European Journal of Phycology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/tejp20

Assessment of nine markers for phylogeny, species and haplotype identification of *Kappaphycus* species and *Eucheuma denticulatum* (Solieriaceae, Rhodophyta)

Ji Tan, Pui-Ling Tan, Sze-Wan Poong, Janina Brakel, Cecilia Rad Menendez, Eka Sunarwidhi Prasedya, Alison R. Sherwood, Flower E. Msuya, Claire Gachon, Juliet Brodie, Azhar Kassim & Phaik-Eem Lim

To cite this article: Ji Tan, Pui-Ling Tan, Sze-Wan Poong, Janina Brakel, Cecilia Rad Menendez, Eka Sunarwidhi Prasedya, Alison R. Sherwood, Flower E. Msuya, Claire Gachon, Juliet Brodie, Azhar Kassim & Phaik-Eem Lim (01 Oct 2024): Assessment of nine markers for phylogeny, species and haplotype identification of *Kappaphycus* species and *Eucheuma denticulatum* (Solieriaceae, Rhodophyta), European Journal of Phycology, DOI: 10.1080/09670262.2024.2396324

To link to this article: https://doi.org/10.1080/09670262.2024.2396324

EUROPEAN JOURNAL OF PHYCOLOGY, 2024
https://doi.org/10.1080/09670262.2024.2396324

Assessment of nine markers for phylogeny, species and haplotype identification of *Kappaphycus* species and *Eucheuma denticulatum* (Solieriaceae, Rhodophyta)

Ji Tan [a], Pui-Ling Tan [b], Sze-Wan Poong [b], Janina Brakel [c,d,e], Cecilia Rad Menendez [c,f], Eka Sunarwidhi Prasedya[g,h], Alison R. Sherwood [i], Flower E. Msuya [j], Claire Gachon [c,k], Juliet Brodie [l], Azhar Kassim[m] and Phaik-Eem Lim [b]

[a]Department of Agricultural and Food Science, Faculty of Science, Universiti Tunku Abdul Rahman, Kampar 31900, Malaysia; [b]Institute of Ocean and Earth Sciences, Universiti Malaya, Kuala Lumpur 50603, Malaysia; [c]Scottish Association for Marine Science, Scottish Marine Institute, Oban PA37 1QA, UK; [d]Helmholtz Institute for Functional Marine Biodiversity (HIFMB), University of Oldenburg Germany, Ammerländer Heerstrasse 231, Oldenburg 26129, Germany; [e]Alfred-Wegener-Institute, Helmholtz Center for Polar and Marine Research, Am Handelshafen 12, Bremerhaven 27570, Germany; [f]Culture Collection of Algae and Protozoa, Scottish Marine Institute, Oban PA37 1QA, UK; [g]Bioscience and Biotechnology Research Centre, Faculty of Mathematics and Natural Sciences, University of Mataram, Mataram 83115, Indonesia; [h]Department of Biology, Faculty of Mathematics and Natural Sciences, University of Mataram 83115, Indonesia; [i]School of Life Sciences, University of Hawaii, 3190 Maile Way, Honolulu, HI, USA; [j]Zanzibar Seaweed Cluster Initiative, Zanzibar 3794, Tanzania; [k]Unité Molécules de Communication et Adaptation des Micro-organismes, UMR 7245, Muséum National d'Histoire Naturelle, CNRS, Paris 75005, France; [l]Research, Natural History Museum, London SW7 5BD, UK; [m]Department of Fisheries Sabah, Ministry of Agriculture, Fisheries and Food Industries Sabah, Kota Kinabalu 88624, Malaysia

ABSTRACT

Molecular studies have contributed to the taxonomy of carrageenan-producing *Kappaphycus* spp. and *Eucheuma denticulatum*. However, unresolved species complexes and the lack of standardization in the use of genetic markers impede the identification of specimens and the delineation of a robust taxonomic framework. Here, nine molecular markers (*cox*1, *cox*2–3 spacer, *cox*2, *cox*3, COB, ITS, *psb*A, UPA and *rbc*L) were used to generate a multilocus phylogeny for 113 fresh eucheumatoid samples and four herbarium specimens. Analyses of species delineation and genetic distances confirmed the monophyly of currently accepted taxa. These analyses suggest that clades previously reported as *K. striatus* KS1 and KS2 are conspecific, and that *E. denticulatum* EDA 'spinosum' and EDB 'endong/cacing' are also conspecific. The results also unveiled possible new taxa from Hawaii and Indonesia. Each molecular marker and combinations thereof were assessed with regard to species identification, ease of amplification and sequencing, and haplotype characterization. All genetic markers recorded at least 94% success in the amplification and sequencing of fresh specimens, with *cox*1 being the most phylogenetically informative. Automatic partitioning, phylogenetic and tree-based assessments showed *cox*1, *cox*2–3 spacer, *cox*2 and *rbc*L were able to correctly identify species while *cox*1+ *rbc*L, COB+*rbc*L, *cox*2 + *rbc*L or *cox*1+ COB+*rbc*L trees best represented the phylogeny with consistently high nodal support. Among individual markers, *cox*1 identified the greatest number of haplotypes, while UPA, partial *rbc*L (750 bp), ITS, *cox*3 and *cox*2–3 spacer were able to retrieve information from herbarium specimens of 12–16 years of age. These molecular results provide a basis for a database essential for the taxonomic framework, cultivar development and germplasm conservation of eucheumatoids.

HIGHLIGHTS

- Mitochondria *cox*1, *cox*2–3 spacer, *cox*2 and plastid *rbc*L can be used for species identification and *cox*1 for haplotype detection of eucheumatoids.
- *cox*1+*rbc*L, COB+*rbc*L, *cox*2+*rbc*L or *cox*1+COB+*rbc*L are the most cost-effective molecular markers for phylogenetic inference.
- The most comprehensive up to date multilocus phylogeny of *Kappaphycus* spp. and *Eucheuma denticulatum* is presented.

ARTICLE HISTORY Received 29 January 2024; Revised 10 August 2024; Accepted 19 August 2024

KEYWORDS Eucheumatoids; herbarium; molecular markers; multilocus phylogeny; taxonomy

Introduction

The use of genetic markers has enabled rapid developments in phycological research related to taxonomy, evolutionary history, population genetics, conservation, DNA barcoding, genetic selection and breeding (Bartolo *et al.*, 2020; McCoy *et al.*, 2020; Brakel *et al.*, 2021). Genetic resources from these studies are invaluable for germplasm conservation and strain improvement, especially in the face of challenges such as climate change, ocean acidification, anthropogenic pollution as well as pests and diseases (Ward *et al.*, 2020, 2022; Yang *et al.*, 2021). Whilst studies that characterize and evaluate genetic resources are more prevalent in the brown algae, e.g. *Undaria pinnatifida* (Avia *et al.*, 2017; Brakel *et al.*, 2021; Shan & Pang, 2021), they are gaining momentum in the red algae (Yang *et al.*, 2016; van Beveren *et al.*, 2022; Borg *et al.*, 2023).

The eucheumatoids *Kappaphycus* and *Eucheuma denticulatum* are commercially important red seaweeds that produce carrageenan, a valuable phycocolloid increasingly used in the food and cosmetic industries (Campo *et al.*, 2009). These seaweeds are mostly cultivated in

CONTACT Phaik-Eem Lim phaikeem@um.edu.my; Sze-Wan Poong szewan@um.edu.my

Published online 01 Oct 2024

Southeast Asia and contributed to 98.88% (11 491 956 tons wet weight) of global carrageenan production in 2019 (Cai *et al.*, 2021). Before the 2000s, these seaweeds were poorly documented and often misidentified due to their morphological plasticity (Doty & Norris, 1985; Zuccarello *et al.*, 2006). The mitochondrial *cox*2–3 spacer region and plastid RuBisCO spacer genetic markers were first utilized to identify specimens of *Kappaphycus* and *Eucheuma* by Zuccarello *et al.* (2006), who laid the groundwork for studies on eucheumatoid phylogeny and haplotype diversity. The ability to correctly identify species of eucheumatoids has spurred developments in taxonomy, phylogeny, DNA barcoding and bioinvasion detection (Brakel *et al.*, 2021; Roleda *et al.*, 2021; Tan *et al.*, 2022*a*, 2022*b*). To date, the mitochondrial *cox*2–3 spacer and *cox*1 markers are most commonly used for genetic studies of *Kappaphycus* and *Eucheuma*.

Nevertheless, the taxonomy of *Kappaphycus* and *Eucheuma* remains problematic due to the lack of resolution at species level (i.e. *Kappaphycus* sp. Hawaii, *Kappaphycus* sp. Africa, *K. malesianus, K. inermis, E. denticulatum* Southeast Asia, *E. denticulatum* Africa etc.) and subspecies levels (i.e. *K. alvarezii* cultivar and *K. alvarezii* Africa, *K. striatus* genotypes 1 and 2, *E. denticulatum* cultivars 'spinosum' and 'endong/cacing') (Zuccarello *et al.*, 2006; Dumilag *et al.*, 2014; Dumilag & Zuccarello, 2022; Tan *et al.*, 2022*a*). Efforts in characterizing these taxa were typically hampered by missing type specimens and/or failure to extract DNA from old herbarium specimens (Dumilag & Lluisma, 2014; Lim *et al.*, 2014; Dumilag & Zuccarello, 2022; Tan *et al.*, 2022*a*), resulting in information gaps in *K. procrusteanus* and the majority of *Eucheuma* species. Of the 25 taxonomically accepted *Eucheuma* spp., only four (*E. denticulatum, E. perplexum, E. serra* and *E. platycladum*) were characterized genetically, with *E. denticulatum* 'spinosum' being the only global cultivar. *E. denticulatum* 'endong/cacing' is a less common cultivar in Southeast Asia with a reportedly thicker and fleshy thallus (Montes *et al.*, 2008; Lim *et al.*, 2014; Tan *et al.*, 2022*a*). The remaining *Eucheuma* spp. have not been reported again after their initial description. These taxonomic issues are compounded by a lack of consensus on the use of genetic markers between researchers, which has hindered comparative analyses of specimens from a wider biogeographic range. To that end, nine genetic markers (i.e. *cox*1, *cox*2–3 spacer, *cox*2, *cox*3, COB, ITS, *psb*A, UPA and *rbc*L) were chosen based on their use in eucheumatoid or red algal research (Freshwater & Rueness, 1994; Zuccarello *et al.*, 2006; Sherwood & Presting, 2007; Tan *et al.*, 2012*a*, 2012*b*; Saunders & Moore, 2013; Zhao *et al.*, 2013). The aims of this study were to (1) identify and recommend the most cost-efficient molecular marker(s) for species and haplotype identification, (2) provide a robust reference database for these markers and (3) present an up to date phylogeny of commercially relevant eucheumatoids using a multilocus dataset.

Materials and methods

Sample collection and DNA extraction

Farmed and wild specimens of *Kappaphycus* and *Eucheuma* were sampled from Malaysia, Indonesia, Fiji and Tanzania. Farmed specimens of *K. alvarezii, K. striatus, K. malesianus, E. denticulatum* 'spinosum' and 'endong/cacing' were collected from cultivation lines of seaweed farms. *E. denticulatum* 'spinosum' and 'endong/cacing' represented two different morphotypes reported from Southeast Asia and are suspected to be different species (Montes *et al.*, 2008; Lim *et al.*, 2014; Tan *et al.*, 2022*a*). Samples were tentatively identified in the field based on their gross morphology and subsequently confirmed through sequence data. Samples were then selected for further analysis based on their morphological variation, geographic distribution and genetic diversity as inferred based on the *cox*1 gene. A total of 113 contemporary specimens, defined as fresh, dried or preserved specimens from which DNA was extracted within a year of collection, were used in the present study. Three specimens of *K. striatus* (F. Schmitz) L.M.Liao and one of *E. denticulatum* (N.L. Burman) Collins & Hervey from Hawaii were also extracted from herbarium specimens collected between 2007–2011. A small portion of the thallus was aseptically excised from each specimen for DNA extraction, which was performed using the *i*genomic Plant DNA Extraction Mini Kit (iNtRON Biotechnology Inc., Korea), or by a cetrimethylammonium bromide (CTAB) based method following Gachon *et al.* (2009). Herbarium specimens were deposited at the University of Malaya Seaweeds and Seagrasses Herbarium (KLU), Pusat Unggulan Biosains dan Bioteknologi (PUBB) Mataram University and/or the Bernice Pauahi Bishop Museum (BISH). The list of species, country of origin, GenBank accession numbers and haplotype information for each marker are provided in Supplementary table S1.

DNA amplification and sequencing

The extracted DNA was used to PCR amplify the *cox*1, *cox*2–3 spacer, *cox*2, *cox*3, COB, ITS, *psb*A, UPA and *rbc*L genetic markers. The primer details and annealing temperature of each marker (Saunders & Druehl, 1992; Freshwater & Rueness, 1994; Zuccarello *et al.*, 1999; Gavio & Fredericq, 2002; Yoon *et al.*, 2002; Geraldino *et al.*, 2006; Sherwood & Presting, 2007; Yang *et al.*, 2008; Tan *et al.*, 2012*a*; Saunders & Moore, 2013) are shown in Supplementary table S2.

PCR amplicons were evaluated using gel electrophoresis of a 1.0% agarose gel stained with SYBR®

Safe DNA gel stain (Invitrogen, USA) and sent to Apical Scientific Sdn. Bhd., Malaysia, or LGC genomics, Germany for purification and DNA sequencing.

Contig assembly and multiple sequence alignment

The DNA sequence trace files of each specimen were used to assemble contigs for each genetic marker in ChromasPro V1.5 (Technelysium Pty Ltd). Two different sets of multiple sequence alignment (MSA) were generated: (1) MSA consisting of sequences from only 95 samples for all nine markers; and (2) MSA of all sequenced DNA from the present study and GenBank DNA sequences. MSA was performed for *cox*1, *cox*2–3 spacer, *cox*2, *cox*3, COB, *psb*A, UPA and *rbc*L using MUSCLE (Edgar, 2004). The ITS (consisting of complete ITS1, 5.8S rRNA gene and partial ITS2) dataset was aligned using MAFFT v.7 (Katoh *et al.*, 2019) with the QINSi refinement method (https://mafft.cbrc.jp/alignment/server/). *Chondrus crispus* (Gigartinales) and *Gracilaria vermiculophylla* (Gracilariales) were used as the outgroup for all datasets based on genetic relatedness (i.e. subclass Rhodymeniophycidae) and availability of DNA data for the nine markers.

A DNA dataset of *Kappaphycus* spp. and *E. denticulatum* consisting of the combination of all nine markers except ITS (hereafter referred to as the concatenated dataset) was also generated. ITS was not included due to species misidentification within our dataset.

Genetic distances

Uncorrected pairwise genetic distances (GD) were estimated for individual and combined markers datasets using PAUP 4.0b10 (Swofford, 2003). Genetic distance information is based on the concatenated dataset unless otherwise specified.

Phylogenetic analysis

To assess the phylogenetic relationships of eucheumatoids, Maximum likelihood (ML) and Bayesian (BI) trees were constructed for every genetic marker plus sequences from GenBank (Supplementary table S3), as well as additional sequences not used in the 95 sequence dataset.

Partition Finder v2.1.1 (Lanfear *et al.*, 2018) was used to generate output files displaying the optimal partitioning schemes and substitution models of each dataset. ML analyses were performed using the IQTREE web server (Trifinopoulos *et al.*, 2016) at http://iqtree.cibiv.univie.ac.at/. The ML tree was reconstructed over 2000 ultrafast bootstrap replicates using the best-fit models of substitution suggested for each partition in IQTREE.

BI analysis was performed in MrBayes v3.2.6 (Ronquist *et al.*, 2012) using the recommended partitions and substitution models from Partition Finder. The Markov chain Monte Carlo (MCMC) method was used with two independent runs of 20 000 000 generations, with each run using four Markov chains, and trees sampled every 100th generation. The convergence of log likelihood values was identified using Tracer v1.7 (Rambaut *et al.*, 2018), and the first 25% of samples were discarded as burn-in. All phylogenetic trees were visualized and annotated using Figtree v1.3.1 (http://tree.bio.ed.ac.uk/software/figtree/).

The DNA sequences of eucheumatoid specimens from Hawaii (i.e. ARS 02860, ARS 03953, ARS 03513 and ARS 08101) were also used for phylogenetic inference whenever possible as they were reportedly genetically distinct from the common eucheumatoids from Southeast Asia (Zuccarello *et al.*, 2006; Lim *et al.*, 2014; Brakel *et al.*, 2021).

Marker assessments

The nine molecular markers were assessed based on the following criteria:

(i) PCR performance.
(ii) Marker performance, i.e. pairwise distance, nucleotide diversity, phylogenetic assessment (via Assemble Species by Automatic Partitioning (ASAP), General Mixed Yule Coalescent (GMYC), multirate Poisson Tree Processes (mPTP) and tree-based approaches) and haplotype diversity.

PCR performance

The ease of PCR amplification of each genetic marker and the marker's ability to sequence fresh specimens and herbarium specimens preserved for 12–16 years was determined. The sequence lengths of the MSA blocks generated for all nine genetic markers were recorded using Geneious Prime v2020.1.2 (https://www.geneious.com).

Marker performance

The ability of each genetic marker to identify specimens at the species level was assessed based on (i) pairwise distances and (ii) phylogenetic trees. In both analyses, the species delimitation of *Kappaphycus* spp. and *E. denticulatum* from the concatenated dataset was used as a reference to assess the congruence of each individual marker as this dataset, the most comprehensive to date, is presumed to best reflect species status of samples.

Percentage pairwise difference, which determines the number of species or Operational Taxonomic Units (OTU) that could be detected using pairwise distance automatic partitioning, was recorded using

Geneious Prime v2020.1.2. The number of parsimony informative sites, a determinator of phylogenetic resolution, was generated using PAUP (Swofford, 2003), while nucleotide diversity (π) was computed using DNAsp.

Pairwise distance and phylogenetic assessment

Species delimitation was determined using three commonly used species delimitation models, i.e. ASAP, General Mixed Yule Coalescent (GMYC; Fujisawa & Barraclough, 2013) and multi-rate Poisson Tree Processes (PTP; Zhang *et al.*, 2013; Kapli *et al.*, 2017). ASAP predicts the 'best' number of partitions and species from a single locus aligned dataset without phylogenetic reconstruction. It also does not require any *a priori* information on number of species, species composition, phylogeny or *a priori* defined intraspecific genetic distances (Puillandre *et al.*, 2020). GMYC adopts a speciation and a neutral coalescent model to delimit species based on classification of branches of an ultrametric tree, while mPTP employs a new algorithm to model the branching processes according to accumulated expected substitution between subsequent speciation events without an ultrametric input tree (Fujisawa & Barraclough, 2013; Zhang *et al.*, 2013; Kapli *et al.*, 2017). These three models were used because each one displays different strengths and weaknesses: ASAP delimits species based solely on a barcode gap, GMYC is prone to oversplitting, and mPTP tends to merge species more frequently (Luo *et al.*, 2018).

ASAP was performed using the default parameters at https://bioinfo.mnhn.fr/abi/public/asap/ (Puillandre *et al.*, 2020). For GMYC, Ultrametric gene trees were inferred using a relax molecular clock model and Yule prior using BEAST v.2.7.6 (Bouckaert *et al.*, 2019). With the outgroups removed, analyses were performed for 50 million generations with trees saved every 1000 generations, and burn-ins discarded using Tracer. Single threshold optimization was used to generate maximum credibility ultrametric trees using the GMYC method in the R programming environment (http://rforge.rproject.org/projects/splits). mPTP was performed using the online server (https://mcmcmptp.hits.org/mcmc/) and the trees previously generated from BEAST.

Tree-based assessment

The tree-based assessment was undertaken to determine the resolution of each genetic marker based on its ability to identify species or OTUs, and how congruent the resulting tree topology was to that of the concatenated tree. ML and BI phylogenetic trees were computed based on the 95 sequence dataset for each genetic marker, a combination of selected best-performing markers, as well as the concatenated dataset. These phylogenetic trees were inferred using the aforementioned protocols.

Haplotype networks

Individual haplotype networks of *Kappaphycus* and *E. denticulatum* were generated using TCS 1.2.1 to estimate relationships between haplotypes (Clement *et al.*, 2000). Gaps, whenever present, were considered as '5th state' in TCS, and connection limits were set at 100 for all datasets. Haplotype diversity was computed using DNAsp. For consistency, the haplotype nomenclature and codes used in the present study follow those of earlier studies (Zuccarello *et al.*, 2006; Tan *et al.*, 2012*b*, 2014; Dumilag & Lluisma, 2014; Lim *et al.*, 2014; Dumilag *et al.*, 2016*a*, 2016*b*, 2018).

Results

Multilocus phylogeny and species delimitation

The phylogenetic tree based on the concatenated dataset revealed four clades within *Kappaphycus*, three of which match the currently accepted species, *K. alvarezii* (Doty) L.M.Liao, *K. striatus* and *K. malesianus* J.Tan, P.E.Lim & S.M.Phang, in addition to a novel clade, highlighted as *Kappaphycus* sp. (Indonesia) (Fig. 1). The currently accepted *E. denticulatum* also formed a clade.

Within the *K. alvarezii* clade, samples of the main commercial cultivar collected from Malaysia were shown to be genetically identical, but 0.02–0.03% different genetically from several genotypes from Indonesia (GRP1, GRP3, PWR2, PWR8, TRT6 and TRT20), and 0.02–0.12% from Tanzanian genotypes (MLN1, MLN2 and MUU27) across 6834 bp (Supplementary table S4). The *K. striatus* clade was delineated into three subclades (KS1a, KS1b and KS2). The commercial *K. striatus* cultivars belonged to KS1a, whereas KS1b and KS2 comprised mostly wild specimens. The p distance between KS1 and KS2 was recorded at less than 0.86%. *Kappaphycus* sp. (Indonesia), which comprised specimens ARW1 and ARW5, was genetically different from other known *Kappaphycus* spp. by 3.35–4.41% and was inferred to be sister to both *K. alvarezii* and *K. striatus*. *K. malesianus* formed a monophyletic lineage and was composed of only one OTU in this dataset, with several wild genotypes genetically dissimilar to the cultivar.

The *E. denticulatum* clade was composed of two subclades ED1 and ED2: ED1 was poorly supported and consisted of wild and cultivated specimens of the 'spinosum' genotype in which the commercial *E. denticulatum* cultivar belongs, while ED2 was composed of the genotypes

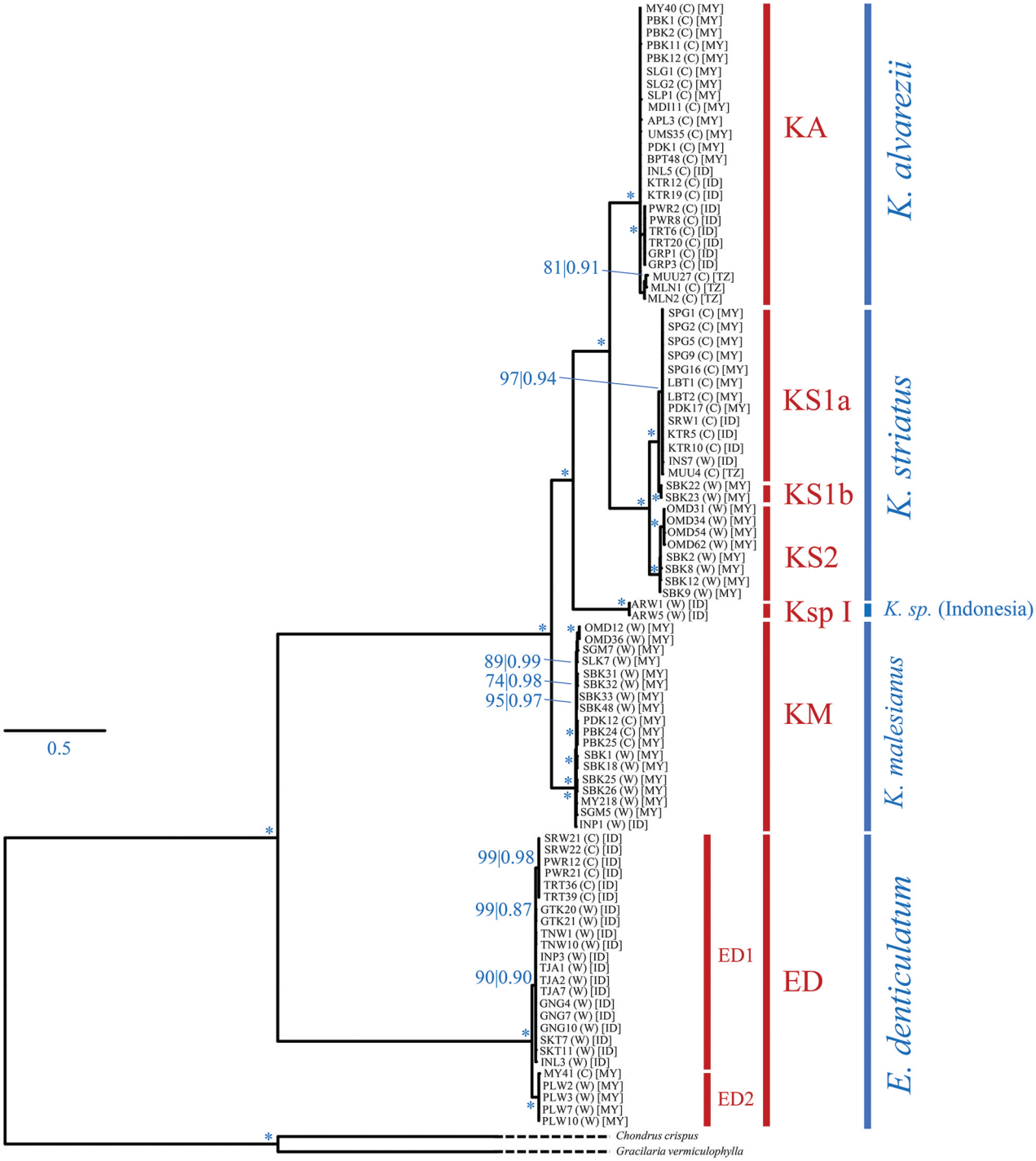

Fig. 1. Bayesian tree of *Kappaphycus* spp. and *Eucheuma denticulatum* based on the concatenated dataset. Numbers at nodes indicate ultrafast ML (UFML) bootstrap support and Bayesian posterior probabilities. Asterisks indicate ML ≥ 95%; PP ≥ 0.99. Nodal support with UFML < 95% and BI < 0.90 are not shown. Blue lines represent the current taxonomic delineation of species, whereas red lines represent Operational Taxonomic Units (OTUs). Letters in parentheses indicate cultivar (C) or wild specimen (W). Letters in square brackets indicate locality of origin: ID, Indonesia; MY, Malaysia; TZ, Tanzania. Scale bar represents number of substitutions per site.

known as 'endong' or 'cacing' which differed genetically by 0.52–0.59%.

Selected genetic distance data of the concatenated dataset and nine individual markers are summarized in Supplementary table S4. Calculation of the genetic distance of the concatenated dataset, using the currently accepted species, revealed an intergeneric genetic distance of 11.92–12.98% between *Kappaphycus* spp. and *E. denticulatum*, and an interspecific genetic distance of 3.60–4.35% within *Kappaphycus*. The ASAP species delineation analysis best reflected that of the concatenated tree (Fig. 1) and identified subclades KS1a, KS1b, KS2, ED1 and ED2 as operational taxonomic units (OTUs) nested within currently accepted species of *Kappaphycus* and *E. denticulatum*.

Marker performance

Utility of individual markers for cost-effective species identification

The ability of each genetic marker to PCR amplify and sequence both fresh and herbarium specimens is summarized in Table 1. All nine genetic markers recorded an amplification success of at least 94.7%. Nested PCR was occasionally required for the amplification of *cox*1 and *rbc*L. Except for specimen MIKO10 from Africa (4 out of 9 markers amplified), amplification failure is random between samples for all markers. The use of *cox*1 recorded the highest number of phylogenetically informative sites at 293, followed by ITS, COB, *cox*3, *cox*2, *rbc*L, *cox*2–3 spacer, *psb*A and UPA. The ITS MSA for *Kappaphycus* and *Eucheuma* were genetically informative and also different in length at 742 bp and 975 bp, respectively.

Amplification and sequencing success were inconsistent for the four 12–16-year-old herbarium specimens (ARS 02860, ARS 03953, ARS 03513 and ARS 08101). UPA was the only marker able to be amplified for all four specimens. For the other markers ITS and *rbc*L sequences were obtained for 3 specimens, *cox*2–3 spacer and *cox*3 for 2 and *cox*2 for 1.

Pairwise distance and phylogenetic assessment

The interpretation of species delimitation from the concatenated dataset described above was used as a reference to assess the performance of each marker (Table 1). Based on ASAP, all genetic markers except ITS were able to consistently identify the eucheumatoid taxa used in the present study (Fig. 2). Among these, only *cox*1 and *cox*2–3 spacer were able to detect the seven OTUs and best reflected the phylogenetic delimitation inferred by the concatenated dataset. *Cox*2 and *cox*3 were genetically variable enough to detect KS1 and KS2, two OTUs within *K. striatus*. Despite being unable to detect additional OTUs, the five main partitions

Table 1. PCR amplification performance and genetic information of markers used in the present study.

	PCR amplification success		MSA block based on the 95-sequence dataset				
Genetic marker	Fresh specimens (n = 113)	Herbarium specimens (12–16 years) (n = 4)	Alignment length (bp)	Number of parsimony-informative sites	Pairwise difference (%)	Nucleotide diversity, π	Anomaly in identification
*cox*1	99.1% amplified and sequenced; nested primers required occasionally	Unsuccessful	1,407	292	8	0.078 ± 0.043	NA
*cox*2	98.2% directly amplified and sequenced	ARS 03953 amplified (15 years)	575	127	8.3	0.086 ± 0.047	NA
*cox*2–3 spacer	99.1% directly amplified and sequenced	ARS 03513 (15 years) and ARS 08101 (12 years) amplified	335	82	9.5	0.093 ± 0.052	NA
*cox*3	96.5% directly amplified and sequenced; difficulty in amplifying certain *E. denticulatum* genotypes	ARS 02860 (16 years) amplified	739	169	8.9	0.089 ± 0.050	Failed to resolve two *E. denticulatum* specimens
COB	94.7% directly amplified and sequenced	Unsuccessful	860	186	8.5	0.085 ± 0.045	NA
ITS	95.5% directly amplified and sequenced	ARS 02860, ARS 03953 and ARS 03513 amplified	1,039 (uneven MSA block) 742 for *Kappaphycus* 975 for *Eucheuma*	219	26.9	0.106 ± 0.059	DNA sequence resolved six *K. alvarezii* specimens as *K. malesianus*
*psb*A	97.3% directly amplified and sequenced	ARS 02860 amplified	893	81	3.4	0.034 ± 0.019	NA
*rbc*L	99.1% amplified and sequenced; nested primers required occasionally	Partial sequence of ARS 02860, ARS 03513 and ARS 08101 amplified	1,388	124	4.1	0.036 ± 0.018	NA
UPA	100% directly amplified and sequenced	ARS 02860, ARS 03513, ARS 08101 and ARS 03953 amplified	387	13	1.3	0.012 ± 0.007	NA

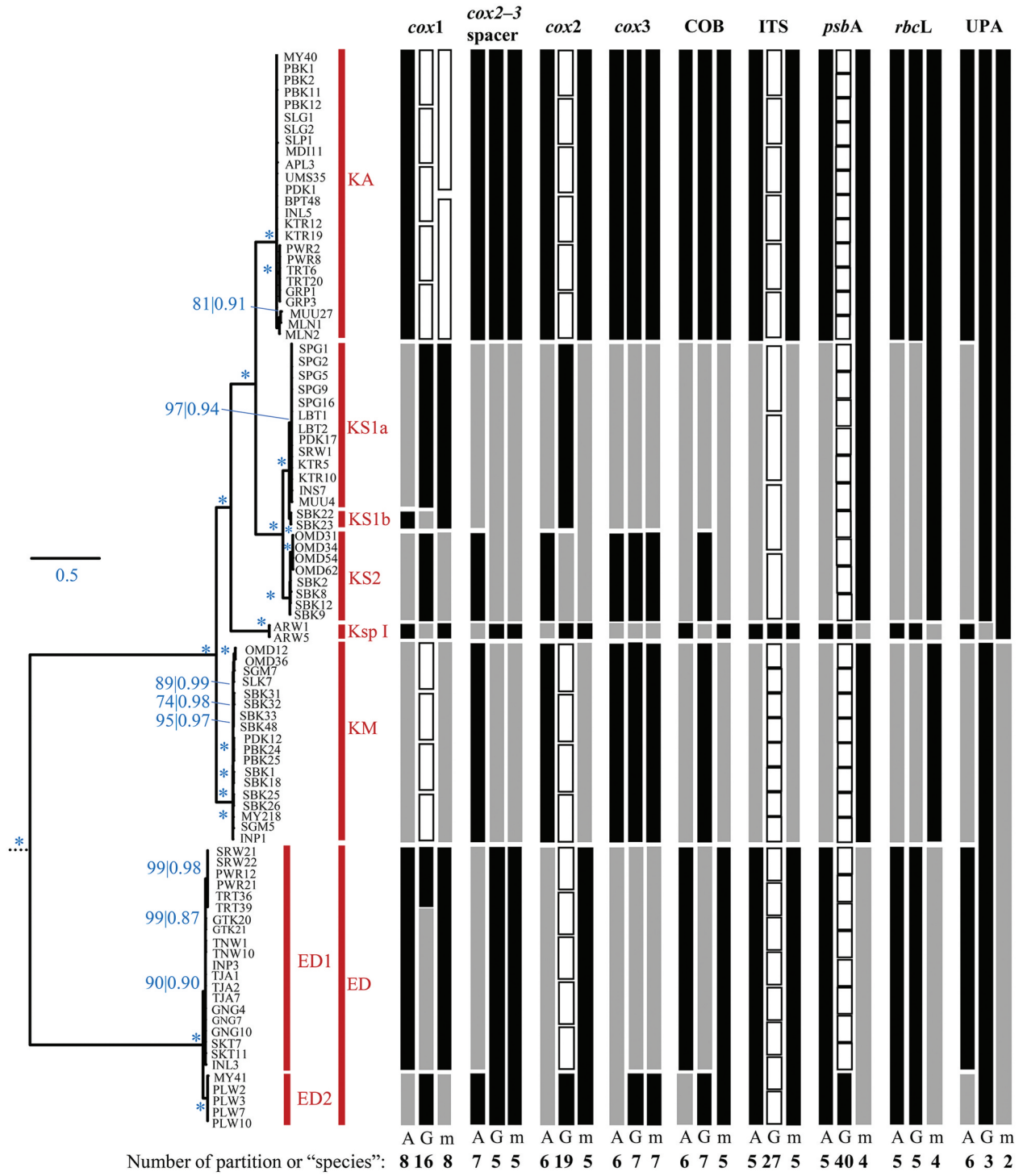

Fig. 2. Simplified Bayesian tree of *Kappaphycus* spp. and *Eucheuma denticulatum* based on the concatenated dataset. Numbers at nodes indicate ultrafast ML (UFML) bootstrap support and Bayesian posterior probabilities. Asterisks indicate ML ≥ 95%; PP ≥ 0.99. Nodal support with UFML < 95% and BI < 0.90 are not shown. Red lines represent Operational Taxonomic Units (OTUs). Black and grey boxes correspond to species groups inferred by delimitation analyses. White boxes indicate number of delimited species which do not correspond to the concatenated tree. A, Assemble Species by Automatic Partitioning (ASAP), G, General Mixed Yule Coalescent (GMYC), m, multi-rate Poisson Tree Processes (mPTP). Scalebar represents number of substitutions per site.

inferred by *psb*A and *rbc*L coincided with current species delimitation.

The number of species delimited using GMYC and PTP was incongruent (Fig. 2). Using the same dataset, GMYC inferred more than 16 species for *cox*1, *cox*2, ITS and *psb*A, while mPTP inferred fewer than five species for *psb*A, *rbc*L and UPA. *Cox*3 and COB recorded the least variation in number of partitions or 'species' between ASAP, GMYC and mPTP approaches, while *cox*1 was inferred with 8 partitions or 'species' by ASAP and mPTP.

Tree-based assessment

The results of the tree-based assessment are summarized in Table 2. The *cox*1 tree (Fig. 3a) best reflected the topology of the concatenated tree (Fig. 1) and this marker can be used to infer the phylogenetic relationships of all *Kappaphycus* spp. and *E. denticulatum* including the three *K. striatus* OTUs (KS1a, KS1b and KS2). The tree also identified specimen MY218 as an additional subclade KM2. The *rbc*L tree (Fig. 3b) was topologically similar to that of the concatenated tree and nodes were well-supported, although the marker was not useful to clearly differentiate subclades KS1 and KS2. The topology of *cox*2 (Fig. 3c) resembled that of *cox*1 but without the ability to identify KS1b (nested within KS1) and KM2. COB (Fig. 3d) performed similarly to *cox*1, except for overall weaker interspecific nodal support, and a poorly resolved taxonomic position of *K. malesianus* and *Kappaphycus* sp. (Indonesia). The use of *cox*2–3 spacer, *cox*3 and psbA (Fig. 3e, 3f and Supplementary fig. 1A, respectively) recorded low interspecific nodal support and these markers were unable to elucidate the relationship between *K. striatus, K. malesianus* and *Kappaphycus* sp. (Indonesia). The UPA tree (Supplementary fig. S1B) was not useful for inferring the monophyly of each species, while the ITS tree (Supplementary fig. S1C) failed to resolve some samples of *K. alvarezii* and *K. malesianus*. Likewise, *cox*3 resolved two *E. denticulatum* specimens from Tanzania (MUU29 and KID30) as *K. alvarezii* (Supplementary table S1).

For *E. denticulatum*, none of the nine markers was able to reliably resolve the 'spinosum' (ED1) and 'cacing' (ED2) genotypes of *E. denticulatum* as two distinct sister subclades as inferred by the concatenated dataset.

Additional assessments were completed using different combinations of the best-performing genetic markers, i.e. *cox*1, *cox*2, COB and *rbc*L (Table 2). Among the possible marker combinations, the *cox*1 + *rbc*L (Supplementary fig. S2A),

Table 2. Taxonomic resolution of individual and combined genetic markers based on tree-based assessment.

Genetic marker	No. of OTU (list of OTUs defined)	Able to delimit species into monophyletic clade?	Phylogeny of eucheumatoid species congruent with concatenated tree?	Interspecific nodal support values
		Individual markers		
*cox*1	8 (KA, KS1a, KS1b, KS2, Ksp I, KM1, KM2, ED)	Yes	Yes	UFML: 70–100% BI: 0.65–1.00
*rbc*L	5 (KA, KS1, Ksp I, KM, ED)	Yes	Yes	UFML: 89–100% BI: 1.00
*cox*2	6 (KA, KS1, KS2, Ksp I, KM, ED)	Yes	Yes	UFML: 71–100% BI: 1.00
COB	8 (KA, KS1a, KS1b, KS2, KM1, KM2, Ksp I, ED)	Yes	No	UFML: NS–88% BI: 0.59–1.00
*cox*2–3 spacer	6 (KA, KM, Ksp I, KS1, KS2, ED)	Yes	No	UFML: NS–92% BI: NS–1.00
*cox*3	4 (KA1, KS1, KS2, Ksp I)	No	No	UFML: NS–100% BI: NS–1.00
*psb*A	6 (KA, KS1, KS2, KM, Ksp I, ED)	No	No	UFML: NS–100% BI: 0.77–1.00
UPA	2 (KA, ED)	No	No	UFML: NA BI: NA
ITS	4 (KA, KS, Ksp I, ED)	No	No	UFML: NA BI: NA
		Combined markers		
*cox*1+ *rbc*L	8 (KA, KS1a, KS1b, KS2, Ksp I, KM1, KM2, ED)	Yes	Yes	UFML: 74–100% BI: 0.90–1.00
COB+ *rbc*L	8 (KA, KS1a, KS1b, KS2, Ksp I, KM1, KM2, ED)	Yes	Yes	UFML: 89–100% BI: 0.71–1.00
cox2+ rbcL	6 (KA, KS1, KS2, Ksp I, KM, ED)	Yes	Yes	UFML: 95–100% BI: 1.00
*cox*1+ COB+ *rbc*L	8 (KA, KS1a, KS1b, KS2, Ksp I, KM1, KM2, ED)	Yes	Yes	UFML: 65–100% BI: 0.52–1.00
*cox*1+ *cox*2	8 (KA, KS1a, KS1b, KS2, Ksp I, KM1, KM2, ED)	Yes	Yes	UFML: 70–100% BI: 0.98–1.00
*cox*1+ COB	8 (KA, KS1a, KS1b, KS2, KM1, KM2, Ksp I, ED)	Yes	No	UFML: 89–100% BI: 0.92–1.00
*cox*2+ COB	8 (KA, KS1a, KS1b, KS2, KM1, KM2, Ksp I, ED)	Yes	No	UFML: 74–100% BI: 0.57–1.00

*NS, No support; NA, Not applicable.

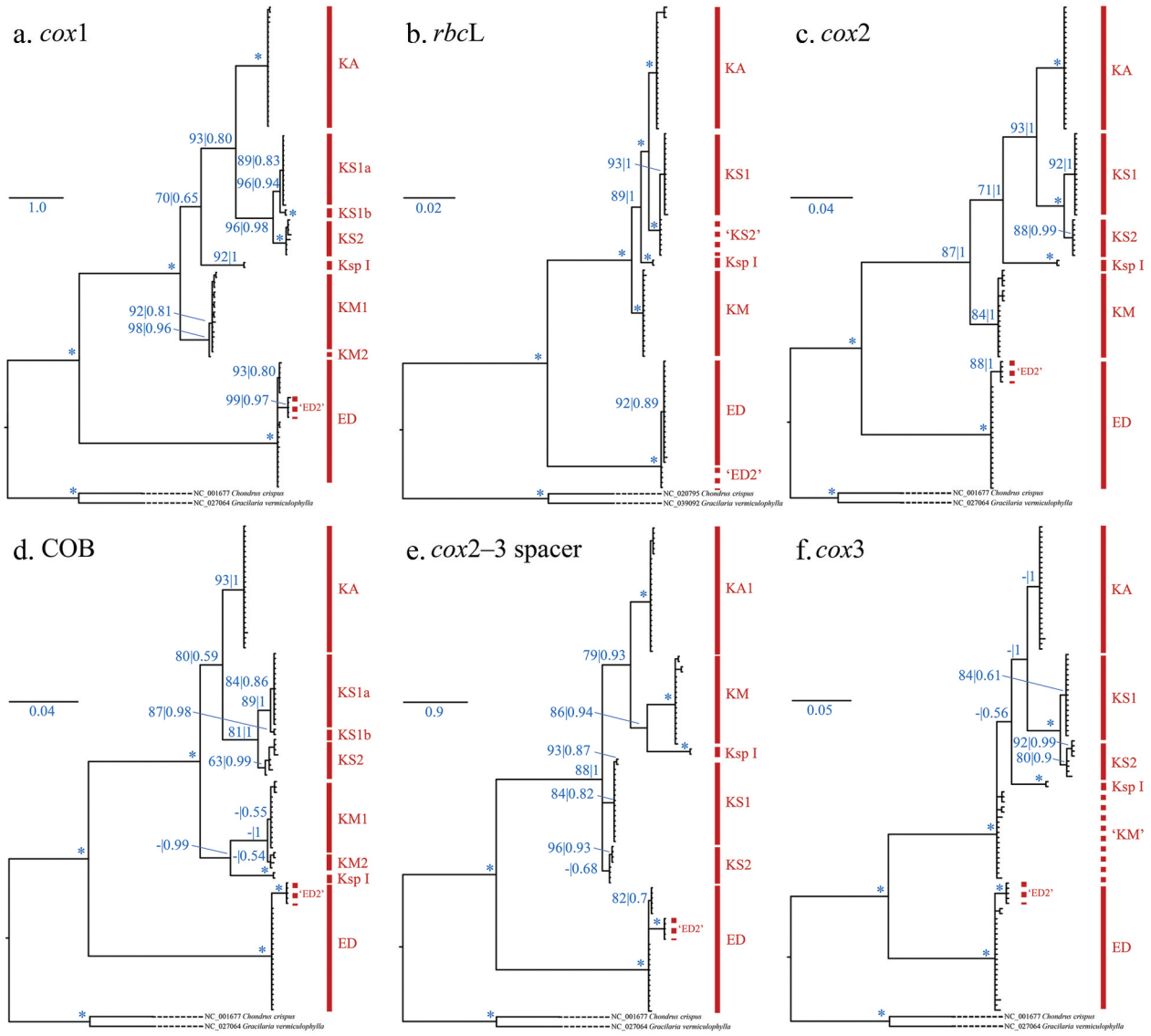

Fig. 3. Bayesian trees of *Kappaphycus* spp. and *Eucheuma denticulatum* based on individual DNA markers (a) *cox*1, (b) *rbc*L, (c) *cox*2, (d) COB, (e) *cox*2–3 spacer and (f) *cox*3. Numbers at nodes indicate ultrafast UFML bootstrap support and Bayesian posterior probabilities. For simplicity, only nodes of main and relevant subclades are annotated. Asterisks indicate ML ≥ 95%; PP ≥ 0.99. Red lines represent Operational Taxonomic Units (OTUs). Dotted red lines indicate unresolved species delimitation. KA, *K. alvarezii*; KS, *K. striatus*; KM, *K. malesianus*; Ksp I, *Kappaphycus* sp. (Indonesia); ED, *E. denticulatum*. Scale bar represents number of substitutions per site.

COB+*rbc*L (Supplementary fig. S2B), *cox*2 + *rbc*L (Supplementary fig. S2C) and *cox*1 + COB + *rbc*L (Supplementary fig. S2D) combinations can be used to infer a tree topology similar to that of the concatenated dataset. However, none of the combinations were able to differentiate *E. denticulatum* subclades ED1 and ED2. The *cox*1 + *rbc*L, COB + *rbc*L and *cox*1 + COB + *rbc*L datasets were useful to identify an additional subclade KM2 of *K. malesianus* but with inconsistent support. The trees of the remaining marker combinations (*cox*1 + *cox*2, *cox*1 + COB, and *cox*2 + COB) are provided in Supplementary fig. S3.

The inclusion of GenBank DNA sequences of *Kappaphycus* spp., *Eucheuma denticulatum* and other related genera (i.e. *Betaphycus, Eucheumatopsis, Kappaphycopsis* and/or *Mimica*) generally decreased the interspecific phylogenetic resolution of individual markers (see Supplementary figs S4–S11 for more details in Supplementary data S2).

Haplotype analyses

The number of haplotypes identified and haplotype diversity for each genetic marker is summarized in Table 3, while haplotype information for specimens used in the present study is in Supplementary table S1. The *cox*1 haplotype network (Fig. 4a) revealed the greatest number of haplotypes for *Kappaphycus* spp. and *E. denticulatum* and the highest haplotype diversity (0.897). This was followed by COB (Fig. 4b), *cox*3 (Fig. 4c) (second highest haplotype diversity, 0.896)

Table 3. Summary of haplotype analyses for individual and combined genetic markers.

Genetic marker	Sequence length analyzed	Number of *Kappaphycus* and *Eucheuma* haplotypes						Haplotype diversity, H_d
		K. alvarezii	*K. striatus*	*K. malesianus*	*K.* sp. (Indonesia)	*E. denticulatum*	Total	
*cox*1	1,407 bp	3	7	10	1	6	**27**	0.897 ± 0.0003
*cox*2–3 spacer	333 bp for *Kappaphycus* 335 bp for *E. denticulatum*	1	4	4	1	5	**15**	0.850 ± 0.0180
*cox*2	575 bp	1	2	3	1	2	**9**	0.836 ± 0.0003
*cox*3	739 bp	3	4	4	1	4	**16**	0.896 ± 0.0002
COB	860 bp	2	7	4	1	2	**16**	0.854 ± 0.0003
ITS*	744 bp for *Kappaphycus* 975 bp for *E. denticulatum*	2	4	5	1	4	**16**	0.870 ± 0.0003
*psb*A	893 bp	1	2	3	1	4	**11**	0.841 ± 0.0003
*rbc*L	1,388 bp	2	3	2	1	2	**10**	0.820 ± 0.0002
UPA	387 bp	1	1	1	1	2	**6**	0.786 ± 0.0002
Mitochondrial markers only	3,916 bp	5	10	13	1	9	**38**	0.933 ± 0.0002
Plastid genes only	2,668 bp	2	3	4	1	3	**12**	0.935 ± 0.0017

*Misidentified haplotypes were not considered.

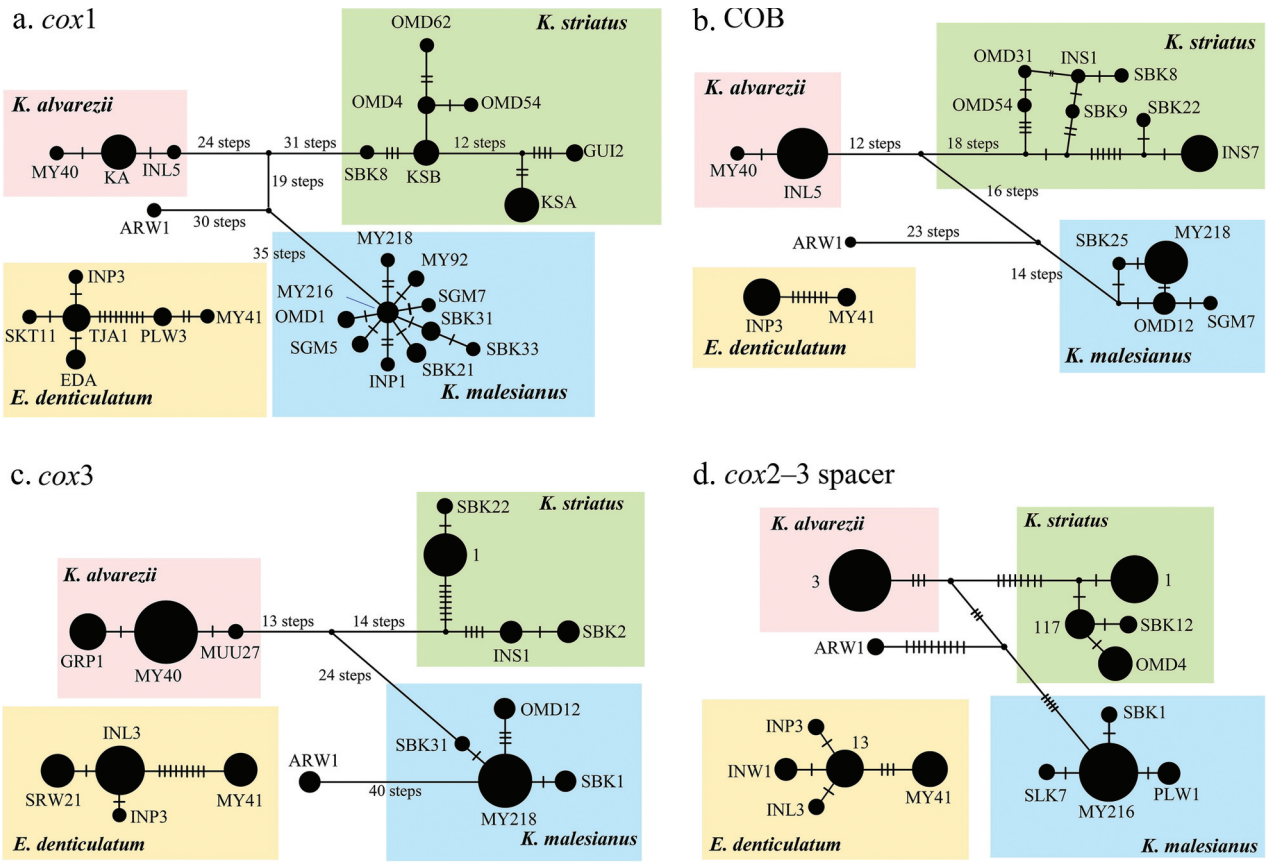

Fig. 4. Haplotype networks of *Kappaphycus* spp. and *Eucheuma denticulatum* based on (a) *cox*1, (b) COB, (c) *cox*3 and (d) *cox*2–3 spacer. The size of each haplotype circle corresponds to its number of constituting samples. Each line between haplotypes indicates one base pair change. Missing haplotypes are indicated by small empty circles. Coloured boxes indicate currently recognized species boundaries based on genetic data. Haplotype networks not drawn to scale.

and ITS with 16 haplotypes each, as well as 15 *cox*2–3 spacer (Fig. 4d) haplotypes. The remaining DNA markers (*psb*A, *cox*2, *rbc*L and UPA) were not very genetically variable, with each recording fewer than 12 haplotypes. All the haplotype networks except ITS were able to demonstrate the gene genealogy of species based on genetic data (Fig. 4). The inclusion of GenBank sequences increased the number of *cox*1 and *cox*2–3 spacer *Kappaphycus* haplotypes to 33 and 38, respectively (Supplementary figs S12 and S13). For further details on haplotype topology see Supplementary data S2.

Discussion

New insights into eucheumatoid phylogeny and diversity

The multi-marker analyses conducted in this study provided new insights into the taxonomy of *Kappaphycus* and *E. denticulatum*. Genetic data identified the specimens from Arakan, Indonesia (ARW1 and ARW5) as good candidates for a new species. The confirmation of two *K. striatus* subclades (KS1 and KS2) (Tan *et al.*, 2012*b*; Lim *et al.*, 2014), and further delineation of subclade KS1 into KS1a and KS1b indicates that the origins, spatial distributions, and reproductive structures of members of KS1 and KS2 require further investigation to determine if they are conspecific or possibly recently diverged species. This was also observed between *E. denticulatum* 'spinosum' and *E. denticulatum* 'endong/cacing', for which minute genetic differences were recorded despite both being reported from Southeast Asia (Montes *et al.*, 2008; Lim *et al.*, 2014).

All nine genetic markers revealed that the Tanzanian *E. denticulatum* cultivars were genetically identical to the global 'spinosum' cultivar, which supported its commercial introduction from the Philippines to Zanzibar and Tanzania by the academia and private entrepreneurs during the late 1980s (Msuya, 2005; Brakel *et al.*, 2021). The presence of relatively large numbers of wild *E. denticulatum* 'spinosum' specimens and related haplotypes reported in Southeast Asia (present study; Lim *et al.*, 2014; Tan *et al.*, 2022*a*) suggested the species could have originated from Southeast Asia, although the type was initially reported to be from South Africa, but probably in error (Silva *et al.*, 1987; Doty, 1988). These Southeast Asian haplotypes are different from the *cox*2–3 spacer haplotypes 8, 46 and 60 reported from East Africa (Madagascar, Mauritius and Tanzania, respectively) by Zuccarello *et al.* (2006), suggesting rich biodiversity in that region. The detailed characterization of wild populations from Africa would better elucidate the taxonomy of eucheumatoids (e.g. *K. inermis, E. denticulatum, E. horridum, E. nodulosum, E. odontophorum, E. serra* etc.), facilitate germplasm biobanking, conservation and development of native cultivars.

The results indicate that *Kappaphycus* and *Eucheuma* are more diverse than previously reported, and more species and haplotypes are likely to be discovered as taxonomic efforts extend to other areas, namely Vietnam, Indonesia, Australia and Africa (Lim *et al.*, 2014; Hurtado *et al.*, 2016). This was supported by at least 11 novel *K. malesianus cox*1 haplotypes recently detected in East Malaysia and the Philippines (Supplemenatry fig. S12), identifying the region as a hotspot for the species (Tan *et al.*, 2022*a*; Dumilag *et al.*, 2023). However, the results also demonstrate that there are many misidentified specimens in GenBank, and caution is therefore needed in the selection of DNA sequences for analyses and interpretation of results. The reference database generated for each marker in the present study will serve as useful reference for future genetic studies.

The herbarium specimens ARS 02860, ARS 03513 and ARS 08101 from Hawaii are likely more closely related to *K. striatus* (Supplementary figs S10 and S11) and have not been reported elsewhere. While this clade has been reported alongside the commercially introduced and described as invasive *K. alvarezii* and *E. denticulatum* (Conklin & Smith, 2005; Conklin *et al.*, 2009), it appears to be a possibility that this clade is endemic to Hawaii, but has been hidden by the bioinvasion of commercial strains (*K. alvarezii* and *E. denticulatum*) in this location.

Marker performance, species identification and phylogenetic inference

The ability of a marker to distinguish species is attributed to the wide intra- and interspecific genetic divergence gap recorded for these markers (Wiemers & Fiedler, 2007; Freshwater *et al.*, 2010; Tan *et al.*, 2012*a*; Čandek & Kuntner, 2015) whereas the ability of a genetic marker to infer phylogenetic relationships between species is important in taxonomy as it can elucidate the evolutionary pathway of species and thus resolve species complexes, especially for eucheumatoids which are morphologically simple or plastic (Jinbo *et al.*, 2011; Tan *et al.*, 2012*a*). The absence of an overlap in intra- and interspecific genetic divergence, which is an important feature of a DNA barcode, greatly reduces the chances of misidentification (Meyer & Paulay, 2005). More than half of the markers assessed in the present study were capable of species identification, as indicated by ASAP and tree-based delimitations. Inconsistencies in GMYC and mPTP delimitations are common in single-locus data and are often attributed to gene flow, sequence divergence and number of species tested (Luo *et al.*, 2018). Despite this, the ability to reliably identify eucheumatoid species extends beyond taxonomic application and can also improve the efficiency of farm management and carrageenan processing via monoculture (Lim *et al.*, 2017). The mitochondrial *cox*1 and *cox*2–3 spacer are probably most suitable for species identification, considering the large number of sequences for these markers already available in GenBank. However, the potential of the more conserved *cox*2 and *rbc*L in identifying other eucheumatoids, i.e. *Kappaphycopsis, Eucheumatopsis, Betaphycus* and *Mimica*, should also be investigated.

The results also indicated the more rapid evolutionary rates of mitochondrial DNA than chloroplast DNA, which is in line with earlier reports (Brown *et al.*, 1979; Clegg *et al.*, 1994; Cho *et al.*, 2020). Single-marker phylogenetic resolution clearly diminished when more taxa were added into the dataset (Supplementary figs S4–8). A separate phylogenetic analysis of the 630-bp COI-5P (Supplementary fig. S13), a genetic marker commonly used for molecular phylogeny of eucheumatoids, also showed a different tree topology for the 95-sequence dataset, with low interspecific nodal support, and the inability to resolve subclades of *K. striatus*. The results of these individual markers support the use of a multi-locus approach to resolve the evolutionary pathway of a species despite an increase in cost. Analyses of datasets involving two genes have identified COB+*rbc*L (Supplementary fig. S2A), *cox*1+ *rbc*L (Supplementary fig. S2B) and *cox*2 + *rbc*L (Supplementary fig. S2C) as being almost equal in representing the tree topology of the concatenated dataset and were also more accurate than the *cox*1-*cox*2–3 spacer tree (Supplementary fig. S14) commonly used for phylogenetic inference and haplotype analysis of eucheumatoids. However, the relative abundance of *cox*1 and *rbc*L DNA sequences in GenBank likely favours the use of *cox*1+ *rbc*L for species identification and phylogenetic studies of rhodophytes beyond eucheumatoids, as *cox*1 is genetically variable and able to distinguish closely related species, while the relatively conserved nature of *rbc*L favours the inference of interspecific and intergeneric relationships (Saunders, 2005; Robba *et al.*, 2006; Yang *et al.*, 2008; Freshwater *et al.*, 2010; Bartolo *et al.*, 2020). A two-marker system is frequently used, and its cost-efficiency in systematics is valuable especially in Southeast Asia where funding for research and development may be limited (Giam & Wilcove, 2012).

The present study has also demonstrated the limitations of two genetic markers: (i) ITS in distinguishing *Kappaphycus alvarezii* and *K. malesianus*; and (ii) *cox*3 in amplifying and identifying the common 'spinosum' cultivar of *E. denticulatum*. This could explain the identification discrepancy reported in earlier studies (Zhao & He, 2011; Thien *et al.*, 2016), and was consistent with Zhao *et al.* (2013) who concluded that ITS is unsuitable for eucheumatoid identification. Secondary structure reconstruction of the 5.8S genetic region of *K. alvarezii* and *K. malesianus* revealed no genetic difference between the two and that the 5.8S structure is conserved, which agreed with findings by Zhao & He (2011). In angiosperms, phylogenetic incongruence between different markers may indicate the signature of hybridization (Xu *et al.*, 2017). It remains unclear if the ITS result indicated a hybrid between *K. alvarezii* and *K. malesianus* given that the ability of eucheumatoids to form hybrids is not well understood despite hybridization being documented in other rhodophytes (Van Der Meer, 1987; Garbary, 1988; Hwang *et al.*, 2019; Kim, 2021). Nevertheless, the high intraspecific genetic heterogeneity of the ITS region could prove useful in the identification and development of novel species-specific eucheumatoid cultivars.

The poor amplification and misidentification of the *E. denticulatum* 'spinosum' genotype (i.e. specimen MUU29 and KID30; Supplementary table S1) by the *cox*3 marker is likely related to poor primer design and specificity as its limitations are only specific to this genotype in addition to no contamination being detected in the present study.

The amplification and sequencing of the four herbarium specimens from Hawaii benefited from shorter DNA regions, i.e. UPA, partial *rbc*L (750 bp), ITS, *cox*3 and *cox*2–3 spacer, of which the more conserved UPA and *rbc*L were the most reliable. Individually, each marker except ITS was able to identify the *K.* '*striatus*' specimens (ARS 02860, ARS 03513 and ARS 08101) from Hawaii as being different from *K. alvarezii* and *K. striatus*, although not able to confidently infer the phylogenetic relationship between these three taxa. As such, a combination of markers is recommended in the amplification and sequencing of herbarium specimens. The herbarium specimens extracted in the present study were considerably newer than many archival eucheumatoid type specimens up to a century old, and therefore do not truly represent the effectiveness of the aforementioned genetic markers. Consistent with the degradation of DNA in herbarium specimens, Saunders & McDevit (2012) have shown UPA and shortened genetic regions COIms (221 bp of COI-5P) and ITS2r (239–261 bp of ITS2) to be able to amplify DNA from dried red algal specimens of 4–11 years old, but unsuitable for older specimens. However, newer findings have suggested the retrieval of DNA from herbarium samples is practical and reproducible and dependent largely on extraction protocol and DNA purity (Telle & Thines, 2008; Staats *et al.*, 2011; Hughey & Gabrielson, 2012; Särkinen *et al.*, 2012; Marinček *et al.*, 2022). Other studies have also shown genome skimming as an effective method to extract DNA information from herbarium plants (Alsos *et al.*, 2020; Nevill *et al.*, 2020). The ability to retrieve genetic information from eucheumatoid type specimens is a critical step in finally resolving the many taxonomic uncertainties plaguing the group, especially those within the genus *Eucheuma*.

Haplotype diversity

The *cox*1 marker was genetically the most variable and was able to detect 27 eucheumatoid haplotypes based

on the 95-sequence dataset used in the present study. *Cox*1 was also more informative compared with COI-5P, which recorded 17 haplotypes (13 *Kappaphycus* and 4 *E. denticulatum*) (Supplementary fig. S15). A comparison with the *cox*1-*cox*2–3 spacer (Supplementary fig. S16), a combined dataset usually used in the identification of eucheumatoid haplotypes, revealed 32 haplotypes (24 *Kappaphycus* and 8 *E. denticulatum*), which was only 15.6% better than the *cox*1 marker, suggesting the use of only *cox*1 for cost-effectiveness. This was evident based on the detection of a novel and distinct haplogroup of *K. malesianus* (Supplementary fig. S12) when recent *cox*1 DNA sequences from the Philippines were included in the analysis (Dumilag *et al*., 2023).

Over the past two decades, studies of eucheumatoid haplotypes have shifted from genetic diversity, phylogeography and bioinvasion detection (Zuccarello *et al*., 2006; Halling *et al*., 2013; Tan *et al*., 2013; Lim *et al*., 2014; Tano *et al*., 2015; Bast *et al*., 2016; Dumilag *et al*., 2016*b*, 2018; Brakel *et al*., 2021; Roleda *et al*., 2021) towards the identification and selection of new eucheumatoid haplotypes to be developed as replacements for the ageing *K. alvarezii* cultivar (Kumar *et al*., 2020; Roleda *et al*., 2021; Tan *et al*., 2022*b*). As *in vitro* cultivation, studies of hybridization and completion of the eucheumatoid life history are still in their infancy (Brakel *et al*., 2021), initial efforts in the introduction of new cultivars would likely be based on the biobanking, selection and assessment of individuals from wild populations. The ability of *cox*1 and to a certain extent, COI-5P in identifying both farmed and wild haplotypes of *K. alvarezii, K. striatus, K. malesianus* and farmed *Kappaphycopsis cottonii* will likely form the baseline for similar research in the future (Pham, 2002; Lim *et al*., 2014; Dumilag *et al*., 2018; Tan *et al*., 2022*a*, 2022*b*). Nevertheless, a robust and standardized framework for the characterization of eucheumatoid haplotypes is critical for efficiency, especially in (i) the selection criteria for specimens from which the main haplotypes are based; (ii) establishment of a standardized haplotype naming convention; and (iii) standardization of length of gene sequenced for analysis. Additionally, the intensification of germplasm biobanking and cultivar development also necessitates the refinement of guidelines for naming and defining eucheumatoid cultivars which are currently loosely based on the International Code of Nomenclature for Cultivated Plants (ICNCP). An updated guideline should consider the triphasic life cycle of eucheumatoids, as well as the effects of commercial cultivar introduction and bioinvasion.

Future recommendations

Analysis of nine genetic markers has revealed those best suited for identification, phylogeny, haplotype analysis and cost-effectiveness for *Kappaphycus* and *Eucheuma*, which will enable extending their use more widely. There is still a considerable amount of work remaining to resolve the taxonomy and phylogenetics of this commercially important red seaweed group. Future study needs to include *Betaphycus, Kappaphycopsis, Mimica* and several unverified *Kappaphycus* spp. (*Kappaphycus* sp. Hawaii, *K. inermis, Kappaphycus* sp. GUI) to determine the ability of these markers, particularly *cox*1, COB and *rbc*L, to infer intergeneric relationships and resolve the taxonomy of the eucheumatoids. Additional DNA sequences of other *Kappaphycus* and *Eucheuma* species are also critical for the taxonomic redescription of many species which have never been encountered again after their original description; an effort that will require further optimization in methods to extract DNA and sequence from critical herbarium specimens.

Acknowledgements

The authors acknowledge infrastructure and resources support from the Higher Institution Centre of Excellence (HiCOE) Programme (Grant No.: IOES 2014-H and HICoE IOES-2023F) by Ministry of Higher Education Malaysia. Collection of samples from Indonesia was supported by a GlobalSeaweedSTAR research grant (GSS/RF/002, UM Ref. No.: IF059-2019). CRM was partly funded by the UK Natural Environment Research Centre (project NE/R017050/1). The authors thank Mr Japson Wong, Mr Ghazali, Mr Sazali, Ms Nurmaisyamimi and officers of Department of Fisheries Sabah for their helpful assistance during sampling trips in Sabah. The authors also thank the Ministry of Fisheries Fiji, Mrs Mere Lakeba, Mr Jone Waraba and Mr Shalendra Singh for their aid in the sampling of specimens from Fiji.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was part of the GlobalSeaweedSTAR project supported by the UK Research and Innovation – Global Challenges Research Fund (GCRF), Biotechnology and Biological Sciences Research Council Grant 2017 No. [BB/PO27806/1 (UM Ref. No. IF015-2019)]. Part of the study was funded by the Safe Seaweed Coalition SSC 2021 – Project Gateway: LS248520, IF054-2022 [SSC 2021 – Project Gateway – CNRS Ref. No.: LS248520, UM Ref. No.: IF054-2022].

Supplementary material

The following supplementary material is accessible via the Supplementary Content tab on the article's online page at https://doi.org/10.1080/09670262.2024.2396324

Supplementary data S1: Tree-based Assessment.

Supplementary data S2: Phylogenetic trees with GenBank sequences.

Supplementary data S3: Haplotype network topology.
Supplementary data S4:. Phylogeny and haplotype networks based on the commonly used COI-5P and *cox*1 +*cox*2–3 spacer.
Supplementary fig. S1. Bayesian trees of *Kappaphycus* spp. and *Eucheuma denticulatum* based on (A) *psb*A, (B) UPA, and (C) ITS. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. For simplicity, only nodes of main and relevant subclades are annotated. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Red lines represent Operational Taxonomic Units (OTUs). Dotted red lines indicate unresolved species delimitation. KA = *K. alvarezii*, KS = *K. striatus*, KM = *K. malesianus*, Ksp I = *Kappaphycus* sp. (Indonesia), ED = *E. denticulatum*.
Supplementary fig. S2. Bayesian trees of *Kappaphycus* spp. and *Eucheuma denticulatum* based on (A) *cox*1+*rbc*L, (B) COB+*rbc*L, (C) *cox*2+*rbc*L, and (D) *cox*1+COB+*rbc*L. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. For simplicity, only nodes of main and relevant subclades are annotated. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Red lines represent Operational Taxonomic Units (OTUs). Dotted red lines indicate unresolved species delimitation. KA = *K. alvarezii*, KS = *K. striatus*, KM = *K. malesianus*, Ksp I = *Kappaphycus* sp. (Indonesia), ED = *E. denticulatum*.
Supplementary fig. S3. Bayesian trees of *Kappaphycus* spp. and *Eucheuma denticulatum* based on (A) *cox*1+*cox*2, (B) *cox*1+COB, and (C) *cox*2+COB. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. For simplicity, only nodes of main and relevant subclades are annotated. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Red lines represent Operational Taxonomic Units (OTUs). Dotted red lines indicate unresolved species delimitation. KA = *K. alvarezii*, KS = *K. striatus*, KM = *K. malesianus*, Ksp I = *Kappaphycus* sp. (Indonesia), ED = *E. denticulatum*.
Supplementary fig. S4. Bayesian tree of *Kappaphycus, Kappaphycopsis cottonii, Mimica arnoldii* and *Eucheuma denticulatum* based on the mitochondrial *cox*1 DNA marker. Numbers at nodes indicate indicate ultrafast ML bootstrap support and Bayesian posterior probabilities. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Blue lines represent the current taxonomic delineation of species, whereas red and orange lines (GenBank sequences) represent Operational Taxonomic Units (OTUs). Letters in parenthesess indicate cultivar (C) or wild specimen (W).
Supplementary fig. S5. Bayesian tree of common eucheumatoids based on the *cox*2–3 spacer. The codes in bold indicate haplotypes of specimens recorded in the present study. Numbers at nodes indicate indicate ultrafast ML bootstrap support and Bayesian posterior probabilities. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Blue lines represent the current taxonomic delineation of species. Letters in square brackets indicate locality of origin: AF = Africa, BR = Brazil, HW = Hawaii, PP = Philippines, ID = Indonesia, MY = Malaysia, US = America, VT = Vietnam.
Supplementary fig. S6. Bayesian tree of common eucheumatoids based on the ITS. The codes in bold indicate specimens collected in the present study. Codes in red color indicate misidentifications. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Blue lines represent the current taxonomic delineation of species. The alphabet I in parentheses (I) indicate and introduced cultivar. Letters in square brackets indicate locality of origin: CH = China, ID = Indonesia, MY = Malaysia, TZ = Tanzania.
Supplementary fig. S7. Bayesian tree of common eucheumatoids based on the *rbc*L. The codes in bold indicate haplotypes of specimens recorded in the present study. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Blue lines represent the current taxonomic delineation of species. Letters in square brackets indicate locality of origin: AS = Australia, CH = China, PP = Philippines, ID = Indonesia, MX = Mexico, MY = Malaysia, NZ = New Zealand, PP = Philippines, US = America.
Supplementary fig. S8. Bayesian tree of *Kappaphycus* spp., *Betaphycus gelatinus* and *Eucheuma denticulatum* based on the UPA dataset. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. The codes in bold indicate haplotypes of specimens recorded in the present study. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. Blue lines represent the current taxonomic delineation of species. Letters in square brackets indicate locality of origin: HW = Hawaii, ID = Indonesia, IN = India.
Supplementary fig. S9. Bayesian tree of *Kappaphycus* spp. and *Eucheuma denticulatum* based on the *cox*2–3 spacer +*cox*3+*rbc*L+UPA dataset. The code in bold indicates a *Kappaphycus* sp. from Hawaii. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Blue lines represent the current taxonomic delineation of species. Letters in parentheses indicate cultivar (C) or wild /introduced specimen (W/I). Letters in square brackets indicate locality of origin: FJ = Fuji, HW = Hawaii, ID = Indonesia, MY = Malaysia, TZ = Tanzania.
Supplementary fig. S10. Bayesian tree of *Kappaphycus* spp. and *Eucheuma denticulatum* based on the *cox*3+*psb*A+*rbc*L +UPA dataset. The code in *bold* indicates a *Kappaphycus* sp. from Hawaii. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Blue lines represent the current taxonomic delineation of species. Letters in brackets indicate cultivar (C) or wild specimen/ introduced (W/I). Letters in square brackets indicate locality of origin: FJ = Fuji, HW = Hawaii, ID = Indonesia, MY = Malaysia, TZ = Tanzania.
Supplementary fig. S11. Bayesian tree of *Kappaphycus* spp. and *Eucheuma denticulatum* based on the *cox*2+UPA dataset. The code in bold indicates an *E. denticulatum* from Hawaii. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Blue lines represent the current taxonomic delineation of species. Letters in parentheses indicate cultivar (C) or wild /introduced specimen (W/I). Letters in square brackets indicate locality of origin: FJ = Fuji, HW = Hawaii, ID = Indonesia, MY = Malaysia, TZ = Tanzania.
Supplementary fig. S12. Haplotype networks of *Kappaphycus* spp. and *Eucheuma denticulatum* based on *cox*1 including sequences from GenBank. The size of each haplotype circle corresponds to its number of constituting samples. Each line between haplotypes indicates one mutation change. Missing haplotypes are indicated by small grey circles. Colored boxes indicate currently recognized species boundaries based on genetic data. Haplotype networks not drawn to scale.
Supplementary fig. S13. Haplotype networks of *Kappaphycus* spp. and *Eucheuma denticulatum* based on

*cox*1 including sequences from GenBank. The size of each haplotype circle corresponds to its number of constituting samples. Each line between haplotypes indicates one mutation change. Missing haplotypes are indicated by small grey circles. Colored boxes indicate currently recognized species boundaries based on genetic data. Haplotype networks not drawn to scale.

Supplementary fig. S14. Haplotype networks of *Kappaphycus* spp. and *Eucheuma denticulatum* based on (A) *cox*2, (B) *psb*A, (C) *rbc*L and (D) UPA. The size of each haplotype circle corresponds to its number of constituting samples. Each line between haplotypes indicates one mutation change. Missing haplotypes are indicated by small empty circles. Colored boxes indicate currently recognized species boundaries based on genetic data. Haplotype networks not drawn to scale.

Supplementary fig. S15. Bayesian tree of *Kappaphycus* spp. and *Eucheuma denticulatum* based on COI-5P. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. Blue lines represent the current taxonomic delineation of species, whereas red lines represent Operational Taxonomic Units (OTUs). Letters in parentheses indicate cultivar (C) or wild specimen (W). Letters in square brackets indicate locality of origin: ID= Indonesia, MY= Malaysia, TZ= Tanzania.

Supplementary fig. S16. Bayesian tree of *Kappaphycus* spp. and *Eucheuma denticulatum* based on *cox*1+*cox*2–3 spacer. Number at nodes indicates ultrafast ML bootstrap support and Bayesian posterior probabilities. Asterisks indicate UFML ≥ 95%; PP ≥ 0.99. Blue lines represent the current taxonomic delineation of species, whereas red lines represent Operational Taxonomic Units (OTUs). Letters in parentheses indicate cultivar (C) or wild specimen (W). Letters in square brackets indicate locality of origin: ID= Indonesia, MY = Malaysia, TZ = Tanzania.

Supplementary fig. S17. Haplotype network of *Kappaphycus* spp. and *Eucheuma denticulatum* based on the 630 bp COI-5P. The size of each haplotype circle corresponds to its number of constituting samples. Each line between haplotypes indicates one mutation change. Missing haplotypes are indicated by small empty circles. Colored boxes indicate currently recognized species boundaries based on genetic data.

Supplementary fig. S18. Haplotype network of *Kappaphycus* spp. and *Eucheuma denticulatum* based on the *cox*1-*cox*2–3 spacer. The size of each haplotype circle corresponds to its number of constituting samples. Each line between haplotypes indicates one mutation change. Missing haplotypes are indicated by small empty circles. Colored boxes indicate currently recognized species boundaries based on genetic data.

Supplementary table S1. Details of samples used in this study.

Supplementary table S2. Primer details of molecular markers used in this study.

Supplementary table S3. Details of GenBank samples used in this study.

Supplementary table S4. Summary of selected uncorrected p-distance parameters based on the nine genetic markers.

Author contributions

J. Tan: data analysis, drafting and editing manuscript; P.L. Tan: sampling and laboratory work; S.W. Poong: sampling, labwork, data analysis, reviewing manuscript, J. Brakel: labwork, data analysis, reviewing manuscript; C. Rad-Menendez: labwork, reviewing manuscript; E.S. Prasedya: sampling; A.R. Sherwood: herbarium specimens, reviewing manuscript; F.E. Msuya: sampling, reviewing manuscript; C. Gachon: original concept, reviewing manuscript; J. Brodie: original concept, reviewing manuscript; A. Kassim: sampling; P.E. Lim: original concept, reviewing manuscript.

ORCID

Ji Tan http://orcid.org/0000-0002-1264-5442
Pui-Ling Tan http://orcid.org/0000-0002-8654-787X
Sze-Wan Poong http://orcid.org/0000-0003-4033-864X
Janina Brakel http://orcid.org/0000-0002-5880-6280
Cecilia Rad Menendez http://orcid.org/0000-0001-5248-0948
Alison R. Sherwood http://orcid.org/0000-0001-5079-9621
Flower E. Msuya http://orcid.org/0000-0001-6578-6268
Claire Gachon http://orcid.org/0000-0002-3702-7472
Juliet Brodie http://orcid.org/0000-0001-7622-2564
Phaik-Eem Lim http://orcid.org/0000-0001-9186-1487

References

Alsos, I.G., Lavergne, S., Merkel, M.K.F., Boleda, M., Lammers, Y., Alberti, A., Pouchon, C., Denoeud, F., Pitelkova, I., Puşcaş, M., Roquet, C., Hurdu, B.-I., Thuiller, W., Zimmermann, N.E., Hollingsworth, P.M. & Coissac, E. (2020). The treasure vault can be opened: large-scale genome skimming works well using herbarium and silica gel dried material. *Plants (Basel)*, **9**: 432.

Avia, K., Coelho, S.M., Montecinos, G.J., Cormier, A., Lerck, F., Mauger, S., Faugeron, S., Valero, M., Cock, J. M. & Boudry, P. (2017). High-density genetic map and identification of QTLs for responses to temperature and salinity stresses in the model brown alga *Ectocarpus*. *Scientific Reports*, **7**: 1–15.

Bartolo, A.G., Zammit, G., Peters, A.F. & Küpper, F.C. (2020). The current state of DNA barcoding of macroalgae in the Mediterranean Sea: presently lacking but urgently required. *Botanica Marina*, **63**: 253–272.

Bast, F., John, A.A. & Bhushan, S. (2016). Molecular assessment of invasive carrageenophyte *Kappaphycus alvarezii* from India based on ITS-1 sequences. *Webbia*, **71**: 287–292.

Borg, M., Krueger-Hadfield, S.A., Destombe, C., Collén, J., Lipinska, A. & Coelho, S.M. (2023). Red macroalgae in the genomic era. *Phytologist*, **240**: 471–488.

Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., Heled, J., Jones, G., Kühnert, D., De Maio, N. & Matschiner, M. (2019). BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *Plos Computational Biology*, **15**(**4**): e1006650.

Brakel, J., Sibonga, R.C., Dumilag, R.V., Montalescot, V., Campbell, I., Cottier-Cook, E.J., Ward, G., Le Masson, V., Liu, T., Msuya, F.E., Brodie, J., Lim, P.E. & Gachon, C.M. (2021). Exploring, harnessing and conserving marine genetic resources towards a sustainable seaweed aquaculture. *Plants People Planet*, **3**: 337–349.

Brown, W.M., George, M.G., Jr. & Wilson, A.C. (1979). Rapid evolution of animal mitochondrial DNA. *Proceedings of the National Academy of Science USA*, **76**: 1967–1971.

Cai, J., Lovatelli, A., Aguilar-Manjarrez, J., Cornish, L., Dabbadie, L., Desrochers, A., Diffey, S., Gamarro Garrido, E., Geehan, J., Hurtado, A.Q., Lucente, D.,

Mair, G., Miao, W., Potin, P., Przybyla, C., Reantaso, M., Roubach, R., Tauati, M. & Yuan, X. (2021). Seaweeds and Microalgae: An Overview For Unlocking Their Potential In Global Aquaculture Development. In *FAO Fisheries and Aquaculture Circular*. No. 1229. Rome, FAO.

Campo, V.L., Kawano, D.F., da Silva, D.B. & Carvalho, I. (2009). Carrageenans: biological properties, chemical modifications and structural analysis - A review. *Carbohydrate Polymers*, **77**: 167–180.

Čandek, K. & Kuntner, M. (2015). DNA barcoding gap: reliable species identification over morphological and geographical scales. *Molecular Ecology Resources*, **15**: 268–277.

Cho, C.H., In Park, S., Ciniglia, C., Yang, E.C., Graf, L., Bhattacharya, D. & Yoon, H.S. (2020). Erratum: potential causes and consequences of rapid mitochondrial genome evolution in thermoacidophilic *Galdieria* (Rhodophyta). *BMC Evolutionary Biology*, **20**: 1–15.

Clegg, M.T., Gaut, B.S., Learn, G.H. & Morton, B.R. (1994). Rates and patterns of chloroplast DNA evolution. *Proceedings of the National Academy of Science USA*, **91**: 6795–6801.

Clement, M., Posada, D. & Crandall, K.A. (2000). TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**: 1657–1659.

Conklin, E.J. & Smith, J.E. (2005). Abundance and spread of the Invasive Red Algae, *Kappaphycus* spp. in Kane'ohe Bay, Hawaii and an experimental assessment of management options. *Biological Invasions*, **7**: 1029–1039.

Conklin, K.Y., Kurihara, A. & Sherwood, A.R. (2009). A molecular method for identification of the morphologically plastic invasive algal genera *Eucheuma* and *Kappaphycus* (Rhodophyta, Gigartinales) in Hawaii. *Journal of Applied Phycology*, **21**: 691–699.

Doty, M.S. (1988). *Prodomus Ad Systematica Eucheumatoideorum: A Tribe of Commercial Seaweeds Related To Eucheuma (Solieriaceae, Gigartinales)*. In *Taxonomy of Economic Seaweeds*, and Abott, I.A., editor), 47–61. California Sea Grant Program, La Jolla, California.

Doty, M.S. & Norris, J.N. (1985). *Eucheuma Species (Solieriaceae, Rhodophyta) That Are Major Sources of Carrageenan*. In *Taxonomy of Economic Seaweeds: With Reference To Some Pacific and Caribbean Species* (Abott, I.A. & Norris, J.N., eds.), 47–61. California Sea Grant College Program, La Jolla.

Dumilag, R.V., Crisostomo, B.A., Aguinaldo, Z.A., Lluisma, A.O., Gachon, C.M.M. & Roleda, M.Y. (2023). Genetic diversity of *Kappaphycus malesianus* (Solieriaceae, Rhodophyta) from the Philippines. *Aquatic Botany*, **187**: 103659.

Dumilag, R.V., Gallardo, W.G.M., Garcia, C.P.C., You, Y. E., Chaves, A.K.G. & Agahan, L. (2018). Phenotypic and mtDNA variation in Philippine *Kappaphycus cottonii* (Gigartinales, Rhodophyta). *Mitochondrial DNA Part A DNA Mapping, Sequencing and Analysis*, **29**: 951–963.

Dumilag, R.V., Liao, L.M. & Lluisma, A.O. (2014). Phylogeny of *Betaphycus* (Gigartinales, Rhodophyta) as inferred from COI sequences and morphological observations on *B. philippinensis*. *Journal of Applied Phycology*, **26**: 587–595.

Dumilag, R.V. & Lluisma, A.O. (2014). Resolving the phylogenetic affinities of *Kappaphycus inermis* within the genus *Kappaphycus* (Gigartinales, Solieriaceae) using mitochondrial and plastid markers. *Phytotaxa*, **162**: 223–231.

Dumilag, R.V., Orosco, F.L. & Lluisma, A.O. (2016a). Genetic diversity of *Kappaphycus* species (Gigartinales, Rhodophyta) in the Philippines. *Systematics and Biodiversity*, **14**: 441–451.

Dumilag, R.V., Salvador, R.C. & Halling, C. (2016b). Genotype introduction affects population composition of native Philippine *Kappaphycus* (Gigartinales, Rhodophyta). *Conservation Genetic Resources*, **8**: 439–441.

Dumilag, R.V. & Zuccarello, G.C. (2022). Phylogeny and genetic diversity of the Philippine eucheumatoid genus *Mimica* (Solieriaceae, Rhodophyta), and the proposal for *Kappaphycopsis* gen. nov. to include the anomalous species of *Kappaphycus, K. cottonii*. *Phycologia*, **61**(**5**): 496–503.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**(**5**): 1792–1797.

Freshwater, D.W. & Rueness, J. (1994). Phylogenetic relationships of some European *Gelidium* (Gelidiales, Rhodophyta) species, based on *rbc*L nucleotide sequence analysis. *Phycologia*, **33**: 187–194.

Freshwater, D.W., Tudor, K., O'Shaughnessy, K. & Wysor, B. (2010). DNA barcoding in the red algal order Gelidiales: comparison of COI with *rbc*L and verification of the "Barcoding gap." *Cryptogamie Algologie*, **31**: 435–449.

Fujisawa, T. & Barraclough, T.G. (2013). Delimiting species using singlelocus data and the generalized mixed Yule coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology*, **62**: 707–724.

Gachon, C.M.M., Strittmatter, M., Müller, D.G., Kleinteich, J. & Küpper, F.C. (2009). Detection of differential host susceptibility to the marine oomycete pathogen *Eurychasma dicksonii* by real-time PCR: not all algae are equal. *Applied and Environmental Microbiology*, **75**: 322–328.

Garbary, D. (1988). Interoceanic hybridization in fully corticated *Ceramium* isolates (Rhodophyta) from Nova Scotia and Washington. *Algae*, **3**: 89–93.

Gavio, B. & Fredericq, S. (2002). *Grateloupia turuturu* (Halymeniaceae, Rhodophyta) is the correct name of the non-native species in the Atlantic known as *Grateloupia doryphora*. *European Journal of Phycology*, **37**: 349–359.

Geraldino, P.J.L., Yang, E.-C. & S-M, B. (2006). Morphology and molecular phylogeny of *Hypnea flexicaulis* (Gigartinales, Rhodophyta) from Korea. *Algae*, **21**: 417–423.

Giam, X. & Wilcove, D.S. (2012). The geography of conservation ecology research in Southseast Asia: current biases and future opportunities. *Raffles Bulletin of Zoology*, **25**: 29–36.

Halling, C., Wikström, S.A., Lilliesköld-Sjöö, G., Mörk, E., Lundsør, E. & Zuccarello, G.C. (2013). Introduction of Asian strains and low genetic variation in farmed seaweeds: indications for new management practices. *Journal of Applied Phycology*, **25**: 89–95.

Hughey, J.R. & Gabrielson, P.W. (2012). Comment on "Acquiring DNA sequence data from dried archival red algae (Florideophyceae) for the purpose of applying available names to contemporary genetic species: a critical assessment. *Botany*, **90**: 1191–1194.

Hurtado, A.Q., Lim, P.E., Tan, J., Phang, S.M., Neish, I.C. & Critchley, A.T. (2016). Biodiversity and Biogeography of Commercial Tropical Carrageenophytes In the Southeast Asian Region. In *Carrageenans: Sources and Extraction Methods, Molecular Structure, Bioactive*

Properties and Health Effects (Pereira, L., ed.), 51–74. Nova Science Publishers, Inc, New York.

Hwang, E.K., Yotsukura, N., Pang, S.J., Su, L. & Shan, T.F. (2019). Seaweed breeding programs and progress in Eastern Asian countries. *Phycologia*, **58**: 484–495.

Jinbo, U., Kato, T. & Ito, M. (2011). Current progress in DNA barcoding and future implications for entomology. *Entomological Science*, **14**: 107–124.

Kapli, P., Lutteropp, S., Zhang, J., Kobert, K., Pavlidis, P., Stamatakis, A. & Flouri, T. (2017). Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. *Bioinformatics*, **33**(**11**): 1630–1638.

Katoh, K., Rozewicki, J. & Yamada, K.D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Bioinformatics*, **20**: 1160–1166.

Kim, N. (2021). Culture study on the hybrid by interspecific crossing between *Porphyra pseudolinearis* and *P. dentata* (Bangiales, Rhodophyta), two dioecious species in culture. *Algae*, **26**(**1**): 79–86.

Kumar, Y.N., Poong, S.W., Gachon, C., Brodie, J., Sade, A. & Lim, P.E. (2020). Impact of elevated temperature on the physiological and biochemical responses of *Kappaphycus alvarezii* (Rhodophyta). *PLOS ONE*, **15**: 1–16.

Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2018). PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution*, **34**: 772–773.

Lim, P.E., Tan, J., Phang, S.M., Nikmatullah, A., Hong, D. D., Sunarpi, H. & Hurtado, A.Q. (2014). Genetic diversity of *Kappaphycus* Doty and *Eucheuma* J.Agardh (Solieriaceae, Rhodophyta) in Southeast Asia. *Journal of Applied Phycology*, **26**: 1253–1272.

Lim, P.-E., Yang, L.-E., Tan, J., Maggs, C.A. & Brodie, J. (2017). Advancing the taxonomy of economically important red seaweeds (Rhodophyta). *European Journal of Phycology*, **52**(**4**): 438–451.

Luo, A., Ling, C., Ho, S.Y.W. & Zhu, C.D. (2018). Comparison of methods for molecular species delimitation across a range of speciation scenarios. *Systematic Biology*, **67**: 830–846.

Marinček, P., Wagner, N.D. & Tomasello, S. (2022). Ancient DNA extraction methods for herbarium specimens: when is it worth the effort? *Applications In Plant Sciences*, **10**: 1–12.

McCoy, S.J., Krueger-Hadfield, S.A. & Mieszkowska, N. (2020). Evolutionary phycology: toward a macroalgal species conceptual framework. *Journal of Phycology*, **56**: 1404–1413.

Meyer, C.P. & Paulay, G. (2005). DNA Barcoding: error rates based on comprehensive sampling. *PLOS Biology*, **3** (**12**): e422.

Montes, H.R., Pobre, K.F.R. & Lluisma, A.O. (2008). Phylogenetic affiliation of the "Endong"/ "Spaghetti" variety of *Eucheuma* as revealed by molecular data. *Philippine Agricultural Scientist*, **91**: 86–93.

Msuya, F.E. (2005). Seaweed farming in Tanzania: farming processes and interactions between farmers and other stakeholders. *Proceedings of the National Stakeholders Workshop*, Bagamoyo, Tanzania, 195–206.

Nevill, P.G., Zhong, X., Tonti-Filippini, J., Byrne, M., Hislop, M., Thiele, K., Van Leeuwen, S., Boykin, L.M. & Small, I. (2020). Large scale genome skimming from herbarium material for accurate plant identification and phylogenomics. *Plant Methods*, **16**: 1–8.

Pham, H.T. (2002). Morphological Variability of Kappaphycus cottonii In Vietnam. In *Taxonomy of Economic Seaweeds With Reference To Some Pacific Species* (Abbott, I.A. & Mk, J., eds.), Vol. **8**, 247–255. California Sea Grant, La Jolla.

Puillandre, N., Brouillet, S. & Achaz, G. (2019). ASAP: assemble species by automatic partitioning. *Molecular Ecology Resources*, **21**(**2**): 609–620.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, **67**: 901–904.

Robba, L., Russell, S.J., Barker, G.L. & Brodie, J. (2006). Assessing the use of the mitochondrial *cox*1 marker for use in DNA barcoding of red algae (Rhodophyta). *American Journal of Botany*, **93**: 1101–1108.

Roleda, M.Y., Aguinaldo, -Z.-Z.A., Crisostomo, B.A., Hinaloc, L.A.R., Projimo, V.Z., Dumilag, R.V. & Lluisma, A.O. (2021). Discovery of novel haplotypes from wild populations of *Kappaphycus* (Gigartinales, Rhodophyta) in the Philippines. *Algae*, **36**: 1–12.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. (2012). Mrbayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, **61**: 539–542.

Särkinen, T., Staats, M., Richardson, J.E., Cowan, R.S. & Bakker, F.T. (2012). How to open the treasure chest? Optimising DNA extraction from herbarium specimens. *PLOS ONE*, **7**: 1–9.

Saunders, G.W. (2005). Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. *Philosophical Transactions of the Royal Society of London B. Biological Sciences*, **360** (**1462**): 1879–1888.

Saunders, G.W. & Druehl, L.D. (1992). Nucleotide sequences of the small-subunit ribosomal RNA genes from selected Laminariales (Phaeophyta): implications for kelp evolution. *Journal of Phycology*, **28**: 544–549.

Saunders, G.W. & McDevit, D.C. (2012). Acquiring DNA sequence data from dried archival red algae (Florideophyceae) for the purpose of applying available names to contemporary genetic species: a critical assessment. *Botany*, **90**: 191–203.

Saunders, G.W. & Moore, T.E. (2013). Refinements for the amplification and sequencing of red algal DNA barcode and RedToL phylogenetic markers: a summary of current primers, profiles and strategies. *Algae*, **28**: 31–43.

Shan, T.F. & Pang, S.J. (2021). Breeding in the economically important brown algae *Undaria pinnatifida*: a concise review and future prospects. *Frontiers in Genetics*, **12**: 2021.

Sherwood, A.R. & Presting, G.G. (2007). Universal primers amplify a 23S rDNA plastid marker in eukaryotic algae and cyanobacteria. *Journal of Phycology*, **43**: 605–608.

Silva, P.C., Meñez, E.G. & Moe, R.L. (1987). Catalog of the benthic marine algae of the Philippines. *Smithsonian Contributions to the Marine Sciences*, **27**: 1–179.

Staats, M., Cuenca, A., Richardson, J.E., Ginkel, R.V., Petersen, G., Seberg, O. & Bakker, F.T. (2011). DNA damage in plant herbarium tissue. *PLOS ONE*, **6**(**12**): e28448.

Swofford, D.L. (2003). *PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods). Version 4*. Sinauer Associates, Sunderland.

Tan, J., Lim, P.-E., Phang, S.-M., Hong, D.D., Sunarpi, H. & Hurtado, A.Q. (2012a). Assessment of four molecular markers as potential DNA barcodes for red algae *Kappaphycus* Doty and *Eucheuma* J.Agardh (Solieriaceae, Rhodophyta). *PLOS ONE*, **7**(**12**): e52905.

Tan, J., Lim, P.E. & Phang, S.M. (2012b). Phylogenetic relationship of *Kappaphycus* Doty and *Eucheuma* J.Agardh (Solieriaceae, Rhodophyta) in Malaysia. *Journal of Applied Phycology*, **25**: 13–29.

Tan, J., Lim, P.E., Phang, S.M., Rahiman, A., Nikmatullah, A., Sunarpi, H. & Hurtado, A.Q. (2014). *Kappaphycus malesianus* sp. nov.: a new species of *Kappaphycus* (Gigartinales, Rhodophyta) from Southeast Asia. *Journal of Applied Phycology*, **26**: 1273–1285.

Tan, J., Tan, P.L., Poong, S.-W., Brakel, J., Gachon, C., Brodie, J., Sade, A., Kassim, A. & Lim, P.E. (2022a). Genetic differentiation in wild *Kappaphycus* Doty and *Eucheuma* J.Agardh (Solieriaceae, Rhodophyta) from East Malaysia reveals high inter- and intraspecific diversity with strong biogeographic signal. *Journal of Applied Phycology*, **34**: 2719–2733.

Tan, P.L., Poong, S.W., Tan, J., Brakel, J., Gachon, C., Brodie, J., Sade, A. & Lim, P.E. (2022b). Assessment of genetic diversity within eucheumatoid cultivars in east Sabah, Malaysia. *Journal of Applied Phycology*, **34**: 709–717.

Tano, S.A., Halling, C., Lind, E., Buriyo, A. & Wikström, S. A. (2015). Extensive spread of farmed seaweeds causes a shift from native to non-native haplotypes in natural seaweed beds. *Marine Biology*, **162**: 1983–1992.

Telle, S. & Thines, M. (2008). Amplification of *cox*2 (620 bp) from 2 mg of up to 129 years old herbarium specimens, comparing 19 extraction methods and 15 polymerases. *PLOS ONE*, **3**(**10**): e3584.

Thien, V.Y., Yong, W.T.L. & Chin, G.J.W.L. (2016). Morphological and molecular studies of undescribed *Kappaphycus* species. *International Journal of Marine Science*, **6**: 1–7.

Trifinopoulos, J., Nguyen, L., Von, H.A. & Minh, B.Q. (2016). W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research*, **44**: W232–W235.

van Beveren, F., Eme, L., López-García, P., Ciobanu, M. & Moreira, D. (2022). Independent size expansions and intro proliferation in red algal plastid and mitochondrial genomes. *Genome Biology and Evolution*, **14**(**4**): evac037.

Van Der Meer, J.P. (1987). Experimental hybridization of *Palmaria palmata* (Rhodophyta) from the northeast and northwest Atlantic Ocean. *Canadian Journal of Botany*, **65**: 1451–1458.

Ward, G.M., Faisan, J.P., Cottier-Cook, E.J., Gachon, C., Hurtado, A.Q., Lim, P.E., Matoju, I., Msuya, F.E., Bass, D. & Brodie, J. (2020). A review of reported seaweed diseases and pests in aquaculture in Asia. *Journal of the World Aquaculture Society*, **51**(**4**): 815–828.

Ward, G.M., Kambey, C.S.B., Faisan, J.P., Tan, P.L., Daumich, C.C., Matoju, I., Stentiford, G.D., Bass, D., Lim, P.E., Brodie, J. & Poong, S.W. (2022). Ice-Ice disease: an environmentally and microbiologically driven syndrome in tropical seaweed aquaculture. *Reviews in Aquaculture*, **14**: 414–439.

Wiemers, M. & Fiedler, K. (2007). Does the DNA barcoding gap exist? A case study in blue butterflies (Lepidoptera: Lycaenidae). *Frontiers in Zoology*, **4**: 1–16.

Xu, B., Zeng, X.M., Gao, X.F., Jin, D.P. & Zhang, L.B. (2017). ITS non-concerted evolution and rampant hybridization in the legume genus *Lespedeza* (Fabaceae). *Scientific Reports*, **7**: 1–15.

Yang, E.C., Boo, S.M., Bhattacharya, D., Saunders, G.W., Knoll, A.H., Fredericq, S., Graf, L. & Yoon, H.S. (2016). Divergence time estimates and the evolution of major lineages in the Florideophyte red algae. *Scientific Reports*, **6**: 21361.

Yang, E.C., Kim, M.S., Geraldino, P.J.L., Sahoo, D., Shin, J.-A. & Boo, S.M. (2008). Mitochondrial *cox*1 and plastid *rbc*L genes of *Gracilaria vermiculophylla* (Gracilariaceae, Rhodophyta). *Journal of Applied Phycology*, **20**: 161–168.

Yang, H.P., Huo, Y.Z., Yee, K.C. & Yarish, C. (2021). Germplasm cryopreservation of macroalgae for aquaculture breeding and natural resource conservation: a review. *Aquaculture*, **544**: 737037.

Yoon, H.S., Hackett, J.D., Pinto, G. & Bhattacharya, D. (2002). The single, ancient origin of chromist plastids. *Proceedings of the National Academy of Science USA*, **99**: 15507–15512.

Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, **29**: 2869–2876.

Zhao, S. & He, P. (2011). Molecular identification based on ITS sequences for *Kappaphycus* and *Eucheuma* cultivated in China. *Chinese Journal of Oceanology and Limnology*, **29**: 1287–1296.

Zhao, X., Pang, S., Shan, T. & Liu, F. (2013). Applications of three DNA barcodes in assorting intertidal red macroalgal flora in Qingdao, China. *Journal of Ocean University of China*, **12**: 139–145.

Zuccarello, G.C., Burger, G., West, J.A. & King, R.J. (1999). A mitochondrial marker for red algal intraspecific relationships. *Molecular Ecology*, **8**: 1443–1447.

Zuccarello, G.C., Critchley, A.T., Smith, J., Sieber, V., Lhonneur, G.B. & West, J.A. (2006). Systematics and genetic variation in commercial *Kappaphycus* and *Eucheuma* (Solieriaceae, Rhodophyta). *Journal of Applied Phycology*, **18**: 643–651.