

A Knowledge Base for fish and fishery products

A.M. Pappalardo¹, F. Guarino¹, A. Messina¹, A. Pulvirenti², R. Giugno², A. Ferro²✉, V. De Pinto¹

¹Dipartimento di Sc. Biologiche, G. e A., Università di Catania e INBB, Istituto Nazionale Biostrutture e Biosistemi, Roma, Italy

²Dipartimento di Biomedicina Clinica e Molecolare, Università di Catania, Italy

Motivations

The species subject of our work belong to different orders of Teleosts, one of which is the order of Clupeiformes consisting in two large families of commercial interest (Clupeidae and Engraulidae), which include the small fish commonly known, respectively, for herring, sardines, shads, and sprats and anchovies. Their molecular identification is useful because a common fraud is their species replacement. In addition, because of globalization, it is possible to find in our markets also other species of anchovy typical of different distribution areas, such as Japanese anchovy (*Engraulis japonicus*), the Peruvian Anchovy (*E. ringens*), the Atlantic Ali (*E. anchoita*) and the Californian anchovy (*E. mordax*). Due to the differentiation of the processed foodstuffs that characterizes the market, it is essential to develop tools for unequivocal and quick identification of the species present in the market even when morphological identification is no longer possible.

Methods

DNA barcoding consists of the utilization of short DNA sequences to identify organisms, in particular the species. The quest for a genetic marker useful to determine unambiguously the species is still a matter of debate. For animals, the best candidate to this role has been proposed to be the cytochrome oxidase I (COI) gene. In vertebrates the mtCOI gene is 1545 bp long, a region of 648 bp close to the beginning of the translated sequence is properly said to be the "barcode". We applied the DNA barcoding technologies, upon a segment of COI to compare fish belonging to Clupeiformes order and the analysis of 5'Dloop sequences to stock identification [1]. The fish was used for extracting genomic DNA, amplifying and sequencing. The sequences obtained were aligned using ClustalX. Relationships among the sequence obtained were examined using Neighbour-joining (NJ) and Bayesian analyses. The NJ tree was constructed using pairwise

distances calculated following the application of Kimura's two-parameter (K2P) correction for multiple substitutions in MEGA v. 4.0. The robustness of internal branches of distance was estimated by bootstrapping with 1000 replicates. Modeltest v. 3.06 was used to select the most appropriate models of sequence evolution for the Bayesian analysis that was implemented in MrBayes v. 3.0 using a Metropolis-coupled, Markov Chain Monte Carlo (MCMC) sampling approach. The DNA Barcoding procedure produces a whole catalogue of I.D. of individual species or populations of fish. A key service for the utilization of analytical data is the organization of a BioBank collecting biological samples, DNA samples, genomic sequences connected also to a Knowledge base that represents sequence collected and is capable to interact with remote databases. The knowledge base will be equipped with a data mining module providing advanced tools to automatically extract new knowledge by highlighting predictive patterns of interest [2].

Results

Our results allow the molecular identification of the fish (through the COI) also in its processed product, and in some case the identification of the geographical origin of the specific fish (through the 5'Dloop). Genetic analysis of the species of interest, are conducted through the innovative technique of DNA Barcoding alone or in combination with the high-resolution melting analysis (HRM) analysis, called Bar-HRM (High Resolution Melting-barcode DNA). This will allow to increase the knowledge of species not yet well studied, but of considerable importance for the Mediterranean area. The genetic analysis of the species of interest are exploited not only for the recognition of species-specific sequences, i.e. the ability to distinguish between two organisms belonging to different species, but also for determining the provenance of the fish (characterization of fish stock).

References

1. A.M. Pappalardo, F. Guarino, S. Reina, A. Messina and V. De Pinto Geographically widespread swordfish barcode stock identification: a case study of its application (2011) PLOS ONE, doi: [10.1371/journal.pone.0025516](https://doi.org/10.1371/journal.pone.0025516)
2. A. Laganà, S. Forte, A. Giudice, M. R. Arena, P. L. Puglisi, R. Giugno, A. Pulvirenti, D. Shasha, Ferro A. (2009). miRò: a miRNA knowledge base. DATABASE, vol. Vol. 2009, bap008; p. *, ISSN: 1758-0463, doi: [10.1093/database/bap008](https://doi.org/10.1093/database/bap008).