Genetic evidence fails to discriminate between *Macroramphosus gracilis* Lowe 1839 and *Macroramphosus scolopax* Linnaeus 1758 in Portuguese waters

Joana Isabel Robalo · C. Sousa-Santos · H. Cabral · R. Castilho · V. C. Almada

Abstract Fish belonging to the genus *Macroramphosus* are distributed throughout the Atlantic, Indian and Pacific oceans. Some authors consider this genus monotypic, *Macroramphosus scolopax* being the only valid species. Other authors consider (based on several morphological and ecological characters) that another species (*Macroramphosus gracilis*) exists and occurs frequently in sympatry with the first one. Intermediate forms are also reported in literature. In this paper, using the mitochondrial control region and the nuclear first S7 intron markers, we failed to find genetic differences between individuals considered to belong to both species as well as the intermediate forms. Our results suggest that in the northeastern Atlantic, *Macroramphosus* is represented by a single species, *M. scolopax*, with different morphotypes interbreeding in the sampling areas.

Introduction

The genus *Macroramphosus* (Lacepède 1803) (family Macroramphosidae Nelson 2006) comprises small, gregarious fish distributed around the Atlantic, Indian and Pacific oceans (Nelson 2006), mainly in latitudes between 20º and 40º north (Ehrich 1986).

There has been considerable debate over the years on the number of valid species in the genus. The existence of two forms, one comprising slender dark fish and another including deep-bodied orange animals with a more developed second dorsal spine, led many authors to accept the existence of at least two species (e.g. Mohr 1937): *Macroramphosus gracilis* (Lowe 1839) and *Macroramphosus scolopax* (Linnaeus 1758), respectively. Ehrich (1976) synonymised all 15 nominal species of the genus under *M. scolopax* assuming that the slender form tends to change to the deep-bodied one during ontogeny, a view supported by Oliveira et al. (1993), based on long-term observations of captive fish. The scarcity of small-sized deep-bodied fish and the presence of intermediate forms (Assis 1992) also argue in favour of this view which was adopted by Ehrich (1986) and Quéro et al. (1990).

In contrast, several recent studies (e.g. Matthiessen et al. 2003; Miyazaki et al. 2004; Marques et al. 2005; Bilecenoglu 2006) consider both species valid and found substantial differences between them in morphology, larval development, behaviour and feeding habitats. All these studies were based on the analysis of sympatric forms, thus allowing the exclusion of possible artefacts caused by comparisons of geographically distinct populations.

This paper is the first attempt to use DNA sequences to assess the validity of these two species. We used two markers, one mitochondrial (control region) and one nuclear (S7 ribosomal protein, first intron).

Materials and methods

Sample collection

Samples were obtained from commercial fishing vessels operating with bottom trawls in the central and southwestern Portuguese coast, at depths of about 200–300 m, in the spring and summer of 2006. All individuals were classified as "*gracilis*", "*scolopax*" or "*intermediate*" according to the criteria described by Lopes et al. (2006). A piece of muscle was stored in 96% ethanol for subsequent DNA analysis.

DNA procedures

Total DNA from 36 "*gracilis*", 31 "*scolopax*" and 20 "*intermediate*" was extracted with a REDExtract-N-Amp kit (Sigma-Aldrich, www.sigma.com). Voucher specimens are deposited in ISPA and CCMAR collections (ethanol preserved tissues).Whenever possible, each fish was sequenced for both fragments.

To amplify the mitochondrial control region and the S7 nuclear gene, the following pairs of primers were used:

- Control region: L-PRO1 5′-ACT CTC ACC CCT AGC TCC CAA AG-3′ and H-DL1 5′-CCT GAA GTA GGA ACC AGA TGC CAG-3′ (Ostellari et al. 1996);
- S7, first intron: S7RPEX1F 5′-TGG CCT CTT CCT TGG CCG TC-3′and S7RPEX2R 5′-AAC TCG TCT GGC TTT TCG CC-3′ (Chow and Hazama 1998).

PCR amplification reactions were performed in a 20 µl total-reaction volume with 10 µl of REDExtract-N-ampl PCR reaction mix (Sigma-Aldrich), 0.8 µl of each primer (10 µM), 4.4 µl of sigma-water and 4 µl of template DNA. For the control region and S7, respectively, the PCR conditions were the following: an initial denaturation at 94°C for 2/3 min was followed by 35 cycles (denaturation at 94°C for 30/45 s, annealing at 55/58º for 30/45 s, and extension at 72°C for 1 min) and a final extension at 72°C for 5/10 min on a BioRad Mycycler thermal cycler. The same primers were used for the sequencing reaction, and the PCR products were purified and sequenced in STABVIDA (http://www.stabvida.net/).

Phylogenetic analyses

Sequences were edited with CodonCode Aligner v. 2.0 (http://www.codoncode.com/) and aligned with Clustal X (Thompson et al. 1997). All sequences have been deposited in GenBank (available at www.ncbi.nlm.nih.gov/) with the following accession numbers: FJ457702–FJ457763 (control region) and FJ457647–FJ457701 (S7).

For all phylogenetic analyses, we included *Hippocampus kuda* as outgroup (GenBank Accession numbers NC_010272 and DQ288388 for the control region and S7 first intron, respectively). For each fragment, maximum parsimony (MP) and minimum evolution (neighbour joining, NJ) were performed separately. MP and NJ analyses were performed with PAUP* 4.0 (Swofford 2002). MP analysis was conducted using a heuristic search strategy with random stepwise addition (1,000 replicates) and TBR branch swapping. Bootstrap analyses (1,000 replicates) were used to assess the relative robustness of branches of the ME and the MP trees (Felsenstein 1985).

For NJ, the best-fit model of nucleotide substitution was selected with the program Modeltest 3.0 (Posada and Crandall 1998) with the Akaike Information Criterion (AIC).

ARLEQUIN software package version 3.1 (Excoffier and Schneider 2005) was used to estimate the genetic diversity in the species studied, to access population differentiation, to perform neutrality tests and to determine the number of haplotypes. Analysis of molecular variance (AMOVA–Excoffier et al. 1992) and pairwise FST were also performed. Intra-group distances were corrected by subtracting the mean intraspecific pairwise distances as implemented in Arlequin. Relationships amongst haplotypes were analysed with a parsimony network estimated by the software TCS version 1.18 (Clement et al. 2000).

Results

Not all samples were screened at the nuclear loci due to unsuccessful PCR amplification or non-repeatable band patterns in some samples. A total of 356 and 551 bp were amplified corresponding to the control region and S7 (values after alignment).

For the control region, 62 fishes were sequenced (20 "*scolopax*", 22 "*gracilis*" and 20 "*intermediate*"); 50 haplotypes were found and are represented in Figs. 1 and 2.

For the S7 intron, 29 fish (11 "*scolopax*", 14 "*gracilis*" and 4 "*intermediate*") corresponded to 3 haplotypes. The most common haplotype was present in 27 of the 29 fish in homozygosity. The remaining two fish (both "*gracilis*") were heterozygous, having one DNA chain equal to the common one and another one that differed from the first by one mutation (in different positions in each fish).

As the number of haplotypes in S7 was too small to perform the vast majority of the analyses presented in the "Materials and methods" section we opted to further explore only the results of the control region data set. Anyway, the almost complete lack of variation in this marker, which in other species proved to be sufficiently variable to be used as a phylogeographic tool (Domingues et al. 2007), strongly argues against the hypothesis that the different morphotypes correspond to different taxa.
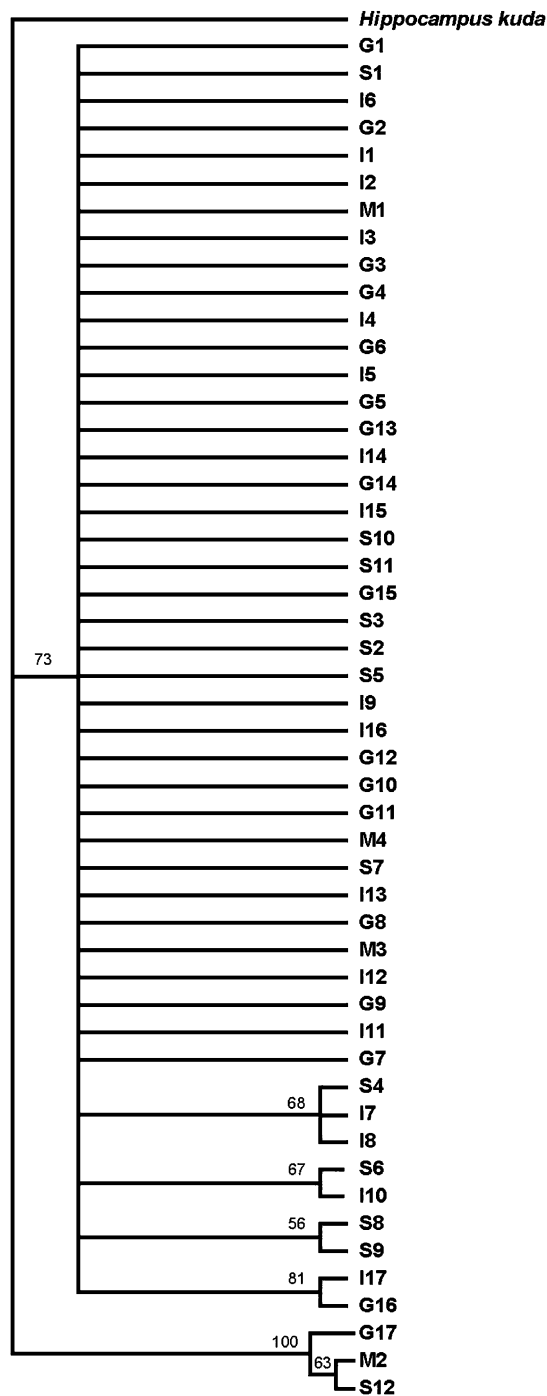
Fig. 1 MP tree of the mitochondrial control region. Tree length = 284; consistency index (CI) = 0.6056; homoplasy index (HI) = 0.3944; CI excluding uninformative characters = 0.3846; HI excluding uninformative characters = 0.6154; retention index (RI) = 0.3333; rescaled consistency index (RC) = 0.2019. The values presented in the nodes correspond to MP. As all the bootstrap values for NJ were lower than 50%, they are not represented. Nodes with no bootstrap support (inferior to 50) in a given method are represented by -. Only haplotypes are represented and are the same considered in TCS. *Letters* represent original classification: *G* "*gracilis*", *I* "*intermediate*", *S* "*scolopax*" and *M* mixed forms [*M1*: *gracilis* (1), intermediate (1); *M2*: *gracilis* (2), *scolopax* (1); *M3*: intermediate (1), *scolopax* (1); *M4*: *scolopax* (4), *intermediate* (1) and *gracilis* (1)]
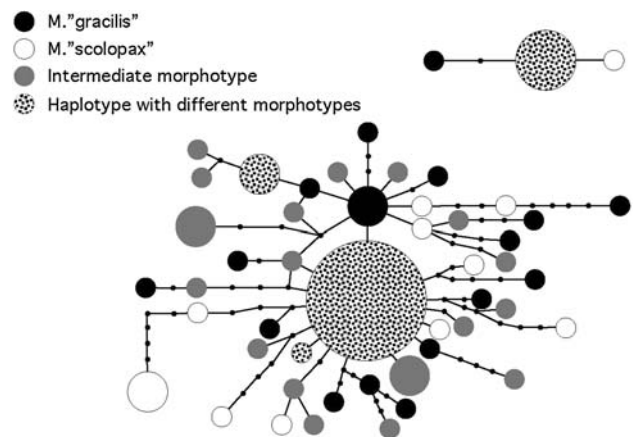

Fig. 2 TCS 95% confidence parsimony network of haplotypes for the control region

Concerning parsimony analyses for the control region, 292 characters were constant, 24 were parsimony-uninformative and 40 were parsimony-informative. As much as 100 trees were retained.

Figures 1 and 2 show that there is no evident differentiation between the forms of *Macroramphosus*. In both figures, only haplotypes are represented. The only differentiation present is the one that separates the haplotypes M2 (three individuals, two "*gracilis*" and one "*scolopax*"), S12 (one individual, "*scolopax*") and G17 (one individual, "*gracilis*") from the remaining ones. This is well represented in the tree (Fig. 1) and in the 95% confidence network made by TCS (Fig. 2) where this group of haplotypes represents an individual network. As the two groups of haplotypes include both "*scolopax*" and "*gracilis*" forms, they do not argue in favour of their genetic separation. They likely represent an ancient polymorphism retained in the population. The remaining haplotypes (47) are grouped in a different network, all individuals being at a maximum of 11 mutations from the ancestral one. Seven haplotypes were present in more that one individual, although the vast majority of haplotypes (43 in 50) represents only one fish. The haplotype that represents more individuals is M4, representing 6 fish (4 "*scolopax*", 1 "*gracilis*" and 1 "*intermediate*"). The large proportion of unique haplotypes could obscure any potential differences amongst the forms. However, neither the phylogenetic trees nor the haplotype network revealed haplotype groups consistently separated according to each fish morphotype. On the contrary, haplotype groups typically represent mixtures of fish of different forms.

The results of the AMOVA considering the three forms were nonsignificant ($p = 0.30792 \pm 0.00912$, percentage of variation between populations = 0.48, percentage of variation within populations = 99.52). The same holds for pairwise FSTs and net average distances (Table 1).

Table 1 Above diagonal pairwise FSTs/corresponding *p* values, below the diagonal net average distances, diagonal average number of pairwise differences within population

	gracilis	*scolopax*	*intermediate*
gracilis	9.09524	−0.01198/0.56757 ± 0.0430	0.01563/0.17117 ± 0.0394
scolopax	−0.09858	7.94737	0.01410/0.14414 ± 0.0309
intermediate	0.12774	0.09895	5.89474

Discussion

The present study failed to detect genetic differences between the forms of *Macroramphosus* found in Portugal, supporting the presence of a single species, *M. scolopax*.

Oliveira et al. (1993) observed spawnings involving partners of the different forms in captivity, thus providing circumstantial evidence favouring interbreeding.

Our results are compatible with two different hypotheses: (a) the speciation might be so recent and incipient that the molecular markers used do not evolve fast enough to capture its signature (microsatelites should be used to test this hypothesis) or (b) two morphological types coexist and interbreed. Such a difference in morphology could derive either from genetic causes, environmental conditions during development, distinct ontogenetic stages or a combination of those factors. This hypothesis seems more plausible and in accordance with our results. The fact that these forms occur consistently in sympatry in several ocean basins is consistent with an ancient differentiation of morphotypes in the snipefish populations. Indeed, it is difficult to imagine that the slender and deep-bodied fish represent two species and were able to disperse globally in a process that was so recent that its signature was not detectable even with the rapidly evolving control region.

It is known that the large "*scolopax*" form tends to feed on the bottom, whilst the "*gracilis*" form is more planktivorous and occurs at a higher level in the water column (e.g. Matthiessen et al. 2003). Rapid evolution of benthic and limnetic forms was documented in fish lineages as different as *Gasterosteus* (e.g. McPhail 1994) and *Coregonus* (e.g. Østbye et al. 2005 and references therein). In both sticklebacks and whitefish there are examples of this type of eco-morphological evolution, which occurred in the last 10,000–12,000 years and even led to the formation of new species. If a similar process of differentiation of benthic and pelagic forms took place in snipefish, but the different forms continued to interbreed, these morphotypes might have persisted in the populations, although they exploited different ecological niches.

The presence of *Macroramphosus* in both sides of the Atlantic and in the Indian and Pacific Oceans makes snipefish a very promising subject for phylogeographic analysis. How many species of *Macroramphosus* do occur worldwide, remains an open question, regardless of the probable interbreeding of the different forms in each population. Indeed, fish in different oceans or in the two margins of the Pacific may have become genetically isolated, representing distinct species.

References

Assis CA (1992) On the systematics of *Macrorhamphosus scolopax* (Linnaeus, 1758) and *Macrorhamphosus gracilis* (Lowe, 1839). I. A preliminary biometrical approach. Bol Soc Port Cienc Nat (2nd Sér.) 25:5–19

Bilecenoglu M (2006) Status of the genus *Macroramphosus* (Syngnathiformes:Centriscidae) in the eastern Mediterranean Sea. Zootaxa 1273:55–64

Chow S, Hazama K (1998) Universal PCR primers for S7 ribosomal protein gene introns in fish. Mol Ecol 7:1247–1263. doi:10.1046/j.1365-294x.1998.00406.x

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

Domingues VS, Santos RS, Brito A, Alexandrou M, Almada VC (2007) Mitochondrial and nuclear markers reveal isolation by distance and effects of Pleistocene glaciations in the northeastern Atlantic and Mediterranean populations of the white seabream (*Diplodus sargus*, L.). J Exp Mar Biol Ecol 346:102–113. doi:10.1016/j.jembe.2007.03.002

Ehrich S (1976) Zur Taxonomie, Ökologie und Wachstum von *Macroramphosus scolopax* (Linnaeus, 1758) (Pisces, Syngnathiformes) aus dem subtropischen. Ber Deut Wiss Komm 24:251–266

Ehrich S (1986) Macroramphosidae. In: Whitehead P, Bauchot M, Hureau J, Nielsen J, Tortonese E (eds) Fishes of the northeastern Atlantic and the Mediterranean (FNAM), vol 2. UNESCO, Paris, p 627

Excoffier L, Schneider S (2005) Arlequin ver. 3.0: an integrated software package for population genetics data analysis. Evol Bioinform Online 1:47–50

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

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evol Int J Org Evol 39:783–791. doi:10.2307/2408678

Lopes M, Murta AG, Cabral HN (2006) Discrimination of snipefish *Macroramphosus* species and boarfish *Capros aper* morphotypes through multivariate analysis of body shape. Helgol Mar Res 60:18–24. doi:10.1007/s10152-005-0010-7

Marques V, Chaves C, Morais A, Cardador F, Stratoudakis Y (2005) Distribution and abundance of snipefish (*Macroramphosus* spp.) of Portugal (1998–2003). Sci Mar 69:563–576

Matthiessen B, Fock H, von Westernhagen H (2003) Evidence for two sympatric species of snipefishes *Macroramphosus* spp. (Syngnathiformes, Centriscidae) on Great Meteor Seamount. Helgol Mar Res 57:63–72

McPhail JD (1994) Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of southwestern British Columbia. In: Bell MA, Foster SA (eds) The evolutionary biology of the threespine stickleback. Oxford Science Publications, Oxford, pp 399–437

Miyazaki E, Sasaki K, Mitani T, Ishida M, Uehara S (2004) The occurrence of two species of *Macroramphosus* (Gasterosteiformes:Macroramphosidae) in Japan: morphological and ecological observations on larvae, juveniles, and adults. Ichthyol Res 51:256–262. doi:10.1007/s10228-004-0227-5

Mohr E (1937) Revision der centriscidae (Acanthopterygii, Centrisciformes). Dana Report 13:1–69

Nelson JS (2006) Fishes of the world, 4th edn. Wiley, New York

Oliveira RF, Almada VC, Gil MF (1993) The reproductive behavior of the longspine snipefish *Macrorhamphosus scolopax* (Syngnathiformes, Macrorhamphosidae). Environ Biol Fishes 36:337–343. doi:10.1007/BF00012410

Østbye K, NÆsje TF, Bernatchez L, Sandlund OT, Hindar K (2005) Morphological divergence and origin of sympatric populations of European whitefish (*Coregonus lavaretus* L.) in Lake Femund, Norway. J Evol Biol 18:683–702. doi:10.1111/j.1420-9101.2004.00844.x

Ostellari L, Bargelloni L, Penzo E, Patarnello P, Patarnello T (1996) Optimization of single-strand conformation polymorphism and sequence analysis of the mitochondrial control region in *Pagellus bogaraveo* (Sparidae, Teleostei): rationalized tools in fish population biology. Anim Genet 27:423–427

Posada D, Crandall KA (1998) Modeltest: testing the model of DNA substitution (v.3.6). Bioinformatics 14:817–818. doi:10.1093/bioinformatics/14.9.817

Quéro JC, Hureau JC, Karrer C, Post A, Saldanha L (eds) (1990) Clofeta I-III, checklist of the fishes of the eastern tropical Atlantic. Junta Nacional de Investigação Científica e Tecnológica, Lisboa

Swofford DL (2002) PAUP*: phylogenetic analysis using parsimony (* and other methods) version 4.0. Sinauer Associated, Sunderland, MA

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882. doi:10.1093/nar/25.24.4876