# Global-Scale Diversity and Distribution Characteristics of Reef-Associated Symbiodiniaceae *via* the Cluster-Based Parsimony of Internal Transcribed Spacer 2 Sequences

GONG Sanqiang<sup>1), 2), 3)</sup>, CHAI Guangjun<sup>1)</sup>, SUN Wei<sup>1)</sup>, ZHANG Fengli<sup>1)</sup>, YU Kefu<sup>4)</sup>, and LI Zhiyong<sup>1), \*</sup>

1) Marine Biotechnology Laboratory, State Key Laboratory of Microbial Metabolism and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai 200240, China

2) Key Laboratory of Tropical Marine Bioresources and Ecology and Guangdong Provincial Key Laboratory of

Applied Marine Biology, South China Sea Institute of Oceanology, Guangzhou 510301, China

3) Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China

4) Coral Reef Research Center of China, Guangxi University, Nanning 530004, China

(Received October 22, 2019; revised July 6, 2020; accepted September 8, 2020) © Ocean University of China, Science Press and Springer-Verlag GmbH Germany 2021

**Abstract** Photosynthetic species belonging to Symbiodiniaceae play important ecological roles in coral reef ecosystems. Nine phylogenetically supported Symbiodiniaceae clades (each containing multiple genetically distinct molecular types) have been detected in reef-associated animals and in seawater or sediments near coral reef areas. Although major advances in research on Symbiodiniaceae diversity have been made, information on the global-scale diversity and distribution characteristics of Symbiodiniaceae clades/types is still limited or must be expanded because most studies on Symbiodiniaceae have mainly focused on a limited number of hosts associated with Symbiodiniaceae or coral reef areas. Here, the diversity and distribution characteristics of Symbiodiniaceae data derived from reef-associated animals and seawater or sediments near coral reef areas were analyzed using 3899 qualified Symbiodiniaceae internal transcribed spacer 2 (ITS2) sequences retrieved from the GenBank database in May 2017. These sequences were annotated as Symbiodiniaceae ITS2 types *via* sequence alignment against database of Symbiodiniaceae ITS2 types. Results showed that photosynthetic Symbiodiniaceae ITS2 sequences could be annotated as 119 Symbiodiniaceae ITS2 types within nine clades (A–I), and Symbiodiniaceae ITS2 types mainly inhabited coral reefs between 30°N and 30°S. Symbiodiniaceae ITS2 types was distinct in the Atlantic Ocean, the Red Sea, the Mediterranean Sea, and the Indo-Pacific Ocean. Although future field research should be performed to verify our results, this study provided a broad perspective of the diversity of Symbiodiniaceae clades/types on a global scale.

Key words Symbiodiniaceae; diversity; reef-dwelling animals; coral reefs; global scale

# **1** Introduction

Photosynthetic Symbiodiniaceae plays important roles in the establishment of the most diverse and productive coral reef ecosystems in oligotrophic marine waters (Coffroth *et al.*, 2006). Symbiodiniaceae is mainly associated with metazoans (*e.g.*, reef-building corals, non-reef-building corals, flatworms, and sponges) and foraminifera by providing them nutrients. In turn, these reef-associated metazoans and foraminifera create a microenvironment and supply some metabolic substrates that allow Symbiodiniaceae to survive at high densities (Blackall *et al.*, 2015). Previous studies revealed the environmental factors that influence the diversity, geographic distribution, and mutual association of Symbiodiniaceae with different hosts (Baird *et al.*, 2009; Takabayashi *et al.*, 2012; Yamashita and Koike, 2013). The breakdown of Symbiodiniaceae– coral associations contributes to remarkable declines in coral reef covers worldwide (Jones, 2008; De'Ath *et al.*, 2012). Consequently, the diversity and geographic distribution of Symbiodiniaceae should be monitored to predict the fate of corals and coral reef ecosystems under marine environmental changes.

Initially, Symbiodiniaceae members were considered a single pandemic species based on morphological analyses (Freudenthal, 1962). Subsequently, our understanding of the diversity of Symbiodiniaceae was significantly changed because of sequence-based molecular analyses (Baker, 2003; Coffroth and Santos, 2005). At present, Symbiodiniaceae taxa are divided into nine phylogenetically supported clades

<sup>\*</sup> Corresponding author. Tel: 0086-21-34204036 E-mail: zyli@sjtu.edu.cn

(A – I), and several Symbiodiniaceae clades have been formally described as a genus based on their genetic and ecological characteristics (*Symbiodinium* [clade A], *Breviolum* [clade B], *Cladocopium* [clade C], *Durusdinium* [clade D], *Effrenium* [clade E], *Fugacium* [clade F], and *Gerakladium* [clade G]; Lajeunesse *et al.*, 2018). Each clade or genus of Symbiodiniaceae includes a number of informally described molecular types (Pochon *et al.*, 2014; Lajeunesse *et al.*, 2018).

The diversity and geographic distribution of Symbiodiniaceae taxa and their associations with reef-associated metazoans and foraminifera have been mainly explored with different molecular markers, *e.g.*, nuclear large subunit (LSU), internal transcribed spacer regions (ITS2 and ITS1), chloroplast large subunit (*cp23s*), cytochrome oxidase b, and noncoding chloroplast *psbA* region (LaJeunesse *et al.*, 2011). However, the ITS2 region currently remains the most commonly employed DNA marker for assessing the diversity of Symbiodiniaceae ITS2 types (Franklin *et al.*, 2012; Tonk *et al.*, 2013; Arif *et al.*, 2014; Green *et al.*, 2014; Prada *et al.*, 2014; Quigley *et al.*, 2014; Thomas *et al.*, 2014; Cunning *et al.*, 2015; Reimer *et al.*, 2017; Ziegler *et al.*, 2017).

Baker (2003) and van Oppen *et al.* (2009) reviewed the global-scale diversity and geographic distribution of Symbiodiniaceae among corals at the clade level and suggested that Symbiodiniaceae clades A-D are common to corals living in the Atlantic Ocean, whereas Symbiodiniaceae clades C and D are dominant in corals inhabiting the tropical Indo-Pacific Ocean.

Lajeunesse (2002, 2003), and Lajeunesse et al. (2005) analyzed the geographic distributions of Symbiodiniaceae types within clades C and B in cnidarians (especially reefbuilding corals) living between the Atlantic Ocean and the Indo-Pacific Ocean at the type level. Franklin et al. (2012) built a global Symbiodiniaceae type data set with an online interface for users to explore and utilize global geospatial and ecological data on Symbiodiniaceae-host symbioses. Tonk et al. (2013) reviewed studies on the diversity of Symbiodiniaceae in invertebrates from the Great Barrier Reef. In addition, previous studies revealed that corals can change or exchange their symbiotic Symbiodiniaceae communities, and Symbiodiniaceae living in seawater or sediments near coral reef areas likely serve as an open pool from which corals can acquire algae when they are under environmental stress (Fautin and Buddemeier, 2004; Keshavmurthy et al., 2014; Hughes et al., 2017). By contrast, studies have reported stable Symbiodiniaceae communities in corals, such as Porites-Symbiodiniaceae associations, which are stable in the Indo-Pacific in broad geographical ranges and temperature gradients and through bleaching events or after transplantation (Baker, 2003).

Furthermore, physiological differences in Symbiodiniaceae clades or types have been described. For example, certain Symbiodiniaceae clades or types are exceptionally thermotolerant (*e.g.*, D1, D1–4, C15, and A3), whereas others are thermosensitive (*e.g.*, C3, C7, B17, and A13; Stat *et al.*, 2013; Tonk, 2013; Hughes *et al.*, 2017). These data suggest that the diversity and distribution of Symbiodiniaceae may depend on biotic (*e.g.*, host species) and abiotic (*e.g.*, temperature, light, nutrients, and geographic affinities) factors (Baker, 2003; Prada *et al.*, 2014).

Although advances in the diversity and distribution of Symbiodiniaceae have been made, global-scale knowledge on Symbiodiniaceae clades/types is still limited because most previous studies focused on a limited number of hosts associated with Symbiodiniaceae or coral reef areas. In addition, Baker's work (2003) should be expanded. Baker (2003) provided a global perspective of the diversity and distribution of Symbiodiniaceae based on nuclear LSU gene sequences of reef-building corals. However, the diversity of Symbiodiniaceae has been analyzed in reef-building corals and other metazoans (e.g., non-reefbuilding corals, flatworm sand sponges), foraminifera, and seawater or sediments near coral reef areas mainly based on the Symbiodiniaceae ITS2 region (Franklin et al., 2012; Tonk et al., 2013; Arif et al., 2014; Green et al., 2014; Prada et al., 2014; Quigley et al., 2014; Thomas et al., 2014; Cunning et al., 2015; Reimer et al., 2017; Ziegler et al., 2017).

This study aimed to provide a broad global perspective of the diversity and distribution of Symbiodiniaceae by annotating the Symbiodiniaceae ITS2 sequences derived from reef-associated metazoans and foraminifera and from seawater or sediments over 28 coral reef areas worldwide in the GenBank database as Symbiodiniaceae ITS2 types *via* sequence alignment against database of Symbiodiniaceae ITS2 types in accordance with previously described methods (Arif *et al.*, 2014).

### 2 Materials and Methods

### 2.1 Retrieval of Symbiodiniaceae ITS2 Sequences from GenBank

We carried out a systematic search in the NCBI nucleotide database in order to obtain ITS2 sequences of Symbiodiniaceae ITS2 in May 2017 using the string '(Symbiodiniaceae and ITS2) and (Symbiodiniaceae and internal transcribed spacer 2)'. The GenBank-formatted ITS2 sequences that matched the string were downloaded from the NCBI nucleotide database and saved as ITS2.GB, which was used as raw data source for the subsequent analyses. Only the Symbiodiniaceae ITS2 sequences derived from the samples of different hosts, seawater, or sediment (including reef sand) were included. In addition, the term 'host-associated Symbiodiniaceae data' referred to Symbiodiniaceae ITS2 sequences derived from hosts (host information from GenBank), and the term 'free-living Symbiodiniaceae data' corresponded to Symbiodiniaceae ITS2 sequences obtained from seawater or sediment samples.

### 2.2 Quality Control and Bioinformatic Analysis of Symbiodiniaceae ITS2 Sequences

Symbiodiniaceae ITS2 sequences from the NCBI nucleotide database were subjected to quality control and bioinformatic analysis with modified pipeline and scripts provided by Arif *et al.* (2014). In brief, Symbiodiniaceae ITS2 sequences were removed if they had a length of <150 nucleotides (nt), contained >6 ambiguous bases, or had a homopolymer length of >6. Chimeric sequences were identified and removed using chimera.uchime and remove.seqs scripts against a database of Symbiodiniaceae ITS2 types (two databases supplied by Arif *et al.*, 2014, and Cunning *et al.*, 2017, were merged).

Symbiodiniaceae ITS2 sequences were first clustered into different clades by applying an 85% cutoff value with the get.oturep script (Arif et al., 2014) and BLASTn (Altschul et al., 1990) against the custom database of Symbiodiniaceae ITS2 types because of difficulties in aligning these sequences. Symbiodiniaceae ITS2 sequences belonging to different clades were aligned separately with MUS-CLE. The aligned Symbiodiniaceae ITS2 sequences were further clustered into operational taxonomic unit (OTU) at a 97% cutoff value. The OTU that represented fewer than three Symbiodiniaceae ITS2 sequences was removed. The most abundant sequence from each OTU was chosen as a representative sequence and annotated as a Symbiodiniaceae ITS2 type with BLASTn (Altschul et al., 1990) against the custom database of Symbiodiniaceae ITS2 types based on 97% similarity.

The number of Symbiodiniaceae ITS2 sequences of different Symbiodiniaceae ITS2 types was used to record the diversity of hosts or coral reef areas from where Symbiodiniaceae ITS2 sequences were derived.

Rarefaction curve analysis was performed using past3 software (http://folk.uio.no/ohammer/past/) to investigate whether the number of Symbiodiniaceae ITS2 sequences exceeded the point where the rarefaction curve reached its asymptote.

The annotated ITS2 sequences of Symbiodiniaceae types were aligned with MUSCLE in MEGA 6.06 (Tamura *et al.*, 2013) and trimmed to have an equal length (152 bp). Phylogenetic trees of ITS2 sequences were constructed with maximum likelihood (ML) and neighbor joining (NJ) methods in MEGA 6.06. The reliability of internal branches was assessed with the bootstrap method with 1000 replicates. The NJ trees were nearly identical to the corresponding ML trees; thus, only the ML trees were addressed further in this study.

The number of Symbiodiniaceae ITS2 sequences in different Symbiodiniaceae ITS2 clades/types among different hosts or coral reef areas was used to perform a principal component analysis (PCA) of the community structure *via* Canoco software for Windows 4.5 (Microcomputer Power, NY, USA). The PCA results were assessed with an MCPP test in Canoco for Windows 4.5 in accordance with the manufacturer's instructions (Ter Braak and Smilauer, 2002).

### 3 Results

### 3.1 Diversity of Symbiodiniaceae Based on ITS2 Sequences

In this study, a total of 3899 Symbiodiniaceae ITS2 sequences were obtained after they were subjected to qua-

lity control and further annotated as 119 Symbiodiniaceae ITS2 types within nine clades (Fig.1). Most Symbiodiniaceae ITS2 sequences belonged to Symbiodiniaceae ITS2 types within clades A, B, C, D, and F, accounting for 96% of the ITS2 sequences and 90% of all Symbiodiniaceae ITS2 types. Specifically, Symbiodiniaceae ITS2 types within clade C showed the highest abundance (2305 ITS2 sequences) and diversity, representing 59% of the total ITS2 sequences and 53% of all the types. Clades A-I had 14, 9, 63, 5, 2, 16, 5, 3, and 2 Symbiodiniaceae ITS2 types, respectively. The diversity of Symbiodiniaceae ITS2 types was further analyzed on the basis of rarefaction curve analysis. The results showed that the number of Symbiodiniaceae ITS2 sequences deposited in GenBank exceeded the point where the rarefaction curve reached its asymptote, but the rarefaction curves based on hosts or seawater samples did not reach their asymptotes.



Fig.1 Number of Symbiodiniaceae ITS2 sequences (a) and number of Symbiodiniaceae ITS2 types (b) in different clades (A–I). The number of Symbiodiniaceae ITS2 sequences was designated among Symbiodiniaceae ITS2 sequences with different hosts and/or habitats information. The Symbiodiniaceae clades and types were clustered and annotated with an operational taxonomic unit-based method against the database of Symbiodiniaceae ITS2 types.

# 3.2 Diversity of Hosts Associated with Symbiodiniaceae

The number of Symbiodiniaceae ITS2 sequences and the host information corresponding to the sequences indicated that the majority (91.0%) of ITS2 sequences belonging to Symbiodiniaceae clades A – I were derived from metazoans and foraminifera (Fig.2), including Cnidaria (69.1%; reef-building corals, soft corals, jellyfishes, and anemones), Mollusca (11.0%; Tridacna, bivalves, and nudibranchs), Porifera (0.4%; sponges), and Foraminifera (9.7%). Specifically, the ITS2 sequences belonging to Symbiodiniaceae clades A, B, C, and D were mainly derived from reef-building corals; the ITS2 sequences belonging to Symbiodiniaceae clades F, H, and I were mainly derived from protists (foraminifera); the ITS2 sequences belonging to Symbiodiniaceae clades G and A were mainly derived from Porifera; and the ITS2 sequences belonging to Symbiodiniaceae clades A, B, C, D, and E were derived from anemones. These results suggested that the hosts associated with Symbiodiniaceae were diverse.

In addition, approximately 9.0% of the Symbiodiniaceae ITS2 sequences belonging to Symbiodiniaceae clades A, B, C, D, E, and G were derived from the seawater or sediment samples collected near coral reef areas (Fig.2).

### 3.3 Global-Scale Geographic Distribution of Symbiodiniaceae

A broad global perspective of the diversity and distribution of the nine Symbiodiniaceae clades was obtained on the basis of the number of Symbiodiniaceae ITS2 sequences belonging to different Symbiodiniaceae clades and the corresponding reef area information (Fig.3). The results showed that the Symbiodiniaceae ITS2 sequences belonging to different clades were mainly derived from 28 coral reef areas across the tropical and subtropical oceans between 30°N and 30°S in the ocean. Symbiodiniaceae ITS2 sequences belonging to clades A, C, D, and F were widely distributed and derived from coral reef areas across the Atlantic Ocean, the Red Sea, the Mediterranean Sea, and the Indo-Pacific Ocean. By contrast, the remaining Symbiodiniaceae ITS2 sequences belonging to the other clades were derived from specific oceanic regions. Specifically, Symbiodiniaceae ITS2 sequences belonging to Symbiodiniaceae clade B were mainly derived from coral reef areas in the Atlantic Ocean, e.g., the coral reefs of the Caribbean (Fig.3).

In Figs.4–7, the broad global perspectives of the diversity and distribution of the 119 Symbiodiniaceae ITS2 types belonging to the nine Symbiodiniaceae clades were



Fig.2 Symbiodiniaceae clades (A-I) associated hosts and Symbiodiniaceae detected in sediments or seawater. The relative abundance of Symbiodiniaceae clades was calculated on the basis of the number of Symbiodiniaceae ITS2 sequences derived from samples of different hosts and sediment or seawater samples.



Fig.3 Global geographic distribution of Symbiodiniaceae at the clade level. Numbers (1–28) represent different coral reef areas in or near the following regions: 1, Gulf of Mexico; 2, Caribbean Sea; 3, Brazil; 4, Southampton; 5, Gulf of Alicante (Spain); 6, Mediterranean Sea; 7, Rea Sea; 8, Persian Gulf; 9, Tanzania; 10, Bay of Bengal; 11, Thailand; 12, South China Sea; 13, South Korea; 14, Japan; 15, Taiwan; 16, Philippines; 17, Singapore; 18, Guam; 19, Palau; 20, Indonesia; 21, Papua; 22, New Guinea; 23, Australia; 24, New Zealand; 25, Hawaii; 26, Samoa; 27, Polynesia; and 28, Ascension Island. Pie charts represent the relative abundances of different Symbiodiniaceae clades found in each coral reef area. The pie charts are presented in one of five sizes according to the occurrence records of Symbiodiniaceae in each coral reef area. The global map was drawn with tmap1.8-1 (https://github.com/mtennekes/tmap) in R.



Fig.4 Phylogenetic tree (A) and principal component analysis (B) of Symbiodiniaceae ITS2 types in clades A and B. A phylogenetic tree (maximum likelihood) was constructed using Symbiodiniaceae ITS2 representative sequences in MEGA 6.06. Symbiodiniaceae was subjected to principal component analysis based on the occurrence records of Symbiodiniaceae in different coral reef areas (empty arrows indicate the hosts associated with Symbiodiniaceae ITS2 types, red arrows refer to free-living Symbiodiniaceae ITS2 types, and circles represent coral reef areas). The host information (h) is presented within parentheses next to the name of each coral reef area. h0, free-living Symbiodiniaceae; h1, *Tridacna* spp.; h2, *Stylophora* spp.; h3, *Pocillopora* spp.; h4, *Nephthea* spp.; h5, *Zoanthus* spp.; h6, *Porites* spp.; h7, *Plexaura* spp.; h8, *Cliona* spp.; h9, *Gorgonia* spp.; h10, *Orbicella* spp.; h11, *Hippopus* spp.; h12, *Acropora* spp.; h13, *Cassiopeia* spp.; h14, *Madracis* spp.; h15, *Eunicea* spp.; h16, *Plexaurella* spp.; h17, *Antillogorgia* spp.; h18, *Montastraea* spp.; h19, *Actinia* spp.; h20, *Millepora* spp. MCPP shows the *P*-value of the Monte Carlo permutation test.

obtained on the basis of the number of Symbiodiniaceae ITS2 sequences belonging to different Symbiodiniaceae ITS2 types and the corresponding reef area information, respectively. The results showed that the ITS2 sequences of Symbiodiniaceae ITS2 types A1.1, A1.2, and A1.3 were mainly derived from the hosts living in the Mediterranean Sea and the Rea Sea, the ITS2 sequences of Symbiodiniaceae ITS2 types A2.1 and A2.3 were mainly derived from the coral reef areas of the Indo-Pacific Ocean, and the ITS2 sequences of Symbiodiniaceae A6 was mainly derived from the coral reef areas of the Rea Sea (near the Mediterranean). The ITS2 sequences belonging to nine phylogenetically related Symbiodiniaceae ITS2 types in clade B were mainly derived from coral reefs in the Atlantic Ocean, e.g., the Caribbean (Fig.4). The ITS2 sequences of Symbiodiniaceae ITS2 types F1.1, F1.2, F1.3, and F1.4 were mainly derived from hosts in coral reef areas in the Indo-Pacific Ocean (near Hawaii and New Guinea). Conversely, the ITS2 sequences of Symbiodiniaceae ITS2 types F2.1, F2.2, and F2.4 were mainly derived from hosts in the Red Sea and the Mediterranean Sea. The ITS2 sequences of Symbiodiniaceae ITS2 types F4.1, F4.2, F4.3, and F4.4 were mainly derived from coral reef areas in the Indo-Pacific Ocean (near Guam), and the ITS2 sequences

of Symbiodiniaceae ITS2 types F5.1 and F5.2 were mainly derived from coral reef areas in the Caribbean. The ITS2 sequences of Symbiodiniaceae ITS2 types in clade I (i.e., types I1 and I2) were mainly derived from foraminifera living near Hawaii (Fig.5). The ITS2 sequences of the Symbiodiniaceae ITS2 types in clade H (e.g., H1.1, H1.2, and H2) were mainly derived from foraminifera in coral reefs in the Indo-Pacific Ocean (near Guam and New Guinea) and the Caribbean (H1.2; Fig.6). Notably, the ITS2 sequences of Symbiodiniaceae C1, C2, and C3 within clade C were derived from numerous animal hosts in different coral reef areas in the Atlantic Ocean, the Red Sea, and the Indo-Pacific Ocean (Fig.7B). These results suggested that Symbiodiniaceae C1, C2, and C3 were globally distributed ITS2 types, whereas the other Symbiodiniaceae ITS2 types within clade C had limited distributions (Fig.7C).

Further rarefaction curve analysis of Symbiodiniaceae ITS2 sequences and observed Symbiodiniaceae ITS2 type in coral reef areas in the Atlantic Ocean, the Mediterranean Sea, the Rea Sea, the Indo-Pacific Ocean, and Australian waters show that the number of Symbiodiniaceae ITS2 sequences from the four reef areas exceeded the point where the rarefaction curve reached its asymptote.



Fig.5 Same as Fig.4 but for clades D, E, G, and I. The host information (h) is presented within parentheses next to the name of each coral reef area. h0, free-living Symbiodiniaceae; h3, *Pocillopora* spp.; h5, *Zoanthus* spp.; h6, *Porites* spp.; h8, sponges (*Cliona* spp.); h12, *Acropora* spp.; h18, *Montastraea* spp.; h24, *Goniastrea* spp.; h26, *Montipora* spp.; h30, *Leptoseris* spp.; h33, foraminifera (h33-1, *Sorites* spp.; h33-2, *Marginopora* spp.); h34, *Palythoa* spp.; h36, *Galaxea* spp.; h37, anemone (unknown); h39, *Oulastrea* spp.; and h40, *Pteraeolidia* spp. MCPP shows the *P*-value of the Monte Carlo permutation test.



Fig.6 Same as Fig.4 but for clades F and H. The host information (h) is presented within parentheses next to the name of the coral reef area. h18, *Montastraea* spp.; h26, *Montipora* spp.; h33, foraminifera (*Sorites* spp.); h41, *Meandrina* spp.; and h42, *Sinularia* spp. MCPP shows the *P*-value of the Monte Carlo permutation test.



Fig.7 Phylogenetic tree (A) and principal component analysis (B) of Symbiodiniaceae ITS2 types C1, C2, and C3 (B) and other Symbiodiniaceae ITS2 types (C) in clade C. A phylogenetic tree (maximum likelihood) was constructed using Symbiodiniaceae ITS2 representative sequences in MEGA 6.06. Symbiodiniaceae was subjected to principal component analysis based on the occurrence records of Symbiodiniaceae in different coral reef areas (empty triangles (B) indicate Symbiodiniaceae ITS2 types C1, C2, and C3; empty arrows (C) refer to the hosts associated with Symbiodiniaceae ITS2 types; and circles represent coral reef areas). The host information (h) is presented within parentheses next to the name of the coral reef area. h0, free-living Symbiodiniaceae; h1, *Tridacna* spp.; h2, *Stylophora* spp.; h3, *Pocillopora* spp.; h5, *Zoanthus* spp.; h6, *Porites* spp.; h12, *Acropora* spp.; h18, *Montastraea* spp.; h21, *Coscinaraea* spp.; h22, *Entacmaea* spp.; h23, *Favites* spp.; h24, *Goniastrea* spp.; h25, *Marginopora* spp.; h26, *Montipora* spp.; h27, *Seriatopora* spp.; h28, *Symphyllia* spp.; h29, *Turbinaria* spp.; h30, *Leptoseris* spp.; h31, black corals (h31-1, *Antipathes* spp.; h31-2, *Aphanipathes* spp.); h32, *Agaricia* spp.; h33, foraminifera (h33-1, *Sorites* spp.; h33-2, *Marginopora* spp.); h34, *Palythoa* spp.; h35, *Plesiastrea* spp.; h36, *Galaxea* spp.; h37, *Condylactis* spp.; and h38, *Protopalythoa* spp. MCPP shows the *P*-value of the Monte Carlo permutation test.

Springer

These results revealed the geographic segregation of Symbiodiniaceae clades, especially Symbiodiniaceae ITS2 types, in the Atlantic Ocean, the Mediterranean Sea, the Red Sea, and the Indo-Pacific Ocean reef areas.

In summary, this study provided a broad overview of the diversity and distribution of Symbiodiniaceae clades/ types on a global scale.

## 4 Discussion

Recognizing the diversity of Symbiodiniaceae is challenging because neither sex differences nor significant morphological differences are readily observable among Symbiodiniaceae types (Santos and Coffroth, 2003). With the practical difficulties associated with analyzing the diversity of Symbiodiniaceae according to biological and morphological features, the phylogenetic species concept has been used to understand the diversity of Symbiodiniaceae (Correa and Baker, 2009), whose members are hypothesized to differentiate into potentially ecologically distinct entities (Correa and Baker, 2009). In accordance with the phylogenetic species concept, nine phylogenetically supported Symbiodiniaceae clades, known as clades A-I, are classified on the basis of largely conserved gene markers (Pochon et al., 2014). Symbiodiniaceae clades are further subdivided into subclade types or putative 'species' by using less conserved gene markers (Lajeunesse and Thornhill, 2011). The less conserved gene ITS2 marker has been the most widely used to survey the diversity of Symbiodiniaceae in reef-associated metazoans and foraminifera and in environmental samples near coral reefs in different regions (Arif et al., 2014).

This study showed that 119 Symbiodiniaceae ITS2 types within nine clades (A-I) mainly inhabited coral reefs between 30°N and 30°S. Franklin et al. (2012) conducted a similar study, but they reported 409 Symbiodiniaceae subclade types. This difference between Franklin's study and our study might be caused by the different methods used to collect and process the ITS2 sequences. Raw Symbiodiniaceae ITS2 sequences were collected by Franklin et al. (2012) to represent types. In the present study, the qualified Symbiodiniaceae ITS2 sequences were clustered and annotated as Symbiodiniaceae ITS2 types at a 97% similarity cutoff value (Arif et al., 2014). According to Arif et al. (2014), ITS2 sequence similarities with >97% are considered indicative of the same Symbiodiniaceae type because they likely represent nonspecific amplifications of closely related types or intragenomic sequence divergences. We acknowledged that some Symbiodiniaceae types could not be differentiated with a cutoff value of >97% similarity. For instance, the clade C Symbiodiniaceae, which includes many types, may require higher cutoff values to differentiate the diversity of Symbiodiniaceae at the type level (Arif et al., 2014). A minimum entropy decompositionbased approach (referred to as the metahaplotype approach) has been used to consolidate the high diversity to a smaller set of core sequence nodes based on biologically informative sequence positions rather than the fixed similarity thresholds of the OTU approach (Eren et al., 2015; Smith

*et al.*, 2017). Therefore, our results provided deflated rather than inflated Symbiodiniaceae diversity estimates.

Symbiodiniaceae clades A and F were distributed globally (Fig.3), whereas the Symbiodiniaceae ITS2 types within the two clades were geographically segregated (Figs.4 and 6). Even for the globally distributed Symbiodiniaceae clade C, the geographic segregation of certain Symbiodiniaceae ITS2 types within this clade was also suggested (Fig.7C). For example, Symbiodiniaceae ITS2 types C7, C8, and C36 were mainly associated with hosts living in coral reef areas near Brazil and a host from an aquarium in Chicago. Symbiodiniaceae ITS2 types C4 and C14 were mainly related to hosts in the Red Sea. Symbiodiniaceae ITS2 types C5, C9, C10, C11, C12, C13, C23, C24, C25, C26, and C41 were mainly linked to hosts living in the coral reef areas in the Indo-Pacific Ocean (near Japan). Our study, some previous studies (Lajeunesse, 2002, 2003; Lajeunesse et al., 2005; van Oppen et al., 2009), and one review (Baker, 2003) revealed that the Symbiodiniaceae ITS2 types within clade B were mainly distributed in the Atlantic Ocean. This result implied that they are endemic Symbiodiniaceae ITS2 types. Lajeunesse et al. (2005) also suggested that the distinct geographic distributions of Symbiodiniaceae are likely related to the hypothesis that the closure of the Central American Seaway by the uplift of the Central American Isthmus occurred 3.1-3.5 million years ago. Consequently, a barrier formed between the Atlantic Ocean and the Indo-Pacific Ocean and likely favored the growth of distinct communities of Symbiodiniaceae in the two oceans. The distinctness of the Symbiodiniaceae community in the Red Sea and the Mediterranean Sea may be due to specific physical geographic features in the region (creating landlocked marine ecosystems) and regional environmental conditions, e.g., 'higher sea surface temperatures (SST) and salinity than those in other ocean regions' (Sawall et al., 2014).

The flexible associations of Symbiodiniaceae with different hosts were observed in this study. For example, the following Symbiodiniaceae types were associated with different hosts at the phylum or genus level: A3.1, A5.1, A5.2, A4.2, and B1.1 (Fig.4); D1.1, 1.2, D1.3, E1, E2, E3, G1, and G3.1 (Fig.5); and C1, C2, and C3 (Fig.7). Specifically, this study revealed foraminifera-hosted Symbiodiniaceae in clades C, F, G, H, and I, whereas Baker (2003) reported foraminifera-hosted Symbiodiniaceae in clades C, F, and G. In addition, as the main component of the coral reef ecosystem, sponges could be associated with multiple Symbiodiniaceae clades; for example, they were related to Symbiodiniaceae ITS2 types belonging to clades A and G. This finding might have been observed because the survival of sponges depends on food particles from seawater (Buddemeier and Fautin, 1993) and the nutrition acquired from Symbiodiniaceae.

Corals and their supported coral reefs have suffered unprecedented declines in species abundance and community degradation over the last few decades because of the effects of climate change and anthropogenic disturbances, and further declines and degradation are projected in the near future (Hughes *et al.*, 2017). An emerging

'nugget of hope' for these threats to corals and their supported coral reefs is related to the interactions of corals with the ecological, physiological, and functional diversity of Symbiodiniaceae clades/types. For example, extremely high temperatures can lead to coral bleaching (a result of the breakdown of the interactions of coral hosts with Symbiodiniaceae); however, corals with more heattolerant Symbiodiniaceae (i.e., heat-tolerant Symbiodiniaceae in clade D) less likely bleach than other corals, and bleached corals gain Symbiodiniaceae with a high heat tolerance after recovery (van Oppen et al., 2009). Our results showed that corals are the most flexible hosts mainly associated with Symbiodiniaceae ITS2 types within clades A, B, C, D, and F. According to the 'adaptive bleaching hypothesis' proposed by Buddemeier et al. (1993), such flexible associations between Symbiodiniaceae and coral hosts may serve as an adaptive mechanism of corals under the effects of marine environmental changes (Putnam et al., 2012). Our study, previous studies, and reviews revealed that the Symbiodiniaceae belonging to clade B mainly existed in the Atlantic Ocean (Baker, 2003; Lajeunesse, 2003), implying that its members might be important for the local acclimation of coral-Symbiodiniaceae associations. A previous study suggested that Symbiodiniaceae clade D is typically observed as weedy or opportunistic algae that preferentially occupy recently stressed or marginal habitats, inshore regions, and bleached corals (Stat et al., 2013). These results showed that the Symbiodiniaceae belonging to clade D mainly occur in the Indo-Pacific Ocean (specifically, in corals living near Thailand, the South China Sea, and Japan). Therefore, corals living in the Indo-Pacific Ocean were likely threatened, but coral species mainly associated with clade D Symbiodiniaceae might have a more positive status, as previous studies suggested that corals composed of Symbiodiniaceae with a high heat tolerance (*i.e.*, heat-tolerant Symbiodiniaceae in clade D) were less likely to bleach.

The term 'free-living Symbiodiniaceae' is defined as follows: A) Symbiodiniaceae living outside a host while retaining the ability to form symbioses with hosts (Hirose et al., 2008), including 'transiently free-living' Symbiodiniaceae that has been recently expelled from their hosts and may not persist in a free-living state (Yamashita et al., 2013) or B) novel Symbiodiniaceae types not currently known to engage in symbioses permanently in a free-living state (Takabayashi et al., 2012). According to adaptive bleaching hypothesis, free-living Symbiodiniaceae may provide an open pool from which hosts under environmental stress can obtain symbionts (Johnson, 2011). In addition, free-living Symbiodiniaceae can affect the community composition of Symbiodiniaceae associated with different hosts (Manning and Gates, 2008; Takabayashi et al., 2012; Cunning et al., 2015) because Symbiodiniaceae may be routinely expelled by reef-associated metazoans, e.g., jellyfish and reef-building corals (Kinzie et al., 2001; Baird et al., 2009; van Oppen et al., 2009). This family may also be exogenously acquired by juvenile and adult corals via horizontal transfer (Stat et al., 2008; Baird et al., 2009). Our study partially supported these ideas as most

Symbiodiniaceae ITS2 types belonging to clades A, B, C, and D were observed in various reef-associated hosts and seawater or sediments near coral reef areas. Furthermore, some Symbiodiniaceae ITS2 types (*i.e.*, most Symbiodiniaceae ITS2 types within clades E and G) mainly existed in a free-living state.

### **5** Conclusions

The diversity and geographic distribution of Symbiodiniaceae were analyzed on the basis of the Symbiodiniaceae ITS2 sequences derived from metazoans and foraminifera associated with reefs and from seawater and sediments near coral reef areas in the GenBank database. In particular, the possible geographic segregation of Symbiodiniaceae ITS2 types implied that different Symbiodiniaceae ITS2 types might have different ecological functions. For example, the Symbiodiniaceae ITS2 types belonging to clade B were mainly from the Atlantic Ocean, implying that they were important for the local acclimation of host-Symbiodiniaceae associations. By contrast, the Symbiodiniaceae ITS2 types belonging to clade D were mainly found in the Indo-Pacific Ocean, suggesting that they might be essential for the maintenance of host-Symbiodiniaceae associations under environmental stress. However, the 3899 gualified Symbiodiniaceae sequences from GenBank were obtained from reef-associated metazoans and foraminifera and from seawater and sediment samples near coral reef areas by different researchers using various methods at different periods; consequently, sampling or method biases were unavoidable. Thus, any conclusions were premature at this stage. Nevertheless, our results provided a broad overview of Symbiodiniaceae clade/ type diversity and distribution on a global scale. An international program of the global sampling and investigation of Symbiodiniaceae via a consistent research strategy should be developed to fully understand the diversity and distribution of Symbiodiniaceae clades/types on a global scale.

1) Multiple gene barcodes (*e.g.*, ITS2, ITS1, rbcL, and matK) should be assessed and used to survey Symbiodiniaceae type diversity, and 28S rDNA should be utilized as the benchmark barcode of different Symbiodiniaceae clades (Pochon *et al.*, 2012, 2014).

2) High-resolution methods, including quantitative PCR and high-throughput DNA sequencing, should be considered, as recent studies revealed that rare Symbiodiniaceae clades/types are common in different hosts and environmental samples, which cannot be detected with traditional methods (*i.e.*, DGGE or RFLP; Arif *et al.*, 2014).

3) Known and potential hosts associated with Symbiodiniaceae and seawater, sand, and sediments near reef areas should be uniformly sampled from different coral reefs areas.

4) Environmental factors (*e.g.*, temperature, pH, light, inorganic and organic pollutants, and sampling depth) in different coral reefs areas should be considered.

5) A standardized nomenclature system of Symbiodiniaceae types and a related database should be constructed on the basis of the above aspects to monitor the dynamics of Symbiodiniaceae on a global scale and determine the effects of environmental factors on coral reefs.

### Acknowledgements

This work was funded by the National Key Basic Research Program of China (No. 2013CB956103), the Key Ecological Processes and Health Regulation Principles of Marine Ecosystem in Guangdong-Hong Kong-Macao Greater Bay Area (No. GML2019ZD0405), the CAS Pioneer Hundred Talents Program, and the South China Sea Institute of Oceanography, CAS (Nos. Y8SL031001 and Y9Y B021001).

### References

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J., 1990. Basic local alignment search tool. *Journal of Molecular Biology*, 215: 403-410.
- Arif, C., Daniels, C., Bayer, T., Hinestroza, E. B., Barbrook, A., Howe, C. J., Lajeunesse, T. C., and Voolstra, C. R., 2014. Assessing *Symbiodinium* diversity in scleractinian corals *via* nextgeneration sequencing-based genotyping of the ITS2 rDNA region. *Molecular Ecology*, 23: 4418-4433.
- Baird, A. H., Guest, J. R., and Willis, B. L., 2009. Systematic and biogeographical patterns in the reproductive biology of scleractinian corals. *Annual Review of Ecology Evolution and Systematics*, 40: 551-571.
- Baker, A. C., 2003. Flexibility and specificity in coral-algal symbiosis: Diversity, ecology, and biogeography of Symbiodinium. Annual Review of Ecology Evolution and Systematics, 34: 661-689.
- Blackall, L. L., Wilson, B., and van Oppen, M. J., 2015. Coralthe world's most diverse symbiotic ecosystem. *Molecular Ecology*, 24: 5330-5347.
- Buddemeier, R. W., and Fautin, D. G., 1993. Coral bleaching as an adaptive mechanism: A testable hypothesis. *BioScience*, 43: 320-326.
- Coffroth, M. A., and Santos, S. R., 2005. Genetic diversity of symbiotic dinoflagellates in the genus *Symbiodinium*. *Protist*, 156: 19-34.
- Coffroth, M. A., Lewis, C. F., Santos, S. R., and Weaver, J. L., 2006. Environmental populations of symbiotic dinoflagellates in the genus *Symbiodinium* can initiate symbioses with reef cnidarians. *Current Biology*, 16: 985-987.
- Correa, A. M. S., and Baker, A. C., 2009. Understanding diversity in coral-algal symbiosis: A cluster-based approach to interpreting fine-scale genetic variation in the genus *Symbiodinium. Coral Reefs*, 28: 81-93.
- Cunning, R., Gates, R. D., and Edmunds, P. J., 2017. Using highthroughput sequencing of ITS2 to describe *Symbiodinium* metacommunities in St. John, U.S. Virgin Islands. *PeerJ*, **5**: e3472.
- Cunning, R., Yost, D. M., Guarinello, M. L., Putnam, H. M., and Gates, R. D., 2015. Variability of *Symbiodinium* communities in waters, sediments, and corals of thermally distinct reef pools in American Samoa. *PLoS One*, **10**: e0145099.
- De'Ath, G., Fabricius, K. E., Sweatman, H., and Puotinen, M., 2012. The 27-year decline of coral cover on the Great Barrier Reef and its causes. *Proceedings of the National Academy of Sciences of the United States of America*, **109**: 17995-17999.

- Eren, A. M., Morrison, H. G., Lescault, P. J., Reveillaud, J., Vineis, J. H., and Sogin, M. L., 2015. Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *The ISME Journal*, 9: 968.
- Fautin, D. G., and Buddemeier, R. W., 2004. Adaptive bleaching: A general phenomenon. *Hydrobiologia*, **530**: 459-467.
- Franklin, E. C., Stat, M., Pochon, X., Putnam, H. M., and Gates, R., 2012. GeoSymbio: A hybrid, cloud-based web application of global geospatial bioinformatics and ecoinformatics for *Symbiodinium*–Host symbioses. *Molecular Ecology Resources*, 12: 369-373.
- Freudenthal, H. D., 1962. Symbiodinium gen. nov. and Symbiodinium microadriaticum spp. nov., a Zooxanthella: Taxonomy, life cycle, and morphology. Journal of Protozool, 9: 45-52.
- Green, E. A., Davies, S. W., Matz, M. V., and Medina, M., 2014. Quantifying cryptic *Symbiodinium* diversity within *Orbicella faveolata* and *Orbicella franksi* at the Flower Garden Banks, Gulf of Mexico. *PeerJ*, **2**: e386.
- Hirose, M., Reimer, J. D., Hidaka, M., and Suda, S., 2008. Phylogenetic analyses of potentially free-living *Symbiodinium* spp. isolated from coral reef sand in Okinawa, Japan. *Marine Biology*, **155**: 105-112.
- Hughes, T. P., Kerry, J. T., Lvarez-Noriega, M., Lvarez-Romero, J. G., and Wilson, S. K., 2017. Global warming and recurrent mass bleaching of corals. *Nature*, 543: 373-377.
- Johnson, M. D., 2011. The acquisition of phototrophy: Adaptive strategies of hosting endosymbionts and organelles. *Photo*synthesis Research, **107**: 117-132.
- Jones, R. J., 2008. Coral bleaching, bleaching-induced mortality, and the adaptive significance of the bleaching response. *Marine Biology*, **154**: 65-80.
- Keshavmurthy, S., Meng, P. J., Wang, J. T., Kuo, C. Y., Yang, S. Y., and Hsu, C. M., 2014. Can resistant coral–*Symbiodinium* associations enable coral communities to survive climate change? A study of a site exposed to long-term hot water input. *PeerJ*, 2: e327.
- Kinzie, R. A., Takayama, M., Santos, S. R., and Coffroth, M. A., 2001. The adaptive bleaching hypothesis: Experimental tests of critical assumptions. *Biology Bulletin*, **200**: 51-58.
- Lajeunesse, T. C., 2002. Diversity and community structure of symbiotic dinoflagellates from Caribbean coral reefs. *Marine Biology*, 141: 387-400.
- Lajeunesse, T. C., 2005. 'Species' radiations of symbiotic dinoflagellates in the Atlantic and Indo-Pacific since the Miocene– Pliocene transition. *Molecular Biology and Evolution*, **22**: 570-581.
- Lajeunesse, T. C., and Thornhill, D. J., 2011. Improved resolution of reef-coral endosymbiont (*Symbiodinium*) species diversity, ecology, and evolution through psbA non-coding region genotyping. *PLoS One*, 6: e29013.
- Lajeunesse, T. C., Loh, W., van Woesik, R., Hoegh-Guldberg, O., Schmidt, G. W., and Fitt, W. K., 2003. Low symbiont diversity in southern Great Barrier Reef corals, relative to those of the Caribbean. *Limnology and Oceanography*, **48**: 2046-2054.
- Lajeunesse, T. C., Parkinson, J. E., Gabrielson, P. W., Jeong, H. J., Reimer, J. D., Voolstra, C. R., and Santos, S. R., 2018. Systematic revision of Symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Current Biology*, 28: 1-11.
- Manning, M. M., and Gates, R. D., 2008. Diversity in populations of free-living *Symbiodinium* from a Caribbean and Pacific reef. *Limnology and Oceanography*, 53: 1853-1861.

- Pochon, X., Putnam, H. M., and Gates, R. D., 2014. Multi-gene analysis of *Symbiodinium* dinoflagellates: A perspective on rarity, symbiosis, and evolution. *PeerJ*, **2**: e394.
- Pochon, X., Putnam, H. M., Burki, F., and Gates, R. D., 2012. Identifying and characterizing alternative molecular markers for the symbiotic and free-living dinoflagellate genus *Symbiodinium. PLoS One*, 7: e29816.
- Prada, C., McIlroy, S. E., Beltrán, D. M., Valint, D. J., Ford, S. A., Hellberg, M. E., and Coffroth, M. A., 2014. Cryptic diversity hides host and habitat specialization in a gorgonian-algal symbiosis. *Molecular Ecology*, 23: 3330-3340.
- Putnam, H. M., Stat, M., Pochon, X., and Gates, R. D., 2012. Endosymbiotic flexibility associates with environmental sensitivity in scleractinian corals. *Proceedings of the Royal Society B–Biological Sciences*, 279: 4352-4361.
- Quigley, K. M., Davies, S. W., Kenkel, C. D., Willis, B. L., Matz, M. V., and Bay, L. K., 2014. Deep-sequencing method for quantifying background abundances of *Symbiodinium* types: Exploring the rare *Symbiodinium* biosphere in reef-building corals. *PLoS One*, **9**: e94297.
- Reimer, J. D., Herrera, M., Gatins, R., Roberts, M. B., Parkinson, J. E., and Berumen, M. L., 2017. Latitudinal variation in the symbiotic dinoflagellate *Symbiodinium* of the common reef zoantharian *Palythoa tuberculosa* on the Saudi Arabian coast of the Red Sea. *Journal of Biogeography*, 44: 661-673.
- Santos, S. R., and Coffroth, M. A., 2003. Molecular genetic evidence that dinoflagellates belonging to the genus *Symbiodinium* are haploid. *Biology Bulletin*, **204**: 10-20.
- Sawall, Y., Al-Sofyani, A., Banguera-Hinestroza, E., and Voolstra, C. R., 2014. Spatio-temporal analyses of *Symbiodinium* physiology of the coral *Pocillopora verrucosa* along largescale nutrient and temperature gradients in the Red Sea. *PLoS One*, 9: e103179.
- Smith, E. G., Ketchum, R. N., and Burt, J. A., 2017. Host specificity of *Symbiodinium* variants revealed by an ITS2 metahaplotype approach. *The ISME Journal*, **11**: 1500-1503.
- Stat, M., Loh, W. K. W., Hoegh-Guldberg, O., and Carter, D. A., 2008. Symbiont acquisition strategy drives host-symbiont associations in the southern Great Barrier Reef. *Coral Reefs*, 27:

763-772.

- Stat, M., Pochon, X., Franklin, E. C., Bruno, J. F., Casey, K. S., and Selig, E. R., 2013. The distribution of the thermally tolerant symbiont lineage (*Symbiodinium* clade D) in corals from Hawaii: Correlations with host and the history of ocean thermal stress. *Ecology and Evolution*, **3**: 1317-1329.
- Takabayashi, M., Adams, L. M., Pochon, X., and Gates, R. D., 2012. Genetic diversity of free-living *Symbiodinium* in surface water and sediment of Hawai'i and Florida. *Coral Reefs*, 31: 157-167.
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, **30**: 2725-2729.
- Ter Braak, C. J. F., and Smilauer, P., 2002. CANOCO reference manual and CanoDraw for Windows user's guide: Software for canonical community ordination (version 4.5). http://www. canoco.com.
- Thomas, L., Kendrick, G. A., Kennington, W. J., Richards, Z. T., and Stat, M., 2014. Exploring *Symbiodinium* diversity and host specificity in *Acropora* corals from geographical extremes of western Australia with 454 amplicon pyrosequencing. *Molecular Ecology*, 23: 3113-3126.
- Tonk, L., Sampayo, E. M., Weeks, S., Magno-Canto, M., and Hoegh-Guldberg, O., 2013. Host-specific interactions with environmental factors shape the distribution of *Symbiodinium* across the Great Barrier Reef. *PLoS One*, **8**: e68533
- van Oppen, M. J. H., Baker, A. C., Coffroth, M. A., and Willis, B. L., 2009. Bleaching resistance and the role of algal endosymbionts. In: *Coral Bleaching: Patterns, Processes, Causes and Consequences.* van Oppen, M. J. H., and Lough, J. M., eds., Springer, Berlin, Heidelberg, 83-102.
- Yamashita, H., and Koike, K., 2013. Genetic identity of freeliving *Symbiodinium* obtained over a broad latitudinal range in the Japanese coast. *Phycological Research*, **61**: 68-80.
- Ziegler, M., Arif, C., Burt, J. A., Dobretsov, S., Roder, C., and Lajeunesse, T. C., 2017. Biogeography and molecular diversity of coral symbionts in the genus *Symbiodinium* around the Arabian Peninsula. *Journal of Biogeography*, **44**: 674-686.

(Edited by Chen Wenwen)