Using internal transcribed spacers 2 (ITS2) to identify seaweed species from Tomini Bay and Banten Bay

Using internal transcribed spacers 2 (ITS2) to identify seaweed species from Tomini Bay and Banten Bay

N Dharmayanti[1*], Abinawanto[2] and A Anti[1]

[1]Study Program of Fisheries Processing Technology, Jakarta Fisheries University, Indonesia
[2]Study Program of Biology Mathematics and Natural Science Faculty, Universitas Indonesia, Jakarta, Indonesia

*E-mail: niken.stp@gmail.com

Abstract. Samples collected from two coastal areas by diving and using a boat. The fresh seaweed morphologically identified while the dried samples collected to be used for DNA testing. Previous research has been shown that there is an influence from environmental parameters on seaweed composition so that the water quality testing was carried out to determine the dominance of wild seaweed water. We conducted genomes extraction, PCR amplification, DNA sequencing testing, alignment with MEGA 7 application and made phylogeny tree, which are distributed at Tomini bay *Dictyota* sp., *Halimeda* sp., *Padina* sp., *Polysiphonia* sp., and Banten bay as *Sargassum* sp. This study evaluated the feasibility of using the ITS2 marker to identify some wild seaweed species, from Tomini Bay and Banten Bay waters. Five wild seaweed species collected from these waters were submitted for DNA barcoding. Currently available universal primers for ITS2 region amplification were sufficient to be successfully amplified and the result of DNA amplification continue to be sequenced. In conclusion, the data presented here indicates that ITS2 are useful markers for DNA barcoding of wild seaweed, with advantage for ITS2 due to the larger availability of analytical tools and reference barcodes deposited at databases for this marker.

Keywords*:* DNA, ITS2 markers, wild seaweed

1. Introduction

The identity of five wild seaweed species has been determined by amplification of the Internal Transcribed Spacer 2 of the nuclear rDNA (ITS2) markers. This allows the detection of specificity patterns for each species and provides an alternative method, when morphological or biological species concept's methods fail. Total genomic DNA suitable for amplification was extracted from the dried tip of a leaf's alga following the CTAB method. We then proceeded to the clasification and DNA testing result of the collected seaweeds.

Seaweed inhabits rocky substrates of the infralittoral zone where they play an important ecological role in providing habitat, food and shelter as well as spawning and nursery grounds for a wide variety of organisms [1, 2]. However, in recent years, seaweed stands have been significantly reduced, [3-7]. Despite its ecological importance, the taxonomy of species within the genus is still poorly understood and no studies have yet explored the genetic diversity of wild seaweed species in the Tomini and Banten Bay. The ecomorphs differ greatly from the average morphology of the species. That makes it difficult to distinguish which morphological variation is an adaptation to ecological conditions, which

is a characteristic of a different species. Species determination is made even more difficult by the lack of comprehensive identification keys for the genus. The only currently available key for wild seaweed species [8] based on morphology is outdated and was written in a period when sampling possibilities, and thus also determination, were limited. Difficulties arise while identifying the species, as several names may have been applied to the various populations or ecomorphs belonging to a single species, i.e. the morphological variability of the species has been overlooked. The situation is similar for other Sargassaceae genera such as, for example, the genus Sargassum [9]. The current method of determining relations within the genus based primarily on morphological characteristics, is shown as being insufficient to unambiguously separate some species of this genus. Because of this, the focus is increasingly shifting to molecular methods of studying marine algae phylogeny, particularly brown algae, and this genus. In order to overcome morphological ambiguities in the identification of speceis, molecular markers are nowadays used [10].

Several different molecular markers have been used in recent studies of brown algae molecular taxonomy and phylogeny. The most common ones are nuclear ribosomal markers, ITS, SSU and LSU [11-13], ITS2 [10], plastid markers, rbcL and psaA [14, 15], and mitochondrial markers, mt 23S and mt spacer [2, 16]. Hovewer, most of these markers have been used for delineation among the higher taxonomic units such as genera and families and rarely for distinguishing between the species. This is especially the case in the order Fucales, the family Sargassaceae, and the genus Cystoseira, which require highly variable markers to determine relations between the taxa [17]. High genetic variations in the populations of the genus Sargassum, and suggest that 2 tis2 those variations the populations should be considered separately when studying the ecology of the genus [5]. Recently, researchers have underlined the usefulness of the mitochondrial 23S in the delineation of *Sargassaceae* genera and the potential of the mt23S-tRNAVal spacer at below genus level [2]. Molecular markers have to be tested to better understand the taxonomy and evolution of the seaweed species research. The purpose of this research is identifying the appropriate ITS2 molecular marker with which species distinction of the wild seaweed could be achieved. We choose to test molecular markers in a preliminary study of four wild seaweed Tomini Bay's species and one wild seaweed Banten Bay's species. The data was then compared to published sequences of Tomini and Banten Bay's spesies.

2. Methods

Sample collection and DNA preparation; The individual of each species were sampled from the rocky bottom between 20 and 80 cm of depth. Four wild seaweeds were collected from the Tomini Bay at Avulua Village 10 km north to Parigi Moutong with coordinate 0^{o}39'53.4"S 120^{o}04'25.7"E in April 2018, and wild seaweed *Sargassum* sp from the Banten Bay at Pulau Lima waters (6°0'5.6376"S – 106°9'15.3684 "E) in February 2018. Morphological analysis was guided with characteristics described previously [8]. Immediately after collection, a part of the branch least overgrown by epiphytes was separated from each specimen. Visible epiphytes were then mechanically removed from the sampled part, the tip of talus was washed with distilled water and stored in silica gel. Total DNA was isolated from silica dried algal tissue using Qiagen Plant mini kit and additionally purified using Qiagen Plasmid mini kit. PCR primers, PCR and sequencing Primers published by [2] for the mt 23S and the mt23S-tRNAVal spacer were used for PCR amplification and sequencing. Primers, PCR conditions and reaction ingredients are listed in table 1. The thermo cycler used in PCR amplification of all regions was Applied Biosystems 2720. The amplified DNA fragments were purified using QIAEX II gel extraction kit 150 (Qiagen). Sequencing was done by Macrogen in Korea using an ABI PRISM 3100 Avant Genetic Analyzer.

Table 1. Primers used to isolate molecular markers with annealing temperature (AT) an bibliographic source [18].

No	Markers	AT	Primers	Primer sequences
1	ITS2 Nuclear	55	5.8S-BF 25 BR-2	5'-GATGAAGAACGCAGCGAAATGCGAT-3' 5'-TCCTCCGCTAGTATATGCTTAA-3'
2	COX3 Mitocondria	42	CAF4A: CAR4A:	ATGTTTACTTGGTGRAGRGA CCCCACCARTAWATNGTNAG

Data analysis raw sequences were edited and trimmed manually using the version of Molecular Evolutionary Genetics Analysis v. 7 (MEGA 7). The number of sequences determined by using the GenBank database which is hosted at the NCBI (National Center for Biotechnology Information, USA) [Link: http://www.ncbi.nlm.nih.gov/]. The NCBI hosts a number of biological databases for example whole-genome databases for human, mouse, chimp, seaweed or algae and another specific genome of organisms.

The number of polymorphic sites and phylogenetically informative sites were determined using the DnaSP v5 [19]. Alignment of sequences was processed using ClustalX [20] and visualized in GeneDoc [21] for each marker independently. A total of 11 sequences for other Mediterranean Cystoseira species available on GenBank and published by [2] were included in the analyses. The most suitable model of nucleotide evolution was determined by the Akaike Information Criterion, AIC, [22] as implemented in JmodelTest [23]: for the 23S Rdna the selected model was TrN+G with gamma=0.013, and for the mt spacer the selected model was HKY, with ti/tv=1.4910. Bayesian analyses were performed using MrBayes 3.1.1 [24]. The parameters of the Markov Chain Monte Carlo (MCMC) analysis comprised two runs (four chains each) for 500000 Species generations, with the sample frequency set to 100 and discarding first 1250 trees as the burn in. Thus, the posterior probabilities of the clades were determined from 3750 trees and the 50% majority-rule consensus tree was built.

3. Results and Discussion

Data analysis of sequences after trimmed by using MEGA X, resulting a vary sequences that equals for their pairs. Here the sequences based on their correct pairs.

1. POLYSIPHONIA #3217322_1N_ITS2-5_8S-BF

GGTCACAAGTGACGCTGCTTGCCGCCGAGGCGGCCCGCGTGGCGGTCGAGCCAGACAAC
TGACTGACTGACTGACCATCGTGGGTCGGACGCGCAAAGCCTAATAGCGGGGCCTGGAC
ATCACCCGCACACCTCGCCCCGGTGGGGACAAGGCGGGGCGCCGAGTCTCGATCGCCCG
TCGCGTACCCCACGCCCCCCCCCCTGCCAAGAGAAATGGCGGGAGAGAGGAATGGGGAC
GGGCAAGAAACATTGGCCCTCCACTAGTCGCAATGCCGCAGTCATACTTCTTAGGCCGGC
TCACGAGACGAAGCTCGCGGCCGGCTCATAGTCCGCCGCCCCCGAGCGCGCGCGAGTCA
CCATTTCTGGCGTCTCAAGCGGGCCACGCACGGGGGGGCGAGTGGGGTGGACAGACACT
CCGACAAGCATGCTCCCAGGAGTATCCCGGAAGCGCAAGATGCGTTCAAAGGTTTGATG
ATTCAC

2. POLYSIPHONIA #3217324_2N_600bp_ITS2-5_8S-BF

GAGATCCCGTAGGAAATAAGGACTGCTGCTTTTTTTTCTAGGAAATAAAAAAACAGCATT
GACAGCCTAAAATACCCACATTTCTGTAGATATACTGTCGATCTGAGCGTTGTAATGTAA
AAATCGAAAGCCATGATCACTCATTTATTTTGAACCTTTGATATCTCTCTCCACCTCAACA
AAGCAAGAACAAGTAAACTCTGAATGACAAGTTCAATTCATGCCTGCCGGCTCAAAAGA
ATCAATTCATTTTTGTAAACAAAAACAAAGCGGCGCAAAAAACACTGGAAATACACTTT
CTACAACTAGAATTTCAGGTGACCCCGCTTGTAGGTTATTCAAGCCTGGTAGCAAGGACG
CACCCATCAGCTGCGATTTCGGTGCGGTCAGACCGCTAAAATCGCAAAGTTATGTTGTGA
AATTTTACCTACACATGGGACACTCACACAGGTATGCTCGCAGATTAAACCGCGAGCGCC
ATTTGCGTTCAAAAATTCGATGATTCA

3. HALIMEDA #3217326_2N_1400bp_ITS2-5_8S-BF
CCAAAAGCAGGACCGCTACCAACGCCACACCTCGCCTCCACAAACACACACTCAACACT
CCGACTCAGACGGACCACTGCAAATGCCAAGCCCTAACAAACCCGCAAGCACGCACACA
CACACACACGCACACACACCCACCAGTGCAGGACTCAACTTCAGCACTCACACCCGCCC
GCTTTCGCGAACAAGCACAAACACATCCATGAATTCATCTACACCAACAAGCGCGCGCA
CGCACGCACACACACATACACACTCACACGCAAAGCTCAAACACAGAACTTCGACATACAC
AGCACCGCCACCAGTTCACACTCCGCACACCAACTGCTCCGATCAAAGATTGCTCTCCTT
ACCTTCACAGCCAGTGAATTGTAGTAACCGATCCCCAACAACGAAGCCATGCGCACAGG
TCGCCCTGCACACCTGCCTCTATTCCACATCCCGTCGCCTCACTAAGTTCACCAGAACTA
AAGAATGCCTCAACTCGCTTCCACAAAGCAGGACTCAGCTTTGGCTTGCGCCTCCACGAA

4. PADINA #3217328_3N_ITS2-5_8S-BF
CCACTATCCCACAGCAGTCACTGGATCGTAGTTACTTCCCAGAGACTGAGCAAAAAAAA
GGGGGAAGGGATGGGATAACGTTTTCCAACTCTGTCGACCGAAACAACAGAAGTTTCTC
TTCGGAAAACTCTCATCACCCCCCTCCCCTTCTCTTTCTTGCCAGGACCCAGAACCGCAA
AAGACAACCAAGCTAATATAGTACAATACACACACACAAATCCATTTCTCCCTTCTTGTA
TAATATTCCTTTCCACAGTTCACTGAAAAAGTGTAACCGGTACTAGGAACCAAAAAGTGA
GAAAGTGGGCGCGGCTTTGGGTTCTCAAGTATTCCATAAGCCAGATGTTTGTATATTTTA
GCAAAACACACAACTCTTGCGGGATGCGGTGTCGACACGCTGTACCAACTGGCCGTGTC
CAGTTGATACTATTAAAAGGCAAAGAGAATGTCACCTTTCTCCTGGTTTCGCACACTCAC
GAAAAGTCAAAACGCTTGGGGAAAAACAGTAACAACACTGGTTTCT

5. DICTYOTA #3217330_4N_ITS2-5_8S-BF
CGGAGGTCCGTAAAGACGACGCGCTGCCCGATAGCAAGCAAGCAAGCAAGCAAGCTAA
CTAACGGGCAGTCACTTACCTGACCGCGGTCGGAGGGGTCTCGGCGGGCCGGACGGTCG
GGTGAAAGGGCAGCACCGCAACAGCAGAAATCGATCAATCACTCGATCAATGGATCAAA
CAATCAACACAACCTCGCCGCCACGGGCCACACGGCCACCCAACTTCCTACTCTCGCTCT
CTGCTACTCGGCGCGCAACCCGACTTGCGCCGGACTAACGCTGCCGCGTATACGATCCGA
GACCCTTCCTCCGGGTTGGACAGTTTTTGCTACTCTTCGCGGGGCTCGGCGATGGTCGTTC
GCACGTCGGTAGAGGAGACGAAGAGACAGCTAGGATAACAACAACAACAACAACAACAACA
ACAACAACAACAACAACAGCACTACACCATCATCGTCATCATCATCGTCATCATCATCAT
CACCATCATCATCATCAGCATCCCTACTACCACCGGGAGAACGGCCACCACCGGCGCTCG
CATAGTCCGCCCAAGTACTCACCCTTTCCCCGCCGCGGCCAAGTCATCTTGGTACCGCGA
GAAGCGGTGCCGCGATGACGAGGCTAGGCGAAGACCGAGTGTGCACAAGGACGAGTGG
GGTGAACGAAGACTCCGACAAGCATGCTCCCAGGGATATCCCGGAAGCGCAAGATGCGT
TCAAAGTTTTGATGATTCGC

6. SARGASSUM #3217332_SAR_ITS2-5_8S-BF
CATAATCATAATCACCGGCCTCTCACCGGTACTCTCGAGACTACCTACCGTCCTAAGCCT
CAAGTCCACCCGCGCACATCCCCGGCGACCCGTCGTGTCCCCGCCCCGCGATAGGGGGC
GCCGTCAACAACGGATCCTCCGGAGGACGCAAGGTGAACTTGCCTCCAGCGACGCCTCC
AGCGCCGCGAAGACAACAGAGGCCCAGACAATCGGTAACTCTCTAGGCTTGGGTGGACT
CAGGGGACGAGCAGGGAGCATGTACAAAAATACACCGCACTACCGCTCCCCGGAACACT
CATATTCCGCTCCACCGCCTTGAGGTCGAGATCGAACCCAAAGCTGTGGGCGAGTTTTCG
CCTCCTCCTCCCCGACAAGCATGCTCCCAGGCATATCCCGGGAGCGCAAGATGCGTTCGA
AGTTTCGATGATTCACGAGAT

A total of four sequences for the mt23S Rdna (394 pb) and 15 sequences for the mt intergenic spacer (mt23S-tRNAVal) (333 pb) were successfully amplified and sequenced. The sequences were deposited in the GenBank.

Genetic analysis of the gene as the samples collected from Tomini Bay and Lima Island revealed a uniqe similarities to other species which is different from estimated sample as in described by their

morphologycal characteristics as well as from their herbarium. Analysis of the rbcL gene in these samples revealed some species as we see in figure 1.

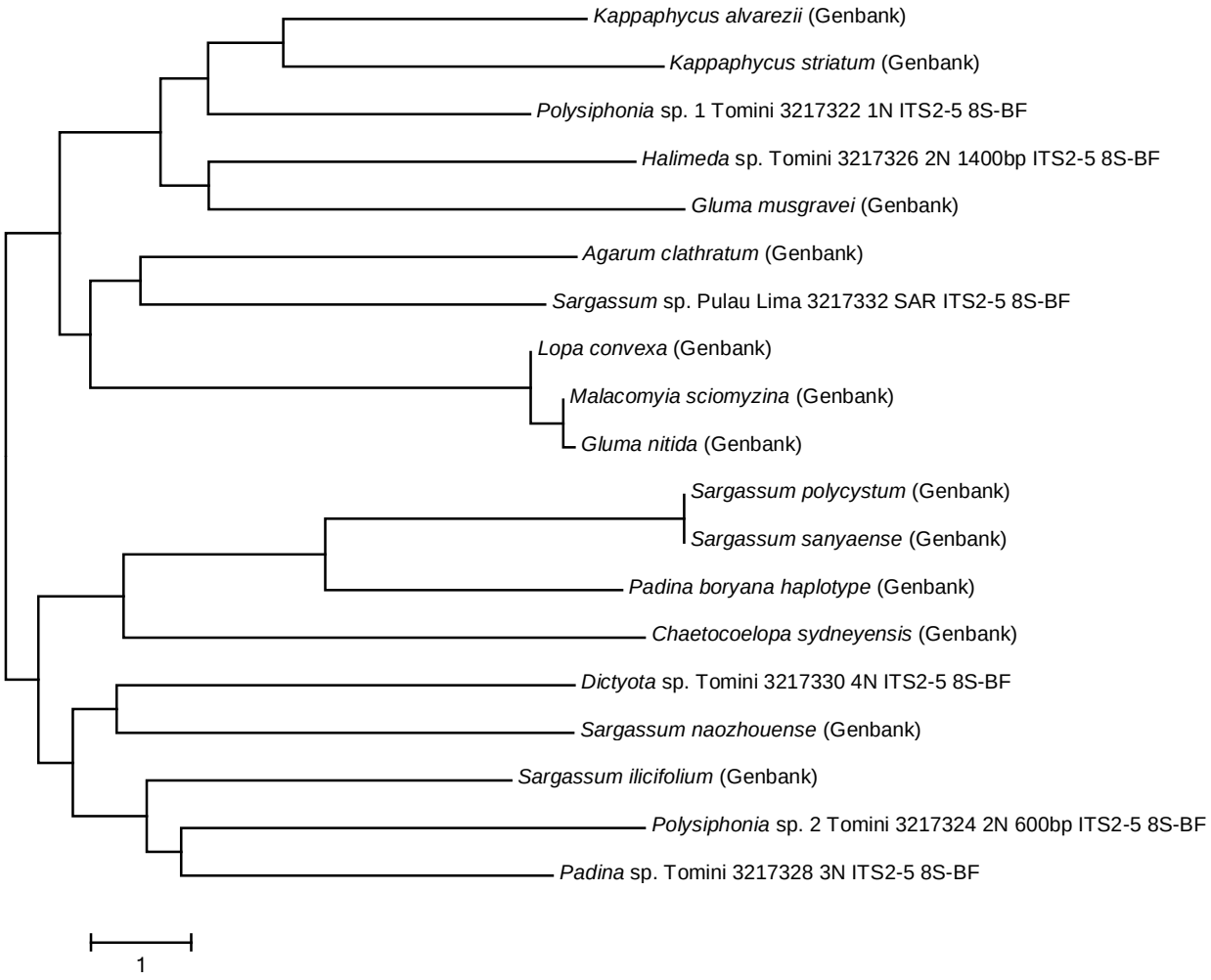

Figure 1. Neighbor joining consensus trees using partial sequences of GenBank-NCBI.

Blast of the rbcL Tomini bay sample against the GenBank database did not support precise genetic identification for all samples. One sample showed a 100% similarity, although they come from a different clade that is *Sargassum* sp from Lima Island. In the other branch (clade) Padina sp from Tomini Bay was located within a very well supported branch (clade) with Padina boryana. More morphological and molecular studies have to be carried out before a definitive conclusion can be made.

4. Conclusion

ITS2 are useful marker for DNA barcoding of wild seaweed, with advantage for ITS2 due to the larger availability of analytical tools and reference barcodes deposited at databases for this marker.

References

[1] Garreta A G 2000 Cystoseira C Agardh In: Gomez Garreta A (Editor) Flora phycologica Iberica Fucales. *University of Murcia, Publications Services, Barcelona*. **1** 99–101
[2] Draisma S G A, Ballesteros E, Rousseau F and Thibaut T 2010 DNA sequence data demonstrate the polyphyly of the genus Cystoseira and other Sargassaceae genera (Phaeophyceae) *J. Phycol.* **46** 1329-1345
[3] Thibaut T and Ballesteros E 2005 First report of Cystoseira squarrosa de Notaris (Fucophyceae, Fucales) from Spanish coastal waters. *Cryptogam Algol.* **26** 203-207
[4] Susini M, Mangialajo L, Thibaut T and Meinesz A 2007a Development of a transplantation technique of Cystoseira amentacea var. stricta and Cystoseira compressa *Hydrobiologia*. **580** 241-244
[5] Susini M, Thibaut T, Meinesz A and Forcioli D 2007b A preliminary study of genetic diversity in Cystoseira amentacea (C. Agardh) Bory var. stricta Montagne (Fucales, Phaeophyceae) using random amplified polymorphic DNA *Phycologia* **46** 605–611
[6] Mangialajo L, Ruggieri N, Asnaghi V, Chiantore M, Povero P and Cattaneo-Vietti R 2007 Ecological status in the Ligurian Sea: the effect of coastline urbanization and the importance of proper reference sites *Mar. Pollut. Bull.* **55** 30-41
[7] Mangialajo L, Chiantore M and Cattaneovietti R 2008 Loss of fucoid algae along a gradient of urbanization, and structure of benthic assemblages *Mar. Ecol. Prog. Ser*. **358** 63-74
[8a] Ercegović A 1952 Jadranske cistozire (Cystoseira of Adriatic). Institute of Oceanograpy and Fisheries *Split.* 208
[8b] Ercegović A 1959 Les facteurs de selection et d'isolement dans la genese de quelques especes d'algues Adriatiques (Factors of selection and isolation in the genesis of some algae model identification *IEEE Trans. Autom. Control* **19** 716-723
[9] Mattio L and Payri C E 2011 190 Years of Sargassum taxonomy, facing the advent of DNA phylogenies *Bot. Rev.* **77** 31-70
[10] Jegou C, Culioli G, Kervarec N, Simon G and Stiger-Pouvreau V 2010 LC/ESI-MS n and 1H HR-MAS NMR analytical methods as useful taxonomical tools within the genus Cystoseira C. Agardh (Fucales; Phaeophyceae) *Talanta* **83** 613-622
[11] Rousseau F, Burrowes R, Peters A F, Kuhlenkamp R and De Reviers B 2001 A comprehensive phylogeny of the Phaeophyceae based on nrDNA sequences resolves the earliest divergences *Life Sci*. **324** 305-319
[12] Leclerc M C, Barriel V, Lecointre G and De Reviers B 1998 Low divergence in rDNA ITS sequences among five species of Fucus (Phaeophyceae) suggests a very recent radiation *J. Mol Evol*. **46** 115-120
[13] Harvey J B J and Goff L J 2006 A reassessment of species boundaries in Cystoseira and Halidrys (Phaeophyceae, Fucales) along the North American west coast *J. Phycol*. **42** 707-720
[14] Bittner L C E, Payri A, Couloux C, Cruaud B, De Reviers and Rousseau F 2008 Molecular pylogeny of the Dictyotales and their position within the Phaeophyceae, based on nuclear, plastid and mitochondrial DNA sequence data *Mol. Phylogenet. Evol.* **49** 211-226
[15] Cho G Y, Rousseau F, De Reviers B and Boo S M 2006 Phylogenetic relationships within the Fucales (Phaeophyceae) assessed by the photosystem I coding psaA sequences *Phycologia*. **45** 512-519
[16] Coyer J A, Hoarau G, Oudot–Le Secq M, Stam W T and Olsen J L 2006 A mtDNA - based phylogeny of the brown algal genus Fucus (Heterokontophyta; Phaeophyta) *Mol. Phylogenet. Evol.* **39** 209–222
[17] De Reviers B, Rousseau F and Draisma S G 2007 Classification of the Phaeophyceae from past to present and current challenges. In: J. Brodie and J. Lewis (Editors). Unraveling the algae: the past, the present, and future of algal systematics *CRC Press. Boca Raton.* 267–279

[18] Camacho O, Mattio L, Draisma S, Fredericǫ S and Diaz-Pulido G 2015 Morphological and molecular assessment of *Sargassum* (Fucales, Phaeophyceae) from Carribbean Colombia, including the proposal of *Sargassum giganteum* sp. nov, *Sargassum schnetteri* comb. nov. and *Sargassum* section *Cladophyllum* sect. nov *Systematics and biodiversity* **13** 105-130
[19] Librado P and Rozas J 2009 DnaSP v5: A software for comprehensive analysis of DNApolymorphism data *Bioinformatics* **25** 1451-1452
[20] Larkin M A, Blackshields G, Brown N P, Chenna R, Mcgettigan P A, Mcwilliam H, Valentin F, Wallace I M, Wilm A, Lopez R, Thompson J D, Gibson T J and Higgins D G 2007 Clustal W and Clustal X version 2.0 *Bioinformatics* **23** 2947-2948
[21] Nicholas K B, Jr Nicholas H B and Ii Deerfield D W 1997 GeneDoc: Analysis and Visualization of Genetic Variation. *Embnew. news* **4** 14
[22] Akaike H 1974 A new look at the statistical model identification *IEEE Trans*. Autom Control **19** 716-723
[23] Posada D 2008 jModelTest: Phylogenetic Model Averaging *Mol. Biol.Evol*. **25** 12531256
[24] Huelsenbeck J P and Ronquist F R 2001 MrBayes: Bayesian inference of phylogenetic trees *Bioinformatics* **17** 754-745