ORIGINAL RESEARCH

Assessing species diversity of Coral Triangle artisanal fisheries: A DNA barcode reference library for the shore fishes retailed at Ambon harbor (Indonesia)

Gino Limmon[1] | Erwan Delrieu-Trottin[2,3] | Jesaya Patikawa[1] | Frederik Rijoly[1] |
Hadi Dahruddin[4] | Frédéric Busson[2,5] | Dirk Steinke[6] | Nicolas Hubert[2]

[1]Pusat Kemaritiman dan Kelautan, Universitas Pattimura (Maritime and Marine Science Center of Excellence), Ambon, Indonesia

[2]Institut de Recherche pour le Développement, UMR 226 ISEM (UM-CNRS-IRD-EPHE), Montpellier, France

[3]Museum für Naturkunde, Leibniz-Institut für Evolutions-und Biodiversitätsforschung an der Humboldt-Universität zu Berlin, Berlin, Germany

[4]Division of Zoology, Research Center for Biology, Indonesian Institute of Sciences (LIPI), Cibinong, Indonesia

[5]UMR 7208 BOREA (MNHN-CNRS-UPMC-IRD-UCBN), Muséum National d'Histoire Naturelle, Paris, France

[6]Department of Integrative Biology, Centre for Biodiversity Genomics, University of Guelph, Guelph, ON, Canada

Correspondence
Nicolas Hubert, UMR 5554 ISEM (IRD, UM, CNRS, EPHE), Université de Montpellier, Montpellier Cedex, France.
Email: nicolas.hubert@ird.fr

Abstract

The Coral Triangle (CT), a region spanning across Indonesia and Philippines, is home to about 4,350 marine fish species and is among the world's most emblematic regions in terms of conservation. Threatened by overfishing and oceans warming, the CT fisheries have faced drastic declines over the last decades. Usually monitored through a biomass-based approach, fisheries trends have rarely been characterized at the species level due to the high number of taxa involved and the difficulty to accurately and routinely identify individuals to the species level. Biomass, however, is a poor proxy of species richness, and automated methods of species identification are required to move beyond biomass-based approaches. Recent meta-analyses have demonstrated that species richness peaks at intermediary levels of biomass. Consequently, preserving biomass is not equal to preserving biodiversity. We present the results of a survey to estimate the shore fish diversity retailed at the harbor of Ambon Island, an island located at the center of the CT that display exceptionally high biomass despite high levels of threat, while building a DNA barcode reference library of CT shore fishes targeted by artisanal fisheries. We sampled 1,187 specimens and successfully barcoded 696 of the 760 selected specimens that represent 202 species. Our results show that DNA barcodes were effective in capturing species boundaries for 96% of the species examined, which opens new perspectives for the routine monitoring of the CT fisheries.

KEYWORDS

fisheries assessment, reference library, species delimitation, species diagnostic, Wallacea

1 | INTRODUCTION

Located at the boundary between the Indian and Pacific Ocean, the Coral Triangle (CT) is home to some 4,350 marine fish species (Froese & Pauly, 2014). It encompasses both the Philippine and the Indonesian archipelagoes, constituting the largest diversity anomaly in the world's oceans with local species richness peaking at about 2,500 fish species per 5° × 5° grid cell. This rich biodiversity

is decreasing rapidly (Bellwood & Meyer, 2009; Gaboriau, Leprieur, Mouillot, & Hubert, 2018; Pellissier et al., 2014) as a consequence of ocean warming (Garciá Molinos et al., 2016), overfishing, habitat degradation, and other unsustainable human activities. The future of the CT's biodiversity has received increased attention over the two last decades (Allen, 2008; Cinner et al., 2016; Roberts, 2002). Our ability to anticipate and potentially mitigate biodiversity loss is of prime importance and will serve as an example for our ability to preserve an exceptionally rich area. Species loss in the CT is not only a conservation concern but also poses a food security challenge similar to one faced by most tropical and biodiversity-rich countries (Lal, 2004; Lobell et al., 2008; Schmidhuber & Tubiello, 2008). In a region where fish consumption accounts for more than 20% of the daily animal protein intake and fish supply largely relies on one of the largest fisheries worldwide (FAO, 2018), sustainable management of fish resources is of broad concern. The issue is amplified by the global decline of marine fish stocks, which the current rate, questions both the sustainability of ocean resource harvesting, and the resilience of marine ecosystems (Hughes et al., 2003; Tittensor et al., 2010).

Most large-scale fisheries studies for the Indo-Pacific ocean have focused on linking biomass through space and/or time to varying parameters affecting them (Cinner, Graham, Huchery, & Macneil, 2013; Cinner et al., 2016; Maire et al., 2016). In ecosystems home to thousands of species and tenths of closely related and morphologically similar species, biomass represents a straightforward shortcut to address ecosystem functioning and fisheries dynamics. It can be used to monitor fisheries through time and detect potential stock collapses (Cinner et al., 2016). While this approach allows to document large-scale fisheries trends and to identify regions where those jeopardize food security, it is of limited value for the exploration of ecological dynamics driving species coexistence and ecosystem resilience (Chase & Leibold, 2002; Gravel et al., 2011; Massol et al., 2011; Yachi & Loreau, 1999). In addition, biomass estimates are usually derived from fisheries statistics or underwater visual census through allometric length–weight relationships, further converted into units of mass per area, a proxy that only account for a fraction of the coral reef biodiversity. This level of granularity and the increasing consumer demand for better food traceability requires reliable and routine species level identification of fish products.

DNA barcoding, the use of a 650 base pair sequence of the mitochondrial Cytochrome Oxidase I gene as an internal species tag (Hebert, Cywinska, Ball, & Waard, 2003), provides the necessary ability to assign unknown samples to known species independent of the source of samples. It also offers an unprecedented level of resolution in cases where traditional morphology failed to reliably assign unknowns to knowns and thereby enables a number of applications for marine fisheries and ecosystem research such as the detection of sea food market substitutions (Chin, Adibah, Danial Hariz, & Siti Azizah, 2016; Di Pinto et al., 2015; Holmes, Steinke, & Ward, 2009; Pardo et al., 2018; Wong & Hanner, 2008), the detection of previously unrecorded species (Collet et al., 2018; Kiszka

et al., 2018) or the exploration of early life stage dynamics at the species level (Hubert, Espiau, Meyer, & Planes, 2015; Kimmerling et al., 2018; Steinke, Connell, & Hebert, 2016). The utility of DNA barcoding, however, always depends on the taxonomic coverage of an associated DNA barcode reference library. Recent DNA barcoding studies of reef fish diversity in the Indo-Pacific Ocean unraveled historical taxonomic conflicts, discovered cryptic or unrecognized diversity, and confirmed the need for more complete reference libraries for this region (Durand, Hubert, Shen, & Borsa, 2017; Hubert et al., 2012, 2017; Randall & Victor, 2015; Steinke, Zemlak, & Hebert, 2009; Winterbottom, Hanner, Burridge, & Zur, 2014).

In order to determine the fish diversity harvested by artisanal fisheries in the CT, we set out to build a DNA barcode inventory of the shore fish species retailed at Ambon Island. This approach not only accounts for potential cryptic diversity but also helps to establish a sustainable resource for fisheries monitoring and food traceability. Ambon Island is located in the centre of the CT and has been identified as one of a few bright spots in the Indo-Pacific where biomass is still higher than expected especially when considering overall environmental conditions and socioeconomic drivers of the region (Cinner et al., 2016). Ambon Island can serve as a model for artisanal fisheries in less perturbed parts of the CT, and the diversity of landed fish determined by our study is expected to provide valuable information on fisheries trends elsewhere.

2 | MATERIALS AND METHODS

2.1 | Sampling, morphological identification, and acquisition of DNA barcodes

A total of 1,187 specimens were collected at Ambon harbor between March and December 2016 (Figure 1). Specimens were selected at the harbor based on the morphological diversity available at each time the fish stalls were visited. Provided a species was sampled on day one, the same species was sampled again if encountered later in order to build a comprehensive coverage of the taxonomic diversity retailed. This was done to ensure that closely related species, potentially requiring a more careful examination of diagnostic morphological characters, were not overlooked. Morphological identification was initially done following (Erdmann & Allen, 2012) and subsequently (postsequencing) confirmed through a Barcode Index Number (BIN) discordance report as implemented in the Barcode of Life Datasystem (BOLD—Ratnasingham & Hebert, 2007) in order to spot potential misidentification. Upon detection of discordances, secondary morphological identifications were performed using original species descriptions of the taxa under scrutiny. Specimens were photographed and individually labeled, and voucher specimens were preserved in a 5% formalin solution. A fin clip or a muscle biopsy was taken for each specimen prior to preservation and fixed in a 96% ethanol solution for further genetic analyses. Both tissue and voucher specimens were deposited at the collection of the
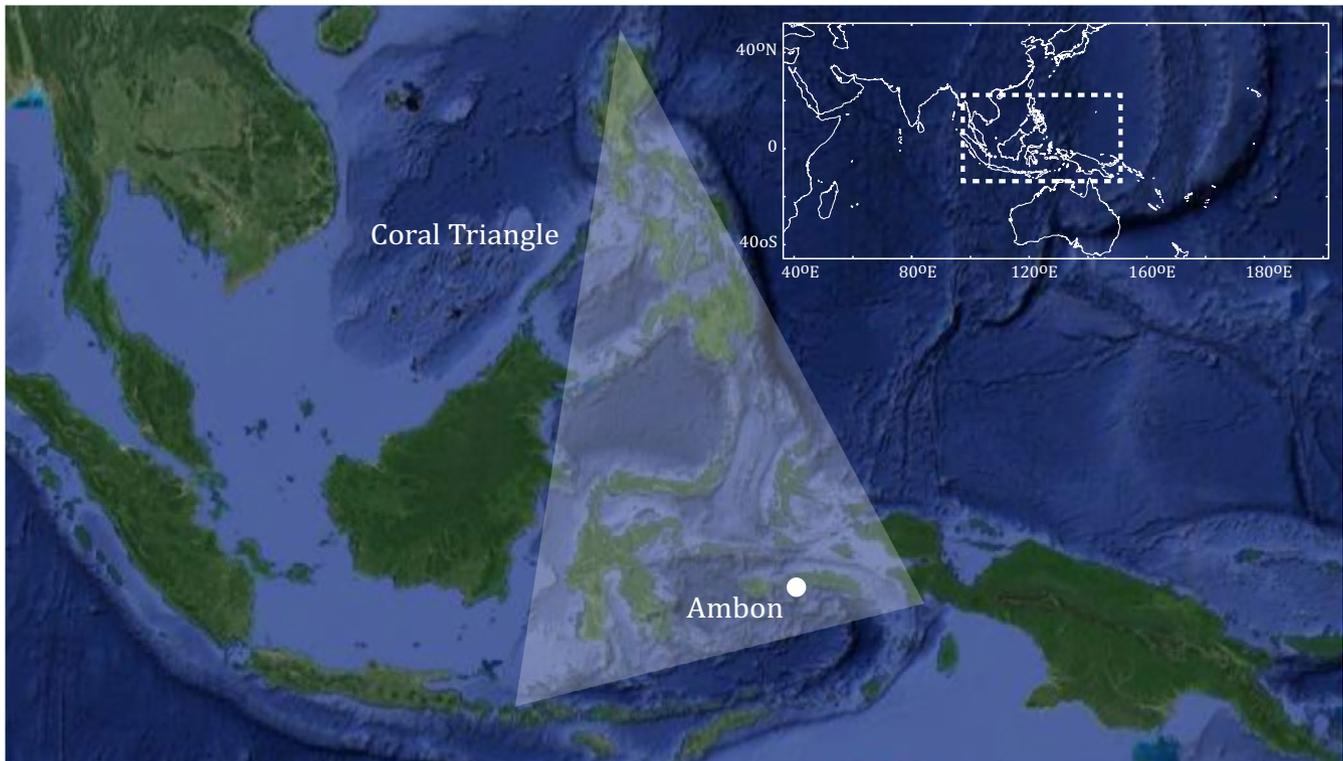
FIGURE 1 Map of the Coral Triangle and Ambon Island

Maritime and Marine Science Center of Excellence at the University of Pattimura.

A total of 760 specimens were selected for sequencing in order to cover as much as possible of the intraspecific genetic diversity by selecting specimens originating from different sampling events in order to validate the identification hypotheses initially produced. Genomic DNA was extracted using a Qiagen DNeasy 96 tissue extraction kit following the manufacturer's specifications. The standard 652-bp segment from the 5' region of the cytochrome oxidase I (COI) was subsequently amplified under the following thermal conditions: Two min at 95°C; 35 cycles of 0.5 min at 94°C, 0.5 min at 52°C, and 1 min at 72°C; 10 min at 72°C; then held at 4°C. The 12.5 µl PCR reaction mixes included 6.25 µl of 10% trehalose, 2.00 µl of ultrapure water, 1.25 µl 10X PCR buffer (200 mM Tris-HCl [pH 8.4], 500 mM KCl), 0.625 µl MgCl (50 mM), 0.125 µl of each primer cocktail (0.01 mM, using primer cocktails C_FishF1t1 and C_FishR1t1 or C_VF1LFt1 and C_VR1LRt1 (Ivanova et al. 2007), 0.062 µl of each dNTP (10 mM), 0.060 µl of Platinum® Taq Polymerase (Invitrogen), and 2.0 µl of DNA template. PCR amplicons were visualized on a 1.2% agarose gel E-Gel® (Invitrogen) and bidirectionally sequenced using sequencing primers M13F or M13R and the BigDye® Terminator v.3.1 Cycle Sequencing kit (Applied Biosystems, Inc.) on an ABI 3730xl capillary sequencer following manufacturer's instructions (for more detail and alternatives see Steinke, Prosser, & Hebert, 2016). Bidirectional sequences were assembled and edited using CodonCode Aligner software (CodonCode Corporation) prior to their upload.

DNA barcodes, photographs, sequences, and collection data were deposited on BOLD in the data set "DNA barcode reference library for the commercial shore fishes of Ambon Island" (https://doi.org/10.5883/DS-BMFAMB). Sequences were also submitted to GenBank, and Accession Numbers are accessible through the BOLD individual records.

2.2 | Species delimitation and genetic distances

DNA sequence divergence was calculated using the Kimura 2-parameter (K2P) model. The midpoint rooted Neighbor-joining (NJ) tree of K2P distances was constructed to provide a graphic representation of the species divergence as implemented in the Sequence Analysis module of BOLD (Ratnasingham & Hebert, 2007). Sequence divergence below and above species boundaries was determined by calculating the maximum intraspecific distance and the distance to the closest phylogenetic neighbor in the data set.

Four sequence-based methods were used for species delimitation. From now on, species identified based on traditional morphological characters will be referred to as species while species delimited by DNA sequences will be referred to as Operational Taxonomic Unit (OTU), which represent diagnosable molecular lineages (Avise, 1989; Moritz, 1994; Vogler & DeSalle, 1994). The species delimitation methods used are based on different assumptions but they all focus on the detection of transition points between mutation/drift (within species) and speciation/extinction (between species) dynamics (Hubert & Hanner, 2015). Each method is susceptible to pitfalls which is why our final delimitation scheme was based on a 50% consensus among methods in order to produce a robust delimitation

(Kekkonen & Hebert, 2014; Kekkonen, Mutanen, Kaila, Nieminen, & Hebert, 2015). OTUs were delimited using the following algorithms: (a) Refined Single Linkage (RESL) as implemented in BOLD and used to produce Barcode Index Numbers (BIN) (Ratnasingham & Hebert, 2013), (b) Automatic Barcode Gap Discovery (ABGD) (Puillandre, Lambert, Brouillet, & Achaz, 2012), (c) Poisson Tree Process (PTP) in its multiple rates version as implemented in the stand-alone software mptp_0.2.3 (Kapli et al., 2017; Zhang, Kapli, Pavlidis, & Stamatakis, 2013), and (d) General Mixed Yule-Coalescent (GMYC) in its multiple rates version as implemented in the R package Splits 1.0-19 (Fujisawa & Barraclough, 2013). The mPTP algorithm uses a phylogenetic tree as an input file; thus, a maximum likelihood (ML) tree was first reconstructed using RAxML (Stamatakis, Ludwig, & Meier, 2005) based on a GTR + Γ substitution model. An ultrametric and fully resolved tree was reconstructed using a Bayesian approach implemented in BEAUti and BEAST 2.4.8 (Bouckaert et al., 2014) to be further used for OTU delimitation using the mGMYC algorithm. Two Markov chains of 50 millions each were run independently using a strict-clock model, based on a canonical 1.2% of genetic divergence per million years (Bermingham, McCafferty, & Martin, 1997), and with a HKY + I + Γ substitution model. Trees were sampled every 10,000 states after an initial burnin period of

10 million, and both runs were combined using LogCombiner 2.4.8 (Bouckaert et al., 2014). A maximum credibility tree was constructed using TreeAnnotator 2.4.7 (Bouckaert et al., 2014). Duplicated sequences were pruned prior to the Bayesian analysis.

3 | RESULTS

We were able to obtain 696 sequences representing 202 species, 73 genera and 24 families from 760 sampled individuals (92%). Amplification failures were randomly distributed among species, and all species were successfully sequenced. Sequence length for all barcodes was >600 bp and no codon stops were detected suggesting that these sequences correspond to functional coding regions. Of the 24 families sampled, the families Serranidae (groupers), Lutjanidae (Snappers) displayed the highest species richness with 31 and 26 species, respectively (Figure 2a). Both families were followed by eight additional families (Nemipteridae, Scarida, Holocentridae, Labridae, Acanthuridae, Siganidae, Lethrinidae, and Mullidae) with a species richness larger than ten species (Figure 2a). Of all genera obtained, the Snapper genus *Lutjanus* displayed the highest species richness with 18 species, followed by *Siganus* with 13 species, the
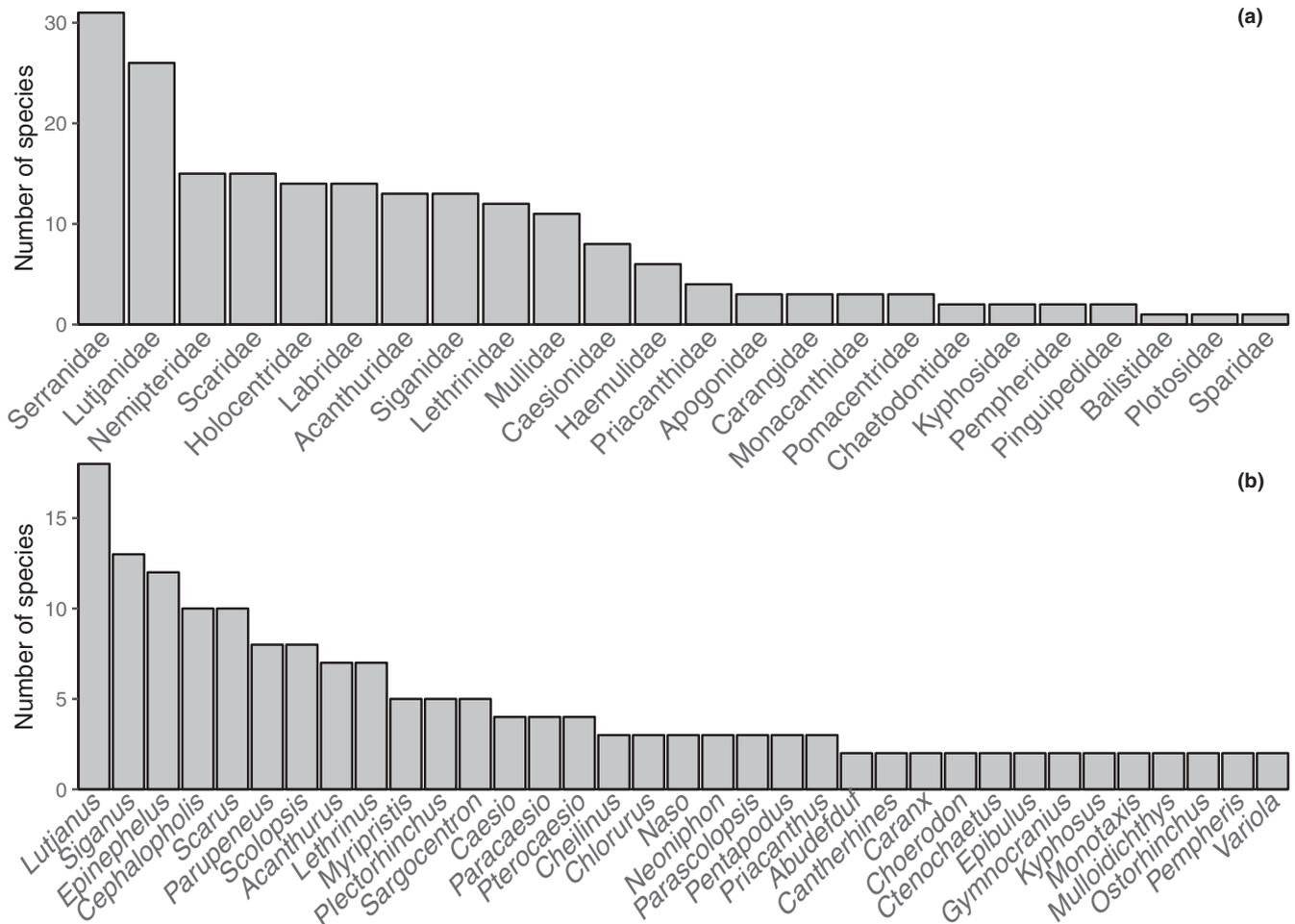

FIGURE 2 Ranking of families (a) and genera (b) according to the number of species observed is this survey

groupers *Epinephelus* and *Cephalopholis* with 12 and 10 species, respectively (Figure 2b).

Intraspecific distances ranged from 0% to 7.12% with an average of 0.32% while interspecific distances within genus ranged from 0% to 21.83% (average 12.05%; Table 1). The distributions of the maximum intraspecific distance and the distance to the nearest neighbor overlapped (Figure 3a,b); however, nearest neighbor distances were 30-fold higher on average than maximum intraspecific distances (Table S2). Maximum intraspecific distances exceeded nearest neighbor distance only for six species *Siganus canaliculatus* (7.12%), *Lethrinus ornatus* (3.86%), *Siganus vermiculatus* (3.64%), *Lutjanus decussatus* (3.64%), *Nemipterus furcosus* (2.36%), and *Paracaesio xanthura* (2.21%) and a barcoding gap was generally observed (Figure 3c) (Table S2 and Figure S1). By contrast, K2P distances to the nearest neighbor <2% were observed for eleven species including six cases of haplotype sharing (*Lutjanus decussatus*, *L. semicinctus*, *Siganus canaliculatus*, *S. fuscescens*, *S. guttatus*, and *S. lineatus*) and five cases of low K2P distances (*Siganus vermiculatus*, *S. corallinus*, *S. doliatus*, *Pterocaesio capricornis*, and *P. tile*).

The OTU delimitation analyses yielded varying numbers of OTUs depending on the algorithm used (Figure 4 and Table S3). Numbers of delimited OTUs were 206 for RESL, 217 for ABGD, 160 for mPTP, and 216 for mGMYC (Figure 4 and Table S3). The consensus delimitation scheme yielded 208 OTUs for 202 nominal species, highlighting several conflicts between OTUs and species delimitation (Figure 4 and Table S3). A total of eight species with multiple OTUs were detected with K2P genetic distances ranging from 1.72 in *Priacanthus hamrur* to 3.86 in *Lethrinus ornatus* (Table 2). Three OTUs, however,

were shared by more than one species (OTU192 including *S. analiculatus* and *S. fuscescens*, OTU92 including *L. decussatus* and *L. semicinctus*, and OTU196 including *Siganus. guttatus*, *S. lineatus*, and *S. vermiculatus*) for a total of seven species displaying mixed genealogies (Figure S1 and Table S2).

4 | DISCUSSION

The usefulness of DNA barcoding to identify fish species is widely accepted with nearly 90% of the species analyzed having their boundaries aptly captured by DNA barcodes in previous studies (April, Mayden, Hanner, & Bernatchez, 2011; Hubert et al., 2008; Knebelsberger, Dunz, Neumann, & Geiger, 2015; Shen et al., 2019; Weigt et al., 2012). In the present study, only six cases of shared haplotypes between species were detected. Hence, 97% of the species analyzed (196 out of 202) were accurately delimitated by their DNA barcodes. Only 4% of the species with more than one individual sampled showed maximum intraspecific K2P genetic distances lower than the distance to the nearest neighbor (6 out of 153); however, these species were characterized by collections of private haplotypes indicating that they can readily by identified through DNA barcodes (Collins & Cruickshank, 2014). This trend was confirmed by the ratio between the maximum intraspecific and the nearest neighbor genetic distances, the latter being 30-fold higher on average than the former. This ratio varies substantially between fish studies (7–30) largely depending on the spatial scope and taxonomic coverage (April et al., 2011; Hubert et al., 2017; Pereira, Pazian,

TABLE 1 Summary statistics of the genetic distances (K2P) with increasing taxonomic levels

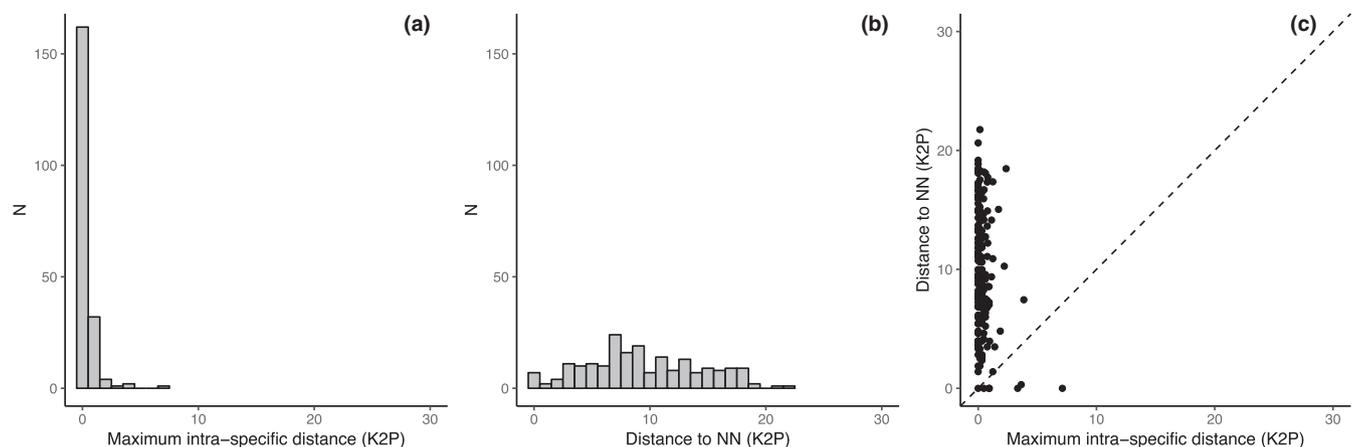| Label	n	Taxa	Comparisons	Min dist(%)	Mean dist(%)	Max dist(%)
Within species	643	153	1,268	0.00	0.32	7.12
Within genus	601	35	7,555	0.00	12.05	21.83
Within family	608	17	9,203	7.21	18.28	26.15

FIGURE 3 Distribution of genetic distances below and above species boundaries. (a) Distribution of maximum intraspecific distances (K2P). (b) Distribution of nearest neighbor distances (K2P). (c) Relationship between maximum intraspecific and nearest neighbor distances. Points above the diagonal line indicate species with a barcode gap
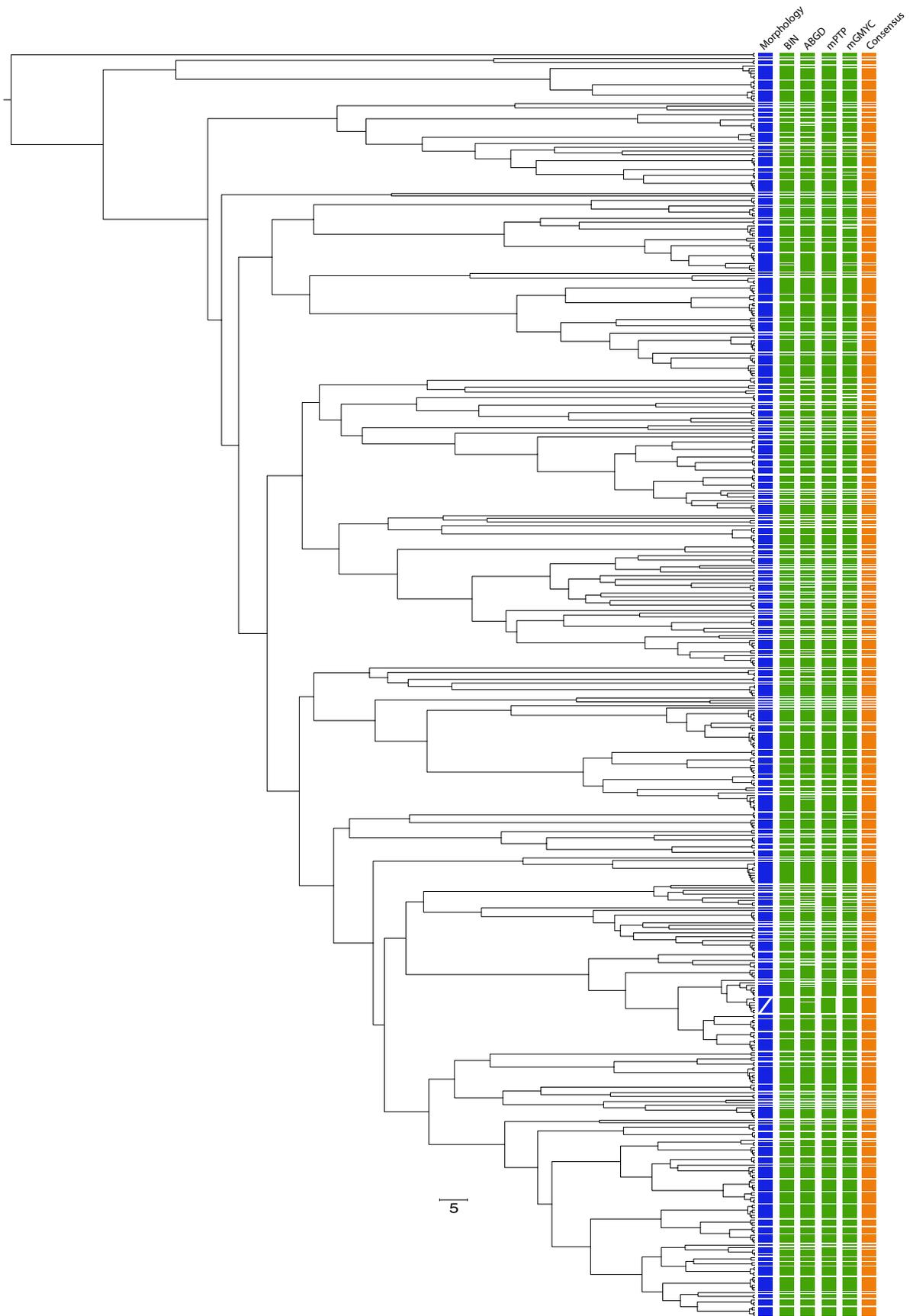
FIGURE 4 Bayesian Chronogram for the 676 DNA barcodes obtained based on a 1.2% of genetic divergence per Million years including species delimitation scheme based on morphological characters (blue), DNA-based species delimitation schemes derived from RESL, ABGD, mPTP and mGMYC (green), and the consensus delimitation scheme (orange)

TABLE 2 Summary statistics of the 16 OTUs among the 8 species with more than a single OTU including their BIN, maximum intraspecific distance and distance to the nearest neighbor

	Dist. Max. intra.	Dist. Near. Neigh.
Lethrinus ornatus	3.86	7.45
OTU81 (BIN:AAK9232)	0.8	3.86
OTU82 (BIN:AAE7226)	—	3.86
Lutjanus decussatus	3.34	0
OTU91 (BIN:AAF0336)	0	3.34
OTU92 (BIN:ADF5202)	0.15	0
Nemipterus furcosus	2.36	18.47
OTU117 (BIN:ACE4155)	0.46	2.36
OTU118 (BIN:ADF2810)	0.46	2.36
Paracaesio xanthura	2.21	10.27
OTU131 (BIN:AAI6821)	—	2.21
OTU132 (BIN:ABZ2393)	0.33	2.21
Plectorhinchus lineatus	1.87	4.81
OTU152 (BIN:AAF2980)	0.31	1.87
OTU153 (BIN:AAF2980)	—	1.87
Priacanthus hamrur	1.72	15.05
OTU158 (BIN:AAB1643)	0.78	1.72
OTU159 (BIN:AAB1643)	—	1.72
Siganus vermiculatus	3.64	0.31
OTU197 (BIN:ABB2342)	—	0.31
OTU202 (BIN:AAJ6703)	—	3.64
Siganus caniculatus	7.12	0
OTU193 (BIN:AAB8846)	0	0
OTU194 (BIN:ACR6962)	—	7.12

Hanner, Foresti, & Oliveira, 2011). The ratio recovered for this study is high indicating that most species consist of tightly defined clusters of DNA barcodes. It may decrease with increasing spatial and taxonomic scale given the high levels of intraspecific genetic distances that have been previously reported for Indo-Pacific reef fishes at the regional scale (Hubert et al., 2012, 2017; Steinke et al., 2009; Winterbottom et al., 2014). This trend for Indo-Pacific reef fishes is due to high levels of cryptic diversity especially when species with distribution ranges spanning across the Indian and Pacific ocean are included (Hubert et al., 2017; Zemlak, Ward, Connell, Holmes, & Hebert, 2009). This study confirms, however, that at the local scale, DNA barcoding—less prone to inflated intraspecific genetic distances due to taxonomic uncertainties and cryptic diversity—can be successfully applied to automated identifications of fishes. We detected several cases of cryptic diversity in 4% of the species under scrutiny (16 OTUs delineated within 8 species)' by 'We detected several cases of cryptic diversity (16 OTUs delineated within 8 species), representing 4% of the species under scrutiny. This rate is lower than in previous assessments at regional scale in the Indian and Pacific oceans where 10% of species contained cryptic and

highly divergent lineages (Hubert et al., 2012, 2017). Several cases of shared haplotypes were also detected within *Lutjanus* and particularly *Siganus* species. Those potentially result from introgressive hybridization and/or overlapping morphological characters. Several cases were expected as mitochondrial introgression was previously reported for *Siganus* based on a comprehensive assessment of genetic diversity at mitochondrial and nuclear markers (Kuriiwa, Hanzawa, Yoshino, Kimura, & Nishida, 2007; Ravago-Gotanco, Cruz, Josefa Pante, & Borsa, 2018). In such cases that constitute a limit where sister-species co-occur and hybridize (Ravago-Gotanco et al., 2018), mitochondrial markers will inflate the number of false positives (i.e. considering heterospecifics as conspecifics). The amount of false positives detected does not exceed 4% of the specimens analyzed suggesting that, if cautiously interpreting the individual assignment for those species, automated identification can be performed (Collins & Cruickshank, 2014).

This study highlights that artisanal fisheries in the CT harvest a substantial amount of species that originate from a diversity of shore habitats including coral reefs (Serranidae, Holocentridae, Scaridae, Pomacentridae, and Chaetodontidae), seaweed beds (Siganidae, Mullidae, and Labridae), and open waters (Carangidae). The retailed coral reef families account for most of the species richness, for example, the Serranidae (groupers), Acanthuridae (Surgeon fish), and Holocentridae (Soldier fish) that rank first, fifth, and seventh in species number collected, respectively. On the other hand, species-rich coral reef families such as Pomacentridae and Chaetodontidae were rarely encountered at the fish stalls and only a few species were collected for each family. Our result suggests a market trend toward species from the families Serranidae (*Epinephelus* and *Cephalopholis*) and Lutjanidae (*Lutjanus*) as both rank first and second in terms of number of species landed, respectively. This trend is of particular concern for the sustainability of the CT fisheries as groupers (Serranidae) have already been identified as severely threatened by overfishing (Sadovy de Mitcheson et al., 2013). Both families (with 138 Serranidae and 59 Lutjanidae species) are also not among the most species-rich in the Indonesian archipelago (Froese & Pauly, 2014), which amplifies the pressure through increased fisheries.

This study confirms the potential of DNA barcoding for the automated identification of Indo-Pacific shore fishes (Delrieu-Trottin et al.,2019; Durand et al., 2017; Hubert et al., 2017, 2015, 2012; Jaafar, Taylor, Mohd Nor, Bruyn, & Carvalho, 2012; Randall & Victor, 2015; Steinke et al., 2009; Ward, Costa, Holmes, & Steinke, 2008; Winterbottom et al., 2014) and opens new perspectives for the monitoring of CT fisheries. By providing the first reference library available to date for commercial shore fishes of the CT, this study enables the DNA-based assignment of unknown individuals to known species. In extremely diversified biomes, identification of fish species through morphological characters is a difficult task that relies on a few specialists worldwide. As a consequence of the worldwide loss of taxonomists, Indo-Pacific shore fish taxonomy has largely became a black box for many groups (Hubert & Hanner, 2015). This development might have also led to the biomass-based approach to the study of shore fisheries trends in the Indo-Pacific

(Cinner et al., 2013, 2016; Maire et al., 2016). While biomass is certainly a useful proxy for determining fisheries trends and their impact on ecosystems dynamics, it has its limits, particularly in the context of extreme species richness as observed in the CT. Biomass and species richness are not correlated for the Indo-Pacific as a whole. Biomass peaks at peripheral areas of the CT while species richness peaks for intermediate levels of biomass, a trend that has been suggested based on mathematical models (Mouillot & Mouquet, 2006). This trend shows that ecological dynamics for most species-rich ecosystems are complex and likely rely on a large array of drivers including their biogeographical history, community assembly dynamics and persistence through time (Gaboriau et al., 2018; Gaither & Rocha, 2013; Hubert et al., 2012; Pellissier et al., 2014). Thus, monitoring ecosystem recovery by measuring biomass might be misleading given that species richness is not directly linked to it but rather associated with a few productive species. In this context, this DNA barcodes reference library opens new possibilities for the monitoring of CT artisanal fisheries by enabling the assessment of their dynamics at the species level and further allows environmental DNA (eDNA) approaches in the CT. Alternative molecular markers have been used for eDNA purposes such as 12S and 16S (Miya et al., 2015), Cytochrome b (Ficetola, Miaud, Pompanon, & Taberlet, 2008), and COI (Hajibabaei, Shokralla, Zhou, Singer, & Baird, 2011) during the last decade. We advocate here for the use of standardized procedures in DNA barcoding along others (Andújar, Arribas, Yu, Vogler, & Emerson, 2018) and recommend the use of the COI gene for eDNA considering the extensive coverage of the COI libraries available (BOLD), the appropriate rate of substitution of this protein-coding gene for species diagnostic and the development of new sequencing protocols that avoid PCR-based enrichment (Mariac et al., 2018).

5 | CONCLUSIONS

This study provides the first DNA barcode reference library of the CT shore fishes targeted by artisanal fisheries. It contains 202 species including collection and sequence data for 696 specimens. Species boundaries were aptly captured by DNA barcodes in 96% of the species examined, which demonstrates the effectiveness of DNA barcodes for further automated identification of unknown specimens. This library opens new perspectives for monitoring of artisanal fisheries in the CT and enables species level surveys. Considering the extreme species richness of the CT marine ecosystems and the difficulties to accurately and routinely identify species, this library allows further detailed studies in fisheries and coastal management.

CONFLICT OF INTEREST

None declared.

AUTHORS' CONTRIBUTIONS

GL and NH designed the study. GL, JP, and FR conducted the market survey. GL, JP, HD, DS, and NH conducted the genetic analyzes. GL, JP, FR, EDT, and NH performed the morphological identifications. GL, EDT, and NH analyzed the data. EDT and FB curated the DNA barcode records in BOLD. NH and GL wrote the initial draft of the manuscript, and all authors commented and approved the final version of the manuscript.

DATA AVAILABILITY STATEMENT

All collecting and sequence data are available on the Barcode of Life Datasystem (BOLD) in the project "Barcoding Indonesian Fishes - DNA Barcoding of the marine fishes landed at Ambon harbour" and are also available on Genbank (see Table S1).

ORCID

Erwan Delrieu-Trottin (iD) https://orcid.org/0000-0002-4120-9316
Dirk Steinke (iD) https://orcid.org/0000-0002-8992-575X
Nicolas Hubert (iD) https://orcid.org/0000-0001-9248-3377

REFERENCES

Allen, G. R. (2008). Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. *Aquatic Conservation: Marine and Freshwater Ecosystems*, *18*, 541–556. https://doi.org/10.1002/aqc.880

Andújar, C., Arribas, P., Yu, D. W., Vogler, A. P., & Emerson, B. C. (2018). Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, *27*(20), 3968–3975. https://doi.org/10.1111/mec.14844

April, J., Mayden, R. L., Hanner, R. H., & Bernatchez, L. (2011). Genetic calibration of species diversity among North America's freshwater fishes. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 10602–10607. https://doi.org/10.1073/pnas.1016437108

Avise, J. C. (1989). *Molecular markers, natural history and evolution*. New York, NY: Springer.

Bellwood, D. R., & Meyer, C. P. (2009). Searching for heat in a marine biodiversity hotspot. *Journal of Biogeography*, 36, 569–576. https://doi.org/10.1111/j.1365-2699.2008.02029.x

Bermingham, E., McCafferty, S., & Martin, A. P. (1997). Fish biogeography and molecular clocks: Perspectives from the Panamanian isthmus. In T. D. Kocher, & C. A. Stepien (Eds.), *Molecular systematics of fishes* (pp. 113–128). San Diego, CA: Academic Press.

Bouckaert, R. R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Chase, J. M., & Leibold, M. A. (2002). Spatial scale dictates the productivity-biodiversity relationship. *Nature*, 416, 427–430. https://doi.org/10.1038/416427a

Chin, T. C., Adibah, A. B., Danial Hariz, Z. A., & Siti Azizah, M. N. (2016). Detection of mislabelled seafood products in Malaysia by DNA barcoding: Improving transparency in food market. *Food Control*, 64, 247–256. https://doi.org/10.1016/j.foodcont.2015.11.042

Cinner, J. E., Graham, N. A. J., Huchery, C., & Macneil, M. A. (2013). Global effects of local human population density and distance to markets on the condition of coral reef fisheries. *Conservation Biology*, 27, 453–458. https://doi.org/10.1111/j.1523-1739.2012.01933.x

Cinner, J. E., Huchery, C., MacNeil, M. A., Graham, N. A. J., McClanahan, T. R., Maina, J., … Mouillot, D. (2016). Bright spots among the world's coral reefs. *Nature*, 535(7612), 416–419. https://doi.org/10.1038/nature18607

Collet, A., Durand, J. D., Desmarais, E., Cerqueira, F., Cantinelli, T., Valade, P., & Ponton, D. (2018). DNA barcoding post-larvae can improve the knowledge about fish biodiversity: An example from La Reunion, SW Indian Ocean. *Mitochondrial DNA Part A*, 29(6), 905–918. https://doi.org/10.1080/24701394.2017.1383406

Collins, R. A., & Cruickshank, R. H. (2014). Known knowns known unknowns unknown unknowns and unknown knowns in DNA barcoding: A comment on Dowton et al *Systematic Biology*, 63(6), 1005–1009. https://doi.org/10.1093/sysbio/syu060

Delrieu-Trottin, E., Williams, J. T., Pitassy, D., Driskell, A., Hubert, N., Viviani, J., & Planes, S. (2019). A DNA barcode reference library of French Polynesian shore fishes. *Scientific Data*, 6(1), 114. https://doi.org/10.1038/s41597-019-0123-5

Di Pinto, A., Marchetti, P., Mottola, A., Bozzo, G., Bonerba, E., Ceci, E., … Tantillo, G. (2015). Species identification in fish fillet products using DNA barcoding. *Fisheries Research*, 170, 9–13. https://doi.org/10.1016/j.fishres.2015.05.006

Durand, J.-D., Hubert, N., Shen, K.-N., & Borsa, P. (2017). DNA barcoding grey mullets. *Reviews in Fish Biology and Fisheries*, 27(1), 233–243. https://doi.org/10.1007/s11160-016-9457-7

Erdmann, M., & Allen, G. R. (2012). *Reef fishes of the East Indies*. Perth, Australia: UH Press.

FAO (2018). *La situation mondiale des pêches et de l'aquaculture 2018. Atteindre les objectifs de développement durable*. Rome.

Ficetola, G. F., Miaud, C., Pompanon, F., & Taberlet, P. (2008). Species detection using environmental DNA from water samples. *Biology Letters*, 4, 423–425. https://doi.org/10.1098/rsbl.2008.0118

Froese, R., & Pauly, D. (2014). *FishBase*. http://www.fishbase.org, electronic version Accessed 11/2019.

Fujisawa, T., & Barraclough, T. G. (2013). Delimiting species using single-locus data and the generalized mixed yule coalescent approach: A revised method and evaluation on simulated data sets. *Systematic Biology*, 62(5), 707–724. https://doi.org/10.1093/sysbio/syt033

Gaboriau, T., Leprieur, F., Mouillot, D., & Hubert, N. (2018). Influence of the geography of speciation on current patterns of coral reef fish biodiversity across the Indo-Pacific. *Ecography*, 40, 1295–1306. https://doi.org/10.1111/ecog.02589

Gaither, M. R., & Rocha, L. A. (2013). Origins of species richness in the Indo-Malay-Philippine biodiversity hotspot: Evidence for the centre of overlap hypothesis. *Journal of Biogeography*, 40, 1638–1648. https://doi.org/10.1111/jbi.12126

Garciá Molinos, J., Halpern, B. S., Schoeman, D. S., Brown, C. J., Kiessling, W., Moore, P. J., … Burrows, M. T. (2016). Climate velocity and the future global redistribution of marine biodiversity. *Nature Climate Change*, 6, 83–88. https://doi.org/10.1038/nclimate2769

Gravel, D., Bell, T., Barbera, C., Bouvier, T., Pommier, T., Venail, P., & Mouquet, N. (2011). Experimental niche evolution alters the strength of the diversity-productivity relationship. *Nature*, 469, 89–94. https://doi.org/10.1038/nature09592

Hajibabaei, M., Shokralla, S., Zhou, X., Singer, G. A. C., & Baird, D. J. (2011). Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PLoSONE*, 6, e17497. https://doi.org/10.1371/journal.pone.0017497

Hebert, P. D. N., Cywinska, A., Ball, S. L., & de Waard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270, 313–321. https://doi.org/10.1098/rspb.2002.2218

Holmes, B. H., Steinke, D., & Ward, R. D. (2009). Identification of shark and ray fins using DNA barcoding. *Fisheries Research*, 95, 280–288. https://doi.org/10.1016/j.fishres.2008.09.036

Hubert, N., Dettai, A., Pruvost, P., Cruaud, C., Kulbicki, M., Myers, R. F., & Borsa, P. (2017). Geography and life history traits account for the accumulation of cryptic diversity among Indo-West Pacific coral reef fishes. *Marine Ecology Progress Series*, 583, 179–193. https://doi.org/10.3354/meps12316

Hubert, N., Espiau, B., Meyer, C., & Planes, S. (2015). Identifying the ichthyoplankton of a coral reef using DNA barcodes. *Molecular Ecology Resources*, 15(1), 57–67. https://doi.org/10.1111/1755-0998.12293

Hubert, N., & Hanner, R. (2015). DNA barcoding, species delineation and taxonomy: A historical perspective. *DNA Barcodes*, 3, 44–58. https://doi.org/10.1515/dna-2015-0006

Hubert, N., Hanner, R., Holm, E., Mandrak, N. E., Taylor, E., Burridge, M., … Bernatchez, L. (2008). Identifying Canadian freshwater fishes through DNA barcodes. *PLoS ONE*, 3(6), e2490. https://doi.org/10.1371/journal.pone.0002490

Hubert, N., Meyer, C. P., Bruggemann, H. J., Guérin, F., Komeno, R. J. L., Espiau, B., … Planes, S. (2012). Cryptic diversity in indo-pacific coral-reef fishes revealed by DNA-barcoding provides new support to the centre-of-overlap hypothesis. *PLoS ONE*, 7(3), e28987. https://doi.org/10.1371/journal.pone.0028987

Hughes, T. P., Baird, A. H., Bellwood, D. R., Card, M., Connolly, S. R., Folke, C., … Roughgarden, J. (2003). Climate change, human impacts, and the resilience of coral reefs. *Science*, 301(5635), 929–933. https://doi.org/10.1126/science.1085046

Ivanova, N. V., Zemlak, T. S., Hanner, R. H., & Hebert, P. D. (2007). Universal primer cocktails for fish DNA barcoding. *Molecular Ecology Notes*, 7(4), 544–548. https://doi.org/10.1111/j.1471-8286,2007.01748.x

Jaafar, T. N. A. M., Taylor, M. I., Mohd Nor, S. A., De Bruyn, M., & Carvalho, G. R. (2012). DNA barcoding reveals cryptic diversity within commercially exploited Indo-Malay carangidae (Teleosteii: Perciformes). *PLoS ONE*, 7, e49623. https://doi.org/10.1371/journal.pone.0049623

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, 1630–1638.

Kekkonen, M., & Hebert, P. D. N. (2014). DNA barcode-based delineation of putative species: Efficient start for taxonomic workflows. *Molecular Ecology Resources*, 14, 706–715. https://doi.org/10.1111/1755-0998.12233

Kekkonen, M., Mutanen, M., Kaila, L., Nieminen, M., & Hebert, P. D. N. (2015). Delineating species with DNA barcodes: A case of

taxon dependent method performance in moths. *PLoS ONE*, *10*(4), e0122481. https://doi.org/10.1371/journal.pone.0122481

Kimmerling, N., Zuqert, O., Amitai, G., Gurevich, T., Armoza-Zvuloni, R., Kolesnikov, I., … Sorek, R. (2018). Quantitative species-level ecology of reef fish larvae via metabarcoding. *Nature Ecology and Evolution*, *2*, 306–316. https://doi.org/10.1038/s41559-017-0413-2

Kiszka, J. J., Mouillot, D., Bonnin, L., Manel, S., Kulbicki, M., Boussarie, G., … Mariani, S. (2018). Environmental DNA illuminates the dark diversity of sharks. *Science Advances*, *4*, eaap9661. https://doi.org/10.1126/sciadv.aap9661

Knebelsberger, T., Dunz, A. R., Neumann, D., & Geiger, M. F. (2015). Molecular diversity of Germany's freshwater fishes and lampreys assessed by DNA barcoding. *Molecular Ecology Resources*, *15*, 562–572. https://doi.org/10.1111/1755-0998.12322

Kuriiwa, K., Hanzawa, N., Yoshino, T., Kimura, S., & Nishida, M. (2007). Phylogenetic relationships and natural hybridization in rabbitfishes (Teleostei: Siganidae) inferred from mitochondrial and nuclear DNA analyses. *Molecular Phylogenetics and Evolution*, *45*, 69–80. https://doi.org/10.1016/j.ympev.2007.04.018

Lal, R. (2004). Soil carbon sequestration impacts on global climate change and food security. *Science*, *304*(5677), 1623–1627. https://doi.org/10.1126/science.1097396

Lobell, D. B., Burke, M. B., Tebaldi, W. P., Mastrandrea, C., Falcon, M. D., & Naylor, R. L. (2008). Needs for food security in 2030 region. *Science*, *319*, 607–610.

Maire, E., Cinner, J., Velez, L., Huchery, C., Mora, C., Dagata, S., … Mouillot, D. (2016). How accessible are coral reefs to people? A global assessment based on travel time. *Ecology Letters*, *19*, 351–360. https://doi.org/10.1111/ele.12577

Mariac, C., Vigouroux, Y., Duponchelle, F., García-Dávila, C., Nunez, J., Desmarais, E., & Renno, J. F. (2018). Metabarcoding by capture using a single COI probe (MCSP) to identify and quantify fish species in ichthyoplankton swarms. *PLoS ONE*, *13*(9), 1–15. https://doi.org/10.1371/journal.pone.0202976

Massol, F., Gravel, D., Mouquet, N., Cadotte, M., Fukami, T., & Leibold, M. (2011). Linking community and ecosystem dynamics through spatial ecology. *Ecology Letters*, *14*, 313–323. https://doi.org/10.1111/j.1461-0248.2011.01588.x

Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., … Iwasaki, W. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: Detection of more than 230 subtropical marine species. *Royal Society Open Science*, *2*, 150088.

Moritz, C. (1994). Defining "Evolutionary Significant Units" for conservation. *Trends in Ecology and Evolution*, *9*, 373–375.

Mouillot, D., & Mouquet, N. (2006). Species richness peaks for intermediate levels of biomass in a fractal succession with quasi-neutral interactions. *Oikos*, *115*, 349–357. https://doi.org/10.1111/j.2006.0030-1299.14894.x

Pardo, M. Á., Jiménez, E., Viðarsson, J. R., Ólafsson, K., Ólafsdóttir, G., Daníelsdóttir, A. K., & Pérez-Villareal, B. (2018). DNA barcoding revealing mislabeling of seafood in European mass caterings. *Food Control*, *92*, 7–16. https://doi.org/10.1016/j.foodcont.2018.04.044

Pellissier, L., Leprieur, F., Parravicini, V., Cowman, P. F., Kulbicki, M., Litsios, G., … Mouillot, D. (2014). Quaternary coral reef refugia preserved fish diversity. *Science*, *344*(6187), 1016–1019. https://doi.org/10.1126/science.1249853

Pereira, L. H. G., Pazian, M. F., Hanner, R., Foresti, F., & Oliveira, C. (2011). DNA barcoding reveals hidden diversity in the Neotropical freshwater fish *Piabina argentea* (Characiformes: Characidae) from the Upper Parana Basin of Brazil. *Mitochondrial DNA*, *22*, 87–96.

Puillandre, N., Lambert, A., Brouillet, S., & Achaz, G. (2012). ABGD, automatic barcode gap discovery for primary species delimitation. *Molecular Ecology*, *21*, 1864–1877. https://doi.org/10.1111/j.1365-294X.2011.05239.x

Randall, J. E., & Victor, B. C. (2015). Descriptions of thirty-four new species of the fish genus *Pempheris* (Perciformes: Pempheridae), with a key to the species of the western Indian Ocean. *Journal of Ocean Science Foundation*, *18*, 1–77.

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes*, *7*, 355–364.

Ratnasingham, S., & Hebert, P. D. N. (2013). A DNA-based registry for all animal species: The barcode index number (BIN) system. *PLoS ONE*, *8*, e66213. https://doi.org/10.1371/journal.pone.0066213

Ravago-Gotanco, R., de la Cruz, T. L., Josefa Pante, M., & Borsa, P. (2018). Cryptic genetic diversity in the mottled rabbitfish Siganus fuscescens with mitochondrial introgression at a contact zone in the South China Sea. *PLoS ONE*, *13*, 1–27. https://doi.org/10.1371/journal.pone.0193220

Roberts, C. M. (2002). Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science*, *295*(5558), 1280–1284. https://doi.org/10.1126/science.1067728

Sadovy de Mitcheson, Y., Craig, M. T., Bertoncini, A. A., Carpenter, K. E., Cheung, W. W. L., Choat, J. H., … Heemstra, P. C. (2013). Fishing groupers towards extinction: a global assessment of threats and extinction risks in a billion dollar fishery. *Fish and Fisheries*, *14*(2), 119–136. https://doi.org/10.1111/j.1467-2979.2011.00455.x

Schmidhuber, J., & Tubiello, F. N. (2008). Global food security under climate change. *Proceedings of the National Academy of Sciences of United States of America*, *104*, 19703–19708. https://doi.org/10.1073/pnas.0701976104

Shen, Y., Hubert, N., Huang, Y., Wang, X., Gan, X., Peng, Z., & He, S. (2019). DNA barcoding the ichthyofauna of the Yangtze River: Insights from the molecular inventory of a mega-diverse temperate fauna. *Molecular Ecology Resources*, *19*, 1278–1291. https://doi.org/10.1111/1755-0998.12961

Stamatakis, A., Ludwig, T., & Meier, H. (2005). RAxML-II: A program for sequential, parallel and distributed inference of large phylogenetic trees. *Concurrency and Computation: Practice and Experience*, *17*, 1705–1723. https://doi.org/10.1002/cpe.954

Steinke, D., Connell, A. D., & Hebert, P. D. N. (2016). Linking adults and immatures of South African marine fishes. *Genome*, *59*(11), 959–967. https://doi.org/10.1139/gen-2015-0212

Steinke, D., Prosser, S. W. J., & Hebert, P. D. N. (2016). DNA barcoding of marine metazoans. In S. J. Bourlat (Ed.), *Marine genomics: Methods and protocols* (pp. 155–168). New York, NY: Springer.

Steinke, D., Zemlak, T. S., & Hebert, P. D. N. (2009). Barcoding nemo: DNA-based identifications for the ornamental fish trade. *PLoS ONE*, *4*, e6300. https://doi.org/10.1371/journal.pone.0006300

Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., & Worm, B. (2010). Global patterns and predictors of marine biodiversity across taxa. *Nature*, *466*, 1098–1103. https://doi.org/10.1038/nature09329

Vogler, A. P., & DeSalle, R. (1994). Diagnosing units of conservation management. *Conservation Biology*, *6*, 170–178. https://doi.org/10.1046/j.1523-1739.1994.08020354.x

Ward, R. D., Costa, F. O., Holmes, B. H., & Steinke, D. (2008). DNA barcoding of shared fish species from the North Atlantic and Australasia: Minimal divergence for most taxa, but *Zues faber* and *Lepidopus caudatus* each probably constitute two species. *Aquatic Biology*, *3*, 71–78.

Weigt, L. A., Baldwin, C. C., Driskell, A., Smith, D. G., Ormos, A., & Reyier, E. A. (2012). Using DNA barcoding to assess caribbean reef fish biodiversity: Expanding taxonomic and geographic coverage. *PLoS ONE*, *7*, e41059. https://doi.org/10.1371/journal.pone.0041059

Winterbottom, R., Hanner, R., Burridge, M., & Zur, M. (2014). A cornucopia of cryptic species – A DNA barcode analysis of the gobiid genus Trimma (Percomorpha, Gobiiformes). *ZooKeys*, *381*, 79–111.

Wong, E. H. K., & Hanner, R. H. (2008). DNA barcoding detects market substitution in North American seafood. *Food Research International*, *41*, 828–837. https://doi.org/10.1016/j.foodres.2008.07.005

Yachi, S., & Loreau, M. (1999). Biodiversity and ecosystem productivity in a fluctuating environment: The insurance hypothesis. *Proceedings of the National Academy of Sciences of the United States of America*, *96*, 1463–1468. https://doi.org/10.1073/pnas.96.4.1463

Zemlak, T. S., Ward, R. D., Connell, A. D., Holmes, B. H., & Hebert, P. D. N. (2009). DNA barcoding reveals overlooked marine fishes. *Molecular Ecology Resources*, *9*, 237–242. https://doi.org/10.1111/j.1755-0998.2009.02649.x

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. *Bioinformatics*, *29*, 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Limmon G, Delrieu-Trottin E, Patikawa J, et al. Assessing species diversity of Coral Triangle artisanal fisheries: A DNA barcode reference library for the shore fishes retailed at Ambon harbor (Indonesia). *Ecol Evol*. 2020;10:3356–3366. https://doi.org/10.1002/ece3.6128