

Compiling a DNA barcode reference library for benthic molluscs of the Southern European Atlantic coast.

Maria Fais^{1*}, Claudia Hollatz¹, Jorge Lobo^{1, 2}, Sofia Duarte¹, Luisa M. Borges^{1, 3, 4} and Filipe O. Costa¹

¹ University of Minho, Centre of Molecular and Environmental Biology (CBMA), Portugal

² Lisbon New University, Marine and Environmental Sciences Centre (MARE), Portugal

³ Helmholtz-Zentrum Geesthacht, Germany

⁴ L3 Scientific Solutions, Germany

The investigation and monitoring of biological diversity as greatly benefitted from the implementation of DNA barcodes as universal standardized molecular tools for species identification (Costa and Antunes 2012). DNA barcodes provide a means for rapid and accurate identification of species from small pieces of tissue, regardless of the life stage or gender, without the need for taxonomic expertise. They also provide a unique tool for discrimination of cryptic species, which have been increasingly reported in diverse animal groups over the last decades. Recently, the emergence of high-throughput sequencing (HTS) technology enabled the application of the DNA barcode concept to the investigation of the composition of species in complex communities (DNA metabarcoding), like marine zooplanktonic and zoobenthic assemblages, or even in environmental samples of marine water or sediments (Environmental barcoding) (Costa and Antunes 2012).

Nevertheless, the accomplishment of such great potential for biomonitoring of marine communities is dependent on the availability of comprehensive reference libraries of DNA barcodes, linking morphologically identified species with their respective set of diagnostic DNA barcodes. Despite much effort has been devoted to the completion of reference libraries for many animals groups, available libraries for marine invertebrates are still incipient or poorly represented in certain regions, which is the case of benthic invertebrates from the NE Atlantic (e.g. Borges et al. 2016). Since molluscs are among the most diverse taxonomic groups of marine benthic invertebrates, they are a crucial component to be considered in any reference library for zoobenthic organisms, which aims to support metabarcoding applications to benthic monitoring.

Although Mollusca is a well-studied phylum, a great effort is still needed to collect sequences for getting a minimally comprehensive DNA barcode database. Additionally, reference databases must be carefully scrutinized for the taxonomic congruence of morphology and molecular-based identities, as well as for consistency of sequence data generated independently by distinct research groups (Costa et al. 2012). Therefore, the process of compilation of reference libraries is a valuable opportunity to revise taxonomic ambiguities, and detect potential hidden diversity. In this study we report on an initial reference library of DNA barcodes for benthic molluscs of the Southern European Atlantic coast, by generating and compiling DNA barcodes from representative species of the classes Bivalvia, Gastropoda and Polyplacophora.

Marine gastropods, bivalves and polyplacophora specimens were collected between 2009 and 2015, at sites along the coasts of mainland Portugal and two Azorean Islands, and morphologically identified according to Borges et al. (2016). DNA was extracted from each specimen using one of the following kits: GenElute Mammalian Genomic DNA Mini Prep kit (Sigma) or E.Z.N.A. Mollusc DNA kit (Omega Bio-tek). A 658 bp fragment of the mitochondrial cytochrome c oxidase subunit I (COI-5P) gene was amplified using one the following primer pairs: LCO1490 and HCO2198, dgLCO1490 and dgHCO2198 and LoboF1 and Lobo R1, in 25 μ L PCR reactions. PCR reactions and cycling conditions can be found in Lobo et al. (2013) and Borges et al. (2016). Successful PCR products were purified using the EXOSAP method and sequenced bidirectionally, using external sequencing service suppliers.

The complementary strands of COI sequences were edited and aligned manually using MEGA 7. A multiple alignment was performed using Clustal W, checked for possible occurrence of indels, and subsequently translated into amino acid sequences to inspect for the presence of stop codons or unusual sequence patterns. After basic editing, sequence similarity searches for species identification were carried out using the Barcode of Life Data Systems (BOLD) identification engine and BLASTN search at GenBank, from the National Center for Biotechnology Information (NCBI).

In order to compare and validate our results, we selected sequences from BOLD and GenBank from the same morphospecies or from the same genera. Within-species distances were calculated using the Kimura 2-Parameter model (K2P), for each Mollusca morphospecies. The neighbour-joining (NJ) tree method was used for building a phenogram using the K2P model, and the bootstrap support for the nodes was determined using 1000 iterations.

Sequences for COI-5P were obtained for 177 specimens, assigned to 58 morphospecies (34 Gastropoda, 18 Bivalvia and 6 Polyplacophora). These morphospecies were distributed by: 10 orders, 15 families and 22 genera among the Gastropoda; 4 orders, 11 families and 14 genera among the Bivalvia and 2 orders, 5 families and 5 genera among the Polyplacophora. To this dataset, we added publicly available COI-5P sequences (GenBank and BOLD) of the same or taxonomically close species, to inspect for DNA barcode discriminatory ability and data congruence, in a total of 325 sequences.

Overall, within-species K2P mean distance varied between 0.8 to 2.9% for Gastropoda and Bivalvia, respectively. In general, there was

a good match between our sequences and sequences of independent studies on public databases for 18 species of Gastropoda, among which 8 species from geographically distant populations (e.g. *Gibbula cineraria* and *Nucella lapillus*). For the species *Nassarius incrassatus* and *Siphonaria pectinata* we observed comparatively high intraspecific divergence (2.5 and 3.1%, respectively), which corresponded to geographically sorted populations. On the other hand, higher intraspecific divergences were found among the Polyplacophora and the Bivalvia attaining a maximum of 5.0% for *Leptochiton algesirensis* and 19.7% for *Ruditapes decussatus*, respectively. However, the limited number of DNA barcode sequences publicly available for these two species, prevents additional conclusions on the relevance of these findings, in particular for *R. decussatus*.

The barcode reference library reported here expects to contribute to a more accurate taxonomic resolution within the Mollusca. This initial study already emphasised some cases of species ambiguity where a higher intraspecific divergence was found (e.g. *R. decussatus*). These species are worthwhile to further investigation, by increasing and optimising the effort in sampling and molecular identification protocols. Nevertheless, these reference libraries are fundamental steps for the development of accurate metabarcoding approaches to gain insights into the diversity of complex communities in benthic ecosystems.

References

- Borges, L.M.S., Hollatz, C., Lobo, J., Cunha, A.M., Vilela, A.P., Calado, G., Coelho, R., Costa, A.C., Ferreira, M.S.G., Costa, M.H. and Costa FO. (2016). With a little help from DNA barcoding: investigating the diversity of Gastropoda from the Portuguese coast. *Sci. Rep.* 6:20226. doi: 10.1038/srep20226
- Costa, F.O. and Antunes, P.M. (2012). "The contribution of the Barcode of Life initiative to the discovery and monitoring of Biodiversity", in *Natural Resources, Sustainability and Humanity - A Comprehensive View*, ed. A. Mendonça et al. (Dordrecht: Springer), 37-68. doi: 10.1007/978-94-007-1321-5_4
- Costa, F.O., Landi, M., Martins, R., Costa, M.H., Costa, M.E., Carneiro, M., Alves, M.J., Steinke, D. and Carvalho, G.R. (2012). A ranking system for reference libraries of DNA barcodes: application to marine fish species from Portugal. *PLoS ONE* 7:e35858. doi: 10.1371/journal.pone.0035858
- Lobo, J., Costa, P.M., Teixeira, M.A.L., Ferreira, M.S.G., Costa, M.H. and Costa, F.O. (2013). Enhanced primers for amplification of DNA barcodes from a broad range of marine metazoans. *BMC Ecol.* 13:34. doi: 10.1186/1472-6785-13-34

Keywords: Mollusca, Portuguese Coast, DNA barcode, Reference library, COI-5P

Conference: XIX Iberian Symposium on Marine Biology Studies, Porto, Portugal, 5 Sep – 9 Sep, 2016. **Presentation Type:** Poster Presentation

Topic: 1. ECOLOGY, BIODIVERSITY AND VULNERABLE ECOSYSTEMS

Citation: Fais M, Hollatz C, Lobo J, Duarte S, Borges LM and Costa FO (2016). Compiling a DNA barcode reference library for benthic molluscs of the Southern European Atlantic coast. *Front. Mar. Sci. Conference Abstract: XIX Iberian Symposium on Marine Biology Studies*. doi: 10.3389/conf.FMARS.2016.05.00132

Received: 30 Apr 2016; **Published Online:** 03 Sep 2016.

* **Correspondence:** Mrs. Maria Fais, University of Minho, Centre of Molecular and Environmental Biology (CBMA), Braga, 4710-350, Portugal, maria.fais@bio.uminho.pt