Marine Genomics xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Marine Genomics

journal homepage: www.elsevier.com/locate/margen

RAD SNP markers as a tool for conservation of dolphinfish *Coryphaena hippurus* in the Mediterranean Sea: Identification of subtle genetic structure and assessment of populations sex-ratios

Francesco Maroso [a,*,1], Rafaella Franch [a,1], Giulia Dalla Rovere [a], Marco Arculeo [b], Luca Bargelloni [a]

[a] Department of Compared Biomedicine and Food Science, University of Padova, viale dell'Università, 16, 35020 Legnaro, Italy
[b] Department STEBICEF, University of Palermo, CoNISMa, via Archirafi, 18, 90123 Palermo, Italy

ARTICLE INFO

Article history:
Received 19 May 2016
Received in revised form 12 July 2016
Accepted 13 July 2016
Available online xxxx

Keywords:
2bRAD
Genetic differentiation
Outliers
Sex determination markers

ABSTRACT

Dolphinfish is an important fish species for both commercial and sport fishing, but so far limited information is available on genetic variability and pattern of differentiation of dolphinfish populations in the Mediterranean basin. Recently developed techniques allow genome-wide identification of genetic markers for better understanding of population structure in species with limited genome information. Using restriction-site associated DNA analysis we successfully genotyped 140 individuals of dolphinfish from eight locations in the Mediterranean Sea at 3324 SNP loci. We identified 311 sex-related loci that were used to assess sex-ratio in dolphinfish populations. In addition, we identified a weak signature of genetic differentiation of the population closer to Gibraltar Strait in comparison to other Mediterranean populations, which might be related to introgression of individuals from Atlantic. No further genetic differentiation could be detected in the other populations sampled, as expected considering the known highly mobility of the species. The results obtained improve our knowledge of the species and can help managing dolphinfish stock in the future.

1. Introduction

Dolphinfish, *Coryphaena hippurus* Linnaeus, 1758, is an important target species for artisanal, recreational, and commercial fisheries. It is considered a mid-trophic level pelagic fish (Merten et al., 2016), with high potential for dispersal. Dolphinfish are found in tropical and subtropical waters, including the Mediterranean Sea. It is considered a fast growing species with an estimated maximum longevity of four years in Caribbean Sea (Beardsley, 1967; Rose and Hassler, 1974), even if in the Mediterranean no individuals exceeding two years of age have been found (Gatt et al., 2015). It is available for fishing during the summer season (Massutí and Morales-Nin, 1997), with sport fishing targeting larger individuals, and commercial fishing aiming at the juvenile stage (25–60 cm fork length (FL)), which corresponds to individuals aged between 2 and 8 months (Benseddik et al., 2011). While the negative effects of fishing juveniles are known (e.g. reduction of future yield and recruitment for the species), there is no specific minimum size regulation for this species (though some exceptions exist, as for example in Sardinia, with a 60 cm minimum size, www.sardegnaambiente.it/documenti/19_4_20080215151247.pdf). Especially for dolphinfish this can be a major issue, considering the different behavior of fish of different age. It is known that under fish aggregating devices (FADs), used by commercial and sometimes also by recreational fishermen, mostly female and young males are found, while adult males prefer open waters as they move between female dominated rafts (Oxenford, 1999). Thus, fishing around FADs could lead to alteration in the sex ratio at particular life stages. Sexual dimorphism is present but morphological differences arise only when sexual maturity is reached, usually from May to October of the first year, at 60 cm FL (Gatt et al., 2015), in both sexes. Sexual dimorphism is evident in large individuals as males develop a typical bulging squared-off forehead, which is not present in females (Massutí and Morales-Nin, 1997). For younger specimens, sex can only be determined by histological analysis of gonads, which is a time consuming and often difficult task.

Restriction enzyme Associated DNA (RAD) refers to a family of genotyping techniques that use the cutting activity of restriction enzymes and selection of resulting fragments to obtain a reduced representation library (RRL) of individual DNA that will subsequently be sequenced. Indeed, for the aims of many genetic approaches (e.g. population genetic studies) the information provided by only a small portion of the entire genome is sufficient, and requires less sequencing effort to achieve enough depth of coverage for reliable genotyping. These library preparation techniques, combined with high throughput (up to 1.5 terabases for the latest Illumina technologies machines), of Next

* Corresponding author.
E-mail address: francesco.maroso@studenti.unipd.it (F. Maroso).
[1] These two authors contributed equally to the work.

http://dx.doi.org/10.1016/j.margen.2016.07.003
1874-7787/

Generation Sequencing (NGS) machines, allow the simultaneous analysis of multiple individuals at the same time (which is typical for population genetic studies) at reduced costs. One of the approaches derived from the original RAD is the 2bRAD (Wang et al., 2012), which exploits the cutting activity of type IIB restriction enzymes to cut specific sites in the genome and retrieve uniform length fragments (centered at the enzyme's recognition sequence), shared by all the individuals analyzed. Among the advantages of this particular RAD technique, the relatively simple laboratory approach (e.g. no need for shearing, no agarose gel size selection) and its flexibility are the two most frequently addressed. In particular, the availability of different combinations of enzymes-adaptors can be used to adapt the number of markers analyzed according to the needs of any specific study or to the species addressed. From a bioinformatic point of view, the homogeneous length of the fragments and the presence of the restriction enzyme recognition site in the center of the sequences facilitate clustering steps, especially in species lacking reference genome resources. This approach has already been proved effective in fish population genetic analysis (e.g. in tuna by Pecoraro et al. (2016)), and allowed the identification of previously undiscovered population genetic structure.

In the present study, for the first time, a large set of SNP markers was identified with 2bRAD and used to study the population genetic variability of dolphinfish, providing a robust tool for determining sex and showing preliminary evidence for subtle genetic divergence within the Mediterranean basin despite the large potential for dispersal of the species. This work was carried out in the framework of the Ritmare project (http://www.ritmare.it/en/), the Italian flagship research project on marine biology for the period 2012–2015.

2. Material and methods

2.1. Sampling design, libraries preparation and sequencing

Fin clips from 169 juvenile dolphinfish (FL range 36–64 cm) from eight different landing localities across the Mediterranean Sea (Fig. 1) were collected.

Genomic DNA (gDNA) was extracted using either commercial kits (Invisorb® Spin Tissue Mini Kit (Invitek, STRATEC Biomedical, 242 Germany) and Real Genomics Tissue DNA Extraction kit (RBC Bioscience, Taiwan)) or the SSTNE buffer, a modified TNE buffer added of spermidine and spermine (Pardo et al., 2005).

Genomic libraries were constructed following the 2bRAD protocol first introduced by Pecoraro et al. (2016), with minor modifications. In brief, gDNA (300 ng) was digested with 2 U of the enzyme CspCI (New England Biolabs, NEB, Ipswich, Massachusetts, USA) for 1 h at 37 °C. The digested DNA was ligated in a 25 μL total volume reaction consisting of 0.4 μM for each of the two library-specific adaptors, 0.2 mM ATP (New England Biolabs) and 1 U T4 DNA ligase (SibEnzyme Ltd., Academ town, Siberia). To reduce marker density, one adaptor with fully degenerate 3′ overhangs NN and one with reduced 3′degeneracy NG were chosen. Sample-specific barcodes were designed with Barcode Generator (http://comailab.genomecenter.ucdavis.edu/258index.php/Barcode_generator) and introduced by PCR with platform-specific barcode-bearing primers (P6-BC). In order to minimize PCR amplification bias (Mastretta-Yanes et al., 2015), 2b-RAD tags were amplified splitting in three wells a 60 μL mixture consisting of 12.5 μL of ligated DNA, 0.5 μM each primer P4 and P6-BC (Eurofins Genomics S.r.l, Italy), 0.2 μM each primer P5 and P7 (Eurofins Genomics), 0.3 mM dNTP (New England Biolabs), 1× Phusion HF buffer and 1 U Taq Phusion high-fidelity DNA polymerase (ThermoFisher Scientific). Cycling conditions were: 98 °C for 4 min; 98 °C for 5 s, 60 °C for 20 s, 72 °C for 5 s for 14 for 5 cycles, 72 °C for 5 min. The reduced number of amplification cycles (n = 14) is crucial to decrease the amount of PCR amplification errors and the ratio of GC rich fragments. PCR products were purified with the SPRIselect purification kit (Beckman Coulter, Pasadena, California, USA), to exclude any low-molecular weight DNA remaining after PCR amplification. The concentration of purified individual libraries was quantified using Qubit® ds DNA BR Assay Kit (Invitrogen–ThermoFisher Scientific) and Mx3000P qPCR instrument. The quality of a subset of purified libraries was checked on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, California, USA), before sending for sequencing. Samples were equally pooled into three libraries and sequenced on an Illumina HiSeq2500 platform with 50 bp single-end read module at the Genomix4Life S.r.l. facilities (Baronissi, Salerno, Italy). To assess the robustness of the method, technical replicates (TRs) for 14 specimens were prepared by constructing three independent libraries for each replicated individual.

2.2. SNP discovery and filtering

Standard demultiplexing and quality filtering of raw data were performed by the sequencing service provider following Illumina protocols. Subsequently, a custom-made script (available upon request) was used to retain only reads with the *Csp*CI recognition site and trim them to 32 base pairs (bp) long fragments (centering on the recognition site).

Stacks' pipeline 'denovo_map.pl' (version 1.35) was used to cluster obtained reads and identify SNPs across samples (Catchen et al., 2013, 2011). Demultiplexed reads were first clustered on a single-sample basis (subroutine *ustacks*), with minimum coverage (parameter – m) of 5× and maximum number of three mismatches between reads (parameter – M). *cstacks* was then used to merge tags from single individuals and define a catalog of tags with maximum number of three

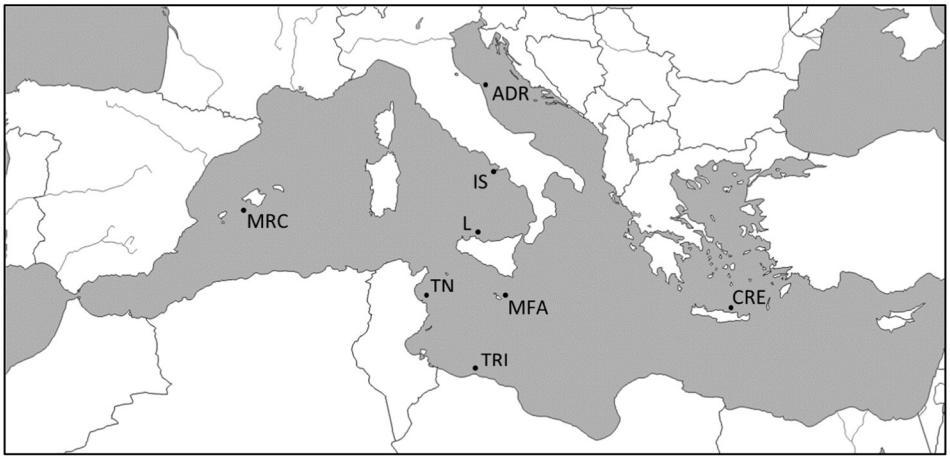

Fig. 1. Sampling locations. Mediterranean sites surveyed in the present study (MRC = Majorca, IS = Ischia, L = Porticello, TN = Tunisia, MFA = Malta, TRI = Libya, ADR = Ancona, CRE = Crete).

Please cite this article as: Maroso, F., et al., RAD SNP markers as a tool for conservation of dolphinfish *Coryphaena hippurus* in the Mediterranean Sea: Identification of subtle genetic structure and assessment of populations sex-ratios, Mar. Genomics (2016), http://dx.doi.org/10.1016/j.margen.2016.07.003