

In silico DNA barcoding by COI gene of *Crassostrea virginica* in the Atlantic Ocean and Caribbean Sea

Francia Amairani Avalos-Saldaña^{1,2}, Fátima Yedith Camacho-Sánchez^{1,2}, Erick de Jesús de-Luna-Santillana¹, Karina Janett Juárez-Rendón¹, and Miguel Ángel Reyes-López^{1,*}

¹ Instituto Politécnico Nacional, Centro de Biotecnología Genómica, Conservation Medicine Lab.

² Universidad Autónoma de Tamaulipas, UAM Reynosa-Aztlán, Ingeniería Industrial.

*Correspondence: mreyesl@ipn.mx; Tel.: +52-55-57296000 ext. 88751

- (1) **Background:** *Crassostrea virginica* belongs to a class of bivalve mollusks that play vital roles in the marine ecosystem. Its economic and biological importance makes it a key species in the Atlantic Ocean, the Caribbean Sea, and the Gulf of Mexico. Then, its analysis using DNA barcodes would be useful to understand the distribution and genetic diversity and to generate conservation and production programs for *C. virginica* in regions where these practices are not carried out.
- (2) **Methods:** All possible *C. virginica* COI gene sequences in BOLD were downloaded, aligned, and trimmed using BioEdit v.7.7.1.0 and ClustalW v.1.81. For the haplotypes and their networks, Network 10.2.0.0 and PopArt were used by MJN. DnaSP v6 was used for sequence grouping and population genetic analysis. The *p*-Distance and Tajima-Nei values were obtained by Mega v11 software. Finally, to obtain the phylogenetic tree, IQTree v. 2.2.2.6 by ML and FigTree v1.4.4.
- (3) **Results:** The 3 most abundant haplotypes were AY5422 (58,937%), GB8915 (9,661%), and GB1513 (5,314%). Four groups of sequences were formed by the origin of capture, Canada, the USA, Cuba, and Panama. The results showed two haplotype networks by region and frequency and a total of 42 different haplotypes were found. The haplotype network showed that they are grouped by regions where Canada and the US are mixed, as well as Cuba and Panama. Genetic diversity values were close to zero for individuals from the same geographic populations and high values, in contrast, among the 4 described groups ($0.01 < F_{ST} > 0.9$, respectively). The Tajima-Nei analysis showed values from ~0.001 to ~0.029 in intra- and inter-populations, respectively.
- (4) **Conclusions:** This work showed that in the 4 total populations studied, there were high genetic divergence values among them, potentially missing other regions of the Gulf of Mexico that should be studied to generate protection and commercialization programs for *C. virginica*.

Keywords: *Crassostrea virginica*; DNA barcode; Haplotypes