SPRFMO
South Pacific Regional Fisheries Management Organisation

11th MEETING OF THE SCIENTIFIC COMMITTEE

11 to 16 September 2023, Panama City, Panama

SC11 – SQ06
Population genetic analysis of the D gigas along its distribution range based on the mtDNA ND2 gene

Republic of Peru

SPRFMO
South Pacific Regional Fisheries Management Organisation

11th MEETING OF THE SCIENTIFIC COMMITTEE

Panamá, 11-16 September 2023

POPULATION GENETIC ANALYSIS OF THE JUMBO FLYING SQUID *Dosidicus gigas* ALONG ITS DISTRIBUTION RANGE BASED ON THE mtDNA ND2 GENE

by

Giovanna Sotil & Paul Guarnizo
Instituto del Mar del Peru (IMARPE)

This report contains information on the jumbo flying squid stock and fishery in Peruvian jurisdictional waters that, we reiterate, the delegation of Peru, in use of its discretionary powers, voluntarily provides for the purpose of information and support to the scientific research work within the Scientific Committee of the SPRFMO. In doing so, while referring to Article 5 of the Convention on the Conservation and Management of High Seas Fishery Resources in the South Pacific Ocean and reiterating that Peru has not given the express consent contemplated in Article 20 (4) (a) (iii) of the Convention, Peru reaffirms that the decisions and conservation and management measures adopted by the SPRFMO Commission are not applicable within Peruvian jurisdictional waters.

2023

SUMMARY

As it is well known, the jumbo flying squid is a straddling species widely distributed in the Eastern Pacific Ocean, where it sustains coastal and high seas fisheries of great commercial importance. Several markers have being studied for understand its population genetics throughout its distribution range, as well as its population structure and related phenotypic size-groups. In this sense, a population genetic analysis based on the mtDNA ND2 gene analysis is presented with the aim to (a) reevaluate the genetic differentiation among organisms for the three phenotype sizes-groups collected in different latitudinal and longitudinal geographic areas along Peruvian jurisdictional waters; and (b) evaluate the population genetic structure of the species along its wider Eastern Pacific distribution range. After evaluating the presence of stop-codons in the aligned sequences obtained according to our sampling design in Peruvian jurisdictional waters, pairwise Fst analysis showed significant differences, mainly between central oceanic (medium and large phenotype size-groups); while group comparisons done with AMOVA showed a significant difference between coastal and oceanic groups. This was also observed with SAMOVA, where a differentiation among coastal and oceanic groups (Fct and Fst p values < 0.05) was recorded, also related to the large phenotype group. On the other hand, when sequences of organisms collected in our study and in those of others from southern hemisphere (n=594) were compared with those of organisms from the northern hemisphere (n=239), a clear and significant discrimination was observed with SAMOVA analysis (Fct and Fst, $p < 0.05$) between both groups. The median-joining haplotype network showed two main haplogroups formed by the southern organisms, while different small haplogroups for the northern organisms were observed.

1. INTRODUCTION

The jumbo flying squid, *Dosidicus gigas*, is an ecologically and economicaly important straddling species widely distributed in the Eastern Pacific Ocean, where it sustains highly valuable coastal and high seas fisheries. It has a wide distribution range, expanding from around 40° N to 47° S in the Eastern Pacific Ocean, with higher concentrations observed off Peru in the southern hemisphere and off Mexico and California in the northern hemisphere (Nigmatullin et al. 2001), with three phenotypic size-groups reported throughout it range, and all three ocurring in Peruvian waters, although varying greatly through time in their dominance and relative abundance (Csirke et al. 2018, Arguelles et al. 2019).

Different studies of population structure using mitochondrial markers of different species exist in the literature and, in particular for *D. gigas,* mtDNA genes have also been used for population analysis of the species at different geographical zones. Although, the cytochrome oxidase subunit I (COI) gene is a mtDNA gene used for phylogenetic and phylogeographic analysis in several species, it shows complications for population analysis in the jumbo flying squid because, as also was reported for other squids species, it is present in duplicate copies in the mitogenomes, in addition to its low mutational rate.

Another mtDNA gene studied in *D. gigas* is the NAD dehydrogenase 2 gene (ND2), which occurs as a single copy in its mitogenome. Sanchez et al. (2020) used partial sequences of the ND2 gene obtained from organisms collected from northern and southern hemisphere, combined with previously reported sequences by Staaf et al.

(2010). Although they identified three historical lineages of *D. gigas* and show signatures of demographic expansion during the late Pleistocene, no genetic structure was concluded due to limited sampling strategies.

In this sense, ND2 gene sequences were analyzed considering the previous revision of stop-codons, to (a) reevaluate the genetic differentiation among organisms for the three phenotype size-groups collected in different latitudinal and longitudinal geographic areas along Peruvian jurisdictional waters; and (b) evaluate the population genetic structure along the wider Eastern Pacific distribution range of the species.

2. MATERIALS AND METHODS

2.1. mtDNA ND2 gene sequences

A total of 128 mtDNA ND2 gene sequences were selected from mature (stages III and IV) organisms, which were collected according to the protocol described in SC7-SQ10, on 2019. Samples included the three phenotypes size-groups (small, medium and large,) along species distribution range in the Peruvian jurisdictional waters (Table 1).

Table 1. Samples collected in the Peruvian jurisdictional waters, according to the protocol described in SC7-SQ10

Groups		n	Range
Total	-	128	
			(Mantle size)
Phenotype-size	Small	49	193 – 341 cm
	Medium	44	500 – 665 cm
	Large	35	770 – 900 cm
Latitudinal distribution	North	27	
	Central	48	
	South	53	
			(Distance from coast)
Longitudinal distribution	Coastal	74	23 – 87 mn
	Oceanic	54	146 – 200 mn

A second set of sequences was considered, where those obtained in our study were analyzed including others from the wider Eastern Pacific (northern and southern hemispheres) which were previously reported by Staaf et al. (2010), and available in GenBank. A total of 934 sequences (1041 bp) were aligned and considered for the analysis.

2.2. ND2 analysis

Sequences obtained in our study were aligned using Clustal X, and the presence of stop-codons was verified prior to performing any analysis. Then AMOVA analysis was done considering three groups' criteria (latitudinal distribution, longitudinal distribution, phenotype-sizes). Pairwise Fst analysis was calculated between 15 groups formed

according to different combinations of size-latitudinal-longitudinal distribution, and the genetic differentiation among groups was tested using Arlequin 3.5.2.2 (Excoffier 2010) with 1000 iterations. Spatial analysis of molecular variation (SAMOVA) was evaluated using SAMOVA 2.0 (Dupanloup et al., 2002), in order to detect a geographic genetic structure. The number of groups evaluated ranged from k=2 to 5, and the maximum F_{CT} (genetic diversity between groups) and F_{sT} (genetic diversity within populations) values were calculated to determine the best way to group the population.

Then, sequences obtained from northern and southern hemispheres, were compared including those obtained in our study, but no discrimination among phenotype sizes could be performed due to absence of that information. A total of 28 groups were obtained, and genetic diversity indices, including the number of haplotypes (H), number of polymorphic sites (S), haplotype diversity (Hp), and nucleotide diversity (e) were calculated for each population using Arlequin 3.5.2.2. The relationship between haplotypes was evaluated with a haplotype network built using the Median-Joining algorithm (Bandelt 1999), in PopArt v1.7 (Leigh & Bryant 2015). Also, SAMOVA analysis was evaluated considering k=2, while F_{CT} and F_{CT} were registered.

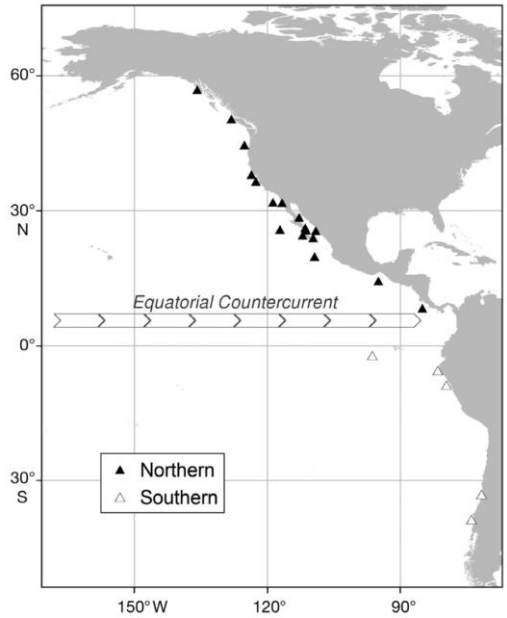

Figure 1. Map obtained from Staaf et al. (2010) pointing locations (in black and white triangles) of origin of ND2 gene sequences used in this study. Sequences were available in GenBank database.

3. RESULTS

3.1. ND2 gene analysis from organisms collected in the Peruvian jurisdictional waters

The ND2 gene sequences obtained in our study (and reported during the SC10, on 2022), were evaluated looking for stop-codons. Only one stop-codon was observed at the end, and in the same position, of all ND2 sequences. Even though, we deleted those extremes and finally 1039 bp were used for further analysis.

The pairwise Fst distance analysis was performed among 6 groups related to its latitudinal and longitudinal geographical distribution. Significant differences between

organisms from the central oceanic and those from the southern, were observed (Table 2). Also, the analysis was performed considering 15 groups organized related to latitude, longitude and phenotype size-groups; it was observed that the group of large organisms from the central oceanic zone (COG) showed the highest pairwise distances among other groups, significant differences (Figure 2).

AMOVA analysis performed under 3 different hypothesis showed a Fct value with significant difference when coastal and oceanic groups were compared (Table 3). Also, this differentiation was observed when SAMOVA was evaluated considering from k = 2 to 5, remarking that it was always observed a discrimination of the central oceanic group among others. Indeed, when SAMOVA was evaluated with k = 2, the central oceanic (CO) organisms were discriminated from the others with significant value for Fst (Figure 3). Whereas, when 4 groups were tested, organisms from the coastal distribution were grouped together, while the oceanic groups were discriminated according to their three (from the north, central and south) collected areas with significance in Fct and Fst values ((Table 4).

Table 2. Pairwise Fst distance values between 6 groups of *D. gigas*, discriminated according their latitudinal and longitudinal distribution. Bold numbers indicate significant differences. CC= central coastal; CO= central oceanic; NC= north coastal; NO= north oceanic; SC= south coastal; SO= south oceanic.

	Central		North		South	
	Coastal (CC)	Oceanic (CO)	Coastal (NC)	Oceanic (NO)	Coastal (SC)	Oceanic (SO)
CC	0					
CO	0.02249	0				
NC	-0.01174	0.04189	0			
NO	0.01143	0.04298	0.01545	0		
SC	-0.00897	**0.05463**	-0.00558	0.01848	0	
SO	0.00536	**0.03696**	0.00874	0.00183	0.01739	0

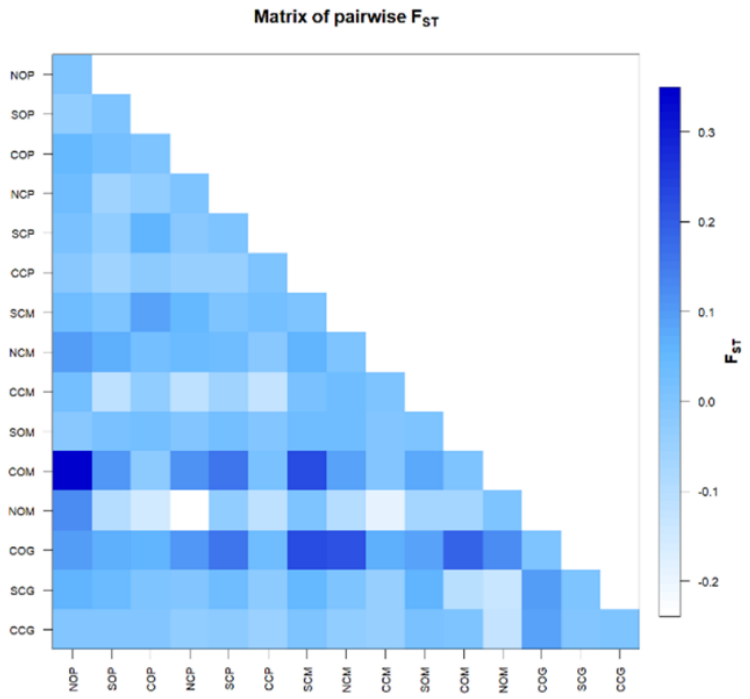

Figure 2. Heatmap of pairwise Fst distance analysis performed considering 15 groups of *D. gigas*, related to their latitudinal and longitudinal distribution, and phenotype sizes. CC= central coastal; CO= central oceanic; NC= north coastal; NO= north oceanic; SC= south coastal; SO= south oceanic; G= large; M=medium; P=small phenotype.

Table 3. AMOVA analysis considering different grouping criteria for the evaluation of *D. gigas*, based on mtDNA ND2 gene. Bold numbers indicates values with p<0.05, observed when coastal and oceanic organisms were compared.

Source of variation	Variance components	Percentage of variation	Fixation indices (p-value)
a. Grouping hypothesis: North vs Central vs South			
Among groups	0.00201	0.19	Fct = 0.0019 (0.41251)
Among population within groups	0.01679	1.58	Fsc = 0.01585 (0.11339)
Within populations	1.04269	98.23	Fst = 0.01772 (0.13099)
b. Grouping hypothesis: Coastal vs Oceanic			
Among groups	0.01549	1.45	**Fct = 0.0145 (0.02542**)
Among population within groups	0.01001	0.94	Fsc = 0.00951 (0.26295)
Within populations	1.04269	97.61	Fst = 0.02387 (0.1349)
c. Grouping hypothesis: Small vs Medium vs Large size			
Among groups	0.00557	0.52	Fct = 0.00524 (0.2913)
Among population within groups	0.01421	1.34	Fsc = 0.01345 (0.12414)
Within populations	1.04269	98.14	Fst = 0.01862 (0.1349)

Table 4. Spatial analysis of molecular variance (SAMOVA) for *D. gigas* for ND2 haplotypes, considering increasing k (number of groups) values. CC= central coastal; CO= central oceanic; NC= north coastal; NO= north oceanic; SC= south coastal; SO= south oceanic.

# groups	Group definition	Variance	% total		Fixation index	P
2	**[CO]** and **[CC, NC, NO, SC, SO]**					
	Among groups	0.043	3.64	Fct	0.036	0.167
	Among populations within groups	0.004	0.35	Fsc	0.003	0.277
	Within populations	1.122	96	**Fst**	**0.04**	**0.045**
3	**[CO], [SO]** and **[CC, NC, NO, SC]**					
	Among groups	0.036	3.14	Fct	0.031	0.061
	Among populations within groups	-0.004	-0.38	Fsc	-0.004	0.492
	Within populations	1.122	97.24	Fst	0.027	0.038
4	**[CO], [NO], [SO]** and **[CC, NC, SC]**					
	Among groups	0.039	3.38	**Fct**	**0.034**	**0.05**
	Among populations within groups	-0.012	-0.93	Fsc	-0.01	0.743
	Within populations	1.122	97.55	**Fst**	**0.024**	**0.032**
5	**[CO], [NO], [SO], [SC]** and **[CC, NC]**					
	Among groups	0.036	3.16	Fct	0.032	0.072
	Among populations within groups	-0.016	-1.38	Fsc	-0.014	0.677
	Within populations	1.122	98.22	Fst	0.018	0.05

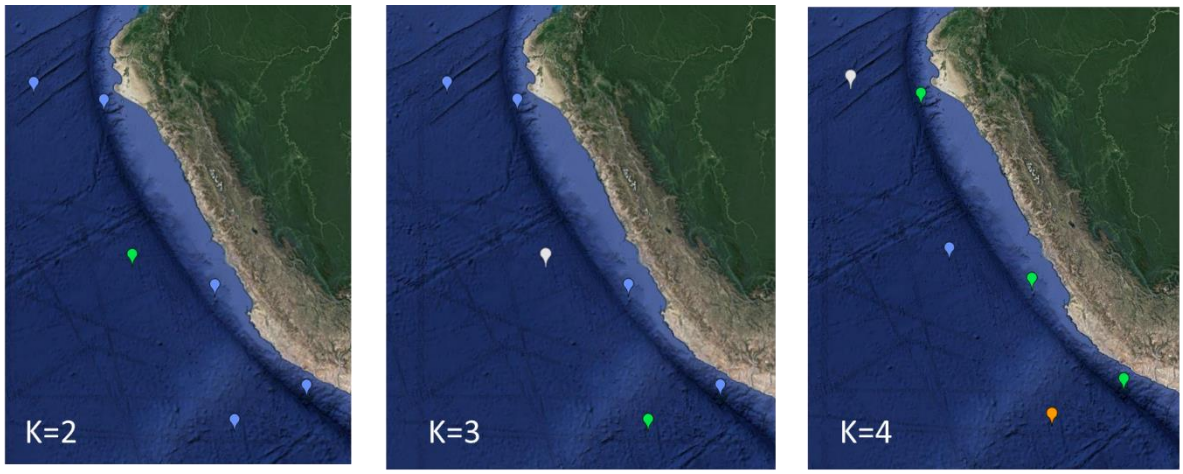

Figure 3. Representation of the spatial analysis of molecular variance (SAMOVA) for *D. gigas, considering K from* 2 to 4, considering groups according to their latitudinal and longitudinal geographical distribution. For k=2, the central oceanic (CO) organisms (green point in map) differentiated from the others (blue points). For k=3, the CO (white point) differentiated from the south oceanic organisms (green point) and the others (blue point); while with k=4 the oceanic organisms (white, blue and orange points) discriminated from the coastal ones (green point).

3.2. Analysis of ND2 gene sequences along Pacific distribution

From sequences obtained from GenBank and reported by Staaf et al. (2010), and those obtained in our study, 28 groups were formed (with georeferencing available also). The pairwise Fst distance showed differentiation between organisms from northern and southern hemisphere (Figure 4). In the same way, when SAMOVA was evaluated with k = 2, the same discrimination was observed (Figure 5). The median-joining network

showed that organisms from southern hemisphere were distributed in two haplogroups, while smallest haplogroups were observed for the northerns (Figure 6).

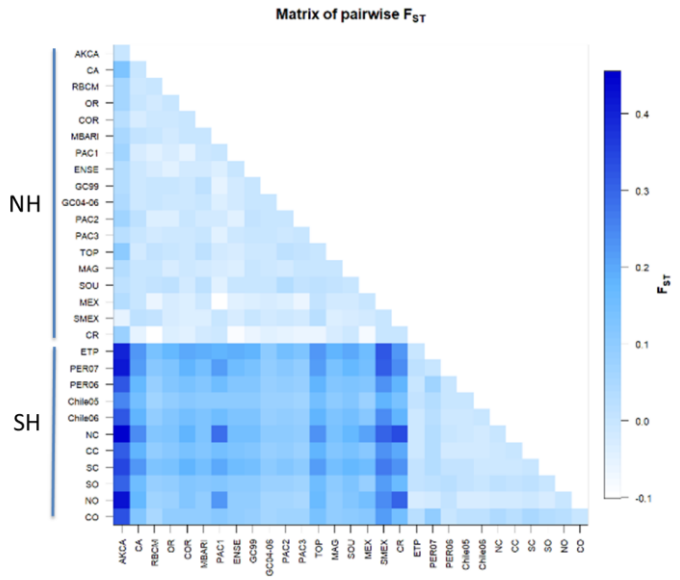

Figure 4. Heatmap of pairwise Fst distance analysis performed considering groups of *D. gigas* collected in the northern (NH) and southern (SH) hemisphere from Eastern Pacific. ND2 sequences used were obtained from our study and Staaf et al, (2010) available in GenBank.

Figure 5. Representation of the spatial analysis of molecular variance (SAMOVA) for *D. gigas,* considering k=2, where the northern (green points) differentiated significantly from the southern (blue points)**.**

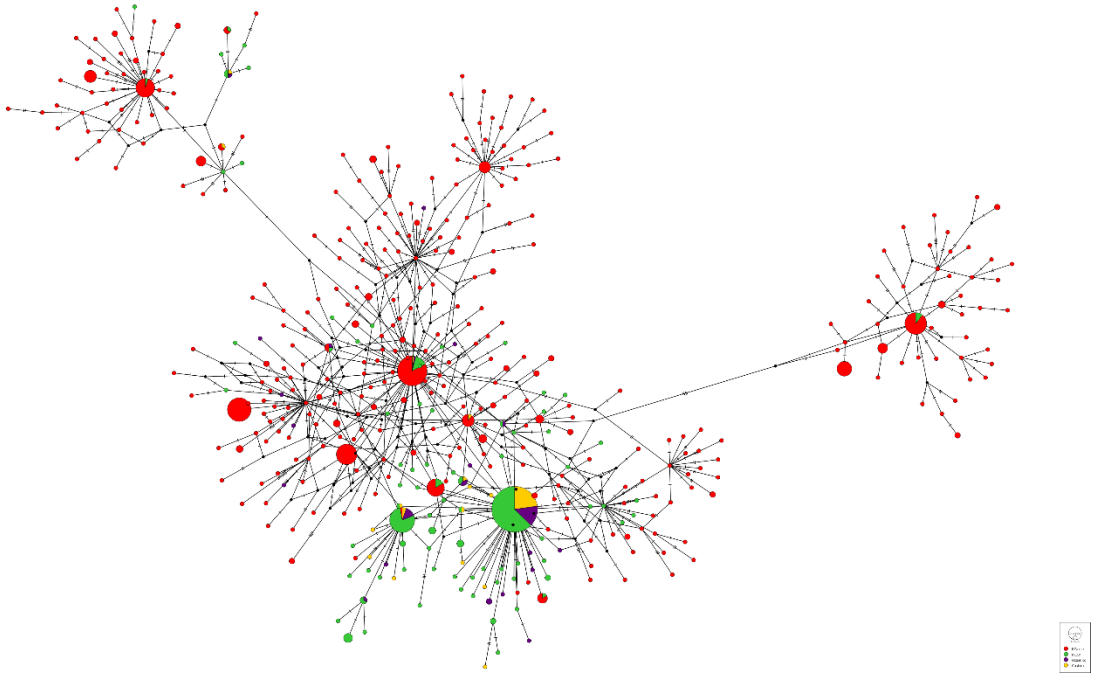

Figure 6. Haplotype network based on mtDNA ND2, considering sequences from the northern (red circles) and southern (green, yellow and purple circles) hemisphere distribution of *D. gigas* in the Eastern Pacific Ocean.

4. CONCLUSIONS

After the several statistical evaluations described above, we can confirm that, based on the ND2 gene analysis of organisms of the Peruvian jurisdictional waters, significant differences are observed between the coastal and oceanic organisms, mainly in the central oceanic ones (medium and large phenotype size-groups).

In addition, we conclude that there is a significant differentiation between the northern and hermisphere organisms.

Given the above, spatio-temporal differences should be considered in further evaluations, in order to differentiate more clearly any populations or stocks, due to different response of organisms to environmental changes which could affect their selection, migration, etc

5. REFERENCES

Arguelles, J., **J. Csirke**, D. Grados, R. Tafur & J. Mendoza. 2019. Changes in the predominance of phenotypic groups of jumbo flying squid *Dosidicus gigas* and other indicators of a possible regime change in Peruvian waters. Working paper presented to the 7th meeting of the Scientific Committee of the SPRFMO, La Havana, Cuba, 5-12 October 2019. SPRFMO Doc. SC7-SQ03: 19p

Bandelt H, Forster P, Röhl A. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol Biol Evol 16(1):37–48.

Csirke J, Argü elles J, Alegre A, Ayó n P, Bouchon M, Castillo G, Castillo R, Cisneros R, Guevara-Carrasco R, Lau L, Mariátegui L, Salazar C, Tafur R, Torrejón J, Yamashiro C. 2018. Biología, estructura poblacional y pesquería de pota o calamar gigante (*Dosidicus gigas*) en el Perú . Bol Inst Mar Perú . 33(2): 302-364

Dupanloup, I., Schneider, S. and Excoffier, L. (2002), A simulated annealing approach to define the genetic structure of populations. Molecular Ecology, 11: 2571-2581. https://doi.org/10.1046/j.1365-294X.2002.01650.x

Excoffier L, Lischer HEL. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. Molecular Ecology Resources. 10: 564-567.

Leigh, JW, Bryant D. 2015. PopART: Full-feature software for haplotype network construction. Methods Ecol Evol 6(9): 1110–1116

Nigmatullin CM, Nesis KN, Arkhipkin A. 2001. A review of the biology of the jumbo squid Dosidicus gigas (Cephalopoda: Ommastrephidae). Fisheries Research. 54: 9-19.

Sanchez, G., Kawai, K., & Umino, T. (2020). Patterns of mitochondrial and microsatellite DNA markers describe historical and contemporary dynamics of the Humboldt squid Dosidicus gigas in the Eastern Pacific Ocean. Reviews in Fish Biology and Fisheries, 30(4), 661-677. https://doi.org/10.1007/s11160-020-09609-9

Staaf DJ, Ruiz-Cooley RI, Elliger C, Lebaric Z, Campos B, Markaida U, Gilly WF. 2010. Ommastrephid squids *Sthenoteuthis oualaniensis* and *Dosidicus gigas* in the eastern Pacific show convergent biogeographic breaks but contrasting population structures. Marine Ecology Progress Series, 418, 165-178.

Sotil G, Cueva D, Espinel M and Mariátegui L. 2019. Protocol for the collection of jumbo flying squid *Dosidicus gigas* muscle tissue for molecular analysis used by the Instituto del Mar del Peru (IMARPE). Working paper presented to the 7th meeting of the Scientific Committee of the SPRFMO, La Havana, Cuba, 7-12 October 2019. SPRFMO document SC7-SQ10: 8p. http://www.sprfmo.int/assets/2019-SC7/Meeting-Docs/SC7-SQ10-Protocol-for-collecting-jumbo-flying-squid-muscle-tissue-for-molecular-analysis.pdf

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A. 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Datasets. Mol. Biol. Evol. 34: 3299-3302. DOI: 10.1093/molbev/msx248