Mediterranean Marine Science

Vol 17, No 2 (2016)

DNA BARCODING OF FISH SPECIES FROM THE MEDITERRANEAN COAST OF ISRAEL

A. SHIRAK, LIOR DOR, E. SEROUSSI, M. RON, G. HULATA, D. GOLANI

doi: 10.12681/mms.1384

Research Article

Mediterranean Marine Science
Indexed in WoS (Web of Science, ISI Thomson) and SCOPUS
The journal is available on line at http://www.medit-mar-sc.net
DOI: http://dx.doi.org/10.12681/mms.1384

DNA barcoding of fish species from the Mediterranean coast of Israel

A. SHIRAK[1], L. DOR[1,2], E. SEROUSSI[1], M. RON[1], G. HULATA[1] and D. GOLANI[2]

[1] Institute of Animal Science, Agricultural Research Organization (ARO), Volcani Center, Bet Dagan 50250, Israel
[2] Department of Evolution, Systematics and Ecology, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

Corresponding author: shiraka@agri.gov.il

Handling Editor: Panagiotis Kasapidis

Received: 4 June 2015; Accepted: 2 December 2015; Published on line: 16 May 2016

Abstract

Accurately classified genomic data in the Barcode of Life Data System (BOLD) database is vital to protecting and conserving marine biodiversity in the Mediterranean Sea. The taxonomic classifications of 468 fish of 50 Mediterranean species were analyzed using the BOLD Identifier tool for variation in the mitochondrial gene encoding cytochrome oxidase subunit I (COI). Within species, nucleotide maximum composite likelihood was low with a mean of 0.0044 ± 0.0008. Three presumptive species had significantly higher values: *Arnoglossus* spp. (0.07), *Torquigener flavimaculosus* (0.013) and *Boops boops* (0.028). However, the *Arnoglossus* spp. samples were subclassified into two groups that were ultimately identified as two different species: *Arnoglossus laterna* and *Arnoglossus thori*. For the different species, BLAST searches against the BOLD database using our DNA barcoding data as the query sequences categorized the most similar targets into groups. For each analyzed species, the similarity of the first and second threshold groups ranged from 95 to 99% and from 83 to 98%, respectively. Sequence-based classification for the first threshold group was concordant with morphology-based identification. However, for 34 analyzed species (68%), overlaps of species between the two threshold groups hampered classification. Tree-based phylogenetic analysis detected more than one cluster in the first threshold group for 22 out of 50 species, representing genetic subgroups and geographic origins. There was a tendency for higher conservation and a lower number of clusters in the Lessepsian (Red Sea) migrant versus indigenous species.

Keywords: COI barcoding, sequence-based classification, Lessepsian migration.

Introduction

The conserved sequence of the 5' region of the mitochondrial cytochrome oxidase subunit I (COI) gene has been proposed as a platform for the universal DNA barcoding of life (Hebert *et al*., 2003). To date, about 213,000 of the estimated 4.5 million species (4.7%) have been formally characterized with DNA barcodes and registered in the Barcode of Life Data System (BOLD, www.barcodinglife.org). BOLD was developed as a collaborative online workbench that has evolved into a resource for the DNA barcoding community, including tools for bioinformatics and taxonomical identification based on cluster analysis. Under this system, the Fish Barcode of Life initiative (FISH-BOL, www.fishbol.org) focuses on fish, which comprise half of the global vertebrate species. Rapid growth of the FISH-BOL database has been facilitated by a universal primer cocktail that is capable of amplifying the COI fragment in more than 93% of fish species (Ivanova *et al.*, 2007), and more than 10,600 of the estimated 32,000 fish species (33.1%) have been barcoded so far.

An important prerequisite for the protection and conservation of marine biodiversity is building an inventory of taxonomic references that relies on genetic measurements. Most fish species exhibit profound phenotypic changes during development, and juveniles of related species are hard to distinguish by morphology. In this context, the identification of fish species using DNA barcoding provides new perspectives in fish ecology and systematics (Hanner & Gregory, 2007; Costa & Carvahlo, 2007), as well as in biodiversity surveys, and the phylogenetic and evolutionary history of life (Hebert & Gregory, 2005). DNA barcoding aims to provide an efficient method for species-level identification using an array of species-specific molecular tags derived from a region of the COI gene (Hubert *et al.*, 2008). The resolution of genetic methods is dependent on the sequence divergence among individuals of closely related groups relative to that within groups. COI barcoding is characterized by high intraspecific conservation and high sequence differences between closely related species, subspecies and even populations. In addition, as a mitochondrial gene with multiple copies per cell, the COI fragments can even be amplified from DNA which has been degraded due to aging, harsh storage conditions or food processing (Hajibabaei *et al.,* 2006; Shirak *et al.*, 2012). The worldwide problem of food fraud requires new methods of regulation at the DNA level, such as species identification using mitochondrial DNA barcoding and individual traceability using genomic markers (Ogden, 2008; Chiu *et al.,* 2015). Recent successful identification of fish fraud

using COI barcoding has demonstrated the utility of the procedure and analysis using BOLD inventory (Cutarelli *et al.,* 2014; Galal-Khallaf *et al.,* 2014).

The construction of the Suez Canal enabled the invasion of aquatic species from the Red Sea to the Mediterranean, with numerous migrant fish species of Indo-Pacific origin being observed in the eastern Mediterranean. The influx of species from the Red Sea into the Mediterranean has been denoted "Lessepsian migration" (Por, 1978). Lessepsian migration provides a unique model for investigating the compound processes of migration, invasion and colonization (Golani, 1998; Bariche *et al.,* 2015). The objectives of the present study were to assess, using BOLD, the concordance between DNA barcode- and morphology-based taxonomical classifications of 468 Mediterranean fish, and to compare COI sequence conservation between indigenous and migrant species.

Materials and Methods

Specimen sampling, morphology classification and nomenclature

Fish samples were collected at the port of Jaffa, south of Tel Aviv, from catches in the Mediterranean. Samples were collected over 5 years on four different dates: 15 Feb 2009, 17 Sep 2009, 15 Jan 2010 and 20 Jun 2014. Fish were photographed and measured according to the BOLD specifications (http://www.barcodinglife.org). Fin clips were collected and stored in 70% EtOH for further DNA extraction. The 468 sampled fish were rejects from the fishermen's catches due to their small size or toxicity, and consisted of no more than 6 individuals from each species per collection date (Table S1). Based on morphology, 459 fish were assigned to 48 species, and 9 fish were generally classified to their genus (*Arnoglossus* spp. and *Solea* spp.). Two previously published *Epinephelus* spp. (species 37 and 45, Table S1) were thoroughly explored in a different study (Dor *et al.*, 2014) and thus omitted from this report.

A code was assigned to each sample that was added to the laboratory collection voucher and deposited in FishBol and GenBank (Table S2). Individual codes were composed from: (a) initial letters of genus and species (e.g., LiMo for *Lithognathus mormyrus*); (b) serial number of the species according to Table S1, and (c) capital letter of the serial sample number (1 to 6) according to the four collection dates, respectively (i.e., A–F, G–L, M–R and S–X). For example, the third sample from the second collection date of *Diplodus annularis* was assigned the code DiAn13I (species 13, Table S1).

DNA extraction, COI fragment amplification and sequencing

DNA was extracted from 100–200 mg of caudal fin using MasterPure DNA Purification Kit (Madison, WI, USA, www.epicentre.com) according to the manufacturer's protocol. The final DNA extract was dissolved in 50 µl of ddH_2O and kept at -4ºC. DNA samples were diluted (1:10) and 1 µl was used as the DNA template for PCR amplification of the 663-bp COI fragment. Primer cocktail COI-3 and M13 nested primers were used for amplification and sequencing, respectively, according to a previously published protocol (Ivanova *et al.*, 2007). PCR products were subjected to 1.5% agarose gel electrophoresis. The band of the relevant size was excised from the gel, purified with DNA Montage Gel Extraction Kit (Millipore, Bedford, MA, USA) and then sequenced on an ABI3730XL capillary sequencer.

Sequence alignment and analysis

For each sample, the forward and reverse trace files of the sequenced COI PCR product were assembled using the *GAP4* program (Staden *et al.*, 1999), and the resulting 663-bp sequences were deposited in GenBank via the BOLD system. All 468 DNA sequences were aligned using ClustalW (http://www.genome.jp/tools/clustalw). Nearest-neighbor analysis was conducted with BOLD. Within-species distance was computed by nucleotide maximum composite likelihood using MEGA version 6 based on at least two samples per species (Tamura *et al.*, 2013).

BOLD-based taxonomic classification

A single sequence of the first sample of each species was used as the query sequence for BLASTN search in the BOLD system. The 99 closest matches were examined to detect the most similar taxonomic classification, the related taxonomic groups, and their geographic origin. The first similarity threshold was defined as the lower boundary of sequence similarity that brackets the group of sequences with the highest similarity to the sequence under study, with a distinct gap from the second similarity group.

Comparison between migrant and indigenous species

The range of variability in percentage of COI sequences in the first threshold and the number of phylogeny tree clusters were computed for each species (Table S3). Unpaired Student t-test was used to compare variability in percentage of COI sequences and number of phylogeny tree clusters between migrant and indigenous species.

Results

Table 1 presents 468 fish samples taxonomically classified into 50 species. The number of samples per presumed species varied between 1 and 24. Analysis of COI nucleotide maximum composite likelihood showed low within-species distance in most species, with a mean

Table 1. Fish samples by taxonomic classification.

Family	Taxonomic classification	Common name	Origin[1]	Total No.	Distance[2]
Torpedinidae	*Torpedo torpedo*	Eyed electric ray	1	2	0.0015
Rajidae	*Raja miraletus*	Twineye Skate	1	4	0.0018
Clupeidae	*Sardinella aurita*	Round sardinella	1	6	0.001
Clupeidae	*Herklotsichthys punctatus*	Spotback herring	2	6	0.0022
Congridae	*Ariosoma balearicum*	Bandtooth conger	1	12	0.0003
Synodontidae	*Saurida undosquamis*	Brushtooth lizardfish	2	14	0.001
Synodontidae	*Synodus saurus*	Atlantic lizardfish	1	1	ND
Plotosidae	*Plotosus lineatus*	Striped eel catfish	2	23	0.0004
Atherinidae	*Atherina boyeri*	Big-scale sand smelt	1	6	0.0056
Triglidae	*Trigloporus lastoviza*	Streaked gurnard	1	1	ND
Serranidae	*Serranus hepatus*	Brown comber	1	16	0.0029
Serranidae	*Serranus cabrilla*	Comber	1	7	0.0079
Apogonidae	*Apogon smithi*	Smith's carinalfish	2	21	0
Apogonidae	*Apogonichthyoides pharaonis*	Bullseye cardinalfish	2	1	ND
Apogonidae	*Ostorhinchus fasciatus*	Broadbanded cardinalfish	2	6	0.0018
Carangidae	*Decapterus russelli*	Indian scad	2	14	0.0052
Carangidae	*Trachurus mediterraneus*	Mediterranean scad	1	12	0.0034
Leiognathidae	*Equulites klunzingeri*	Klunzinger's ponyfish	2	19	0.0005
Nemipteridae	*Nemipterus randalli*	Randall's threadfin bream	2	20	0
Haemulidae	*Pomadasys stridens*	Striped grunt	2	2	0
Sparidae	*Pagrus caeruleostictus*	Bluespotted seabream	1	6	0.0043
Sparidae	*Lithognathus mormyrus*	Striped seabream	1	7	0.0018
Sparidae	*Pagellus erythrinus*	Common pandora	1	14	0.0021
Sparidae	*Pagellus acarne*	Axillary seabream	1	16	0.0005
Sparidae	*Diplodus annularis*	Annular seabream	1	11	0.0016
Sparidae	*Boops boops*	Bogue	1	18	0.0277*
Centracantidae	*Spicara maena*	Blotched picarel	1	12	0.0006
Centracantidae	*Spicara smaris*	Picarel	1	4	0.0041
Mullidae	*Upeneus pori*	Por's goatfish	2	13	0.0012
Mullidae	*Mullus barbatus*	Red mullet	1	6	0.0024
Mullidae	*Upeneus moluccensis*	Gold band goatfish	2	2	0
Uranoscopidae	*Uranoscopus scaber*	Atlantic stargazer	1	4	0
Callionymidae	*Callionymus filamentosus*	Blotchfin dragonet	2	18	0.0011
Gobiidae	*Oxyurichthys petersii*	Peter's goby	2	1	ND
Gobiidae	*Gobius niger*	Black goby	1	1	ND
Siganidae	*Siganus rivulatus*	Rabbitfish	2	2	0.0031
Sphyraenidae	*Sphyraena chrysotaenia*	Green- gold barracuda	2	1	ND
Trichiuridae	*Trichiurus lepturus*	Largehead hairtail	1	1	ND
Scombridae	*Scomber colias*	Atlantic chub mackerel	1	6	0.004
Citharidae	*Citharus linguatula*	Atlantic spotted flounder	1	18	0.0013
Bothidae	*Arnoglossus* spp.	Scaldfish	1	8	0.0702*
Bothidae	*Bothus podas*	Wide-eyed flounder	1	18	0.0005
Soleidae	*Soles* spp.	Sole	1	1	ND
Cynoglossidae	*Cynoglossus sinusarabici*	Pelada Del Mar Rojo	2	19	0.0025
Monacanthidae	*Stephanolepis diaspros*	Reticulated leatherjacket	2	21	0.0023
Tetraodontidae	*Lagocephalus sceleratus*	Silverstripe blaasop	2	2	0
Tetraodontidae	*Lagocephalus guntheri*	Half-smooth golden pufferfish	2	17	0.0012
Tetraodontidae	*Lagocephalus suezensis*	Suez pufferfish	2	24	0
Tetraodontidae	*Torquigener flavimaculosus*	Yellowspotted pufferfish	2	4	0.013*
			Total:	**468**	
			Average		**0.0044**

[1]Origin: 1 – indigenous and 2 – Lessepsian migrant species.
[2]Nucleotide maximum composite likelihood; ND – not determined.
*Significantly different from the average nucleotide distance of 41 species ($p < 0.001$).

of 0.0044 ± 0.0008 (Fig. 1). Three presumptive species had significantly higher within-species distance values: *Arnoglossus* spp. (0.07), *Torquigener flavimaculosus* (0.013) and *Boops boops* (0.028). Sequences of the eight individuals that we classified as *Arnoglossus* spp. were subclassified by DNA barcoding-based analysis into two groups at a distance of 0.4, indicating the large taxonomic distances between clusters. The two major clusters consisted of six ArSp5 individuals (A, B, H, I, M and N) belonging to *Arnoglossus laterna*, and two other ArSp5 samples (G and O) belonging to *Arnoglossus thori* (Fig. 2). By comparison, the phylogenetic tree of *Decapterus russelli* was characterized by a 100-fold smaller distance of 0.004 (Fig. 3). This phylogenetic tree was comprised of sequences from the Far East and the Persian Gulf in a distant branch, and two other close branches that included our Mediterranean Sea coast sequences and a sequence originating from Africa. A single SoSp individual (11A) that we classified as belonging to the presumptive *Solea* species, was further classified by BLAST analysis as *Buglossidium luteum*.

BLAST search of our DNA barcoding sequences against the BOLD database revealed lower boundary values of similarity for the different species, ranging from 95 to 100% for the first group of sequences that showed the highest similarity score (first threshold), and from 83 to 99% for the next group of sequences with modest similarity (second threshold). The range of boundaries of similarity values for the first and second thresholds are presented in Table S3. For each threshold category, the number of individuals in BOLD with morphology-based species classification identical to that of our sampled

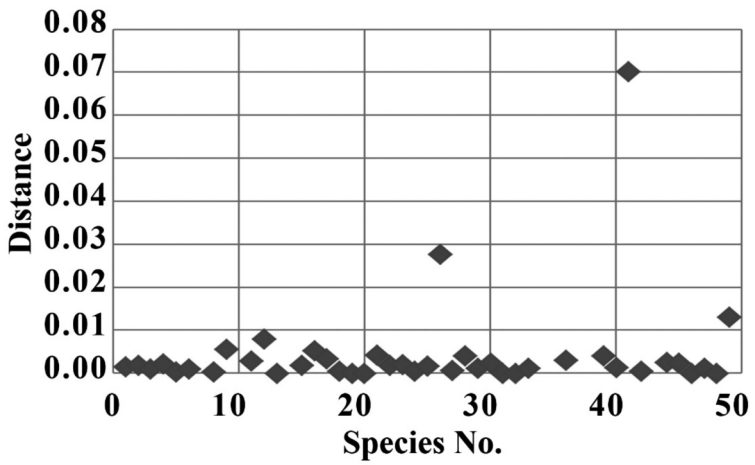

Fig. 1: Mean nucleotide distance values of 49 species. The determined distance values (Table 1) are plotted by species serial number.

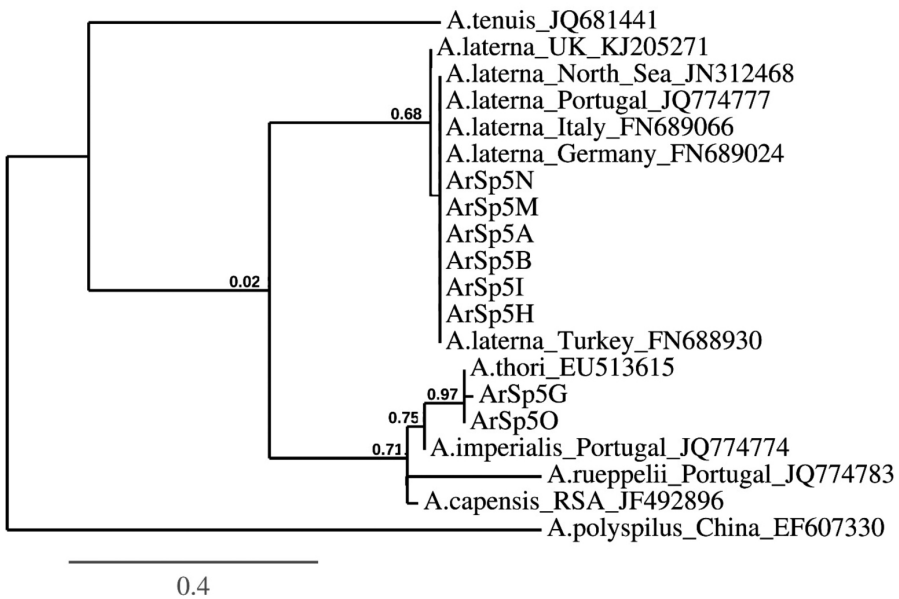

Fig. 2: Phylogenetic tree of *Arnoglossus* spp. The eight samples classified in this study are denoted by prefix ArSp5. Bootstrap values (n = 1000 replicates) are given for each node having 50% or greater support.

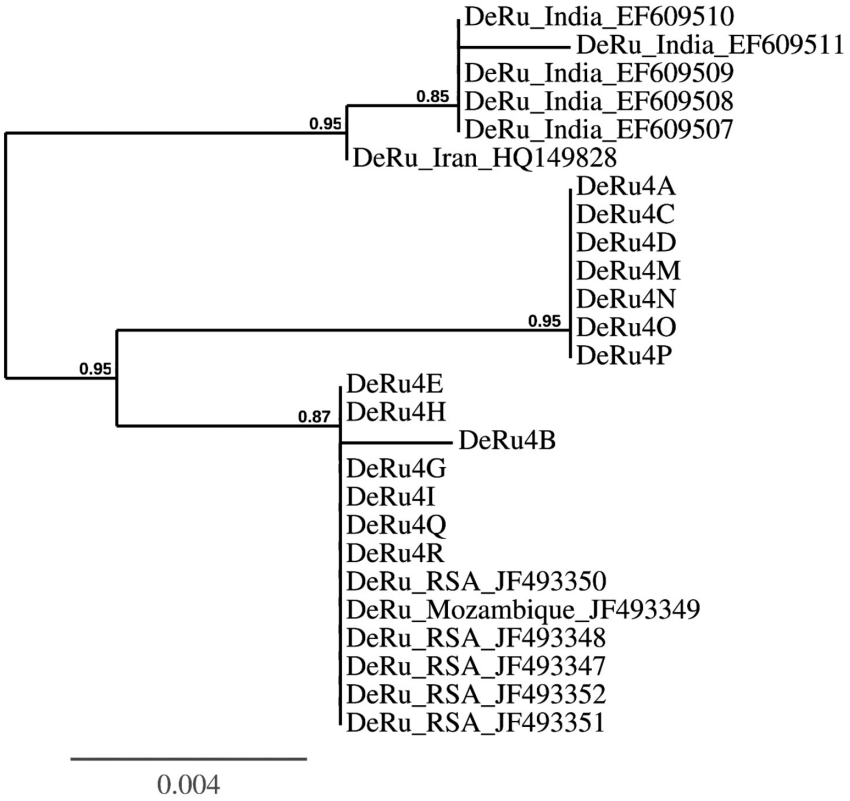

Fig. 3: Phylogenetic tree of *Decapterus russelli*. The 14 samples classified in this study are denoted by prefix DeRu4. Bootstrap values (n = 1000 replicates) are given for each node having 50% or greater support.

individuals are presented in brackets and denoted with a superscript letter "a" in Table S3. The closely related species are denoted with superscript letter "b". For the highest similarity group, DNA barcoding and taxonomic classification in BOLD were in complete concordance with our respective data. For all 50 species, our presumptive taxonomic classification was exclusively or predominantly evident within the first threshold similarity group. The gap between the lower boundary of the first threshold group and the upper boundary of the second threshold group varied between species from 1 to 14%, with an average of 5.9% (Table S3). Nevertheless, the resolving power for identification of the presumptive species in the first threshold group and the related species in the second threshold group, with no overlap, was realized for only 16 of the 50 species (32%). For our samples that were classified to *Lagocephalus guentheri*, *Oxyurichthys petersi*, *Sardinella aurita*, and *Torquigener flavimaculosus*, there were two to three different taxonomic classifications within the first threshold group, one of which corresponded to our presumptive taxonomic classification. Conversely, for seven species, *Diplodus annularis*, *Scomber colias*, *Ariosoma balearicum*, *Plotosus lineatus, Trichiurus lepturus*, *Synodus saurus* and *Raja miraletus*, the same taxonomic name was evident in both the first and second threshold similarity groups. In the case of *Callionymus filamentosus*, there were more than 99 sequences in BOLD; however, analysis using the BOLD identifier is restricted to 99 sequences, and thus additional sequences could not be analyzed. Within the group of analyzed sequences, no related taxonomic group was reported.

In the second threshold group, there were four cases in which the sequence- and morphology-based taxonomies were contradictory. For *Pagellus acarne*, *Citharus linguatula*, *Herklotsichthys punctatus* and *Pomadasys stridens*, the closest species were *Oblada melanura*, *Atherinomorus vaigiensis*, *Pseudocorynopoma doriae* and *Isacia conceptionis*, respectively. Whereas *Pagellus acarne* and *Oblada melanura*, and *Pomadasys stridens* and *Isacia conceptionis* belonged to different Sparidae or Haemulidae families, respectively, *Citharus linguatula* and *Atherinomorus vaigiensis*, and *Herklotsichthys punctatus* and *Pseudocorynopoma doriae* belonged to very distant taxonomic groups (Table S3).

Tree-based analysis detected more than one cluster within the first threshold group for 22 out of the 48 analyzed species, although intraspecies distance was low in most of them. The number of clusters varied between species from 1 to 5 with a mean of 1.6. In most species, our sequences colocalized in clusters with sequences reported from other European and North African Mediterranean countries. Unexpectedly, for the sequences of *Trigloporus lastoviza*, *Upeneus moluccensis* and *Gobius*

niger, our samples also colocalized with clusters from the UK, Far East, and North and Baltic seas. Sequences from Turkey and Portugal showed different patterns of clustering: colocalization with Mediterranean sequences (*Spicara maena*, *Trachurus mediterraneus* and *Trigloporus lastoviza*), separate clusters (*Lithognathus mormyrus*, *Atherina boyeri*, *Spicara smaris*, *Serranus cabrilla* and *Upeneus moluccensis*), or both (*Diplodus annularis*). In most species, sequences originating from the Far East, Persian Gulf, Australia, and North and South America produced separate clusters (*Decapterus russelli, Equlites klunzingeri*, *Upeneus pori*, *Ariosoma balearicum*, *Lagocephalus sceleratus, Apogon smithi, Lagocephalus guentheri* and *Sardinella aurita*).

A comparison between migrant and indigenous species showed a tendency toward a lower number of clusters (1.48 ± 0.68 vs. 1.80 ± 1.00; $p = 0.10$) and a higher range of similarity (0.125 ± 0.141 vs. 0.359 ± 0.586; $p = 0.09$) in the former versus the latter.

Discussion

The power of the barcoding method to identify distinct genetic groups representing taxonomic boundaries is dependent on sequence divergence between and within closely related groups. In the present study, we utilized the BOLD Identifier engine to determine species and minor taxonomic groups of 468 Mediterranean fish specimens. The concordance of DNA- and morphology-based identifications in our study supports the predicted promise of the COI sequence identification platform for taxonomy assessment (Hebert *et al.*, 2003; Hebert & Gregory, 2005; Kochzius *et al.,* 2010; Mofan *et. al.,* 2011; Keskin & Atar, 2013; Landi *et al.,* 2014; Bariche *et al.,* 2015). Based on 1% and 2% intraspecies sequence differences in freshwater and marine species, respectively, 87% of the studied species were identified using BOLD (Keskin & Atar, 2013; Landi *et al.,* 2014). In our study, the BOLD Identifier successfully detected boundaries between the presumptive species and closely related taxonomic groups for 42 of the 50 species (84%). The gap between the two thresholds of similarity for species with undetected boundaries was 1 to 4%. Thus, the two threshold groups should be united into a single group and a third threshold of similarity may be needed to infer the boundary of related taxonomic groups.

Concordance was detected between sequence- and morphology-based taxonomy in the identification of closely related species, except for the following four cases:

(1) the Sparidae classification into three subfamilies (Sparinae, Boopsinae and Denticinae) is based mainly on the tooth structure, which is a controversial criterion (De La Herran *et al.*, 2001). According to morphology, both species, i.e. *Pagellus acarne* and *Oblada melanura* belong to the Sparidae family, but *Pagellus acarne* is related to the Sparinae subfamily and *Oblada melanura* to the Boopsinae subfamily. Based on COI sequence analysis, examination of the Sparinae species *Pagrus caeruleostictus*, *Lithognathus mormyrus*, *Pagellus erythrinus* and *Diplodus annularis* revealed that *Pagellus acarne* is similar to *Oblada melanura*. Therefore, our results support the suggestion that the Sparidae classification should be based on tooth structure. Furthermore, misidentification has been suggested in *O. melanura* samples contained in BOLD (Keskin & Atar, 2013).

(2) Based on COI sequence similarity using BOLD (83%), *Atherinomorus vaigiensis* was the closest species to *Citharus linguatula*. Both fish are distant species that belong to the taxonomic unit Percomorphaceae. This similarity to *Atherinomorus vaigiensis* was higher than to sinistral Citharinae (*Citharoides macrolepidotus*, *Citharoides macrolepis*) and to dextral Brachypleurinae (*Brachypleura novaezeelandiae* and *Lepidoblepharon ophthalmolepis*), which are related to *Citharus linguatula* (Hoshino, 2000). Hence, the taxonomic status of *Citharus linguatula* also needs revising.

(3) Searching for the closest species to *Herklotsichthys punctatus* unexpectedly revealed a single report of *Pseudocorynopoma doriae* that belongs to a different order (Characiformes), with 85% sequence similarity. In BOLD, *Herklotsichthys punctatus* was significantly variable (over 20% difference) from six sequences of *Herklotsichthys quadrimaculatus* and *Herklotsichthys spilurus*; therefore, the *Herklotsichthys* group needs further investigation to correctly classify its taxonomy.

(4) *Isacia conceptionis* was identified as the closest species (87–88% similarity) to *Pomadasys stridens*, which also belongs to the family Haemulidae. This is an unexpected result since BOLD includes more than 80 reports for COI sequences of different *Pomadasys* species with <80% similarity. The *Pomadasys* species are defined mainly based on osteological characteristics and are a principally monophyletic group based on joint analyses of two mitochondrial (including COI) and three nuclear genes (Sanciangco *et al.*, 2011). However, these analyses indicate the equivocal classification that arises from a single gene and an effect of geographic distribution among all Haemulidae. Most of these mismatches in barcoding and current taxonomy have been recently reported by Bariche *et al.* (2015).

Only three out of the 41 species (*Arnoglossus* spp., *Boops boops* and *Torquigener flavimaculosus*) displayed nucleotide maximum composite likelihood values higher than 1%. Multiple taxonomic clusters per species with diverse geographic origins may be responsible for such high intraspecific differences. However, two clusters in BOLD were only found for one of the three species (*Torquigener flavimaculosus*). Landi *et al.* (2014) found five species with 2 to 10% intraspecific differences and each one of the species with two or more clusters. This might result from collecting samples from a wide region of the Mediterranean Sea.

In a number of cases, taxonomic names of the presumed closest species were arguable. Thus, separation of three pairs of related species (*Apogon smithi* and *Apogon ellioti*, *Trachurus mediterraneus* and *Trachurus japonicus*, *Sphyraena chrysotaenia* and *Sphyraena pinguis*) into distinct taxonomic units is controversial (Dutt & Radhakrishna-Rao, 1980; Caputo *et al.*, 1996: Doiuchi & Nakabo, 2005). COI analysis of marine fish species revealed low intraspecies differences for the genus *Trachurus* (Keskin & Atar, 2013; Landi *et al.,* 2014). Similarly, the morphologically closely related shark species, *Carcharhinus altimus* and *Carcharhinus plumbeus*, were not distinguished using COI sequences due to low interspecific variability (Moftah *et al.,* 2011).

Experience with taxonomic identification of fish species is highly variable among taxa and geographic regions. Some reports provide an excellent and reproducible basis for fish identification, whereas others suggest the existence of genetically distant subspecies (Hebert & Gregory, 2005; Keskin & Atar, 2013). A final definition of "species, subspecies and populations" in closely related groups depends on the biology of the species, primarily their reproductive behavior. For example, the cichlid species of the genera *Oreochromis*, *Sarotherodon* and *Tilapia* are taxonomically classified based on the type of parental fry care, i.e., mouth versus external rearing or biparental versus maternal care. We detected lower nucleotide differences between *Sarotherodon galilaeus* and *Oreochromis aureus* than between different *Oreochromis* species, indicating the complexity of taxonomic classification (Shirak *et al.*, 2009). Conventional taxonomy is carried out mostly by examination of specimens' morphometric and meristic characteristics (Cadrin, 2000). The general difficulty of morphology-based identification at early life stages (Victor *et al.*, 2009) was resolved for *Arnoglossus* (Fig. 2) and *Solea* species by identifying genetic subgroups using the BOLD identifier. Three additional species, *Arnoglossus laterna*, *Arnoglossus thori* and *Buglossidium luteum*, were added to the BOLD database. Thus, DNA barcoding is a useful method for indicating those species that require taxonomic reanalysis (Hajibabaei *et al.,* 2007; Lara *et al.*, 2010). Barcoding results can assist in the comparative examination of individuals with different sequence variants to identify tiny differences in morphology and different geographic origins (Moftah *et al.,* 2011). The best candidates in this study for this type of examination were *Decapterus russelli*, *Atherina boyeri*, *Serranus cabrilla*, *Torquigener flavimaculosus* and *Boops boops*. For three species, *Lithognathus mormyrus*, *Spicara smaris*, and *Upeneus moluccensis*, sequences derived from the Mediterranean Sea were localized in separate clusters of the taxonomic tree from those sequences originating from Turkey and Portugal (Keskin & Atar, 2013; Landi *et al.,* 2014). These results may indicate the role of territorial behavior in the same species and subspecies. Conversely, the existence of a reproductive barrier between two separate subspecies that are distributed in the same geographic region can be studied by the analysis of nuclear DNA markers (Martinez-Takeshita *et al.,* 2015) or other mitochondrial markers (Kochzius *et al.,* 2010; Cutarelli *et al.,* 2014) to reveal cryptic diversity (Rougerie *et al.*, 2014).

The tendency toward a lower number of clusters and a higher range of similarity in the migrant versus indigenous species might be explained by a common origin of the limited number of migrants that successfully penetrated and adapted to different conditions of the Mediterranean Sea. The implicated low genetic variability among Red Sea (Lessepsian) species is in accordance with the expected reduction of genetic variability in newly established populations, termed "bottleneck" or "founder effect". The Mediterranean population was presumably established by a small number of founder individuals that constituted only a fraction of the source population of each species (Bernardi *et al.*, 2010).

The present study exemplifies the universal use of a primer cocktail for different taxonomic groups of fish species. Recently, barcoding sequences of Mediterranean fish species have been deposited in the BOLD system by different Israeli research groups, although most of the sequences still have private status. Thus, this research is a preliminary study toward the long-term goal of barcoding all fish species of the Israeli Mediterranean and Red Sea fauna for biodiversity and conservation of fish in Israel.

Acknowledgements

This is a contribution from the ARO, Institute of Animal Science, Bet Dagan, Israel, No. 713/15.

References

Bariche, M., Torres, M., Smith, C., Sayar, N., Azzurro, E. *et al.,* 2015. Red Sea fishes in the Mediterranean Sea: a preliminary investigation of a biological invasion using DNA barcoding. *Journal of Biogeography*, 42(12), 2363-2373

Bernardi, G., Golani, D., Azzurro, E., 2010. The genetics of Lessepsian bioinvasions. In: *Fish Invasions of the Mediterranean Sea: Change and Renewal.* Golani, D. and B. Appelbaum Golani (eds.). Sofia *and Moscow ,* Pensoft *Publishers* pp. 71-84.

Cadrin, S. X., 2000. Advances in morphometric identification of fishery stocks. *Reviews in Fish Biology and Fisheries*, 10(1), 91–112.

Caputo, V., Marchegiani, F., Olmo, E., 1996. Karyotype differentiation between two species of carangid fishes, genus *Trachurus* (Perciformes: Carangidae). *Marine Biology*, 127(2), 193–199.

Chiu, T.H., Kuo, C.-W., Lin, H.C., Huang, D.S., Wu, P.L., 2015. Genetic diversity of ivory shell (*Babylonia areolata*) in Taiwan and identification of species using DNA-based assays. *Food Control*, *48*, 108–116.

Costa, F. O., Carvalho, G. R., 2007. The Barcode of Life Initiative: synopsis and prospective societal impacts of DNA barcoding of Fish. *Genomics Society and Policy*, 3(2), 29–40.

Cutarelli, A., Amoroso, M. G., De Roma, A., Girardi, S., Galiero, G. *et al.,* 2014. Italian market fish species identification and commercial frauds revealing by DNA sequencing. *Food Control*, *37*, 46–50.

De La Herrán, R., Ruiz Rejón, C., Ruiz Rejón, M., Garrido-Ramos, M. A., 2001. The molecular phylogeny of the Sparidae (Pisces, Perciformes) based on two satellite DNA families. *Heredity*, 87(6), 691–697.

Doiuchi, R., Nakabo, T., 2005. The Sphyraena obtusata group (Perciformes: Sphyraenidae) with a description of a new species from southern Japan. *Ichthyological Research*, 52(2), 132–151.

Dor, L., Shirak, A., Gorshkov, S., Ron, M., Hulata, G., 2014. Development of genetic markers for the white grouper (*Epinephelus aeneus*). *Aquaculture*, 420-421, S104–S110.

Dutt, S., Radhakrishna Rao, M. P. R., 1980. *Apogon andhrae*, a new species of apogonid fish from the Bay of Bengal. *Journal of Natural History*, 14(5), 743–748.

Galal-Khallaf, A., Ardura, A., Mohammed-Geba, K., Borrell, Y. J., Garcia-Vazquez, E., 2014. DNA barcoding reveals a high level of mislabeling in Egyptian fish fillets. *Food Control*, 46, 441–445.

Golani, D.,1998. Impact of Red Sea fish migrants through the Suez Canal on the aquatic environment of the Eastern Mediterranean. *Bulletin Series Yale School of Forestry and Environmental Studies*, *103*, 375-387.

Hajibabaei, M., Janzen, D. H., Burns, J. M., Hallwachs, W., & Hebert, P. D. N., 2006. DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences of the United States of America*, 103(4), 968–71.

Hajibabaei, M., Singer, G. A.C., Clare, E. L., Hebert, P. D. N., 2007. Design and applicability of DNA arrays and DNA barcodes in biodiversity monitoring. *BMC Biology*, 5(1), 24.

Hanner, R. H., Gregory, T. R., 2007. Genomic diversity research and the role of biorepositories. *Cell Preservation Technology*, 5(2), 93–103.

Hebert, P. D. N., Gregory, T. R., 2005. The promise of DNA barcoding for taxonomy. *Systematic Biology*, 54(5), 852–859.

Hebert, P. D. N., Cywinska, A., Ball, S. L., de Waard, J. R., 2003. Biological identifications through DNA barcodes. *Proceedings Royal Society / Biological Sciences*, 270(1512), 313–21.

Hoshino, K., 2000. *Citharoides orbitalis*, a new sinistral citharid flounder from western Australia (Pleuronectoidei, pleuronectiformes). *Ichthyological Research*, 47(3-4), 321–326.

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E. *et al*., 2008. Identifying Canadian freshwater fishes through DNA barcodes. *PloS One*, 3(6), e2490.

Ivanova, N. V., Zemlak, T. S., Hanner, R. H., Hebert, P. D. N., 2007. Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes*, *7*(4), 544–548.

Keskin, E., Atar, H. H., 2013. DNA barcoding commercially important fish species of Turkey. *Molecular Ecology Resources*, 13(5), 788–797.

Kochzius, M., Seidel, C., Antoniou, A., Botla, S. K., Campo, D. *et al*., 2010. Identifying Fishes through DNA Barcodes and Microarrays. *PloS One*, *5*(9), e12620.

Landi, M., Dimech, M., Arculeo, M., Biondo, G., Martins, R. *et al.,* 2014. DNA barcoding for species assignment: the case of Mediterranean marine fishes. *PloS One*, *9*(9), e106135.

Lara, A., Ponce de León, J. L., Rodríguez, R., Casane, D., Côté, G. *et al*., 2010. DNA barcoding of Cuban freshwater fishes: Evidence for cryptic species and taxonomic conflicts. *Molecular Ecology Resources*, 10(3), 421–430.

Martinez-Takeshita, N., Purcell, C. M., Chabot, C. L., Craig, M. T., Paterson, C. N. et al., 2015. A tale of three tails: cryptic speciation in a globally distributed marine fish of the genus *Seriola*. *Copeia*, *103*(2), 357–368.

Moftah, M., Abdel Aziz, S. H., Elramah, S., Favereaux, A., 2011. Classification of sharks in the Egyptian Mediterranean waters using morphological and DNA barcoding approaches. *PloS One*, *6*(11), e27001.

Ogden R., 2008. Fisheries forensics: the use of DNA tools for improving compliance, traceability and enforcement in the fishing industry. *Fish and Fisheries*, *9*(4), 462–472.

Por, F. D. 1978. *Lessepsian Migration - The influx of Red Sea biota into the Mediterranean.* Springer. Springer-Verlag Berlin Heidelberg.

Rougerie, R., Kitching, I. J., Haxaire, J., Miller, S. E., Hausmann, A. *et al*., 2014. Australian Sphingidae-DNA barcodes challenge current species boundaries and distributions. *PloS One*, 9(7), e101108.

Sanciangco, M.D., Rocha, L.A., & Carpenter, K. E., 2011. A molecular phylogeny of the Grunts (Perciformes: Haemulidae) inferred using mitochondrial and nuclear genes. *Zootaxa*, 50(2966), 37–50.

Shirak A,, Barkai O, Lernau O, Kahanov Y, Seroussi E. *et al*., 2012. Barcoding analysis of fish bones from a shipwreck found at Dor, Israel. *The Israeli Journal of Aquaculture – Bamidgeh*, 65, 873-878.

Shirak, A., Cohen-Zinder, M., Barroso, R. M., Seroussi, E., Ron, M. *et al*., 2009. DNA Barcoding of Israeli indigenous and introduced cichlids. *The Israeli Journal of Aquaculture – Bamidgeh*, 61, 83–88.

Staden, R., Beal, K., Bonfield, J., 1999. The Staden Package, 1998. In: *Bioinformatics Methods and Protocols*. Misener & S., Krawetz (eds.) Vol. 132, pp. 115–130. Humana Press.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MBE MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), 1–12.

Victor, B. C., Hanner, R., Shivji, M., Hyde, J., Caldow, C., 2009. Identification of the larval and juvenile stages of the Cubera Snapper*, Lutjanus cyanoptems*, using DNA barcoding. *Zootaxa*, 36(2215), 24–36.