COMPARISON OF EUROPEAN SARDINE (SARDINA PILCHARDUS, WALBAUM 1792) GREEK HAPLOTYPES WITH THOSE FOUND IN THE GLOBAL DISTRIBUTION OF THE SPECIES


1Department of Food Science and Technology, School of, International Hellenic University, Alexander Campus, 57400 Sindos, Thessaloniki, Greece
2Department of Nursing, School of Health Sciences, International Hellenic University, Alexander Campus, 57400 Sindos, Thessaloniki, Greece (e-mail: gminos@ihu.gr; phone: +30-2310-013-830)
3Department of Ichthyology and Aquatic Environment, School of Agricultural Sciences, University of Thessaly, 38446 Volos, Greece (e-mail: exadact@uth.gr; phone: +30-2421-093-073)

*Corresponding author
(e-mail: imsiri@ihu.gr; phone: +30-2310-013-381)

(Received 9th May 2021; accepted 12th Aug 2021)

Abstract. The identification of fish stocks is the first step in management and conservation processes. The European Commission has established among others the Protected Geographical Indication (PGI) protection label in order to identify an agricultural product, raw or processed, for which quality, reputation or other characteristics are linked to its geographical origin. In the present study, the Greek Sardina pilchardus haplotypes found with three mitochondrial segments (COI, cyt b, D-loop), were compared with the ones previously submitted in the databases, to test the uniqueness of the discovered Greek haplotypes for a future PGI definition. For all the three mtDNA markers, the discovered Greek haplotypes were found to be common with species' haplotypes all over the world. These results reinforce the aspect that a single sardine individual cannot be classified as coming from a certain population (Mediterranean or Atlantic), as most of the data (present and former) support a single evolutionary unit for sardines.

Keywords: product traceability, country of origin, phylogeography, Clupeidae, DNA barcoding

Introduction

European sardine (Sardina pilchardus, Walbaum 1792) is a small pelagic clupeoid fish, inhabiting the Mediterranean Sea and part of the eastern Atlantic, from the North Sea to Senegal, with peripheral populations around the Azores, Madeira and the Canary Islands (Parrish et al., 1989). Adults usually swim close to the littoral zone, and display daily vertical movement capacity (Olivar et al., 2001). Spawning occurs in open waters and larvae remain in plankton for long periods of time (Olivar et al., 2001). Sardines show schooling and migratory behavior, as well as great dispersal capabilities both at the larval and adult stages. In the Mediterranean and Black Sea, landings between 2014 and 2016 are dominated by small pelagic fish (about 50%), mainly European anchovy and sardine (22% and 16%, respectively). Sardine (S. pilchardus) is the second main species landed (189,500 t on average between 2014 and 2016). In the same area, from the top twenty most important species by value, sardine is the fourth, accounting 9% of the total landing value (74,489,374 $) (FAO, 2018). In Greece, it is one of the two most
commercially important fish (after anchovy), with a total catch of 13,581.1 t in 2019, representing 16.6% of the total catch of marine species (Hellenic Statistical Authority, 2020).

The identification of fish stocks is the first step in management and biodiversity conservation processes (Waldman et al., 2005). Microsatellites are nuclear markers with higher mutation rates that have proved to be efficient and informative for detecting fine-scale population structure in marine pelagic fishes (Ruggeri et al., 2012, 2013). On the other hand, mitochondrial DNA is maternally inherited, lacks recombination, and shows relatively fast evolutionary rates, which make this molecular marker particularly suitable for inferring phylogeographic patterns (Avise, 2000). This marker is also appropriate for detecting historical vicariant or genetic bottleneck events, and has been very useful in describing present day phylogeography of taxa (Avise, 2000). In general, mtDNA genes show more population structure than nuclear genes do (Birky et al., 1989). This makes them particularly suitable indicators for the population genetic differentiation of marine organisms, which are generally high gene-flow species.

In 1992, according to regulation 2081/92, the European Union first adopted the system for the protection of geographical indications of agricultural products and foodstuffs, and according to the regulation 2082/92, the rules on the certificates of specific character for agricultural products and foodstuffs. In 2006, in order to improve the system, the above regulations were replaced by regulations (EC) 510/06 and (EC) 509/06, respectively, without changing their scope and feasibility. "Geographical indication" is a name which identifies a product, raw or processed: originating in a specific place, region or country, the quality, reputation or other characteristics of which are essentially attributable to its geographical origin and at least one of its production steps takes place in the defined geographical area. One hundred and twenty-five Greek products have been characterized as Protected Designation of Origin (PDO)/Protected Geographical Indication (PGI), while there is only one in the category "Fresh fish, molluscs, and crustaceans and products derived from them", Avgotaraho Messolonghiou (PDO - Council Regulations 1107/96 and 1263/96). Avgotaraho or botargo is the cured grey mullet roe. The mature roe pouches of the female grey mullet, Mugil cephalus, are processed as follow: removed from the fish, pressed, salted, sun-dried and dipped several times into bee's wax to preserve the product. It is produced mainly along the western coast of Greece; the most famous is produced outside Messolonghi.

In a previous study (Imsiridou et al., 2019), the PCR and further Sanger sequencing analysis of three mitochondrial segments (cytochrome oxidase subunit I - COI, cytochrome b - cyt b, control region - D-loop) were applied for three Greek S. pilchardus populations, to assess the potential of using the population differentiation of the species as a tool for its future PGI definition. Three S. pilchardus populations were collected by professional fishermen from Amvrakikos Gulf (39° N, 20.8° W), the Ionian Sea (39.55° N, 20.1° W) (Western Greece) and Kalloni Bay in the Aegean Sea (39.15° N, 26.183333° W) (Northern Greece) during July 2017. In total, 641 bp at the 5’ end of the COI gene, 294 bp at the 5’ end of the cyt b gene and 473 bp at the 5’ end of the D-loop region were sequenced for the majority of the individuals. For all the three segments, the statistical analysis revealed low levels of genetic structuring among populations. Additionally, most of the genetic variation was present within samples and the total Fst value was Fst=0.01232, which means that only 1.23% of genetic heterogeneity was apportioned among populations. Thirty-three haplotypes were detected for COI gene,
twenty-five haplotypes were detected for cytb gene and fifty-three haplotypes were found for the D-loop region, in the three S. pilchardus populations. All haplotypes were deposited to GenBank under the accession numbers MH141137-MH141169 (COI haplotypes), MH127862-MH127879 (cytb haplotypes) and MH141170-MH141222 (D-loop haplotypes).

The different haplotypes inside a population represent the various nucleotide sequences its individuals have and consequently the uniqueness of its specimens. In the present study, the species' haplotypes found for each mitochondrial segment in the survey of Imsiridou et al. (2019), were compared with the ones previously submitted in the databases, to test the uniqueness of the discovered Greek haplotypes and consequently their geographic uniqueness, for further characterization of the Greek S. pilchardus populations as Protective Geographic Indication (PGI) products.

Materials and Methods

From the Nucleotide section of GenBank (http://www.ncbi.nlm.nih.gov/), all the registered S. pilchardus haplotypes for each mitochondrial segment (COI, cytb, D-loop) were retrieved. From the generated page of each haplotype, the relevant publication was recovered. Publications related to fish species identification, mislabeling of seafood samples and cooked products were excluded, because the actual S. pilchardus haplotypes should be included in the present study. Finally, the following haplotypes were chosen for each mitochondrial segment:

For COI gene: KJ768296, KJ768297 (Landi et al., 2014 - eighteen samples analyzed), KJ205156 - KJ205158 (Knebelberger et al., 2014 - three samples analyzed), KC501213 - KC501232 (Keskin and Atar, 2013 - twenty samples analyzed), JQ775102 - JQ775107 (Costa et al., 2012 - six samples analyzed), JQ774896 - JQ774900 (Costa et al., 2012 - five samples analyzed), JQ774708 - JQ774712 (Costa et al., 2012 - five samples analyzed), KY176600 - KY176602 (unpublished), HQ340604 (Ardura et al., 2011), EF609451 (Ward and Holmes, 2007 - one sample analyzed), JQ623979 (unpublished), MG729573 - MG729589 (unpublished).

For cytb gene: JQ621900 (unpublished), JQ585750 - JQ585753 (unpublished), JQ237098 - JQ237114 (unpublished), AF291854 - AF291860 (Tinti et al., 2002 - three hundred and seven samples analyzed).

For the D-loop region: DQ139463 - DQ139723 (Atarhouch et al., 2006 - two hundred and sixty-one samples analyzed).

Then, the NCBI haplotypes from GenBank were converted to BioEdit v.7.0.5.3 files (Hall, 1999). Subsequently, the GenBank haplotypes were inserted, together with the discovered Greek haplotypes, into the same BioEdit files and were aligned with the CLUSTAL W algorithm (Thompson et al., 1997), as implemented in the software BioEdit v.7.0.5.3. As a result, three different BioEdit datasets were created, one for each segment. The three BioEdit datasets were processed with the DAMBE6 software package (Xia et al., 2017), which clusters the identical haplotypes. Then, from the "source" feature of GenBank, the locations of all the haplotypes were identified. Median networks for all the three groups of haplotypes of the total dataset were generated with the median joining algorithm (Bandelt et al., 1999), using the NETWORK 4.5.1.6 program.

The maximum likelihood approach was applied based on the Tamura-Nei model (Tamura and Nei, 1993), using MEGA7 software (Kumar et al., 2016), to depict the
phylogenetic relationships among haplotypes. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 200.0000)). The rate variation model allowed for some sites to be evolutionarily invariable. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position.

Results and Discussion

For COI gene: The thirty-three Greek haplotypes for the COI gene were aligned with sixty-three COI haplotypes from GenBank, with the CLUSTAL W algorithm, and visually confirmed. The total dataset included ninety-six sequences and was analyzed with the DAMBE6 software package, in order to detect the common haplotypes. Eight Greek haplotypes (H1, H4, H7, H25, H27, H30, H31, H32) were found to be identical with haplotypes from different Mediterranean and Atlantic Ocean areas (Portugal, North Sea) (Table 1, Figure 1, Figure 2). Figure 3A depicts the phylogenetic relationships among COI haplotypes (data from Table 1).

Figure 1. Map showing the global distribution of the Greek haplotypes. COI-N: North Sea - Atlantic Ocean = H7; COI-P: Portugal - Atlantic Ocean = H1, H25, H30; COI-Ad: Croatia - Adriatic Sea = H1, H4, H25, H27, H31, H32; COI-An: Turkey, Antalya - Mediterranean Sea = H1; COI-Am: Amvrakikos Gulf, Greece = H1, H7; COI-K: Kalloni Bay, Greece = H1, H4, H25, H27, H30, H31, H32; COI-I: Ionian Sea, Greece = H1, H4; Cyt-T: Tunisian coast - Mediterranean Sea = H1, H10; Cyt-Ad: Adriatic Sea = H1, H6, H10; Cyt-Am: Amvrakikos Gulf, Greece = H1, H6; Cyt-K: Kalloni Bay, Greece = H1, H10; Cyt-I: Ionian Sea, Greece = H1; D-loop-Da: Dakhla-Atlantic coast of Morocco = H2; D-loop-Na: Nador-Mediterranean coast of Morocco = H2, H15; D-loop Am: Amvrakikos Gulf, Greece = H2, H15; D-loop-K: Kalloni Bay, Greece = H2
Table 1. Distribution of Greek and GenBank haplotypes per gene, for species Sardina pilchardus

<table>
<thead>
<tr>
<th>COI gene</th>
<th>Greek haplotypes</th>
<th>Location identified</th>
<th>Common GenBank haplotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H1</td>
<td>H4</td>
<td>H7</td>
</tr>
<tr>
<td></td>
<td>Amvrakikos Gulf, Ionian Sea, Kalloni Bay</td>
<td>Kalloni Bay</td>
<td>MG729573</td>
</tr>
<tr>
<td></td>
<td>H4</td>
<td>H7</td>
<td>H25</td>
</tr>
<tr>
<td></td>
<td>Amvrakikos Gulf</td>
<td>Kalloni Bay</td>
<td>MG729580</td>
</tr>
<tr>
<td></td>
<td>H25</td>
<td>H27</td>
<td>H30</td>
</tr>
<tr>
<td></td>
<td>H27</td>
<td>H30</td>
<td>H31</td>
</tr>
<tr>
<td></td>
<td>H30</td>
<td>H31</td>
<td>H32</td>
</tr>
<tr>
<td></td>
<td>H31</td>
<td>H32</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cyt b gene</th>
<th>Greek haplotypes</th>
<th>Location identified</th>
<th>Common GenBank haplotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H1</td>
<td>H6</td>
<td>H10</td>
</tr>
<tr>
<td></td>
<td>Amvrakikos Gulf, Ionian Sea, Kalloni Bay</td>
<td>Kalloni Bay</td>
<td>AF291854/ JQ37101</td>
</tr>
<tr>
<td></td>
<td>H6</td>
<td>H10</td>
<td>AF291855/ JQ585751</td>
</tr>
<tr>
<td></td>
<td>H10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D-loop</th>
<th>Greek haplotypes</th>
<th>Location identified</th>
<th>Common GenBank haplotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H2</td>
<td>H15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Amvrakikos Gulf, Kalloni Bay</td>
<td>Amvrakikos Gulf</td>
<td>DQ139479</td>
</tr>
<tr>
<td></td>
<td>H15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Location identified</th>
<th>Common GenBank haplotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dakhla-Atlantic coast of Morocco/ Nador-Morocco</td>
<td>DQ139597</td>
</tr>
</tbody>
</table>
Imsiridou et al.: Comparison of European sardine (Sardina pilchardus, Walbaum 1792)) Greek haplotypes with those found in the global distribution of the species

Figure 2. Median-joining network of COI haplotypes of the total dataset (Greek haplotypes-black area, GenBank haplotypes-white area). The area of each circle is proportional to the number of individuals exhibiting that haplotype. Each line in the network represents one mutational step.

Figure 3. Maximum likelihood trees, one per each mtDNA part (COI, cyt b, D-loop), illustrating the genetic relationships among haplotypes. The model used was the Tamura-Nei model. Numbers above branches indicate bootstrap values among 1,000 replicates.
These findings are in agreement with previous microsatellite data (Gonzalez and Zardoya, 2007; Kasapidis et al., 2012) which do not support the existence of the two *S. pilchardus* subspecies. The degree of genetic differentiation between Atlantic and Mediterranean populations has been the subject of study for many marine organisms. In many cases, the Mediterranean and the adjacent Atlantic coasts support very similar faunas (Borsa et al., 1997; Bargelloni et al., 2003). For a proportion of species presently occurring in both the Atlantic and the Mediterranean, the Strait of Gibraltar represents a corridor sufficient for gene flow; thus, an ongoing exchange of individuals between the two seas is taking place (see reviews in Borsa et al., 1997). The fact that the Gibraltar Strait and the Almería-Oran front may or not act as barrier to gene flow for different marine pelagic species, has been attributed to differences in life-history traits (e.g. dispersal capacity), and to the existence of distinct past demographical events (e.g. bottlenecks) (Gonzalez and Zardoya, 2007).

For *cytb* gene: The twenty-five *cytb* Greek haplotypes were compared with twenty-seven *cytb* haplotypes from GenBank, with the CLUSTAL W algorithm. The total dataset included fifty-two sequences. Three Greek haplotypes (H1, H6, H10) were found to be common with haplotypes from different regions of the Mediterranean Sea (*Table 1, Figure 1, Figure 4*). The relationships among *cytb* haplotypes are shown in *Figure 3B* (data from *Table 1*).

![Median-joining network of cytb haplotypes of the total dataset](image)

*Figure 4. Median-joining network of cytb haplotypes of the total dataset (Greek haplotypes-black area, GenBank haplotypes-white area). The area of each circle is proportional to the number of individuals exhibiting that haplotype. Each line in the network represents one mutational step.*

A previous survey (Parrish et al., 1989) based on phenotypic variation mainly in gill raker counts and head length, has proposed the existence of two subspecies, that show a different distribution pattern with respect to the Atlantic Ocean-Mediterranean Sea transition: *Sardina pilchardus pilchardus* in the Eastern Atlantic from the North Sea to Southern Portugal, and *S. p. sardina* in the Mediterranean Sea and the Northwest African coast. In the present study, the Greek *cytb* haplotypes were found to be common.
with species haplotypes from the Mediterranean Sea (Adriatic Sea) and Northwest African coast (Tunisian coast). This result supports the existence of *S. p. sardina* subspecies in the Mediterranean Sea, as proposed, based on meristic (Parrish et al., 1989) and mtDNA studies (Atarhouch et al., 2006). Some mitochondrial data might be reflecting past isolation of sardine populations into two distinct groupings during Pleistocene (Atarhouch et al., 2006), whereas microsatellite data reveal the existence of present day gene flow among populations (Gonzalez and Zardoya, 2007; Kasapidis et al., 2012).

Moreover, the Greek Ionian Sea haplotypes were revealed to be common with GenBank Adriatic Sea haplotypes. This finding is consistent with a previous microsatellite analysis of European sardine from Adriatic and Ionian seas (Ruggeri et al., 2013), which indicated a lack of genetic differentiation between the two basins.

For the D-loop region: The fifty-three Greek haplotypes for the D-loop region were aligned with two hundred and sixty-one *D-loop* haplotypes from Genbank, with the CLUSTAL W algorithm. The total dataset included three hundred and fourteen sequences. Two Greek haplotypes (H2, H15) were found to be common with haplotypes from the Atlantic and the Mediterranean coast of Morocco (Table 1, Figure 1, Figure 3C). Median network for the control region haplotypes were generated with the median joining algorithm using the NETWORK 4.5.1.6 program (results not shown).

In the present survey, the Greek COI haplotypes were found to be common with haplotypes from different Mediterranean and Atlantic Ocean regions. Furthermore, the Greek D-loop haplotypes were revealed to be identical with Atlantic and Mediterranean sardine haplotypes. Previous studies also failed to find a barrier to gene flow for sardines in the Atlantic Ocean and the Mediterranean Sea, and to detect significant geographical structuring between sardine population samples in the two basins (Atarhouch et al., 2006; Gonzalez and Zardoya, 2007; Laurent et al., 2007; Chlaida et al., 2009; Baibai et al., 2012). Consequently, for sardine populations the Gibraltar Strait does not present a phylogeographic boundary, as there is no sign of Atlantic-Mediterranean division.

The study of population genetic variation of marine pelagic fish species has proven to be particularly challenging, because of the biological peculiarities of these fishes, including large effective population sizes and high dispersal capacities, as well as because of the apparent lack of physical barriers to gene flow in the marine realm (Avise, 1998; Graves, 1998; Hauser et al., 1998). In the marine environment, natural barriers are often absent and the seawater constitutes a potential means for dispersal, favoring intermixing of individuals over the species range. This phenomenon could commonly lead to high levels of current gene flow as a result of the species high dispersal abilities, both at larval and adult stages (Armason and Palsson, 1996; Vis et al., 1997). Only a few migrants per generation are often sufficient to prevent detectable genetic differentiation between conspecific stocks (Hartl and Clark, 1997). For the three mtDNA markers used in the previous study (Imsiridou et al., 2019), the discovered Greek haplotypes were found to be common with species' haplotypes all over the world. This is evidence of the high gene flow among *S. pilchardus* populations and suggests high panmictia among them. All of these results reinforce the aspect that a single sardine individual cannot be classified as coming from a certain population (Mediterranean or Atlantic), as most of the data (present and former) support a single evolutionary unit for sardines.
Conclusions

For the three mtDNA markers (COI, cytb, D-loop) used from a previous survey, the discovered Greek haplotypes were found to be common with species’ haplotypes all over the world. This is evidence of the high gene flow among S. pilchardus populations. There have been many studies addressing the genetic structure of sardine in different parts of its global distribution, using different molecular markers. The results of these studies suggest low levels of genetic differentiation. Present and former results reinforce the aspect that a single sardine individual cannot be classified as coming from a certain population (Mediterranean or Atlantic), as most of the data support a single evolutionary unit for sardines. Conclusively, from the data so far, the Greek S. pilchardus populations could not be identified as Protective Geographic Indication products, provided that the Greek haplotypes are spread throughout the global distribution of the species. Nevertheless, these findings need to be enriched by more mtDNA or microsatellite data of additional sampling sites in the Ionian and Aegean basins, and further comparison of the new haplotypes with the existing ones, in order to strengthen or definitively reject the idea of characterizing the Greek sardine as a PGI product.

REFERENCES


revealed by a microsatellite analysis of European sardine (Sardina pilchardus) archived samples. – Canadian Journal of Fisheries and Aquatic Sciences 69: 1698-1709.


