ACADEMIA JOURNAL OF BIOLOGY 2024, 46(2): 47–61
DOI: 10.15625/2615-9023/19316

COMPARISON BETWEEN SCOT AND CBDP TECHNIQUES IN ASSESSMENT GENETIC DIVERSITY AND VARIATION OF TWO POPULATIONS OF BIGFIN REEF SQUID (*Sepioteuthis lessoniana* d'Orbigny) IN CON DAO AND PHU QUOC ISLANDS, VIETNAM

Le Ngoc Trieu[1,*], Thai Thach Bich[1], Nguyen Van Long[2,3]

[1]Faculty of Biology, Dalat University, Lam Dong, Vietnam
[2]Institute of Oceanography, Vietnam Academy of Science and Technology, 01 Cau Da, Nha Trang City, Vietnam
[3]Graduated University of Science and Technology, Vietnam Academy of Science and Technology, 18 Hoang Quoc Viet, Ha Noi, Vietnam

Received 31 October 2023; accepted 31 May 2024

ABSTRACTS

The Start Codon Targeted Polymorphism (SCoT) and CAAT box-derived polymorphism (CBDP) techniques were used to analyze the genetic diversity and variation of two bigfin reef squid populations in waters surrounding the Con Dao and Phu Quoc islands of Vietnam for technical comparison. The two used techniques reflected different levels of pairwise genetic similarity among individuals depended on the investigated population. Gene differentiation (G_{ST}) between the two investigated populations was 0.0767 and 0.0373 leaded to the genetic distance between them was 0.0381 and 0.0228, and the gene flow was *N*m = 6.0195 and 12.9061 migrants per generation between the populations based on SCoT and CBDP techniques, respectively. Genetic variation within individuals of both populations (*WP*) played the key role in the total genetic variation of whole species in surveyed geographic regions with the distribution of 91.44% based on SCoT data and 93.76% based on CBDP data, the distribution of genetic variation among populations (*AP*) was small. For whole species in the surveyed region, the CBDP markers showed higher genetic diversity, while the SCoT markers reflected the differentiation and genetic distance between the two investigated populations better. Overall, the abilities to detect polymorphisms and the number of revealed loci using SCoT markers were better than using CBDP markers, while the ability to distinguish samples and the primer combination to detect the differences among investigated samples using CBDP markers were better than using SCoT markers, and the overall utility was comparable between these two marker systems. The results from this study prove that the CBDP technique can also be used in studies of animal population genetics.

Keywords: Bigfin reef squid, SCoT, CBDP, genetic diversity, marker systems.

Citation: Le Ngoc Trieu, Thai Thach Bich, Nguyen Van Long, 2024. Comparison between scot and CBDP techniques in assessment genetic diversity and variation of two populations of bigfin reef squid (*Sepioteuthis lessoniana* d'Orbigny) in Con Dao and Phu Quoc islands, Vietnam. *Academia Journal of Biology*, 46(2): 47–61. https://doi.org/10.15625/2615-9023/19316

[*]*Corresponding author email*: trieuln@dlu.edu.vn

INTRODUCTION

The bigfin reef squid, scientifically known as *Sepioteuthis lessoniana* d'Orbigny has been considered to be the most widely distributed loliginid squid in the Indo-West Pacific region (Jereb et al., 2010). In Vietnam, bigfin reef squid was found in the waters of the Gulf of Tonkin, the Southern Central, the Southern East and the Southern West (Duc et al., 1983; Aoki et al., 2008; Cheng et al., 2013) and has been one of the most economically important seafood resources in fisheries.

In its distribution areas, this species usually occurs in water from the surface to a depth of about 100m, in coastal habitats with the bottom layer covered by seagrass beds and coral reefs (Jereb et al., 2010). Such habitats are relatively common in the Southwestern and Southeastern sea waters of Vietnam, especially in coastal habitats surrounding the two large islands, Phu Quoc and Con Dao, respectively. These two islands are separated by about 345 km as the crow flies and the shortest sea distance between them is about 410 km. The planktonic paralarvae of *S. lessoniana* are widely distributed after hatching by coastal and ocean currents (Young & Harman, 1988; Zeidberg & Hamner, 2002) while juveniles and adults live in coastal and consistently move inshore to start mating and spawning (Jereb et al., 2010). These indicated that adult squid individuals from Con Dao and Phu Quoc were almost impossible to mate directly together. Therefore, bigfin reef squid individuals from Con Dao island and from Phu Quoc island must belong to two separate populations.

Bigfin reef squid in particular and squid in general are crucial components in the food chains and webs as well as the biodiversity of marine ecosystems, and are seafood of high economic value for fisheries. Despite playing such important roles, the species *S. lessoniana* has faced two risks, including overexploitation and loss of suitable habitat as a result of climate change (Pecl & Jackson, 2007; Edwards, 2020). These risks can lead to species extinction and loss of valuable marine resources without effective strategies for conservation and sustainable exploitation. Such strategies are often established based on a fundamental conception of the importance of intraspecific genetic diversity, which is determined by the genetic diversity of the constituent populations and by the gene flow among them, and that intraspecific genetic diversity can be measured at markers scattered throughout the genome (Teixeira & Huber, 2021).

Molecular markers are widely applied in many fields such as the identification of taxa and agricultural cultivars, forensics, phylogeny, genetic maps, breeding (Schlötterer, 2004; Henry, 2012; Trieu et al., 2017), and especially in population genetics studies for conservation biology (Minh et al., 2019; Ali et al., 2020; Trieu et al., 2022).

Along with the development of molecular biology, more and more new types of molecular markers along with associated DNA fingerprinting techniques have been established and improved. In general, molecular markers can be divided into dominant and co-dominant markers. Co-dominant markers have the common advantage of being able to directly recognize the homozygous and heterozygous status of alleles of the investigated loci. Besides, they also have disadvantages such as high development cost, low genome coverage, high time consuming laborious for primer development (in Microsatellite or Simple sequence repeat technique), high quality and quantity of DNA and radiolabeled probes required (in Restriction Fragment length polymorphism technique), sequencing technique required (in Express sequence tag and Single nucleotide polymorphism techniques), etc. Dominant markers (with the techniques of Amplified fragment Length Polymorphism, Inter simple sequence repeats, Randomly amplified polymorphic DNA, etc.) have the common disadvantage of not being able to directly recognize the status of alleles of the investigated loci. However, they also have advantages such as being highly polymorphic, used for any organisms without sequence information, good genome coverage, low cost,

easy and quick implementation (Dhingani et al., 2015).

Although not directly reflecting the allele status of the locus, dominant markers have been used in population genetic diversity assessment due to their high sensitivity in expressing the differences among the investigated individuals. With these advantages, many dominant markers have been developed recently, such as Start Codon Targeted Polymorphism (SCoT) (Collard & Mackill, 2008), CAAT box–derived polymorphism (CBDP) (Singh et al., 2013). SCoT and CBDP markers target the start codon and CAAT box of functional genes, respectively. Both of these markers are widely used in studies of population genetics in plants. In animals, SCoT markers have also been used in studies of population genetics (Zeinab & Mohammad, 2017; Ali et al., 2020; Ali, 2020), but CBDP markers have not been used. In the gene structure in eukaryotes, the CAAT box is located upstream of the transcription initiation site and signals the binding site for the RNA transcription factor, namely NF-Y subunits. The sequence of 5'-GGTTA-3' is conserved for the CAAT box (Romier et al., 2002). To demonstrate that CBDP markers can be applied in population genetic studies in animals, in the present study, two DNA fingerprinting data sets of *S. lessoniana* populations generated by CBDP and SCoT techniques were compared together. For *S. lessoniana* species, previous studies on population genetic diversity were conducted by two groups of methods, including using mitochondrial DNA sequence data (Aoki et al., 2008; Cheng et al., 2013) and using DNA fingerprinting data generated via the Allozyme marker (Pratoomchat et al., 2001), Microsatellite marker (Tomano et al., 2013; Zheng et al., 2012). To demonstrate that CBDP markers can be applied in population genetic studies in animals, in the present study, two DNA fingerprinting data sets of *S. lessoniana* populations generated by CBDP and SCoT techniques are compared together. The aims of this study are to determine the properties and efficiency of the SCoT and CBDP techniques in estimating the genetic diversity and variation of two *S. lessoniana* populations in Con Dao and Phu Quoc.

MATERIALS AND METHODS

Materials

Samples were collected in October 2020 by night fishing with lure-hooks. Of fifty-nine collected samples in whole species in the surveyed region, twenty-eight representative samples of the Con Dao population were caught in the waters around Con Dao island and assigned as Sles_036, Sles_037, Sles_039 - Sles_049, Sles_052 - Sles_055, Sles_057, Sles_058, Sles_060 - Sles_065, Sles_067, Sles_068 and Sles_070. Thirty-one representative samples of the Phu Quoc population were caught in the waters around Phu Quoc island and assigned as Sles_072 - Sles_091 and Sles_093 - Sles_103. The fishing localities of these samples were not more than 10 km offshore. Samples were not numbered consecutively as above due to the exclusion of different species individuals during fishing.

Collected specimens were tentacular tissues of captured individuals, individually labeled and preserved in 95% ethanol at -20 ^{o}C until DNA extraction.

DNA extraction

Total genomic DNA was extracted from ethanol-preserved tentacular tissue using the modified CTAB method as described by Adamkewicz and Harasewych (1996). Three DNA samples were extracted from each individual sample. The DNA concentration and purity were determined using the spectrophotometry method (Weising et al., 2005) using a NanoScan2 system (Analytik Jena). The DNA samples with OD_{260}/OD_{280} values between 1.8 and 2.0 were kept in the freezer at -20 ^{o}C until subsequently being used.

DNA fingerprinting

In the current study, original DNA fingerprinting data sets were obtained by using SCoT and CBDP techniques. PCRs were performed in 50 µL reactions containing 25 µL My Red HS Taq mix (Bioline), 0.2 µM

primer and 2 µL DNA template (15 ng/µL). The PCR was performed by using an Eppendorf Mastercycler Pro S thermalcycler. The thermal conditions used in the SCoT technique were as follows: initial denaturation at 94 °C for 5 min; 36 cycles of 94 °C for 15 s, 50 °C for 15 s, 72 °C for 45 s; final extension at 72 °C for 10 min [modified from the method of Collard & Mackill (2008)]. The thermal conditions used in the CBDP technique were: initial denaturation at 94 °C for 5 min; 6 cycles of 94 °C for 45 s, 35 °C for 45 s, 72 °C for 90 s; 30 cycles of 94 °C for 45 s, 51 °C for 45 s, 72 °C for 90 s; final extension at 72 °C for 10 min [modified from the method of Singh et al. (2013)].

To choose the primers for DNA fingerprinting, 15 primers were tested for each of the two used techniques. After the screening, 10 primers for each technique which met the criteria to induce PCR products and generate at least one polymorphic band based on five random samples taken from 2 surveyed populations were chosen to use in the study. The sequence and amplification features of these primers are shown in Table 1.

Table 1. Primers used in the study and their amplification features in total 59 samples of whole species in the surveyed geographic region

No.	Primer code	Sequence 5’ to 3’	Number of induced bands	Number of polymorphic bands	Percentage of polymorphic bands
SCoT technique					
1	SCoT 6	CAACAATGGCTACCACGC	18	16	88.89
2	SCoT 9	CAACAATGGCTACCAGCA	15	15	100.00
3	SCoT 12	ACGACATGGCGACCAACG	28	22	78.57
4	SCoT 13	ACGACATGGCGACCATCG	22	20	90.91
5	SCoT 18	ACCATGGCTACCACCGCC	19	17	89.47
6	SCoT 19	ACCATGGCTACCACCGGC	5	4	80.00
7	SCoT 22	AACCATGGCTACCACCAC	16	15	93.75
8	SCoT 29	CCATGGCTACCACCGGCC	2	1	50.00
9	SCoT 30	CCATGGCTACCACCGGCG	12	11	91.67
10	SCoT 34	CATGGCTACCACCGGCCC	15	12	80.00
	Total		152	133	
	Average		15.2		87.50
CBDP technique					
1	CBDP 1	TGAGCACGATCCAATAGC	15	15	100.00
2	CBDP 2	TGAGCACGATCCAATAAT	13	12	92.31
3	CBDP 3	TGAGCACGATCCAATACC	13	11	84.62
4	CBDP 4	TGAGCACGATCCAATAAG	11	10	90.91
5	CBDP 5	TGAGCACGATCCAATCTA	15	14	93.33
6	CBDP 6	TGAGCACGATCCAATCGA	11	9	81.82
7	CBDP 7	TGAGCACGATCCAATCGA	13	13	100.00
8	CBDP 8	TGAGCACGATCCAATCGG	13	11	84.62
9	CBDP 9	TGAGCACGATCCAATGAT	9	8	88.89
10	CBDP 10	TGAGCACGATCCAATGTT	14	11	78.57
	Total		127	114	
	Average		12.7		89.76

To obtain DNA fingerprints for investigated samples, PCR products were separated in 2% agarose gel. Electrophoresis was performed at 60 V for three hours using a TBE buffer. The gel then was stained with Ethidium bromide (0.5 µg/mL), and photographed under 254/312 nm wavelength lights using UVP GelStudio Plus System (Analitik Jena, Germany).

Data analysis

Since both CBDP and SCoT markers are dominant, each observed band was assumed to represent the genotype at a single bi-allelic locus (Williams et al., 1990). The bands were scored as 1 (presence) or 0 (absence) to construct the binary data matrix. The analysis was performed separately for two DNA fingerprinting data sets induced by SCoT and CBDP techniques and the obtained results from each of the theses were compared together.

POPGENE 32 software was used to calculate genetic diversity and variation parameters of the percentage of polymorphic bands (*PPB*), the expected heterozygosity (H_e), Shannon index (*I*), the gene differentiation (G_{ST}), the genetic distance between investigated populations (*D*) and gene flow between them (*N*m) (Yeh et al., 2000).

Pairwise genetic similarity coefficients among investigated samples were calculated and UPGMA dendrograms for their genetic relationship were established by using NTSYSpc 2.1 software (Rohlf, 2004).

Analysis of molecular variance (AMOVA) was carried out using the GenAlEx 6.5 program (Peakall & Smouse, 2012) to describe the distribution of genetic variation among (*AP*) and within (*WP*) investigated populations.

To compare the usability of the two techniques, the data for whole species in the surveyed region (59 samples) was used. The comparative criteria are the number of revealed loci per used primer; level of genetic diversity reflection; level of genetic difference among investigated samples and parameters of gene differentiation and gene flow between investigated populations. Besides, the efficiency of the marker was evaluated via the ability to detect polymorphisms, which was calculated as Polymorphism information content (*PIC*); the number of revealed loci, which was recognized through the efficiency of the primer-marker system as Effective multiplex ratio (*EMR*); the ability to distinguish individuals, which was calculated by Discriminating power (*DP*); the ability of the primer combination to detect the differences between investigated samples, which was expressed by Resolving power (*R*); and the overall utility of marker system, which was measured by calculated Marker index (*MI*) using iMEC - an Online marker efficiency calculator which was available at https://i-rscope.shinyapps.io/iMEC/ (Amiryousefi et al., 2018).

RESULTS

Genetic diversity of two populations and whole species in investigated geographic region

DNA fingerprinting data as matrices of the presence and absence of induced band for investigated samples based on SCoT and CBDP techniques can be consulted in the Mendeley Data website with doi: 10.17632/ykvbz9b44t.1 (Trieu, 2022).

The genetic diversity parameters of the two investigated populations and whole species in surveyed geographic regions are shown in Table 2.

The parameters from Table 2 show that the SCoT data reflected a slightly higher genetic diversity than the CBDP data in the Con Dao population while in the Phu Quoc population, the CBDP technique reflected a much higher genetic diversity than SCoT data.

In the Con Dao population, the values of H_e and *I* parameters were almost comparable between the SCoT and CBDP techniques, but the *PPB* parameters based on the SCoT technique were higher. In the Phu Quoc population, the values of all three diversity parameters based on the CBDP technique were higher than those based on the SCoT technique.

Table 2. Genetic diversity of Con Dao and Phu Quoc populations and whole species in investigated geographic region

Parameters	Con Dao population (28 individuals)		Phu Quoc population (31 individuals)		Whole species (59 individuals)	
	SCoT	CBDP	SCoT	CBDP	SCoT	CBDP
H_e	0.2286	0.2217	0.1480	0.2299	0.2018	0.2348
I	0.3512	0.3396	0.2331	0.3518	0.3218	0.3620
PPB (%)	82.05	77.86	60.26	82.44	85.90	87.02

Note: *PPB*: Percentage of polymorphic bands; H_e: the expected heterozygosity; *I*: Shannon index.

At the species level in the surveyed geographic region with 59 individuals, genetic diversity parameters in Table 2 showed that the CBDP technique reflected a much higher genetic diversity than the SCoT technique and the values of genetic diversity parameters for whole species were at the intermediate level comparing to those of the two constituent populations.

Properties of SCoT and CBDP data sets in reflecting the genetic structure of two populations and whole species in the surveyed geographic region

Con Dao population

The pairwise genetic similarity coefficients among individuals of the Con Dao population based on SCoT data were in the range of 0.4934–0.9737 with an average of 0.7935, based on CBDP data were in the range of 0.5276–0.9764 with an average of 0.8028. Accordingly, in this population, the level of reflection of genetic differences among individuals was almost equal between the two used techniques.

Based on these genetic similarity coefficients, the dendrograms for the genetic relationship among individuals of the Con Dao population were constructed and shown in Figure 1 & Figure 2.

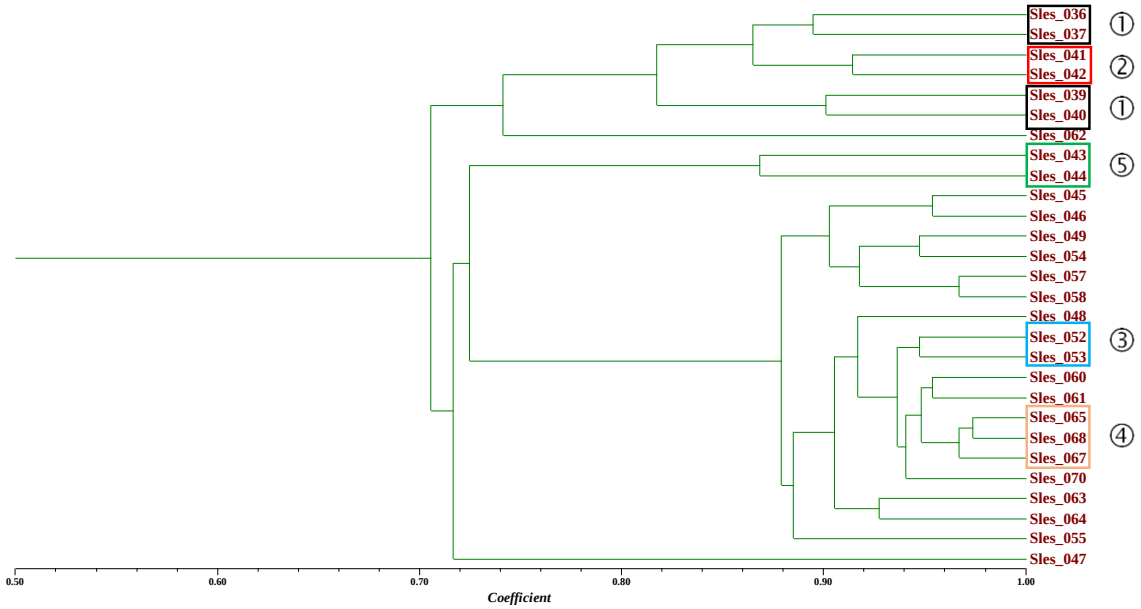

Figure 1. UPGMA Dendrogram for the genetic relationship among individuals of the Con Dao population based on SCoT data

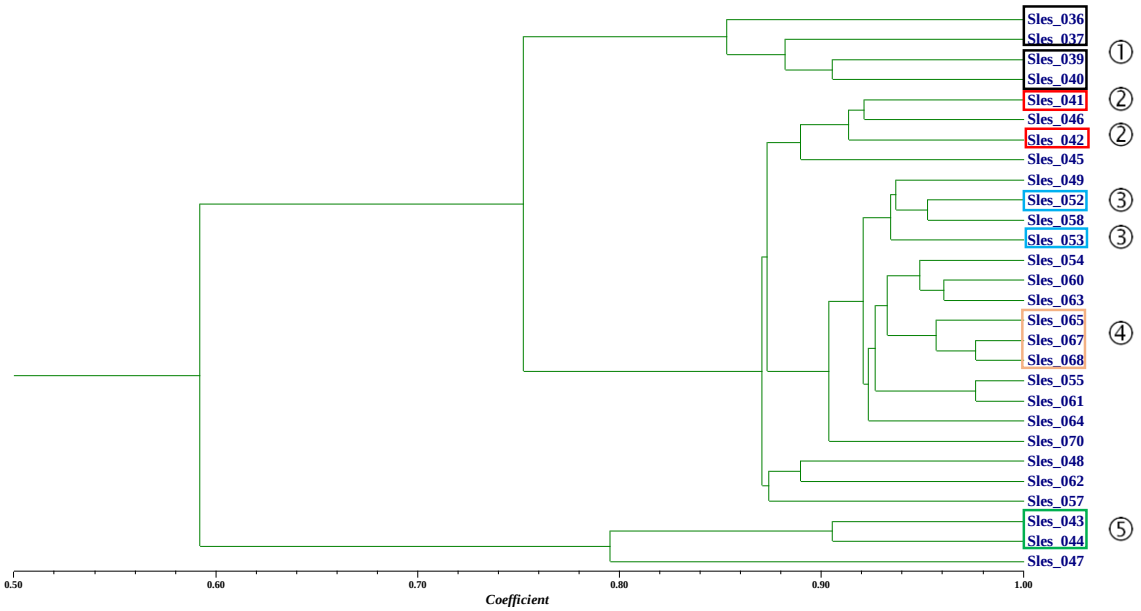

Figure 2. UPGMA Dendrogram for the genetic relationship among individuals of the Con Dao population based on CBDP data

Figures 1 & 2 show that the clustering among samples in the Con Dao population was different depending on the type of marker data, however, in both of the dendrograms based on SCoT data and on CBDP data, the samples belonged to each of the five following groups which were closely clustered together. These five sample groups were: Sles_036, Sles_037, Sles_039, Sles_040 (1); Sles_041, Sles_042 (2); Sles_052, Sles_053 (3); Sles_065, Sles_067, Sles_068 (4); and Sles_043, Sles_044 (5) and marked by different color rectangular frames in the dendrograms. The ability to separate groups (1) and (2) of the CBDP technique is better than that of the SCoT technique. The three samples of Sles_043, Sles_044 and Sles_047 tended to be separate from the rest of the population and formed their own clade.

Phu Quoc population

The pairwise genetic similarity coefficients among individuals of the Phu Quoc population based on SCoT data were in the range of 0.6645–0.9671 with an average of 0.8653, on CBDP data were in the range of 0.4724–0.9685 with an average of 0.7913. Accordingly, in the Phu Quoc population, CBDP data reflected the genetic differences among individuals better than SCoT data.

Based on these genetic similarity coefficients, the dendrograms for genetic relationships among individuals of the Phu Quoc population were constructed and shown in Figures 3 & 4.

Figures 3 & 4 show that the clustering among samples in the Phu Quoc population was different depending on the type of used marker data, however, in both of the dendrograms based on SCoT data and on CBDP data, the samples belonged to each of four following groups which were closely clustered together. These four sample groups were: Sles_080, Sles_081 (1); Sles_101, Sles_102 (2); Sles_094, Sles_095 (3); and Sles_083, Sles_084 (4) and marked by different color rectangular frames in the dendrograms. The two samples of Sles_083 and Sles_084 tended to be separate from the rest of the population and formed their own clade.

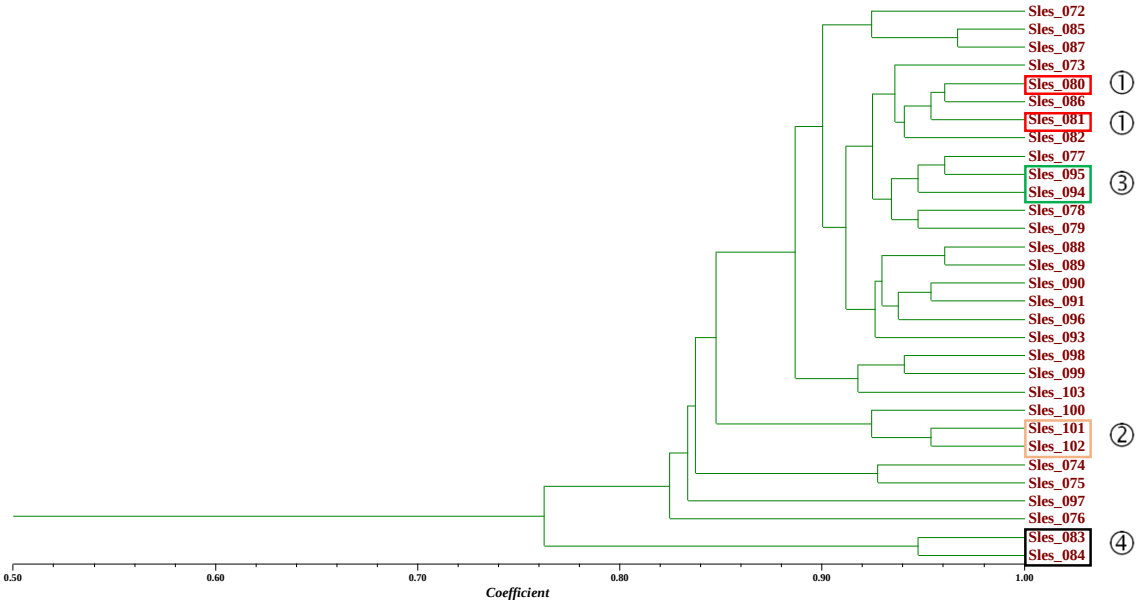

Figure 3. UPGMA Dendrogram for the genetic relationship among individuals of the Phu Quoc population based on SCoT data

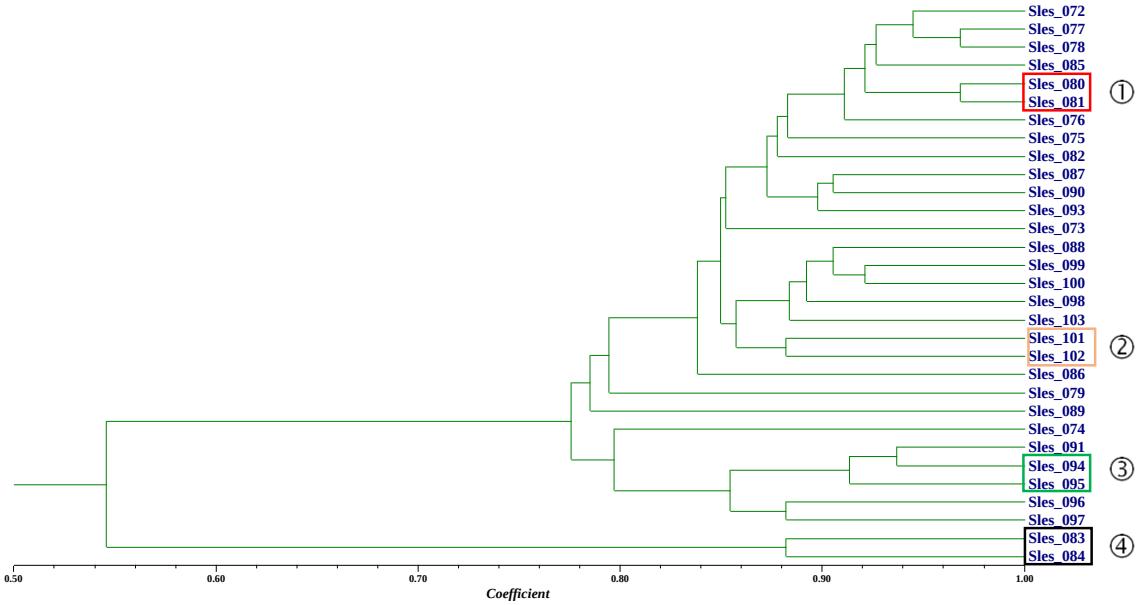

Figure 4. UPGMA Dendrogram for the genetic relationship among individuals of the Phu Quoc population based on CBDP data

Whole species in the investigated geographic region

The pairwise genetic similarity coefficients among 59 individuals from both Con Dao and Phu Quoc populations based on SCoT data were in the range of 0.4934–0.9737 with an average of 0.8223, on CBDP data were in the range of 0.4724–0.9764 with an average of 0.7892. These results showed that in whole species in surveyed geographic regions, CBDP

data reflected the genetic differences among individuals better than SCoT techniques.

Based on these genetic similarity coefficients, the dendrograms for the genetic relationship among individuals of whole species in the surveyed geographic region were constructed and shown in Figures 5 & 6.

Figures 5 & 6 show that the clustering among samples in the whole species in the surveyed geographic region was different depending on the type of marker data, however, in all of the dendrograms based on SCoT data, CBDP data and combined data, the samples belonged to each of seven following groups which were closely clustered together. These seven sample groups were: Sles_036, Sles_037, Sles_039, Sles_040 (1); Sles_041, Sles_042 (2); Sles_080, Sles_081 (3); Sles_094, Sles_095 (4); Sles_098, Sles_099 (5); Sles_101, Sles_102 (6); and Sles_043, Sles_044, Sles_083, Sles_084 (7) and marked by different color rectangular frames in the dendrograms.

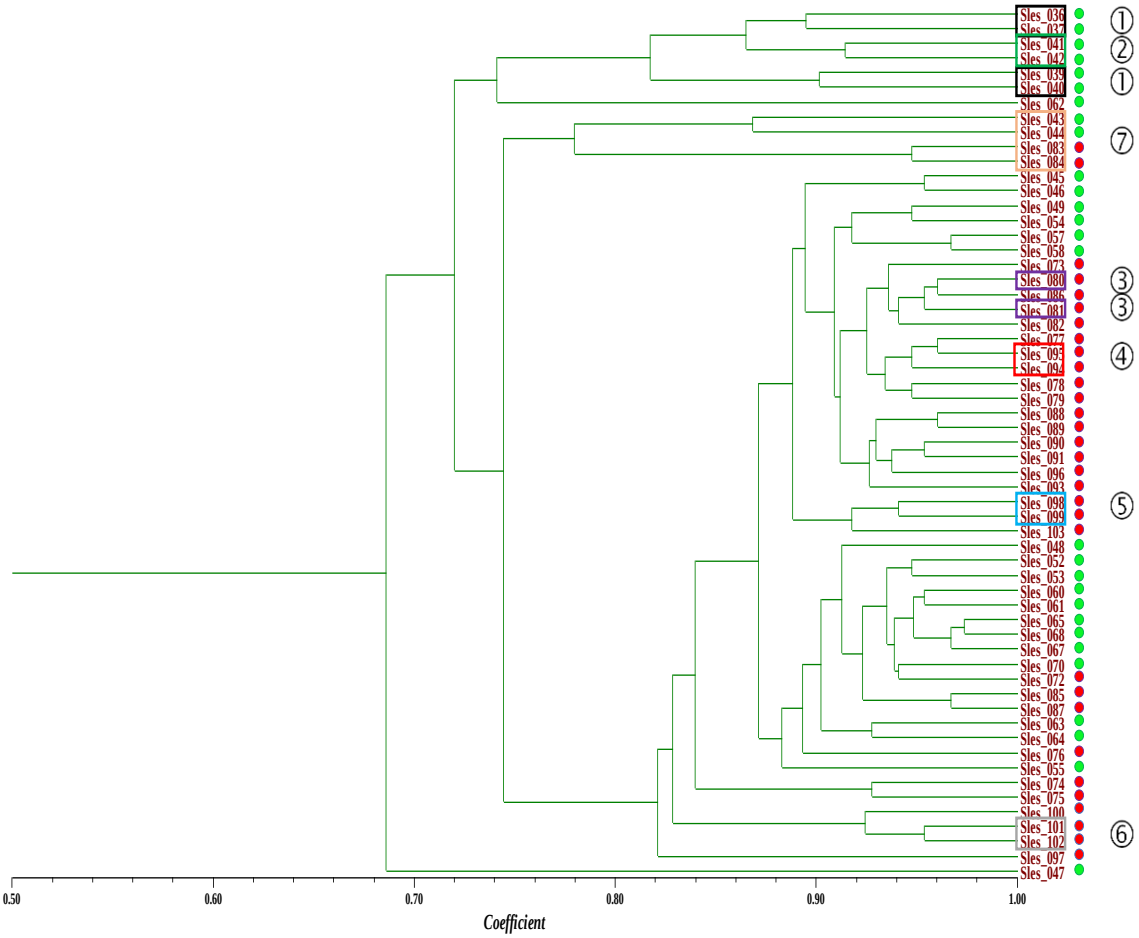

Figure 5. UPGMA Dendrogram for the genetic relationship among individuals of whole species in surveyed geographic regions based on SCoT data.

Note: The individuals from the Con Dao population were marked as the red circles and the individuals from the Phu Quoc population were marked as the green circles

The four samples of Sles_43, Sles_44, Sles_83, Sles_84 and the sample of Sles_47 tended to be separate from the rest of the total investigated samples and formed their own

clade, especially based on the CBDP data set. For the remaining fifty-four samples, both SCoT and CBDP data showed interlaced clustering among samples from the two surveyed populations.

Parameters of the gene differentiation, the genetic distance between investigated populations, gene flow between them and the distribution of genetic variation in whole species in the surveyed region are indicated in Table 3.

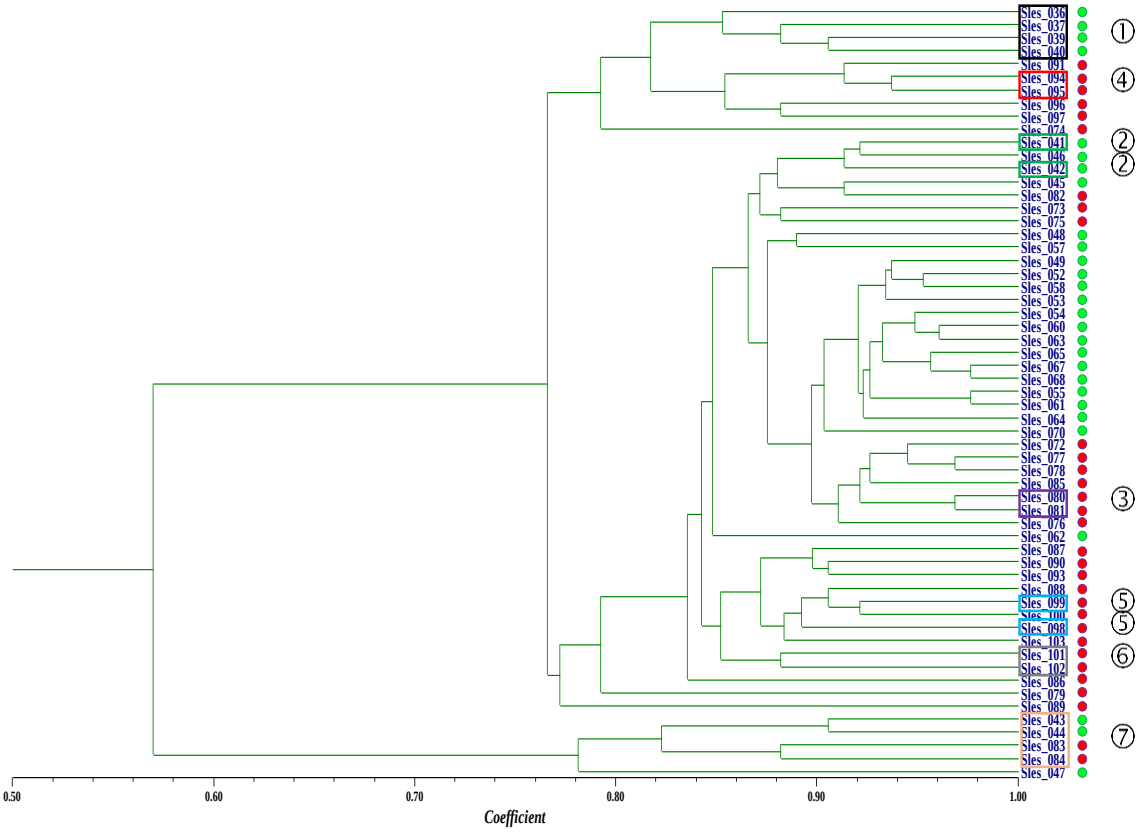

Figure 6. UPGMA Dendrogram for the genetic relationship among individuals of whole species in surveyed geographic regions based on CBDP data.

Note: The individuals from the Con Dao population were marked as the red circles and the individuals from the Phu Quoc population were marked as the green circles

Table 3. The gene differentiation, the genetic distance between investigated populations, gene flow between them and distribution of genetic variation in whole species in the surveyed region

Parameters	Based on SCoT data	Based on CBDP data
Genetic distance between investigated populations (*D*)	0.0381	0.0228
Gene differentiation (G_{ST})	0.0767	0.0373
Gene flow between investigated populations (*N*m)	6.0195	12.9061
Distribution of genetic variation among investigated populations (*AP*), %	8.56	6.24
Distribution of genetic variation within individuals from both investigated populations (*WP*), %	91.44	93.76

The results from Table 3 showed the existence of genetic differentiation and also genetic exchange between the Con Dao population and the Phu Quoc population. Analysis of molecular variance (AMOVA) revealed high genetic variation within populations and low genetic differentiation among two populations in investigated regions.

The efficiency of SCoT and CBDP markers used in the current study

Besides the comparative criteria mentioned in the above sections, the efficiency of the use of SCoT and CBDP markers is also shown via the calculated parameters for the two independent data sets for fifty-nine presentative samples of whole species in the surveyed region, which are presented in Table 4.

The results from Table 4 shows that the CBDP technique reflected higher *PIC* and *E* values than the SCoT technique, but the SCoT technique reflected higher *DP* and *R* values than the CBDP technique, the *MI* values based on the two used techniques were equal.

Table 4. The efficiency of used primers and techniques in genetic diversity estimation of whole species in the surveyed geographic region

Primer code	Polymorphism information content (*PIC*)	Effective multiplex ratio (*E*)	Discriminating power (*DP*)	Resolving power (*R*)	Marker index (*MI*)
CBDP technique					
CBDP 1	0.3777	8.4407	0.6836	3.6610	0.0045
CBDP 2	0.3797	5.4915	0.8219	4.4746	0.0023
CBDP 3	0.3749	6.9322	0.7160	3.1186	0.0050
CBDP 4	0.3820	6.5085	0.6503	2.9492	0.0045
CBDP 5	0.3748	7.9831	0.7170	2.9492	0.0050
CBDP 6	0.3876	6.8136	0.6167	3.5593	0.0048
CBDP 7	0.3786	5.5932	0.8152	5.2203	0.0016
CBDP 8	0.3797	5.4915	0.8219	3.6610	0.0044
CBDP 9	0.3795	3.8136	0.8209	2.6441	0.0017
CBDP 10	0.4057	9.5932	0.5307	2.8814	0.0050
Average of CBDP technique	**0.3820**	**6.6661**	**0.7194**	**3.5119**	**0.0039**
SCoT 6	0.3711	9.7119	0.7091	0.7091	0.0045
SCoT 9	0.4024	4.6780	0.9030	0.9030	0.0023
SCoT 12	0.4062	19.5932	0.5105	0.5105	0.0050
SCoT 13	0.4382	17.3051	0.3814	0.3814	0.0045
SCoT 18	0.3860	11.9831	0.6024	0.6024	0.0050
SCoT 19	0.3763	2.9153	0.6609	0.6609	0.0048
SCoT 22	0.4242	4.0000	0.9377	0.9377	0.0016
SCoT 29	0.3698	1.0339	0.7349	0.7349	0.0044
SCoT 30	0.4190	3.1695	0.9305	0.9305	0.0017
SCoT 34	0.3879	9.5763	0.5927	0.5927	0.0050
Average of SCoT technique	**0.3981**	**8.3966**	**0.6963**	**0.6963**	**0.0039**

DISCUSSION

The ability to reflect the genetic diversity of SCoT and CBDP data sets were different depended on the investigated population. This suggested that the ability of molecular marker systems to reflect genetic diversity depends not only on the specification of the used marker but also on the actual genetic structure of the studied population.

Population genetic diversity in bigfin reef squid has been reported in previous studies. Using the allozyme technique, Pratoomchat et al. (2001) showed that the observed heterozygosity and *PPB* of the populations at Nagasaki (Japan) and at Rayon (Thailand) were 0.28; 45.45% and 0.23; 50%, respectively. The heterozygosity values of the Con Dao population based on the SCoT markers and that of the Phu Quoc population based on the CBDP markers were nearly equivalent to the population at Rayon while the heterozygosity values of the Con Dao population based on the CBDP markers and of the Phu Quoc population based on the SCoT markers were lower than those at Rayon. Both used techniques reflected the lower heterozygosity of both investigated populations but significantly higher percentage of polymorphic bands compared to those of the population at Nagasaki.

Using the SSR technique, Tomano et al. (2013) showed that the heterozygosity of the population in Mugi (Japan) was 0.697 and Zheng et al. (2012) indicated the expected heterozygosity of the population in Hainan island was 0.852. Comparing to the results of these studies, the genetic diversity levels of the Con Dao and Phu Quoc populations based on SCoT and CBDP data sets were relatively low. However, these comparisons may not reflect the true correlation because the results are compared based on different frames of reference.

The results on the pairwise genetic similarity coefficients showed that the data sets based on SCoT and CBDP techniques reflected different levels of discrimination among individuals depended on the investigated population. This showed that although there was a general trend in comparing the genetic diversity of populations, the ability of these two techniques to distinguish individuals within a given population depends on the specific circumstance. The CBDP technique reflected the genetic differences among individuals better than the SCoT technique in the Phu Quoc population but the genetic differences among individuals were the same between the two used techniques in the Con Dao population.

Although general clustering among individuals in each of the populations and the whole species in the surveyed region were different between SCoT and CBDP techniques, the dendrograms still showed the similarity of the two used techniques. That similarity was the appearance of groups of individuals which closely clustered in both techniques. This phenomenon can be explained by two hypotheses, the first is the individuals belonging to each of these groups were genetically related (close relatives) and the second is the individuals belonging to each of these groups were migratory individuals or descendants of individuals migrating from the same remote populations.

One possible explanation for the separate clade with five samples of Sles_043, Sles_044, Sles_083, Sles_084 and Sles_047 is that these samples were taken from migrants or were the descendants of migrants that may had migrated across large geographic barriers from other populations by the spread of planktonic paralarvae. In other words, these samples were the consequence of relatively recent gene flows, this hypothesis can be confirmed by all their COI gene sequences were 99.63–100% homologous to the corresponding sequence recorded for bigfin reef squid in Bali, Indonesia (GenBank accession number of KF052372), (data not shown). Except for the 5 samples mentioned above, even though both the SCoT and CBDP datasets did not show absolute clustering of individuals by population, there were large clusters of individuals from the same population in Figure 5 (SCoT data based) and Figure 6 (CBDP data based).

Technically, the results from Table 3 showed that both SCoT and CBDP data sets reflected the existence of genetic differentiation and also genetic exchange between the Con Dao population and the Phu Quoc population but at different levels. The SCoT data reflected the genetic distance and genetic differentiation between the two investigated populations at a higher level than the CBDP data, and as a corollary, the number of individuals migrating between the two populations in a generation (*N*m) based on the SCoT data was lower than on CBDP data.

The genetic distance between Con Dao and Phu Quoc populations in the current study (*D* = 0.0381 and 0.0228 based on SCoT and CBDP data, respectively) was significantly higher than between Nagasaki and Rayon *S. lessoniana* populations (*D* = 0.003, using allozyme technique) in the study of Pratoomchat et al. (2001), although the geographical distance between the Nagasaki and Rayon populations is much farther than between the Con Dao and Phu Quoc populations.

Based on both SCoT and CBDP data sets, analysis of molecular variance (AMOVA) revealed high genetic variation within populations and low genetic differentiation among two populations in investigated regions. However, the distribution of genetic variation among investigated populations based on SCoT data was higher than based on CBDP data, this is completely consistent with the results that the SCoT data reflected the genetic distance and genetic differentiation between the two investigated populations at a higher level than the CBDP data.

The average values of *PIC* using the SCoT technique were higher than the CBDP technique, these indicated that with the same number of used primers, the ability to detect polymorphisms of SCoT markers were better than CBDP markers, this was also consistent with the experimental results, i.e. the number of polymorphic loci of SCoT technique was higher than CBDP technique (Table 1). These differences between the two used techniques may be due to the sequences of genes annealed by the CBDP primers were more variable than the sequences of genes annealed by the SCoT primers. According to Chesnokov and Artemyeva (2015), the higher the value of *E* means the more efficient the "primer-marker system", so in the current study, with higher average values of *E*, SCoT technique revealed more loci than CBDP technique, and this again was in accordance with the results in Table 1 (152 loci in SCoT technique and 127 loci in CBDP technique). The higher average values of *DP* and *R* using the CBDP technique showed that it possessed the better ability to distinguish individuals and the primer combination to detect the differences between investigated samples using the CBDP technique were more efficient than the SCoT technique. The average *MI* value of SCoT and CBDP techniques was equal, indicating that the overall utility of these two marker systems was comparable.

CONCLUSION

For the bigfin reef squid species in the surveyed region, the ability to reflect the genetic diversity of SCoT and CBDP data sets were different depended on the investigated population. Between the Con Dao population and the Phu Quoc population existed a low to moderate genetic differentiation and a genetic exchange through gene flow based on both used techniques. The SCoT and CBDP techniques have different advantages in reflecting the genetic diversity and variation of the two investigated populations in the surveyed region, but they shared the same level of overall utility in this study. The results obtained from this study indicate that the CBDP technique can also be used in studies of animal population genetics and could be considered for use in similar studies, especially on bigfin reef squid.

REFERENCES

Adamkewicz S. L., Harasewych M. G., 1996. Systematics and biogeography of the geneus Donax (Bivalvia: Donacidae) in eastern North America. *American Malacological Bulletin*, 13(1): 97–103.

Ali H. A. A., 2020. Potential of Start Codon Targeted (SCoT) Markers and SDS-PAGE to Estimate Genetic Diversity and Relationships among Three Gastropods Species from The Mediterranean Sea, Port Said, Egypt. *Egyptian Journal of Aquatic Biology & Fisheries*, 24(7): 133–143.

Ali H. A. A., Hala E. A., Osama A. A., Mohamed K., 2020. Hassan Genetic Relationship Between Two Species of Genus *Dicentrarchus* Based on SCoT Markers and SDS - PAGE. *Egyptian Journal of Aquatic Biology & Fisheries*, 24(5): 393–402.

Amiryousefi A., Hyvönen J., Poczai P., 2018. iMEC: Online Marker Efficiency Calculator. *Applications in Plant Sciences*, 6(6): e01159. https://doi.org/10.1002/aps3.1159

Aoki M., Imai H., Naruse T., Ikeda Y., 2008. Low genetic diversity of oval squid, *Sepioteuthis cf. lessoniana* (Cephalopoda: Loliginidae), in Japanese waters, inferred from a mitochondrial DNA non-coding region. *Pacific Science*, 62: 403–411.

Cheng S. H., Anderson F, E., Bergman A. M., Mahardika G. N., Muchlisin Z. A., Dang B. T., Calumpong H. P., Mohamed K. S., Sasikumar G., Venkatesan V., Barber P. H., 2013. Molecular evidence for co-occurring cryptic lineages within the *Sepioteuthis cf. lessoniana* species complex in the Indian and Indo-West Pacific Oceans. *Hydrobiologia*, 725(1): 165–188.

Chesnokov Y. V., Artemyeva A. M., 2015. Evaluation of the measure of polymorphism information of genetic diversity. *Agricultural Biology*, 50(5): 571–578.

Collard B. C. Y., Mackill D. J., 2008. Start Codon Targeted (SCoT) Polymorphism: A Simple, Novel DNA Marker Technique for Geneerating Genee-Targeted Markers in Plants. *Plant Molecular Biology Reporter*, 27 (1): 86–93.

Dhingani R. M., Umrania V. V., Tomar R. S., Parakhia M. V., Golakiya B. A., 2015. Introduction to QTL mapping in plants. *Annals of Plant Sciences*, 4: 1072–1079.

Duc N. X., Duong T. M., Thuy L. D., 1983. Species composition and exploited production of squid resources in the South Vietnam sea. *Academia Journal of Biology*, 1(1): 21–24. https://doi.org/10.15625/0866-7160/v1n1.3064

Edwards A. J., 2020. Impact of climatic change on coral reefs, mangroves, and tropical seagrass ecosystems. Doeke Eisma (Eds), Climate Change: Impact on Coastal Habitation. Lewis Publishers, Boca Raton, pp. 209–234.

Henry R. J., 2012. Molecular Markers in Plants. Wiley, pp. 216.

Jereb P., Vecchione M., Roper C. F. E., 2010. Family Loliginidae. In P. Jereb & CFE Roper, eds. Cephalopods of the world. An annotated and illustrated catalogue of species known to date. Volume 2. Myopsid and Oegopsid Squids. FAO Species Catalogue for Fishery Purposes. Rome, FAO 2(4): 38–117.

Minh N. T. A., Tien T. V., Hau H. V., Trieu L. N. Chinh V. T., Vinh T. T., Duy N. V., 2019. Genetic diversity and variation of *Huperzia serrata* (Thunb. Ex Murray) Trevis. population in Vietnam revealed by ISSR and SCoT markers. *Biotechnol. Biotechnol. Equip.*, 33(1): 1525–1534.

Peakall R., Smouse P. E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research - an update. *Bioinformatics*, 28: 2537–2539.

Pecl G. T., Jackson G.D., 2007. The potential impacts of climate change on inshore squid: biology, ecology and fisheries. *Reviews in Fish Biology and Fisheries*, 18(4): 373–385.

Pratoomchat B., Natsukari Y., Miki I., Chalermwat K., 2001. Allozyme determination of genetic diversity in Japanese and Thai populations of oval

squid (*Sepioteuthis lessoniana* Lesson, 1830). *La mer*, 39: 133–139.

Rohlf F. J., 2004. NTSYSpc numerical taxonomy and multivariate analysis system version 2.1 - User guide. *Applied Biostatistics Inc*., pp. 42.

Romier C., Cocchiarella F., Mantovani R., Moras D., 2002. The NF-YB/NF-YC Structure Gives Insight into DNA Binding and Transcription Regulation by CCAAT Factor NF-Y. *Journal of Biological Chemistry*, 278(2): 1336–1345. https://doi.org/10.1074/jbc.M209635200

Schlötterer C., 2004. The evolution of molecular markers–Just a matter of fashion?. *Nature Reviews Genetics*, 5: 63–69.

Singh A. K., Rana M. K., Singh S., Kumar S., Kumar R., Singh R., 2013. CAAT box - derived polymorphism (CBDP): a novel promoter - targeted molecular marker for plants. *Journal of Plant Biochem. Biotechnol*., 23(2): 175–183.

Teixeira J. C., Huber C. D., 2021. The inflated significance of neutral genetic diversity in conservation genetics. *Proc. Natl. Acad. Sci. USA*, 118(10). https://doi.org/ 10.1073/pnas.2015096118

Tomano S., Ahmad-Syazni K., Ueta Y., Ohara K., Umino T., 2013. Eleven Novel Polymorphic Microsatellite Loci for Oval Squid *Sepioteuthis Lessoniana* (Shiro - Ika Type). *International Journal of Molecular*, 14: 19971–19975.

Trieu L. N., Bich T. T., Ket N. V., 2022. Population genetic Diversity of *Camellia Dilinhensis* on the Di Linh plateau of Vietnam revealed by ISSR and SCoT markers. *Dalat University Journal of Science*, 12(3): 43–55.

Trieu L. N., Phong N. H., Dat M. T., Bich T. T., Tien N. T., Bao L. D. V., Quang N. K., Nhu P. N. Q., 2017. Genetic diversity investigation and molecular markers establishment for identification of several initially selective avocado (*Persea americana* Miller) strains in Lamdong province. *Dalat University Journal of Science*, 6(4): 467–480. https://doi.org/ 10.37569/DalatUniversity.6.4.89(2016)

Trieu L.N., 2022. Data for research on genetic diversity, variation and structure of two populations of bigfin reef squid (*Sepioteuthis lessoniana* d'Orbigny) in Con Dao and Phu Quoc islands, Vietnam. Mendeley Data, V1. https://doi.org/ 10.17632/ykvbz9b44t.1

Weising K., Nybom H., Wolff K., Kahl G., 2005. DNA fingerprinting in plants: Principles, methods, and applications (2nd ed.). CRCPress, Taylor & Francis Group, pp. 472.

Williams J. G. K., Kubelik A. R., Livak K. J., Rafalski J. A., Tingey S. V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18: 6531–6535.

Yeh F. C., Yang R., Boyle T. J., Ye Z., Xiyan J. M., 2000. PopGene 1.32, Microsoft windows-based freeware for population genetic analysis, version 1.32. Molecular Biology and Biotechnology Centre: University of Alberta, Edmonton, Canada, pp. 28.

Young R. E., Harman R. F., 1988. "Larva", "paralarva" and "subadult" in cephalopod terminology. *Malacologia*, 29: 201–208.

Zeidberg L., Hamner W., 2002. Distribution of squid paralarvae, *Loligo opalescens* (Cephalopoda: Myopsida), in the Southern California Bight in the three years following the 1997–1998 El Niño. *Marine Biology*, 141: 111–122. https://doi.org/ 10.1007/s00227-002-0813-x

Zeinab A. M., Mohammad A., 2017. Using Start Codon Targeted (SCoT) polymorphism for Genetic Diversity Analysis of three Red Sea fishes (Family: Carangidae). *IOSR Journal of Pharmacy and Biological Sciences*, 12(1): 50–56.

Zheng X. D., Zuo Z. R., Su X. J., Li Q., 2012. Isolation and characterization of 24 polymorphic microsatellite loci in *Sepioteuthis lessoniana*. *Genetics and Molecular Research*, 11(4): 3961–3965.