Evidence of high genetic connectivity for the longnose spurdog Squalus blainville in the Mediterranean Sea

KOUSTENI V. Faculty of Biology, Department of Zoology-Marine Biology, University of Athens, Panepistimiopolis, Ilisia, 15784 Athens, Greece

KASAPIDIS P. Institute of Marine Biology, Biotechnology and Aquaculture, Hellenic Centre for Marine Research (HCMR), P.O. Box 2214, 71003 Heraklion, Greece

KOTOULAS G. Institute of Marine Biology, Biotechnology and Aquaculture, Hellenic Centre for Marine Research (HCMR), P.O. Box 2214, 71003 Heraklion, Greece

MEGALOFONOU P. Faculty of Biology, Department of Zoology-Marine Biology, University of Athens, Panepistimiopolis, Ilisia, 15784 Athens, Greece

https://doi.org/10.12681/mms.1222

To cite this article:

Evidence of high genetic connectivity for the longnose spurdog *Squalus blainville* in the Mediterranean Sea

V. KOUSTENI1, P. KASAPIDIS2, G. KOTOLAS2 and P. MEGALOFONOU3

1 Faculty of Biology, Department of Zoology-Marine Biology, University of Athens, Panepistimiopolis, Ilisia, 15784 Athens, Greece
2 Institute of Marine Biology, Biotechnology and Aquaculture, Hellenic Centre for Marine Research (HCMR), P.O. Box 2214, 71003 Heraklion, Greece
3 Corresponding author: bkousten@geol.uoa.gr

Handling Editor: Fabrizio Serena

Received: 19 January 2015; Accepted: 20 December 2015; Published on line: 14 March 2016

Abstract

*Squalus blainville* is one of the least studied Mediterranean shark species. Despite being intensively fished in several locations, biological knowledge is limited and no genetic structure information is available. This is the first study to examine the genetic structure of *S. blainville* in the Mediterranean Sea. Considering the high dispersal potential inferred for other squalid sharks, the hypothesis of panmixia was tested based on a 585 bp fragment of the mitochondrial DNA cytochrome c oxidase subunit I gene from 107 individuals and six nuclear microsatellite loci from 577 individuals. Samples were collected across the Ionian, Aegean and Libyan Seas and off the Balearic Islands. Twenty three additional sequences of Mediterranean and South African origin were retrieved from GenBank and included in the mitochondrial DNA analysis. The overall haplotype diversity was high, in contrast to the low nucleotide diversity. Low and non-significant pairwise *F*ₚₛ and *F*ₚₛ values along with a Bayesian cluster analysis suggested high connectivity with subsequent genetic homogeneity among the populations studied, and thus a high dispersal potential for *S. blainville* similar to other squalids. The historical demography of the species was also assessed, revealing a pattern of population expansion since the middle Pleistocene. These findings could be considered in species-specific conservation plans, although sampling over a larger spatial scale and more genetic markers are required to fully elucidate the genetic structure and dispersal potential of *S. blainville*.

Keywords: *Squalus blainville*, Genetic structure, COI gene, Microsatellite loci, Mediterranean Sea.

Introduction

The squalids comprise the second most diverse order of sharks (Squaliformes) (Musick *et al.*, 2004), with seven families and over 130 species (Compagno *et al.*, 2005), mostly benthopelagic with a great variety in size (Compagno, 1984a). Other than their apparent high biodiversity, there is limited information about their population structure (Straube *et al.*, 2011), which hampers the assessment of species-specific conservation efforts, with the latter being focused mainly on teleosts (Forrest & Walters, 2009). Actually, from 1994, a European Union programme (MEDIT) has evaluated the demersal resources of the Mediterranean, including the genus *Squalus* (Bertrand *et al.*, 1997).

Only few squalid species have been genetically studied so far, exhibiting various levels of genetic differentiation. Strong genetic structure was recorded for the spiny dogfish *Squalus acanthias* across the equatorial Pacific (Franks, 2006; Ward *et al.*, 2007; Hauser, 2009; Verissimo *et al.*, 2010), but not between the south Pacific and Atlantic Oceans (Verissimo *et al.*, 2010). Strong genetic differentiation, but only at nuclear level, was also revealed between the Atlantic and Pacific population samples of the longnose velvet dogfish *Centroscymnus crepidater* (Cunha *et al.*, 2012). On the other hand, a single genetic stock was identified within a wide area in the Hawaiian Archipelago for the shortspine spurdog *Squalus mitsukurii* (Daly-Engel *et al.*, 2010) and among several distant locations across the eastern Atlantic for the Portuguese dogfish *Centroscymnus coelolepis* (Verissimo *et al.*, 2011) and the leafscale gulper *Centrophorus squamosus* (Verissimo *et al.*, 2012). Finally, the southern lanternshark *Etmopterus granulosus* was proved not to be endemic to Chile, but widely distributed in the southern hemisphere, being synonymous to New Zealand’s lanternshark *Etmopterus baxteri*. Low levels of population differentiation suggest that *E. granulosus* migrates between these sites, which are separated by thousands of kilometres (Straube *et al.*, 2011). Thus, the dominant pattern emerging from these few studies for squalid species is that of high gene flow among populations where abyssal depths and great distances do not, by themselves, constitute barriers to gene flow.

The longnose spurdog *S. blainville* (Risso, 1827) is a demersal, medium-sized and long-lived squalid shark
Squalus blainville is widely distributed in the continental shelves and upper slopes between 300–500 m of depth (Cannizzaro et al., 1995) down to about 700 m (Whitehead et al., 1984). It has been reported to occupy the whole investigated depth range of the MEDITS programme (10–800 m), even if it is not distributed in a homogeneous way (Serena et al., 2009). It is considered an opportunistic scavenger, feeding primarily on crustaceans and teleosts (Martinho et al., 2012). Through aplacental viviparity, it produces few embryos during all seasons of the year, ranging between 180 and 210 mm of total length (Kousteni & Megalofonou, 2011). It is a commercial species obtained from incidental catches of various demersal fisheries and retained for human consumption (Compagno, 1984a). However, no official records exist for its landings (Kousteni & Megalofonou, 2011).

Squalus blainville is one of the most poorly studied Mediterranean shark species, although it comprises a very common by-catch of bottom trawl fisheries (Damaslas & Vassilopoulou, 2011). Actually, it is listed as data deficient by the IUCN (Nieto et al., 2015), since, despite its relatively well-studied biology (e.g. Sion et al., 2003; Kousteni & Megalofonou, 2011, 2015; Martinho et al., 2012), there is insufficient information on its distribution, population dynamics and genetic structure to meaningfully inform extinction risk and population management models.

Currently, there is growing concern over the extent of shark catches in the Mediterranean Sea, as their inherent characteristics make them intrinsically vulnerable to fishing pressure (Dulvy et al., 2014). Although slow growth, late sexual maturity and low fecundity are common features among chondrichthyans, they are especially pronounced in squids, which are characterized by a very long population doubling time of a few decades to over a century, making them highly vulnerable to overexploitation (Simpfendorfer & Kyne, 2009). In a recent study (Kousteni & Megalofonou, 2015), the minimum and maximum age recorded for S. blainville was 1 and 28 y, respectively, indicating not only the high longevity of the species, but also the fact that it enters the trawl fishery from its first year of life. The extremely slow growth rate (k = 0.03 and 0.08 y⁻¹ for females and males, respectively) and late maturing age (T₉₀ = 17.03 and 11.31 y for females and males, respectively) (Kousteni & Megalofonou, 2015), reveal that S. blainville is an even more slow-growing species than believed to be (Cannizzaro et al., 1995; Marouani et al., 2012). These features, in conjunction with the low fecundity of the species (ovarian = 1–7 number of ripe oocytes; uterine = 1–6 number of embryos per litter) (Kousteni & Megalofonou, 2011), may intensify the fishing pressure exerted on S. blainville populations.

In this study, we present information on the genetic diversity and population structure of S. blainville, mainly from the eastern Mediterranean Sea, by analyzing a fragment of the mitochondrial DNA (mtDNA) cytochrome c oxidase subunit I (COI) gene and six nuclear microsatellite loci. Considering the potential for long-distance dispersal suggested for squalid sharks in previous studies (Daly-Engel et al., 2010; Verissimo et al., 2010, 2011, 2012; Straube et al., 2011; Cunha et al., 2012), but missing basic information on migratory and reproductive behaviour, we test the null hypothesis of genetic homogeneity among the sampled populations of S. blainville. By adding published COI data for S. blainville from other parts of the Mediterranean Sea and from South Africa, we offer useful information for resolving the reported, but debated, wide distribution of S. blainville. Moreover, we perform a phylogenetic analysis including public sequences for two other Squalus species (S. megalops and S. mitsukurii) that have been confused with S. blainville, in order to confirm the taxonomic status of our sampled populations.

Materials and Methods

Sampling and DNA extraction

A total of 577 individuals with total length ranging from 180 to 779 mm (mean = 451 ± 127), were caught using three types of bottom-fishing gears (trawls, longlines and gillnets) from December 2004 to June 2012 at depths between 80 and 744 m. Sampling was determined mainly by the availability of commercial fishing operations. Species identification followed Compagno (1984a). Samples were obtained on board vessels or landing sites at different locations in the eastern Mediterranean Sea (the north Aegean Sea (nAEG), the Cyclades Islands (CYC), the Myrtoan Sea (MYR), the Cretan Sea (CRE), the Libyan Sea (LIB) and the Ionian Sea (ION)), and the western Mediterranean Sea (the Balearic Islands (BAL)) (Fig. 1, Table 1).

Individual samples consisted of fin clips preserved in 95% ethanol and stored at 4°C. Total genomic DNA (gDNA) was extracted from approximately 25 mg of each individual sample following the salting-out protocol by Miller et al. (1988).
Mitochondrial cytochrome c oxidase I PCR and sequencing

A fragment of 738 bp of the mtDNA COI gene was amplified by polymerase chain reaction (PCR). The M13-tailed primers FishF2_t1 and FishR2_t1 were used to generate longer sequencing reads (Ivanova et al., 2007). PCR amplifications were carried out in a volume of 10 μL and contained 1 μL gDNA template, 5 μL trehalose (10%), 1.9 μL ultra-pure water, 1 μL 10x PCR buffer, 0.5 μL MgCl₂ (50mM), 0.2 μL dNTPs (10mM), 0.15 μL of

Table 1. Geographic origin and sampling year for S. blainville sample collections included in the present study.

<table>
<thead>
<tr>
<th>Sampling site</th>
<th>Code</th>
<th>LONG (DD)</th>
<th>LAT (DD)</th>
<th>Sampling year</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mediterranean Sea</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cretan Sea (Aegean)</td>
<td>CRE</td>
<td>25.44–27.67 E</td>
<td>35.66–35.89 N</td>
<td>2011</td>
<td>Present study</td>
</tr>
<tr>
<td>Libyan Sea</td>
<td>LIB</td>
<td>23.76–24.27 E</td>
<td>34.82–35.09 N</td>
<td>2004, 2005</td>
<td>Present study</td>
</tr>
<tr>
<td>Cyprus (Levantine Basin)</td>
<td>LEV</td>
<td>33.13–33.18 E</td>
<td>34.35–34.36 N</td>
<td>2009</td>
<td>ELASMOMED (2010)*</td>
</tr>
<tr>
<td>Sicily</td>
<td>SIC</td>
<td>11.87 E</td>
<td>37.31 N</td>
<td>not specified</td>
<td></td>
</tr>
<tr>
<td>Tuscany (Ligurian Sea)</td>
<td>LIG</td>
<td>9.65 E</td>
<td>43.23 N</td>
<td>2008</td>
<td>ELASMOMED (2010)*</td>
</tr>
<tr>
<td>Balearic Islands</td>
<td>BAL</td>
<td>2.42–2.64 E</td>
<td>39.29–38.98 N</td>
<td>2012</td>
<td>Present study</td>
</tr>
<tr>
<td>Atlantic-Indian Ocean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Africa (Agulhas)</td>
<td>sAFR</td>
<td>21.00 E</td>
<td>-36.00 S</td>
<td>2006</td>
<td>Steinke et al. (2011)*</td>
</tr>
</tbody>
</table>

LONG, longitude; LAT, latitude; DD, decimal degrees; *included in the mtDNA analysis only.

Mitochondrial cytochrome c oxidase I PCR and sequencing

A fragment of 738 bp of the mtDNA COI gene was amplified by polymerase chain reaction (PCR). The M13-tailed primers FishF2_t1 and FishR2_t1 were used to generate longer sequencing reads (Ivanova et al., 2007). PCR amplifications were carried out in a volume of 10 μL and contained 1 μL gDNA template, 5 μL trehalose (10%), 1.9 μL ultra-pure water, 1 μL 10x PCR buffer, 0.5 μL MgCl₂ (50mM), 0.2 μL dNTPs (10mM), 0.15 μL of
each primer (10 μM) and 0.5 units Taq polymerase. PCR amplifications were performed using a T100 thermal cycler (Biorad, Hercule, CA, USA) with initial denaturation of 2 min at 95°C, 35 cycles of 30 s at 94°C, 45 s at 51°C and 45 s at 72°C, and a final extension step for 10 min at 72°C.

Agarose gel electrophoresis was used to check the integrity of each PCR product (1 μL) and the remaining volume was purified by ethanol precipitation. The forward and – in cases where the full sequence was not obtained – the reverse strands of the amplicons were sequenced using the ABI Big Dye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems Inc., Foster City, CA, USA) and resolved on an ABI 3730xl DNA sequencer (Applied Biosystems Inc., Foster City, CA, USA). PCR reactions were repeated for each primer (10 μM) and 0.5 units Taq polymerase. PCR amplifications were performed using a T100 thermal cycler (Biorad, Hercule, CA, USA) with initial denaturation of 2 min at 95°C, 35 cycles of 30 s at 94°C, 45 s at 51°C and 45 s at 72°C, and a final extension step for 10 min at 72°C.

The accuracy in nucleotide base assignment of the sequencing reactions was performed with an initial denaturation of 3 min at 96°C, followed by 40 cycles of 20 s at 96°C, 15 s at 55°C and 4 min at 60°C.

The 107 sequences obtained in this study were deposited in GenBank (accession numbers: KU198487–KU198565 and KU198567–KU198594). Additionally, 23 sequences were retrieved from GenBank and included in the mtDNA analysis: 3 from the north Aegean Sea (nAEG; accession numbers: HQ603895–HQ603896 and KP192409) by Zambournis et al. (2010) and Gafas et al. (2015), 5 off Cyprus (LEV; accession numbers: GU805914–GU805917, GU805924) by ELASMOMED (2010), 3 off Malta (MAL; accession numbers: KJ709923–KJ709925) by Landi et al. (2014), 4 off Sicily (SIC; accession numbers: GU805902, GU805904–GU805905, GU805907) and 5 off Tuscany (LIG; accession numbers: GU805857–GU805860, GU805862) by ELASMOMED (2010), and 3 off South Africa (sAFR; accession numbers: JF494583–JF494585) by Steinke et al. (2011). Finally, the 130 mtDNA COI sequences represented eleven sample collections according to their geographic origin.

Nuclear microsatellite loci genotyping

Twelve nuclear microsatellite loci, described for S. acanthias in McCauley et al. (2004) (DFH429, DFH434, DFJ451, DFT289, DFU273, DFU285 and DFV296) and in Verissimo et al. (2010) (Saca3853, Saca4234, Saca6396 and SacaGA11), were initially tested for cross-species amplification in S. blainville. The forward primers were fluorescently labelled with FAM, HEX, ROX or TAMRA. Finally, a total of 577 individuals were genotyped at seven nuclear microsatellite loci (DFH429, DFH434, DFT289, DFU285, Saca3853, Saca6396 and SacaGA11), which were chosen after testing for consistent amplification. The seven loci were amplified in two different multiplex PCR reactions (A: DFH429, DFU285, Saca3853 and Saca6396; B: DFH434, DFT289 and SacaGA11) using the QIAGEN Multiplex PCR Kit (Qiagen, Chatsworth, CA, USA) and protocol, and sized on an ABI 3730xl along with the GeneScan 500 LIZ size standard (Applied Biosystems Inc., Foster City, CA, USA). PCR reactions were repeated for about 20% of the individuals in order to fill in missing data because of amplification or electrophoresis failures and to confirm genotypes.

Genotypes were scored manually using the STRand v2.4.59 software package (Veterinary Genetics Laboratory, University of California, Davis, CA, USA). In order to minimize microsatellite alleles miscalling, the binning

Table 2. Genetic diversity indices for S. blainville sample collections inferred from the mtDNA COI fragment (585 bp).

<table>
<thead>
<tr>
<th>Sample collection</th>
<th>N</th>
<th>Hap</th>
<th>P</th>
<th>H ± SD</th>
<th>N ± SD</th>
<th>k ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>nAEG</td>
<td>31</td>
<td>13</td>
<td>13</td>
<td>0.834 ± 0.029</td>
<td>0.0035 ± 0.002</td>
<td>1.96 ± 1.138</td>
</tr>
<tr>
<td>CYC</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1.000 ± 0.500</td>
<td>0.0034 ± 0.004</td>
<td>2.01 ± 1.737</td>
</tr>
<tr>
<td>CRE</td>
<td>18</td>
<td>7</td>
<td>7</td>
<td>0.726 ± 0.094</td>
<td>0.0027 ± 0.002</td>
<td>1.56 ± 0.980</td>
</tr>
<tr>
<td>LIB</td>
<td>22</td>
<td>7</td>
<td>7</td>
<td>0.593 ± 0.117</td>
<td>0.0022 ± 0.002</td>
<td>1.30 ± 0.847</td>
</tr>
<tr>
<td>LEV</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>0.900 ± 0.161</td>
<td>0.0041 ± 0.003</td>
<td>2.41 ± 1.564</td>
</tr>
<tr>
<td>ION</td>
<td>32</td>
<td>10</td>
<td>9</td>
<td>0.796 ± 0.051</td>
<td>0.0028 ± 0.002</td>
<td>1.66 ± 1.000</td>
</tr>
<tr>
<td>MAL</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>1.000 ± 0.272</td>
<td>0.0046 ± 0.004</td>
<td>2.68 ± 1.930</td>
</tr>
<tr>
<td>SIC</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>0.833 ± 0.222</td>
<td>0.0026 ± 0.002</td>
<td>1.51 ± 1.124</td>
</tr>
<tr>
<td>LIG</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0.700 ± 0.218</td>
<td>0.0021 ± 0.002</td>
<td>1.20 ± 0.911</td>
</tr>
<tr>
<td>BAL</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>0.700 ± 0.218</td>
<td>0.0027 ± 0.002</td>
<td>1.61 ± 1.132</td>
</tr>
<tr>
<td>sAFR</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1.000 ± 0.272</td>
<td>0.0069 ± 0.006</td>
<td>4.02 ± 2.740</td>
</tr>
<tr>
<td>Overall</td>
<td>130</td>
<td>31</td>
<td>29</td>
<td>0.763 ± 0.034</td>
<td>0.0029 ± 0.002</td>
<td>1.69 ± 0.999</td>
</tr>
</tbody>
</table>

N, number of individuals; Hap, number of haplotypes (unique haplotypes); P, number of polymorphic sites; H, haplotype diversity; π, nucleotide diversity; k, mean number of nucleotide differences between haplotypes; SD, standard deviation. The sample collections codes are explained in Table 1.
of alleles was assisted using the FLEXIBIN v2 software package (Amos et al., 2007). MICROCHECKER v2.2.3 (Van Oosterhout et al., 2004) was used to check for null alleles, large allele dropout and stutter peaks. All 577 genotypes were grouped into seven sample collections according to their geographic origin.

**Phylogenetic analysis of three Squalus species**

A neighbour-joining (NJ) tree (Saitou & Nei, 1987), using the Kimura 2-parameter (K2P) distance (Kimura, 1980) with 1,000 bootstrap pseudo-replications, was created with the MEGA software package to infer the phylogenetic relationships between the three Squalus species (S. blainville, S. megalops and S. mitsukurii). The mtDNA COI haplotypes of S. blainville recognized in this study were matched with the public mtDNA COI sequences (N = 23) of the same species and with those (N = 16) of S. megalops off Australia (accession numbers: DQ108268–DQ108271, EU399029–EU399031) by Ward et al. (2005, 2008) and off Japan (accession number: GU130698) by Straube et al. (2010), and of S. mitsukurii off South Africa (accession numbers: JF494587–JF494592) by Steinke et al. (2011) and off Uruguay (accession numbers: EU074610–EU074611) by Mabragana et al. (2011).

**Genetic diversity**

The genetic variability of the mtDNA COI sequences (number of singleton and parsimony informative sites) and the genetic p-distance among haplotypes were estimated in MEGA. Molecular diversity indices for the mtDNA COI fragment, including the number of haplotypes (Hap), the number of polymorphic sites (P), haplotype (h) and nucleotide (π) diversity, and the mean number of nucleotide differences between haplotypes (k) were calculated for each sample collection and overall in ARLEQUIN v3.5.1.2 (Excoffier et al., 2005).

For nuclear microsatellite data, deviations from Hardy-Weinberg equilibrium (HWE) for each locus within each sample collection and tests for linkage disequilibrium between each pair of loci within and among all sample collections, were performed in GENEPOP v4.2 (Raymond & Rousset, 1995) with Markov chain lengths obtained through 10,000 dememorization steps followed by 100 of 5,000 iterations per batch. Indices of genetic diversity, including the number of alleles per locus (A), observed (H_o) and expected (H_e) heterozygosity were calculated for each locus within each sample collection and overall in ARLEQUIN. Allelic richness (A_r) was calculated in FSTAT v2.9.3.2 (Goudet, 2001), while the number of private alleles (A_p) was assessed in GENALEX v6.5 (Peakall & Smouse, 2006). The inbreeding coefficient (F_is) and its significance for each sample collection were estimated in GENEPOP. Sequential Bonferroni correction was used to adjust the levels of significance (Rice, 1989).

**Genetic differentiation and population structure**

The divergence among mtDNA COI haplotypes was depicted with a median-joining network (Bandelt et al., 1999) constructed using the NETWORK v.4.6.1.1 software package (Fluxus Technology Ltd., Suffolk, UK). The pairwise genetic structure was assessed by calculating both Φst values, based on the mean number of pairwise differences among mtDNA COI sequences (Slatkin, 1995) and Fst values, based on the microsatellite allele frequencies (Weir & Cockerham, 1984), in ARLEQUIN. Statistical significance was assessed using 1,000 permutations, following the sequential Bonferroni adjustments (Rice, 1989) and a less strict correction method suggested by Benjamini and Hochberg (1995).

For each type of marker, the hierarchical analysis of molecular variance (AMOVA) (Excoffier et al., 1992) was implemented in ARLEQUIN. Considering the geographic origin of the individuals, three a priori groupings were tested: (a) one group (pamminixa), including the Aegean (AEG) sample collections (nAEG, CYC and CRE), as well as the LIB and ION sample collections, (b) two groups, the first including the AEG and LIB, and the second the ION, and (c) three groups, the first including the AEG, the second including the ION and the third including the LIB. The LEV, MAL, SIC, LIG, BAL and SAFR sample collections were excluded from the analysis because of their small size.

A Bayesian cluster analysis was performed in STRUCTURE v2.3 (Pritchard et al., 2000) to assess the potential number of genetic clusters present within the dataset without prior knowledge of the geographic origin of the individuals. By minimizing linkage disequilibrium and deviations from HWE expectations, a cluster is equivalent to a population or gene pool. Based on the admixture model with correlated allelic frequencies, as some mixing between sample collections is expected, five independent runs with a number of potential genetic clusters (K), ranging from one to five, were carried out. A burn-in period of 250,000 steps followed by a run phase of one million Markov-chain Monte Carlo (MCMC) simulations was used. The generated results were imported into the STRUCTURE HARVESTER software package (Earl & von Holdt, 2012) and the ad hoc statistic ΔK was calculated, based on the second order rate of change of the log probability of data between the successive K values (Evanno et al., 2005). The K value, where ΔK had the highest value was identified as the most probable number of clusters (Pritchard et al., 2000).

**Demographic history**

All demographic analyses were conducted using the entire mtDNA dataset in ARLEQUIN. Deviations
from selective neutrality were tested with the $F_s$ estimator (Fu, 1997), which is expected to be significantly negative under a population expansion scenario. A mismatch distribution of pairwise nucleotide differences among the mtDNA COI haplotypes was compared with the expectation under a sudden-expansion model (Rogers & Harpending, 1992). In samples from expanding populations, the pairwise differences between haplotypes form a unimodal pattern, while in samples from populations in demographic equilibrium, a multimodal pattern of numerous sharp peaks is observed. Harpending’s raggedness index ($r$), quantifying the smoothness of the observed mismatch distributions, was computed to determine the goodness of fit to a unimodal distribution (Harpending, 1994). The statistical significance was tested using 1,000 permutations. A significant raggedness index value ($P < 0.05$) was taken as evidence for rejecting the sudden population expansion scenario (Schneider & Excoffier, 1999).

The relative time since lineage expansion ($t$) was estimated based on the equation $t = 2\mu t$, where $\mu$ is the mode of the mismatch distribution and $\tau$ is the mutation rate for the entire sequence, assumed to be equal for all populations (Rogers & Harpending, 1992). Given the lack of species-specific mutation rate for the COI gene, the rate of $2.38 \times 10^{-9}$ substitutions per site per year reported for the COI gene in angel sharks (Stelbrink et al., 2010) was used.

Finally, in order to reconstruct past population dynamics, a Bayesian skyline plot model was computed in BEAST v1.7.5 (Drummond et al., 2012). A strict molecular clock was assumed under the HKY+I substitution model of molecular evolution determined by MODELTEST v2.1.4 (Darriba et al., 2012) using the substitution rate of Stelbrink et al. (2010). Each subset was run twice for 100 million iterations of the MCMC chains, following a burn-in of 10 million iterations and sampled every 1,000 steps, starting from an UPGMA tree and applying default Bayesian priors. The Bayesian skyline reconstructions were conducted in TRACER v1.5 (Rambaut & Drummond, 2007). The time to the most recent common ancestor ($t_{MRCA}$) was inferred for the Bayesian skyline plot model.

Results

Phylogenetic analysis of three Squalus species

The K2P neighbour-joining tree for the mtDNA COI haplotypes of S. blainville obtained in this study, along with the public mtDNA COI sequences of S. blainville, S. megalops and S. mitsukurii, is shown in Fig. S1 that is included in the Supplementary Material. The sequences of S. blainville and of the Australasian S. megalops form a distinct monophyletic clade with high bootstrap support. Squalus mitsukurii forms another monophyletic clade, while the Japanese S. megalops individual seems to be phylogenetically closer to the S. mitsukurii clade than to the Australasian S. megalops. A haplotype network of all S. blainville individuals, together with the Australasian S. megalops ones, shows that, although closely related, the latter form a distinct clade (Fig. S2 in Supplementary Material). This indicates that these two taxa may represent recently diverged species or even a single species.

Our results clearly illustrate the taxonomic uncertainties within the genus Squalus and confirm that further study is required to elucidate their taxonomic status and relationships. In any case, it is clear that the S. blainville individuals analyzed in this study belong to the same species.

Genetic diversity

A total of 31 different haplotypes were defined (Table 2) out of 130 mtDNA COI sequences, based on 29 polymorphic sites (19 singleton and 10 parsimony informative sites). Twenty-three haplotypes were recognized as private. The presence of private haplotypes, even in the sample collections with few individuals (LEV, SIC, BAL and sAFR), along with the medium to high percentages of private haplotypes at other sites (29% in LIB and 40–54% in nAEG, CRE and ION), are indicative of the high genetic diversity of the species. The genetic $p$-distance between haplotypes was low and ranged from 0.17 to 1.03%. The haplotype diversity was high in almost all sample collections, while the nucleotide diversity was low, reflecting the low number of polymorphic sites distinguishing the different haplotypes.

Statistically significant deviations from HWE after Bonferroni correction were detected in three out of the seven amplified loci: DFH429 (in nAEG), Saca6396 (in CYC) and DFU285 (in nAEG, MYR and ION). Pairwise comparisons between loci did not reveal any linkage disequilibrium. Evidence for null alleles and stutter peaks were revealed for the locus DFU285 only, which was excluded from further analysis. Finally, the following six loci were used: DFH429, DFH434, DFT289, Saca3853, Saca6396 and SacaGA11.

The allelic richness had similar values in all sample collections. The genetic diversity, measured as $H_{s}$, was also similar among the sample collections, reaching its highest values in the MYR, CRE and LIB (Table 3). The number of alleles per locus ($A$) ranged from 5 (DFT289) to 39 (SacaGA11). A total of 103 alleles were recorded, corresponding to a mean number of 17.17 alleles per locus. The allelic richness was low for the DFT289 (1.39) and DFH434 (2.77) loci, while it showed medium to high values in the other loci (range = 4.80–8.10), with the ScaGA11 locus being the most polymorphic (Table S1 in Supplementary Material).
Genetic differentiation and population structure

The haplotype network had a star-like shape (Fig. 1) and consisted of two central haplotypes (hapl1 and 2) separated by two mutational steps. The hapl1 was observed in all sample collections, except the sAFR, and shared by 44.6% of the individuals, while the hapl2 was observed in all collections, except LEV, MAL and sAFR, and shared by 18.5% of the individuals. The third most abundant haplotype (hapl3) showed 6.9% frequency in the whole dataset, while the remaining haplotypes were less frequent (< 5%). Eight haplotypes were found at more than one location and 22 out of the 23 private haplotypes were found in only one individual. Most of the private haplotypes (N = 13) were recorded in the nAEG (Table S2 in Supplementary Material). The overall haplotype distribution in the network did not show any evidence of phylogeographic structure.

For mtDNA, pairwise $\Phi_{ST}$ values ranged from zero to 0.269 and were found to be non-significant, with the exception of LIB vs SIC, which were significantly different after Bonferroni, but not Benjamini and Hochberg correction. However, the SIC sample size is small for reliable inferences to be made. For nuclear microsatellite data, $F_{ST}$ values were practically zero in all pairwise comparisons (Table 4). Both $\Phi_{ST}$ and $F_{ST}$ values indicated that the sample collections comprised a single panmictic population.

The results of AMOVA analysis were generally concordant between mtDNA and nuclear microsatellite data. The null hypothesis of panmixia was not rejected ($\Phi_{ST} = -0.0119$ and $F_{ST} = 0.0009$, $P > 0.05$). The among-group differentiation was zero and not significant when considering either two ($\Phi_{CT} = -0.0026$ and $F_{CT} = -0.0003$, $P > 0.05$) or three groups ($\Phi_{CT} = 0.0146$ and $F_{CT} = -0.001$, $P > 0.05$).

The genetic homogeneity of the sample collections that was inferred from the previous analyses was also consistent with the outcome of the STRUCTURE analysis. Although the most probable number of genetic clusters $K$ was equal to 4 (average log probability of data $\text{Ln}[P(D/K)] = -13,188.1 \pm 179.9$), it was statistically weak. Each cluster was represented with the same proportion in all sample collections reflecting the absence of population structure (Pritchard et al., 2000).

Demographic history

Since no genetic structure was found, Fu’s neutrality test, mismatch distribution analysis and skyline plots were computed for the entire mtDNA dataset. Fu’s index was statistically significantly negative ($F_{S} = -23.66$, $P < 0.001$), indicating an excess of rare mutations, which combined with the non-significant raggedness index ($r = 0.027$, $P = 0.92$) is equivalent to a recent population expansion for $S$. blainville. The frequency graph of the observed and expected pairwise differences between haplotype types showed a unimodal mismatch distribution pattern, thus confirming a sudden population expansion distribution for the whole collection of samples (Fig. 2a).

Through mismatch distribution analysis, population expansion was estimated at about 756 thousand years before the present (ky BP) (318,519–1,290,771 at 95% confidence interval), supporting a Pleistocene population expansion event scenario for $S$. blainville. The Bayesian skyline plot also confirmed sudden population growth, which was estimated to have started less than 400 ky BP, and thus later than predicted by the mismatch distribution analysis, but within its confidence interval limits (Fig. 2b).

Discussion

To our knowledge, this is the first study assessing the genetic diversity and structure of the longnose spurdog $S$. blainville, in the Mediterranean Sea. Based on two types of molecular markers, no genetic structure was revealed for the species, since both haplotype (for mtDNA) and genotype (for nuclear microsatellite loci) variation were homogeneously distributed among the sample collections, without any discernible geographic pattern. This was more evident for the mtDNA dataset, which represented a wider geographic distribution in the Mediterranean Sea. The close relationship of the South African haplotypes to the Mediterranean ones showed that the distribution of $S$. blainville extends at least to South Africa. The observed genetic homogeneity of $S$. blainville in the Mediterranean indicated that the species probably has a high dispersal potential similar to other squalid sharks. Furthermore, the star-like haplotype network and the presence of private haplotypes at almost all sampling sites were consistent with a relatively recent historical demographic expansion. Both mismatch distribution and skyline plot analyses suggested a Pleistocene population expansion event, corresponding to the general view that much of the contemporary phylogeographic pattern of the northern hemisphere biota reflects the climatic fluctuations during the Pleistocene glacial-interglacial cycles (Hewitt, 2000). However, more extended sampling and the use of additional nuclear markers are needed to fully elucidate the genetic structure and historical processes that formed the present phylogeographic pattern of the species.

The genetic diversity, based on both mitochondrial and nuclear DNA markers, has been examined in several chondrichthyans, with some bias against squalid sharks (Table S3 in Supplementary Material). Even though these studies differed in the sampling method used, the molecular markers applied and the biology/ecology of the studied species, a pattern of high haplotype and low nucleotide diversity was observed in most cases, which usually corresponds to a relatively recent bottleneck, fol-
allowed by rapid population growth and accumulation of mutations (Grant & Bowen, 1998). In *S. blainville*, this was evident from the shallow haplotype network where mostly single nucleotide differences were found between the two central haplotypes and the derived ones. The low nucleotide diversity is not uncommon in chondrichthans (e.g. Chevolot et al., 2006; Schultz et al., 2008; Ahonen et al., 2009) because of the slow mutation rates in their mtDNA (Martin et al., 1992).

In *S. blainville*, only half of the initially tested nuclear microsatellite loci, which were originally designed for *S. acanthias* by McCauley et al. (2004) and Verissimo et al. (2010), were successfully amplified. A subset of these loci has also been used in population genetics studies of other squalids (Table S3 in Supplementary Material). The observed and expected heterozygosity for *S. blainville* were within the range of previously reported values for other species ($H_o = 0.38–0.83$ and $H_e = 0.49–0.84$, respectively), and actually among the highest ones (Table S3 in Supplementary Material), indicating a rather large effective population size.

Both mitochondrial and microsatellite DNA analyses pointed to a single genetic stock for *S. blainville*, at least in the eastern Mediterranean Sea where sufficient sampling was performed. This is in agreement with the results obtained for other squalid sharks and reflects their great dispersal potential. For instance, in *S. acanthias*, no genetic differentiation was detected at a larger spatial scale, between the south Pacific and Atlantic Oceans, which was attributed to the well-described transoceanic movements of the species. However, a strong genetic subdivision was found across the equatorial Pacific, which appeared to coincide with local differences in the life-history traits of the species and was consistent with the presence of a physical barrier created by warm-temperate and tropical waters in low latitudes (Verissimo et al., 2010). The high dispersal potential of squalids was also pronounced in *S. mitsukurii*, which was found to form a single breeding population across a distance of 2,000 km in the Hawai-

Fig. 2: Haplotype mismatch distribution based on the observed and expected pairwise nucleotide differences between *S. blainville* haplotypes under the sudden expansion model (a). Bayesian skyline plot derived from the mtDNA COI fragment (585 bp) for all sample collections. The x axis represents the time since the present in years and the y axis represents the estimated effective population size multiplied by the generation time. The solid line is the median estimate and the dotted lines show the 95% HPD limits. The time to the most recent common ancestor ($t_{MRCA}$) is shown by a vertical dashed line (b).
Table 3. Genetic diversity indices for *S. blainville* sample collections inferred from 6 nuclear microsatellite loci: DFH429, DFH434 and DFT289 described in McCauley et al. (2004), and Saca3853, Saca6396 and SacaGA11 described in Veríssimo et al. (2010).

<table>
<thead>
<tr>
<th>Sample collection</th>
<th>N</th>
<th>A</th>
<th>A_e</th>
<th>H_o</th>
<th>H_e</th>
<th>F_{IS}</th>
</tr>
</thead>
<tbody>
<tr>
<td>nAEG</td>
<td>183</td>
<td>15.83</td>
<td>5.23</td>
<td>0.69</td>
<td>0.68</td>
<td>0.017</td>
</tr>
<tr>
<td>CYC</td>
<td>136</td>
<td>15.17</td>
<td>5.13</td>
<td>0.68</td>
<td>0.67</td>
<td>0.009</td>
</tr>
<tr>
<td>MYR</td>
<td>50</td>
<td>12.83</td>
<td>5.33</td>
<td>0.70</td>
<td>0.69</td>
<td>0.015</td>
</tr>
<tr>
<td>CRE</td>
<td>31</td>
<td>11.33</td>
<td>5.28</td>
<td>0.70</td>
<td>0.72</td>
<td>-0.022</td>
</tr>
<tr>
<td>LIB</td>
<td>37</td>
<td>10.83</td>
<td>5.05</td>
<td>0.70</td>
<td>0.68</td>
<td>0.018</td>
</tr>
<tr>
<td>ION</td>
<td>135</td>
<td>15.17</td>
<td>5.14</td>
<td>0.68</td>
<td>0.65</td>
<td>0.036</td>
</tr>
<tr>
<td>BAL</td>
<td>5</td>
<td>4.67</td>
<td>5.14</td>
<td>0.65</td>
<td>0.60</td>
<td>0.089</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>577</td>
<td>12.26</td>
<td>5.19</td>
<td>0.69</td>
<td>0.67</td>
<td>0.023</td>
</tr>
</tbody>
</table>

N, number of individuals; A, mean number of alleles; A_e, mean allelic richness; H_o, mean observed heterozygosity; H_e, mean expected heterozygosity; F_{IS}, mean inbreeding coefficient. The sample collections codes are explained in Table 1.

Table 4. Pairwise $\Phi_{ST}$ values for mtDNA (below diagonal) and pairwise $F_{IS}$ values for nuclear microsatellite data (above diagonal) among *S. blainville* sample collections.

<table>
<thead>
<tr>
<th>nAEG</th>
<th>CYC</th>
<th>MYR</th>
<th>CRE</th>
<th>LIB</th>
<th>LEV</th>
<th>ION</th>
<th>MAL</th>
<th>SIC</th>
<th>LIG</th>
<th>BAL</th>
<th>sAFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>nAEG</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.001</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CYC</td>
<td>-0.296</td>
<td>-0.001</td>
<td>-0.002</td>
<td>0.001</td>
<td>-0.002</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MYR</td>
<td>-</td>
<td>-</td>
<td>-0.004</td>
<td>0.006</td>
<td>-0.002</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>CRE</td>
<td>-0.028</td>
<td>-0.278</td>
<td>-</td>
<td>-0.002</td>
<td>0.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LIB</td>
<td>0.002</td>
<td>-0.141</td>
<td>-0.022</td>
<td>-</td>
<td>0.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LEV</td>
<td>-0.047</td>
<td>-0.409</td>
<td>-0.038</td>
<td>0.014</td>
<td>-</td>
<td>0.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ION</td>
<td>-0.021</td>
<td>-0.297</td>
<td>-0.021</td>
<td>0.009</td>
<td>-0.047</td>
<td>-</td>
<td>-</td>
<td>0.005</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>MAL</td>
<td>-0.070</td>
<td>-0.436</td>
<td>-0.052</td>
<td>0.014</td>
<td>-0.218</td>
<td>0.065</td>
<td>-</td>
<td>0.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>SIC</td>
<td>0.068</td>
<td>-0.350</td>
<td>0.134</td>
<td>0.269*</td>
<td>-0.013</td>
<td>0.085</td>
<td>0.029</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LIG</td>
<td>-0.054</td>
<td>-0.232</td>
<td>-0.078</td>
<td>-0.079</td>
<td>-0.071</td>
<td>-0.041</td>
<td>-0.204</td>
<td>0.236</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>BAL</td>
<td>-0.082</td>
<td>-0.459</td>
<td>-0.085</td>
<td>0.019</td>
<td>-0.137</td>
<td>-0.081</td>
<td>-0.147</td>
<td>-0.002</td>
<td>0.093</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>sAFR</td>
<td>0.045</td>
<td>-0.365</td>
<td>0.058</td>
<td>0.120</td>
<td>-0.077</td>
<td>0.075</td>
<td>-0.154</td>
<td>0.048</td>
<td>-0.025</td>
<td>-0.066</td>
<td>-</td>
</tr>
</tbody>
</table>

*Significant value after Bonferroni correction ($P$-value < 0.001), but not after Benjamini and Hochberg correction. The sample collection codes are explained in Table 1.

Thus developed non-oceanodromous behaviour, avoiding crossing open seas or abyssal plains. On the other hand, *S. blainville* is able to explore deeper waters, and furthermore, by living at the edge of the continental shelf, which by definition is discontinuous compared to the inner part of the shelf, it may have to cross open water masses frequently. This behaviour would offer a much higher potential for dispersal to *S. blainville*, thus enhancing population connectivity. Moreover, the drop of the sea level during the Pleistocene glaciations may have affected the shallower habitats where *S. canicula* breeds to a greater extent, enforcing its genetic discontinuity between the Aegean and Ionian Seas (Kousteni et al., 2014), but not the deeper habitats that *S. blainville* prefers.

The two species also differ in their maximum size, a variable intimately associated with the vagility of chondrichthians (Musick et al., 2004), and in their reproductive strategy. *Squalus blainville* reaches larger sizes and gives birth directly to actively swimming embryos (Compagno, 1984a), while *S. canicula* reaches smaller sizes...
(Compagno, 1984b) and it lays its embryos in protective egg cases anchored to solid structures on the substratum (Wheeler, 1978). These traits support a higher dispersal potential for *S. blainville* in contrast to *S. canicula*, which is known to exhibit a high degree of site fidelity (Rodríguez-Cabello et al., 2004). Strong genetic structure was also observed between the western and eastern Mediterranean sample collections of *S. canicula* (Kousteni et al., 2014), but no comparisons can be made with *S. blainville* because of the small number of individuals obtained off the Balearic Islands.

Nevertheless, it should be added that the lack of genetic differentiation among *S. blainville* sample collections does not exclude the presence of different harvest stocks. No detectable genetic differentiation between samples is consistent with any model of population structure ranging from complete panmixia to the exchange of as little as 1% individuals (Ward, 2000). Thus, a subtle genetic structure might exist, but the relatively small number of microsatellite loci used in this study may have not been sufficient to reveal it.

**Conclusion**

This study provides a first insight into the phylogeographic pattern and genetic structure of the longnose spurdog *S. blainville* in the Mediterranean Sea, indicating high genetic connectivity and lack of genetic differentiation in this area. However, the restricted geographic coverage of samples across the distribution of the species and the limited number of nuclear markers used cannot provide a conclusive picture of the evolutionary history and stock structure of the species. Broadening the geographic representation of the species in conjunction with analyzing more genetic and other natural markers (e.g. parasites), as well as mark-recapture data, would provide a better perspective of its population structure and clarify whether the species has indeed high dispersal potential similar to other squalid sharks. Given that the identification of genetic stocks could be used to implement separate assessment of fishing levels for each of them (Ward, 2000), if the managed stock does not correspond to the true biological stock, the effectiveness of management plans may be compromised (Verissimo et al., 2011). Finally, it is clear that a global assessment of the molecular taxonomy of the species is necessary to clarify its distribution and taxonomic status.

**Acknowledgements**

This study has been conducted within the framework of the PhD thesis of Vasiliki Kousteni and co-financed by the European Union (European Social Fund-ESF) and Greek national funds through the Operational Programme ‘Education and Lifelong Learning’ of the National Strategic Reference Framework (NSRF)-Research Funding Programme: Hercule II: investing in a knowledge society through the European Social Fund. We deeply thank Mrs. Chryssi Mytilineou and Dr Smith Christopher J. for providing a part of the samples from the Ionian Sea through the CoralFISH project (European Community’s FP7 under grant N° 213144; Co-financing was provided to HCMR from the Greek General Secretariat for Research and Technology). Dr Beatriz Guijarro is also acknowledged for collecting and providing the fin samples off the Balearic Islands.

**References**


Mabragana, E., Diaz de Astarloa, J.M., Hanner, R., Zhang, J., Gonzalez Castro, M., 2011. DNA barcoding identifies...
argentine fishes from marine and brackish waters. *PloS ONE*, 6(12), e28655.


