

Combined COI barcode-based methods to avoid mislabelling of threatened species of deep-sea skates

L. Carugati¹ , R. Melis¹ , A. Cariani² , A. Cau¹ , V. Crobe² , A. Ferrari² , M. C. Follesa¹ , M. L. Geraci^{3,4} , S. P. Iglésias⁵ , P. Pesci¹ , F. Tinti²  & R. Cannas¹ 

¹ Department of Life and Environmental Sciences, University of Cagliari, Cagliari, Italy

² Department of Biological, Geological and Environmental Sciences (BiGeA), University of Bologna, Bologna, Italy

³ Department of Biological, Geological and Environmental Sciences (BiGeA) – Marine biology and fisheries laboratory, University of Bologna, Fano (PU), Italy

⁴ Institute for Biological Resources and Marine Biotechnologies (IRBIM), National Research Council (CNR), Mazara del Vallo (TP), Italy

⁵ Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Station Marine de Concarneau, Concarneau, France

Keywords

conservation; deep sea; *Dipturus*; DNA barcoding; species delimitation; specimen assignment; species identification; skates.

Correspondence

Rita Cannas, Department of Life and Environmental Sciences, University of Cagliari, Via T. Fiorelli 1, 09126 Cagliari, Italy. Tel. +390706758004
Email: rcannas@unica.it

Editor: Jeff Johnson
Associate Editor: Matt Fujita

L. Carugati and R. Melis are equally contributed.

Received 07 December 2020; accepted 20 May 2021

doi:10.1111/acv.12716

Abstract

Skates are characterised by conservative body morphology which hampers identification and leads to frequent taxonomic confusion and market mislabelling. Accurate specimen classification is crucial for reliable stock assessments and effective conservation plans, otherwise the risk of extinction could be unnoticed. The misclassification issue is evident for the genus *Dipturus*, distributed worldwide, from the continental shelf and slope to the deep sea. In this study, barcode cytochrome oxidase I gene (COI) sequences were used along with species delimitation and specimen assignment methods to improve taxonomy and zoogeography of species of conservation interest inhabiting the Atlantic Ocean and Mediterranean Sea. In this study, we provided new evidence of the occurrence of *D. nidarosiensis* in the Central-Western Mediterranean Sea and the lack of Atlantic-Mediterranean genetic divergence. The Atlantic endangered species *D. laevis* and *D. batis* clustered together under the same molecular operational taxonomic unit (MOTU) with any delimitation methods used, while the assignment approach correctly discriminated specimens into the two species. These results provided evidence that the presence of the barcode gap is not an essential predictor of identification success, but the use of different approaches is crucially needed for specimen classification, especially when threshold- or tree-based methods result less powerful. The analyses also showed how different putative, vulnerable, species dwelling across South-Western Atlantic and South-Eastern Pacific are frequently misidentified in public sequence repositories. Our study emphasised the limits associated to public databases, highlighting the urgency to verify and implement the information deposited therein in order to guarantee accurate species identification and thus effective conservation measures for deep-sea skates.

Introduction

Skates are cartilaginous fish of the family Rajidae (order Rajiformes) which includes more than 150 described species, belonging to 17 genera (Last *et al.*, 2016a). Most of them are benthic organisms with restricted geographical distributions, living between the shoreline and >4000 m depth (Orlov, Cotton and Byrkjedal, 2006; Ebert and Compagno, 2007). Despite their apparent restrictive habitat preferences (e.g., soft bottom substrates), skates exhibit a high degree of biodiversity and endemism (Ebert and Compagno, 2007).

However, due to their conserved dorsoventrally flattened body morphology, the identification of skates based solely on morphological characters can be problematic (Geraci *et al.*, 2017).

The identification of species boundaries is fundamental for undertaking effective conservation and management measures. The difficulties encountered among this family have led to a high degree of mislabelling in scientific trawl surveys (Cariani *et al.*, 2017) and commercial landings (Iglésias *et al.*, 2010; Griffiths *et al.*, 2013). Skates are rarely specifically targeted by fisheries, but their high susceptibility to by-

catch make them some of the most vulnerable marine fishes (Stevens *et al.*, 2000; Dulvy *et al.*, 2016). Accurate species identification of individuals is thus needed and is critical for reliable conservation plans and stock assessments, otherwise declines, especially of threatened taxa, could be unnoticed.

The misclassification problem is particularly evident in the genus *Dipturus* (Griffiths *et al.*, 2010; Iglésias *et al.*, 2010) which includes more than 40 species (Last *et al.*, 2016a), which are widely distributed from temperate to tropical regions, from continental shelves and slopes to the deep sea (Ebert and Compagno, 2007; Follesa *et al.*, 2019). According to the current state of knowledge (Last *et al.*, 2016a), in the European waters (considering Mediterranean and Black Sea, North-Eastern Atlantic Ocean), *Dipturus* currently encompasses three species of conservation concern: the longnosed skate, *Dipturus oxyrinchus* (Linnaeus, 1758), the Norwegian skate, *Dipturus nidarosiensis* (Storm, 1881), and the common skate, *Dipturus batis* (Linnaeus, 1758).

The longnosed skate is distributed in both the Mediterranean and North-Eastern Atlantic, to depths of 900 m (Serena, 2005; Ebert and Stehmann, 2013; Ellis *et al.*, 2015; Mulas *et al.*, 2015). The Norwegian skate was thought to inhabit only the North-Eastern Atlantic and Bay of Biscay (Priede *et al.*, 2010; Rodríguez-Cabello, Pérez and Sánchez, 2013), but its occurrence has been recently reported in different areas of the Western Mediterranean (Cannas *et al.*, 2010; Cariani *et al.*, 2017; Ramirez-Amaro *et al.*, 2017; Carbonara *et al.*, 2019). Their presence in the Mediterranean Sea has been further confirmed by the recent finding of paintings in manuscripts by Cuvier and Valenciennes (Ms-535), unpublished in their *Histoire naturelle des poissons*, published between 1829 and 1849. The two original paintings represent both *D. oxyrinchus* and *D. nidarosiensis* collected off Nice (France) in 1829 (Figure 1). For both species, it is still debated the potential genetic distinctiveness and connectivity between Atlantic and Mediterranean populations (Griffiths *et al.*, 2011; Ramirez-Amaro *et al.*, 2017; Melis *et al.*, 2020). The low genetic differentiation among Atlantic and Mediterranean *D. oxyrinchus* specimens seems to exclude the possibility that the two basins host totally isolated populations or even different species as proposed in the past (Melis *et al.*, 2020 and references therein), but additional data are needed to confirm this picture also for *D. nidarosiensis*.

Previous molecular studies (Griffiths *et al.*, 2010; Iglésias *et al.*, 2010) pointed out that *D. batis* is a complex of two cryptic species: the flapper skate, *D. intermedius* (Parnell, 1937) and the common blue skate, *D. batis* (Linnaeus, 1758), also sometimes referred to as *D. cf. flossada* (Risso, 1826). Considering the long historical taxonomic confusion under the name *Dipturus batis*, in the present study we use the name '*D. Batis*' as defined by Last *et al.* (2016a,b) and Fricke *et al.*, (2020), to designate the smaller of the two species formerly confused in the common skate complex, synonym to the definition of *D. cf. flossada* by Iglésias *et al.* (2010).

Both species were historically present in most of the European Atlantic and Mediterranean Sea waters, but nowadays they are mostly restricted around the British Isles: *D.*

intermedius mainly appears along the Western and Northern coasts of Scotland (Neat *et al.*, 2015), while *D. batis* dwells mostly in the Celtic Sea (Lynghammar *et al.*, 2014). Both species are still very occasionally found in other parts of the North Atlantic Ocean.

Previous studies (Compagno and Ebert, 2007) reported that *D. batis* was often confused with the barndoor skate, *D. laevis* (Mitchill, 1818), which is distributed in the North-Western Atlantic (Leavitt *et al.*, 2018), usually caught as bycatch and classified as 'Endangered' (EN), and with the slime skate, *D. pullopunctatus* (Smith, 1964), which is a common endemic Southern African skate, classified at 'Least Concern' (LC) in the International Union for Conservation of Nature (IUCN) Red List (Pollom *et al.*, 2020). In addition, the slime skate presents an overlap in the geographical distribution with *D. nidarosiensis*, which has been found also off South-Eastern Africa and frequently confused with other skates (Last *et al.*, 2016a; Geraci *et al.*, 2019).

Reporting species-specific landings became European law in 2009 (EC No 43/2009); however, such legislation requires an unambiguous classification of the individuals. DNA barcoding has been proven a smart and integrative taxonomic toolbox that can effectively contribute to species identification among closely related skates (Serra-Pereira *et al.*, 2011; Lynghammar *et al.*, 2014; Frodella *et al.*, 2016; Cariani *et al.*, 2017). This approach relies on the concept that the genetic variation, that is the 'barcode gap', between species is larger than that within species (although this term should be treated with caution, see e.g., Moritz and Cicero 2004; Meyer and Paulay, 2005). Previous studies demonstrated the effectiveness of COI to uncover cryptic diversity and to contribute to solve doubts regarding species-complex in skates (de Astarloa *et al.*, 2008; Smith *et al.*, 2008; Griffiths *et al.*, 2010; Coulson *et al.*, 2011; Pasolini *et al.*, 2011; Cariani *et al.*, 2017).

Currently, one of the main limitations in skate species identification appears to be the invalid and/or outdated information contained in public databases (Cerutti-Pereyra *et al.*, 2012). Delimitation methods and specimen assignment approaches have proved to be helpful in solving species identification issues, especially in the presence of taxonomically difficult lineages (Piemontese *et al.*, 2020). In integrative taxonomic and systematic studies, species delimitation methods are increasingly applied to delineate species-level entities and ascertain the number of species in a sample (Dellicour and Flot, 2018 and references therein). Whereas, species assignment methods have proved to be helpful in the reliable assignment of an unknown query sequence to its correct species, which remains one of the most important methodological problem for DNA barcoding (Zhang *et al.*, 2012a). Here, we highlighted how the joint use of these two approaches could provide important data in outlining species relationships, improve the accuracy of specimen assignment and help to feed databases with correctly assigned entries (Delrieu-Trottin *et al.*, 2020).

In this study, DNA sequences were used along with two complementary approaches (species delimitation and specimen assignment methods) in order to (1) provide new

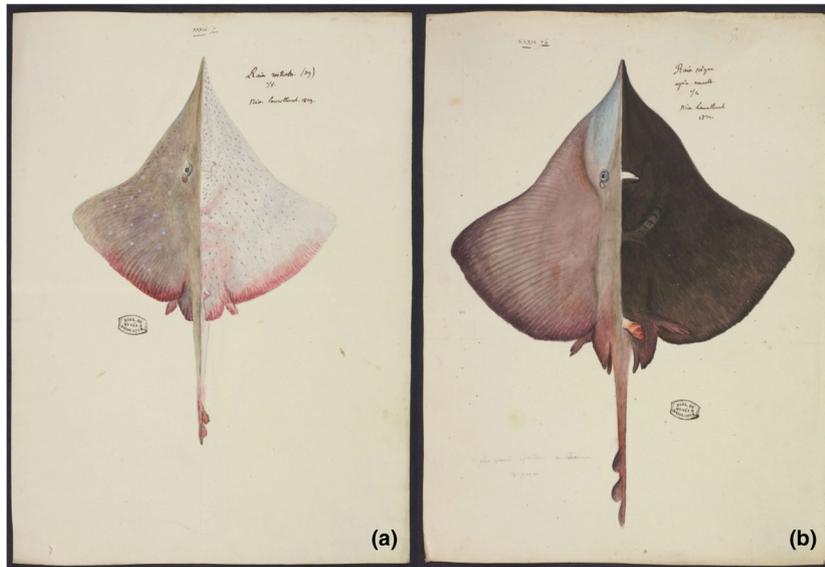


Figure 1 Original paintings of two specimens of *D. oxyrinchus* and *D. nidarosiensis* collected off Nice, France (Mediterranean Sea). *D. oxyrinchus* has been reported under the name 'Raia rostrata' (a), while *D. nidarosiensis* under the name 'Raie nègre, espèce nouvelle' (b). Paintings from 1829 have been found in the manuscript Ms-535 of Cuvier and Valenciennes conserved at the Library of the Muséum national d'Histoire naturelle (Paris, France). Photograph: © Muséum national d'Histoire naturelle. *D. nidarosiensis* is morphologically distinguished from *D. oxyrinchus* mainly for its rhombic disc and its very elongate and narrowly pointed snout. Whereas, the longnosed skate presents a more deeply concave anterior disc and narrower snout.

insights into the Atlantic-Mediterranean *D. oxyrinchus* and *D. nidarosiensis*, (2) improve the taxonomy and zoogeography of *D. intermedius* and *D. batis* and (3) identify and critically discuss species assignment errors or outdated information occurring in public databases which may jeopardise proper conservation and management of the investigated taxa.

Materials and methods

Sampling

Overall, 105 sequences of six putative species belonging to the genus *Dipturus* were retrieved from individuals collected across Central-Western Mediterranean Sea and Atlantic Ocean, at depths ranging from 55 to 1573 m, during commercial fishing cruises, scientific trawl surveys or obtained from landings or fish markets (Figure 2). Details related to sequences generated in the present study are reported in Tables 1 and S1. After identification, muscle tissue samples were collected and preserved in 96% ethanol and kept at -20°C until laboratory analyses. The morphometric data were provided by the samplers (C. Osio and S.P. Iglésias, for the North-Western and North-Eastern Atlantic, respectively, and Rob Leslie for the South-Eastern Atlantic samples).

Genetic data analysis

DNA extraction, amplification and sequencing analysis are described in the Supplementary Material. Sequence trace files were checked and manually edited in MEGA v7 (Kumar,

Stecher and Tamura, 2016). A more comprehensive barcode dataset was built by adding 314 COI sequences attributed to *D. oxyrinchus*, *D. nidarosiensis*, *D. batis*, *D. intermedius*, *D. laevis* and *D. pullopunctatus* retrieved from GenBank® (GenBank; Clark *et al.*, 2016) and the Barcode of Life Data Systems (BOLD; Ratnasingham and Hebert, 2013) (Table S2).

Newly generated and retrieved sequences, for a total of 419 records, constituted the 'DIPTURUS' dataset. Sequences were aligned using the ClustalW multiple sequence alignment algorithm (Thompson, Higgins and Gibson, 1994) and haplotypes were retrieved using DnaSP v6.12 (Rozas *et al.*, 2017). The relationships among haplotypes were investigated using the TCS method (named after Templeton, Crandall and Sing 1992, and fully described in Clement *et al.*, (2002), using the PopART software for building the network of haplotypes (Leigh & Bryant, 2015). To estimate the occurrence of population structuring between the Mediterranean and Atlantic basins (for *D. oxyrinchus* and *D. nidarosiensis*), the Analysis of Molecular Variance (AMOVA) was carried out, performing conventional F-statistics for haplotype frequencies and 10000 permutations, by using Arlequin v3.5 software (Excoffier and Lischer, 2010).

To identify and critically discuss species assignment errors or outdated information occurring in public sequence repositories, the 'DIPTURUS' dataset was further extended by adding COI sequences retrieved from GenBank and BOLD public databases assigned to all the other *Dipturus* species, including the genera *Zearaja*, *Spiniraja*, *Dentiraja* (Tables S3), generating a new dataset called 'DIPTURUS_EXT', composed by 828 sequences. We also included sequences deposited as *Zearaja*, because they are now currently renamed as *Dipturus* (Concha

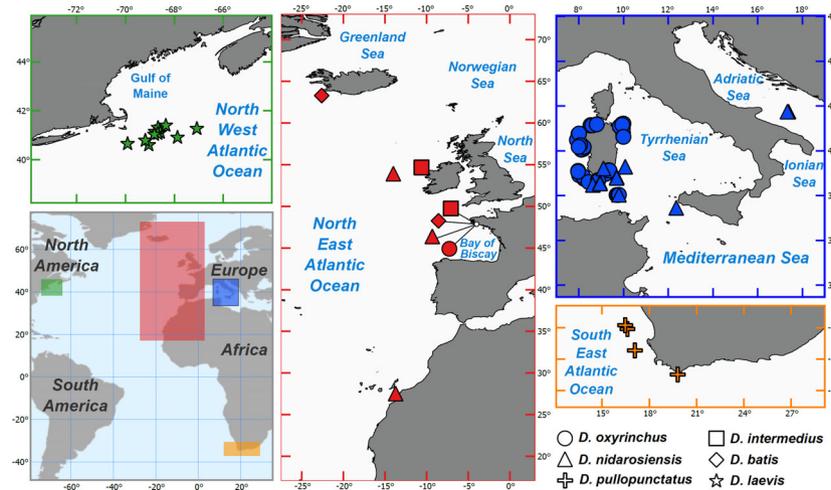


Figure 2 Sampling areas and sites of *Dipturus* specimens collected. All specimens were collected from commercial fishing cruises or scientific trawl surveys with exception of *D. pullopunctatus* specimens that were obtained from landings or fish markets.

Table 1 Number of individual sequences (N), individual specimen sex (M = male, F = female, u = unsexed) and related sampling area of the six *Dipturus* target species generated and analysed in the present paper. For each species, the GenBank accession numbers of sequences retrieved are given

Species	N	Sex (M/F/u)	Sampling area	Reference	GenBank accession
<i>D. oxyrinchus</i>	11	2/6/3	Mediterranean Sea	Present paper	MK675197-MK675207
<i>D. nidarosiensis</i>	54	15/31/8	Mediterranean Sea	Present paper	MK675121-MK675169; MW074310-MW074314
<i>D. batis</i>	4	3/0/1	North-Eastern Atlantic Ocean	Present paper	MW074309; MK675171-MK675173
<i>D. intermedius</i>	8	1/8/2	North-Eastern Atlantic Ocean	Present paper	MK675174-MK675176; MW074302-MW074306
<i>D. laevis</i>	8	3/3/2	North-Eastern Atlantic Ocean	Present paper	MK675177-MK675182; MW074307-MW074308
<i>D. pullopunctatus</i>	14	6/6/2	North-Western Atlantic Ocean	Present paper	MK675183-MK675196
	6	4/5/0	South-Eastern Atlantic Ocean	Present paper	MW074315-MW074320

et al., 2019). We also included sequences of *Dentiraja* spp. and *Spiniraja whiteleyi* because they were originally deposited as *Dipturus* but currently renamed (Last, Weigmann and Yang, 2016b). Finally, we retrieved additional 1462 sequences assigned to the family Rajidae, to further check on the presence of *Dipturus*' sequences erroneously labelled (Table S4). A sequence of *Torpedo marmorata* was used as outgroup (KT307460; Cariani et al., 2017). All the haplotypes were analysed using tree-based approaches: Neighbour-Joining (NJ, implemented in MEGA), Bayesian inference (BI, using MrBayes v3.2.7; Ronquist et al., 2012) and Maximum Likelihood method (ML, with PhyML; Guindon et al., 2010). DNA barcodes were further used for both species' delineation and specimen assignment analyses, whose combined use is potentially beneficial for species identification (Delrieu-Trottin et al., 2020).

Species delimitation analyses

Species identified based on morphology are referred to as 'species', while species delimited by DNA barcode are

referred to as Molecular Operational Taxonomic Unit (henceforth MOTU; Moritz, 1994; Vogler and DeSalle, 1994).

In the present study, MOTUs were delimited using three different algorithms: (a) Refined Single Linkage (RESL) as implemented in BOLD and used to produce Barcode Index Numbers (BINs; Ratnasingham and Hebert, 2013), (b) Automatic Barcode Gap Discovery (ABGD; Puillandre et al., 2012), (c) tree-based Bayesian Poisson Tree Processes (bPTP) method (Zhang et al., 2013). More details are provided in the Supplementary Material. A final delimitation criterion was established based on a 66% consensus among the three methods (Garcia-Melo et al., 2019).

Specimen assignment method using distance

To test the accuracy of COI for species identification, the proportion of correct identifications was calculated using the 'Best Close Match Analysis' (henceforth BCMA, Meier et al., 2006), which is a distance-based analysis incorporating a threshold, carried out in the R package Spider v1.5

(Brown *et al.*, 2012). For this method, successful identification requires that a query sequence should have the closest barcode match with a sequence from the same nominal species, while the identification is considered ambiguous or incorrect if the former has closest matches with barcodes of different species (Meier *et al.*, 2006). We here used the function `threshOpt()` implemented in the same R package to identify the optimal threshold value for our dataset following the author's tutorial indications (Brown and Collins, 2011). Further, we checked for the presence of a barcoding gap that is the lack of overlap between the distributions of the maximum intraspecific (MI) and the nearest neighbour (NN) genetic distances (Meyer and Paulay, 2005). The occurrence of a distinct gap between intraspecific and interspecific variability was assessed by plotting the MI distance against the distance to the NN for each putative species (Collins and Cruickshank, 2013); MI and NN distances were calculated using the Kimura's two parameter model (K2P; Kimura, 1980).

Other specimen assignment methods

One of the most important methodological problem in the field of DNA barcoding remains the reliable assignment of an unknown query sequence to its correct species. Considering issues associated to the use of distance measures to infer species affiliation (Zhang *et al.*, 2008), we here applied new specimen assignment methods recently proposed for allocating specimens to species. Four specimen assignment methods implemented in the R package `BarcodingR` v1.0.2 (Zhang *et al.*, 2017) were used: (a) the back-propagation neural networks method (BP) (Zhang *et al.*, 2008; Zhang and Savolainen, 2009); (b) the fuzzy set-based approach method (FZ) (Zhang *et al.*, 2012a); (c) the alignment-free kmer-based method (FZKMER) (Zhang *et al.*, 2012b) and (d) the Bayesian-based (Ba) method (Jin *et al.*, 2013). Both FZ and FZKMER methods assign to each potential identification an 'FMF value' in the range of 0–1, which indicates the likelihood of the assignment (Zhang *et al.*, 2012b), while BP and

Ba assign to each identification a 'bp.probabilities' or 'Bayesian.prob', respectively, also in a range of 0–1. For the sake of clarity, we will refer hereafter to FMF values, bp.probabilities and Bayesian.prob as 'probabilities'. As the different methods can lead to conflicting results, a consensus has been established for the multiple methods using the `consensus.identify` function of the `barcodingR` package. We considered that a consensus emerged if at least three methods were converging. More details are provided in the Supplementary material.

Results

The final alignment of the '*DIPTURUS*' barcode dataset consisted of 555 bp and included 41 haplotypes (DIPH1–DIPH41; Tables S1 and S2). The TCS haplotype network revealed the occurrence of four main haplogroups (Groups I, II, III and IV) and two single haplotypes (Figure 3). Group I included exclusively specimens morphologically assigned to *D. oxyrinchus*, while Groups II, III and IV were formed by specimens morphologically identified as *D. nidarosiensis*, with the only exception of DIPH13 which corresponded to a sequence erroneously attributed to *D. oxyrinchus* (Vargas-Caro, 2018) and DIPH20 recorded in GenBank as *D. springeri* but reported in BOLD as *D. nidarosiensis* (Table S2). Group II included three haplotypes (DIPH29–31), shared by individuals 'morphologically' identified as *D. intermedius* (or *D. batis*). Group III included 12 haplotypes shared by specimens 'morphologically' identified as *D. batis* (DIPH26–28) and *D. laevis* (DIPH33–40). The haplotype DIPH25 was shared by individuals of both species (11 common blue skate individuals and 2 barndoor skates). The network showed the presence of two haplotypes (DIPH32 and DIPH41) out of the main four delimited groups. DIPH41 was shared by individuals morphologically identified as *D. pullopunctatus* collected in the South-Eastern Atlantic. DIPH32 was found in a single individual caught in the Indian Ocean (IND) and deposited in BOLD as *D. intermedius*. Both Groups I and IV presented haplotypes shared by individuals coming from

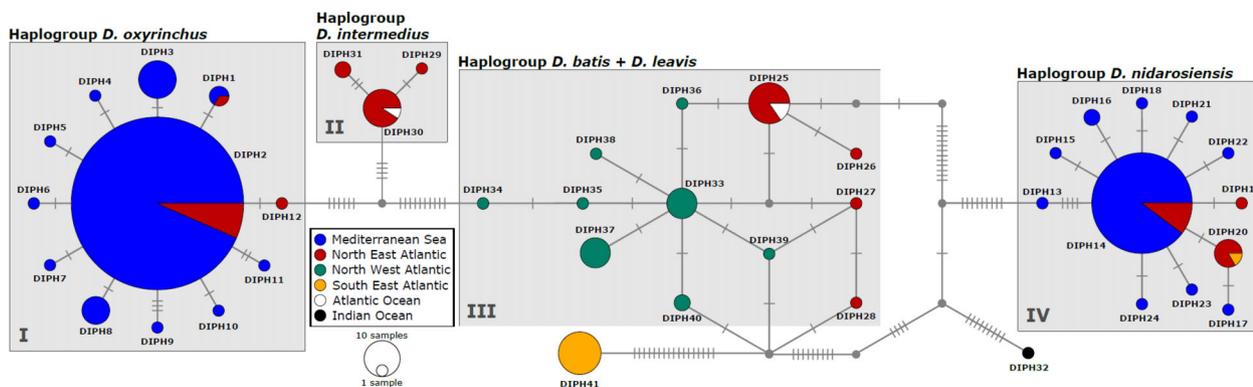


Figure 3 TCS network of haplotypes for COI sequences (both newly obtained and downloaded from public databases) belonging to individuals of *D. oxyrinchus*, *D. nidarosiensis*, *D. laevis*, *D. batis*, *D. intermedius* and *D. pullopunctatus*. The grey boxes correspond to haplogroups I to IV. The size of the circle represents the frequency of each haplotype; each segment represents one mutational event.

the Mediterranean Sea and the Atlantic Ocean. In addition, for both *D. oxyrinchus* and *D. nidarosiensis*, AMOVA analysis confirmed the lack of genetic differentiation between Mediterranean and Atlantic sites ($F_{ST} = 0$; $P = 1.0$).

The extended dataset 'DIPTURUS_EXT' gave a final alignment of 514 bp with 120 haplotypes. Within the genus *Dipturus* we found the presence of mislabelled sequences, deposited in the public repositories as *Raja* or *Rajiformes* spp. (Table S3). After adding sequences from the family Rajidae (Table S4), we obtained a final alignment of 514 bp, with a total of 453 haplotypes. The COI-based taxonomy at the family level accounted for three clusters corresponding to the subfamilies Rostrojini, Amblyrajini and Rajini, and the latter included all the sequences assigned to *Dipturus* (Figure S1; ML and BA trees not shown).

The molecular-based taxonomy at *Dipturus* level (Figure 4) revealed that haplotypes corresponding to Group I of the haplotype network formed a monospecific group with public sequences assigned to *D. oxyrinchus* and those of the Group IV clustered with sequences assigned in public repositories to *D. nidarosiensis*. The three haplotypes of Group II assigned to *D. intermedius* clustered separately from the haplotype DIPH32 (deposited as *D. intermedius*), which in contrast clustered together with DENH94 and DENH95 assigned to *Dentiraja healdi*. Haplotypes of Group III (assigned to *D. laevis* and *D. batis*) clustered in a single cluster (supported by the 94% of bootstrap replicates) together with public sequences of the two nominal species. Finally, all the trees showed the presence of the single haplotype DIPH41 which included all the sequences assigned to *D. pullopunctatus*.

Refined Single Linkage algorithm (RESL, Figure 4) attributed a single BIN to each of the nominal species, except a few cases where a unique BIN included haplotypes assigned to two species (i.e., *D. laevis* – *D. batis* and *Dentiraja endeavouri* – *Dentiraja polyommata*) or to several nominal species (i.e., the unique BIN AAB5857 included haplotypes attributed to *D. argentinensis*, *D. chilensis*, *D. nasutus*, *D. maugeanus* and *D. trachyderma*). The species delimitation analyses, performed on the 'DIPTURUS_EXT' dataset which included 41 nominal species, yielded a varying number of MOTUs: 32, 30 and 32 MOTUs using RESL, ABGD and bPTP, respectively. A consensus number of 31 MOTUs was obtained from the three methods (Figure 4).

The BCMA carried out on the 'DIPTURUS' dataset, using 0.3% as the optimal threshold, showed that 96.4% of specimens were correctly recognised to their species (Table S5); errors in species identification amounted to 3.6% and mostly concerned sequences morphologically identified as *D. batis* (2.9%) and to a lesser extent to those classified as *D. laevis* (0.7%). The same analysis performed on the 'DIPTURUS_EXT' dataset, using 0.1% as the optimal threshold, revealed that 92% of sequences were correctly identified (Table S6); most of the 8% of errors concerned sequences morphologically classified as *D. chilensis* (2.5%), *D. argentinensis* (1.8%) and *D. batis* (1.3%) (Table S6). The analysis of the 'barcoding gap' performed on the 'DIPTURUS' dataset revealed that the genetic distance between each conspecific individual was smaller than to any allospecific individual for all species except for *D. laevis* and

D. batis (Figure 5A). When the analysis was performed on 'DIPTURUS_EXT' dataset (Figure 5B), besides *D. laevis* and *D. batis*, the barcode gap was also lacking for *D. argentinensis*, *D. brevicaudatus*, *D. chilensis*, *Dentiraja lemprieri*, *D. polyommata*, *D. nasutus*, *Dipturus* sp. A, *Dipturus springeri* and *D. trachyderma*.

The COI identification efficacy of the newly generated sequences was 100% based on the results of BP, FZ and Ba assignment methods and 96% in the FZKMER approach. Probabilities associated to each specimen assignment varied between the four methods (Table S7). Despite this, a consensus emerged for all the specimens analysed, with all the methods converging in 96% of the cases. When the species assignment analyses were performed on the 'DIPTURUS_EXT' dataset, the identification was confirmed for 99% of the cases based on the results of BP, FZ and FZKMER methods and for the 97% looking at the results of Ba approach. In eight cases, the four methods provided the same results regarding the wrongly identified specimens (Table S8). Despite differences among methods, a consensus emerged for all the specimens and in 97% of the cases the four methods converged.

Discussion

DNA-based identification methods, such as DNA barcoding, have been used across broad taxonomic ranges and represent an especially useful tool for species identification, especially when morphological classification may be problematic. In the last half of century, DNA barcoding approach not only proved to be useful to assign unknown specimens to known species (Bazin et al., 2012) but also led to the development of multiple DNA-based methods to delineate species (Brown et al., 2012; Puillandre et al., 2012; Ratnasingham and Hebert, 2013). This study highlights that a joint use of species delimitation methods and specimen assignment approaches is beneficial not only to improve the accuracy of specimen assignment but also to build curate reference libraries which currently host a substantial portion of taxonomic misidentification (Hubert et al., 2010; Hubert and Hanner, 2015).

Our case study focussed on skates that are among the most complex organisms for species identification due to their conserved body morphology and thus plagued by misidentifications. Here we report that in most cases, clear identification is possible, indicating the efficacy of combined COI-based methods in a problematic group. As applied here, the proposed workflow could be useful for a wide range of taxa to successfully assign an unknown specimen to a known species, with a certain level of confidence. The applied methods prove to be extremely useful when tree- and threshold-based methods are less informative, that is when the barcode gaps lacks, and when working on species showing strikingly conserved morphological attributes. In this framework, considering the limits of traditional morphological identification, the correct assignment of specimens to the species level using molecular analyses is critical for the accurate status assessments and conservation of species, whatever the system considered.

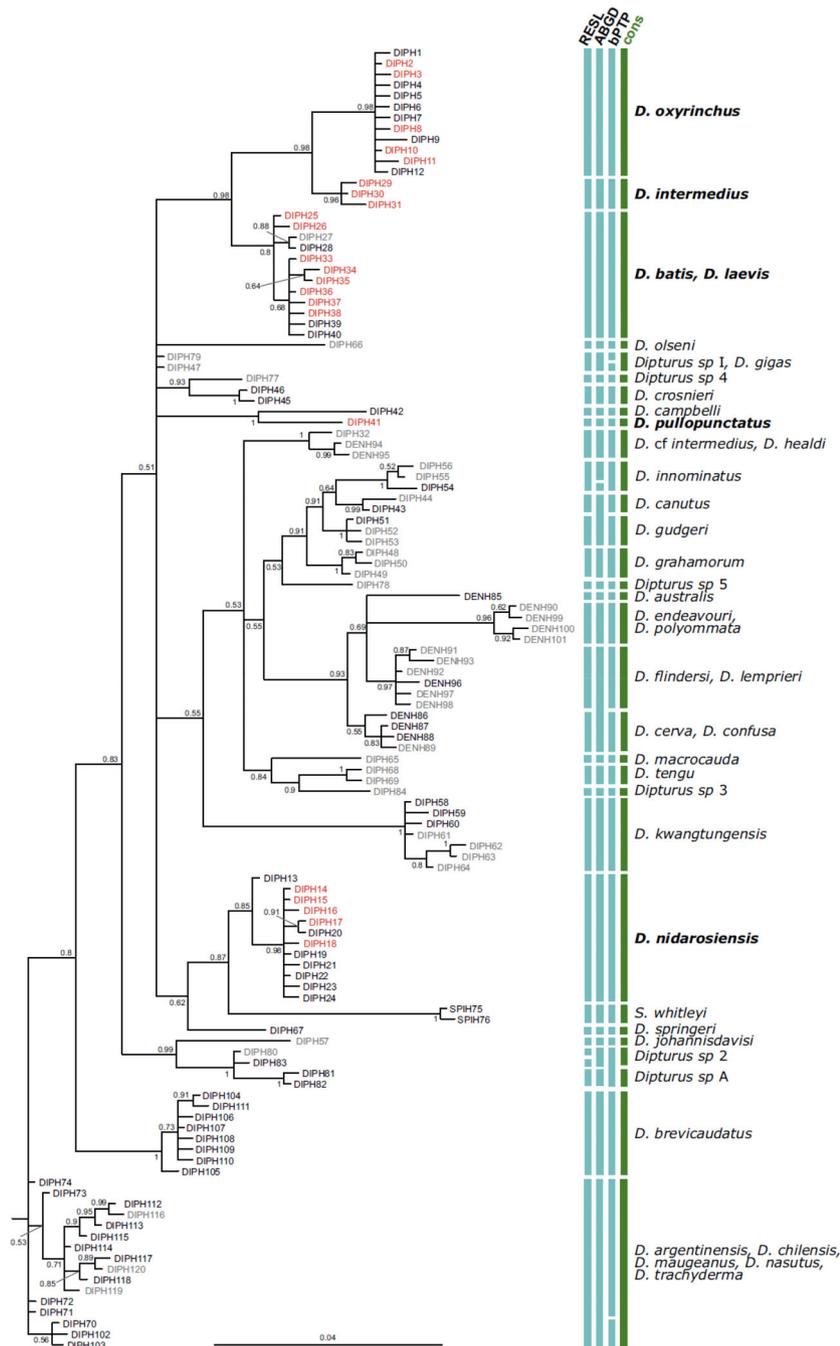


Figure 4 Bayesian tree of haplotypes from *Dipturus* sequences; near the node are the posterior probabilities (>50%). On the right, MOTU delimitation schemes according to the three species delimitation methods implemented (aqua green) and the resulting consensus MOTU (green). Names in red indicate the haplotypes obtained in the present study (Table S1). Names in black represent haplotypes from published sequences deposited in the GenBank and BOLD. All the other haplotypes (grey) are from public but unpublished (i.e., deposited in the database but not published yet in a scientific paper) sequences.

Species delimitation and effectiveness of specimen assignment methods in the genus *Dipturus*

Our study revealed that any of the delimitation methods used cannot retrieve all the 41 *Dipturus* putative species, but 80%

as maximum, represented by 32 MOTUs recovered by RESL and bPTP. Discordant cases could be due to intrinsic limitations of the methods and problems in identification (Garcia-Melo *et al.*, 2019). Among the delimitation methods used, ABGD performed worst and the main error type produced is over lumping. ABGD, indeed, is known to under split

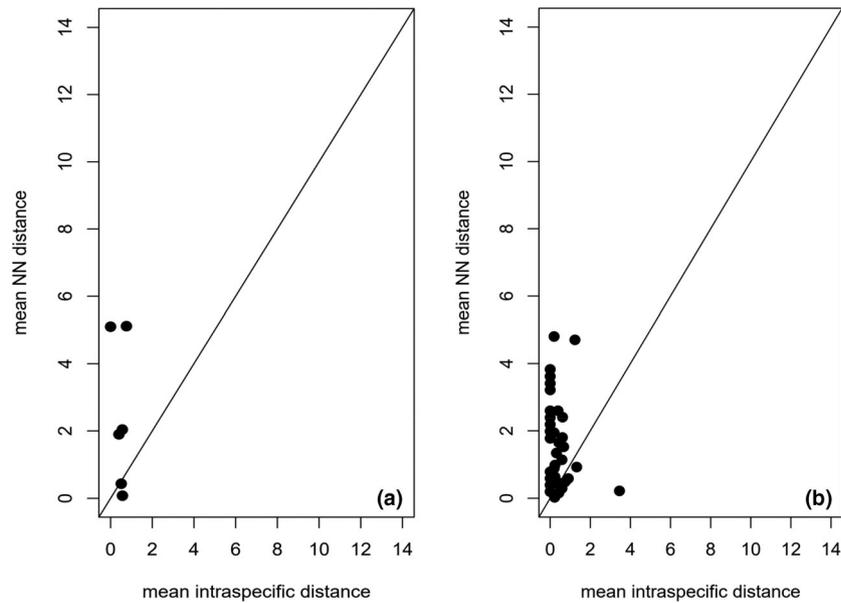


Figure 5 Comparison of NN (Nearest Neighbour or minimum interspecific distance) and MI (Maximum Intraspecific) distances for *D. oxyrinchus*, *D. nidarosiensis*, *D. intermedius*, *D. batis*, *D. laevis* and *D. pullopunctatus* (a) and for all the species belonging to 'DIPTURUS-EXT' dataset (b). Equal intra- and inter-specific variation is marked by the black line. Points above the black line indicate 'species' with 'barcode gap'.

species, and poorly perform on datasets with many species and faster speciation rates (Dellicour and Flot, 2018). By contrast, bPTP cannot discriminate all the morphological species, since it is typically influenced by the under sampling of rare species and over sampling of species with small intraspecific variation, as *D. oxyrinchus* and *D. nidarosiensis* (Zhang *et al.*, 2013).

As concerns the results of additional analyses, BCMA returned a percentage of ambiguous identifications, mostly linked to species lacking the barcode gap and also to wrong taxonomic identifications which could bias data analyses and the interpretation of results. However, the use of specimen assignment methods solved most of these uncertainties and shed light on errors reported in public databases. A series of misidentification problems were found for the following species dwelling in two near-by areas, i.e., South-Western Atlantic and South-Eastern Pacific: *D. argentinensis*, *D. brevicaudatus*, *D. chilensis*, *D. maugeanus*, *D. nasutus*, and *D. trachyderma*. All these species, except *D. brevicaudatus*, fall in the same BIN AAB5857 and the same MOTU with any delimitation method used. In particular, *D. brevicaudatus* (also known as *Zearaja brevicaudata*) has been resurrected a few years ago thanks to a study (Gabbanelli *et al.*, 2018) providing evidences that specimens from South-Western Atlantic (i.e., Argentina) were separated from those dwelling in the South-Eastern Pacific. Our results, however, showed that a few sequences of specimens caught in the South-Eastern Pacific were assigned to *D. brevicaudatus*. Considering the currently high levels of exploitation of *D. brevicaudatus* and its 'Vulnerable' (VU) status, there is the urgent need to further study its abundance trend and geographic range. Our results also highlighted that a sequence of *D.*

trachyderma was recognised as *D. argentinensis*. The rough-skin skate *D. trachyderma* is typically confused with all the other smaller skates of the genus *Dipturus* dwelling off the South America, i.e., the yellownose skate *D. chilensis* and the Argentine skate *D. argentinensis* (Last *et al.*, 2016a). The external morphology of these skates is remarkably similar, especially in early life stages, leading to frequent mislabelling (Vargas-Caro *et al.*, 2017). Both *D. chilensis* and *D. trachyderma* are important components of commercial elasmobranchs in South American fisheries (Agnew *et al.*, 2000; Estalles *et al.*, 2011; Bustamante *et al.*, 2012) and classified as 'Vulnerable' (VU). Thus, their correct identification is crucial for the effectiveness of fishery monitoring and conservation (Vargas Caro *et al.*, 2017). Finally, we reported that two unpublished (i.e., present in public database, but not published yet in a scientific paper) sequences of *D. nasutus* were identified as *D. chilensis*. However, *D. chilensis* is morphologically different to *D. nasutus*, and the two specimens came from New Zealand, out of the typical range of *D. chilensis* (Last *et al.*, 2016a; Concha *et al.*, 2019). *D. nasutus* is at 'Least Concern' (LC) for IUCN, but since it is becoming an increasingly important commercial fish species in New Zealand, its populations needed to be more carefully examined and we need to deepen the knowledge on the range distribution of species living in this wide area.

The difficulties encountered in identifying species and the lack of curation of international repositories could lead to an increasing number of misidentified records in elasmobranchs (Hobbs *et al.*, 2006; Verissimo *et al.*, 2017; Wannell *et al.*, 2020), and the family Rajidae is not an exception (Cerutti-Pereira *et al.*, 2012). There is the urgent need to check and update the information reported in public

databases and feed them with new verified entries. The expected increase of reference libraries will likely disclose taxonomic ambiguities, guarantee the correct species identification and the reliable assessment of resource status. This could pave the way for more effective conservation of vulnerable skates, whose risk of decline could be underestimated if not correctly identified.

Zoogeography of the Atlantic-Mediterranean *D. oxyrinchus* and *D. nidarosiensis*

While it is consolidated that *D. oxyrinchus* has a wide geographical distribution, the zoogeography of *D. nidarosiensis* is frequently a matter of debate. Our findings provided new evidence of its presence in the Mediterranean Sea. Indeed, despite being historically considered endemic of the Atlantic Ocean, the occurrence of *D. nidarosiensis* has been recently reported in several Mediterranean areas as the Southern Sardinia coast (Cannas *et al.*, 2010; Follesa *et al.*, 2012; Cariani *et al.*, 2017; Marongiu *et al.*, 2017; Porcu *et al.*, 2017), off Algeria (Cariani *et al.*, 2017), Alboran Sea (Ramírez-Amaro *et al.*, 2017), in the Southern Adriatic and Ionian Seas (Cariani *et al.*, 2017; Carbonara *et al.*, 2019) and the Strait of Sicily (Geraci *et al.*, 2019). All these areas are characterised by the presence of numerous deep-sea canyons which provide more heterogeneous habitat than the adjacent slopes, potentially representing a favourable ecosystem for deep-sea chondrichthyan species (Rey *et al.*, 2010; Baro, Rueda and Diaz del Rio, 2012; Ramírez-Amaro *et al.*, 2016). The distribution of *D. nidarosiensis* in the Mediterranean could be even wider but possibly still unreported because of misidentification issues (i.e., confused with *D. oxyrinchus*; Follesa *et al.*, 2011; Cariani *et al.*, 2017; Ramírez Amaro *et al.*, 2017; Cannas *et al.*, 2010; Iglésias, Toulhoat and Sellos, 2010) or technical constraints related to its deep bathymetric distribution. *Dipturus nidarosiensis* is indeed one of the deepest living skates, found till 1573 m in the present study, dwelling beyond the range of either commercial fishing or bottom trawl surveys; because of this, it has been supposed to be rarely caught so far. In order to fill the gaps in knowledge on this species and other deep-sea elasmobranchs, there is an urgent need to further monitor by-catches of deep-sea fisheries, especially those coming from shrimp' fishing vessels, which are recognised as the main source of skates by-catches (CIESM, 2003; Tudela, 2004; Bensch *et al.*, 2009). Nevertheless, in light of multiple records in most of the Western Mediterranean basin, the distribution of the Norwegian skate should be revised and updated in the IUCN Red List and in all the fish databases.

Network analysis showed the presence of shared haplotypes between Atlantic and Mediterranean specimens, likely confirming their low genetic differentiation. In addition, AMOVA analysis confirmed the lack of genetic differentiation between the two areas for both species, in line with what has been previously reported (Cannas *et al.*, 2010; Ramirez-Amaro *et al.*, 2017; Carbonara *et al.*, 2019; Melis *et al.*, 2020).

Although our results could be biased by the low number of public sequences from the Atlantic Ocean, the lack of genetic differentiation between the Atlantic and Mediterranean individuals could be explained by the capability of big skates to travel for longer distances than previously thought. Recent studies, indeed, based on satellite tagging, have reported that individual skates can travel up to 2100 km (Farrugia *et al.*, 2016; Bird *et al.*, 2020), despite what has been previously observed using fishery-based tagging (Wearmouth and Sims, 2009; King and McFarlane, 2010). Such results contrast with increasing findings of distinct populations in other elasmobranchs species such as deep-sea sharks for which limited gene flow has been reported between the Atlantic and Mediterranean, potentially linked to bathymetric constraints limiting their dispersal potential (e.g., Catarino *et al.*, 2015, 2017; Gubili *et al.*, 2016).

Further research is needed to further quantify the connectivity of deep-sea skates between the Mediterranean and Atlantic Ocean keeping the feeding of current databases with new sequences as a priority. Both satellite tagging and genetic studies (with the use of additional nuclear markers, such as microsatellites or single-nucleotide polymorphisms) should be performed in order to address the important issues of connectivity and thus genetic conservation of these species.

D. intermedius and the close relatedness between *D. batis* and *D. laevis*

The phylogenetic reconstruction showed that all the haplotypes of the Group II, morphologically identified as *D. intermedius*, do form a strongly supported clade, differentiated with respect to other *Dipturus* species, supporting *D. intermedius* as a distinctive taxon. In addition, all the sequences attributed to the flapper skate derived from individuals caught in the North-Eastern Atlantic, confirming its range distribution in North-Western Scotland (Neat *et al.*, 2015).

The haplotype network (Figure 2) and tree (Figure 4) showed that haplotypes from specimens morphologically identified as *D. batis* and *D. laevis* clustered together in the same Group III and in the same BIN/MOTU with any delimitation methods used. The inability to resolve *D. batis* from *D. laevis* is not an unexpected result: a previous study showed the close evolutionary distance between these species (Ball *et al.*, 2016) which were included in the same monophyletic group in their COI phylogenetic analysis. However, this finding was never properly discussed, getting almost unnoticed. Moreover, the BCMA approach returned an overall 4% of ambiguous identification entirely due to the inability to assign sequences attributed to these two species. This result can be explained in two ways. Firstly, the ambiguous identification with the BCMA could be due to the increasing geographic scale of sampling which could increase the chances of encountering closely related species (Lukhtanov *et al.*, 2009), as in our case. Indeed, the increase of geographical sampling range could lead to a high genetic distance within the same species and low distance to the

nearest neighbour, thus making the identification more difficult (Bergsten *et al.*, 2012). Secondly, errors produced by the BCMA, that is false negatives in our case, are reported to occur when little or no sequence variation in the barcoding fragment is found between two different species; such errors are critical in the barcoding approach because the method results less powerful to delimit species boundaries (Wiemers and Fiedler, 2007). Conversely, the four species assignment methods correctly assigned all the sequences morphologically attributed to *D. batis* and *D. laevis*. The only exception was represented by two public sequences deposited as *D. laevis* which were indeed assigned to *D. batis*. In the light of these results, haplotype DIPH25 is not shared between the two species but actually formed only by *D. batis* sequences.

The contrasting results obtained by using the tree-based method and the BCMA vs the assignment approaches could be explained considering a series of issues linked to DNA barcoding (Zhang *et al.*, 2012a). First, in the BCMA, the query sequence is identified by the closest barcode match at a genetic distance below a defined threshold (Li *et al.*, 2016). However, the use of thresholds in species assignments has been extensively debated because of the lack of strong biological support and universality to all animal species (Meyer and Paulay, 2005; Hickerson, Meyer and Moritz, 2006; Ward, Hanner and Hebert, 2009). In our case, the threshold approach is not completely reliable because the barcode gap, which guarantees the correct species identification, did not exist for the two investigated species. A second issue is related to sampling: the 14 available sequences of *D. batis* included in this study should not be enough to cover all the genetic diversity of a population (Zhang *et al.*, 2010). Another issue is related to errors encountered with tree-based methods, which assume that species are monophyletic, and this may be unrealistic, especially for recently diverged species (Hudson and Coyne 2002; Hickerson *et al.*, 2006; Knowles and Carstens 2007; Lou and Golding 2010), which likely is the case of *D. laevis* and *D. batis*.

On the contrary, specimen assignment methods based on artificial intelligence (Zhang *et al.*, 2008), fuzzy-set-theory (Zhang *et al.*, 2012a) non-tree-based Bayesian methods (Jin *et al.*, 2013) successfully identified specimens belonging to both species and seem to outperform tree- and threshold-based methods when the barcode gap lacks. Thus, in our data the presence of the barcode gap is not considered as a strict predictor of the identification success, as also previously reported in other case studies (i.e., Little, Knopf and Schulz, 2013; Li *et al.*, 2016; Gibbs, 2018).

D. batis and *D. laevis* should be easily morphologically distinguished (Last *et al.*, 2016a) and overall, they do not have any overlap in their geographical distribution, at least as far as we know. Indeed, still few data are available regarding the distribution of these species, especially for *D. batis* (Lynghammar *et al.*, 2014). Nevertheless, we cannot exclude that *D. batis* and *D. laevis* could be single species with an ampho-Atlantic distribution, whose observed morphological variations constitute the intraspecific population variability corresponding to the beginning of an allopatric

speciation. Considering the status of conservation concern of both species, more studies integrating morphological and molecular analyses are needed in order to shed light on their distribution and relationship.

Conclusion

This study provides evidences that the joint use of DNA barcoding with species delimitation methods and specimen assignment approaches proved to be extremely useful to improve the accuracy of species identification and build curate reference libraries. Our results also pinpoint out that the lack of barcode gap does not mean a certain failure in species identification, but the need of combining different COI-based methods especially when dealing with species characterised by conserved body morphology and when tree- or threshold-based methods results less powerful.

Besides problems arisen using a single, mitochondrial marker, the application of DNA barcoding is further complicated by databases' limitations. The number of published, verified entries is still too low for some species, especially for the latest described ones, and there is the need to check and update the information as well as to feed databases with new verified entries. The expected increase of reference libraries will likely disclose taxonomic ambiguities, guarantee the correct species identification and the assessment of resource status. This could pave the way for more effective fishery management and conservation of vulnerable deep-sea skates, whose risk of decline could be underestimated if not correctly identified.

Finally, we here provided additional evidences of the occurrence of *D. nidarosiensis* in the Central-Western Mediterranean Sea and also confirmed the lack of genetic distinctness between Mediterranean and Atlantic individuals of longnosed and Norwegian skates. We also reported the close evolutionary distance between *D. batis* and *D. laevis* and the need of further studies to shed light on their phylogenetic relationship.

Acknowledgements

The authors express their profound gratitude and deep regard to Emeritus Professor Angelo Cau who provided funds, equipment, constant encouragement, and valuable scientific guidance throughout the duration of the project. They are also thankful to the staff of the Marine Biology Group (students, post-docs and permanent staff) for helping in sampling during the commercial and scientific surveys at sea. The authors express their gratitude also to Chato Osio for providing tissue samples of *D. laevis* analysed in this study. This study was funded by the Autonomous Region of Sardinia within the frame of the research project 'Approccio multidisciplinare per la conservazione e gestione della selacofauna del Mediterraneo' (LR7 CRP-25321). The author LC acknowledges the support from PON AIM 1854833– PON Ricerca e Innovazione 2014-2020 – Azione I.2 - D.D. n. 407, 27/02/2018 – Attraction and International Mobility'. The authors also thank the curators of the library of the

Muséum national d'Histoire naturelle for allowing the examination and digitisation of manuscripts.

References

- Agnew, D.J., Nolan, C.P., Beddington, J.R. & Baranowski, R. (2000). Approaches to the assessment and management of multispecies skate and ray fisheries using the Falkland Islands fishery as an example. *Can. J. Fish. Aquat. Sci.* **57**, 429–440.
- Ball, R.E., Serra-Pereira, B., Ellis, J., Genner, M.J., Iglésias, S., Johnson, A.F., Jones, C.S., Leslie, R., Lewis, J., Mariani, S., Menezes, G., Neat, F., Noble, L.R., Sims, D.W. & Griffiths, A.M. (2016). Resolving taxonomic uncertainty in vulnerable elasmobranchs: are the Madeira skate (*Raja maderensis*) and the thornback ray (*Raja clavata*) distinct species? *Conserv. Genet.* **17**, 565–576.
- Baro, J., Rueda, J.L. & Díaz del Río, V. (2012). South Iberian submarine canyons in the Alboran sea: geohabitats, associated communities and fisheries resources. In Würtz, M. (ed) *Mediterranean submarine canyons: ecology and governance* (pp 145–156). Málaga, Spain: IUCN.
- Bazinet, A.L. & Cummings, M.P. (2012). A comparative evaluation of sequence classification programs. *BMC Bioinformatics* **13**, 92.
- Bensch, A., Gianni, M., Gréboval, D., Sanders, J.S. & Hjort, A. (2009). Worldwide review of bottom fisheries in the high seas. FAO Fisheries and Aquaculture Technical Paper, 522 Rev. 1, 1–145.
- Bergsten, J., Bilton, D.T., Fujisawa, T., Elliott, M., Monaghan, M.T., Balke, M., Hendrich, L., Geijer, J., Herrmann, J., Foster, G.N., Ribera, I., Nilsson, A.N., Barraclough, T.G. & Vogler, A.P. (2012). The effect of geographical scale of sampling on DNA barcoding. *Syst. Biol.* **61**, 851–869.
- Bird, C., Burt, G.J., Hampton, N., McCully Phillips, S.R. & Ellis, J.R. (2020). Fifty years of tagging skates (Rajidae): using mark-recapture data to evaluate stock units. *J. Mar. Biol. Assoc. U.K.* **100**, 121–131.
- Brown, S.D.J. & Collins, R.A. (2011). Spider: Species Identity and Evolution in R. A Tutorial.
- Brown, S.D.J., Collins, R.A., Boyer, S., Lefort, M.-C., Malumbres-Olarte, J., Vink, C.J. & Cruickshank, R.H. (2012). SPIDER: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Mol. Ecol. Resour.* **12**, 562–565.
- Bustamante, C., Vargas-Caro, C., Oddone, M.C., Concha, F., Flores, H., Lamilla, J. & Bennett, M.B. (2012). Reproductive biology of *Zearaja chilensis* (Chondrichthyes: Rajidae) in the south-east Pacific Ocean. *J. Fish Biol.* **80**, 1213–1226.
- Cannas, R., Follesa, M.C., Cabiddu, S., Porcu, C., Salvadori, S., Iglésias, S.P., Deiana, A.M. & Cau, A. (2010). Molecular and morphological evidence of the occurrence of the Norwegian skate *Dipturus nidarosiensis* (Storm, 1881) in the Mediterranean Sea. *Mar. Biol. Res.* **6**, 341–350.
- Carbonara, P., Cannas, R., Donnalioia, M., Melis, R., Porcu, C., Spedicato, M.T., Zupa, W. & Follesa, M.C. (2019). On the presence of *Dipturus nidarosiensis* (Storm, 1881) in the Central Mediterranean area. *PeerJ* **7**, e7009.
- Cariani, A., Messinetti, S., Ferrari, A., Arculeo, M., Bonello, J.J., Bonnici, L., Cannas, R., Carbonara, P., Cau, A., Charilaou, C., El Ouamari, N., Fiorentino, F., Follesa, M.C., Garofalo, G., Golani, D., Guarniero, I., Hanner, R., Hemida, F., Kada, O., Lo Brutto, S., Mancusi, C., Morey, G., Schembri, P.J., Serena, F., Sion, L., Stagioni, M., Tursi, A., Vrgoc, N., Steinke, D. & Tinti, F. (2017). Improving the conservation of mediterranean chondrichthyans: the ELASMOMED DNA barcode reference library. *PLoS One* **12**, e0170244.
- Catarino, D., Knutsen, H., Veríssimo, A., Olsen, E.M., Jorde, P.E., Menezes, G., Sannæs, H., Stanković, D., Company, J.B., Neat, F., Danovaro, R., Dell'Anno, A., Rochowski, B. & Stefanni, S. (2015). The Pillars of Hercules as a bathymetric barrier to gene flow promoting isolation in a global deep-sea shark (*C. entoscyrnus coelolepis*). *Mol. Ecol.* **24**, 6061–6079.
- Catarino, D., Stanković, D., Menezes, G. & Stefanni, S. (2017). Insights into the genetic structure of the rabbitfish *Chimaera monstrosa* (Holocephali) across the Atlantic-Mediterranean transition zone. *J. Fish. Biol.* **91**, 1109–1122.
- Cerutti-Pereyra, F., Meekan, M.G., Wei, N.W.V., O'Shea, O., Bradshaw, C.J. & Austin, C.M. (2012). Identification of rays through DNA barcoding: an application for ecologists. *PLoS One* **7**, e36479.
- CIESM (2003). Mare Incognitum? Exploring Mediterranean deepsea biology. CIESM Workshop Monographs. n 23, 128 pp Available at: www.ciesm.org/publications/Heraklion03.pdf.
- Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J. & Sayers, E.W. (2016). GenBank. *Nucleic Acids Res.* **44**(D1), D67–D72.
- Clement, M., Snell, Q., Walker, P., Posada, D. & Crandall, K. (2002). TCS: estimating gene genealogies. In Parallel and Distributed Processing Symposium, International (Vol. 2, pp. 0184–0184).
- Collins, R.A. & Cruickshank, R.H. (2013). The seven deadly sins of DNA barcoding. *Mol. Ecol. Resour.* **13**, 969–975.
- Compagno, L.J. & Ebert, D.A. (2007). Southern African skate biodiversity and distribution. In Ebert, D.A. & Sulikowski, J.A. (Eds.), *Biology of Skates* (pp 19–39). Dordrecht: Springer.
- Concha, F.J., Caira, J.N., Ebert, D.A. & Pompert, J.H. (2019). Redescription and taxonomic status of *Dipturus chilensis*. (Guichenot, 1848), and description of *Dipturus lamillai* sp. nov. (Rajiformes: Rajidae), a new species of long-snout skate from the Falkland Islands. *Zootaxa*, **4590**, 501–524.
- Coulson, M.W., Denti, D., Van Guelpen, L., Miri, C., Kenchington, E. & Bentzen, P. (2011). DNA barcoding of Canada's skates. *Mol. Ecol. Resour.* **11**, 968–978.
- Cuvier, G. & Valenciennes, A. (1829). Manuscript Ms-535 - Squales, in manuscripts Ms 480-558 - Histoire naturelle des

- poissons, par Cuvier et Valenciennes. Bibliothèque du Muséum national d'Histoire naturelle. [Papier et calque. 78 liasses. Aquarelle, gravure, dessins au crayon et à l'encre. 300 × 220 mm. Matériaux de cet ouvrage publié de 1829 à 1849.]
- De Astarloa, J.D., Mabragana, E., Hanner, R. & Figueroa, D.E. (2008). Morphological and molecular evidence for a new species of longnose skate (Rajiformes: Rajidae: *Dipturus*) from Argentinean waters based on DNA barcoding. *Zootaxa* **1921**, 35–46.
- Dellicour, S. & Flot, J.F. (2018). The hitchhiker's guide to single-locus species delimitation. *Mol. Ecol. Resour.* **18**, 1234–1246.
- Delrieu-Trottin, E., Durand, J.D., Limmon, G., Sukmono, T., Sugeha, H.Y., Chen, W.J., Busson, F., Borsa, P., Dahruddin, H., Sauri, S., Fitriana, Y., Syamsul, M., Zein, A., Hocdé, R., Pouyaud, L., Keith, P., Wowor, D., Steinke, D., Hanner, R. & Hubert, N. (2020). Biodiversity inventory of the grey mullets (Actinopterygii: Mugilidae) of the Indo-Australian Archipelago through the iterative use of DNA-based species delimitation and specimen assignment methods. *Evol. Appl.* **13**, 1451–1467.
- Dulvy, N., Allen, D., Ralph, G. & Walls, R. (2016). The conservation status of sharks, rays and chimaeras in the Mediterranean Sea. [Brochure]. Malaga, Spain: IUCN.
- Ebert, D.A. & Compagno, L.J.V. (2007). Biodiversity and systematics of skates (Chondrichthyes: Rajiformes: Rajoidei). *Environ. Biol. Fishes* **80**, 111–124.
- Ebert, D.A. & Stehmann, M.F.W. (2013). *Sharks, batoids, and chimaeras of the North Atlantic FAO Species Catalogue for Fishery Purposes*, Vol. No. (7) (p 523). Rome: FAO.
- Ellis, J., Abella, A., Serena, F., Stehmann, M.F.W. & Walls, R. (2015). *Dipturus oxyrinchus*. The IUCN Red List of Threatened Species 2015: e.T63100A48908629. <https://doi.org/10.2305/IUCN.UK.2015-1.RLTS.T63100A48908629.en>
- Estalles, M., Coller, N.M., Perier, M.R. & Di Giacomo, E.E. (2011). Skates in the demersal trawl fishery of San Matias Gulf, Patagonia: species composition, relative abundance and maturity stages. *Aquat. Living Resour.* **24**, 193–199.
- Farrugia, T.J., Goldman, K.J., Tribuzio, C. & Seitz, A.C. (2016). First use of satellite tags to examine movement and habitat use of big skates *Beringraja binoculata* in the Gulf of Alaska. *Mar. Ecol. Progr. Ser.* **556**, 209–221.
- Follesa, M.C., Cannas, R., Cabiddu, S., Cau, A., Mulas, A., Porcu, C. & Cau, A. (2012). Preliminary observations of the reproductive biology and diet for the Norwegian skate *Dipturus nidarosiensis* (Rajidae) from the Central Western Mediterranean Sea. *Cybium* **36**, 473–477.
- Follesa, M.C., Marongiu, M.F., Zupa, W., Bellodi, A., Cau, A., Cannas, R., Colloca, F., Djurovic, M., Isajlovic, I., Jadaud, A., Manfredi, C., Mulas, A., Peristeraki, P., Porcu, C., Ramirez-Amaro, S., Salmerón Jiménez, F., Serena, F., Sion, L., Thasitis, I., Cau, A. & Carbonara, P. (2019). Spatial variability of Chondrichthyes in the northern Mediterranean. *Sci. Mar.* **83**(S1), 81–100.
- Follesa, M.C., Porcu, C., Cabiddu, S., Mulas, A., Deiana, A.M. & Cau, A. (2011). Deep-water fish assemblages in the central-western Mediterranean (south Sardinian deepwaters). *J. Appl. Ichthyol.* **27**, 129–135.
- Fricke, R., Eschmeyer, W.N. & Van der Laan, R. (eds) 2020. Eschmeyer's catalog of fishes: genera, species, references. (<http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp>). Electronic version accessed 07 October 2020.
- Frodella, N., Cannas, R., Velonà, A., Carbonara, P., Farrell, E.D., Fiorentino, F., Follesa, M.C., Garofalo, G., Hemida, F., Mancusi, C., Stagioni, M., Ungaro, N., Serena, F., Tinti, F. & Cariani, A. (2016). Population connectivity and phylogeography of the Mediterranean endemic skate *Raja polystigma* and evidence of its hybridization with the parapatric sibling *R. montagui*. *Mar. Ecol. Progr. Ser.* **554**, 99–113.
- Gabbanelli, V., de Astarloa, J.M.D., Gonzalez-Castro, M., Vazquez, D.M. & Mabragaña, E. (2018). Almost a century of oblivion: Integrative taxonomy allows the resurrection of the longnose skate *Zearaja breviceaudata* (Marini, 1933) (Rajiformes; Rajidae). *C. R. Biol.* **341**(9–10), 454–470.
- García-Melo, J.E., Oliveira, C., Da Costa Silva, G.J., Ochoa-Orrego, L.E., Garcia Pereira, L.H. & Maldonado-Ocampo, J.A. (2019). Species delimitation of neotropical Characins (Stevardiinae): implications for taxonomy of complex groups. *PLoS One* **14**, e0216786.
- Geraci, M.L., Di Lorenzo, M., Falsone, F., Scannella, D., Di Maio, F., Colloca, F., Vitale, S. & Serena, F. (2019). The occurrence of Norwegian skate, *Dipturus nidarosiensis* (Elasmobranchii: Rajiformes: Rajidae), in the Strait of Sicily, central Mediterranean. *Acta Ichthyol. Piscat.* **49**, 203–208.
- Geraci, M.L., Ragonese, S., Norrito, G., Scannella, D., Falsone, F. & Vitale, S. (2017). A tale on the demersal and bottom dwelling chondrichthyes in the south of Sicily through 20 years of scientific survey. In Rodrigues-Filho, L.F. and de Luna, S.J.B. (eds) *Chondrichthyes—Multidisciplinary approach* (pp 13–37). London, UK: IntechOpen.
- Gibbs, J. (2018). DNA barcoding a nightmare taxon: assessing barcode index numbers and barcode gaps for sweat bees. *Genome* **61**, 21–31.
- Griffiths, A.M., Miller, D.D., Egan, A., Fox, J., Greenfield, A. & Mariani, S. (2013). DNA barcoding unveils skate (Chondrichthyes: Rajidae) species diversity in 'ray' products sold across Ireland and the UK. *PeerJ* **1**, e129.
- Griffiths, A.M., Sims, D.W., Cotterell, S.P., El Nagar, A., Ellis, J.R., Lynghammar, A., McHugh, M., Neat, F.C., Pade, N.G., Queiroz, N., Serra-Pereira, B., Rapp, T., Wearmouth, V.J. & Genner, M.J. (2010). Molecular markers reveal spatially segregated cryptic species in a critically endangered fish, the common skate (*Dipturus batis*). *Proc. R. Soc. B* **277**, 1497–1503.
- Griffiths, A.M., Sims, D.W., Johnson, A., Lynghammar, A., McHugh, M., Bakken, T. & Genner, M.J. (2011). Levels of

- connectivity between longnose skate (*Dipturus oxyrinchus*) in the Mediterranean Sea and the north-eastern Atlantic Ocean. *Conserv. Genet.* **12**, 577–582.
- Gubili, C., Macleod, K., Perry, W., Hanel, P., Batzakas, I., Farrell, E.D., Lynghammare, A., Mancusi, C., Mariani, S., Menezes, G.M., Neath, F., Scarcella, G. & Griffiths, A.M. (2016). Connectivity in the deep: phylogeography of the velvet belly lanternshark. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **115**, 233–239.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W. & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321.
- Hickerson, M.J., Meyer, C.P. & Moritz, C. (2006). DNA barcoding will often fail to discover new animal species over broad parameter space. *Syst. Biol.*, **55**, 729–739.
- Hobbs, C.A., Potts, R.W., Walsh, M.B., Usher, J. & Griffiths, A.M. (2019). Using DNA barcoding to investigate patterns of species utilisation in UK shark products reveals threatened species on sale. *Sci. Rep.*, **9**, 1–10.
- Hubert, N., Delrieu-Trottin, E., Irisson, J.O., Meyer, C. & Planes, S. (2010). Identifying early stages of coral reef fishes through DNA barcoding: A test case with the families Acanthuridae and Holocentridae. *Mol. Phylogenetics Evol.* **55**, 1195–1203.
- Hubert, N. & Hanner, R. (2015). DNA barcoding, species delineation and taxonomy: a historical perspective. *DNA Barcodes* **3**, 44–58.
- Hudson, R.R. & Coyne, J.A. (2002). Mathematical consequences of the genealogical species concept. *Evolution* **56**, 1557–1565.
- Iglésias, S.P., Toulhoat, L. & Sellos, D.Y. (2010). Taxonomic confusion and market mislabelling of threatened skates: important consequences for their conservation status. *Aquat. Conserv.: Mar. Freshw. Ecosyst.* **20**, 319–333.
- Jin, Q., Han, H., Hu, X., Li, X., Zhu, C., Ho, S.Y., Ward, R.D. & Zhang, A.B. (2013). Quantifying species diversity with a DNA barcoding-based method: Tibetan moth species (Noctuidae) on the Qinghai-Tibetan Plateau. *PLoS One* **8**, e64428.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120.
- King, J.R. & McFarlane, G.A. (2010). Movement patterns and growth estimates of big skate (*Raja binoculata*) based on tag-recapture data. *Fish. Res.* **101**, 50–59.
- Knowles, L.L. & Carstens, B.C. (2007). Delimiting species without monophyletic gene trees. *Syst. Biol.* **56**, 887–895.
- Kumar, S., Stecher, G. & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874.
- Last, P., Naylor, G., Séret, B., White, W., de Carvalho, M. & Stehmann, M. (Eds.). (2016a). *Rays of the World*. Clayton South: CSIRO Publishing.
- Last, P.R., Weigmann, S. & Yang, L. (2016b). *Changes to the nomenclature of the skates (Chondrichthyes: Rajiformes). Rays of the World: Supplementary Information* (pp 11–34). Clayton: CSIRO Special Publication.
- Leavitt, J.S., Huntsberger, C.J., Smolowitz, R.J. & Siemann, L.A. (2018). The seasonal distribution and abundance of barndoor skate on Georges Bank based on scallop dredge surveys. *Fish. Res.* **199**, 202–211.
- Leigh, J. & Bryant, D. (2015). PopART: full-feature software for haplotype network construction. *Methods Ecol Evol* **6**, 1110–1116.
- Li, Q., Kong, L., Yu, H., Zheng, X., Yu, R., Dai, L., Sun, Y., Chen, J., Liu, J., Ni, L., Feng, Y., Yu, Z., Zou, S. & Lin, J. (2016). DNA barcoding reveal patterns of species diversity among northwestern Pacific molluscs. *Sci. Rep.* **6**, 33367.
- Little, D.P., Knopf, P. & Schulz, C. (2013). DNA barcode identification of Podocarpaceae—the second largest conifer family. *PLoS One* **8**, e81008.
- Lou, M. & Golding, G.B. (2010). Assigning sequences to species in the absence of large interspecific differences. *Mol. Phylogenet. Evol.* **56**, 187–194.
- Lukhtanov, V.A., Sourakov, A., Zakharov, E.V. & Hebert, P.D.N. (2009). DNA barcoding Central Asian butterflies: increasing geographical dimension does not significantly reduce the success of species identification. *Mol. Ecol. Resour.* **9**, 1302–1310.
- Lynghammar, A., Christiansen, J.S., Griffiths, A.M., Fevolden, S.-E., Hop, H. & Bakken, T. (2014). DNA barcoding of the northern Northeast Atlantic skates (Chondrichthyes, Rajiformes), with remarks on the widely distributed starry ray. *Zool. Scr.* **43**, 485495.
- Marongiu, M.F., Porcu, C., Bellodi, A., Cannas, R., Cau, A., Cuccu, D., Mulas, A. & Follesa, M.C. (2017). Temporal dynamics of demersal chondrichthyan species in the central western Mediterranean Sea: the case study in Sardinia Island. *Fish. Res.* **193**, 81–94.
- Meier, R., Shiyang, K., Vaidya, G. & Ng, P.K. (2006). DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Syst. Biol.* **55**, 715–728.
- Melis, R., Vacca, L., Bellodi, A., Cau, A., Porcu, C., Follesa, M.C. & Cannas, R. (2020). Insights into population genetics, connectivity and demographic history of the longnosed skate *Dipturus oxyrinchus* (Linnaeus, 1758) in the western Mediterranean Sea. *Aquat. Conserv.: Mar. Freshw. Ecosyst.* **30**, 86–103.
- Meyer, C.P. & Paulay, G. (2005). DNA barcoding: error rates based on comprehensive sampling. *PLoS Biol.* **3**, e422.
- Moritz, C. (1994). Defining 'Evolutionary Significant Units' for conservation. *Trends Ecol. Evol.* **9**, 373–375.
- Moritz, C. & Cicero, C. (2004). DNA Barcoding: promise and pitfalls. *PLoS Biol.* **2**, e354.
- Mulas, A., Bellodi, A., Cannas, R., Cau, A., Cuccu, D., Marongiu, M.F., Porcu, C. & Follesa, M.C. (2015). Diet

- and feeding behaviour of longnosed skate *Dipturus oxyrinchus*. *J. Fish Biol* **86**, 121–138.
- Neat, F., Pinto, C., Burrett, I., Cowie, L., Travis, J., Thorburn, J., Gibb, F. & Wright, P.J. (2015). Site fidelity, survival and conservation options for the threatened flapper skate (*Dipturus cf. intermedius*). *Aquat. Conserv.: Mar. Freshw. Ecosyst.* **25**, 6–20.
- Orlov, A., Cotton, C. & Byrkjedal, I. (2006). Deepwater skates (Rajidae) collected during the 2004 cruises of R.V. “G.O. Sars” and M.S. “Loran” in the Mid-Atlantic Ridge area. *Cybiu* **30**, 35–48.
- Pasolini, P., Ragazzini, C., Zaccaro, Z., Cariani, A., Ferrara, G., Gonzalez, E.G., Landi, M., Milano, I., Stagioni, M., Guarniero, I. & Tinti, F. (2011). Quaternary geographical sibling speciation and population structuring in the Eastern Atlantic skates (suborder Rajoidea) *Raja clavata* and *R. traeleni*. *Mar. Biol.* **158**, 2173–2186.
- Piemontese, L., Giovannini, I., Guidetti, R., Pellegrini, G., Dioli, P., Maistrello, L., Rebecchi, L. & Cesari, M. (2020). The species identification problem in mirids (Hemiptera: Heteroptera) highlighted by DNA barcoding and species delimitation studies. *Eur. Zool.* **87**, 310–324.
- Pollom, R., Ebert, D.A., Leslie, R. & Winker, H. (2020). *Dipturus pullopunctatus*. The IUCN Red List of Threatened Species 2020 (p e.T44609A124435129.). 10.2305/IUCN.UK.2020-2.RLTS.T44609A124435129.en.
- Porcu, C., Marongiu, M.F., Bellodi, A., Cannas, R., Cau, A., Melis, R., Mulas, A., Soldovilla, G., Vacca, L. & Follesa, M.C. (2017). Morphological descriptions of the eggcases of skates (Rajidae) from the central-western Mediterranean, with notes on their distribution. *Helgol. Mar. Res.* **71**, 10.
- Priede, I., Godbold, J.A., King, N.J., Collins, M.A., Baile, Y.D.M. & Gordon, J.D.M. (2010). Deep-sea demersal fish species richness in the Porcupine Seabight, NE Atlantic Ocean: global and regional patterns. *Mar. Ecol.* **31**, 247–260.
- Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. (2012). ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* **21**, 1864–1877.
- Ramírez-Amaro, S., Ordines, F., Puerto, M.Á., García, C., Ramon, C., Terrasa, B. & Massutí, E. (2017). New morphological and molecular evidence confirm the presence of the Norwegian skate *Dipturus nidarosiensis* (Storm, 1881) in the Mediterranean Sea and extend its distribution to the western basin. *Mediterr. Mar. Sci.* **18**, 251–259.
- Ramírez-Amaro, S., Ordines, F., Terrasa, B., Esteban, A., García, C., Guijarro, B. & Massutí, E. (2016). Demersal chondrichthyans in the western Mediterranean: assemblages and biological parameters of their main species. *Mar. Freshwater Res.* **67**, 636–652.
- Ratnasingham, S. & Hebert, P.D. (2013). A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. *PLoS One* **8**, e66213.
- Rey, J., Coelho, R., Lloris, D., Séret, B. & Gil de Sola, L. (2010). Distribution pattern of *Galeus atlanticus* in the Alboran Sea (south western Mediterranean) and some sexual character comparison with *Galeus melastomus*. *Mar. Biol. Res.* **6**, 364–372.
- Rodríguez-Cabello, C., Pérez, M. & Sánchez, F. (2013). New records of chondrichthyans species caught in the Cantabrian Sea (southern Bay of Biscay). *J. Mar. Biol. Assoc. U.K.* **93**, 1929–1939.
- Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542.
- Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E. & Sánchez-Gracia, A. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **34**, 3299–3302.
- Serena, F. (2005). Field identification guide to the sharks and rays of the Mediterranean and Black Sea. FAO Species Identification Guide for Fishery Purposes. FAO: Rome. 2005. 97p.
- Serra-Pereira, B., Moura, T., Griffiths, A.M., Gordo, L.S. & Figueiredo, I. (2011). Molecular barcoding of skates (Chondrichthyes Chondrichthyes: Rajidae) from the southern Northeast Atlantic. *Zool. Scr.* **40**, 76–84.
- Smith, P.J., Steinke, D., Mcveagh, S.M., Stewarts, A.L., Struthers, C.D. & Roberts, C.D. (2008). Molecular analysis of Southern Ocean skates (Bathyraja) reveals a new species of Antarctic skate. *J. Fish Biol.* **73**, 1170–1182.
- Stevens, J.D., Bonfil, R., Dulvy, N.K. & Walker, P.A. (2000). The effects of fishing on sharks, rays, and chimaeras (Chondrichthyans), and the implications for marine ecosystems. *ICES J. Mar. Sci.* **57**, 476–494.
- Templeton, A.R., Crandall, K.A. & Sing, C.F. (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* **132**, 619–633.
- Thompson, J.D., Higgins, D.G. & Gibson, T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.
- Tudela, S., Simard, F., Skinner, J. & Guglielmi, P. (2004). The Mediterranean deep-sea ecosystems: a proposal for their conservation. In *The Mediterranean deep-sea ecosystems: an overview of their diversity, structure, functioning and anthropogenic impacts, with a proposal for conservation* (pp 39–47). Málaga: International Union for the Conservation of Nature (IUCN) and Rome, World Wide Fund for Nature (WWF).
- Vargas-Caro, C., Bustamante, C., Bennett, M.B. & Ovenden, J.R. (2017). Towards sustainable fishery management for skates in South America: The genetic population structure of *Zearaja chilensis* and *Dipturus trachyderma* (Chondrichthyes, Rajiformes) in the south-east Pacific Ocean. *PLoS One* **12**, e0172255.

- Veríssimo, A., Zaera-Perez, D., Leslie, R., Iglésias, S.P., Séret, B., Grigoriou, P., Steriotti, A., Gubili, C., Barriá, C., Duffy, C., Hernández, S., Batjakas, I.E. & Griffiths, A.M. (2017). Molecular diversity and distribution of eastern Atlantic and Mediterranean dogfishes *Squalus* highlight taxonomic issues in the genus. *Zool. Scr.* **46**, 414–428.
- Vogler, A.P. & DeSalle, R. (1994). Diagnosing units of conservation management. *Cons. Biol.* **6**, 170–178.
- Wannell, G.J., Griffiths, A.M., Spinou, A., Batista, R., Mendonça, M.B., Wosiacki, W.B., Fraser, B., Wintner, S., Papadopoulos, A.I., Krey, G. & Gubili, C. (2020). A new minibarcode assay to facilitate species identification from processed, degraded or historic ray (batoidea) samples. *Conserv. Genet. Resour.* **12**, 659–668.
- Ward, R.D., Hanner, R. & Hebert, P.D.N. (2009). The campaign to DNA barcode all fishes, FISH-BOL. *J. Fish Biol.* **74**, 329–356.
- Wearmouth, V.J. & Sims, D.W. (2009). Movement and behaviour patterns of the critically endangered common skate *Dipturus batis* revealed by electronic tagging. *J. Exp. Mar. Biol. Ecol.* **380**(1–2), 77–87.
- Wiemers, M. & Fiedler, K. (2007). Does the DNA barcoding gap exist?—a case study in blue butterflies (Lepidoptera: Lycaenidae). *Front. Zool.* **4**, 8.
- Zhang, A.B., Feng, J., Ward, R.D., Wan, P., Gao, Q., Wu, J. & Zhao, W.Z. (2012b). A new method for species identification via protein-coding and non-coding DNA barcodes by combining machine learning with bioinformatic methods. *PLoS One* **7**, e30986.
- Zhang, A.B., Hao, M.D., Yang, C.Q. & Shi, Z.Y. (2017). BarcodeR: An integrated R package for species identification using DNA barcodes. *Methods Ecol. Evol.* **8**, 627–634.
- Zhang, A.B., He, L.J., Crozier, R.H., Muster, C. & Zhu, C.D. (2010). Estimating sample sizes for DNA barcoding. *Mol. Phylogenetics Evol.* **54**, 1035–1039.
- Zhang, A.B., Muster, C., Liang, H.B., Zhu, C.D., Crozier, R., Wan, P., Feng, J. & Ward, R.D. (2012a). A fuzzy-set-theory-based approach to analyse species membership in DNA barcoding. *Mol. Ecol.* **21**, 1848–1863.
- Zhang, A.B. & Savolainen, P. (2009). BPSI2.0: a C/C++ interface program for species identification via DNA barcoding with a BP-neural network by calling the Matlab engine. *Mol. Ecol. Resour.* **9**, 104–106.
- Zhang, A.B., Sikes, D.S., Muster, C. & Li, S.Q. (2008). Inferring species membership using DNA sequences with back-propagation neural networks. *Syst Biol.* **57**, 202–215.
- Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* **29**, 2869–2876.

Supporting information

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