



# BIOLOGICAL IDENTIFICATIONS THROUGH DNA BARCODES: THE CASE OF THE CRUSTACEA

Filipe O. Costa<sup>1</sup>, Jeremy R. deWaard<sup>2</sup>, Sujevan Ratnasingham<sup>2</sup>, Robert Dooh<sup>2</sup>, Mehرداد Hajibabaei<sup>2</sup>, Paul D.N. Hebert<sup>2</sup>

<sup>1</sup>Instituto do Mar (IMAR), Universidade Nova de Lisboa, 2829-516 Caparica, Portugal  
<sup>2</sup>Biodiversity Institute of Ontario, University of Guelph, ON N1G2W1 Canada

## INTRODUCTION

The ideal DNA-based identification system would employ a single gene for the placement of any species in the full taxonomic hierarchy from kingdom to species. The general ability of a 650 base pair section of the mitochondrial cytochrome c oxidase I (COI) gene to provide this resolution has been demonstrated for large taxonomic assemblages of animals such as birds (1) and lepidopterans (2). However, the effectiveness of this approach has not been examined for the subphylum Crustacea, which is the most ancient and structurally diverse group of arthropods. In this study we examine the ability of COI to provide taxonomic resolution in this group, examining three divergent levels in the taxonomic hierarchy – orders, families and species.

## MATERIALS AND METHODS

One hundred and fifty species, each belonging to a different family, and including representatives from the 23 most prominent of 47 crustacean orders, were sequenced to establish a COI ordinal profile. Family and species level resolution was analyzed in separate datasets which comprised members of the order Decapoda and of the genus *Gammarus* (Amphipoda).

We isolated DNA using the GenElute kit (Sigma) and PCR amplified a 648 base pair fragment of the 5' end of COI. We used ABI 377 and 3730 sequencers and checked sequence quality with a PHRED-based algorithm. Additional data were extracted from GenBank. We used a K2P distance model (3) to calculate divergences. Neighbor-joining (NJ) trees were built using K2P distances for graphical representation of the patterns of COI divergences among species. Levels of COI nucleotide divergences among conspecific versus congeneric individuals were compared in Decapoda, *Gammarus* and *Pandalus* (Decapoda), using the tools provided by the Barcode of Life Database (visit [www.barcodinglife.org](http://www.barcodinglife.org)).

## RESULTS AND DISCUSSION

The amino acid divergence profile for representatives of 150 crustacean families showed high cohesion of taxonomic groups. Most formed monophyletic assemblages according to established higher-taxonomic categories - orders, superorders or subclasses (Figure 1). For example, members of the orders Anostraca, Stomatopoda, Mysida, Isopoda, Cumacea, and Amphipoda formed cohesive groups. The orders Sessilia and Pedunculata grouped as superorder Thoracica. The subclasses Copepoda and Phyllopoda were aggregated in distinctive assemblages. There were a few deviations from this pattern; members of the order Decapoda split into two groups, although still forming a conspicuous cluster together with Stomatopoda.

A similar pattern of association at the family level was obvious in the NJ plot of nucleotide divergences for members of the order Decapoda (Figure 2). All 15 families represented by more than one species formed a cohesive cluster, except the Portunidae. This NJ tree for Decapoda illustrates the ability of COI sequences to provide species-level resolution. In all 3 examples provided in Table 1, the mean intraspecific divergences are much lower than congeneric ones (Figure 3).

Table 1 – Pairwise nucleotide divergences of 5'-COI based on K2P distance.

Pairwise Divergence	n Comparisons	Min Distance	Mean Distance	Max Distance	Standard Error
<b>Order Decapoda (150 representatives)</b>					
Within a Species (n=56)	272	0	0.68	5.4	0.048
Within a Genus (n=17)	915	1.42	17.95	30.46	0.17
Within a Family	931	11.02	22.29	34.28	0.141
<b>Genus Gammarus (Amphipoda) (11 representatives)</b>					
Within a Species (n=11)	151	0	0.97	3.09	0.066
Within a Genus	1124	5.58	25.33	31.39	0.108
<b>Genus Pandalus (Decapoda) (6 representatives)</b>					
Within a Species (n=6)	24	0	0.48	1.86	0.112
Within a Genus	186	9.85	16.49	20.82	0.176

Fig. 1 – NJ tree showing 5'-COI amino acid profile for 150 crustacean species, each belonging to a different family.

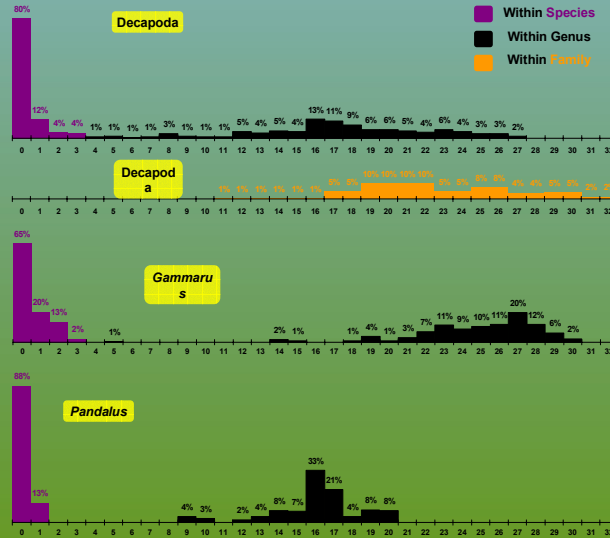
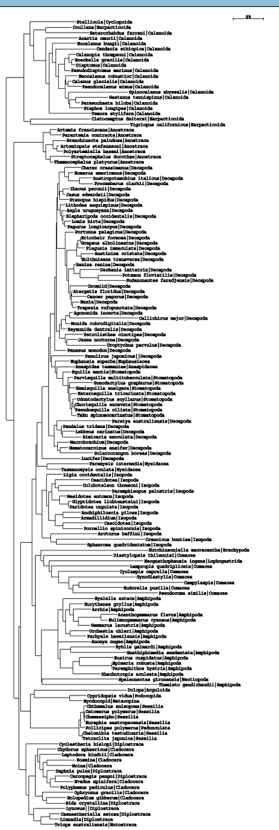


Fig. 3 – Frequency distribution of pairwise nucleotide divergences from the three sample datasets reported in table 1.

## CONCLUSIONS

Our results show that sequence diversity in COI can assign species to major partitions in the taxonomic hierarchy of the Crustacea. Moreover, they indicate that COI analysis provides an effective approach for species identifications of Crustacea.

The differences between conspecific and congeneric COI divergences are much higher in crustaceans than those observed in other animal groups, such as birds and insects, suggesting that species are old or that rates of evolution are high. Hence, application of COI barcodes to species diagnosis in crustaceans will match or exceed those in other groups.

## REFERENCES

- Hebert, P.D.N., Stoeckle, M.Y., Zemlak, T.S., Francis, C.M. (2004). *PLoS Biology* 2, 1657-1663.
- Hebert, P.D.N., A. Cywinska, S.L. Ball and J.R. deWaard (2003). *Proc. R. Soc. Lond. B* 270, 313-322.
- Kimura, M. (1980). *J. Mol. Evol.* 16, 111-120.

## ACKNOWLEDGMENTS

We thank James Boutiller for the collection and identification of decapods and Sonya Uzunova and Monika Nomant for providing the *Gammarus* samples. Fundação Calouste Gulbenkian (Portugal) sponsored the participation of F. Costa in this conference. The Fundação para a Ciência e a Tecnologia (Portugal) provided a postdoctoral fellowship (BPD/11583/2002) to F. Costa. Research support was provided by NSERC and by the Gordon and Betty Moore Foundation.

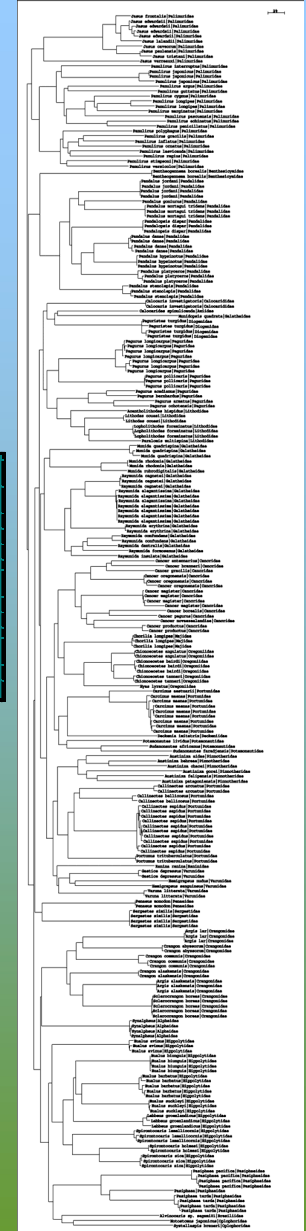


Fig. 2 – NJ tree based on 5'-COI nucleotide distances (K2P) for 113 decapod species representing 25 families.