Subrata Trivedi · Abid Ali Ansari
Sankar K. Ghosh · Hasibur Rehman
Editors

DNA Barcoding in Marine Perspectives

Assessment and Conservation of Biodiversity

A Search for a Single DNA Barcode for Seagrasses of the World

Barnabas H. Daru and Kowiyou Yessoufou

Abstract It has recently been predicted that 91 % of marine species diversity is still unknown. Given that the future of marine habitats is threatened by anthropogenic activities and climate change, there is a pressing need to accelerate the documentation of marine biodiversity. The traditional morphological biodiversity screening could be aided by molecular approach such as DNA barcoding. In this study, we search for single DNA marker that could be used as DNA barcode for all seagrasses, irrespective of the lineages and the geographical locations. We found that the nuclear *phyB* followed by the plastid *matK* emerged as the best candidates. Although both markers have their own strengths and limitations, we suggest they could be prioritised in seagrass biodiversity assessment pending future improvements.

Keywords DNA barcoding · *phyB* · *matK* · Marine biodiversity · Cymodoceaceae · Ruppiaceae

1 Introduction

How many species are there and how do we recognize them? A recent prediction of species richness estimated that global ecosystems harbor 8.7 million species, including 2.2 million marine species (Mora et al. 2011). Of this impressive diversity, 86 % of terrestrial and 91 % of marine species are currently unknown (Mora et al. 2011), raising an urgent need for accelerating the process and increasing our commitment for biodiversity assessment. This need for biodiversity

B.H. Daru (✉)
Department of Organismic and Evolutionary Biology,
Harvard University, 22 Divinity Avenue, Cambridge, MA 02138, USA
e-mail: barnabas_daru@fas.harvard.edu; darunabas@gmail.com

K. Yessoufou
Department of Environmental Science, University of South Africa,
Florida Campus, Florida 1710, South Africa

© Springer International Publishing Switzerland 2016

S. Trivedi et al. (eds.), *DNA Barcoding in Marine Perspectives*,
DOI 10.1007/978-3-319-41840-7_19

barnabas_daru@fas.harvard.edu

assessment is even more pressing given the current extinction crisis driven by an unprecedented rate of species loss estimated to be 1,000–10,000 times greater than that recorded in the past (Millennium Ecosystem Assessment 2005; Barnosky et al. 2011). Traditionally, we assess species diversity using morphological features, a long-standing approach that can be very tedious and questionable owing to the potential subjectivity attached to it since the relevance of morphological features to be used is to the discretion of the taxonomist. Although, this approach is irreplaceable, it has its own limitations (e.g. see Packer et al. 2009). For example it took >250 years to describe less than a 1/4 of the world species (see review in Radulovici et al. 2010). Given the unprecedented rate of species extinction, we cannot afford to wait again for more than two centuries to know a tiny proportion of earth's diversity; a quicker and integrative approach (that combines perhaps morphology and molecular data) that can help accelerate biodiversity assessment, the discovery of new species including cryptic species become a matter of urgency.

DNA barcoding has been proposed as an important molecular tool that provides complementary information overlooked in morphology-based biodiversity assessment (Hebert et al. 2003). It is a technique that uses short sequences of DNA to either confirm species identity or assign unknown biological materials (e.g. plants, animals and fungi at any stage of life cycle) to corresponding species or higher taxonomic groups or reveal cryptic species (morphologically similar but genetically distinct species). The technique has witnessed a great application for assessing biodiversity (Smith et al. 2005; Papadopoulou et al. 2015; van der Bank and Greenfield 2015; see also Trivedi et al. 2016 for a comprehensive review). However, more attention has been given to terrestrial ecosystems (see Fig. 1 in Radulovici et al. 2010), although oceans cover more than 70 % of our planet and are potentially as species-rich as terrestrial ecosystems. For example, of the currently known 35 animal phyla, 14 are marine endemics (Briggs 1994; Gray 1997). In general, marine ecosystems provide unique ecosystem services to humanity: foods (e.g. fish, prawns, etc.), biotechnological and non-living resources, as well as indicator of environmental health and ecosystem functioning (food webs), erosion control and carbon sinks (e.g. mangroves) etc. Given major anthropogenic factors that threaten marine ecosystems (e.g. habitat loss, overharvesting, global warming, pollution, invasive species, etc.), there is a need to know the ecosystem engineers that ensure the provision of goods and services for humanity in oceans.

The application of DNA barcoding to assess marine biodiversity is increasingly generating renewed interest (see reviews in Radulovici et al. 2010; Trivedi et al. 2016). However, of the few studies that show interests into marine biodiversity (compared to terrestrial biodiversity), most have focused on marine animals (Radulovici et al. 2010; Trivedi et al. 2016, see also Marine Barcode of Life MarBOL, www.marinebarcoding.org; accessed March 20, 2015), resulting potentially in comparatively poorer knowledge of marine plant diversity. In the present study, we focus on seagrasses, an ecologically important plant taxonomic group in marine ecosystems.

Seagrasses belong to the monocot order Alismatales comprising 72 species represented in 13 genera and five families (Les et al. 1997; den Hartog and Kuo

barnabas_daru@fas.harvard.edu

2006). They have a wide range of vegetative and floral diversity (Fig. 1), and are widely distributed along all marine coastlines worldwide from intertidal to subtidal depths (den Hartog 1970; Green and Short 2003), providing key ecosystem services such as primary productivity, nutrient cycling etc. (Hemminga and Duarte 2000; Duarte 2002; Les et al. 2002; Orth et al. 2006; McGlathery et al. 2007). They are also well-known in traditional medicine for the valuable secondary compounds (e.g. phenolic acids, rosmarinic acid and zosteric acid) widely used as an antioxidant and effective antifouling agent (Trivedi et al. 2016). Nonetheless, these marine plants are undergoing a rapid decline in both species richness and geographic cover: we are losing seagrasses at a rate of 110 km^2 per year (Waycott et al. 2009), prompting the need for documenting the diversity of seagrasses before we lose what we do not yet know about them. However, species boundaries among the lineages are still not resolved (Tomlinson and Posluszny 2001; den Hartogand Kuo 2006). The real challenge lies with the fact that they are submerged plants with high prevalence of cryptic species (Briggs 2003; Trivedi et al. 2016). Because they occur submerged in marine water, they may have acquired adaptations such as reduced morphology in both vegetative and floral structures, making morphological identification difficult. In fact, seagrass species in the field or archived in herbaria are often devoid of

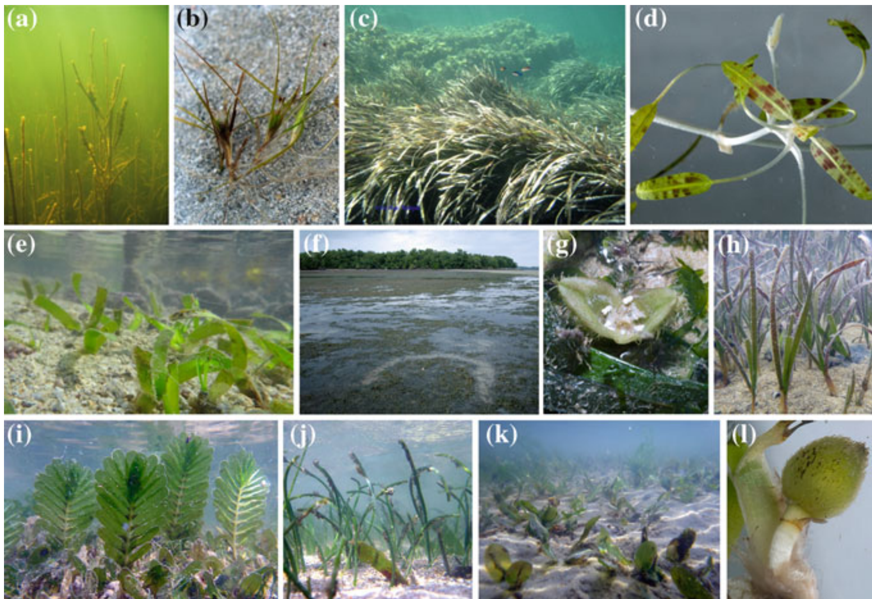

Fig. 1 Representatives of seagrass species showing variation and diversity in the group. **a** *Cymodocea nodosa*; **b** *Lepilaena australis*; **c** *Posidonia oceanica*; **d** *Halophila beccarii*; **e** *Thalassia hemprichii*; **f** Seagrass meadow—a feeding trail for sea cow (*Dugong dugon*) in ChekJawa, Singapore; **g** *Enhalus acoroides* (male flowers); **h** *Cymodocea rotundata*; **i** *Halophila spinulosa*; **j** *Syringodium isoetifolium*; **k** *Halophila ovalis*; **l** shoot of *Thalassia hemprichii*.—Photographs: **a, b** courtesy Y. Ito; **c** J.Á. Rodríguez; **d–l** R. Tan

barnabas_daru@fas.harvard.edu

diagnostic flowers (Trivedi et al. 2016; Personal observations). This calls for an urgent need for a fast, reliable, and cost-efficient technique for recognition and identification of seagrasses especially by non-experts (Cocheret de la Morinière et al. 2003).

Lucas et al. (2012) showed the importance of DNA barcoding in delimiting species boundaries for seagrasses in India. For 14 species examined using *matK* and *rbcL*, sequence divergence for discriminating species is higher for *matK* than *rbcL*. Another study showed the success of DNA barcoding in identifying six seagrass species in the gut of rabbit fish *Siganus fuscescens* in Moreton Bay, Australia (Chelsky Budarf et al. 2011). Other studies have focused on single clades using different markers e.g. *trnK* and *rbcL* for *Zostera* (Les et al. 2002), ITS for *Halophila* (Waycott et al. 2002), 5.8S rDNA and ITS2 for *Halophila* (Uchimura et al. 2008), ITS1, *matK*, *rbcL*, *psbA*-*trnH* for Zosteraceae (Coyer et al. 2013). As indicated in these studies, they focus either on a single geographic location or a single genus of seagrasses, leaving a knowledge gap on whether a single DNA barcode could help screen seagrass diversity irrespective of the geographic locations or genera.

In this study, we explored this possibility by first assessing the potential of nine markers to discriminate seagrass species of the world, and second, assess the efficacy of barcodes across major seagrass clades.

2 Materials and Methods

2.1 Taxon Sampling

We retrieved from GenBank/EBI all available sequences of seagrasses for nine molecular markers, *atp1*, *cob*, ITS, *matK*, NAD5, *phyB*, *rbcL*, *rpoB* and *trnH-psbA*. These sequences are from 44 species belonging to all the five seagrass families Cymodoceaceae, Hydrocharitaceae, Posidoniaceae, Ruppiaceae, and Zosteraceae (Appendix A). Our sampling comprised 95 specimens (see Appendix A). The sequences were aligned using SeaViewv.4 (Gouy et al. 2010) and manually adjusted using Mesquitev.2.5 (Maddison and Maddison 2008).

2.2 Barcoding Analyses

First, we evaluated the performance of the various plant DNA regions in discriminating seagrass species by applying three criteria commonly used in DNA barcoding literature: the barcode gap of Meyer and Paulay (2005), the level of sequence divergence and the discriminatory power. Barcode gap was assessed by comparing intra-specific variation (i.e. the amount of genetic variation within species) to inter-specific variation (between species). A good barcode should

barnabas_daru@fas.harvard.edu

exhibit a significant gap, meaning that sequence divergence within species should be significantly lower than between species. Statistical significance between intra- and inter-specific variation was assessed using Wilcoxon test in R (R Core Team 2013). In addition, we calculated the distribution of range, mean and standard deviation of both intra- and inter-specific distances.

Second, we identified the best DNA barcode using two distance-based methods; near neighbour and best close match (Meier et al. 2006) using the functions *near Neighbour* and *best Close Match* respectively, implemented in the R package SPIDER (Brown et al. 2012). This was done by combining all sequences. Prior to the evaluation of discriminatory power of each barcode candidate, we determined the distance threshold i.e. the optimised genetic distance for species delimitation, given that the 1 % threshold suggested by BOLD does not hold for every organism (Meyer and Paulay 2005). This distance cut-off was identified using the function *local Minima* implemented in SPIDER, which evaluates the transition between intra- and inter-specific distances (Brown et al. 2012). The optimised threshold was used especially in best close match method.

Lastly, given the possibility that the performance of marker could vary between taxonomic levels (Gere et al. 2013), we further assessed the performance of the core barcode within two families, Cymodoceaceae and Ruppiaceae; the other seagrass families were not considered here due to lack of sufficient DNA sequences.

3 Results and Discussion

Information on aligned sequence length, number of species, mean number of substitutions per nucleotide for all DNA regions considered singly or in combination, the range and means of intra- and interspecific distances are summarized in Table 1. The mean interspecific distance for the single and combined regions are lower than 1 %, ranging from 0.011 in *cob* to 0.77 in ITS. The mean intraspecific variation for each and combined DNA regions was also low ranging from 0.00008 in *rbcL* + *matK* + *cob* to 0.024 in ITS (Table 1).

We show that the ranges and mean intraspecific distances for all markers when considered singly or in combination with the core barcodes (*matK* + *rbcL*), are significantly lower than interspecific distances (Wilcoxon test, $p < 0.01$; Fig. 2), suggesting the presence of barcode gap. Comparison of the proportion of sequences with barcode gap showed that *trnH-psbA* (88 %) followed by *rpoB* (78 %) had the highest proportion with the lowest proportion found in *cob* (0 %), *rbcL* + *matK* + *atp1* (0 %), and *rbcL* + *matK* + *cob* (0 %) (Table 2).

We calculated the optimised genetic distance (threshold distance) that is appropriate for species delimitation. The thresholds range from 0.0014 for *rpoB* to 0.29 for ITS. Using these cut-offs, for the best close match method, *phyB* exhibited the highest species identification rate of 71 % for single regions, which improved to

barnabas_daru@fas.harvard.edu

Table 1 Summary statistics of all DNA markers used to delimit seagrasses of the world

Gene region	No. of seq (# spp)	Seq length	K	Range (inter)	Mean inter (±SD)	Range (intra)	Mean intra (±SD)	Threshold (%)
atp1	25 (19)	1,062	0.011	0–0.083	0.022 ± 0.029	0–0.00096	0.00019 ± 0.0004	1.88
cob	25 (18)	1,031	0.17	0–0.044	0.011 ± 0.009	0–0.00097	0.00018 ± 0.00039	2.56
ITS	52 (20)	720	0.053	0–0.12	0.77 ± 0.45	0–0.08	0.024 ± 0.026	29.25
matK	186 (42)	550	0.057	0–0.48	0.12 ± 0.073	0–0.33	0.0055 ± 0.01	4.11
NAD5	22 (16)	1,121	0.0076	0–0.055	0.015 ± 0.013	0–0.017	0.0044 ± 0.0062	1.025
phyB	69 (10)	1,050	0.042	0–0.22	0.12 ± 0.078	0–0.039	0.018 ± 0.016	8.17
rbcL	152 (43)	771	0.28	0–0.21	0.048 ± 0.027	0–0.10	0.0036 ± 0.01	2.077
rpoB	99 (10)	517	0.012	0–0.15	0.026 ± 0.044	0–0.004	0.0024 ± 0.0019	0.14
trnH-psbA	88 (25)	485	0.045	0–0.89	0.44 ± 0.24	0–0.28	0.046 ± 0.041	17.44
rbcL + *matK*	113 (46)	1,321	0.15	0–0.44	0.093 ± 0.06	0–0.19	0.0045 ± 0.013	2.88
rbcL + *matK* + *atp1*	24 (18)	2,383	0.033	0–0.13	0.05 ± 0.034	0–0.00043	0.0001 ± 0.00018	5.44
rbcL + *matK* + *cob*	25 (18)	2,352	0.011	0.00044–0.16	0.051 ± 0.042	0–0.00043	0.00008 ± 0.0002	4.78
rbcL + *matK* + ITS	52 (20)	2,041	0.035	0–0.68	0.26 ± 0.13	0–0.11	0.012 ± 0.012	9.97
rbcL + *matK* + NAD5	22 (16)	2,442	0.0098	0.0004–0.087	0.038 ± 0.023	0–0.0083	0.0017 ± 0.003	2.24
rbcL + *matK* + *phyB*	44 (10)	2,371	0.023	0.0008–0.16	0.074 ± 0.055	0–0.023	0.0096 ± 0.0083	5.63
rbcL + *matK* + *rpoB*	39 (10)	1,838	0.020	0.0006–0.15	0.043 ± 0.053	0–0.017	0.0033 ± 0.0026	2.13
rbcL + *matK* + *trnH-psbA*	59 (22)	1,809	0.038	0–0.36	0.16 ± 0.09	0–0.034	0.012 ± 0.0079	6.58
Combined[a]	27 (18)	7,968	0.0094	0.0003–0.33	0.075 ± 0.068	0–0.0092	0.0029 ± 0.0034	17.21

[a]Combined, *atp1* + *cob* + ITS + *matK* + NAD5 + *phyB* + *rbcL* + *rpoB* + *trnH-psbA*

barnabas_daru@fas.harvard.edu

Table 2 Percentage barcode gap using best close match method (Meier et al. 2006)

Gene region	Number of sequences without gap	Proportion of sequences with gap (%)
atp1	15	40
cob	25	0
ITS	23	56
matK	61	67
NAD5	6	73
phyB	18	74
rbcL	89	41
rpoB	22	78
trnH-psbA	11	88
rbcL + *matK*	84	26
rbcL + *matK* + *atp1*	24	0
rbcL + *matK* + *cob*	25	0
rbcL + *matK* + ITS	22	58
rbcL + *matK* + NAD5	20	9
rbcL + *matK* + *phyB*	24	45
rbcL + *matK* + *rpoB*	18	54
rbcL + *matK* + *trnH-psbA*	27	54
Combined	23	15

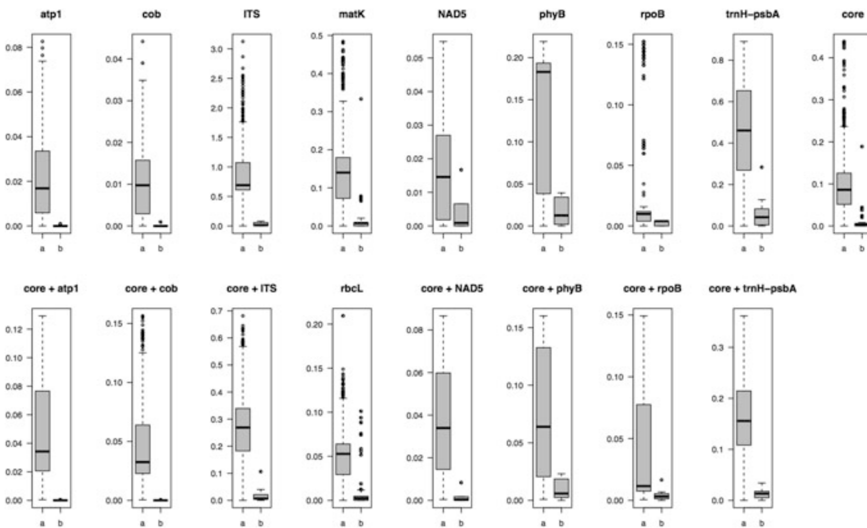

Fig. 2 Comparison of the distribution range of inter- and intra-specific distances using boxplot. The bottom and top of boxes show the first and third quartiles respectively, the median is indicated by the horizontal line, the range of the data by the vertical dashed line and outliers (points outside 1.5 times the interquartile range) by circles. a = interspecific, b = intraspecific

barnabas_daru@fas.harvard.edu

86 % when combined with the core barcodes (i.e. for *phyB* + *rbcL* + *matK*). This was followed by *matK* (52 % for single regions) and 77 % for *rbcL* + *matK* + *rpoB*. The core barcodes alone yielded an identification success of 62 % (Table 3). Similarly, for the near neighbour method, *phyB* followed by *rpoB* yielded the highest identification rates for the single regions (86 % and 85 %, respectively), which improved markedly when combined with the core barcodes (*phyB* + *matK* + *rbcL* = 91 % and *rpoB* + *matK* + *rbcL* = 82 %).

Lastly, at the family level, we found that the combination of *phyB* and the core barcodes (*matK* + *rbcL*) improve species discrimination in Ruppiaceae from 86 to 88 % (*phyB* alone vs *matK* + *rbcL* + *phyB,* respectively), and 78 to 80 % in Cymodoceaceae for *phyB* alone vs *matK* + *rbcL* + *phyB,* respectively (Table 4).

Several criteria have been defined for the identification of the best DNA barcode candidate (Hebert et al. 2004; Kress and Erickson 2007; Lahaye et al. 2008; CBOL Plant Working Group 2009). Firstly, it must provide maximal discrimination between species, and this ability to discriminate depends on the existence of a barcode gap (Meyer and Paulay 2005). All the nine markers tested exhibit significant barcode gap, indicating that they are all good candidates for DNA barcode of seagrasses. To identify the best candidate, we tested their discriminatory power using two distance methods, the near neighbour and best close match methods. In both methods, *phyB* followed by *matK* yielded the best identification rates, thus

Table 3 Identification efficacy of potential DNA barcodes using distance based methods

Gene region	Near neighbor		Best close match			
	True (%)	False (%)	Ambiguous (%)	Correct (%)	Incorrect (%)	No ID (%)
atp1	36	64	16	20	56	8
cob	28	72	40	0	48	12
ITS	65	35	27	35	15	23
matK	81	19	33	52	12	3
NAD5	14	86	36	5	45	14
phyB	86	14	13	71	9	7
rbcL	66	34	51	28	16	5
rpoB	85	15	74	21	2	3
trnH-psbA	52	48	34	33	20	13
rbcL + *matK*	65	35	9	62	26	3
rbcL + *matK* + *atp1*	38	62	0	38	29	33
rbcL + *matK* + *cob*	44	56	0	44	24	32
rbcL + *matK* + ITS	62	38	0	60	19	21
rbcL + *matK* + NAD5	50	50	0	50	50	0
rbcL + *matK* + *phyB*	91	9	0	86	5	9
rbcL + *matK* + *rpoB*	82	18	0	77	10	13
rbcL + *matK* + *trnH-psbA*	66	34	0	63	25	12
Combined	56	44	0	56	18	26

barnabas_daru@fas.harvard.edu

Table 4 Comparisons of the core barcodes (*matK* + *rbcL*) and best barcode within seagrass families, Ruppiaceae and Cymodoceaceae

Family	DNA regions	No of seq	Mean inter (±SD)	Threshold (%)	Best close match			
					Ambiguous (%)	Correct (%)	Incorrect (%)	No ID (%)
Ruppiaceae	core	35	0.0094 ± 0.0037	0.37	0	86	14	0
	core + *phyB*	34	0.025 ± 0.018	1.18	0	88	3	9
Cymodoceaceae	core	27	0.03 ± 0.017	0.84	0	78	22	0
	core + *phyB*	10	0.032 ± 0.031	5.19	0	80	10	10

barnabas_daru@fas.harvard.edu

making them the priority markers for further analyses. We then assessed their performance in combination with the core barcodes. The combination core + *phyB* emerged as the best candidate in both near neighbor and best close match. The core barcodes alone perform poorly and this has already been reported in many cases for different plant taxonomic groups (Hollingsworth et al. 2009; Pettengill and Neel 2010; Roy et al. 2010; Wang et al. 2010; Clement and Donoghue 2012).

Phytochrome B (phyB) is a low copy nuclear DNA marker active in light-grown plants, and plays a key role in regulating circadian rhythm in plants (Somers et al. 1998). Previous phylogenetic studies have shown its utility in resolving relationships in angiosperms (Mathews et al. 2000; Simmons et al. 2001), and in detecting polyploids and hybrids in some seagrass lineages (Ito et al. 2010, 2013). Given that hybridization is very common in aquatic monocots, including seagrasses (Les and Philbrick 1993), our study lends support to the utility of *phyB* as a barcode candidate for identifying a complex taxonomic group like seagrasses. In addition to *phyB, matK* emerged second best in species identification. Although *matK* region has been initially proposed as best plant barcode (Lahaye et al. 2008), some studies have identified potential pitfalls against its suitability (e.g. lack of universal primers; Chase et al. 2007). However more recent studies revealed that such drawback was unjustified for seagrasses. Overall, the nuclear *phyB* and the plastid *matK* are single best candidates that can be used to assess or screen the diversity of seagrasses, but each of both has its own strength and limitations.

Molecular and morphological data do not always concord with regard to species delimitation and this has also been reported for seagrasses (e.g. see Les et al. 2002; Kato et al. 2003; Tanaka et al. 2003 versus den Hartog and Kuo 2006 for the *Zostera capricorni* complex in Australia/New Zealand). Potential reasons for this include mechanisms such as different ecotypes for a single species, ongoing speciation and incomplete lineage sorting or hybridisation through introgression (Coyer et al. 2008). These mechanisms and introgression in particular, obscure taxonomic delimitation caused by deep intraspecific splits in gene trees, resulting in species appearing as paraphyletic or polyphyletic (Pentinsaari et al. 2014). However, introgression concerns less frequently nuclear markers compared to chloroplast (Rieseberg et al. 1991) and mitochondrial markers (Pentinsaari et al. 2014), giving an advantage for nuclear gene (here *phyB*). Also, introgression is more likely to occur between closely related species (Rieseberg et al. 1991; Coyer et al. 2008), suggesting that the differentiation between higher taxa (e.g. genera or families) than species is likely to be more efficient (see Lucas et al. 2012). Our evaluation of the performance of the core barcode at family level for Ruppiaceae and Cymodoceaceae confirms this with a discriminatory power of 86 % and 78 %, respectively. The core barcode performs poorly on all dataset but performs better when its use is limited to diversity within a family. The difference in the barcoding performance between the two could reflect differences in evolutionary history, incomplete lineage sorting, different ecological types, or hybridisation (Coyer et al.

barnabas_daru@fas.harvard.edu

2008, 2013). Ruppiaceae is a monogeneric family widely distributed in brackish waters along tropical and temperate coastlines of the world (Verhoeven 1979), characterised by species with highly similar morphology, and high level of introgression due to polyploidisation and hybridisation (Ito et al. 2010). Similarly, the family Cymodoceaceae is another seagrass lineage with reportedly high level of hybridisation (Ito and Tanaka 2011).

4 Conclusion

Seagrasses are submerged angiosperms that provide important ecosystem services such as nutrient recycling, high primary productivity, and sources of medicinal molecules. However, we are losing them at an alarming rate, in term of diversity and geographical ranges (Waycott et al. 2009; Daru and le Roux 2016), prompting the need for accelerating the screening of seagrass diversity as part of the global campaign for documenting biodiversity. In this need, molecular techniques could complement traditional taxonomic approach, and efforts to identify appropriate marker as DNA barcodes for seagrasses has attracted much attention (Les et al. 2002; Waycott et al. 2002; Uchimura et al. 2008; Lucas et al. 2012). The search for single marker for the entire seagrasses is more convenient as it is cheaper and less time-consuming than the search for multiple markers for each seagrass lineage. Pending future studies with additional sampling and DNA markers, we proposed that the nuclear *phyB* and, secondarily the plastid *matK* as suitable single DNA barcode for genetic identification of seagrass species.

Appendix

See (Table A.1).

barnabas_daru@fas.harvard.edu

Table A.1 Voucher information and GenBank/EBI accession numbers for seagrasses used in this study. APG, Angiosperm phylogeny group

APG III Family	Taxon	GenBank accession numbers								
		atp1	*cob*	*ITS*	*matK*	*NAD5*	*phyB*	*rbcL*	*rpoB*	*trnH-psbA*
Cymodoceaceae	*Amphibolis antarctica*	KF488552	KF488541	–	KF488499	KF488529	–	KF488485	–	–
		–	–	–	–	–	–	U80686	–	–
	Amphibolis griffithii	HQ317985	HQ317978	–	KF488500	HQ267476	–	HQ901574	–	–
	Cymodocea nodosa	DQ859094	DQ859130	–	KF488501	HQ267481	–	KF488486	–	–
	Cymodocea rotundata	–	KF488542	AF102272	KF488502	KF488530	–	KF488487	–	–
		JQ031762	JQ031761	–	JN225358	KF488531	–	U80688	–	–
		KF488553	KF488543	–	JQ031760	KF488532	–	JN225334	–	JN225311
		KF488554	KF488544	–	KF488503	KF488533	–	JQ031763	–	FJ648790
		KF488555	KF488545	–	KF488504	–	–	KF488488	–	–
		–	–	–	KF488505	–	–	KF488489	–	–
	Cymodocea serrulata	AY277801	DQ859131	–	JN225359	KF488534	–	JN225335	–	JN225310
		DQ859095	KF488546	–	KF488506	KF488535	–	KF488491	–	–
		KF488556	KF488547	–	KF488507	–	–	KF488492	–	–
	Halodule pinifolia	KF488557	KF488548	–	JN225368	KF488536	AB571211	AB571198	–	AB571183
		KF488558	KF488549	–	JN225369	KF488537	AB571212	AB571199	–	AB571184
		–	–	–	KF488508	–	AB571213	AB571200	–	AB571185
		–	–	–	KF488509	–	AB571214	AB571201	–	AB571186
	Halodule uninervis	KF488559	KF488550	–	JN225370	KF488538	AB571216	AB571206	–	AB571191
		–	–	–	JN225371	–	AB571219	AB571207	–	AB571192
		–	–	–	KF488510	–	AB571220	AB571208	–	AB571193
	Halodule wrightii	–	–	–	JN225379	–	AB571224	AB571196	–	JN225331
		–	–	–	–	–	AB571225	AB571197	–	–
	Syringodium filiforme	DQ859116	DQ859154	–	KF488511	KF488539	–	KF488496	–	–
		–	–	–	–	–	–	U03727	–	–

(continued)

barnabas_daru@fas.harvard.edu

Table A.1 (continued)

APG III Family	Taxon	GenBank accession numbers								
		atp1	*cob*	*ITS*	*matK*	*NAD5*	*phyB*	*rbcL*	*rpoB*	*trnH-psbA*
	Syringodium isoetifolium	HQ31798		–		HQ26750		AB507901		FJ648793
		–	–	–	JN225372	–	–	JN225342	–	JN225314
		–	–	–	KF488512	–	–	KF488497	–	–
	Thalassodendron ciliatum	KF48856	–	–	KF488513	KF48854	–	KF488498	–	–
	Thalassodendron pachyrhizum	–	–	–	–	–	–	U80692	–	–
Hydrocharitaceae	*Enhalus acoroides*	JF975469	JF975452	AY870347	AB002569	–	–	AB004889	JF975516	–
	Halophila ovalis	JF975473	JF975453	AB243970	AB002570	–	–	AB004890	JF975517	–
	Thalassia hemprichii	JF975480	JF975463	–	AB002577	–	–	AB004897	JF975528	–
Posidoniaceae	*Posidonia australis*	DQ859111	DQ859148	GQ927720	KF488514	–	–	HQ901573	–	–
	Posidonia kirkmanii	–	–	GQ927724	GQ927728	–	–	–	–	–
	Posidonia oceanica	DQ859112	DQ859149	GQ927725	GQ927729	–	–	JQ995767	–	JX028527
		–	–	JQ937106	JQ990941	–	–	U80719	–	JX028527
		–	–	–	KF488515	–	–	–	–	–
Ruppiaceae	*Ruppia cirrhosa*	DQ859114	DQ859151	AB728740	AB728682	HQ267503	AB728706	AB728688	AB728694	AB728724
		–	–	AB728748	KC505607	–	AB728707	DQ859175	KC505613	AB728730
		–	–	AB728749	KF488516	–	–	JN113275	–	JN113267
	Ruppia maritima	–	–	AB728734	AB507905	HQ317975	AB508028	AB507865	AB507945	AB728718
		–	–	AB728735	AB507906	–	AB508031	AB507866	AB507946	AB728720
		–	–	AB728736	AB507907	–	AB508034	AB507867	AB507947	AB728721
		–	–	AB728737	AB507908	–	AB508035	AB507868	AB534790	AB728722
		–	–	AB728738	AB507909	–	AB508037	AB507869	AB534791	AB728723
	Ruppia megacarpa	–	–	JQ034337	AB507929	–	AB508065	AB507889	AB507969	KC505612
		–	–	–	AB507930	–	AB508066	AB507890	AB507970	–
		–	–	–	AB507931	–	AB508067	AB507891	AB507971	–

(continued)

barnabas_daru@fas.harvard.edu

Table A.1 (continued)

APG III Family	Taxon	GenBank accession numbers								
		atp1	*cob*	*ITS*	*matK*	*NAD5*	*phyB*	*rbcL*	*rpoB*	*trnH-psbA*
	Ruppia occidentalis	–	–	–		–		AB507894		–
		–	–	–	AB507935	–	AB508071	AB507895	AB507975	–
	Ruppia polycarpa	–	–	–	AB507938	–	AB508074	AB507898	AB507978	–
	Ruppia tuberosa	–	–	–	AB507939	–	AB508075	AB507899	AB507979	–
		–	–	–	AB507940	–	AB508076	AB507900	AB507980	–
Zosteraceae	*Phyllospadix iwatensis*	–	–	JQ766110	AB096172	–	–	–	–	JX028522
		–	–	–	JQ990933	–	–	–	–	JX028523
	Phyllospadix japonicus	–	–	–	JQ990932	–	–	JQ995760	–	–
	Phyllospadix scouleri	DQ859110	DQ859147	–	–	HQ267497	–	DQ859172	–	–
	Phyllospadix torreyi	–	–	AY077985	EF198333	–	–	JQ995764	–	JX028524
		–	–	EF198346	JQ990934	–	–	U80731	–	–
	Zostera asiatica	–	–	EF198347	AB096161	–	–	AB125352	–	JX028519
		–	–	–	EF198338	–	–	JQ995761	–	–
		–	–	–	JQ990931	–	–	–	–	–
	Zostera caespitosa	–	–	–	AB096162	–	–	AB125351	–	–
		–	–	–	JQ990937	–	–	–	–	–
	Zostera capensis	–	–	–	AB096165	–	–	AM235166	–	JX028515
		–	–	–	JQ990930	–	–	–	–	–
	Zostera capricorni	–	–	AY077995	AB096167	DQ406964	–	AY077963	–	JX028513
	Zostera caulescens	–	–	–	AB096163	–	–	AB125350	–	–
		–	–	–	JQ990936	–	–	–	–	–
	Zostera japonica	–	–	EF198356	AB096166	–	–	AB125353	–	JX028514
		–	–	EF198357	EF198335	–	–	AY077964	–	JX028516
		–	–	EF198358	EF198336	–	–	JQ995758	–	JX028517

(continued)

barnabas_daru@fas.harvard.edu

Table A.1 (continued)

APG III Family	Taxon	GenBank accession numbers								
		atp1	*cob*	*ITS*	*matK*	*NAD5*	*phyB*	*rbcL*	*rpoB*	*trnH-psbA*
		–	–	–	EF198337	–	–	–	–	–
		–	–	–	JQ990922	–	–	–	–	–
		–	–	–	JQ990923	–	–	–	–	–
	Zostera marina	DQ859121	DQ859160	AF102274	AB096164	HQ267511	–	AB125348	–	DQ786516
		–	–	AY077986	EF198339	HQ317970	–	AB125349	–	JN225326
		–	–	EF198349	EF198341	–	–	JN225352	–	JN225327
		–	–	EF198350	EF198342	–	–	JN225353	–	JN225328
	Zostera minima	–	–	–	AJ581456	–	–	–	–	–
	Zostera mucronata	–	–	–	AB096168	–	–	U80732	–	–
		–	–	AY077993	JQ990938	–	–	–	–	–
	Zostera muelleri	–	–	AY077997	AB096169	–	–	JQ995757	–	GU906231
		–	–	AY077998	JQ990921	–	–	–	–	GU906232
	Zostera nigricaulis	–	–	–	JQ990919	–	–	JQ995756	–	JX028510
	Zostera noltii	–	–	JQ677022	JN894021	–	–	JN890769	–	JX028518
		–	–	JQ677023	JN894022	–	–	JN890770	–	JN225329
		–	–	JQ677024	JQ990924	–	–	JN225350	–	JN225330
		–	–	AF102275	AB096170	–	–	JN225351	–	–
		–	–	AY077992	EF198334	–	–	U80733	–	–
	Zostera novazelandica	–	–	–	AB096173	–	–	–	–	–
	Zostera pacifica	–	–	EF198348	EF198340	–	–	–	–	JX028520
		–	–	–	JQ990929	–	–	–	–	–
	Zostera polychlamys	–	–	–	JQ990920	–	–	JQ995759	–	JX028511
	Zostera tasmanica	HQ317987	HQ317980	AY077987	AB096171	HQ267487	–	U80730	–	–

barnabas_daru@fas.harvard.edu

References

Barnosky AD, Matzke N, Tomiya S, Wogan GOU, Swartz B, Quental TB et al (2011) Has the Earth's sixth mass extinction already arrived? Nature 471:51–57

Briggs JC (2003) Marine centres of origin as evolutionary engines. J Biogeogr 30:1–18

Briggs JC (1994) Species-diversity—land and sea compared. Syst Biol 43:130–135

Brown SDJ, Collins RA, Boyer S, Lefort M-C, Malumbres-Olarte J, Vink CJ et al (2012) Spider: An R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. Mol Ecol Resour 12:562–565

CBOL Plant Working Group (2009) A DNA barcode for land plants. Proc Natl Acad Sci USA 106:12794–12797

Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madriñán S, Petersen G et al (2007) A proposal for a standardised protocol to barcode all land plants. Taxon 56:295–299

Chelsky Budarf A, Burfeind DD, Loh WKW, Tibbetts IR (2011) Identification of seagrasses in the gut of a marine herbivorous fish using DNA barcoding and visual inspection techniques. J Fish Biol 79:112–121

Clement WL, Donoghue MJ (2012) Barcoding success as a function of phylogenetic relatedness in *Viburnum*, a clade of woody angiosperms. BMC Evol Biol 12:73

Cocheret de la Morinière E, Pollux BJA, Nagelkerken I, Hemminga MA, Huiskes AHL, van der Velde G (2003) Ontogenetic dietary changes of coral reef fishes in the mangrove-seagrass-reef continuum: stable isotopes and gut-content analysis. Mar Ecol Prog Ser 246:279–289

Coyer JA, Hoarau G, Kuo J, Tronholm A, Veldsink J, Olsen JL (2013) Phylogeny and temporal divergence of the seagrass family Zosteraceae using one nuclear and three chloroplast loci. Syst Biodivers 11:271–291

Coyer JA, Miller KA, Engle JM, Veldsink JH, Stam WT, Olsen JL (2008) Eelgrass meadows in the California Channel Islands and adjacent coast reveal a mosaic of two species, evidence for introgression and variable clonality. Ann Bot 101:73–87

Daru BH, le Roux PC (2016) Marine protected areas are insufficient to conserve global marine plant diversity. Glob Ecol Biogeogr 25:324–334

den Hartog C, Kuo J (2006) Taxonomy and biogeography of seagrasses. In: Larkum AWD, Orth RJ, Duarte CM (eds) Seagrasses: biology, ecology and conservation. Springer, Dordrecht, pp 1–23

den Hartog C (1970) Seagrasses of the world. North Holland, Amsterdam

Duarte CM (2002) The future of seagrass meadows. Environ Conserv 29:192–206

Gere J, Yessoufou K, Daru BH, Mankga LT, Maurin O, van der Bank M (2013) Incorporating *trnH-psbA* to the core DNA barcodes improves significantly species discrimination within southern African Combretaceae. ZooKeys 365:127–147

Gouy M, Guindon S, Gascuel O (2010) SeaView Version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol Biol Evol 27:221–224

Gray JS (1997) Marine biodiversity: patterns, threats and conservation needs. Biodivers Conserv 6:153–175

Green EP, Short FT (eds) (2003) World atlas of seagrasses. University of California Press, Los Angeles

Hebert PD, Cywinska A, Ball SL, de Waard JR (2003) Biological identifications through DNA barcodes. Proc Biol Sci 270:313–321

Hebert PD, Stoeckle MY, Zemlak TS, Franci CM (2004) Identification of birds through DNA barcodes. PLoS Biol 2:e312

Hemminga MA, Duarte CM (2000) Seagrass ecology: an introduction. Cambridge University Press, Cambridge

Hollingsworth ML, Clark A, Forrest LL, Richardson J, Pennington RT, Long DG et al (2009) Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. Mol Ecol Resour 9:439–457

barnabas_daru@fas.harvard.edu

Ito Y, Ohi-Toma T, Murata J, Tanaka N (2013) Comprehensive phylogenetic analyses of the *Ruppiamaritima* complex focusing on taxa from the Mediterranean. J Plant Res 126:753–762

Ito Y, Ohi-Toma T, Murata J, Tanaka N (2010) Hybridization and polyploidy of an aquatic plant, *Ruppia* (Ruppiaceae), inferred from plastid and nuclear DNA phylogenies. Am J Bot 97:1156–1167

Ito Y, Tanaka N (2011) Hybridisation in a tropical seagrass genus, *Halodule* (Cymodoceaceae), inferred from plastid and nuclear DNA phylogenies. Telopea 13:219–231

Kato Y, Aioi K, Omori Y, Takahata N, Satta Y (2003) Phylogenetic analyses of *Zostera* species based on *rbcL* and *matK* nucleotide sequences: implications for the origin and diversification of seagrasses in Japanese waters. Genes Genet Syst 78:329–342

Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. PLoS ONE 2:e508

Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G et al (2008) DNA barcoding the floras of biodiversity hotspots. Proc Natl Acad Sci USA 105:2923–2928

Les D, Philbrick CT (1993) Studies of hybridization and chromosome number variation in aquatic angiosperms: evolutionary implications. Aquat Bot 44:181–228

Les DH, Cleland MA, Waycott M (1997) Phylogenetic studies in Alismatidae. II: evolution of marine angiosperms (‘seagrasses’) and hydrophily. Syst Bot 22:443–463

Les DH, Moody ML, Jacobs SWL, Bayer RJ (2002) Systematics of seagrasses (Zosteraceae) in Australia and New Zealand. Syst Bot 27:468–484

Lucas C, Thangaradjou T, Papenbrock J (2012) Development of a DNA barcoding system for seagrasses: successful but not simple. PLoS ONE 7:e29987

Maddison WP, Maddison DR (2008) Mesquite: a modular system for evolutionary analysis. Version 2.5. http://mesquiteproject.org

Mathews S, Tsai RC, Kellogg EA (2000) Phylogenetic structure in the grass family (Poaceae): evidence from the nuclear gene phytochrome B. Am J Bot 87:96–107

McGlathery KJ, Sundbäck K, Anderson IC (2007) Eutrophication in shallow coastal bays and lagoons: the role of plants in the coastal filter. Mar Ecol Prog Ser 348:1–18

Meier R, Shiyang K, Vaidya G, Ng PKL (2006) DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. Syst Biol 55:715–728

Meyer CP, Paulay G (2005) DNA barcoding: error rates based on comprehensive sampling. PLoS Biol 3:e422

Millennium Ecosystem Assessment (2005) Ecosystems and human well-being: synthesis. Island Press, Washington, DC

Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B (2011) How many species are there on Earth and in the ocean? PLoS Biol 9:e1001127

Orth RJ, Curruthers TJB, Dennison WC, Duarte CM, Fourqurean JW, Heck KL et al (2006) A global crisis for seagrass ecosystems. Bioscience 56:987–996

Packer L, Gibbs J, Sheffield C, Hanner R (2009) DNA barcoding and the mediocrity of morphology. Mol Ecol Resour 9:42–50

Papadopoulou A, Chesters D, Coronado I, De la Cadena G, Cardoso A, Reyes JC et al (2015) Automated DNA-based plant identification for large-scale biodiversity assessment. Mol Ecol Resour 15:136–152

Pentinsaari M, Mutanen M, Kaila L (2014) Cryptic diversity and signs of mitochondrial introgression in the *Agrilusviridis* species complex (Coleoptera: Buprestidae). Eur J Entomol 111:475–486

Pettengill JB, Neel MC (2010) An evaluation of candidate plant DNA barcodes and assignment methods in diagnosing 29 species in the genus *Agalinis* (Orobanchaceae). Am J Bot 97:1381–1406

R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Radulovici AE, Archambault P, Dufresne F (2010) DNA barcodes for marine biodiversity: moving fast forward? Diversity. 2:450–472

barnabas_daru@fas.harvard.edu

Rieseberg LH, Choi HC, Ham D (1991) Differential cytoplasmic versus nuclear introgression in *Helianthus*. J Hered 82:489–493
Roy S, Tyagi A, Shulka V, Kumar A, Singh UM, Chaudhary LB et al (2010) Universal plant DNA barcode loci may not work in complex groups: a case study with Indian *Berberis* species. PLoS ONE 5:e13674
Simmons MP, Clevinger CC, Savolainen V, Archer PH, Mathews S, Doyle JJ (2001) Phylogeny of the Celastraceae inferred from phytochrome B gene sequences and morphology. Am J Bot 88:313–325
Smith MA, Fisher BL, Hebert PDN (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Phil Trans R Soc B 360:1825–1834
Somers DE, Devlin PF, Kay SA (1998) Phytochromes and cryptochromes in the entrainment of the *Arabidopsis* circadian clock. Science 282:1488–1490
Tanaka N, Kuo J, Omori Y, Nakaoka M, Aioi K (2003) Phylogenetic relationships in the genera*Zostera* and *Heterozostera* (Zosteraceae) based on *matK* sequence data. J Plant Res 116:273–279
Tomlinson PB, Posluszny U (2001) Generic limits in the seagrass family Zosteraceae. Taxon 50:429–437
Trivedi S, Aloufi AA, Ansari AA, Ghosh SK (2016) Role of DNA barcoding in marine biodiversity assessment and conservation: an update. Saudi J Biol Sci. doi:10.1016/j.sjbs.2015.01.001
Uchimura M, Faye EJ, Shimada S, Inoue T, Nakamura Y (2008) A reassessment of *Halophila* species (Hydrocharitaceae) diversity with special reference to Japanese representatives. Bot Mar 51:258–268
van der Bank H, Greenfield R (2015) A pioneer survey and DNA barcoding of some commonly found gastropod molluscs on Robben Island. ZooKeys 481:15–23
Verhoeven JTA (1979) The ecology of *Ruppia* dominated communities in western Europe. I. Distribution of *Ruppia* representatives in relation to their autecology. Aquat Bot 6:197–268
Wang W, Wu Y, Yan Y, Ermakova M, Kerstetter R, Messing J (2010) DNA barcoding of the Lemnaceae, a family of aquatic monocots. BMC Plant Biol 10:205
Waycott M, Freshwater DW, York RA, Calladine A, Kenworthy WJ (2002) Evolutionary trends in the seagrass genus *Halophila* (thouars): Insights from molecular phylogeny. Bull Mar Sci 71:1299–1308
Waycott M, Duarte CM, Carruthers TJB, Orth RJ, Dennison WC, Olyarnik S et al (2009) Accelerating loss of seagrasses across the globe threatens coastal ecosystems. Proc Natl Acad Sci USA 106:12377–12381

barnabas_daru@fas.harvard.edu