition of four evolutionary significant units (ESUs) for *L. planeri* and one ESU for *L. fluviatilis*. These findings have important consequences for the interpretation of refugia biological diversity and have major implications on the conservation of these threatened species.

**Electronic supplementary material** The online version of this article (doi:10.1007/s10592-011-0209-x) contains supplementary material, which is available to authorized users.

C. S. Mateus (✉) · P. R. Almeida · B. R. Quintella  
Centro de Oceanografia, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisbon, Portugal  
e-mail: csmateus@fc.ul.pt

C. S. Mateus · P. R. Almeida  
Departamento de Biologia, Escola de Ciência e Tecnologia,  
Universidade de Évora, Largo dos Colegiais,  
7000 Évora, Portugal

C. S. Mateus · M. J. Alves  
Museu Nacional de História Natural, Universidade de Lisboa,  
Rua da Escola Politécnica 58, 1250-102 Lisbon, Portugal

C. S. Mateus · M. J. Alves  
Centro de Biologia Ambiental, Faculdade de Ciências da  
Universidade de Lisboa, Campo Grande,  
1749-016 Lisbon, Portugal

B. R. Quintella  
Departamento de Biologia Animal, Faculdade de Ciências  
da Universidade de Lisboa, Campo Grande,  
1749-016 Lisbon, Portugal

**Keywords** Iberian Peninsula · Allopatric speciation · Species pairs · Cryptic species · Critically endangered · Conservation units

## Introduction

The population genetic structure of organisms is expected to reflect historical patterns of dispersal and isolation (Avise 2000; Wilson et al. 2004). Comparing the genetic diversity levels in marine, freshwater and anadromous fishes, there is evidence that freshwater fish tend to show higher population differentiation than marine fish, with the data from anadromous species indicating that this group occupies an intermediate position (Gyllensten 1985; Ward et al. 1994). These differences can be endorsed primarily to differences in average levels of gene flow, with gene flow in marine species thought to be less restricted than in freshwater species. On one hand, physical barriers to adult movement will be less pronounced in the marine

environment than in freshwater habitats and on the other hand it is possible that historical founder effects and drift, brought about by the effects of Pleistocene glaciations, have impinged to a greater degree of population differentiation on freshwater species than on marine species (Gyllensten 1985; Ward et al. 1994).

The lampreys constitute a good model to investigate how intrinsic factors (life-history type and dispersal capability) interact with historical extrinsic factors to shape genetic structuring, as many genera present anadromous species and closely related freshwater residents. It is generally assumed that the freshwater lampreys have evolved from the migratory parasitic form and become non-parasitic (reviewed in Docker 2009).

The Iberian Peninsula seems to have played a major role as a glacial refugium for the European river lamprey *Lampetra fluviatilis* (L.) and the European brook lamprey *L. planeri* (Bloch), as Iberian populations reveal higher genetic diversity when compared to those of central and northern Europe (Espanhol et al. 2007). *L. fluviatilis* and *L. planeri* are 'paired species', i.e. the larvae are morphologically similar but the adults adopt different life history types (Zanandrea 1959). The parasitic and anadromous *L. fluviatilis* is most easily distinguished in its sexually mature stage from the non-parasitic *L. planeri* by its generally greater body size (Hardisty 1986a). Phylogeographical analysis revealed that the two taxa are not reciprocally monophyletic, suggesting that loss of the migratory ability may have occurred multiple times (Espanhol et al. 2007). Many Iberian populations of *L. planeri* are apparently composed of private mitochondrial DNA (mtDNA) haplotypes, suggesting some time of independent evolutionary history for these populations (Espanhol et al. 2007; Pereira et al. 2010). As suggested for other brook lamprey species, these isolated populations may represent a complex of incipient or cryptic resident species, despite their highly conserved body form (cf. Docker 2009). This finding is in agreement with the realization that many species display a strong population substructure within glacial refugia (reviewed in Gómez and Lunt 2006) and indicates that phylogeography of *Lampetra* within the Iberian glacial refugium warrants further investigation.

Here, sequence variation in two mtDNA genes is used to evaluate the relationship between *L. planeri* and *L. fluviatilis* within Iberian river basins and the existence of divergent allopatric evolutionary lineages. We also investigate the processes that have shaped genetic structure in the genus *Lampetra*. Our results from population genetic, phylogenetic and phylogeographical analyses indicate a complex and dynamic evolutionary history of expansion and fragmentation in multiple, independent lineages, with important conservation implications.

## Materials and methods

### Sampling, extraction, amplification and DNA sequencing

In total, we collected 233 individuals of *Lampetra*, comprising 66 adults of *L. planeri*, 16 adults of *L. fluviatilis* and 151 ammocoetes of unknown specific status in 21 sites throughout the distributional range of the species in the Iberian Peninsula, covering all the major basins (Fig. 1; Table 1). Tissue samples were deposited in the zoological collections 'Museu Bocage' (MB85) of Museu Nacional de História Natural, Portugal (Table 1).

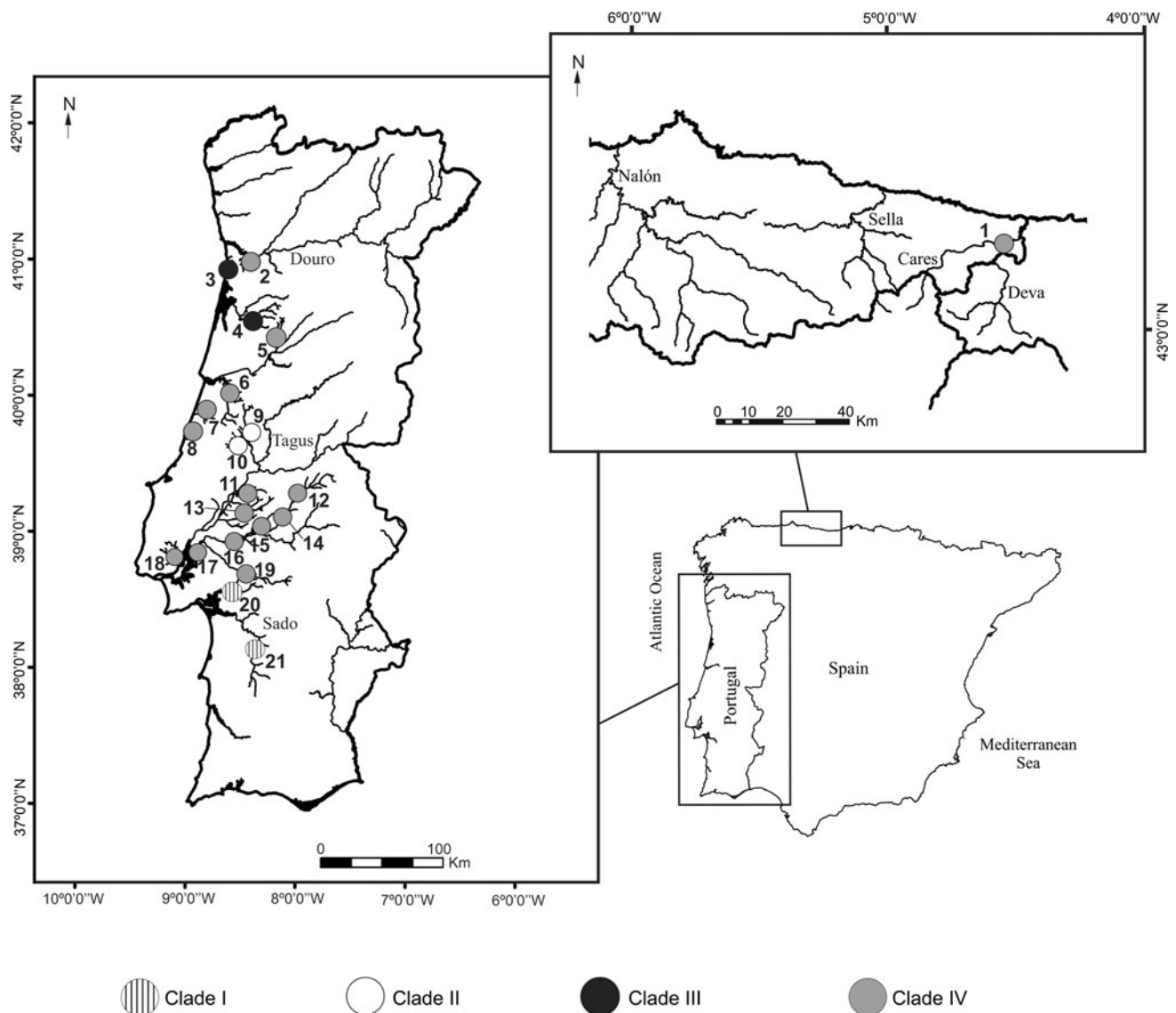
We extracted the total genomic DNA from muscle tissue preserved in alcohol pro-analyses by the conventional SDS-proteinase K/phenol–chloroform protocol. We quantified DNA samples using NanoDrop ND-1000 Spectrophotometer and established standard working stocks of 40 ng  $\mu\text{l}^{-1}$  in sterile water for all individuals.

We amplified the mitochondrial genes ATPase subunits 6 and 8 (ATPase 6/8) and cytochrome *b* (*cyt b*) by polymerase chain reaction (PCR) in a thermocycler Biometra Tgradient. A total of 1,173 bp of the *cyt b* gene were amplified using the primers LampLA and LampPRO and the internal primers LampLB and LampCB2-H (Espanhol et al. 2007). The primers used for the amplification of the 829 bp of the genes ATPase subunits 6 and 8 were ATPfor and ATPrev (Espanhol et al. 2007). PCR reactions were performed in a final volume of 25  $\mu\text{l}$ , with 1  $\mu\text{l}$  of total genomic DNA, 1.5 mM  $\text{MgCl}_2$ , 0.2 mM dNTP mix, 0.5  $\mu\text{M}$  of each primer, 1 U of *Taq* DNA polymerase (Fermentas) and 1 $\times$  of the reaction buffer supplied. PCR conditions were as follows: an initial denaturation step of 94°C for 3 min followed by 30 cycles consisting of denaturation at 94°C for 1 min, primer annealing at 60°C for 1 min, extension at 72°C for 2 min, and a final extension step of 72°C for 2 min. The resulting PCR products were purified using the ExoSAP kit (Fermentas) and sequenced using an ABI PRISM 3730 DNA Analyser at MacroGen ([www.macrogen.com](http://www.macrogen.com)).

### Data analysis

We aligned and edited the DNA sequences manually, using *Sequencher* V4.8 (Gene Codes Corporation, Ann Arbor, USA). Sequences from *Petromyzon marinus* (L.) from the EMBL database (U11880) and *Eudontomyzon mariae* (Berg) from the EMBL database (AM051061) were used as outgroups.

For either gene, we performed a  $\chi^2$  test of homogeneity to test the assumption of base-compositional homogeneity. Prior to combining the two genes into single analyses, we implemented the incongruence length difference test (ILD)



**Fig. 1** Geographical distribution of the 21 sampling sites (*dots*) assayed in this study and of the four clades suggested by the phylogenetic analysis. Sites are numbered as in Table 1. A detailed distribution of all haplotypes is described in Table 2

to access the significance of incongruence between the two data sets. Both analyses were implemented in PAUP\* version 4.0b10 (Swofford 2002). All further analyses were performed in the concatenated alignment of both genes.

Levels of gene diversity were described as haplotype (*h*) and nucleotide ( $\pi$ ) diversities. The definition of the haplotypes and the estimation of the levels of gene diversity were attained with the software package *DnaSP* version 4.50 (Rozas et al. 2003).

Haplotypes were connected on a network obtained using the 95% parsimony criterion implemented in the program TCS version 1.21 (Clement et al. 2000). We performed the phylogenetic analysis by three methods, maximum parsimony (MP), neighbour-joining (NJ) and maximum likelihood (ML), all implemented by using PAUP\*. We used the

Modeltest version 3.8 software (Posada and Crandall 1998) associated with PAUP\* to select the most appropriate evolutionary model of nucleotide substitution and its parameters, according to the Bayesian information criterion (BIC), following Posada and Buckley (2004). We calculated the maximum-likelihood (ML) and neighbour-joining (NJ) phylogenetic trees according to the selected model, HKY + *G* substitution model (gamma distribution shape parameter = 0.2102; base frequencies: *A* = 0.3029, *C* = 0.2424, *G* = 0.1180, *T* = 0.3366; transition/transversion ratio = 5.0898). For the three methods, the optimal trees were found by a heuristic search with tree-bisection-reconnection (TBR) as the branch-swapping algorithm. Initial trees were obtained via stepwise addition with 100 replicates of random addition sequence and gaps were

**Table 1** Collection data for the individuals analyzed in this study

Locality no.	Drainage	River/Tributary	<i>N</i>	Species	Voucher no.
1	Deva*	Deva	8	Unknown	(MB85) 8698-8700; 8716-8720
2	Douro	Inha	10	Unknown	(MB85) 5673, 5674; 9630-9637
3	Esmoriz	Esmoriz	15	8 Unknown; 7 <i>L. plan</i>	(MB85) 5497; 5502-5505; 5507-5510; 5544-5546; 9424-9426
4	Vouga	Águeda	10	6 Unknown; 4 <i>L. plan</i>	(MB85) 5649-5652; 5655; 5657; 9457-9460
5	Mondego	Criz	9	<i>L. plan</i>	(MB85) 5619-5622; 5624-5628
6	Mondego	Anços	10	<i>L. plan</i>	(MB85) 5606-5610; 5612-5614; 5617; 5618
7	Lis	Lis	14	9 unknown; 5 <i>L. plan</i>	(MB85) 5639-5645; 5647; 5648; 9359-9363
8	Ribeiras do Oeste	Ribeira de S. Pedro	13	8 unknown; 5 <i>L. plan</i>	(MB85) 5629; 5631-5636; 5638; 9392-9396
9	Tagus	Nabão	9	Unknown	(MB85) 5596-5603; 5605
10	Tagus	Ribeira do Olival	13	8 Unknown; 5 <i>L. plan</i>	(MB85) 5587–5592; 5594; 5595; 9461–9465
11	Tagus	Ulme	32	Unknown	(MB85) 5466-5486; 5576-5586
12	Tagus	Longomel	9	<i>L. plan</i>	(MB85) 5566-5571; 5573-5575
13	Tagus	Muge	7	Unknown	(MB85) 5556-5559; 5562; 5564; 5565
14	Tagus	Sôr	6	Unknown	(MB85) 5972-5977
15	Tagus	Erra	15	8 Unknown; 2 <i>L. fluv</i> ; 5 <i>L. plan</i>	(MB85) 5487-5495; 5498; 5499; 5547; 9515-9517
16	Tagus	Sorraia	1	<i>L. plan</i>	(MB85) 5501
17	Tagus	Ponta de Erva	13	<i>L. fluv</i>	(MB85) 5669-5671; 6175-6184
18	Tagus	Tagus	1	<i>L. fluv</i>	(MB85) 5971
19	Tagus	Canha	11	10 Unknown; 1 <i>L. plan</i>	(MB85) 5496; 5548-5552; 5554; 5555; 5968-5970
20	Sado	Marateca	19	14 Unknown; 5 <i>L. plan</i>	(MB85) 5452-5465; 9530-9534
21	Sado	Sado	8	Unknown	(MB85) 5659-5665; 5667

Sampled localities are presented from north to south and locality numbers correspond to locations as in Fig. 1. Sample sizes (*N*) and specific status are also presented. Voucher numbers correspond to zoological collections ‘Museu Bocage’ (MB85) of Museu Nacional de História Natural, Portugal

*L. plan*, *L. planeri*; *L. fluv*, *L. fluviatilis*

\* Spanish river basin

treated as missing data. Bootstrap proportions (Felsenstein 1985) were obtained to assess node robustness, using PAUP\*. In MP and NJ analyses, 1,000 heuristic pseudoreplicates were generated, each consisting of 100 heuristic TBR searches of random addition sequence. In ML analysis, 500 heuristic pseudoreplicates were generated with TBR searches of as-is addition sequence.

We performed Bayesian analyses using MrBayes version 3.1.2 (Ronquist and Huelsenbeck 2003) to produce a population of trees and parameter values used then to estimate a 50% majority-rule consensus tree. We estimated the probabilities of the phylogenetic trees by a Metropolis-coupled, Markov chain Monte Carlo sampling algorithm (MCMCMC). For each analysis, a total of  $2 \times 10^6$  samples were taken (2 separate concurrent runs), with successive samples separated by 100 generations. Model selection was carried out separately for each mtDNA data set with MrModeltest version 2.2 (Nylander 2004), allowing different values for each parameter for each partition.

We also performed an analysis of molecular variance, AMOVA (Excoffier et al. 1992). This analysis accomplishes three components of genetic variation: among groups ( $F_{CT}$ ), among populations within each group ( $F_{SC}$ ), and within populations ( $F_{ST}$ ). Molecular variance was first partitioned into two hierarchical levels, where individuals were assembled into two different groups, reflecting their specific taxonomic status (AMOVA I). Consequently, the 151 ammocoetes of unknown specific status were not included in this analysis. A second analysis was performed in which individuals from the same locality were treated as individual populations to test for overall genetic subdivision (AMOVA II), regardless their specific taxonomic status. Finally, molecular variance was partitioned into two hierarchical levels, where localities were assembled into different groups reflecting the results from the phylogenetic analyses (AMOVA III), regardless their specific taxonomic status. The significance of the observed variances for each hierarchical comparison was tested by 10,000 permutations.

We quantified genetic differentiation among populations by computing pairwise  $F_{ST}$  estimates calculated using conventional  $F$ -statistics based on mtDNA haplotype frequencies and the pairwise difference distance method. Significance of pairwise population comparisons was assessed by 1,000 permutations. All analyses were conducted using Arlequin version 3.11 (Excoffier et al. 2005).

We performed Fu's  $F_s$  (Fu 1997) and Tajima's  $D$  (Tajima 1989) statistics using *DnaSP* version 4.50. Fu's  $F_s$  tends to be negative under an excess of recent mutations and a significant negative value will be taken as evidence of population growth and/or selection. A positive value of  $F_s$  is evidence for a deficiency of alleles, as would be expected from a recent population bottleneck or from overdominant selection. Tajima's  $D$  statistic is influenced by changes over time in the size of populations, population structure and the action of natural selection. The test uses the nucleotide diversity and the number of segregating sites observed in a sample of DNA sequences to make two estimates of the amount of variation. In a constant-size neutral equilibrium population, the expectation of Tajima's  $D$  is nearly zero because the expectations of both estimates are the same. When some kind of balancing selection is acting, Tajima's  $D$  tends to be positive. On the other hand, purifying selection can generate negative values of Tajima's  $D$ .

## Results

The incongruence length difference test showed that the two genes were not incongruent (ILD  $P = 0.45$ ). For both genes, the null hypothesis of homogeneity in base composition across sequences was not rejected by the  $\chi^2$  test (ATPase:  $\chi^2 = 4.06$ ,  $d.f. = 165$ ,  $P = 1.00$ ; cyt  $b$ :  $\chi^2 = 1.31$ ,  $d.f. = 165$ ,  $P = 1.00$ ). All further results refer to concatenated alignments of both genes.

The 233 samples were grouped in 56 composite haplotypes. Haplotype codes follow Espanhol et al. (2007). No haplotype was observed exclusively in *L. fluviatilis*, while nine were detected only in *L. planeri* (Table 2). Considering the geographical distribution of the haplotypes, 47 were found at single localities, nine (H6, H31, H37, H38, H47, H60, H62, H65, H66) were observed in two or more localities and none was detected at all localities (Table 2). Haplotype H31 was the most frequently observed haplotype ( $N = 44$ ) and was found in seven localities of the Tagus, Mondego and Douro river basins. The otherwise most frequent haplotype H6 ( $N = 31$ ) was found in six localities of the Tagus, Lis and Mondego river basins. The 56 haplotypes were distinguishable by 82 polymorphic sites ( $S$ ), 42 in cyt  $b$

and 40 in ATPase 6/8 (see Online Resource, ESM 1). Overall haplotype diversity ( $h$ ) was 0.929 and nucleotide diversity ( $\pi$ ) was 0.00633. Tajima's  $D$  was not significantly different from zero ( $D = -0.30872$ ;  $P > 0.10$ ) but Fu's  $F_s$  produced a significant negative value ( $F_s = -50.260$ ).

The tree topologies obtained from analysis of the mtDNA data by the three methods MP, NJ and ML were highly concordant and revealed four clades (I–IV), which were well supported by bootstrap and Bayesian credibility values. Clade IV presents further well supported subdivision (subclades IV-A to IV-C) (Fig. 2). As previously observed in Espanhol et al. (2007), clades recovered are not species specific: clades I, II and III are composed of adults of *L. planeri* and ammocoetes of unknown specific status, while clade IV includes ammocoetes and both migratory and resident adults. Clades do not apparently overlap geographically: clade I includes the samples from Sado basin, represented by 14 haplotypes; clade II includes the individuals from River Nabão and its tributary Ribeira do Olival; clade III includes the populations from Esmoriz and Águeda rivers; and clade IV shows a wider distribution from Tagus river basin to the northern Spanish River Deva. Subclade IV-A groups individuals from the Tagus, Ribeiras do Oeste, Mondego, and Douro river basins; subclade IV-B includes individuals from the Mondego and Tagus river basins; and subclade IV-C groups individuals from the Tagus, Lis, Mondego, and Deva river basins (Fig. 2; Table 2). The parsimony network of haplotypes (Fig. 3) revealed that clade IV has a double star-like structure, with two dominant haplotypes H31 and H6 in the centre of subclades IV-A and IV-C, respectively.

In AMOVA I, where populations were grouped according to their specific status, variance between species was virtually null, while most of the variance was distributed among localities within species (79.05%,  $P < 0.001$ ; Table 3). This result supports the non-existence of specific clades for *L. planeri* and *L. fluviatilis* as revealed in the phylogenetic analysis and the assumption that the two taxa do not form reciprocal monophyletic groups. In AMOVA II, where individuals from the same locality, regardless their specific taxonomic status, were grouped in the same population, statistically significant amounts of the molecular variance (76.52%,  $P < 0.001$ ) was attributed to differences among localities. When molecular variance was partitioned into two hierarchical levels reflecting the results of the phylogenetic analysis (AMOVA III), most of the variance (71.41%,  $P < 0.001$ ) was distributed among groups, while variance among localities within each group accounted for 12.92% ( $P < 0.001$ ) and variance within localities for 15.67% ( $P < 0.001$ ) (Table 3).

**Table 2** Distribution of haplotypes across samples, according to their specific taxonomic status and across the 21 sampled localities, regardless the specific taxonomic status of the samples

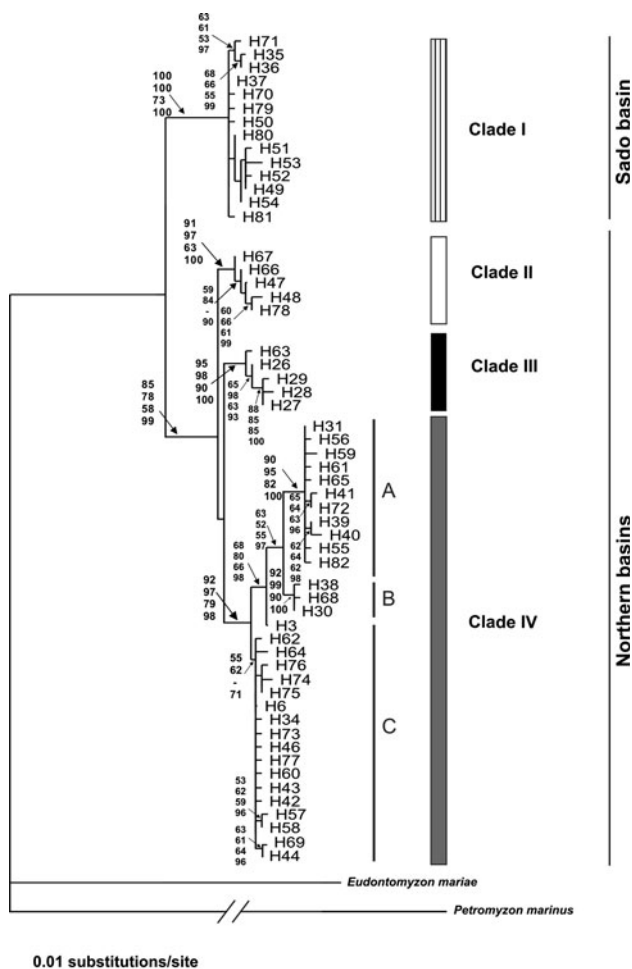
Haplotype	Species		Locality																					NS <sub>H</sub>
	<i>L. fluv</i>	<i>L. Am plan</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
H3	1			1																				1
H6	7	13	11	8	10													1	1	6	5			31
H26		7	7	14																				14
H27		4	4	8																				8
H28		1	1	1																				1
H29		1	1	1																				1
H30		2			2																			2
H31		9	35		1					11	4	7	4	8										44
H34		1	1																	2				2
H35			1																		1			1
H36			2																			2		2
H37			3																		1	2		3
H38		3	1		3					1														4
H39		5	7							12														12
H40			1							1														1
H41		2																						2
H42		1									2													2
H43			1		1																			1
H44		1																						1
H46																								2
H47		4																			1			1
H48										7	10													17
H49										1														1
H50																						6		6
H51																						1		1
H52																						1		1
H53																						2		2
H54																						2		2
H55																								15
H56											15													1
H57											1													1
H58											2													2
H59		1									2													2

**Table 2** continued

Haplotype	Species		Locality																			NS <sub>H</sub>			
	<i>L. fluv</i>	<i>L. plan</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		20	21	
H60	5	2															2	4		1			7		
H61		1															1						1		
H62	4	2															1	3	1	1			6		
H63		1			1																		1		
H64		1																			1		1		
H65		3											3			1							4		
H66		2							1	1													2		
H67		1								1													1		
H68		3				3																	3		
H69		1					1																1		
H70		1																					1		
H71		2																					2		
H72		1														1							1		
H73		1														1							1		
H74		2																					2		
H75		3																					3		
H76		1																					1		
H77		1																					1		
H78		1																					1		
H79		2																				2	2		
H80		2																				2	2		
H81		1																				1	1		
H82		1														1							1		
NS	16	66	151	8	10	15	10	9	10	14	13	9	13	32	9	7	6	15	1	13	1	11	19	8	233
NH	3	20	47	5	2	2	3	2	5	4	2	3	4	6	3	1	3	6	1	3	1	6	10	5	
NH <sub>P</sub>	0	9	34	5	1	2	3	1	3	3	2	1	2	4	1	0	1	2	0	0	0	3	9	4	

For each species/locality, number of samples (NS), number of haplotypes (NH) and number of private haplotypes (NH<sub>P</sub>) are presented. Number of samples in each haplotype is also indicated (NS<sub>H</sub>)

*L. fluv*, *L. fluvialitilis*; *L. plan*, *L. planeri*; Am, ammocoete



**Fig. 2** Maximum-likelihood phylogenetic tree of 2002 bp of the ATPase 6/8 and cyt *b* mtDNA genes. Numbers in terminal nodes refer to the haplotype numbers as in Table 2. Main geographical division is indicated as two lineages, Sado basin and Northern basins and inside each division the respective clades are presented: clades I to IV. Clade IV is divided in three subclades: A, B and C. Numbers are the bootstrap support values equal to or higher than 50% obtained from maximum parsimony, neighbour-joining, and maximum likelihood and the Bayesian credibility value, respectively

Pairwise comparisons among geographical samples revealed that for 72% of the pairwise  $F_{ST}$  values there are significant differences ( $P < 0.05$ ) in allele frequencies (see Online Resource, ESM 2). These results are congruent with the results of the AMOVA analysis.

Haplotype diversity in the four clades ranged from 0.407 in clade II (River Nabão and Ribeira do Olival) to 0.932 in clade I (Sado basin; Table 4). Nucleotide diversity was lower, ranging from 0.00035 in clade II to 0.00254 in clade IV (Table 4).

Nucleotide sequences are available at the EMBL database under the accession numbers AJ937923, AJ937926, AJ937946–AJ937951, AJ937954–AJ937957, FN641825–FN641863 and FR669668–FR669672.

## Discussion

### Genetic diversity and population structuring

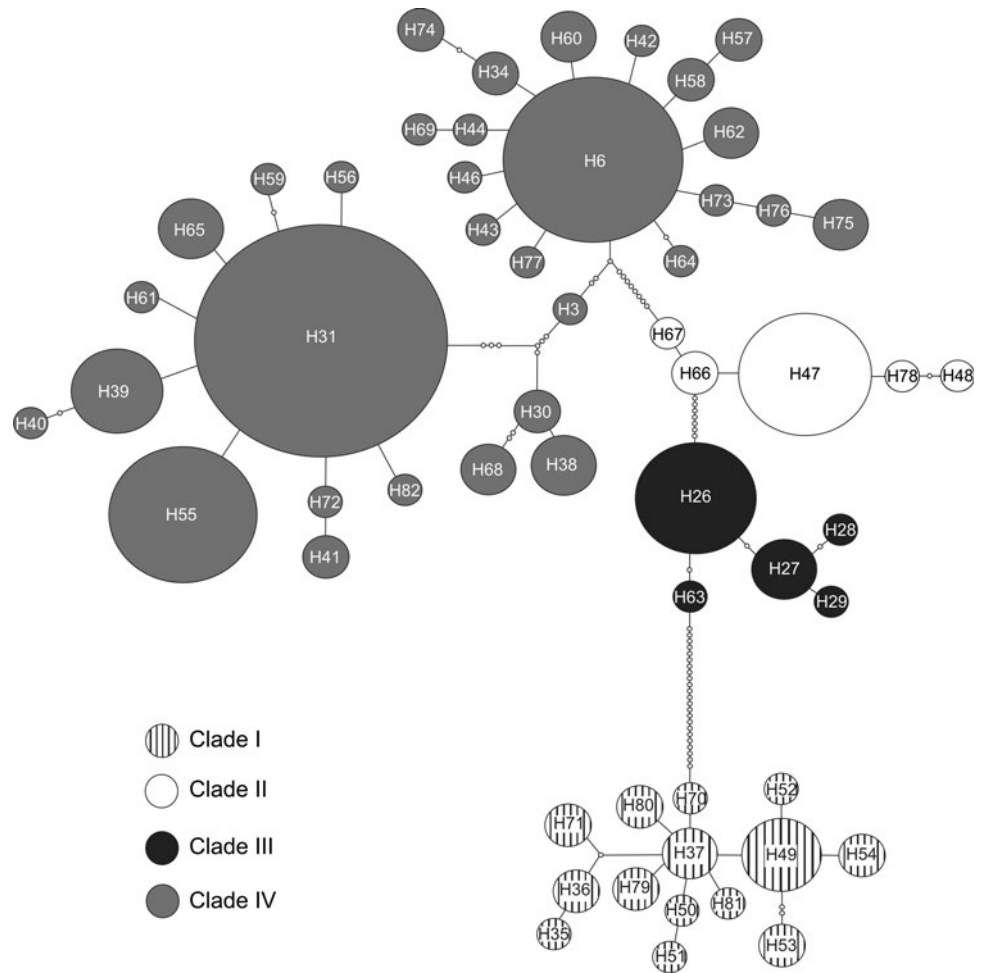
This study revealed a significant level of phylogenetic structuring in the mitochondrial cyt *b* and ATPase genes of the genus *Lampetra* in the Iberian Peninsula. Clades recovered are not species specific but instead follow a geographic pattern, confirming the results of Espanhol et al. (2007) who concluded that *L. planeri* and *L. fluviatilis* do not form reciprocal monophyletic groups. The present results suggest the existence of four highly divergent clades with apparent allopatric distribution. Clades I–III apparently only include adults of the resident form and they have restricted distribution, each being confined to a few localities of the same or close small river basins. Clade IV shows a wider distribution, extending from throughout Tagus river basin to Deva river basin, in northern Spain, and it includes both adults of the migratory and resident forms. Clades II and IV both occur in the Tagus river basin, but so far they have been found in different rivers; nevertheless, they may be sympatric at some localities that remained undetected in the present sampling. The AMOVA analysis showed concordant results, revealing substantial levels of genetic fragmentation both between localities and between groups of localities, independently of the taxonomic specific status (Table 3).

The present results suggest the past occurrence of repeated landlocking of diadromous forms in freshwater isolates, leading to the lost of the migratory ability. The new founded resident populations would have become isolated from other resident and migratory populations (discussed below), with very little or no subsequent gene exchange, allowing for the accumulation over time of new local mutations. Reduced gene flow combined with small population sizes in small streams promoted allopatric divergence, leading to monophyly in most populations. The most ancient long-term barriers to the gene flow involved the populations from Sado basin and the northern basins. Before the establishment of the exorheic network in the Plio-Pleistocene, most river systems drained to a large number of inland lakes (Calvo et al. 1993). Since the uplifting of the Arrábida Chain in the Late Miocene (Antunes et al. 1995; Choffat 1908) and probably the posterior establishment of the Cascais and Setúbal canyons (Alves et al. 2000; Coppier and Mougnot 1982) Tagus and Sado basins remained independent basins. The differentiation of a Sado phylogroup has been also observed in other freshwater fishes, namely in *Chondrostoma lusitanicum* (Mesquita et al. 2001) and *Squalius pyrenaicus* (Sousa-Santos et al. 2007).

Clade IV is widely distributed in the Tagus basin, with the exception of populations from River Nabão, which are



**Fig. 3** Haplotype network inferred by the criterion of parsimony with TCS 1.21 representing the 56 haplotypes obtained in this study. Haplotype numbers are identified as in Table 2. The cladogram was estimated under the 95% statistical limits of parsimony. Circle size represents haplotype frequency. Each line in the network represents a single mutational change and empty circles indicate hypothetical, missing haplotypes



**Table 3** Analysis of molecular variance (AMOVA)

Source of variation	<i>d.f.</i>	Sum of squares	Variance components	Percentage of variation	<i>P</i>	Fixation indices
<b>AMOVA I</b>						
Among species	1	77.660	-0.40648	-3.79	>0.05	$F_{CT}$ : -0.03795
Among populations within species	13	578.443	8.46731	79.05	<0.001	$F_{SC}$ : 0.76163
Within populations	67	177.555	2.65007	24.74	<0.001	$F_{ST}$ : 0.75258
Total	81	833.659	10.71090			
<b>AMOVA II</b>						
Among populations	20	1961.794	8.73964	76.52		
Within populations	212	568.450	2.68137	23.48	<0.001	$F_{ST}$ : 0.76522
Total	232	2530.245	11.42101			
<b>AMOVA III</b>						
Among groups	3	1523.157	12.22215	71.41	<0.001	$F_{CT}$ : 0.71410
Among populations within groups	17	438.637	2.21197	12.92	<0.001	$F_{SC}$ : 0.45204
Within populations	212	568.450	2.68137	15.67	<0.001	$F_{ST}$ : 0.84334
Total	232	2530.245	17.11548			

In AMOVA I individuals of *Lampetra fluviatilis* and *L. planeri* were assembled into two different groups, in AMOVA II individuals from the same locality were grouped in the same population, independently of the taxonomic status, and in AMOVA III populations were grouped into the four clades suggested by the phylogenetic analyses, independently of the taxonomic status

apparently monophyletic for the clade II (discussed above). The Nabão river is the only tributary of the right bank of Tagus basin where *L. planeri* is known to occur. The strong divergence of these populations is probably related with events from the Late Miocene that extended through the Pliocene. Different tectonic movements (subsidence and uplift) of both banks produced distinct systems with own characteristics. The subsidence of the right bank in the beginning of the Superior Miocene promoted the existence of lower areas, protected from sea level changes, with retention of small lakes. In this system, drainage was predominantly endorheic, retaining the water with no outflow to other bodies of water. In the left bank, however, the fluvial system was opened, with exorheic drainage, disabling the existence of lakes (Barbosa 1995). Populations from the right bank may have remained isolated in the endorheic lakes through this period and would have time to diverge from the other populations. Tectonic movements between banks remained through the Pliocene, with inversion of the tectonic subsidence to the left bank, as marked by the accumulation of Pliocene sediments (Barbosa 1995). The dissimilarity of ecological conditions between the tributaries of both banks may have promoted the isolation and differentiation of populations when the Tagus river basin gained its present configuration.

The differentiation of the populations from the Vouga river basin and its close neighbour the Esmoriz river is somehow surprising, considering the paleogeological evidences (Rodrigues and Dias 1989) and previous phylogeographic studies with other freshwater fishes (Aboim et al. 2009; Sousa-Santos et al. 2007), which suggest recent connections between these basins and the Douro and the Mondego drainages. In fact, in the Pleistocene, connections between the Mondego and Douro and the Mondego and Vouga drainages were still possible (Rodrigues and Dias 1989), allowing the dispersal of freshwater fishes between these basins (Sousa-Santos et al. 2007). This high differentiation, together with the high structuring within the Tagus river basin (discussed above), suggests limited dispersal capabilities in continuous freshwater systems (further discussed below).

The star-like structure and the geographical distribution of haplotypes within clade IV are consistent with a scenario of dispersal and demographic expansion. The Fu's  $F_s$  significant negative value for clade IV indicates that this expansion was recent (Table 4). This phylogenetic lineage is apparently the only one that still includes the migratory form, *L. fluviatilis*, and postglacial sea dispersal by the anadromous form, followed by demographic expansion and establishment of resident populations, has been postulated to explain the widespread distribution of clade IV in central and northern Europe (Espanhol et al. 2007). Nevertheless, structuring within clade IV suggests that movements of the migratory form were probably restricted during the glacial times, favouring population differentiation. The most ancestral haplotype of subclade IV-A (H31) is most common in the left tributaries of the Tagus basin, suggesting that this subclade has differentiated most probably in this basin, having attained its current distribution through long distance colonization. In fact, this subclade is present in the Douro basin, in the Mondego basin and also in the Ribeira de S. Pedro. Subclade IV-B seems to have differentiated in the left tributaries of the Mondego basin, and its restricted distribution suggests reduced dispersal. In fact, only one haplotype from this clade has been observed outside the Mondego basin, in the River Ulme (Tagus basin). Finally, the most ancestral haplotype of subclade IV-C (H6) is most common in the right tributary of the Mondego basin but also in some left tributaries of the Tagus basin (Erra, Sorraia and Canha), making it difficult to infer the geographic origin of this subclade. This subclade is also present in the neighbour basin Lis, and also in the distant Spanish Deva river basin, suggesting long distance colonization. The observation that subclade IV-C is apparently absent from the right tributaries of the Tagus basin reinforces this hypothesis, as it suggests that gene flow between populations from the Mondego and Tagus basins was established not by inland connections through river captions, as it has been suggested for other freshwater fishes (Sousa-Santos et al. 2007), but instead by sea.

The high levels of genetic diversity and population structuring attained in this study for the Iberian Peninsula

**Table 4** Summary of genetic variability, Tajima's  $D$  and Fu's  $F_s$  neutrality tests in the four clades based on *cyt b* and ATPase 6/8 mtDNA genes

Clade	$N$	$NH$	$h$	$\pi$	$S$	$k$	Tajima's $D$	Fu's $F_s$
I	27	14	0.932	0.00124	14	2.934	-1.35071 NS	-14.785***
II	22	5	0.407	0.00035	5	2.400	0.00000 NS	-2.680 NS
III	25	5	0.603	0.00071	7	3.200	-0.33192 NS	-2.116 NS
IV	159	32	0.869	0.00254	37	6.569	-1.10353 NS	-35.376***

$N$  sample size,  $NH$  number of haplotypes,  $h$  haplotype diversity,  $\pi$  nucleotide diversity,  $S$  polymorphic sites,  $k$  average number of pairwise nucleotide differences

\*\*\* Significant at the 0.1% level; NS, not significant ( $P > 0.05$ )

can be explained by the persistence of multiple glacial refugia. Our findings are in agreement with a number of other phylogeographical studies (e.g., Alexandrino et al. 2002; Gante et al. 2009; Martínez-Solano et al. 2006; Paulo et al. 2001; Ribera and Vogler 2004), which provided evidence of considerable genetic divergence within the peninsula and suggested that a strong fragmented distribution may be considered a set of interglacial refugia. Haplotype diversity was higher in clades I and IV, which also showed the highest values of nucleotide diversity (Table 4). The high levels of genetic diversity observed in the Sado clade are somehow surprising considering its present restricted distribution, and suggest that environmental conditions in this basin may have allowed the stability of population through time, despite the climatic crisis of Upper Pliocene which was responsible for the disorganization of the Sado drainage network (Pimentel 1997), and have apparently caused declines in other freshwater fishes (Sousa-Santos et al. 2007).

#### On the dispersal ability of *Lampetra fluviatilis* and *L. planeri*

The Tagus basin is the only river basin where *L. fluviatilis* has been recorded in the Iberian Peninsula (Almaça and Collares-Pereira 1988; Doadrio 2001), and the sampled individuals all belong to subclade IV-C. This subclade is present in the Tagus, in the Lis, in the Mondego and in the Deva river basins, but is apparently absent from all other basins, suggesting that the migratory form stopped at some point visiting these basins to spawn. Even between the former river basins, gene flow should have become restricted as they all show unique derived haplotypes, in particular River Deva, where all five haplotypes are exclusive of that population (Table 2). Also, the majority of the population pairwise  $F_{ST}$  between these localities revealed significant differences (see Online Resource, ESM 2). Two hypotheses may account for this scenario: the development of inappropriate local conditions to the migratory form; and/or the reduction of the dispersal ability of *L. fluviatilis*. The absence of the other migratory lamprey, *Petromyzon marinus* in some of these basins (Cabral et al. 2005) provide some support to the first hypothesis. As for the second hypothesis, there are evidences that lampreys are largely affected by temperature (Hardisty and Potter 1971; Potter 1980), which may have caused populations at lower latitudes to abandon anadromy while temperature raised during the interglacials (Espanhol et al. 2007). In particular, the persistence of *L. fluviatilis* in the Tagus basin may have been possible due to the size of its estuary (c. 300 km<sup>2</sup>), allowing individuals to remain in the estuary during the adult stage feeding on estuarine species, a known behaviour for this species (Hardisty 1986b).

The presence of private haplotypes in most *L. planeri* populations (Table 2) and the fact that the majority of the population pairwise  $F_{ST}$  revealed significant differences (see Online Resource, ESM 2) suggest that the resident form presents very low dispersal ability within river basins. This pattern is particularly evident for the two populations sampled in the Sado basin (localities 20 and 21), which share only one haplotype among the 14 identified and present a pairwise  $F_{ST}$  value significant at the 0.1% level; or for the population from Ulme river in the Tagus basin (locality 11), whose the most common haplotype (H55) is apparently absent elsewhere in the basin. In fact, the spawning migrations of brook lampreys are known to be limited, since the spawning grounds are usually located only a short distance upstream from the silt beds inhabited by the ammocoetes and transforming stages (Hardisty 1986b). Furthermore, transformed brook lampreys are unable to feed, and the limited energy supply stored in their tissues might prevent efficient long-distance journeys during the few months of life as an adult (Schreiber and Engelhorn 1998).

#### Existence of cryptic species and implications for conservation

Genetics is an important focus of conservation biology as measuring genetic variation and interpreting these data in a phylogeographic and population genetics context enables us to understand the evolutionary context of species and the development of improved management strategies (Hurt and Hedrick 2004). The assessment of biodiversity within and among populations is central to identifying and prioritizing areas for monitoring, management and protection and the main goal of management should be to maintain levels of gene flow and maximum gene diversity, as inferred from molecular data (Crandall et al. 2000; Moritz and Faith 1998). Particular emphasis should be placed on those populations with highly diverged haplotypes and unique environmental traits (Hurt and Hedrick 2004). In 1986, Oliver Ryder referred to the evolutionarily significant unit (ESU) as a population unit that merits separate management and has a high priority for conservation. Moritz (1994) suggests the distinction between two types of conservation units, the ESUs which are concerned with historical population structure, mtDNA phylogeny and long-term conservation needs, and the management units (MUs) which address current population structure, allele frequencies and short-term management issues.

Conservation units have already been proposed for *L. planeri* (Pereira et al. 2010), but the low overall level of divergence and the low phylogenetic resolution observed in that study, due to the use of a single marker, suggest that a more indepth evaluation is needed. The present study revealed high levels of mtDNA divergence and clear

phylogeographical patterns of genetic structuring. The high genetic diversity attained for the Iberian glacial refugia is even more obvious when compared with the distribution of the haplotypes of samples from across Europe (see Espanhol et al. 2007 and Online Resource, ESM 3).

*L. planeri* and *L. fluviatilis* remain widely distributed across Europe, and, in terms of the current conservation status, they are globally considered *Least Concern* according to the IUCN Red List of Threatened Species due to a markedly recover following earlier pollution problem in central and western Europe (Freyhof and Kottelat 2008a; Freyhof and Kottelat 2008b). Nevertheless, both species are considered threatened in the Iberian Peninsula. In Portugal they are currently included in the *Critically Endangered* category of the red list of endangered species (Cabral et al. 2005). According to the red list of continental fish in Spain *L. fluviatilis* is considered *Regionally Extinct* and *L. planeri* *Critically Endangered* (Doadrio 2001).

The highly divergent clades recovered for *L. planeri* within the Iberian Peninsula are evidence for a long history of local independent evolution, suggesting that they should be considered significant for conservation. These Iberian populations have higher levels of divergence than populations from across Europe, which haplotypes are embedded in Clade IV, the widest distributed clade (see Online Resource, ESM 3). Accordingly, we suggest the definition of four evolutionarily significant units (ESUs) for *L. planeri* in the Iberian Peninsula, namely populations from clades I, II, III and IV. Clades I, II and III are exclusive to the Iberian Peninsula (Sado basin, River Nabão and Esmoriz/Vouga basins, respectively) and clade IV is distributed across Europe (see Online Resource, ESM 3). As suggested by Docker (2009), isolated populations of brook lampreys that are genetically very distinct may represent cryptic species. In fact, the number of brook lamprey species in the genus *Lampetra* may be underestimated as differentiated populations are often considered the same species due to their relatively conserved body form (Boguski 2009; Martin 2006). As most species-level characters in lamprey taxonomy are from the adult stage, morphological analysis of adult specimens representing each ESU identified in the present study is under way, and may provide further clarification on this issue.

The identification of cryptic species has important implications for conservation and natural resource protection and management (Bickford et al. 2007; Cook et al. 2008). These species require special consideration in conservation planning because the prevalence of cryptic complexes in already endangered nominal species presents a dual problem: species already considered endangered or threatened might be composed of multiple species that are even more rare than previously supposed; and the different species might require different conservation strategies

(Bickford et al. 2007). To consider the genetically unique brook lamprey populations as individual species would maximize their need for protection, as each putative cryptic species raise more serious conservation concerns, considering its extremely reduced distribution.

The proposed ESUs also include populations with some degree of divergence, as they are almost entirely composed of private haplotypes (Table 2) and the majority of the population pairwise  $F_{ST}$  revealed significant differences (see Online Resource, ESM 2). These populations should be managed separately for the conservation of biodiversity, constituting independent management units (MUs). These are the case of populations from Deva, Esmoriz, Águeda, Anços, Ribeira de S. Pedro, Nabão, Ribeira do Olival, Marateca and Sado.

The fact that the sympatric populations of *L. planeri* and *L. fluviatilis* are both included in the same clade (IV) raises the question whether the two species should be considered together or separately for conservation purposes. Several authors have questioned the validity of the classification of these species pair as two separate *taxa*, pointing out the possibility of two ecotypes instead (e.g., Eneqvist 1937; Schreiber and Engelhorn 1998) but the uncertainty about this issue (cf. Espanhol et al. 2007) makes the inclusion of individuals of both species in the same conservation units somehow premature. Although paired species presumably have very similar habitat requirements and similar or identical vulnerabilities in the larval stage, these differ considerably following metamorphosis. In particular, migratory and parasitic adults will be impacted by barriers to migration and by depletion of their prey base, factors that presumably would have little or no effect on nonparasitic brook lampreys. Consequently, until the taxonomic issue is fully understood it is important to conserve phenotypic diversity, protecting both parasitic and brook lampreys (Docker 2009). In view of that, we suggest that populations of the migratory *L. fluviatilis* should constitute a separate ESU. In this unit we include not only the threatened Iberian population but also populations from across Europe.

**Acknowledgments** We thank C. Moritz, R. Butlin, T.M. Azevêdo, N. Mesquita and H.F. Gante for valuable discussion. The Associate Editor L. Beheregaray and one anonymous reviewer provided detailed suggestions and comments that considerably improved the quality of this manuscript. We are grateful to all colleagues that assisted in fieldwork, especially C. Alexandre, N. Andrade, H. Monteiro and C. Maia who helped collecting the samples from Portugal. R. Rodríguez-Muñoz helped to acquire the license to collect the samples from Spain and provided valuable help during the sampling process. We also thank the collecting permits provided by Ministério da Agricultura, do Desenvolvimento Rural e das Pescas (Autoridade Florestal Nacional) in Portugal and the Asturian regional government (Consejería del Medio Ambiente, Ordenación del Territorio e Infraestructuras, Departamento de Biodiversidad y Paisaje) in Spain, with the help of Teresa Sanchez Corominas. This research was funded

by Fundação para a Ciência e Tecnologia, COMPETE (Programa Operacional Factores de Competitividade), QREN (Quadro de Referência Estratégico Nacional) and FEDER (Fundo Europeu de Desenvolvimento Regional) (project PTDC/BIA-BDE/71826/2006), and a PhD grant to C.S. Mateus (SFRH/BD44029/2008).

## References

- Aboim MA, Cunha C, Coelho MM (2009) Redistribution of the geographical ranges of the Iberian cyprinid genus *Pseudochondrostoma* based on a phylogenetic analysis: implications for the historical rearrangements of the north-western Iberian drainages. *J Fish Biol* 74:1337–1346. doi:10.1111/j.1095-8649.2009.02189.x
- Alexandrino J, Arntzen JW, 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. doi:10.1038/sj/hdy/6800010
- Almaça C, Collares-Pereira MJ (1988) On the occurrence of the European river lamprey *Lampetra fluviatilis* L., in the Iberian Peninsula. *Cybium* 12:9–15
- Alves TM, Gawthorpe RL, Hunt D, Monteiro JH (2000) Tertiary evolution of the São Vicente and Setúbal submarine canyons, Southwest Portugal: insights from seismic stratigraphy. *Ciências da Terra (UNL)* 14:243–256
- Antunes MT, Elderfield H, Legoinha P, Pais J (1995) Datações isotópicas com Sr do Miocénico do flanco Sul da Serra da Arrábida. *Com Inst Geol e Mineiro* 81:73–78
- Avise JC (2000) *Phylogeography: the history and formation of species*. Harvard University Press, Cambridge
- Barbosa BP (1995) Alostratigrafia e litostratigrafia das unidades continentais da bacia Terciária do Baixo-Tejo. Dissertation, University of Lisbon
- Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winke K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends Ecol Evol* 22:148–155. doi:10.1016/j.tree.2006.11.004
- Boguski, DA (2009) The genetic diversity of brook lampreys genus *Lampetra* (Petromyzontidae) along the Pacific coast of North America. Dissertation, University of Manitoba
- Cabral MJ, Almeida J, Almeida PR, Dellinger T, Ferrand N, Oliveira ME, Palmeirim JM, Queiroz AI, Rogado L, Santos-Reis M (eds) (2005) *Livro Vermelho dos Vertebrados de Portugal*. Instituto de Conservação da Natureza, Lisbon
- Calvo JP, Daams R, Morales J et al (1993) Up-to-date Spanish continental neogene synthesis and paleoclimatic interpretation. *Rev Soc Geol España* 6:29–40
- Choffat P (1908) *Essai sur la tectonique de la chaîne de l'Arrábida*. Comunicações dos Serviços Geológicos de Portugal, Lisbon, p 89
- Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Mol Ecol* 9:1657–1659. doi:10.1046/j.1365-294x.2000.01020.x
- Cook BD, Page TJ, Hughes JM (2008) Importance of cryptic species for identifying 'representative' units of biodiversity for freshwater conservation. *Biol Conserv* 141:2821–2831. doi:10.1016/j.biocon.2008.08.018
- Coppier G, Mougnot D (1982) Stratigraphie sismique et évolution géologique des formations néogènes et quaternaires de la plateforme continentale portugaise au Sud de Lisbonne. *Bull Soc Géol France* 24:421–431
- Crandall KA, Bininda-Emonds ORP, Mace GM, Wayne RK (2000) Considering evolutionary processes in conservation biology. *Trends Ecol Evol* 15:290–295. doi:10.1016/S0169-5347(00)01876-0
- Doadrio I (2001) *Atlas y Libro Rojo de los Peces Continentales de España*. Dirección General de Conservación de la Naturaleza, Museo Nacional de Ciencias Naturales, Madrid, p 364
- Docker MF (2009) A Review of the evolution of nonparasitism in lampreys and an update of the paired species concept. In: Brown LR, Chase SD, Mesa MG, Beamish RJ, Moyle PB (eds) *Biology, management, and conservation of lampreys in North America*. American Fisheries Society, Symposium 72, Bethesda, pp 71–114
- Eneqvist P (1937) The brook lamprey as an ecological modification of the river lamprey. On the river and brook lampreys of Sweden. *Arkiv Zool* 29:1–22
- Espanhol R, Almeida PR, Alves MJ (2007) Evolutionary history of lamprey paired species *Lampetra fluviatilis* (L.) and *Lampetra planeri* (Bloch) as inferred from mitochondrial DNA variation. *Mol Ecol* 16:1909–1924. doi:10.1111/j.1365-294X.2007.03279.x
- Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
- Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol Bioinf Online* 1:47–50
- Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791
- Freyhof J, Kottelat M (2008a) *Lampetra planeri*. In: IUCN 2010. IUCN Red List of Threatened Species. Version 2010.4. <www.iucnredlist.org>
- Freyhof J, Kottelat M (2008b) *Lampetra fluviatilis*. In: IUCN 2010. IUCN Red List of Threatened Species. Version 2010.4. <www.iucnredlist.org>
- Fu Y (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925
- Gante HF, Micael J, Oliva-Paterna FJ, Doadrio I, Dowling TE, Alves MJ (2009) Diversification within glacial refugia: tempo and mode of evolution of the polytypic fish *Barbus sclateri*. *Mol Ecol* 18:3240–3255. doi:10.1111/j.1365-294X.2009.04264.x
- Gómez A, Lunt DH (2006) Refugia within refugia: patterns of phylogeographic concordance in the Iberian Peninsula. In: Weiss S, Ferrand N (eds) *Phylogeography of Southern European Refugia*. Springer, Netherlands, pp 155–188
- Gyllensten U (1985) The genetic structure of fish: differences in the intraspecific distribution of biochemical genetic variation between marine, anadromous and freshwater species. *J Fish Biol* 26:691–699. doi:10.1111/j.1095-8649.1985.tb04309.x
- Hardisty MW (1986a) *Lampetra fluviatilis* (Linnaeus, 1758). In: Holčík J (ed) *The freshwater fishes of Europe*. 1/I Petromyzontiformes. Aula-Verlag, Wiesbaden, pp 249–278
- Hardisty MW (1986b) General introduction to lampreys. In: Holčík J (ed) *The freshwater fishes of Europe*. 1/I Petromyzontiformes. Aula-Verlag, Wiesbaden, pp 19–83
- Hardisty MW, Potter IC (1971) Paired species. In: Hardisty MW, Potter IC (eds) *The biology of lampreys Vol. 1*. Academic Press, London, pp 249–277
- Hurt C, Hedrick P (2004) Conservation genetics in aquatic species: general approaches and case studies in fishes and springsnails of arid lands. *Aquat Sci Res Across Bound* 66:402–413. doi:10.1007/s00027-004-0726-5
- Martin HR (2006) *Intraspecific Phylogeography of the Least Brook Lamprey (Lampetra aepyptera)*. Dissertation, University of Ohio
- Martínez-Solano I, Teixeira J, Buckley D, García-París M (2006) Mitochondrial DNA phylogeography of *Lissotriton boscai* (Caudata, Salamandridae): evidence for old, multiple refugia in

- an Iberian endemic. *Mol Ecol* 15:3375–3388. doi:[10.1111/j.1365-294X.2006.03013.x](https://doi.org/10.1111/j.1365-294X.2006.03013.x)
- Mesquita N, Carvalho G, Shaw P, Crespo E, Coelho MM (2001) River basin-related genetic structuring in an endangered fish species, *Chondrostoma lusitanicum*, based on mtDNA sequencing and RFLP analysis. *Heredity* 86:253–264. doi:[10.1046/j.1365-2540.2001.00776.x](https://doi.org/10.1046/j.1365-2540.2001.00776.x)
- Moritz C (1994) Defining “evolutionary significant units” for conservation. *Trends Ecol Evol* 9:373–375. doi:[10.1016/0169-5347\(94\)90057-4](https://doi.org/10.1016/0169-5347(94)90057-4)
- Moritz C, Faith DP (1998) Comparative phylogeography and the identification of genetically divergent areas for conservation. *Mol Ecol* 7:419–429. doi:[10.1046/j.1365-294x.1998.00317.x](https://doi.org/10.1046/j.1365-294x.1998.00317.x)
- Nylander JAA (2004) MrModeltest 2.2. Program distributed by the author. Department of Systematic Zoology. Evolutionary Biology Centre, Uppsala University, Sweden
- Paulo OS, Dias C, Bruford MW, Jordan WC, Nichols RA (2001) The persistence of Pliocene populations through the Pleistocene climatic cycles: evidences from the phylogeography of an Iberian lizard. *Proc R Soc B* 268:1625–1630. doi:[10.1098/rspb.2001.1706](https://doi.org/10.1098/rspb.2001.1706)
- Pereira AM, Robalo JJ, Freyhof J, Maia C, Fonseca JP, Valente A, Almada VC (2010) Phylogeographical analysis reveals multiple conservation units in brook lampreys *Lampetra planeri* of Portuguese streams. *J Fish Biol* 77:361–371. doi:[10.1111/j.1095-8649.2010.02675.x](https://doi.org/10.1111/j.1095-8649.2010.02675.x)
- Pimentel NL (1997) O terciário da bacia do Sado. Sedimentologia e análise tectono-sedimentar. Dissertation, University of Lisbon
- Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Syst Biol* 53:793–808
- Posada D, Crandall KA (1998) Model test: testing the model of DNA substitution. *Bioinformatics* 14:817–818. doi:[10.1093/bioinformatics/14.9.817](https://doi.org/10.1093/bioinformatics/14.9.817)
- Potter IC (1980) The Petromyzoniformes with particular reference to paired species. *Can J Fish Aquat Sci* 37:1596–1615
- Ribera I, Vogler AP (2004) Speciation of Iberian diving beetles in *Pleistocene refugia* (Coleoptera, Dytiscidae). *Mol Ecol* 13: 179–193. doi:[10.1046/j.1365-294X.2003.02035.x](https://doi.org/10.1046/j.1365-294X.2003.02035.x)
- Rodrigues A, Dias JMA (1989) Evolução pós-glaciária da plataforma continental Portuguesa a Norte do Cabo Mondego. *Anais do Instituto Hidrográfico* 10:39–50
- Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574. doi:[10.1093/bioinformatics/btg180](https://doi.org/10.1093/bioinformatics/btg180)
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497. doi:[10.1093/bioinformatics/btg359](https://doi.org/10.1093/bioinformatics/btg359)
- Ryder OA (1986) Species conservation and systematics: the dilemma of subspecies. *Trends Ecol Evol* 1:9–10. doi:[10.1016/0169-5347\(86\)90059-5](https://doi.org/10.1016/0169-5347(86)90059-5)
- Schreiber A, Engelhorn R (1998) Population genetics of a cyclostome species pair river lamprey (*Lampetra fluviatilis* L.) and brook lamprey (*Lampetra planeri* Bloch). *J Zool Syst Evol Res* 36:85–99. doi:[10.1111/j.1439-0469.1998.tb00781.x](https://doi.org/10.1111/j.1439-0469.1998.tb00781.x)
- Sousa-Santos C, Collares-Pereira MJ, Almada V (2007) Reading the history of a hybrid fish complex from its molecular record. *Mol Phylogenet Evol* 45:981–996. doi:[10.1016/j.ympev.2007.05.011](https://doi.org/10.1016/j.ympev.2007.05.011)
- Swofford DL (2002) PAUP\*. Phylogenetic Analysis Using Parsimony (\* and Other Methods). Version 4.0b10. Sinauer Associates, Sunderland
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595
- Ward RD, Woodwark M, Skibinski DOF (1994) A comparison of genetic diversity levels in marine, freshwater and anadromous fishes. *J Fish Biol* 44:213–232. doi:[10.1111/j.1095-8649.1994.tb01200.x](https://doi.org/10.1111/j.1095-8649.1994.tb01200.x)
- Wilson AJ, Gíslason D, Skúlason S, Snorrason SS, Adams CE, Alexander G, Danzmann RG, Ferguson MM (2004) Population genetic structure of Arctic Charr, *Salvelinus alpinus* from northwest Europe on large and small spatial scales. *Mol Ecol* 13:1129–1142. doi:[10.1111/j.1365-294X.2004.02149.x](https://doi.org/10.1111/j.1365-294X.2004.02149.x)
- Zanandrea G (1959) Speciation among lampreys. *Nature* 184:380. doi:[10.1038/184380a0](https://doi.org/10.1038/184380a0)

**Online Resource – *Conservation Genetics***

**MtDNA markers reveal the existence of allopatric evolutionary lineages in the threatened lampreys *Lampetra fluviatilis* (L.) and *Lampetra planeri* (Bloch) in the Iberian glacial refugium**

C.S. Mateus<sup>1,2,3,4,\*</sup>, P.R. Almeida<sup>1,2</sup>, B.R. Quintella<sup>1,5</sup> and M.J. Alves<sup>3,4</sup>

<sup>1</sup>Centro de Oceanografia, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

<sup>2</sup>Departamento de Biologia, Escola de Ciência e Tecnologia, Universidade de Évora, Largo dos Colegiais, 7000 Évora, Portugal.

<sup>3</sup>Museu Nacional de História Natural, Universidade de Lisboa, Rua da Escola Politécnica 58, 1250-102 Lisboa, Portugal.

<sup>4</sup>Centro de Biologia Ambiental, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

<sup>5</sup>Departamento de Biologia Animal, Faculdade de Ciências da Universidade de Lisboa, Campo Grande, 1749-016 Lisboa, Portugal.

\*Author to whom correspondence should be addressed. Tel.: +351 213921886; fax: +351 213969784; email: csmateus@fc.ul.pt





**ESM 2** Pairwise  $F_{ST}$  values and significance among the 21 sampled localities, calculated using conventional  $F$ -statistics based on mtDNA haplotype frequencies and the pairwise difference distance method

Locality	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1		***	***	***	***	***	*	***	***	***	***	***	***	***	***	NS	***	NS	NS	***	***
2	0.8841		***	***	***	***	***	***	***	***	NS	*	NS	NS	NS	NS	***	NS	***	***	***
3	0.9136	0.9430		***	***	***	***	***	***	***	***	***	***	***	***	NS	***	NS	***	***	***
4	0.9675	0.9901	0.4631		***	***	***	***	***	***	***	***	***	***	***	NS	***	NS	***	***	***
5	0.1454	0.7368	0.8646	0.9171		**	NS	***	***	***	***	***	***	***	***	NS	*	NS	NS	***	***
6	0.6150	0.6080	0.6755	0.7140	0.5294		***	***	***	***	***	*	***	*	***	NS	***	NS	**	***	***
7	0.1592	0.5918	0.7863	0.8284	-0.0521	0.4652		***	***	***	***	**	***	***	***	NS	NS	NS	NS	***	***
8	0.7446	0.4411	0.9037	0.9399	0.6612	0.5658	0.5769		***	***	***	NS	***	***	***	NS	***	NS	***	***	***
9	0.8725	0.9205	0.6607	0.7748	0.8154	0.6020	0.7331	0.8737		NS	***	***	***	***	***	NS	***	NS	***	***	***
10	0.7798	0.8432	0.5898	0.6721	0.7276	0.5600	0.6625	0.8107	-0.0529		***	***	***	***	***	NS	***	NS	***	***	***
11	0.6077	0.0703	0.8922	0.9171	0.5533	0.6243	0.5236	0.3422	0.8738	0.8282		*	NS	NS	NS	NS	***	NS	***	***	***
12	0.5595	0.1523	0.8199	0.8573	0.4663	0.3418	0.4090	0.1815	0.7607	0.6979	0.1770		NS	NS	**	NS	***	NS	**	***	***
13	0.8746	-0.0396	0.9372	0.9911	0.7030	0.5598	0.5532	0.4004	0.9079	0.8230	0.0408	0.1002		NS	NS	NS	***	NS	***	***	***
14	0.8251	0.0556	0.9275	0.9818	0.6608	0.5282	0.5263	0.3180	0.8921	0.8072	0.0429	0.0368	0.0278		NS	NS	***	NS	***	***	***
15	0.5550	0.1279	0.8976	0.9334	0.4601	0.5785	0.4161	0.3685	0.8692	0.8048	0.0707	0.1837	0.0894	0.0795		NS	***	NS	***	***	***
16	-0.2088	0.9718	0.8945	0.9830	-1.0000	0.3194	-0.7051	0.6429	0.8119	0.6773	0.4518	0.2051	1.0000	0.8667	0.3042		NS	NS	NS	NS	NS
17	0.3847	0.9259	0.9322	0.9793	0.0700	0.6823	0.1463	0.7925	0.9092	0.8272	0.6316	0.6363	0.9263	0.8888	0.5943	-0.5714		NS	*	***	***
18	0.2517	0.9753	0.8978	0.9835	-0.4375	0.3485	-0.4276	0.6732	0.8176	0.6881	0.5155	0.2619	1.0000	0.8824	0.3965	1.0000	0.2143		NS	NS	NS
19	0.1806	0.5753	0.7541	0.7989	-0.0217	0.4018	-0.0639	0.5541	0.6879	0.6164	0.5235	0.3641	0.5289	0.4975	0.4122	-0.6378	0.1932	-0.4602		***	***
20	0.8643	0.9080	0.9185	0.9404	0.8406	0.8120	0.7994	0.8822	0.8904	0.8475	0.8661	0.8274	0.8994	0.8890	0.8646	0.8362	0.8924	0.8421	0.7888		***
21	0.7939	0.8526	0.8023	0.8274	0.7565	0.6380	0.7173	0.8238	0.7155	0.6791	0.8533	0.7176	0.8259	0.8046	0.8228	0.6363	0.8490	0.6453	0.6768	0.5230	

Below diagonal, pairwise  $F_{ST}$  values; above diagonal, significance of the pairwise  $F_{ST}$  values

\*Significant at the 5% level; \*\*significant at the 1% level; \*\*\*significant at the 0.1% level; NS, not significant

**ESM 3** Neighbour-joining phylogenetic tree of 68 mitochondrial haplotypes of *Lampetra* (56 from this study and 12 (highlighted) from European populations of *L. planeri* and *L. fluviatilis* included in Espanhol et al. (2007)). Numbers in terminal nodes refer to the haplotype numbers. For the haplotypes from European populations, species, country and river basin are indicated. The abbreviations for the species are: *L. fluv*, *L. fluviatilis*; *L. plan*, *L. planeri*. Clades I to IV and subclades IV-A to IV-C refer to clades obtained in the ML phylogenetic tree of 56 haplotypes (Fig. 2) from this study. The evolutionary model of nucleotide substitution and its parameters was calculated according to the Bayesian Information Criterion (BIC) (HKY + I + G; proportion of invariable sites = 0.4841; gamma distribution shape parameter = 0.7302; base frequencies: A = 0.3049, C = 0.2405, G = 0.1218, T = 0.3328; transition/transversion ratio = 5.2256). Numbers are the bootstrap support values equal to or higher than 50% obtained from neighbour-joining

