Cryptogamie, Algologie, 2010, 31 (4): 467-485

Assessment of five markers as potential barcodes for identifying *Sargassum* subgenus *Sargassum* species (Phaeophyceae, Fucales)

Lydiane MATTIO & Claude PAYRI*

Institut de Recherche pour le Développement, COREUS2 (UR227), BPA5 Anse Vata, 98848 Nouméa, Nouvelle-Calédonie

Abstract – DNA barcoding has been the focus of numerous publications, but only limited studies are available for marine macroalgae and even less are specific to Phaeophyceae. The case study presented here assesses the potential of five different markers for use as DNA barcodes in the genus *Sargassum*: the nuclear ITS-2, a portion of the chloroplastic *RubisCO* operon and a mitochondrial spacer (mtsp), COI and *cox*3. To assess and compare the identification success of the five markers we used three criteria based on distance methods: Best Match, Best Close Match and All Species Barcodes applied to five datasets representing 13 closely related species of *Sargassum* subgenus *Sargassum*. Results demonstrated the inadequacy of ITS-2 and *RubisCO* as barcode markers while they suggested the potential of the mitochondrial markers. Additional research is needed based on numerically and geographically larger datasets to further assess the identification success of these markers.

ITS-2 / *cox*3 / COI / mitochondrial spacer / partial *RubisCO* operon / barcoding / Best Match / Best Close Match / All Species Barcodes

Résumé – Le barcode ADN fait l'objet de nombreuses études, mais un nombre limité sont disponibles pour des macroalgues marines et moins encore concernent les Phaeophyceae. L'étude de cas, présentée ici, évalue le potentiel de cinq marqueurs comme barcodes ADN pour le genre *Sargassum*. Il s'agit du marqueur nucléaire ITS-2, une portion de l'opéron chloroplastique *RubisCO* et des marqueurs mitochondriaux mtsp, COI et *cox*3. Pour évaluer et comparer le potentiel d'identification des cinq marqueurs, trois critères basés sur des méthodes de distance ont été utilisés: *Best Match, Best Close Match* et *all Species Barcodes* appliqués à cinq ensembles de données représentant 13 espèces proches du sous-genre *Sargassum*. Les résultats ont démontré l'inadéquation de l'ITS-2 et de la *RubisCO* comme marqueurs barcodes tandis qu'ils soulignent l'intérêt des marqueurs mitochondriaux. Des recherches supplémentaires basées sur de plus grands ensembles (numériques et géographiques) de données sont nécessaires pour évaluer la performance de ces marqueurs dans l'identification des espèces.

ITS-2 / *cox*3 / COI / marqueur mitochondrial / opéron chloroplastique *RubisCO* / barcoding / Best Match / Best Close Match /All Species Barcodes

* Correspondence and reprints: lydianemattio@gmail.com

INTRODUCTION

DNA barcoding has been the focus of numerous studies since its inception by Herbert *et al.* (2003), but relatively few studies have focussed on marine macroalgae and even less on Phaeophyceae. Identifying species and naming them correctly is a challenge in several morphologically plastic groups of brown algae. This is especially true for the genus *Sargassum* (Phaeophyceae, Fucales) for which about a thousand taxa have been described over the last 190 years. Even if significant taxonomic revisions have recently clarified the status and morphological variability of several species (Mattio *et al.*, 2010), mainly in the Pacific basin, identifying *Sargassum* species accurately usually requires thorough and sometimes subjective morphological examinations of several specimens. In this context, the principle and promises of DNA barcoding are appealing.

A prerequisite for DNA barcoding is the establishment of comprehensive datasets against which unknown sequences may be confidently matched to those belonging to accurately identified species. These datasets should contain multiple sequences per species, for as many localities and sister species as possible (Meyer & Paulay, 2005; Meier *et al.*, 2006), and be identified by specialists – which represents a lengthy and costly project. Ideally, a DNA barcoding marker should be easy to amplify by PCR (Polymerase Chain Reaction), short enough to be sequenced in both directions by a single primer pair, and be more variable between than within species (Kress *et al.*, 2005). The ideal marker would show a gap between the intra- and inter-specific variability – even for closely related species; the more the intra-specific variability and the inter-specific divergence overlap, the less efficient the barcoding marker is (Meyer & Paulay, 2005).

The 650 bp-long mitochondrial COI sequence was proposed by Herbert *et al.* (2004) as a marker for the identification of all livings organisms. Authors have demonstrated the usefulness of COI as a barcode marker for various organisms such as insects (Herbert *et al.*, 2004; Pons *et al.*, 2006), marine gastropods (e.g. Meyer & Paulay, 2005), and Rhodophyta (e.g. Saunders, 2005; Sherwood, 2009; Le Gall & Saunders, 2010) revealing several cryptic species (e.g. Saunders, 2008, 2009; Robba *et al.*, 2006). But other studies have pointed to the inadequacy of this marker in several groups such as Diptera (Meier *et al.*, 2006), flowering plants (Kress *et al.*, 2005), Bryophyta (Liu *et al.*, 2010) and corals (Shearer & Coffroth, 2008). For the Phaeophyceae, published studies have demonstrated that COI successfully differentiates among species of *Fucus* (Kucera & Saunders, 2008) and Laminariaceae (McDevit & Saunders, 2009), as well as for a wide variety of other brown algae from Canada (McDevit & Saunders, 2009). In a fourth study, however, only limited success was obtained for species of the genus *Alaria* due to putative hybridization and a complex species history (Lane *et al.*, 2007). Further research is needed to assess the potential of COI as putative barcoding marker in Phaeophyceae.

Sargassum represents a challenging group to test the potential of barcode markers in Phaeophyceae, in part because low sequence divergence has been reported among closely related species of the *S.* sect. *Sargassum* (Mattio *et al.*, 2008, 2009). This group has been recently revised, the species now well circumscribed, using combined morphological and molecular approaches. These studies have produced a number of sequences and significant datasets are available for the nuclear ITS-2 (Internal Transcribed Spacer 2), the chloroplast partial *RubisCO* operon (ribulose 1.5 bisphosphate carboxylase/oxygénase) and the mitochondrial *cox*3 (cytochrome oxydase 3). However, the potential of these

markers as barcodes, relative to each other and the widely championed COI, has not yet been tested. The aim of the present study, therefore, was to assess and compare the identification success of the previous markers within the genus *Sargassum*. In addition, Draisma *et al.* (in press) have suggested that the mitochondrial spacer (noted mtsp hereafter) used by Coyer *et al.* (2006) be considered for use as a DNA barcode marker in the genus *Sargassum* because it is short and variable. It is thus also considered in this study. For marker assessment and comparison, we use the three criteria identified by Meier *et al.* (2006): *Best Match* (BM), *Best Close Match* (BCM) and *All Species Barcodes* (ASB) applied to worldwide sequence datasets of 13 sister species belonging to *Sargassum* subgenus *Sargassum*.

MATERIAL & METHODS

Dataset and taxonomical identification

All sequences available for *Sargassum* subgenus *Sargassum* species were retrieved from GenBank and our personal alignments. To gain additional comparative data, a number of additional samples were extracted, PCR amplified and sequenced following Mattio *et al.* (2010) using the primers listed in Table 1. The chloroplast partial *RubisCO* operon included a portion of the *rbc*L, the *rbc*L-S spacer and a portion of the *rbc*S. The fragment for the mitochondrial mtsp included 18 bp of the 3' end of the 23S gene to the tRNA Val gene, encompassing also the tRNA Lys gene at the 5' end of the tRNA Val gene (Coyer *et al.*, 2006). Twenty one additional sequences were obtained for ITS-2, 14 for the partial *RubisCO*, 16 for *cox*3 while 50 sequences were newly obtained for the mitochondrial mtsp and 27 for COI. It must be noted here that COI sequences proved to be the most difficult to generate successfully with about 1/3 of the resulting sequences belonging to organisms co-extracted with the *Sargassum* samples, mainly microscopic Phaeophyceae.

Only sequences from species for which the species taxonomy had previously been revised and settled by relevant combined morphological examination and DNA analyses (monophyly) as detailed in Mattio *et al.* (2008, 2009, 2010) were considered. For *Sargassum* sect. *Sargassum* species, very little sequence divergence was found with all markers. However, they represent clear distinct morphological entities which were considered as distinct species (see Mattio *et al.* 2008, 2009, 2010). Sequences with uncertain identification or for which the morphological analyses were not possible (i.e. specimens for which only silica gel samples were available or for data downloaded from GenBank lacking associated herbarium vouchers for morphological analysis) were discarded. Species for which only one sequence was available in the dataset were also discarded. The five resulting alignments included data for 132 ITS-2, 87 partial *RubisCO*, 50 mtsp, 23 COI and 94 *cox*3 sequences.

Our alignments are relatively limited compared to the thousands of sequences used by Meier *et al.* (2006) in their study of Diptera or the hundreds of red algae specimens analysed by Le Gall & Saunders (2010) but equivalent to the brown algae dataset of Kucera & Saunders (2008) and McDevit & Saunders (2010) for which good results were obtained. We believe that our datasets respect the basic prerequisite for testing barcode markers by containing multiple

Table 1. List of primers and annealing temperatures corresponding to the five markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the chloroplatic partial *RubisCO* operon (*Rub*), the mitochondrial spacer (*mtsp*), and the mitochondrial Cytochrome Oxydase 1 and 3 (COI and *cox*3)

Markers	*Primers*	*Primers' sequences* [(a)]	*Reference*	*AT (°C)*
ITS-2	5.8S-BF	5'-CGATGAAGAACGCAGCGAAATGCGAT-3'	Yoshida *et al.* 2000	55
	25BR-2	5'-TCCTCCGCTTAGTATATGCTTAA-3'	Yoshida *et al.* 2000	
Rub	3F	5'-CATCGTGCTGGTAACTCTAC-3'	Phillips 1998	41
	S97R	5'-CATCTGTCCATTCWACACTAAC-3'	Peters & Ramirez 2001 [(b)]	
mtsp	mtsp-F	5'-CGTTTGGCGAGAACCTTACC-3'	Coyer *et al.* 2006	50
	mtps-R	5'-TACCACTGAGTTATTGCTCCC-3'	Coyer *et al.* 2006	
COI	GAZF2	5'-CCAACCATAAAGATATWGGTAC-3'	Lane *et al.* 2007	50
	GAZR2	5'-GGATGACCAAARAACCAAAA-3'	Lane *et al.* 2007	
*cox*3	CAF4A	5'-ATGTTTACTTGGTGRAGRGA-3'	Kogame *et al.* 2005	41
	CAR4A	5'-CCCCACCARTAWATNGTNAG-3'	Kogame *et al.* 2005	

(a) W = A or T, R = A or G, Y = C or T, N = A, T, C or G
(b) Modified by J. Buchanan

Table 2. Sequence dataset summary for the five barcode-markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the Chloroplatic partial *RubisCO* operon (*Rub*), the mitochondrial spacer (*mtsp*), and the mitochondrial Cytochrome Oxydase 1 and 3 (COI and *cox*3). The number of sequences per species and per marker is indicated

	ITS-2	*Rub*	*mtsp*	COI	*cox*3
Sequence overlap (bp)	489	609	282	650	434
Number of sequences	132	87	50	23	94
Number of species	13	13	11	7	12
Sargassum aquifolium	39	21	17	5	32
Sargassum carpophyllum	4	3	2	2	3
Sargassum howeanum	4	3	2	2	4
Sargassum ilicifolium	23	13	7	6	14
Sargassum obtusifolium	15	9	6	0	9
Sargassum pacificum	12	7	4	2	6
Sargassum polycystum	9	5	2	0	6
Sargassum polyphyllum	5	3	0	0	4
Sargassum scabridum	7	7	2	3	5
Sargassum spinuligerum	7	10	5	3	7
Sargassum swartzii	2	2	0	0	0
Sargassum sp. 1	3	2	2	0	2
Sargassum sp. 2	2	2	2	0	2

sequences per species (2 to 39, Table 2), which were sampled from several localities in the Indo-Pacific and represent 7 to 13 closely related species of *Sargassum* subgenus *Sargassum* (Table 2 & 3, Fig. 1). Figure 1 shows a Neighbour-Joining analysis of 132 ITS-2 sequences alignment representing 13 species. Congruent results were obtained for *RubisCO*, mtsp, COI and *cox*3 and may be available upon request to the first author. A detailed list of the sequences considered in this study is listed in Table 4 with GenBank accessions.

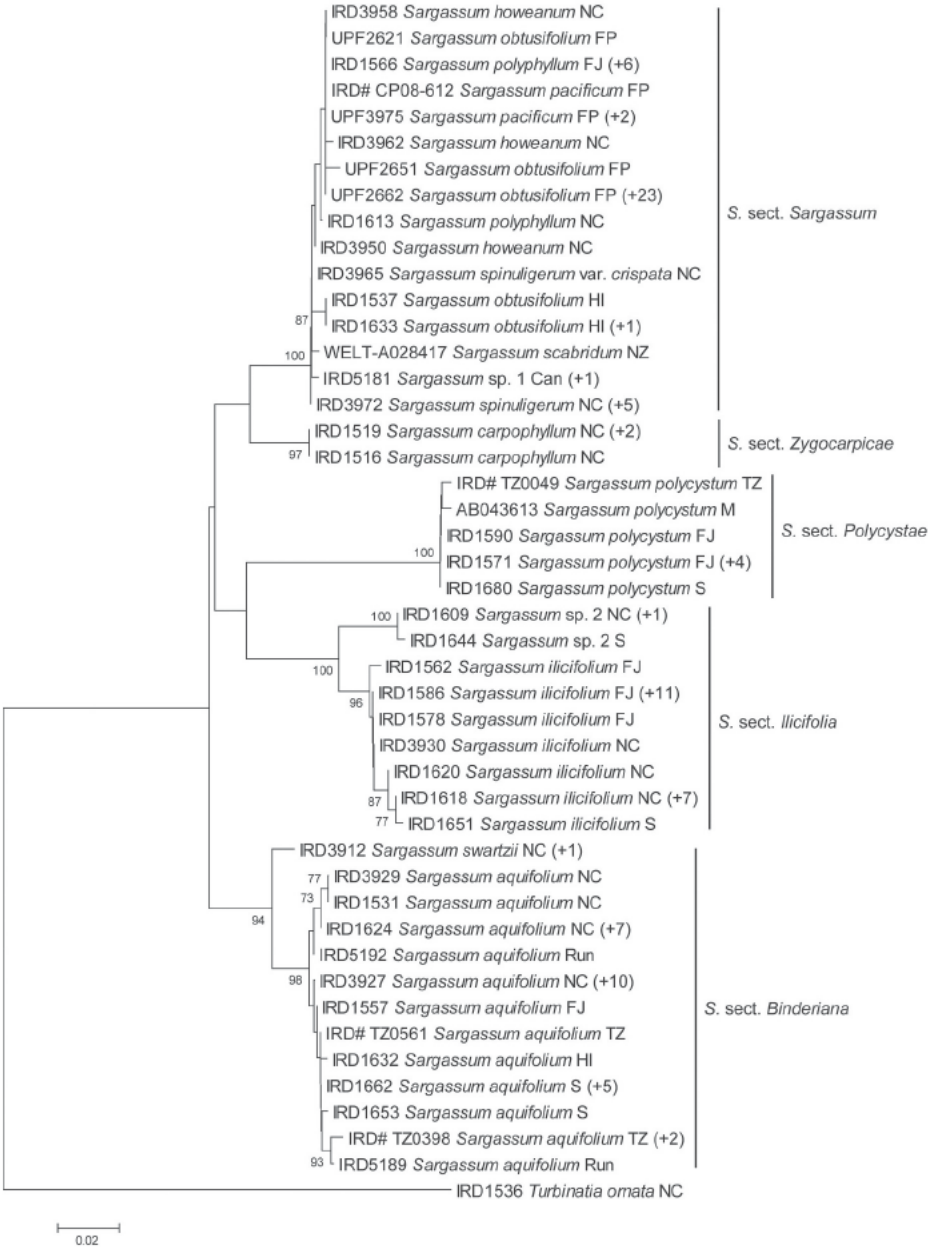

Fig. 1. Neighbour-Joining (NJ – Kimura 2 parameters) analysis based on 132 sequences alignment of the nuclear ITS-2, 489 bp-long (including gaps), representing 13 taxa of *S.* subgen. *Sargassum*. The number of identical sequences is indicated between brackets if relevant and listed in the Table 3. Bootstrap values are indicated when above 70% and were based on 1000 replicates. Root: *Turbinaria ornata*. The topology obtained with the four other markers is similar; trees may be obtained on request to the first author. Localities abbreviations: Can: Canary Is., EI: Easter Is., FJ: Fiji Is., FP: French Polynesia, HI: Hawaii, JP: Japan, K: Kermadec Is. (NZ) M: Malaysia, NC: New Caledonia, NZ: New Zealand, S: the Solomon Is., Sey: Seychelles, TZ: Tanzania, Va: Vanuatu, Viet: Vietnam, WLS: Wallis.

Table 3. Identical sequences as referred in Fig. 1

Sequence in the tree	*Identical sequences not in the tree*
IRD1662 *Sargassum aquifolium* S	IRD# Sh1466 *Sargassum aquifolium* HI IRD3925 *Sargassum aquifolium* NC IRD3949 *Sargassum aquifolium* NC IRD1681 *Sargassum aquifolium* Va UPF3976 *Sargassum aquifolium* FP
IRD1624 *Sargassum aquifolium* NC	IRD3914 *Sargassum aquifolium* NC IRD3916 *Sargassum aquifolium* NC IRD3920 *Sargassum aquifolium* NC IRD1622 *Sargassum aquifolium* NC WELT-A28413 *Sargassum aquifolium* K WELT-A23190 *Sargassum aquifolium* K IRD4041 *Sargassum aquifolium* NC
IRD3927 *Sargassum aquifolium* NC	IRD1682 *Sargassum aquifolium* Va IRD1582 *Sargassum aquifolium* FJ IRD1592 *Sargassum aquifolium* FJ IRD1660 *Sargassum aquifolium* S IRD1677 *Sargassum aquifolium* S IRD1668 *Sargassum aquifolium* S IRD1667 *Sargassum aquifolium* S IRD1666 *Sargassum aquifolium* S IRD1661 *Sargassum aquifolium* S IRD1656 *Sargassum aquifolium* S
IRD# TZ0398 *Sargassum aquifolium* TZ	IRD# TZ0449 *Sargassum aquifolium* TZ IRD# TZ0225 *Sargassum aquifolium* TZ
IRD1519 *Sargassum carpophyllum* NC	AB043067 *Sargassum carpophyllum* JP IRD1511 *Sargassum carpophyllum* NC
IRD3962 *Sargassum howeanum* NC	IRD1611 *Sargassum polyphyllum* NC IRD3953 *Sargassum polyphyllum* NC
IRD1586 *Sargassum ilicifolium* FJ	IRD# TZ0852 *Sargassum ilicifolium* TZ IRD# TZ0223 *Sargassum ilicifolium* TZ IRD# TZ0686 *Sargassum ilicifolium* TZ IRD3931 *Sargassum ilicifolium* NC IRD1589 *Sargassum ilicifolium* FJ IRD1569 *Sargassum ilicifolium* FJ IRD1652 *Sargassum ilicifolium* S IRD1647 *Sargassum ilicifolium* S IRD1645 *Sargassum ilicifolium* S EU169861 *Sargassum ilicifolium* Sey
IRD1618 *Sargassum ilicifolium* NC	IRD3935 *Sargassum ilicifolium* NC IRD3938 *Sargassum ilicifolium* NC IRD3944 *Sargassum ilicifolium* NC IRD3908 *Sargassum ilicifolium* NC IRD1616 *Sargassum ilicifolium* NC IRD1617 *Sargassum ilicifolium* NC IRD3940 *Sargassum ilicifolium* NC
UPF2662 *Sargassum obtusifolium* FP	MER1 *Sargassum obtusifolium* EI MER4 *Sargassum obtusifolium* EI MER3 *Sargassum obtusifolium* EI MER5 *Sargassum obtusifolium* EI WELT-A028408 *Sargassum scabridum* K WELT-A028409 *Sargassum scabridum* K

Table 3. Identical sequences as referred in Fig. 1 *(cont'd)*

Sequence in the tree	*Identical sequences not in the tree*
	WELT-A028410 *Sargassum scabridum* K WELT-A028411 *Sargassum scabridum* K WELT-A028412 *Sargassum scabridum* K UPF2661 *Sargassum obtusifolium* FP UPF2675 *Sargassum obtusifolium* FP UPF2633 *Sargassum obtusifolium* FP UPF2636 *Sargassum obtusifolium* FP UPF2656 *Sargassum obtusifolium* FP UPF2754 *Sargassum pacificum* FP UPF2783 *Sargassum pacificum* FP UPF3972 *Sargassum pacificum* FP UPF2763 *Sargassum pacificum* FP UPF2767 *Sargassum pacificum* FP UPF2743 *Sargassum pacificum* FP UPF2778 *Sargassum pacificum* FP UPF3978 *Sargassum pacificum* FP
IRD1633 *Sargassum obtusifolium* HI	IRD1538 *Sargassum obtusifolium* HI
UPF3975 *Sargassum pacificum* FP	UPF3973 *Sargassum pacificum* FP UPF3974 *Sargassum pacificum* FP
IRD1571 *Sargassum polycystum* FJ	AB043114 *Sargassum polycystum* Viet IRD1626 *Sargassum polycystum* WLS IRD1640 *Sargassum polycystum* Va IRD1642 *Sargassum polycystum* Va
IRD1566 *Sargassum polyphyllum* FJ	IRD3903 *Sargassum spinuligerum* NC IRD3955 *Sargassum howeanum* NC IRD3968 *Sargassum polyphyllum* NC IRD3978 *Sargassum spinuligerum* NC
IRD3972 *Sargassum spinuligerum* NC	IRD# TZ0397 *Sargassum spinuligerum* TZ IRD# TZ0400 *Sargassum spinuligerum* TZ WELT-A028416 *Sargassum scabridum* NZ IRD3961 *Sargassum spinuligerum* var. *crispata* NC IRD3963 *Sargassum spinuligerum* var. *crispata* NC
IRD5181 *Sargassum* sp. 1 Can	IRD5182 *Sargassum* sp. 1 Can
IRD1609 *Sargassum* sp. 2 NC	IRD1634 *Sargassum* sp. 2 Va
IRD3912 *Sargassum swartzii* NC	IRD1532 *Sargassum swartzii* NC

Sequence alignment and distance analyses

Partial *RubisCO*, *cox*3 and COI sequences were aligned by eye with no ambiguities using the BioEdit sequence alignment editor (Hall, 1999). ITS-2 sequences were aligned based on secondary structure as described by Stiger *et al.* (2003). Mitochondrial mtsp sequences were aligned with Clustal W using default parameters (Thompson *et al.*, 1994) and adjusted manually in BioEdit.

All distance analyses were carried out using "TaxonDNA" developed by Meier *et al.* (2006) (available at http://taxondna.sf.net) following the method proposed by Meier *et al.* (2006). In order to test if intra-specific and inter-specific genetic variability were overlapping, the frequency of intra-specific and inter-specific uncorrected pairwise distances was plotted for each dataset. The unique-

Table 4. List of sequences included in distance analyses with species name, collection site, date and collector, Herbarium and GenBank accessions for the five markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the chloroplatic partial *RubisCO* operon (*Rub*), the mitochondrial spacer (*mtsp*), and the mitochondrial Cytochrome Oxydase 1 and 3 (*cox*1 and *cox*3). Sequences with accessions from HQ416015 to HQ416141 were produced in the present study

Species	*Collection site, date and collector*	*Hb accession*	*ITS-2*	*Rub*	*mtsp*	*cox*1	*cox*3
S. aquifolium	Boucan cannot, La Réunion - 2010 - M. Zubia	IRD5189	HQ416073	HQ416018	HQ416089	–	–
S. aquifolium	Cap la Houssaye, La Réunion - 2010 - M. Zubia	IRD5190	–	–	HQ416074	–	–
S. aquifolium	Tuléar - 2008 - P. Chabanet	IRD5036	–	–	HQ416090	–	–
S. aquifolium	Easter Is., Orahe - 2008 - M-E. Ramirez	MER2	–	HQ416026	–	–	HQ416135
S. aquifolium	Easter Is., Ovahe Is. - 2007 - S. Andréfouët	IRD4009	–	–	–	–	EU882244
S. aquifolium	Fiji, Kiuva rf. - 2007 - L. Mattio	IRD1592	EU833433	EU833466	–	–	EU833411
S. aquifolium	Fiji, Makuluva Is. - 2007 - L. Mattio	IRD1557	EU833431	–	–	–	-
S. aquifolium	Fiji, Navutulevu rf. - 2007 - L. Mattio	IRD1582	EU833432	EU833464	HQ416075	–	EU833406
S. aquifolium	Fiji, Navutulevu rf. - 2007 - L. Mattio	IRD1587	–	EU833465	–	–	EU833409
S. aquifolium	French polynesia, Raivavae - 2005 - S. Andréfouët	UPF 3976	EU100795	EU100809	HQ416076	–	EU100833
S. aquifolium	Hawaiian Is., Mahukona - 2006 - T. Sauvage	IRD1546	EU833429	–	–	–	EU833389
S. aquifolium	Hawaiian Is., Mahukona - 2006 - T. Sauvage	IRD1632	EU833430	–	–	–	–
S. aquifolium	Hawaiian Is., Ohau - 2006 - T. Sauvage	IRD# Sh1466	EU100796	EU100821	HQ416077	–	EU100835
S. aquifolium	Kermadec, Raoul Is. - 2004 - R. Stanley	WELT-A23190	EU882250	EU882261	HQ416078	–	EU882240
S. aquifolium	Kermadec, Raoul Is. - 2004 - L. Chadderton	WELT-A28413	EU882251	–	–	–	–
S. aquifolium	New Caledonia, Baie tortues - 2005 - L. Mattio	IRD3927	FJ170433	–	–	–	–
S. aquifolium	New Caledonia, côte oubliée - 2007 - C. Payri	IRD3994	–	FJ170382	–	–	FJ170408
S. aquifolium	New Caledonia, Ile des Pins - 2005 - L. Mattio	IRD1531	EU100800	EU100808	–	–	EU882243
S. aquifolium	New Caledonia, Ile des Pins - 2005 - L. Mattio	IRD1622	EU882252	EU882262	–	HQ416030	EU882241
S. aquifolium	New Caledonia, Ile des Pins - 2005 - C. Payri	IRD3914	FJ170431	–	–	HQ416031	–
S. aquifolium	New Caledonia, Ile des Pins - 2005 - L. Mattio	IRD3916	FJ170432	–	–	HQ416032	–
S. aquifolium	New Caledonia, Ile des Pins - 2005 - L. Mattio	IRD3929	FJ170436	–	HQ416079	HQ416033	FJ170409
S. aquifolium	New Caledonia, Ile des Pins - 2005 - L. Mattio	IRD4041	HQ416053	–	–	HQ416034	–
S. aquifolium	New Caledonia, Ile Nouaré - 2005 - L. Mattio	IRD1624	EU882253	EU882263	HQ416086	–	EU882242
S. aquifolium	New Caledonia, île Nouaré - 2005 - L. Mattio	IRD3920	FJ170437	FJ170383	HQ416087	–	FJ170410
S. aquifolium	New Caledonia, Ouano - 2005 - L. Mattio	IRD3925	FJ170434	FJ170380	–	–	FJ170405
S. aquifolium	New Caledonia, Prony - 2007 - G. Lasne	IRD3948	–	–	–	–	FJ170406
S. aquifolium	New Caledonia, Prony - 2007 - G. Lasne	IRD3949	FJ170435	FJ170381	–	–	FJ170407
S. aquifolium	Solomon Is. Malaita ST848 - 2004 - C. Payri	IRD1661	EU833446	–	–	–	EU833396
S. aquifolium	Solomon Is. Malaita ST848 - 2004 - C. Payri	IRD1662	EU833445	–	–	–	EU833395
S. aquifolium	Solomon Is. Malaita ST850 - 2004 - C. Payri	IRD1660	EU833447	EU833462	–	–	EU833397

Table 4. List of sequences included in distance analyses with species name, collection site, date and collector, Herbarium and GenBank accessions for the five markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the chloroplatic partial *RubisCO* operon (*Rub*), the mitochondrial spacer (*mtsp*), and the mitochondrial Cytochrome Oxydase 1 and 3 (*cox*1 and *cox*3). Sequences with accessions from HQ416015 to HQ416141 were produced in the present study *(cont'd)*

Species	*Collection site, date and collector*	*Hb accession*	*ITS-2*	*Rub*	*mtsp*	*cox*1	*cox*3
S. aquifolium	Solomon Is. Malaita ST852 - 2004 - C. Payri	IRD1653	EU833449	–	–	–	EU833400
S. aquifolium	Solomon Is. Malaita ST852 - 2004 - C. Payri	IRD1656	EU833448	–	–	–	EU833398
S. aquifolium	Solomon Is. Nggela ST842 - 2004 - C. Payri	IRD1666	EU833444	–	–	–	EU833394
S. aquifolium	Solomon Is. Nggela ST842 - 2004 - C. Payri	IRD1667	EU833443	–	–	–	-
S. aquifolium	Solomon Is. Nggela ST842 - 2004 - C. Payri	IRD1668	EU833442	–	–	–	EU833393
S. aquifolium	Solomon Is. Nggela ST842 - 2004 - C. Payri	IRD1677	EU833441	EU833463	–	–	EU833392
S. aquifolium	Tanzania - n.a. - H. Verbruggen	IRD# TZ0398	HQ416054	–	HQ416080	–	HQ416126
S. aquifolium	Tanzania - n.a. - H. Verbruggen	IRD# TZ0225	HQ416055	–	HQ416081	–	-
S. aquifolium	Tanzania - n.a. - H. Verbruggen	IRD# TZ0449	HQ416056	–	HQ416082	–	HQ416127
S. aquifolium	Tanzania - n.a. - H. Verbruggen	IRD# TZ0561	HQ416057	–	HQ416083	–	HQ416128
S. aquifolium	Vanuatu Santo - 2006 - L. Mattio	IRD1681	EU833456	EU833476	HQ416084	–	EU833412
S. aquifolium	Vanuatu Santo - 2006 - L. Mattio	IRD1682	EU833455	EU833477	HQ416085	–	EU833413
S. aquifolium	Hawaii, Phillips et al., 2005	–	–	AY590501	–	–	–
S. aquifolium	Hawaii, Phillips et al., 2005	–	–	AF076689	–	–	–
S. carpophyllum	New Caledonia, Basse Taui - 2005 - L. Mattio	IRD1516	EU100799	EU100806	–	–	EU833416
S. carpophyllum	New Caledonia, Feycinet Is. - 2005 - L. Mattio	IRD1511	EU100797	EU100804	HQ416091	HQ416035	EU833415
S. carpophyllum	New Caledonia, Porc Epic Is. - 2005 - L. Mattio	IRD1519	EU100798	EU100805	HQ416092	HQ416036	EU833417
S. carpophyllum	Japan, Stiger et al., 2000	–	AB043067	–	–	–	–
S. howeanum	New Caledonia, îlot Canard - 2005 - L. Mattio	IRD3958	FJ170439	FJ170385	–	–	FJ170412
S. howeanum	New Caledonia, îlot Maître - 2005 - L. Mattio	IRD3950	FJ170441	FJ170386	HQ416094	HQ416037	FJ170414
S. howeanum	New Caledonia, Ouano - 2005 - L. Mattio	IRD3962	FJ170438	FJ170384	HQ416093	HQ416038	FJ170411
S. howeanum	New Caledonia, Rocher voile - 2005 - L. Mattio	IRD3955	FJ170440	–	–	–	FJ170413
S. ilicifolium	Fiji, Kiuva rf. - 2007 - L. Mattio	IRD1588	–	EU833467	–	–	EU833408
S. ilicifolium	Fiji, Kiuva rf. - 2007 - L. Mattio	IRD1589	EU833439	EU833470	HQ416095	–	–
S. ilicifolium	Fiji, Makuluva Is. - 2007 - L. Mattio	IRD1562	EU833436	EU833468	HQ416096	–	–
S. ilicifolium	Fiji, Makuluva Is. - 2007 - L. Mattio	IRD1569	EU833437	EU833469	HQ416097	–	EU833403
S. ilicifolium	Fiji, Navutulevu rf. - 2007 - L. Mattio	IRD1578	EU833440	–	–	–	EU833404
S. ilicifolium	Fiji, Navutulevu rf. - 2007 - L. Mattio	IRD1586	EU833438	–	–	–	EU833407
S. ilicifolium	New Caledonia, Crouy - 2006 - L. Mattio	IRD3931	FJ170443	FJ170387	HQ416098	–	FJ170416
S. ilicifolium	New Caledonia, Ile des Pins - 2005 - L. Mattio	IRD1617	EU882249	EU882260	–	HQ416039	EU882239
S. ilicifolium	New Caledonia, Ile des Pins - 2005 - L. Mattio	IRD3935	FJ170442	–	–	HQ416040	FJ170415

Table 4. List of sequences included in distance analyses with species name, collection site, date and collector, Herbarium and GenBank accessions for the five markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the chloroplatic partial *RubisCO* operon (*Rub*), the mitochondrial spacer (*mtsp*), and the mitochondrial Cytochrome Oxydase 1 and 3 (*cox*1 and *cox*3). Sequences with accessions from HQ416015 to HQ416141 were produced in the present study *(cont'd)*

Species	*Collection site, date and collector*	*Hb accession*	*ITS-2*	*Rub*	*mtsp*	*cox*1	*cox*3
S. ilicifolium	New Caledonia, Ile des Pins - 2005 - C. Payri	IRD3938	FJ170446	–	–	HQ416041	FJ170417
S. ilicifolium	New Caledonia, Ile des Pins - 2006 - L. Mattio	IRD3940	HQ416058	HQ416029	–	HQ416042	HQ416141
S. ilicifolium	New Caledonia, Ile des Pins - 2005 - C. Payri	IRD3944	FJ170447	–	–	HQ416043	–
S. ilicifolium	New Caledonia, île Nouaré - 2005 - L. Mattio	IRD3908	FJ170448	–	–	–	–
S. ilicifolium	New Caledonia, îlot Canard - 2005 - L. Mattio	IRD3930	FJ170444	–	–	–	–
S. ilicifolium	New Caledonia, Isle of Pines - 2005 - L. Mattio	IRD1616	EU833435	EU833460	–	HQ416044	EU833391
S. ilicifolium	New Caledonia, Nouaré Is. - 2005 - L. Mattio	IRD1618	EU833434	EU833461	–	–	EU833390
S. ilicifolium	New Caledonia, Rocher voile - 2005 - L. Mattio	IRD1620	FJ170445	–	–	–	EU882238
S. ilicifolium	Solomon Is. Malaita ST852 - 2004 - C. Payri	IRD1645	EU833453	–	–	–	–
S. ilicifolium	Solomon Is. Malaita ST852 - 2004 - C. Payri	IRD1647	EU833452	–	–	–	–
S. ilicifolium	Solomon Is. Malaita ST852 - 2004 - C. Payri	IRD1652	EU833451	–	–	–	–
S. ilicifolium	Solomon Is. Nggela ST842 - 2004 - C. Payri	IRD1651	EU833454	–	–	–	–
S. ilicifolium	Tanzania - n.a. - H. Verbruggen	IRD# TZ0223	HQ416059	HQ416023	HQ416099	–	–
S. ilicifolium	Tanzania - n.a. - H. Verbruggen	IRD# TZ0686	HQ416060	HQ416021	HQ416101	–	HQ416129
S. ilicifolium	Tanzania - n.a. - H. Verbruggen	IRD# TZ0852	HQ416061	HQ416022	HQ416100	–	HQ416133
S. ilicifolium	Trois bassins, île de la Réunion - 2010 - M. Zubia	IRD5192	HQ416062	–	HQ416088	–	HQ416138
S. ilicifolium	Seychelles, Rohfritsch *et al.*, 2010	SEY1	EU169861	EU169868	–	–	–
S. obtusifolium	Easter Is., Hanga Oteo - 2008 - R. Garcia Huidobio & M-E. Cordora	MER5	HQ416066	HQ416028	–	–	HQ416137
S. obtusifolium	Easter Is., Hanga Roa - 2008 - M-E. Ramirez	MER3	HQ416064	HQ416027	HQ416102	–	HQ416136
S. obtusifolium	Easter Is., Inter mareal baja - 2007 - M-E. Ramirez	MER4	HQ416065	–	HQ416104	–	–
S. obtusifolium	Easter Is., Playa Anakena - 2008 - M-E. Ramirez	MER1	HQ416063	HQ416025	HQ416103	–	HQ416134
S. obtusifolium	French polynesia, Rapa - 2002 - C. Payri	UPF 2675	EU100786	EU100815	–	–	EU100831
S. obtusifolium	French polynesia, Rapa, Australs - 2002 - C. Payri	UPF 2621	EU100789	–	–	–	–
S. obtusifolium	French polynesia, Rapa, Australs - 2002 - C. Payri	UPF 2633	EU100793	–	–	–	–
S. obtusifolium	French polynesia, Rapa, Australs - 2002 - C. Payri	UPF 2636	EU100790	–	–	–	–
S. obtusifolium	French polynesia, Rapa, Australs - 2002 - C. Payri	UPF 2651	EU100785	EU100819	HQ416105	–	EU100830
S. obtusifolium	French polynesia, Rapa, Australs - 2002 - C. Payri	UPF 2656	EU100792	–	–	–	–
S. obtusifolium	French polynesia, Rapa, Australs - 2002 - C. Payri	UPF 2661	EU100788	EU100818	–	–	EU100829
S. obtusifolium	French polynesia, Rapa, Australs - 2002 - C. Payri	UPF 2662	EU100787	EU100816	HQ416106	–	EU100832
S. obtusifolium	Hawaiian Is., Ho'Okena - 2007 - T. Sauvage	IRD1539	–	–	–	–	EU833386

Table 4. List of sequences included in distance analyses with species name, collection site, date and collector, Herbarium and GenBank accessions for the five markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the chloroplatic partial *RubisCO* operon (*Rub*), the mitochondrial spacer (*mtsp*), and the mitochondrial Cytochrome Oxydase 1 and 3 (*cox*1 and *cox*3). Sequences with accessions from HQ416015 to HQ416141 were produced in the present study *(cont'd)*

Species	*Collection site, date and collector*	*Hb accession*	*ITS-2*	*Rub*	*mtsp*	*cox*1	*cox*3
S. obtusifolium	Hawaiian Is., Maui - 2006 - A. Rivera	IRD1537	EU100794	EU100820	HQ416107	–	EU100834
S. obtusifolium	Hawaiian Is., Molokai SE - 2007 - T. Sauvage	IRD1538	EU833428	–	–	–	–
S. obtusifolium	Hawaiian Is., Molokai W - 2007 - T. Sauvage	IRD1633	EU833427	–	–	–	–
S. obtusifolium	Hawaiian Is., Phillips *et al.* 2005	–	–	AF244328	–	–	–
S. pacificum	French polynesia, Bora Bora - 2003 - C. Payri & V. Stiger	UPF 2754	EU100783	EU100812	HQ416108	–	EU100824
S. pacificum	French polynesia, Bora Bora, Society - 2003 - C. Payri & V. Stiger	UPF 2763	EU100780	–	–	–	–
S. pacificum	French polynesia, Bora Bora - 2003 - C. Payri & V. Stiger	UPF 2767	EU100781	–	–	–	–
S. pacificum	French polynesia, Moorea - 2008 - L. Mattio	IRD# CP08-612	HQ416067	HQ416015	HQ416109	–	–
S. pacificum	French polynesia, Raiatea - 2003 - C. Payri & V. Stiger	UPF 2743	EU100782	–	–	–	–
S. pacificum	French polynesia, Raiatea - 2003 - C. Payri & V. Stiger	UPF 2778	EU100779	–	–	–	–
S. pacificum	French polynesia, Raiatea - 2003 - C. Payri & V. Stiger	UPF 2783	EU100774	EU100811	HQ416110	–	EU100823
S. pacificum	French polynesia, Raiatea - 2003 - C. Payri & V. Stiger	UPF 3972	EU100784	EU100813	HQ416111	–	EU100828
S. pacificum	French polynesia, Raiatea - 2003 - C. Payri & V. Stiger	UPF 3978	EU100778	–	–	–	-
S. pacificum	French polynesia, Tahiti - 2003 - C. Payri & V. Stiger	UPF 3973	EU100776	EU100810	–	HQ416045	EU100826
S. pacificum	French polynesia, Tahiti - 2003 - C. Payri & V. Stiger	UPF 3974	EU100777	EU100814	HQ416125	–	EU100825
S. pacificum	French polynesia, Tahiti - 2003 - C. Payri & V. Stiger	UPF 3975	EU100775	EU100817	–	HQ416046	EU100827
S. polycystum	Fiji, Kiuva rf. - 2007 - L. Mattio	IRD1590	EU833421	EU833472	HQ416112	–	EU833410
S. polycystum	Fiji, Makuluva Is. - 2007 - L. Mattio	IRD1571	EU833422	EU833471	–	–	EU833405
S. polycystum	Solomon Is. Malaita - 2004 - C. Payri	IRD1680	EU833423	–	–	–	EU833399
S. polycystum	Tanzania - n.a. - H. Verbruggen	IRD# TZ0049	HQ416068	HQ416020	HQ416113	–	HQ416130
S. polycystum	Vanuatu Port Vila - 2006 - C. Payri	IRD1640	EU833420	EU833475	–	–	EU833388
S. polycystum	Vanuatu Port Vila - 2006 - C. Payri	IRD1642	EU833419	EU833474	–	–	EU833387
S. polycystum	Wallis, Matu utu - 2005 - C. Chauvet	IRD1626	EU833418	–	–	–	–
S. polycystum	Vietnam, Stiger *et al.*, 2000	–	AB043613	–	–	–	–
S. polycystum	Vietnam, Stiger *et al.*, 2000	–	AB043114	–	–	–	–
S. polyphyllum	Fiji, Makuluva Is. - 2007 - L. Mattio	IRD1566	EU833426	–	–	–	EU833401
S. polyphyllum	Fiji, Navutulevu rf. - 2007 - L. Mattio	IRD1574	–	–	–	–	EU833402
S. polyphyllum	New Caledonia, Ile des Pins - 2005 - C. Payri	IRD3968	FJ170450	–	–	–	–
S. polyphyllum	New Caledonia, îlot M'Bo - 2005 - L. Mattio	IRD3953	FJ170449	–	–	–	–
S. polyphyllum	New Caledonia, Maître Is. - 2005 - L. Mattio	IRD1613	EU833424	EU833458	–	–	EU833385

Table 4. List of sequences included in distance analyses with species name, collection site, date and collector, Herbarium and GenBank accessions for the five markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the chloroplatic partial *RubisCO* operon (*Rub*), the mitochondrial spacer (*mtsp*), and the mitochondrial Cytochrome Oxydase 1 and 3 (*cox*1 and *cox*3). Sequences with accessions from HQ416015 to HQ416141 were produced in the present study *(cont'd)*

Species	*Collection site, date and collector*	*Hb accession*	*ITS-2*	*Rub*	*mtsp*	*cox*1	*cox*3
S. polyphyllum	New Caledonia, Ouano - 2005 - L. Mattio	IRD1611	EU833425	EU833459	–	–	–
S. polyphyllum	Tanzania - n.a. - H. Verbruggen	IRD# TZ0626	–	–	–	–	HQ416132
S. polyphyllum	Hawaii, Phillips *et al.*, 2005	–	–	AY518391	–	–	-
S. scabridum	Kermadec, Meyer Is. - 2005 - C. Duffy	WELT-A028408	FJ170453	FJ170390	–	–	FJ170420
S. scabridum	Kermadec, Meyer Is. - 2005 - C. Duffy	WELT-A028409	FJ170454	FJ170391	–	–	FJ170421
S. scabridum	Kermadec, Meyer Is. - 2005 - C. Duffy	WELT-A028410	FJ170455	FJ170392	–	HQ416047	FJ170422
S. scabridum	Kermadec, Meyer Is. - 2005 - C. Duffy	WELT-A028411	FJ170456	FJ170393	–	HQ416048	FJ170423
S. scabridum	Kermadec, Meyer Is. - 2005 - C. Duffy	WELT-A028412	FJ170457	FJ170394	HQ416115	–	FJ170424
S. scabridum	New Zealand, Auckland - 2005 - W. Nelson	WELT-A028417	FJ170451	FJ170388	HQ416114	–	–
S. scabridum	New Zealand, Whangaparoa - 2005 - T. Farr	WELT-A028416	FJ170452	FJ170389	–	HQ416049	–
S. sp. 1	Canary Islands, Corralejo, Fuerteventura - 2010 - F. Mineur	IRD5181	HQ416069	HQ416016	HQ416116	–	HQ416139
S. sp. 1	Canary Islands, El cotillo, Fuerteventura - 2010 - F. Mineur	IRD5182	HQ416070	HQ416017	HQ416117	–	HQ416140
S. sp.	New Caledonia, Rocher à la voile - 2005 - L. Mattio	IRD1609	EU882248	EU882259	HQ416119	–	EU882237
S. sp.	Solomon Is. Malaita ST848 - 2004 - C. Payri	IRD1644	EU833450	–	–	–	–
S. sp.	Vanuatu Port Vila - 2006 - C. Payri	IRD1634	EU833457	EU833473	HQ416118	–	EU833414
S. spinuligerum	New Caledonia, Basse deTaui - 2005 - L. Mattio	IRD3985	–	FJ170400	–	–	FJ170427
S. spinuligerum	New Caledonia, Ile des Pins - 2005 - C. Payri	IRD3972	FJ170460	FJ170397	–	HQ416050	FJ170425
S. spinuligerum	New Caledonia, Ile des Pins - 2005 - L. Mattio	IRD3974	–	FJ170398	–	–	–
S. spinuligerum	New Caledonia, îlot M'Bo - 2005 - L. Mattio	IRD3903	FJ170461	FJ170399	HQ416120	–	FJ170426
S. spinuligerum	New Caledonia, îlot M'Bo - 2005 - L. Mattio	IRD3963	FJ170463	FJ170402	HQ416121	–	FJ170429
S. spinuligerum	New Caledonia, îlot Maître - 2005 - L. Mattio	IRD3965	FJ170465	FJ170404	–	HQ416052	FJ170430
S. spinuligerum	New Caledonia, îlot Signal - 2005 - L. Mattio	IRD3978	FJ170462	FJ170401	HQ416122	HQ416051	FJ170428
S. spinuligerum	New Caledonia, Rocher voile - 2005 - L. Mattio	IRD3961	FJ170464	FJ170403	–	–	–
S. spinuligerum	Tanzania - n.a. - H. Verbruggen	IRD# TZ0397	HQ416071	HQ416024	HQ416123	–	–
S. spinuligerum	Tanzania - n.a. - H. Verbruggen	IRD# TZ0400	HQ416072	HQ416019	HQ416124	–	HQ416131
S. swartzii	New Caledonia, Ilot Maître - 2005 - L. Mattio	IRD3912	EU882255	EU882264	–	–	–
S. swartzii	New Caledonia, Thio - 2006 - C. Berthot	IRD1532	EU882254	EU100807	–	–	–

ness of a barcode sequence for a given species was assessed for each marker by testing if identical sequences were shared by individuals of different species. Species-specific barcode sequences were constructed as consensus of all con-specific sequences (for each marker independently), and their uniqueness was also tested. The intraspecific sequence variability was summarized using IUPAC (International Union of Pure and Applied Chemistry) codes as suggested by Meier *et al.* (2006).

The three criteria BM, BCM and ASB were calculated using TaxonDNA for each alignment. The principle of *Best Match* (BM) is to find for each sequence, one by one, its closest pairwise match in an alignment. The identification is considered a success if the query sequence and the match are from the same species, mismatched names are considered as failures, while equally good matches from different species are considered as ambiguous. The principle of *Best Close Match* criterion (BCM) is similar to BM except that the relative frequency of all intra-specific distances is plotted to determine a threshold value below 95%. All sequences without a match below the threshold value remain unidentified. The *All Species Barcode* criterion (ASB) assembles a list of all barcode sequences similar to the query sequence using the same threshold as for BCM. Identification is considered a success if the query sequence is associated to all con-specific barcode sequences. It is considered ambiguous if the query sequence is associated to only part of the con-specific barcode sequences. Finally, the query sequence is considered as misidentified if it is associated to barcode sequences which do not belong to the good species.

RESULTS

Intra- versus inter-specific variability and barcode uniqueness

Plots of intra-specific and inter-specific uncorrected pairwise distances for each barcode-marker are shown on Fig. 2. Intra-specific and inter-specific distances were found to overlap for each marker (4.8 to 20.93% of all observations) with the exception of the nuclear ITS-2 for which no overlap is observable within the 0.2-0.4% interval of pairwise distances. The nuclear ITS-2 also represents the marker for which the least observations are present within the total overlap interval (4.8% of all observations) while the partial chloroplastic *RubisCO* shows the highest number of observations within both the total and the 90% overlap (20.93% of all observations, Fig. 2).

Pairwise comparison of all barcodes and species consensus barcodes showed mixed results for each of the markers assessed (Table 5). Both individual barcodes and species consensus barcodes were unique in *mstp* and COI datasets while none was unique in the partial *RubisCO* dataset and only species consensus barcodes were unique in ITS-2 and *cox*3 datasets.

Best Match, Best Close Match and All Species Barcodes

Results for the three criteria (BM, BCM and ASB) are displayed in Table 5 and Fig. 3. They appeared unrelated to the overlap results (e.g. no overlap in ITS-2 pairwise distances while its identification success is the lowest). The ASB criterion had the least success with the maximum ambiguous identifications

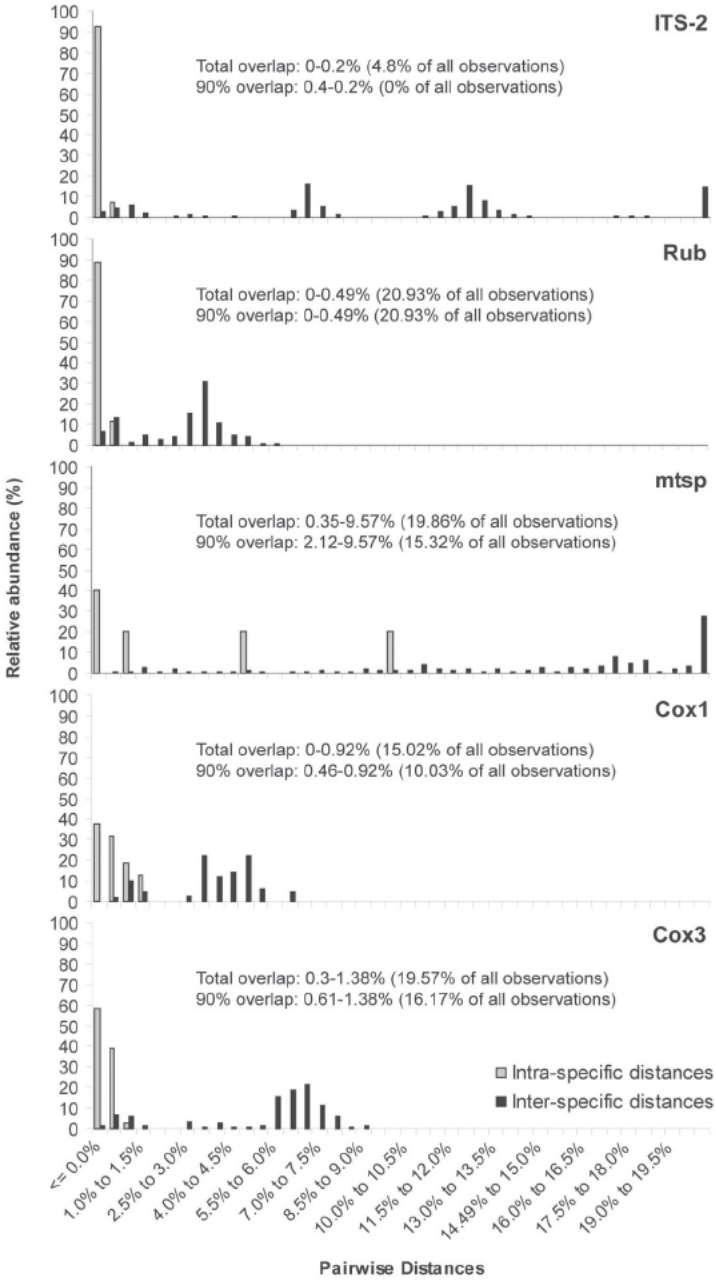

Fig. 2. Plot of intra- and inter-specific genetic distances for the five barcode-markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the Chloroplatic partial *RubisCO* operon (Rub), the mitochondrial spacer (mtsp), and the mitochondrial Cytochrome Oxydase 1 and 3 (COI and *cox*3). Total overlap range (with corresponding percentage of observations it represents) and 90% overlap (largest 5% of the intra-specific and lowest 5% of the inter-specific excluded) are indicated.

Table 5. Identification success based on the three different criteria: *Best Match* (BM), *Best Close Match* (BCM) and *All Species Barcodes* (ASB) for the five barcode-markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the chloroplatic partial *RubisCO* operon (*Rub*), the mitochondrial spacer (*mtsp*), and the mitochondrial Cytochrome Oxydase 1 and 3 (COI and *cox*3). Uniqueness of barcodes per species and uniqueness of species consensus barcodes are also indicated

	Uniqueness of barcodes per species?	*Uniqueness of species consensus barcodes?*	*Criterion*	*Success (%)*	*Ambiguous (%)*	*Misidentified (%)*	*No match (%)*
ITS-2	no	yes	BM	58.33	39.39	2.27	0
			BCM	46.96	34.08	1.51	17.42
			ASB	8.33	73.48	0.75	17.42
Rub	no	no	BM	63.21	35.63	1.14	0
			BCM	62.06	34.48	1.14	2.29
			ASB	45.97	51.72	0	2.29
mtsp	yes	yes	BM	83.01	3.77	13.2	0
			BCM	81.13	3.77	13.2	1.88
			ASB	45.28	50.94	1.88	1.88
COI	yes	yes	BM	91.3	4.34	4.34	0
			BCM	91.3	4.34	4.34	0
			ASB	60.86	34.78	4.34	0
*cox*3	no	yes	BM	74.64	23.4	2.12	0
			BCM	71.27	22.34	2.12	4.25
			ASB	65.95	29.78	0	4.25

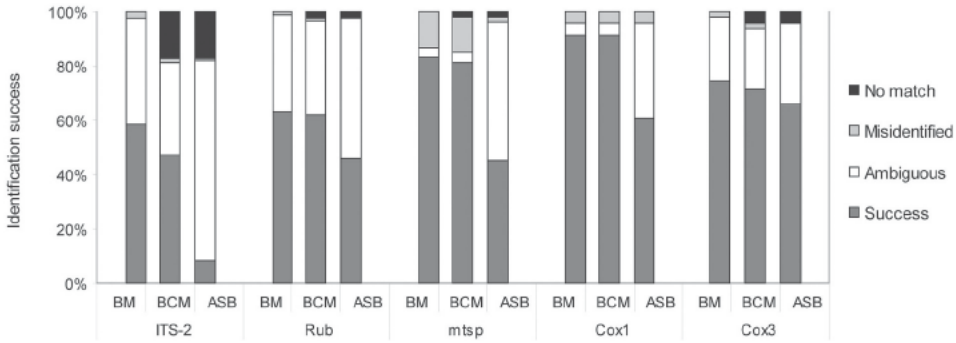

Fig. 3. Identification success based on the three different criteria: Best Match (BM), Best Close Match (BCM) and All Species Barcodes (ASB) for the five barcode-markers assessed in this study: the nuclear Internal Transcribed Spacer 2 (ITS-2), the Chloroplatic partial *RubisCO* operon (Rub), the mitochondrial spacer (mtsp), and the mitochondrial Cytochrome Oxydase 1 and 3 (COI and *cox*3).

combined to the lowest misidentifications scores. BM and BCM criteria showed relatively similar results with success always higher than ambiguous and misidentified scores. Therefore only these two criteria have been considered hereafter for comparisons between markers. The best identification success was obtained for the mitochondrial COI (91.3% for both BM and BCM), followed by the mitochondrial spacer *mtsp* (83.01% for BM and 81.13% for BCM), and the mito-

chondrial *cox*3 (74.64% for BM and 71.27% for BCM) while the nuclear ITS-2 and the partial chloroplastic *RubisCO* obtained less success (Table 5). The mitochondrial *mtsp* and COI were the markers showing the least ambiguous identifications (3.77-4.34%) and no match scores (0-1.88%) but the maximum misidentification scores (4.34-13.2%); while ITS-2, partial *RubisCO* and *cox*3 showed relatively high scores of ambiguous identifications (22.34-39.39%) and no match (2.29-17.42%) and the lowest misidentification scores (1.14-2.27%).

DISCUSSION

The first and most obvious result of this study is that none of the five markers assessed was able to identify accurately (ie. 100% identification success) closely related, but morphologically well circumscribed, species of *Sargassum* subgenus *Sargassum*.

For instance, the COI barcode-marker showed the best global results in distance analyses compared to the four other markers. Both individual COI sequences and species consensus barcodes were unique, identification success using BM and BCM was the highest (91.3%) and ambiguous, misidentification and no match scores were among the lowest (Table 5). Nevertheless, an important proportion of intra- and inter-specific distances were found to overlap (Fig. 2) and the marker proved very difficult to sequence with 1/3 of the obtained sequences not corresponding to the target organisms (*Sargassum*). It must also be underlined that these results were based on a very limited sampling (23 sequences representing 7 species, Table 2) and several authors have demonstrated that identification success of a barcode marker may vary in regard to the size, geographical span and relatedness of species of/in the dataset considered (Meyer & Paulay, 2005; Meier *et al.*, 2006; Lohse, 2009; Bittner *et al.*, 2010; Liu *et al.*, 2010).

Both the mitochondrial *mtsp* and *cox*3 markers have shown relatively good results with unique individual (*mstp* only) and species consensus barcodes, good identification success under BM and BCM criteria (71.27-83.01%), and reasonable 'ambiguous + misidentified + no match' scores (16.97-28.71%). The alignment of *cox*3 sequences was straightforward but the alignment of *mtsp* sequences was not possible without the help of clustal W and significant re-adjustment by eye. Moreover, a large proportion of intra- and inter-specific distances were found to overlap in both markers (15.32-19.86% of all observations). Conversely to the three mitochondrial datasets, ITS-2 and partial *RubisCO* datasets demonstrated low success for identification purposes. Identification success under BM and BCM criteria were below 65% while 'ambiguous + misidentified + no match' scores ranged from 36.77 to 53.01%. The alignment of the partial *RubisCO* sequences was straight forward while that of the nuclear sequences requested significant eye-adjustment to fit the secondary structure of ITS-2.

The very limited barcoding identification success of ITS-2 and partial *RubisCO* markers suggests that they do not represent good candidates for DNA barcoding studies in *Sargassum*. However, the identification success of COI obtained in this study encourages further research to assess its potential in a numerically and geographically larger dataset. If COI was to be considered as barcode marker for *Sargassum*, its practical limits, in particular concerning sequencing efficiency, would have to be overcome with more specific primers. The mitochondrial *mtsp* and *cox*3 could represent alternative candidates. Both

markers' sequences are easy to produce but if *cox*3 sequences are easy to align, the alignment of *mtsp* sequences is time consuming. Both datasets are relatively limited (50 and 94 sequences respectively) and would need to be reassessed based on larger datasets.

Finally, we were surprised that ITS-2 and partial *RubisCO* led to such low identification success since they represent two of the markers usually employed in molecular-assisted taxonomic studies of *Sargassum* (Stiger *et al.*, 2003; Philipps *et al.*, 2005; Mattio *et al.*, 2008, 2009, 2010). We suspected that the low genetic variability reported for these markers in one of the sections of the genus (*S.* section *Sargassum*) may be at the origin of these low scores. To test this hypothesis, analyses of the *Best Close Match* criterion were re-run using modified datasets for which the closely related but morphologically well circumscribed species belonging to *S.* sect. *Sargassum* were considered as a single taxon. All markers (Table 6) show high scores of identification success (71.21-97.7%) combined to lower values in ambiguous, misidentification and no match scores. These results confirm the impact of the low genetic variability of *S.* section *Sargassum* species on the barcoding identification success in the fives markers assessed here.

Table 6. Identification success (%) based on *Best Match* (BM) and *Best Close Match* (BCM) for the five markers assessed in this study based on modified datasets where all taxa belonging to *S.* section *Sargassum* were considered as a single taxon

	Criterion	*Success (%)*	*Ambiguous (%)*	*Misidentified (%)*	*No match (%)*
ITS-2	BM	85.6	12.87	1.51	0
	BCM	71.21	9.84	1.51	17.42
Rub	BM	100	0	0	0
	BCM	97.7	0	0	2.29
mtsp	BM	98.11	0	1.88	0
	BCM	96.22	0	1.88	1.88
COI	BM	100	0	0	0
	BCM	86.95	0	0	13.04
*cox*3	BM	100	0	0	0
	BCM	78.72	0	0	21.27

Barcoding in Phaeophyceae is at its early beginning, searching for the adequate DNA regions. The work presented here is a preliminary assessment of five markers, issued from the three cellular compartments (nuclear, chloroplastic and mitochondrial), focusing on several closely related and well circumscribed species of the genus *Sargassum*. Our datasets contained several sequences per species collected from various localities in the Indo-Pacific but the limited number of sequences available was the major flaw of the study. Despite this bias, results have demonstrated the inefficiency of the nuclear ITS-2 and partial chloroplatic *RubisCO* at identifying species with confidence while they have suggested the potential of the mitochondrial markers (*mtsp*, COI and *cox*3). Additional research is needed based on larger datasets to further assess the identification success of these markers. However, we believe that no good barcoding marker will be identified for the genus *Sargassum* before one can genetically discriminate between the very closely related and morphologically distinct species of *S.* section *Sargassum*. This cannot be achieved without a good understanding of the *Sargassum* taxonomy, implying critical morphological work and significant revisions.

Acknowledgements. We are grateful to Nathalie Duong for her precious help with DNA analyses and sequence processing. The molecular work was carried at the Plateforme du Vivant at IRD in Nouméa while sequences were processed by Macrogen in Corea. We would like to thank all collectors who have helped producing our datasets: Andréfouët A., Cordona M.-E., Duffy C., Farr T., Garcia-Huidobio R., Lasne G., Mineur F., Nelson W., Ramiez M.-E., Rivera A., Sauvage T., Stanley R., Verbruggen H., and Zubia M. Thanks to the reviewers, as well as L. Legall and G. Saunders, for their constructive comments on the manuscript. This work is part of LM post-doctoral studies and CoRéUs program.

REFERENCES

BITTNER L., HALARY S., PAYRI C., CRUAUD C., REVIERS de B., LOPEZ P. & BAPTESTE E., 2010 — Some considerations for analyzing biodiversity using integrative metagenomics and gene networks. *Biology direct* 5:47.

COYER J.A., HOARAU G., OUDOT-LE SECQ M.-P., STAM W.T. & OLSEN J.L., 2006 — A mtDNA-based phylogeny of the brown algal genus *Fucus* (Heterokontophyta; Phaeophyta). *Molecular phylogenetics and evolution* 39: 209-222.

DRAISMA S.G.A., BALLESTEROS E., ROUSSEAU F. & THIBAUT T., 2010 — DNA sequence data demonstrate the polyphyly of the genus *Cystoseira* and other Sargassaceae genera (Phaeophyceae). *Journal of phycology*: *in press*.

HALL T.A., 1999 — BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/nt. *Nucleic acids symposium series* 41: 95-98.

HERBERT P.D.N., CYWINSKA A., BALL S.L. & DEWAARD J.R., 2003 — Biological identification thourgh DNA barcodes. *Proceedings of the royal society of London, B Biological sciences* 270: 313-321.

HERBERT P.D.N., PENTON E.H., BURNS J.M., JANZEN D.H. & HALLWACHS W., 2004 — Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptes fulgerator. *Proceedings of the national academy of science USA* 101: 14812-14817.

KUCERA H. & SAUNDERS G.W., 2008 — Assigning morphological variants of *Fucus* (Fucales, Phaeophyceae) in Canadian waters to recognized species using DNA barcoding. *Botany* 86: 1065-1079.

KRESS W.J., WURDACK K.J., ZIMMER E.A., WEIGT L.A. & JANZEN D.H., 2005 — Use of DNA barcodes to identify flowering plants. *PNAS* 102(23): 8369-8374.

LANE C.E., LINDSTROM A.C. & SAUNDERS G.W., 2007 — A molecular assessment of northeast Pacific *Alaria* species (Laminariales, Phaeophyceae) with reference to the utility of DNA barcoding. *Molecular phylogenetics and evolution* 44: 634-648.

LE GALL L. & SAUNDERS G.W., 2010 — DNA barcoding is apowerful tool to uncover alagal diversity: a case study of the Phyllophoraceae (Gigartinales, Rhodophyta) in Canadian Flora. *Journal of phycology* 46: 374-389.

LIU Y., YAN H.-F., CAO T. & GE X.-J., 2010 — Evaluation of 10 plant barcodes in Bryophyta (Mosses). *Journal of systematics and evolution* 48(1): 36-46.

LOHSE K., 2009 — Can mtDNA barcodes be used to delimit species? A response to Pons *et al.* (2006). *Systematic biology* 58: 439-442.

MACDEVIT D.C. & SAUNDERS G.W., 2009 — On the utility of DNA barcoding for species differentiation among brown macroalgae (Phaeophyceae) including a novel extraction protocol. *Phycological research* 57: 131-141.

MACDEVIT D.C. & SAUNDERS G.W., 2010 — A DNA barcode examination of the Laminariaceae (Phaeophyceae) in Canada reveals novel biogeographical and evolutionary insigths. *Phycologia* 49: 235-248.

MATTIO L., PAYRI C.E. & STIGER-POUVREAU V., 2008 — Taxonomic revision of *Sargassum* (Fucales, Phaeophyceae) from French Polynesia based on morphological and molecular analyses. *Journal of phycology* 44: 1541-1555.

MATTIO L., PAYRI C.E. & VERLAQUE M., 2009 — Taxonomic revision and geographic distribution of subgenus *Sargassum* (Fucales, phaeophyceae) in the western and central Pacific islands based on morphological and molecular analyses. *Journal of Phycology* 45: 1213-1227.

MATTIO L., PAYRI C.E., VERLAQUE M. & DE REVIERS B., 2010 — Taxonomic revision of *Sargassum* sect. *Acanthocarpicae* (Fucales, Phaeophyceae). *Taxon* 59(3): 896-904.

MEIER R., KWONG S., VAIDYA G., & NG P.K.L., 2006 — DNA Barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematic biology* 55: 715-728.

MEYER C.P. & PAULAY G., 2005 — DNA barcoding: Error rates based on comprehensive sampling. *Plos biology* 3(12): 2229-2238.

PHILLIPS, N., SMITH C. & MORDEN C.W., 2005 — Testing systematic concepts of *Sargassum* (Fucales, Phaeophyceae) using portion of the rbcLS operon. *Phycological research* 53: 1-10.s

PONS J., BARRACLOUGH T.G., GOMEZ-ZURITA J., CARDOSO A., DURAN D.P., HAZELL S., KAMOUN S. & SUMLIN W., 2006 — Sequence based species delineation for the DNA taxonomy of undescribed insects. *Systematic biology* 55(4): 595-609.

ROBBA L., RUSSELL S.J., BARKER G.L. & BRODIE J., 2006 — Assessing the use of the mitochondrial cox1 marker for use in DNA barcoding of red algae (Rhodophyta). American *Journal of botany* 93(8): 1101-1108.

ROHFRITSCH A., PAYRI C.E., STIGER V. & BONHOMME F., 2010 — Habitat-related allelic variation revealed by an anonymous DNA locus in reef-dwelling *Turbinaria ornata* (Fucales, Phaeophyceae). *Botanica marina* 53(2): 189-192.

SAUNDERS G.W., 2005 — Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Proceedings of the royal society biology* 360: 1879-1888.

SAUNDERS G.W., 2008 — A DNA barcode examination of the red alagel family Dumontiaceae in Canadian waters reveals substential cryptic species diversity. 1. The foliose *Dilsea-Neodilsea* complex and *Weeksia*. *Botany* 86: 773-789.

SAUNDERS G.W., 2009 — Routine DNA barcoding of Canadian Gracilariales (Rhodophyta) reveals the invasive species Gracilaria vermiculophylla in British Columbia. *Molecular ecology ressources* 9: 140-150.

SHEARER T.L. & COFFROTH M.A., 2008 — Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Molecular ecology resources* 8 : 247-255.

SHERWOOD A.R., 2009 — Phylogeograhy of *Asparagopsis taxiformis* (Bonnemaisoniales, Rhodophyta) in the Hawaiin islands: Two Mtdna markers support three separate introductions. *Phycologia* 47(1): 79-88.

STIGER V., HORIGUCHI T., YOSHIDA T., COLEMAN A.W. & MASUDA M., 2003 — Phylogenetic relationships inferred from ITS-2 nrDNA comparisons within the genus *Sargassum* (Fucales, Phaeophyceae) from the Pacific basin, with an emphasis on the taxonomic subdivision of the genus. *Phycological research* 51: 1-10.

THOMPSON J.D., HIGGINS D.G. & GIBSON T.J., 1994 — CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* 22 (22): 4673-4680.