Deep phylogeographic structure may indicate cryptic species within the Sparid genus Spondyliosoma

McKeown, Niall; Gwilliam, Michael; Healey, Amy; Skujina, Ilze; Potts, Warren; Sauer, W. H. H.; Shaw, Paul

Published in:
Journal of Fish Biology

DOI:
10.1111/jfb.14316

Publication date:
2020

Citation for published version (APA):
McKeown, N., Gwilliam, M., Healey, A., Skujina, I., Potts, W., Sauer, W. H. H., & Shaw, P. (2020). Deep phylogeographic structure may indicate cryptic species within the Sparid genus Spondyliosoma. *Journal of Fish Biology*, *96*(6), 1434-1443. https://doi.org/10.1111/jfb.14316

Received: 26 August 2019 | Accepted: 9 March 2020

DOI: 10.1111/jfb.14316

REGULAR PAPER

JOURNAL OF FISH BIOLOGY fsbi

Deep phylogeographic structure may indicate cryptic species within the Sparid genus *Spondyliosoma*

Niall J. McKeown[1] | Michael P. Gwilliam[1] | Amy J.E. Healey[1] | Ilze Skujina[1] | Warren M. Potts[2] | Warwick H.H. Sauer[2] | Paul W. Shaw[1,2]

[1]Institute of Biological, Environmental, and Rural Sciences, Aberystwyth University, Aberystwyth, UK

[2]Department of Ichthyology and Fisheries Science, Rhodes University, Grahamstown, South Africa

Correspondence
Niall J. McKeown, Institute of Biological, Environmental, and Rural Sciences, Aberystwyth University, Aberystwyth SY23 3DA, UK.
Email: njm2@aber.ac.uk

Abstract

Two geographically nonoverlapping species are currently described within the sparid genus *Spondyliosoma*: *Spondyliosoma cantharus* (Black Seabream) occurring across Mediterranean and eastern Atlantic waters from NW Europe to Angola and *S. emarginatum* (Steentjie) considered endemic to southern Africa. To address prominent knowledge gaps this study investigated range-wide phylogeographic structure across both species. Mitochondrial DNA sequences revealed deep phylogeographic structuring with four regionally partitioned reciprocally monophyletic clades, a Mediterranean clade and three more closely related Atlantic clades [NE Atlantic, Angola and South Africa (corresponding to *S. emarginatum*)]. Divergence and distribution of the lineages reflects survival in, and expansion from, disjunct glacial refuge areas. Cytonuclear differentiation of *S. emarginatum* supports its validity as a distinct species endemic to South African waters. However, the results also indicate that *S. cantharus* may be a cryptic species complex wherein the various regional lineages represent established/incipient species. A robust multilocus genetic assessment combining morphological data and detailing interactions among lineages is needed to determine the full diversity within *Spondyliosoma* and the most adequate biological and taxonomic status.

KEYWORDS

Benguela, microsatellite, mtDNA, phylogeny, Sparid fish, species, taxonomy

1 | INTRODUCTION

The Sparidae (seabreams) are demersal fishes commonly found at a range of depths in temperate and tropical marine waters with maximum species diversity in the NE Atlantic Ocean and Mediterranean Sea (Bauchot & Hurau, 1986). Within this family the genus *Spondyliosoma* is currently recognized as comprising two species, *Spondyliosoma cantharus* (black seabream) and *Spondyliosoma emarginatum* (Steentjie). *S. cantharus* exhibits a wide geographical range from Scandinavia to Angola and occurring around the islands of Madeira, Cape Verde and the Canary Islands in the eastern Atlantic and throughout the Mediterranean Sea (Bauchot & Hurau, 1986). *S. emarginatum* is considered endemic to southern Africa. The gap between the ranges of the two species corresponds with the location of the Benguela Upwelling System (BUS), an established biographic boundary for a number of taxa (Grant & Bowen, 1998; Henriques *et al*., 2012, 2014, 2015). Throughout their respective ranges both species represent important fisheries resources and there are indications that stocks may be overfished (Correia *et al*., 2012).

Genetic studies have provided considerable insight into evolutionary processes in the marine realm as well as identification of

Niall J McKeown and Michael P Gwilliam should be considered joint first authors.

population units for fishery management. There has been only one genetic study including *Spondyliosoma* (Bargelloni *et al*., 2003). This focused on the relationship between Mediterranean and NE Atlantic populations (the latter represented only by a single site in Portugal), and indicated reciprocally monophyletic clades present in the two areas. To date there have been no genetic studies of *S. emarginatum* or of the wider Atlantic distribution of *S. cantharus*.

The limited spatial genetic data available for *Spondyliosoma* underpins a number of prominent knowledge gaps relating to the proper definition of operational units (species to population) that may compromise both conservation of biodiversity and sustainable fishery management. First, the phylogenetic relationship between *S. cantharus* and *S. emarginatum*, and thus their validity as distinct phylogenetic species, remains untested. This relationship is also of interest regarding the role of the BUS as a biogeographic barrier across SE Atlantic coasts. The BUS has been associated with prolonged genetic isolation and speciation in some groups, with other fish species reporting signals of historical and/or recurrent gene flow across it (Henriques *et al*., 2015; Reid *et al*., 2016; Sala-Bozano *et al*., 2009). Second, at the intraspecific level phylogeographic studies can provide insight into historical and recurrent population processes. The Pleistocene glaciations have left pronounced signatures of divergence in, and expansion from, glacial refugia in a number of marine species (Maggs *et al*., 2008). Phylogeographic studies can thus identify endemic lineages that are important and irreplaceable components of a species' evolutionary potential and warrant conservation prioritization. From a fisheries management perspective spatial genetic data can also resolve patterns of population connectivity/isolation useful for the optimization of spatial management strategies (Reiss *et al*., 2009; but see Waples & Gaggiotti, 2006; Hauser & Carvalho, 2008). Phenotypic studies have resolved regional differences among Atlantic *S. cantharus* (Neves *et al*., 2018, 2019) and genetic data could provide information as to what extent such phenotypic differences reflect population isolation or environmental driven variation. The identification of demographically independent populations is important for *Spondyliosoma* as the species' slow growth and habitat specificity may make such populations particularly vulnerable to overexploitation (Neves *et al*., 2017; Pinder *et al*., 2017).

This study aimed to investigate phylogeographic structuring across a large portion of the range of *S. cantharus* and *S. emarginatum* using mitochondrial DNA (mtDNA) sequence variation. As an initial focus was to assess the species status of *S. emarginatum*, nuclear DNA differentiation was also assayed between *S. emarginatum* from South Africa (southern Benguela subsystem) and *S. cantharus* from Angola (the putative southern limit of *S. cantharus* and within the northern Benguela subsystem), *i.e*., the geographically closest representatives of the two species.

2 | MATERIALS AND METHODS

2.1 | Sample collection and inclusion of GenBank data

This study complied with all ethical requirements of the *Journal of Fish Biology* and local authorities. No fish were killed or interfered with in any way as all samples were in the form of ethanol preserved fin clips acquired from local fishers. Samples of *S. emarginatum* were collected from six sites in South African waters, while *S. cantharus* were collected from four sites in Angolan waters (Table 1 and Figure 1). A NE Atlantic *S. cantharus* sample was collected from Cardigan Bay (Wales) (Figure 1). DNA extraction followed the phenol-chloroform method described by Sambrook *et al*., (1989). To increase geographical representation in subsequent analyses of mtDNA cytochrome oxidase I (COI) sequence variation, we also included *S. cantharus* COI sequences deposited on GenBank from barcoding studies around Portugal (Costa *et al*., 2012) and Turkey (Keskin & Atar, 2013). We also included three sequences (GenBank accession no. KJ012436, KJ012438 and KJ012439) from a barcoding study of samples collected at an Italian fish market (Armani *et al*., 2015). Additional samples from Armani *et al*. (2015) had shorter sequences than our final alignment, so were omitted from the main results but their phylogenetic relationships were also tested by truncating our overall sequence alignment accordingly.

2.2 | mtDNA analysis

A fragment of the cytochrome oxidase I gene was PCR amplified using primers SCCOIF (5′-GCTTGAGCCGGAATAGTAG-3′) and SCCOIR (5′-TTGGTAAAGAATTGGGTCTCC-3′) designed from GenBank *S. cantharus* sequences. Similarly, new primers (SCRF5′-CACACATAATGTTAGAGATATAGGA-3′ and SCRR 5′-TGCATAAG TGATTTCATGAGCATAAT-3′) were designed to permit PCR amplification of the first hypervariable region of the mtDNA control region (CR). PCRs for both mtDNA regions comprised 10 μl of BIOMIX (BioLine), 1.0 pMol of primer (both forward and reverse), 6 μl of template DNA and 2 μl of sterile distilled water, giving a total reaction volume of 20 μl. The PCR thermoprofile was: 300 s at 94°C, then 40 cycles of 30 s at 94°C, 30 s at 50°C (CR-51°C) and 60 s at 72°C, with a final extension step of 300 s at 72°C. Amplicons were sequenced with the respective forward primer using BigDye technology and an ABI 3730 DNA analyser (Applied Biosystems).

Analyses were performed using ARLEQUIN (Excoffier & Lischer, 2010) unless stated otherwise. Genetic diversity was estimated using haplotype (h) and nucleotide (π) diversity. A phylogenetic tree among COI haplotypes with a *Spicara maena* (GenBank: AP009164) outgroup was constructed using maximum likelihood (ML) implemented in MEGA v 7 (Kumar *et al*., 2016) using the HKY model identified as most suitable by MRMODELTEST (Nylander, 2004) with the nearest-neighbor-interchange heuristic method and weak branch swap filter. Nodal support values were estimated from 500 nonparametric bootstrap replicates. Phylogenetic relationships were also assessed using median joining networks constructed in NETWORK (www.fluxus-engineering.com/sharenet.htm). Pairwise divergences between and within lineages were estimated using mean K2P distances (Kimura, 1980). Divergence times among lineages (COI) were estimated in BEAST 1.6.1 (Drummond & Rambaut, 2007) using a strict molecular clock and Yule process.

TABLE 1 Sample information including geographical region and ascribed taxon, specific sites and associated numbers of individuals for which mtDNA cytochrome oxidase I, mtDNA control region and microsatellite data were collected

Region	Site/source	Sample size	COI	CR	Microsatellites
South Africa *S. emarginatum*	KwaZulu-Natal (1)	5	5		
	Port Elizabeth (2)	6		6	
	Tsitsikamma (3)	9	6	6	
	Mossel Bay (4)	21	9	15	21
	False Bay (5)	30	9	13	27
	Langebaan (6)	1	1		
Angola *S. cantharus*	Tombua (7)	15	9	5	
	Namibe (8)	30	11	12	29
	Lucira (9)	15	8	5	
	Benguela (10)	29	13	9	27
NE Atlantic *S. cantharus*	Portugal (GenBank)		16		
	Aberystwyth (11)	16	15	16	
Mediterranean *S. cantharus*	Turkey (GenBank)		21		

Note: Numbers in brackets after sites correspond to locations in Figure 1, or in the case of GenBank indicate data obtained from previous studies. COI, cytochrome oxidase I; CR, control region.

Testing using MEGA indicated that sequence data did not deviate from a model of constant evolutionary rate. The Markov Chain was run for 50×10^6 iterations and repeated once with Tracer, used to check for convergence and an effective sample size (ESS) of >200. The maximum clade credibility tree was estimated in TreeAnnotator with the first 10% of trees discarded during the burn-in. To complement this phylogenetic-based analysis, divergence times among populations were also estimated using the coalescent approach implemented in IMa (Hey, 2010) with 1×10^6 burn in generations and 5×10^6 sampling generations to ensure the minimum ESS was >50 (Hey & Nielsen, 2004). Metropolis coupling with a geometric heating scheme for one cold chain and 59 heated chains was used for each run, which was then replicated using a different random number seed. Runs were also performed using a combination of priors for splitting time, maximum population size and migration rate other than the default values. The results converged on the same stationary distributions.

Differentiation between pairs of samples was quantified using pairwise Φ_{ST} with significance assessed by 10,000 permutations. Fu's *Fs* (Fu, 1997) and Tajima's *D* (Tajima, 1989) were used to test for deviations from mutation-drift equilibrium (significance assessed after 10,000 permutations). Mismatch distributions, the frequency distributions of pairwise differences between haplotypes within a sample and simulated distributions under a model of demographic expansion were compared using the sum of squared deviations (SSD) as a test statistic with significance assessed after 10,000 bootstrap replications. The timing of expansions (T) was estimated from $T = \tau/2u$ (Rogers & Harpending, 1992). For mutation rate dependent analyses we employed (a) the widely accepted 1.2% per million years divergence rate (0.006 substitutions per site per million years) for COI (Bermingham *et al*., 1997) and (b) the 11% per million years divergence rate (0.055 substitutions per site per million years) for CR that has been widely used for sparids (Bargelloni *et al*., 2003; Coscia *et al*., 2012; Sala-Bozano *et al*., 2009).

2.3 | Microsatellite analysis

Eighteen microsatellite loci developed previously for a number of Sparid species were tested for PCR amplification in *Spondyliosoma* (Supporting Information Table S1). Following testing, four loci reporting reliable amplification (DsaMS27, DsaMS34 and DsaMS48 from Perez *et al*., (2008) and Dvul84 from Roques *et al*. (2007)) were selected for genotyping of samples from two locations in both South Africa and Angola. Loci were individually amplified by PCR and genotypes separated on an AB3730 DNA sequencer with alleles sized using the software Peak Scanner (Applied Biosystems).

Genetic variation within samples was characterized using the number of alleles (N_A), allelic richness (A_R), observed heterozygosity (H_O) and expected heterozygosity (H_E), all calculated using GENALEX 6.2 (Peakall & Smouse, 2006). Genotype frequency conformance to Hardy–Weinberg expectations (HWE) and genotypic linkage equilibrium between pairs of loci were tested using exact tests (10,000 demorisations, 10,000 batches, 5000 iterations) in GENEPOP 3.3 (Rousset, 2008). Genetic differentiation among samples was quantified using global and pairwise F_{ST} values with significance assessed with P values following 10,000 permutations in FSTAT (Goudet, 1995). F_{ST} values were also estimated using the null allele correction method in FreeNA (Chapuis and Estoup, 2007). The Bayesian clustering method implemented in the program STRUCTURE (Pritchard *et al*., 2000) was used to identify the most probable number of genetic clusters (K) (from a range of 1–5) within the data. The analysis was performed both with and without prior sample information (as recommended by Hubisz *et al*., 2009) and with multiple model assumptions (admixture/no admixture and correlated/noncorrelated allele frequencies, as recommended by Pritchard *et al*., 2000). Each run consisted of a burn-in of 10^6 steps followed by 5×10^6 steps with three runs performed for each K model tested. Optimal models were assessed using L(K).

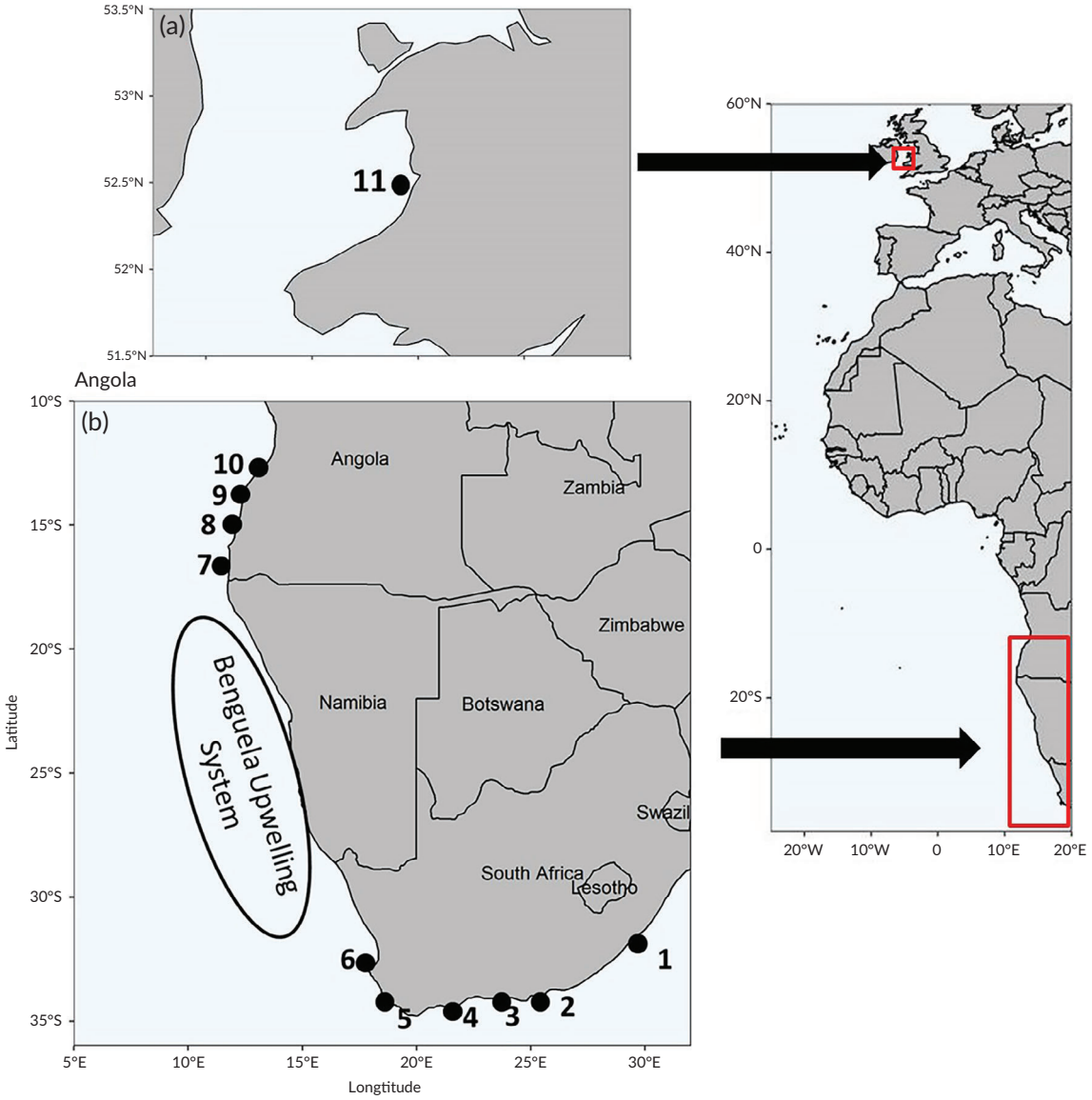

FIGURE 1 Map showing the approximate sample site locations around Southern Africa (b, sites 1–10) and Cardigan Bay (a, site 11). Numbers correspond to samples as in Table 1. The approximate location of the Benguela upwelling system between South Africa and Angola is depicted

3 | RESULTS

3.1 | MTDNA phylogenetics and phylogeography

For COI the final alignment consisted of 550 base pairs across 120 individuals (74 sequenced *de novo* in this study), identifying 24 different haplotypes with an overall haplotype diversity of 0.52 (π = 0.002). The corresponding ML tree revealed four reciprocally monophyletic clades with clear regional affinities (Figure 2). The South African (*S. emarginatum*) and Angolan (*S. cantharus*) samples each formed distinct reciprocally monophyletic clades. The UK and Portuguese samples also clustered into a reciprocally monophyletic clade, with a nonsignificant within-clade Φ_{ST} of 0.002 between these samples. Accordingly, this group is hereafter referred to as the NE Atlantic clade. The remaining Turkish and Italian sequences clustered into a highly divergent and reciprocally monophyletic clade (hereafter referred to as the Mediterranean clade) with the exception of one sequence (Genbank Accession KJ012439) which was identical to Haplotype 13 of the NE Atlantic clade.

The COI phylogeny revealed a clear hierarchical pattern with an initial divergence between the Mediterranean and the three Atlantic clades. BEAST analysis estimated the Mediterranean–Atlantic divergence to have occurred 2.1 million years BP with the subsequent divergence of the three Atlantic clades occurring around 1.3 million years BP. IMa analysis supported a similar time fame of Mediterranean–Atlantic divergence of ~2 million years BP, with subsequent divergence among the Atlantic clades estimated to have occurred ~1 million years BP.

For the control region (CR) a 436 bp sequence was aligned across 87 individuals revealing 76 polymorphic sites and 34 distinct

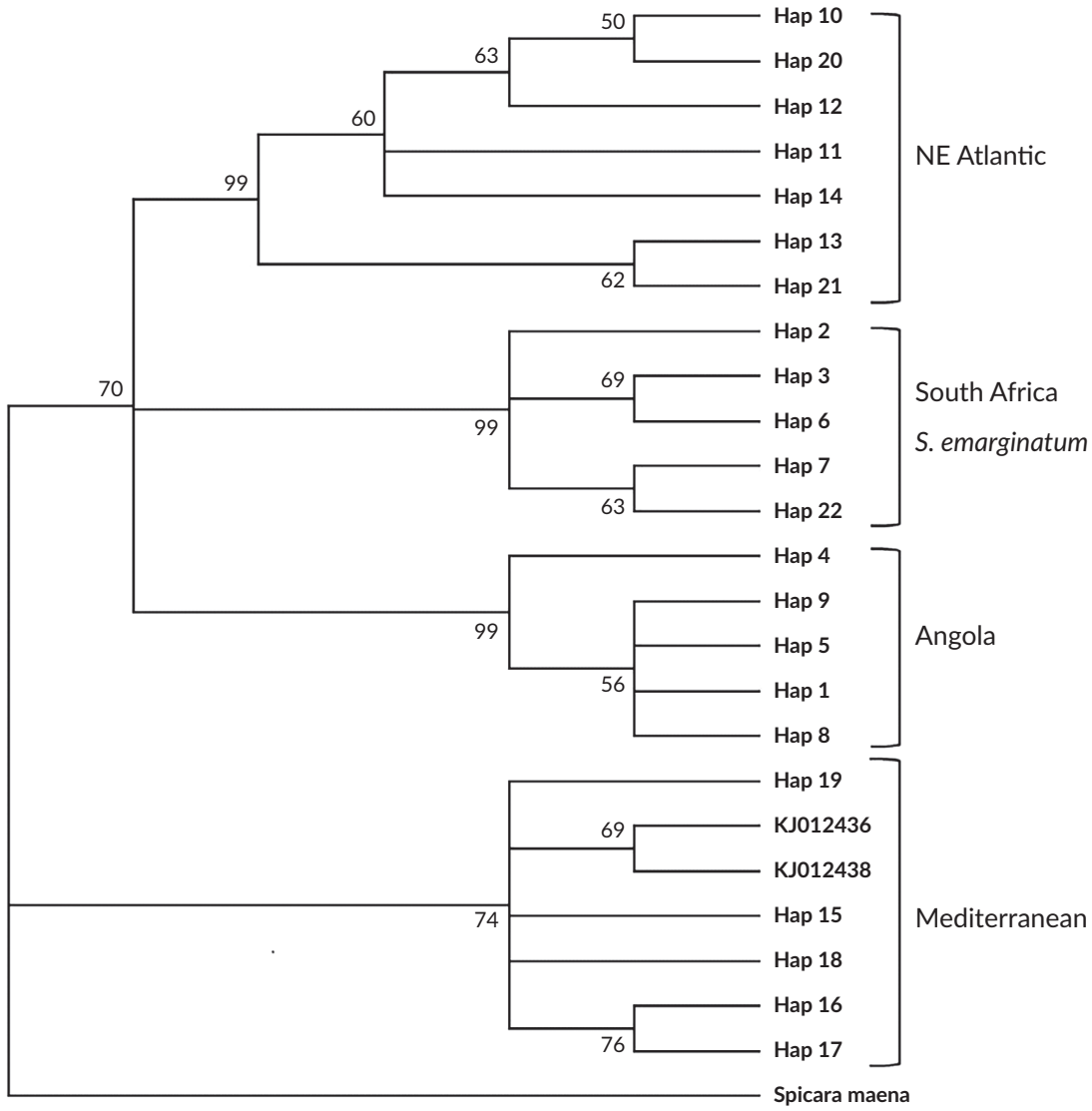

FIGURE 2 ML bootstrap consensus tree of the phylogenetic relationships among distinct COI haplotypes with *S. maena* used as an outgroup root. Node labels denote the percentage bootstrap support above the user inferred 50% cut-off

haplotypes. Overall haplotype diversity was higher (0.70) than for COI. Phylogenetic analysis resolved the same three reciprocally monophyletic and geographically disjunct Atlantic clades as for COI (Figure 3). Sequence divergence between clades was approximately 2–2.5 fold higher for CR than for the COI (Table 2).

For both the COI and CR sequences all Φ_{ST} values between samples within regions (South Africa and Angola) were nonsignificant, and so diversity and demographic test results are reported for each phylogroup (Table 3). Haplotype diversity was generally higher for the Mediterranean clade and lowest for the Angolan clade (Table 3). Mismatch distribution analyses for COI reported conformance to models of population expansions for the Angolan and NE Atlantic clades, but not for the Mediterranean and *S. emarginatum* clades (Table 3). However, as for the other Atlantic clades, the *S. emarginatum* clade did report conformance to an expansion model among CR sequences. Estimated times of expansion events were in all cases considerably more recent for the CR than COI. Significant deviations from neutral expectations (*i.e.*, Fu's *Fs* and Tajima's *D* tests) were generally found for the same phylogroup–gene region combinations for which mismatch distribution analysis supported demographic expansions (Table 3). Accordingly, Fu's *Fs* and Tajima's *D* were both not significant in the case of the Mediterranean group.

3.2 | Microsatellite analysis of South African *S. emarginatum* and Angolan *S. cantharus*

All loci were variable in each sample with the total number of alleles per locus ranging from 18 (DsaMS48) to 28 (DsaMS34), with an average of 21.75 alleles across all loci. Whilst levels of variability differed across loci, variability indices at each locus were similar across all samples (Table 4). Significant deviations from HWE were found in 14 out

of 16 locus/sample comparisons, in all cases due to heterozygote deficits, with nonsignificant test results only for DsaMS27 and Dvul84 in Mossel Bay. There was no evidence of linkage disequilibrium between any locus pair. F_{ST} values were significant in all comparisons between Angolan and South African samples, but not for comparisons within either region with a similar pattern among null allele corrected values (Table 5). Assignment of individuals using STRUCTURE supported this pattern, identifying K = 2 as the optimal model with a clear partitioning of Angolan and *S. emarginatum* individuals (Figure 4).

4 | DISCUSSION

The aim of this study was to assess macrogeographical genetic structure in *Spondyliosoma* and in doing so provide the first genetic-based assessment of the species status of *S. emarginatum*. mtDNA revealed a strong genetic structure comprising four reciprocally monophyletic clades with geographically disjunct distributions. Three of these clades occur among *S. cantharus* and are denoted according to their regional associations as Mediterranean, NE Atlantic and Angolan. The fourth clade comprised the *S. emarginatum* (South African) samples. Integration of samples from GenBank indicate that the NE Atlantic and Mediterranean clades identified here correspond to the Atlantic and Mediterranean clades described by Bargelloni *et al*. (2003). The overall phylogeny revealed the NE Atlantic and Angolan *S. cantharus* clades to be more closely related to the *S. emarginatum* clade than to the Mediterranean clade. The study therefore not only reveals hitherto unidentified lineages and pronounced genetic divergence within Atlantic waters but also reveals *S. cantharus*, as currently described, to be paraphyletic with respect to *S. emarginatum*.

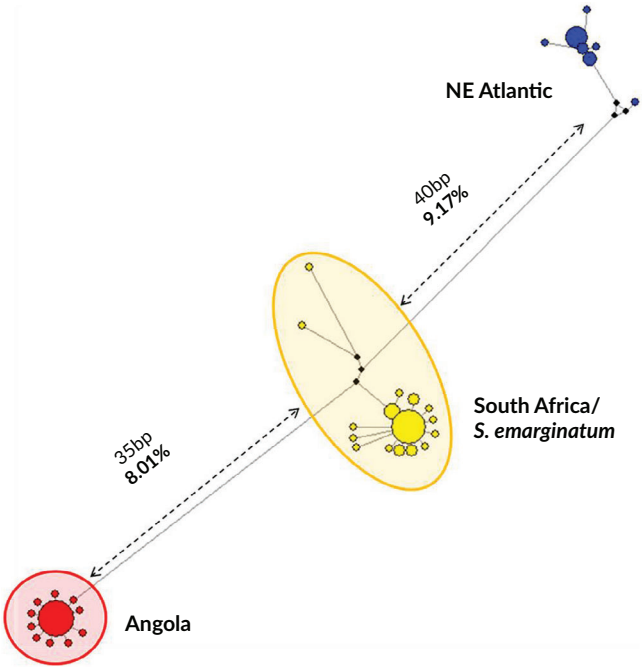

FIGURE 3 Median joining haplotype network among control region haplotypes. Node sizes are proportional to the observed abundance. Minimum number of substitution differences and percentage sequence divergence between clades are reported on the connecting branches

Range contraction and expansion events associated with the Quaternary glaciations (2.5 MYA up until the Last Glacial Maximum 26.5–19KY BP) have profoundly shaped the phylogeographic structure of many marine taxa (Maggs *et al*., 2008). Chronic isolation in disjunct areas of persistence (glacial refuge) has contributed to genetic divergence, regionally associated phylogroups and even speciation. Signatures of such processes are strikingly evident in the phylogeography of *Spondyliosoma*. The data indicate that following an earlier divergence from the Mediterranean clade (~2 million years BP) the three Atlantic clades [NE Atlantic, Angola and *S. emarginatum* (South Africa)] diverged from each other around the same time (~1 million years BP). This chronology aligns well with established patterns of divergence between the Atlantic and Mediterranean (Bargelloni *et al*., 2003; Patarnello *et al*., 2007), within the eastern Atlantic (Durand *et al*., 2005, 2013; Miralles *et al*., 2014; Sala-Bozano *et al*., 2009) and across the Benguela upwelling system (Henriques *et al*., 2014, 2016) reported for other marine taxa. The Atlantic clades also exhibited signatures (unimodal mismatch distributions and negative Fu's F_S and Tajima's D indices) of population demographic expansions after periods of reduced population sizes. Similar patterns in other species have been linked to declines and expansions during Pleistocene glacial and interglacial periods, respectively (Debes *et al*., 2008). There is growing appreciation that the accuracy of estimated times of demographic events may in many cases be compromised by the time-dependency of user-inferred mutation rates (Grant, 2015; Ho et al., 2005, 2007; Hoareau, 2015; McKeown *et al*., 2019). Furthermore, the protoygynous hermaphrodite life cycle of both *Spondyliosoma* species (Goncalves & Erzini, 2000; Mouine *et al*., 2010) may confound estimates assuming a constant ¼Ne for mtDNA compared to nuclear loci (Coscia *et al*., 2016). However, while

TABLE 2 Mean sequence divergence between regional clades for COI (below diagonal) and CR (above diagonal)

	South Africa	Angola	NE Atlantic	Mediterranean
South Africa	*0.004/0.007*	0.046	0.053	–
Angola	0.023	*0.003/0.003*	0.052	–
NE Atlantic	0.020	0.020	*0.004/0.006*	–
Mediterranean	0.032	0.029	0.029	*0.005*

Note: Diagonal values report the mean sequence divergence among haplotypes within clades with the first value for COI, and the second for CR. Divergence calculated using K2P (Kimura, 1980). COI, cytochrome oxidase I; CR, control region.

TABLE 3 Summary indices of regional mtDNA variability and demographic tests for COI and CR, including number of individuals sequenced (*N*), number of haplotypes (*Nhap*), haplotype (*h*) and nucleotide (π) diversity, the probability of mismatch distribution conformance to models of expansion (Mismatch *P*), the mismatch mutation parameter (τ) and associated estimates of time of expansion, Tajima's *D* and Fu's *Fs* indices and associated *P* values of significant deviation from neutral expectations (*critical *P* for Fu's *Fs* = 0.02)

	S. emarginatum		*S. cantharus*					
	South Africa		Angola		NE Atlantic		Mediterranean	
Fragment	COI	CR	COI	CR	COI	CR	COI	
N–*Nhap*	18–5	40–16	41–5	31–11	31–7	16–7	21–5	
h–π	0.557–0.0016	0.769–0.0055	0.188–0.0004	0.548–0.0018	0.579–0.0019	0.792–0.0058	0.757–0.0034	
Mismatch *P*	0.923	0.322	0.407	0.942	0.840	0.392	**0.023**	
τ (expansion time in years)	1.393 (211KYA)	1.322 (27KYA)	3 (454 KYA)	0.801 (16KYA)	2.164 (327 KYA)	1.125 (23KYA)	2.45	
Tajima's *D* (*P*)	−0.487 (0.3)	**−2.4 (P < 0.001)**	**−1.9 (0.004)**	**−2.4 (< 0.001)**	−0.8 (0.217)	−1.6 (**0.047**)	0.41 (0.693)	
Fu's *Fs* (*P*)	−1.2 (0.1)	**−8.1 (<0.001)**	**−5.1 (<0.001)**	**−10.4 (<0.001)**	−2.5 (0.041)*	−1 (0.268)	0.6 (0.649)	

Note: COI, cytochrome oxidase I; CR, control region.

TABLE 4 Summary of nuclear genetic variation for the Mossel Bay, False Bay, Namibe and Benguela samples described using allele number (N_a), allelic richness (A_R), expected (H_E) and observed (H_O) heterozygosities and probability of conformance to the Hardy–Weinberg equilibrium (P_{HWE})

Locus	Indices	MOB	FAL	NAM	BEN
DsaMS27	N_a	16	15	13	13
	A_R	13.523	12.165	10.082	10.282
	H_E	0.908	0.896	0.858	0.854
	H_O	0.952	0.8	0.643	0.64
	P_{HWE}	0.45	**0.043**	**0.004**	**0.001**
DsaMS34	N_a	12	13	20	19
	A_R	10.918	10.737	14.998	15.27
	H_E	0.832	0.81	0.917	0.92
	H_O	0.389	0.524	0.692	0.76
	P_{HWE}	**<0.001**	**0.039**	**<0.001**	**0.017**
DsaMS48	N_a	9	14	11	11
	A_R	9	11.782	9.78	10.153
	H_E	0.814	0.884	0.88	0.861
	H_O	0.571	0.65	0.417	0.579
	P_{HWE}	**0.005**	**0.001**	**<0.001**	**0.004**
Dvul84	N_a	10	13	8	4
	A_R	9.548	10.312	5.545	3.962
	H_E	0.836	0.818	0.57	0.529
	H_O	0.667	0.63	0.556	0.348
	P_{HWE}	0.245	**0.004**	**0.049**	**0.01**

Note. BEN, Benguela; FAL, False Bay; MOB, Mossel Bay; NAM, Namibe; (N_a), allele number; A_R, allelic richness; H_E, expected heterozygosity; H_O observed heterozygosity; P_{HWE}, probability of conformance to the Hardy–Weinberg equilibrium.

mindful of such considerations, the resolved phylogeographic structure clearly supports a prominent role for glacial vicariance in shaping the divergence and regional association of the clades. The lack of overlap detected, albeit only resolved at a macrogeographical scale in this study, points to the role of some contemporary isolating mechanisms which may include factors such as geographical distance, biogeographic barriers, adaptation, and life history.

The BUS has been identified as the prominent driver of divergence across the SE Atlantic region, generating isolated populations as well as endemic lineages and species in South African waters (Gwilliam et al., 2018; Henriques et al., 2012, 2014; Reid *et al*., 2016; Sala-Bozano *et al*., 2009; Schwaninger, 2008; Teske *et al*., 2011). However, for some groups the BUS exhibits varying levels of historical and/or recurrent permeability permitting bi-directional (Henriques *et al*., 2015) or asymmetric (Healey *et al*., 2017) gene flow. Here the mtDNA divergence of South African *S. emarginatum* from the geographically most close Angolan "*S. cantharus*" reveals no evidence of such permeability and confirms the long-term isolation of these groups. The *S. emarginatum* mtDNA also satisfies the reciprocal monophyly criterion of the phylogenetic species concept. In addition, ratios

TABLE 5 Pairwise nuclear F_{ST} between Mossel Bay, False Bay, Namibe and Benguela samples calculated with (above diagonal) and without (below diagonal) correction for null alleles

	MOB	FAL	NAM	BEN
MOB	–	−0.003	0.093	0.087
FAL	−0.006	–	0.103	0.094
NAM	0.096	0.104	–	−0.006
BEN	0.106	0.109	−0.009	–

Note: BEN, Benguela; FAL, False Bay; MOB, Mossel Bay; NAM, Namibe.

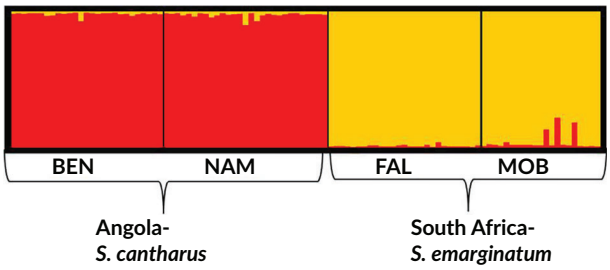

FIGURE 4 Bar plot showing the differential assignment of individuals of *S. cantharus* from Angolan sites (Benguela and Namibe) and *S. emarginatum* from South African sites (False Bay and Mossel Bay) under the optimum model of $K = 2$ using LOCPRIOR. BEN, Benguela; FAL, False Bay; MOB, Mossel Bay; NAM, Namibe

of within/between clade sequence divergence exceed barcode gaps for species delineation suggested for fish (Ward *et al*., 2009). In light of criticisms of species delimitation based solely on mtDNA (reviewed in Hudson & Coyne, 2002; Hudson & Turelli, 2003; Moritz & Cicero, 2004; Sites & Marshall, 2004) confirmatory evidence from other approaches is recommended (Funk & Omland, 2003; Gwilliam *et al*., 2018). Nuclear F_{ST} were also significant in all comparisons between the *S. emarginatum* and Angola samples (with no differentiation within ether group), confirming restricted biparental gene flow between the groups. STRUCTURE based clustering analyses also provided no evidence of hybrids or migrants between these groups. In addition to historical and recurrent genetic isolation, morphological differences between the larval stages of *S. emarginatum* and *S. cantharus* are reported (Beckley & Buxton, 1989; Russell 1976). Furthermore, while adults of both taxa are morphologically similar, *S. emarginatum* appears to be smaller (maximum size 45 cm total length (TL), average size 25 cm TL) than *S. cantharus* (maximum size 60 cm TL, average size 30 cm TL) (Bauchot & Smith 1984). Such size differences against a background of general adult morphological similarity were interpreted as evidence of ecological diversification across the BUS in *Atractoscion aequidens* (Henriques *et al*., 2016). Collectively, the data on *Spondyliosoma* support the genetic, ecological and morphological differentiation of *S. emarginatum* from *S. cantharus* and its recognition as a distinct species.

West African fishes have been less studied using genetic approaches than their northern and southern (*i.e*., South African) counterparts (Durand *et al*., 2013). However, studies have revealed considerable phylogeographic diversity and genetic breaks within this region, supporting a west African glacial refuge (Maggs *et al*., 2008) with more recent studies specifying Angola as a candidate refugial area (Reid *et al*., 2016). The broad phylogeographic structure would be compatible with derivation of the Angolan clade from such an African refuge while the NE Atlantic clade may have emanated from one of the established NE Atlantic refuges (*e.g*., Iberia; Maggs *et al*., 2008). The limited sampling restricts our information as to the ranges of the Angolan and NE Atlantic clades in the waters north of the BUS. However, the data confirm that the cold water BUS represents a southern boundary to the Angolan group. Though speculative at this point, the similarly cold water Canary current may serve as a northern boundary around Senegal (NW Africa), as observed in other groups (Reid *et al*., 2016). If this is the case, the Angolan clade may represent an isolated phylogeographic remnant.

The results of this study have a number of important systematic implications. On the one hand, the data confirm the validity and genetic integrity of *S. emarginatum* as a distinct species. On the other hand, the study resolves similarly and even more highly divergent lineages among individuals all currently described as *S. cantharus*. The question therefore arises as to whether *S. cantharus* comprises a cryptic species complex wherein the various lineages represent established/incipient species. The coarse grain sampling of this study restricts fundamental inferences as to the respective ranges and reproductive isolation among the lineages. In this context, it is interesting that one of the sequences obtained from samples collected at an Italian fish market (Armani *et al*., 2015) clustered with the NE Atlantic clade. While the exact provenance of this specimen is unknown it could point to secondary contact between NE Atlantic and Mediterranean clades within the Mediterranean, the likes of which have been reported for other taxa (Fruciano *et al*., 2011). Such secondary contact may in some cases result in uninhibited gene flow (Sala-Bozano *et al*., 2009). However, gene flow may also continue to be restricted in sympatry, or beyond secondary contact/hybrid zones, so that pre-existing genetic differences are preserved, at least in some areas (Unckless & Orr, 2009). A robust multilocus (nuclear) assessment combining morphological data and particularly detailing interactions among lineages is needed to help determine the full diversity within *Spondyliosoma* and the most adequate biological and taxonomic status.

AUTHOR CONTRIBUTIONS

N.J.M.K. analysed and wrote the manuscript. M.P.G. collected the data. All authors contributed to editing and approved the manuscript.

ORCID

Niall J. McKeown https://orcid.org/0000-0002-7351-5553

REFERENCES

Armani, A., Guardone, L., Castigliego, L., D'Amico, P., Messina, A., Malandra, R., … Guidi, A. (2015). DNA and mini-DNA barcoding for the identification of porgies species (family Sparidae) of commercial interest on the international market. *Food Control*, *50*, 589–596.

Bargelloni, L., Alarcon, J. A., Alvarez, M. C., Penzo, E., Magoulas, A., Reis, C., & Patarnello, T. (2003). Discord in the family Sparidae (Teleostei): Divergent phylogeographical patterns across the Atlantic-Mediterranean divide. *Journal of Evolutionary Biology*, *16*, 1149–1158.

Bauchot, M., & Smith, M. (1984). *FAO species identification sheets for fishery purposes, Western Indian Ocean, fishing area 51*. Rome, Italy: FAO.

Bauchot, M. L., & Hurau, J. C. (1986). Sparidae. In *Fishes of the north – Eastern Atlantic and Mediterranean*. Paris, France: UNESCO.

Beckley, L., & Buxton, C. (1989). Underwater observations of reef fish in and around Algoa-Bay, South Africa. *Transactions of the Royal Society of South Africa*, *47*, 29–38.

Bermingham, E. P., McCafferty, S. S., & Martin, A. P. (1997). Fish biogeography and molecular clocks: Perspectives from the Panamanian Isthmus. In *Molecular systematics of fishes*. Balboa, Republic of Panama: Academic Press.

Chapuis, M. P., & Estoup, A. (2007). Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, *24*, 621–631.

Correia, A. T., Gomes, P., Goncalves, J. M. S., Erzini, K., & Hamer, P. A. (2012). Population structure of the black seabream *Spondyliosoma cantharus* along the south-west Portuguese coast inferred from otolith chemistry. *Journal of Fish Biology*, *80*, 427–443.

Coscia, I., Chopelet, J., Waples, R. S., Mann, B. Q., & Mariani, S. (2016). Sex change and effective population size: Implications for population genetic studies in marine fish. *Heredity*, *117*, 251–258.

Coscia, I., Vogiatzi, E., Kotoulas, G., Tsigenopoulos, C. S., & Mariani, S. (2012). Exploring neutral and adaptive processes in expanding populations of gilthead sea bream, *Sparus aurata* L., in the north-East Atlantic. *Heredity*, *108*, 537–546.

Costa, F. O., Landi, M., Martins, R., Costa, M. H., Costa, M. E., Carneiro, M., … Carvalho, G. R. (2012). A ranking system for reference libraries of DNA barcodes: Application to marine fish species from Portugal. *PLoS One*, *7*, e35858.

Debes, P. V., Zachis, F. E., & Hanel, R. (2008). Mitochondrial phylogeography of the European sprat (*Sprattus sprattus* L., Clupeidae) reveals isolated climatically vulnerable populations in the Mediterranean Sea and range expansion in the Northeast Atlantic. *Molecular Ecology*, *17*, 3873–3888.

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, *7*.

Durand, J. D., Guinand, B., Dodson, J. J., & Lecomte, F. (2013). Pelagic life and depth: Coastal physical features in West Africa shape the genetic structure of the Bonga Shad *Ethmalosa fimbriata*. *PLoS One*, *8*.

Durand, J. D., Tine, M., Panfili, J., Thiaw, O. T., & Lae, R. (2005). Impact of glaciations and geographic distance on the genetic structure of a tropical estuarine fish, *Ethmalosa fimbriata* (Clupeidae, S. Bowdich, 1825). *Molecular Phylogenetics and Evolution*, *36*, 277–287.

Excoffier, L., & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and windows. *Molecular Ecology Resources*, *10*, 564–567.

Fruciano, C., Hanel, R., Debes, P. V., Tigano, C., & Ferrito, V. (2011). Atlantic-Mediterranean and within-Mediterranean molecular variation in *Coris julis* (L. 1758) (Teleostei, Labridae). *Marine Biology*, *158*, 1271–1286.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, *147*, 915–925.

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

Goncalves, J. M. S., & Erzini, K. (2000). The reproductive biology of *Spondyliosoma cantharus* (L.) from the SW coast of Portugal. *Scientia Marina*, *64*, 403–411.

Goudet, J. (1995). FSTAT (version 1.2): A computer program to calculate F-statistics. *Journal of Heredity*, *86*, 485–486.

Grant, W. S. (2015). Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. *Journal of Heredity*, *106*, 333–346.

Grant, W. S., & Bowen, B. W. (1998). Shallow population histories in deep evolutionary lineages of marine fishes: Insights from sardines and anchovies and lessons for conservation. *Journal of Heredity*, *89*, 415–426.

Gwilliam, M. P., Winkler, A. C., Potts, W. M., Santos, C. V., Sauer, W. H. H., Shaw, P. W., & McKeown, N. J. (2018). Integrated genetic and morphological data support eco-evolutionary divergence of Angolan and South African populations of Diplodus hottentotus. *Journal of Fish Biology*, *92*(4), 1163–1176. https://doi.org/10.1111/jfb.13582

Hauser, L. & Carvalho, G. R. (2008). Paradigm shifts in marine fisheries genetics: ugly hypotheses slain by beautiful facts. *Fish and Fisheries9* (4), 333–362.

Healey, A. J. E., McKeown, N. J., Potts, W. M., de Beer, C. L., Sauer, W., & Shaw, P. W. (2017). Phylogeny of the *Sepia officinalis* species complex in the eastern Atlantic extends the known distribution of *Sepia vermiculata* across the Benguela upwelling region. *African Journal of Marine Science*, *39*, 307–313.

Henriques, R., Potts, W. M., Santos, C. V., Sauer, W. H. H., & Shaw, P. W. (2014). Population connectivity and phylogeography of a coastal fish, *Atractoscion aequidens* (Sciaenidae), across the Benguela current region: Evidence of an ancient Vicariant event. *PLoS One*, *9*(2), e87907.

Henriques, R., Potts, W. M., Sauer, W. H., Santos, C. V., Kruger, J., Thomas, J. A., & Shaw, P. W. (2016). Molecular genetic, life-history and morphological variation in a coastal warm-temperate sciaenid fish: Evidence for an upwelling-driven speciation event. *Journal of Biogeography*, *43*, 1820–1831.

Henriques, R., Potts, W. M., Sauer, W. H. H., & Shaw, P. W. (2012). Evidence of deep genetic divergence between populations of an important recreational fishery species, *Lichia amia* L. 1758, around southern Africa. *African Journal of Marine Science*, *34*, 585–591.

Henriques, R., Potts, W. M., Sauer, W. H. H., & Shaw, P. W. (2015). Incipient genetic isolation of a temperate migratory coastal sciaenid fish (*Argyrosomus inodorus*) within the Benguela cold current system. *Marine Biology Research*, *11*, 423–429.

Hey, J. (2010). Isolation with migration models for more than two populations. *Molecular Biology and Evolution*, *27*, 905–920.

Hey, J., & Nielsen, R. (2004). Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of Drosophila pseudoobscura and D-persimilis. *Genetics*, *167*, 747–760.

Ho, S. Y. W., Phillips, M. J., Cooper, A., & Drummond, A. J. (2005). Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Molecular Biology and Evolution*, *22*, 1561–1568.

Ho, S. Y. W., Shapiro, B., Phillips, M. J., Cooper, A., & Drummond, A. J. (2007). Evidence for time dependency of molecular rate estimates. *Systematic Biology*, *56*, 515–522.

Hoareau, T. B. (2015). Late glacial demographic expansion motivates a clock overhaul for population genetics. *Systematic Biology*, *65*, 449–464.

Hubisz, M. J., Falush, D., Stephens, M., & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, *9*, 1322–1332.

Hudson, R. R., & Coyne, J. A. (2002). Mathematical consequences of the genealogical species concept. *Evolution*, *56*, 1557–1565.

Hudson, R. R., & Turelli, M. (2003). Stochasticity overrules the "three-times rule": Genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution*, *57*, 182–190.

Keskin, E., & Atar, H. H. (2013). DNA barcoding commercially important fish species of Turkey. *Molecular Ecology Resources*, *13*, 788–797.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *Journal of Molecular Evolution*, *16*, 111–120.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*, 1870–1874.

Maggs, C. A., Castilho, R., Foltz, D., Henzler, C., Jolly, M. T., Kelly, J., … Wares, J. (2008). Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology*, *89*, S108–S122.

McKeown, N. J., Arkhipkin, A. I., & Shaw, P. W. (2019). Genetic analysis reveals historical and contemporary population dynamics in the longfin squid *Doryteuthis gahi*: Implications for cephalopod management and conservation. *ICES Journal of Marine Science*, *76*, 1019–1027.

Miralles, L., Juanes, F., Pardinas, A. F., & Garcia-Vazquez, E. (2014). Paleoclimate shaped bluefish structure in the northern hemisphere. *Fisheries*, *39*, 578–586.

Moritz, C., & Cicero, C. (2004). DNA barcoding: Promise and pitfalls. *PLoS Biology*, *2*, 1529–1531.

Mouine, N., Ktari, M. H., & Chakroun-Marzouk, N. (2010). Reproductive characteristics of *Spondyliosoma cantharus* (Linnaeus, 1758) in the Gulf of Tunis. *Journal of Applied Ichthyology*, *27*, 827–831.

Neves, A., Vieira, A. R., Sequeira, V., Paiva, R. B., & Gordo, L. S. (2017). Modelling the growth of a protogynous sparid species, *Spondyliosoma cantharus* (Teleostei: Sparidae). *Hydrobiologia*, *797*, 265–275.

Neves, A., Vieira, A. R., Sequeira, V., Paiva, R. B., & Gordo, L. S. (2018). Phenotypic changes in the body of black seabream, *Spondyliosoma cantharus* (Teleostei: Sparidae), along the eastern Atlantic. *Estuarine, Coastal and Shelf Science*, *214*, 31–37.

Neves, A., Vieira, A. R., Sequeira, V., Paiva, R. B., Janeiro, A. I., Gaspar, L. M., & Gordo, L. S. (2019). Otolith shape and isotopic ratio analyses as a tool to study *Spondyliosoma cantharus* population structure. *Marine Environmental Research*, *143*, 93–100.

Nylander, J.A.A., (2004). MrModel test (Uppsala Univeristy). Version 2.

Patarnello, T., Volckaert, F. A., & Castilho, R. (2007). Pillars of Hercules: Is the Atlantic–Mediterranean transition a phylogeographical break? *Molecular Ecology*, *16*, 4426–4444.

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in excel. population genetic software for teaching and research. *Molecular Ecology Notes*, *6*, 288–295.

Perez, L., Infante, C., Ponce, M., Crespo, A., Zuasti, E., Funes, V., … Manchado, M. (2008). Characterization of eight microsatellite markers in the white sea bream, *Diplodus sargus* (Teleostei, Sparidae). *Molecular Ecology Resources*, *8*, 1291–1293.

Pinder, A. C., Velterop, R., Cooke, S. J., & Britton, J. R. (2017). Consequences of catch-and-release angling for black bream *Spondyliosoma cantharus*, during the parental care period: Implications for management. *ICES Journal of Marine Science*, *74*, 254–262.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

Reid, K., Hoareau, T. B., Graves, J. E., Potts, W. M., dos Santos, S. M. R., Klopper, A. W., & Bloomer, P. (2016). Secondary contact and asymmetrical gene flow in a cosmopolitan marine fish across the Benguela upwelling zone. *Heredity*, *117*, 307–315.

Reiss, H., Hoarau, G., Dickey-Collas, M., & Wolff, W. J. (2009). Genetic population structure of marine fish: Mismatch between biological and fisheries management units. *Fish and Fisheries*, *10*, 361–395.

Rogers, A. R., & Harpending, H. (1992). Population-growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution*, *9*, 552–569.

Roques, S., Galarza, J. A., Macpherson, E., Turner, G. F., & Rico, C. (2007). Isolation and characterization of nine polymorphic microsatellite markers in the two-banded sea bream (*Diplodus vulgaris*) and cross-species amplification in the white sea bream (*Diplodus sargus*) and the saddled bream (*Oblada melanura*). *Molecular Ecology Notes*, *7*, 661–663.

Rousset, F. (2008). GENEPOP' 007: A complete re-implementation of the GENEPOP software for windows and Linux. *Molecular Ecology Resources*, *8*, 103–106.

Russell, F. S. (1976). *The eggs and planktonic stages of British marine fishes*. London, UK: Academic Press.

Sala-Bozano, M., Ketmaier, V., & Mariani, S. (2009). Contrasting signals from multiple markers illuminate population connectivity in a marine fish. *Molecular Ecology*, *18*, 4811–4826.

Sambrook, H. (1989). *Molecular cloning: A laboratory manual*. New York, NY: Cold Spring Harbor.

Schwaninger, H. R. (2008). Global mitochondrial DNA phylogeography and biogeographic history of the antitropically and longitudinally disjunct marine bryozoan *Membranipora membranacea* L. (Cheilostomata): Another cryptic marine sibling species complex? *Molecular Phylogenetics and Evolution*, *49*, 893–908.

Sites, J. W., & Marshall, J. C. (2004). Operational criteria for delimiting species. *Annual Review of Ecology, Evolution, and Systematics*, *35*, 199–227.

Tajima, F. (1989). Statistical-method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, *123*, 585–595.

Teske, P. R., von der Heyden, S., McQuaid, C. D., & Barker, N. P. (2011). A review of marine phylogeography in southern Africa. *South African Journal of Science*, *107*, 43–53.

Unckless, R. I., & Orr, H. A. (2009). Dobzhansky-Muller incompatibilities and adaptation to a shared environment. *Heredity*, *102*, 214–217.

Waples, R. S., & Gaggiotti, O. (2006). What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, *15*(6), 1419–1439.

Ward, R. D., Hanner, R., & Hebert, P. D. N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *Journal of Fish Biology*, *74*, 329–356.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: McKeown NJ, Gwilliam MP, Healey AJE, et al. Deep phylogeographic structure may indicate cryptic species within the Sparid genus *Spondyliosoma*. *J Fish Biol*. 2020;96:1434–1443. https://doi.org/10.1111/jfb.14316