Integrating phylogeographic and ecological niche approaches to delimitating cryptic lineages in the blue-green damselfish (*Chromis viridis*)

Shang Yin Liu [Corresp., 1], **Mao-Ning Tuanmu** [2], **Rita Rachmawati** [3, 4], **Gusti Ngurah Mahardika** [5], **Paul H Barber** [4]

[1] Department of Marine Biotechnology and Resources, National Sun Yat-Sen University, Kaohsiung, Taiwan

[2] Biodiversity Research Center, Academia Sinica, Taipei, Taiwan

[3] Center for Fisheries Research, Ministry of Marine Affairs and Fisheries, Jakarta, Indonesia

[4] Department of Ecology and Evolutionary Biology, University of California, Los Angeles, LA, United States

[5] The Indonesian Biodiversity Research Centre, Udayana University, Bali, Indonesia

Corresponding Author: Shang Yin Liu
Email address: syvliu@mail.nsysu.edu.tw

Species delimitation is challenging in sibling species/cryptic lineages because of the absence of clear diagnostic traits. However, integration of different approaches such as phylogeography and ecological niche comparison offers one potential approach to tease apart recently diverged lineages. In this study, we estimate the ecological niche divergence among lineages in *Chromis viridis* in a broad-scale phylogeographic framework to test whether the combination of these two approaches can effectively distinguish recently diverged lineages. Results from Cytb and Rag2 analyses identified two cryptic lineages (*C. viridis A* and *C. viridis B*) that diverged ~3Myr ago. Estimates of ecological niche divergence with 11 environmental parameters across the broad geographic range of these lineages showed overlapping ecological niches and niche conservatism. However, regardless of the incongruence between genetic and ecological niche divergence, the substantial genetic divergence between the two clades of *C. viridis* in both mtDNA and nuclear loci strong suggest that they are cryptic taxa.

Integrating phylogeographic and ecological niche approaches to delimitating cryptic lineages in the blue-green damselfish (*Chromis viridis*).

Shang-Yin Vanson Liu[1*], Mao-Ning Tuanmu[2], Rita Rachmawati[3,4], Gusti Ngurah Mahardika[5], Paul H. Barber[3]

[1] Department of Marine Biotechnology and Resources, National Sun Yat-Sen University, Kaohsiung, Taiwan

[2] Biodiversity Research Center, Academia Sinica, Taipei, Taiwan

[3] Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, USA

[4] Center for Fisheries Research, Ministry of Marine Affairs and Fisheries, Jakarta, Indonesia

[5] The Indonesian Biodiversity Research Centre, Udayana University, Bali, Indonesia

*Correspondence: Shang Yin Vanson Liu, Department of Marine Biotechnology and Resources, National Sun Yat-Sen University, Kaohsiung, Taiwan. E-mail: syvliu@mail.nsysu.edu.tw

Abstract

Species delimitation is challenging in sibling species/cryptic lineages because of the absence of clear diagnostic traits. However, integration of different approaches such as phylogeography and ecological niche comparison offers one potential approach to tease apart recently diverged lineages. In this study, we estimate the ecological niche divergence among lineages in *Chromis viridis* in a broad-scale phylogeographic framework to test whether the combination of these two approaches can effectively distinguish recently diverged lineages. Results from Cytb and Rag2 analyses identified two cryptic lineages (*C. viridis A* and *C. viridis B*) that diverged ~3Myr ago. Estimates of ecological niche divergence with 11 environmental parameters across the broad geographic range of these lineages showed overlapping ecological niches and niche conservatism. However, regardless of the incongruence between genetic and ecological niche

27 divergence, the substantial genetic divergence between the two clades of *C. viridis* in both
28 mtDNA and nuclear loci strong suggest that they are cryptic taxa.
29
30 **Introduction**
31 Comparative phylogeography examines patterns of congruence in phylogenetic breaks across
32 species distribution. It is built on the assumption that the processes driving lineage divergence,
33 speciation, and the evolution of biodiversity mainly involves geographical, historical, and
34 environmental factors that favor isolation and limit gene flow between populations (*Avise, 2000;*
35 *Rocha et al., 2007*).
36 In the marine realm, phylogeographic studies frequently identify cryptic species,
37 morphologically indistinguishable groups that have differences in neutral genetic markers that
38 are equal or greater than those observed between species with diagnostic morphological traits
39 (Knowlton, 1993). In some cases (e.g. *Drew & Barber, 2009; Drew et al., 2010; Hubert et al.,*
40 *2012; Liu et al., 2012*) such cryptic lineages are subsequently described as new species (*Allen &*
41 *Drew, 2012; Liu et al., 2012; Allen, 2015*), indicating that such cryptic lineages represent
42 overlooked biodiversity.
43
44 Systematists use different characteristics to differentiate species, and the variation in these
45 characteristics can arise at different times and rates during the process of lineage diversification
46 (*de Queiroz, 2007*). However, species delimitation is particularly difficult in adaptive or recent
47 radiations, when nascent species boundaries and their defining characteristics can be unclear
48 (*Puebla et al., 2008; Wagner et al., 2012; Victor, 2015*). In such cases, integration of different
49 approaches can increase the ability to detect recently separated lineages (*Leache et al., 2009*)

and can provide stronger evidence for lineage separation when the results are concordant (*Liu et al., 2012; Allen & Erdmann, 2012; Larkin et al., 2016*).

While integrating morphological diversification and genetic divergence is relatively common in addressing recent lineage diversification (e.g. *Allen & Drew, 2012; Allen & Erdmann, 2012*), ecological diversification (i.e. niche divergence) is seldom used to differentiate cryptic lineages/species in the marine realm. Niche diversification is a foundation of speciation (*Pyron et al., 2009*). Therefore, if niche divergence leads to lineage differentiation in the process of populations adapting to new environments (*Wiens, 2004; Wiens & Graham, 2005*), then a niche partitioning should be expected between lineages, particularly among recently diverged overlapping cryptic lineages identified in many phylogeographic studies.

To date, only a handful of studies have found niche divergence between cryptic marine taxa. For example, studies on deep-sea octocorals showed that niche partitioning between lineages is associated with the depth (*Quattrini et al., 2013; Quattrini et al., 2017*) and the marine cyanobacteria (genus *Prochlorococcus*) exhibits niche partitioning in associated with geography and environmental conditions. The absence of niche differentiation in the study of cryptic marine taxa is partially due to the lack of a centralized high-resolution spatial data representing both benthic and pelagic marine environments (*Sbrocco & Barber, 2013*), resources that are now available.

Chromis viridis is widely distributed Indo-Pacific coral reef fish. Both juveniles and adults associate with *Acropora* corals where they school and feed on zooplankton in the water column above coral heads (*Frédérich et al., 2009*). Using phylogenetic analyses based on mtDNA

variation, Froukh & Kochzius (2008) found three cryptic lineages in *C. viridis* across four Indo-Pacific localities, and Messmer et al. (2012) documented the presence of additional cryptic lineages in the Great Barrier Reef. However, it is unclear how geographical and ecological processes contribute to this nascent diversification.

To better understand the processes driving lineage diversification in *C. viridis*, we conducted a broad scale phylogeographic study of *C. viridis* across its distribution range. We then estimated ecological niche divergence between lineages in *C. viridis* in this phylogeographic framework to test whether the combination of these two approaches can effectively distinguish recently diverged lineages.

Materials & Methods

Sample collection and DNA extraction

We collected 252 *Chromis viridis* between 2007 and 2017 from 15 locations across its Indo-Pacific distribution (Table 1; Fig. 1). We collected specimens by hand net and clove oil, either by scuba diving or snorkeling and preserved tissue samples (fin clips, a piece of muscle, or both) in 95% alcohol and stored at 4 °C. Large fish were released following fin clipping, but individuals too small for fin clipping were euthanized with clove oil and preserved whole in 95% ethanol. The sampling depth of all specimens used in present study was less than 10m. This experiment doesn't involve animal experiment and the field sampling process was complied with the regulation drafted by the Animal Care and Use Committee of National Sun Yat-sen University. Additionally, the permission to perform research activities in Indonesia was issued by the Indonesian Government and the Ministry of Research and Technology under research permit No.272/SIP/FRP/SM/VII/2013.

We extracted genomic DNA from tissue fragments using Geneaid Tissue Genomic DNA mini Kit (Geneaid Biotech, Taiwan) following manufacturers' protocol and eluted extracted DNA in TE buffer and stored at -20 °C until amplification by polymerase chain reaction (PCR).

Amplification of genetic markers

We amplified a portion of the mitochondrial cytochrome b (Cytb) gene using universal primers GluDG-L and H16460 (*Palumbi, 1996*). PCRs reactions were 30 µL in volume, containing 10-40 ng template DNA, 3 µL 10X buffer, 0.2 m*M* dNTPs, 1.5 m*M* $MgCl_2$, 10 m*M* each primer, and 0.2 units of Taq polymerase (MDbio, Taipei). The thermocycling profile consisted of initial denaturation at 94 °C for 2 min, followed by 41 cycles of denaturation at 94 °C for 30 s, annealing at 57 °C for 30 s, and extension at 72 °C for 40 s, concluding with a final extension at 72 °C for 2 min.

Because mitochondrial and nuclear genes can have very different histories, we also amplified the nuclear recombination-activating (Rag2) gene for a subset of samples that represented distinct clades recovered in the mtDNA, using primers RAG2F and RAG2R (*Westneat & Alfaro, 2005*). PCR reactions were as above, except that we used a 5m*M* MgCl concentration and the following thermocycling parameters: 39 cycles of denaturation at 94 °C for 30 s, annealing at 56 °C for 30 s, and extension at 72 °C for 40 s, and a final extension at 72 °C for 2 min.

The nucleotide sequences of the PCR products of both loci were determined using ABI 3730XL automated sequencer (Carlsbad CA, U.S.A.) by Genomics (https://www.genomics.com.tw). Nucleotide sequences were assembled and edited using the SEQUENCHER version 4.2 software (Gene Code, Ann Arbor MI, U.S.A.). Sequences were uploaded to GenBank under the accession number MH743228-MH743691.

Phylogenetic analyses

Prior to any analyses, we aligned DNA sequences from each gene region in Clustal W (Thompson et al., 1994) and exported these sequences to MEGA 6 (*Tamura et al., 2013*) to visually inspect all alignments for accuracy. To quantify genetic diversity measures, we calculated standard genetic diversity indexes including haplotype diversity (*h*) and nucleotide diversity (π) in Arlequin 3.5 (*Excoffier & Lischer, 2010*). We then inferred phylogenetic trees from each individual locus using maximum likelihood (ML) and Bayesian inference performed on the CIPRES Science Gateway (Miller et al., 2010), and in BEAST 2.4.5 (*Bouckaert et al., 2014*). We conducted the ML analyses in RAxML version 8.1.24 (*Stamatakis, 2014*) using the GTR+G substitution model which selected by MEGA 6 (*Tamura et al., 2013*) as best-fit substitution model. For Bayesian inference, we used MrBayes version 3.2.2 (*Ronquist et al., 2012*), implementing two parallel runs of four simultaneous Markov chains for 10 million generations, sampling every 1000 generations and using the default parameters. Run parameters employed unlinked substitution models, rated heterogeneity models, and based frequencies across partitions. In the Bayesian analyses the first one million generations (10%) were discarded as burn-in, based on log- likelihood tree scores. Meanwhile, the convergence diagnostic was applied and the stop probability was set to 0.01. Nodal support was evaluated individually for all

trees using non- parametric bootstrapping with 1000 maximum likelihood replicates as employed in RAxML (*Stamatakis, 2014*), and by calculation of posterior probabilities as employed in MrBayes. Lastly, we generated a median-joining haplotype networks based on Cytb and Rag2 sequence datasets by using Popart 1.7 (http://popart.otago.ac.nz).

To date the ages of *Chromis viridis* clades, we arbitrarily subsampled 2 Cytb sequences from each clade; we also included two outgroups, including *Chromis atrilobata* (AY208524.1) and *Chromis multilineata* (AY208533.1) to calibrate date ranges as these two taxa diverged 3.1Mya, (*Quenouille et al., 2004*). We created the XML BEAST input file using the software BEAUti v.1.8.2. (as implemented in BEAST), with the setting of 50 million generations under the uncorrelated relaxed clock model, and sampling trees once every 1000 generations. To test for inter-run variation, we conducted two independent runs in BEAST 2.4.5 (*Bouckaert et al., 2014*), and then checked these runs for convergence with the software Tracer v1.5 (available at http://tree.bio.ed.ac.uk/software/tracer/). After discarding the 20% burn-in, we pooled the remaining tree samples from the two runs into a combined file to calculate the maximum clade credibility tree. From this tree, we calculated the posterior mean divergence ages and 95% credibility intervals (CI) for all nodes using Tree Annotator v1.8.2 (implemented in the BEAST package). To compare the similarity of two gene tree topologies, Phylo.io software (*Robinson et al*., *2016*) was used.

Ecological niche characterization and comparison

To characterize the ecological niches of *C. viridis* (*C. viridis* Clade A) and potential cryptic species (*C. viridis* Clade B), we obtained geophysical, biotic, and environmental data layers for global sea surface from Bio-ORACLE (downloaded on Feb. 20th, 2018; *Tyberghein et al., 2012; Assis et al., 2018*). We then extracted the values of those factors for the locations where *C. viridis* samples were collected. We excluded factors such as ice thickness and sea ice concentration from our analyses because all sample sites were tropical and ice free.

To explore niche differentiation between *C. viridis* Clade A and *C. viridis* Clade B, we first compared the values of individual environmental factors between *C. viridis* Clade A and Clade B sampling sites using Mann-Whitney tests. Next, we plotted niche differences in a two-dimensional non-metric multidimensional scaling (NMDS) space based on the Euclidean dissimilarity of those factors. Due to the different units of the environmental factors, we standardized factor values before running the NMDS analysis. Lastly, we tested equivalency of ecological niches between the species based on environmental factors of *C. viridis* Clade A and Clade*s* B sampling locations in a Principle Components Analysis framework. Due to potential correlations among environmental factors, we ran the principal component analysis (PCA) on all environmental factors and then used the resulting principal components to define niches of the species. Because *C. viridis* mainly occur in coral reefs, we restricted our PCA to regions (grid cells) where the maximum depth was equal or less than 50m based on the bathymetry data obtained from MARSPEC (http://www.marspec.org/; *Sbrocco & Barber, 2013*). To make the spatial resolution of the bathymetry data layer, which is originally at 30-by-30-arc second resolution, consistent with the Bio-ORACLE data, we upscaled it to 5-by-5-arc minute resolution

with each pixel value in the upscaled data layer being the maximum value over 10x10 original pixels within that new pixel.

To measure niche similarity between *C. viridis* Clade A Clade B, we calculated the *D* and *I* statistics developed by Warren et al. (2008). These two indices range from 0 and 1 with higher values indicating more similar niches. To test for niche equivalency between these two cryptic lineages of *C. viridis, we* compared the observed similarity values to null distributions of similarity obtained through a randomization procedure under a hypothesis of an identical niche (*Warren et al., 2008*), using the "dismo" package in R (version 3.4.2).

Results

We sequenced 911 base pairs of mitochondrial Cytb, and 743 base pairs of nuclear gene (Rag2, 743 bp) from 248 individual *Chromis viridis*, representing 15 locations across the Indo-Pacific (Fig. 1). In total, there were 135 unique haplotypes of Cytb; nucleotide diversity (π) ranged from 0.0033 to 0.0544, and haplotype diversity (h) ranged from 0.8824 to 1 among locations. For Rag2, 30 samples would not amplify, resulting in a total of 218 Rag2 sequences, representing fourteen unique haplotypes after omitting those sites with ambiguity codes (e.g. Y and K), additionally, there were 40 parsimony informative sites after alignment and 26 (heterozygotes) out of 40 were contained ambiguous signals. The nucleotide diversity (π) and haplotype diversity (h) of Rag2 ranged from 0 to 0.0154 and 0.9933 to 1, respectively (Table 1). These two genetic diversity indexes could be over-estimated since those ambiguous sites were been considered as variable sites during computation in Arlequin 3.5.

The maximum likelihood tree based on Cytb gene revealed three deeply divergent lineages within *C. viridis*; Clade A, Clade A-1, and Clade B (Fig. 2). In contrast, the corresponding Rag2 likelihood tree only differentiated Clades A and B, with no additional divergence within Clade 1. However, Clades A and B were not concordant across Cytb and Rag2; four mtDNA Clade B samples, one from Komodo and three from Lizard Island, clustered with Clade A based on Rag2. Additionally, one sample from Komodo contained a Clade B Rag2 sequence, but a Clade A Cytb sequence (Table S1).

Whether based on Cytb or Rag2, all Clade B haplotypes came from only three locations; Lizard Island on the Great Barrier Reef, Komodo Island in Indonesia, and the island of Fiji. In contrast, Clade A occurred broadly throughout the Indo-Pacific, including all three Clade B localities. Based solely on mtDNA, Clade A was further divided into two clades; Clade A was broadly distributed, and Clade A-1 was restricted to the Red Sea, although the Red Sea clade was not recovered in the Rag2 phylogeny. The general tree topology of these two genes was highly similar (Figure S1), except the Clade A-1 and those samples clustered with it, which can be revealed only by Cytb.

The patterns from the median joining network tree were identical to the phylogenetic trees based on either Cytb and Rag2. However, a few individuals assigned to Clade B based on Cytb clustered with Clade A based on their Rag2 sequences (Fig. 2).

Results from BEAST indicated that *C. viridis* Clade A and Clade B diverged approximately 3Mya (95% HPD: 0.631-6.291 Mya), with the divergence between Clade A and A-1 dating to 1.165 Mya (95% HPD: 0.247-2.49 Mya), and crown age of A-1 (Red Sea) was dated 0.138 Mya (95% HPD: 0.013-0.332 Mya) (Fig. 3).

Niche divergence tests

Results spanning 11 environmental factors (Table 2) showed no significant difference in ecological niche parameters of Clade A or Clade B of *C. viridis*. Moreover, there was also no significant environmental difference between the locations where two clades co-occur and those where only Clade B occurs (Table 2), suggesting that the environmental factors cannot explain why Clade B co-occurs with Clade A at some locations, but not at others.

The NMDS analysis also showed a lack of niche differentiation. The first two NMDS axes explained 86.7% of the variation in the 11 environmental factors across the 15 sampling sites. The three sampling sites where *C. viridis* B occurs fall at the center of the convex hull of all sampling sites in the two-dimensional NMDS space (Fig. 4), indicating complete niche overlap between Clades A and B. The first four principal components of the environmental factors, which were standardized before the analysis, had eigenvalues higher than 1. Together they explain about 83.1% of the total variation across all the grid cells with the maximum water depth less than or equal to 50m. Based on the component scores corresponding to the locations where Clades A and B were found, niche similarity indices were high, with $D = 0.841$ and $I = 0.973$. Moreover, niche equivalency tests showed that none of the two values were significantly

255 different from the values obtained through a randomization process under the hypothesis of an
256 identical niche (*P*-values for *D* and *I* were 0.311 and 0.294, respectively).
257
258 **Discussion**
259 Range-wide phylogeographic analyses of the blue green damselfish, *C. viridis* revealed two
260 divergent lineages in both mitochondrial Cytb and nuclear Rag2 DNA sequence data. These
261 cryptic lineages were first reported by Froukh and Kochzius (2008) in the Coral Triangle and
262 Red Sea, and subsequently by Messmer et al. (2012) who examined Australia and French
263 Polynesia. However, these studies each only covered a fraction of the geographic range of *C.*
264 *viridis.* By examining patterns across its entire range, this study shows that these cryptic lineages
265 of *C. viridis* are sympatric over a portion of their Pacific range.
266
267 Moreover, previous studies (*Froukh & Kochzius, 2008*; *Messmer et al., 2012*) only used mtDNA
268 markers, providing incomplete insights into genetic structure due to its maternal inheritance
269 (*Prugnolle & De Meeûs, 2002*; *Daly-Engel et al., 2012*). By sequencing both mtDNA and
270 nuclear markers, this study confirms the presence of two cryptic clades in *C. viridis,* Clades A
271 and B. These clades diverged approximately 3Mya similar to the divergence of *C. atrilobata* and
272 *C. multilineata* that were separated by Isthmus of Panama (*Domingues, et al. 2005*). Given the
273 depth of this divergence and the largely concordant differentiation of Clade A and B in Cytb and
274 Rag2, it is likely that these two lineages represent distinct species, raising interesting questions
275 about the sympatry of Clades A and B in part of their range.
276
277 **Origin of lineage diversification**

Although the divergence of Clades A and B of *C. viridis* is clear, the origin of this divergence is not. Within the Indo-Pacific region, the "Indo-Pacific Barrier" has been considered as a soft barrier that could be the main driving force of marine biological provinces in this region (*Gaither et al., 2011*). Barber & Bellwood (2005) and Cowen & Bellwood (2013) examined the importance of this barrier by evaluating the extent of temporal concordance in vicariance in three prominent families of reef fish, including Labridae, Pomacentridae and Chaetodontidae. Both studies showed that the isolation effect of Indo-Pacific barrier on the widely distributed fishes mainly occurred mostly between end Miocene and Early Pliocene (2-6 Myr), and the majority of vicariance events occurred in a narrow time interval at approximately 2.5 Myr.

The date of divergence of Clade A and B broadly conforms to the onset of Plio-Pleisetocene sea level fluctuations (*Voris, 2000*). However, Clade B populations of *C. viridis* occur on islands surround by deep water where lowered sea levels would not result in land barriers that promote vicariance, unlike other marine taxa distributed on opposite sides of the Sunda Shelf (e.g. *Barber et al., 2006*; *Crandall et al., 2008*; *DeBoer et al., 2014a, 2014b*; *Crandall et al., 2014*; *Waldrop et al., 2016*; *Simmonds et al., 2018*). Given the Indo-Pacific wide range of *C. viridis* and a divergence time between two cryptic clades dating to 3Mya, isolation across the Indo-Pacific Barrier is the most likely driver of divergence between Clades A and B. Under this scenario, Clade A would be an Indian Ocean clade that has expanded into the Pacific, where it now overlaps with the Pacific Clade B, a process previously noted in Neritid snails (*Crandall et al., 2008*).

An alternate, but not mutually exclusive explanation for divergence of Clade B comes from recent studies of coral associated marine taxa. Similar to phytophagus insects that undergo ecological divergence associated with host switching (*Berlocher & Feder*, 2002; *Hébert et al*., 2016), recent studies from the marine realm demonstrate that changes in coral host taxa can promote lineage divergence, potentially leading to speciation (*Simmonds et al., 2018*). Samples were not collected in a way that allows us to test this hypothesis, but future studies separating samples by coral host to determine whether individuals from Clades A and B exist in mixed schools in sympatric populations, and if so, whether those schools are associated with different coral hosts.

The geographic distribution of Clade B is curious in that it is observed in Komodo, but not in other populations in the Lesser Sunda Islands (e.g. Amed, Nusa Lembongan). Similarly, Clade B is observed on the Great Barrier Reef and Fiji, but not in New Caledonia, a population located between these too population. One explanation for this disjunct distribution is that Clade B individuals are relatively rare, and that greater sampling intensity would reveal Clade B haplotypes in adjacent ranges, as would be expected. Alternatively, it is possible that Clade A is gradually displacing Clade B populations, and the areas of sympatry represent locations where this process is incomplete. Similar arguments were made to explain sympatry of highly divergent clades of marine snails in the Pacific Ocean (*Crandall et al., 2008*).

In contrast to the divergence of Clade A and B, the divergence of Clade A-1, unique to the Red Sea, is more easily interpreted as the result of vicariance. The Red Sea is a semi-enclosed basin that is frequently invoked in driving population differentiation of reef organisms between the

Red Sea and Indian Ocean (*Dibattista et al., 2013*). Additionally, the 95% HPD of Red Sea Clade is ranged from 0.013 to 0.332 Mya which aligns with the most recent closure of the Red Sea (*Siddall et al.*, 2003). However, while mitochondrial DNA shows a well-supported Red Sea sub clade within *C. viridis* Clade A, this clade is not confirmed with nuclear Rag2 sequence. Lack of concordance between the mtDNA and nuclear phylogenies is not surprising. First, there is substantially more genetic variation in the mtDNA sequences because DNA repair mechanisms are less efficient in mtDNA, resulting in much higher substitution rates (*Alexeyev et al., 2013*). Second, because mitochondrial genome is maternally inherited and haploid, its effective population size is one-fourth that of a nuclear gene, meaning that lineage sorting occurs more rapidly in mtDNA (*Hare, 2001*). Thus, mitochondrial gene trees have a substantially higher probability of accurately recovering recently divergence lineages with short internode distances than do nuclear genes (Moore, 1995). Given that the Red Sea mtDNA clade only dates to 0.138 Mya, it is unsurprising that divergence in this region is not recovered in the Rag2 sequences. Thus, while the Red Sea population is clearly a sub-population of *C. viridis* Clade A, it is unlikely a cryptic species, like Clades A and B.

Cryptic Species

Many phylogeographic studies in the Indo-Pacific find highly divergent clades, often with regions of sympatry (i.e. *Barber et al., 2002, 2006*; *Crandall et al., 2008*; *Gaither et al., 2011*; *Liu et al., 2012*, *DeBoer et al., 2014*; *Bowen et al., 2016*). In many cases, these cryptic clades may represent cryptic species, similar to Liu et al. (2012) where a divergent clade of *Pomacentrus coelestis* in Micronesia was subsequently described as a new species (Liu et al., 2013). The concordant phylogeographic patterns in mtDNA and nuclear DNA, and the depth of

346 this divergence suggests that Clades A and B of *C. viridis* may represent two cryptic species with
347 different geographic ranges; Clade A is a widely distributed Indo-Pacific, while Clade B is only
348 found in Lizard Island, Komodo, and Fiji.
349
350 Comparison of Cytb and Rag2 trees revealed a limited amount of discordance. Given the
351 relatively recent divergence, one potential explanation for this pattern is incomplete lineage
352 sorting (*Tang et al., 2012*). Incomplete lineage sorting is the simplest explanation for individuals
353 with Clade B *C. viridis* haplotypes in the mtDNA tree having Clade A *C. viridis* Rag2 sequences
354 (Table S1). However, incomplete lineage sorting could not explain the sequence of komo4
355 clustered with *C. viridis* clade B in the Rag2 tree but clustered with *C. viridis* Clade A in the
356 mtDNA tree. The later finding could likely explain by the consequence of hybridization.
357
358 Hybridization is in *C. viridis* is unsurprising as closely related Pomacentrids often have
359 overlapping geographic distributions, co-occur in the same microhabitats (e.g. colonies of
360 branching corals) (*Randall et al., 1977*), and spawn sympatrically (*Jan, 1996*). Hybridization has
361 been observed in several sibling species of damselfish, including *Abudefduf*
362 *abdominalis*×*Abudefduf vaigiensis* (*Coleman et al., 2014*), *Amphiprion*
363 *chrysopterus*×*Amphiprion sandaracinos* (*Gainsford et al., 2015*), *Amphiprion*
364 *mccullochi*×*Amphiprion akindynos* (*van der Meer et al., 2012*), *Dascyllus carneus*×*D.*
365 *marginatus* (*DiBattista et al., 2015*), and *D. reticulatus*×*D. aruanus* (*He et al., 2017*). As such,
366 both hybridization and incomplete lineage sorting are likely responsible for the discordant
367 patterns in the two markers.
368

Incongruence between genetic and ecological niche divergence

Hutchinson (1978) proposed that two species cannot occupy the same ecological niche, yet niche conservatism predicts that closely related taxa retain ancestral ecological affiliations and persists in similar environments (*Lord et al., 1995*; *Webb et al., 2002*; *Wiens & Graham, 2005*). Allopatric sister taxa are often characterized by niche conservatism, but because geographic isolation drives speciation (*Peterson et al., 1999*; *Peterson et al., 2001*; *Kozak & Wiens, 2006*) sibling species can't compete for the same niche space, because, by definition, they have non-overlapping geographic ranges.

The alternative to niche conservatism is niche divergence, in which sister taxa occupy different niches (*Dayan & Simberloff, 2005*). Niche divergence is typically associated with sympatric speciation as diversification results from reduction in gene flow associated with divergence in traits with ecological function (e.g. habitat segregation, pollinator divergence, behavioral changes, phenological shifts, and mating system shifts). Under niche divergence, young sister species with high degrees of range overlap should ecologically diverge (*Dayan & Simberloff, 2005*; *Davies et al., 2007*). While ecological modeling and phylogenetics are commonly integrated to understand the relationship between evolutionary and ecological divergence of sibling species (*Kalkvik et al., 2012*; *Schorr et al., 2012*), few studies examine the intra-specific (cryptic lineage) level (*Gutiérrez-Tapia & Palma, 2016*).

Our results show that Clade A and B of *C. viridis* are sympatric, with Clade B haplotypes occurring only at three localities. There are two possible explanations for this pattern. First,

lineage diversification could have occurred in sympatry, as seen in Indo-Pacific coralivorous snails (*Simmonds et al., 2018*). Alternatively, diversification could have occurred in allopatry with secondary overlap in geographic ranges as proposed for neritid snails (*Crandall et al., 2008*). Niche comparisons based on 11 environmental factors showed that Clade A has a broader ecological niche than Clade B does (Fig 4). Despite this difference, there is no significant niche differentiation between these two lineages. As such, diversification is more likely to have arisen with niche conservatism rather than niche differentiation.

Given the absence of niche differentiation, our results suggest that the diversification of Clades A and B most likely occurred in allopatry, and that secondary contact is the most likely explanation for their current distribution. If true, niche conservatism could help explain the disjunct distribution of Clade B. Under this scenario, Clade A and Clade B would compete for the same niche space when they occur in sympatry. If Clade A is a superior competitor, it would gradually displace Clade B populations, potentially explaining the relative rarity of Clade B haptotypes and their disjunct distributions.

Alternatively, Pyron et al. (2015) suggested that niche conservatism could result from soft allopatry where there is low environmental heterogeneity. As noted above, the shallow Sunda and Sahul continental shelves exposed during low sea level stands (*Voris 2000*), forming long land bridges that restricted larval exchange between the tropical Indian Ocean and the western Pacific (reviewed in *Randall, 1998*). However, because deep water passages in what is modern day Indonesia remained open, and because low sea level stands were intermittent, the Indo-Pacific Barrier has been considered a soft dispersal barrier for marine taxa (*Cowman &*

415 *Bellwood, 2013*). The 11 environmental factors we used to define the niches of the two lineages
416 of *C. viridis* only reflect broad-scale environmental conditions. However, there could be subtle
417 variation in environmental factors, or variation could occur at finer scales (e.g. microhabitat),
418 resulting in soft-allopatric divergence.
419
420 It is important to note that the broad environmental variation we examine does not capture the
421 potential diet shifts observed in sibling species (*Goodheart et al., 2017*), novel traits (*Liu et al.,*
422 *2018*), and/or micro habitat preference (*Whitney et al., 2018*) that could act to drive or reinforce
423 lineage diversification. The latter is particularly intriguing given recent studies demonstrating
424 ecological speciation resulting from shifts in coral hosts (*Simmonds et al., 2018*). Further studies
425 such as stable isotope analysis and detailed morphological examinations in a phylogenetic
426 framework are needed to better understand the ecological and morphological divergence between
427 these two lineages.
428
429
430
431 **Conclusions**
432 Although cryptic diversification in widespread marine species is common (*Hubert et al., 2012*),
433 phylogeographic studies typically ignore the potential role of ecological niche partitioning on
434 lineage diversification. The present study shows evolutionary divergence between two lineages
435 of *C. viridis* that have overlapping ecological niches, supporting niche conservatism. It is unclear
436 whether this pattern results from allopatric divergence with secondary contact, or from subtle
437 differentiation in ecological niches not captured by the broad scale environmental data used to
438 compare ecological niches. Regardless of the origins, the substantial genetic divergence between
439 the two clades of *C. viridis* in both mtDNA and nuclear loci strong suggest that they are cryptic

taxa. Because *C. viridis* is highly exploited in the aquarium trade (*Wabnitz, 2003*), we suggest that a higher conservation priority needs to be given on the Clade B lineages restricted to Australia, Indonesia, and Fiji to protect this unique genetic lineage.

Acknowledgements

We thank Dr. S. Cheng, Dr. M-J Ho, Dr. Y-R Cheng, F-T Chang, N. Narendra and A. Sembiring for their assistant in the field, and Dr. V. Messmer, Dr. B. Frédérich, Dr. F. Borsa, Professor S. Planes and S. Johnson for sharing tissue samples. Special thanks to A. C. Bentley, for his curatorial help on obtaining tissue samples from Ichthyology session of Kansas University. This study was funded by Ministry of Science and Technology, Taiwan (Grant Number: MOST 106-2611-M-110-009). All applicable international, national, and institutional guidelines for the use of animals were followed. We declare no conflict of interest.

References

Alexeyev M, Shokolenko I, Wilson G, LeDoux S. 2013. The maintenance of mitochondrial DNA integrity-critical analysis and update. *Cold Spring Harbor Perspectives In Biology* **5:** a012641

Allen GR, Erdmann MV. 2012. *Reef fishes of the East Indies*. Perth, Australia: Tropical Reef Research

Allen GR, Drew J. 2012. *Pomacentrus maafu* a new species of damselfish from the Southwest Pacific. *Aqua: International Journal of Ichthyology* **18：** 171-180

Allen GR, Erdmann MV, Kurniasih E. 2015. *Chrysiptera caesifrons*, a new species of damselfish (Pomacentridae) from the south-western Pacific Ocean. *Journal of the Ocean Science Foundation* **15:** 16-32

Assis J, Tyberghein L, Bosch S, Verbruggen H, Serrão EA, De Clerck O. 2018. Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecology Biogeography* **27:** 277–284

Avise JC. 2000. *Phylogeography: the history and formation of species*. Cambridge, MA: Harvard university press.

Avise JC. 2009. Phylogeography: Retrospect and prospect. *Journal of Biogeography* **36:** 3–15

Barber PH, Palumbi SR, Erdmann MV, Moosa MK. 2002. Sharp genetic breaks among populations of *Haptosquilla pulchella* (Stomatopoda) indicate limits to larval transport: patterns, causes, and consequences. *Molecular Ecology* **11:** 659-674

Barber PH, Palumbi SR, Erdmann MV. 2006. Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the Coral Triangle. *Evolution* **60:** 1825-1839

Berlocher SH, Feder JL. 2002. Sympatric speciation in phytophagous insects: moving beyond controversy? *Annual Review of Entomology* **47:** 773-815

Bowen BW, Karl SA, Pfeiler E. 2007. Resolving evolutionary lineages and taxonomy of bonefishes (*Albula spp*.). In: Ault, J.S. (Ed.), Biology and Management of the World Tarpon.

Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014. BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. *PLoS Computational Biology* **10:** e1003537

Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. 2002. Phylogenies and community ecology. *Annual review of ecology and systematics* **33**: 475-505

Coleman RR, Gaither MR, Kimokeo B, Stanton FG, Bowen BW, Toonen RJ. 2014. Large-scale introduction of the Indo-Pacific damselfish *Abudefduf vaigiensis* into Hawai'i promotes genetic swamping of the endemic congener *A. abdominalis*. *Molecular Ecology* **23:** 5552–5565

Cowman PF, Bellwood DR. 2013. The historical biogeography of coral reef fishes: global patterns of origination and dispersal. *Journal of Biogeography* **40:** 209 - 224

Crandall ED, Frey M, Grossberg RK. Barber, PH. 2008. Contrasting demographic history and phylogeographical patterns in two Indo-Pacific gastropods. *Molecular Ecology* **17:** 611-626

Crandall ED, Treml EA, Liggins L, Gleeson L, Yasuda N, Barber PH, Gert W, Riginos, C. 2014. Return of the ghosts of dispersal past: historical spread and contemporary gene flow in the blue sea star *Linckia laevigata*. *Bulletin of Marine Science* **90:** 399-425

Daly-Engel TS, Seraphin KD, Holland KN, Coffey JP, Nance HA, Toonen RJ, Bowen BW. 2012. Global phylogeography with mixed-marker analysis reveals male-mediated dispersal in the endangered scalloped hammerhead shark (*Sphyrna lewini*). *PLoS One* **7:** e29986

Dayan T, Simberloff D. 2005. Ecological and community-wide character displacement: the next generation. *Ecology Letters* **8:** 875-894

Davies TJ, Meiri S, Barraclough TG, Gittleman JL. 2007. Species co-existence and character divergence across carnivores. *Ecology Letters* **10:** 146–152

DeBoer TS, Naguit MR, Erdmann MV, Ablan-Lagman MCA, Ambariyanto, Carpenter KE, Toha AHA, Barber PH. 2014a. Concordant phylogenetic patterns inferred from mitochondrial and microsatellite DNA in the giant clam *Tridacna crocea*. *Bulletin of Marine Science* **90:** 301-329

DeBoer TS, Naguit MR, Erdmann MV, Ablan-Lagman MCA, Ambariyanto, Carpenter K E, Toha AHA, Barber PH. 2014b. Concordance between phylogeographic and biogeographic boundaries in the Coral Triangle: conservation implications based on comparative analyses of multiple giant clam species *Bulletin of Marine Science* **90:** 277-300

de Queiroz K. 2007. Species concepts and species delimitation. *Systematic Biology* **56:** 879 – 886

DiBattista JD, Berumen ML, Gaither MR, Rocha LA, Eble JA, Choat JH, Craig MT, Sinclair-Taylor TH, Bowen BW. 2013. After continents divide: comparative phylogeography of reef fishes from the Red Sea and Indian Ocean. *Journal of Biogeography* **40:** 1170-1181.

DiBattista JD, Rocha L A, Hobbs J-PA, He S, Priest MA, Sinclair-Taylor TH, Bowen BW, Berumen ML. 2015. When biogeographical provinces collide: hybridization of reef fishes at the crossroads of marine biogeographical provinces in the Arabian Sea. *Journal of Biogeography* **42:** 1601– 1614

Drew JA, Barber PH. 2009. Sequential cladogenesis of *Pomacentrus moluccensis* (Bleeker, 1853) supports the peripheral origin of marine biodiversity in the Indo-Australian Archipelago. *Molecular Phylogenetics and Evolution* **53:** 335–339

Drew JA, Allen GR, Erdmann MV. 2010. Congruence between genes and color morphs in a coral reef fish: population variability in the Indo-Pacific damselfish *Chrysiptera rex* (Snyder, 1909). *Coral Reefs* **29:** 439-444

Frédérich B, Fabri G, Lepoint G, Vandewalle P, Parmentier E. 2009. Trophic niches of thirteen damselfishes (Pomacentridae) at the Grand Recif of Toliara, Madagascar. *Ichthyological Reserch* **56:** 10–17

Froukh T, Kochzius M. 2008. Species boundaries and evolutionary lineages in the blue green damselfishes *Chromis viridis* and *Chromis atripectoralis* (Pomacentridae). *Journal of Fish Biology* **72:** 451-457

Gainsford A, Herwerden L, Jones GP. 2015. Hierarchical behaviour, habitat use and species size differences shape evolutionary outcomes of hybridization in a coral reef fish. *Journal of Evolutionary Biology* **28:** 205–222

Gaither MR, Bowen BW, Bordenave TR, Rocha LA, Newman SJ, Gomez JA, van Herwerden L, Craig MT. 2011. Phylogeography of the reef fish *Cephalopholis argus* (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with contemporary overlap in the Coral Triangle. *BMC evolutionary biology* **11:** 189

579 **Goodheart JA, Bazinet AL, Valdés Á, Collins AG, Cummings MP 2017.** Prey preference
580 follows phylogeny: evolutionary dietary patterns within the marine gastropod group
581 Cladobranchia (Gastropoda: Heterobranchia: Nudibranchia). *BMC evolutionary biology* **17:** 221
582
583 **Gutiérrez-Tapia P, Palma RE. 2016.** Integrating phylogeography and species distribution
584 models: cryptic distributional responses to past climate change in an endemic rodent from the
585 central Chile hotspot. *Diversity and Distributions* **22:** 638-650
586
587 **Hare MP. 2001.** Prospects for nuclear gene phylogeography. *Trends in Ecology & Evolution* **16:**
588 700-706
589
590 **Hébert JB, Scheffer SJ, Hawthorne DJ. 2016.** Evidence for ecological speciation via a host
591 shift in the holly leaf miner, *Phytomyza glabricola* (Diptera: Agromyzidae). *Ecology and*
592 *Evolution* **6:** 6565-6577
593
594 **He S, Johansen JL, Hoey AS, Pappas MK, Berumen ML. 2017.** Molecular confirmation of
595 hybridization between *Dascyllus reticulatus*×*Dascyllus aruanus* from the Great Barrier Reef.
596 *Marine Biodiversity*, 1-10
597
598 **Hubert N, Meyer CP, Bruggemann HJ, Guérin F, Komeno RJL, Espiau B, Causse R,**
599 **Williams JT. Planes S. 2012.** Cryptic diversity in Indo-Pacific coral-reef fishes revealed by
600 DNA-barcoding provides new support to the centre-of-overlap hypothesis. *PLoS one* **7:** e28987
601

602 **Hutchinson GE. 1978.** An Introduction to Population Ecology (Yale Univ Press, New Haven,
603 CT).
604 **Davies JT, Meiri S, Barraclough TG, Gittleman JL. 2007.** Species coexistence and character
605 divergence across carnivores. *Ecology letters* **10:** 146-152
606
607 **Domingues VS, Bucciarelli G, Almada VC, Bernardi G. 2005.** Historical colonization and
608 demography of the Mediterranean damselfish, *Chromis chromis*. *Molecular Ecology* **14:** 4051-
609 4063
610
611 **Kalkvik HM, Stout IJ, Doonan TJ, Parkinson CL. 2012.** Investigating niche and lineage
612 diversification in widely distributed taxa: phylogeography and ecological niche modeling of the
613 *Peromyscus maniculatus* species group. *Ecography* **35:** 54-64
614
615 **Knowlton N. 1993.** Sibling species in the sea. *Annual Review of Ecology and Systematics* **24:**
616 189-216
617
618 **Kocher TD, Conroy JA, McKaye KR, Stauffer JR. 1993.** Similar morphologies of cichlid fish
619 in Lakes Tanganyika and Malawi are due to convergence. *Molecular Phylogenetic and Evolution*
620 **2:** 158–165
621
622 **Kozak KH, Wiens JJ. 2006.** Does niche conservatism promote speciation? A case study in
623 North American salamanders. *Evolution* **60:** 2604-2621
624

Leaché AD, Koo MS, Spencer CL, Papenfuss TJ, Fisher RN, McGuire JA. 2009. Quantifying ecological, morphological, and genetic variation to delimit species in the coast horned lizard species complex (*Phrynosoma*). *Proceedings of the National Academy of Sciences* **106:** 12418-12423

Larkin AA, Blinebry SK, Howes C, Lin Y, Loftus SE, Schmaus CA, Zinser ER, Johnson Z I. 2016. Niche partitioning and biogeography of high light adapted *Prochlorococcus* across taxonomic ranks in the North Pacific. *The ISME journal* **10:** 1555

Lessios HA. 2008. The Great American Schism: divergence of marine organisms after the rise of the Central American isthmus. *Annual Review of Ecology, Evolution, and Systematics* **39:** 63–91

Liu SYV, Dai C-F, Allen GR, Erdmann MV. 2012. Phylogeography of the neon damselfish *Pomacentrus coelestis* indicates a cryptic species and different species origins in the West Pacific Ocean. *Marine Ecology Progress Series* **458:** 155–167

Liu SYV, Ho H-C, Dai C-F. 2014. A new species of *Pomacentrus* (Actinopterygii: Pomacentridae) from Micronesia, with comments on its phylogenetic relationships. *Zoological Studies* **52:** 1–8

Liu SYV, Frederich B, Lavoué S, Chang J, Erdmann MV, Mahardika GN, Barber PH. 2018. Buccal venom gland associates with increased of diversification rate in the fang blenny fish *Meiacanthus* (Blenniidae; Teleostei). *Molecular phylogenetics and evolution* **125:** 138-146

Lord J, Westoby M, Leishman M. 1995. Seed size and phylogeny in six temperate floras: constraints, niche conservatism, and adaptation. *The American Naturalist* **146:** 349-364

Marske K, Leschen R, Buckley T. 2012. Concerted versus independent evolution and the search for multiple refugia: comparative phylogeography of four forest beetles. *Evolution* **66:** 1862–1877

Messmer V, Jones GP, Munday PL, Planes S. 2012. Concordance between genetic and species diversity in coral reef fishes across the Pacific Ocean biodiversity gradient. *Evolution* **66:** 3902-3917

Miller M, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. *Gateway Computing Environments Workshop (GCE)*, 1-8

Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* **49:** 718-726

Palumbi SR. 1996. Nucleic acid II: the polymerase chain reaction. In: Hillis, D.M., Moritz, G., Mable, B.K. (eds.). *Molecular Systematics. Sunderland, Mass.: Sinauer Associates*, 205–247

Peterson AT, Soberón J, Sánchez-Cordero V. 1999. Conservatism of ecological niches in evolutionary time. *Science* **285:** 1265–1267

671

672 **Peterson AT, Sánchez-Cordero V, Soberon J, Bartley J, Buddemeier RW, Navarro-**
673 **Sigüenza AG. 2001.** Effects of global climate change on geographic distributions of Mexican
674 Cracidae. *Ecological modelling* **144:** 21-30

675

676 **Prugnolle F, De Meeûs T. 2002.** Inferring sex-biased dispersal from population genetic tools: a
677 review. *Heredity 88*, 161

678

679 **Puebla O, Bermingham E, Guichard F. Whiteman E. 2007.** Colour pattern as a single trait
680 driving speciation in *Hypoplectrus* coral-reef fishes? *Proceedings of the Royal Society B* **274:**
681 1265–1271

682

683 **Pyron RA, Burbrink FT. 2009.** Lineage diversification in a widespread species: roles for niche
684 divergence and conservatism in the common king snake, *Lampropeltis getula*. *Molecular*
685 *ecology* **18:** 3443-3457

686

687 **Pyron RA, Costa GC, Patten MA, Burbrink FT. 2015.** Phylogenetic niche conservatism and
688 the evolutionary basis of ecological speciation. *Biological Reviews* **90:** 1248-1262

689

690 **Quattrini AM, Georgian SE, Byrnes L, Stevens A, Falco R, Cordes EE. 2013.** Niche
691 divergence by deep-sea octocorals in the genus *Callogorgia* across the continental slope of the
692 Gulf of Mexico. *Molecular Ecology* **22:** 4123-4140

693

694 **Quattrini AM, Gómez CE, Cordes EE. 2017.** Environmental filtering and neutral processes
695 shape octocoral community assembly in the deep sea. *Oecologia* **183:** 221-236
696
697 **Quenouille B, Bermingham E, Planes S. 2004.** Molecular systematics of the damselfishes
698 (Teleostei: Pomacentridae): Bayesian phylogenetic analyses of mitochondrial and nuclear DNA
699 sequences. *Molecular phylogenetics and evolution* **31:** 66-88
700
701 **Rambaut A, Drummond AJ. 2014.** BEAST 2: A Software Platform for Bayesian Evolutionary
702 Analysis. *PLoS Computational Biology* **10:** e1003537
703
704 **Randall HA, Allen GR. 1977.** A revision of the damselfish genus *Dascyllus* (Pomacentridae)
705 with the description of a new species. *Records of the Australian Museum* **31:** 349-385
706
707 **Randall JE. 1998.** Zoogeography of shore fishes of the Indo-Pacific region. *Zoological Studies*
708 **37:** 227-268
709
710 **Robinson O, Dylus D, Dessimoz C. 2016.** Phylo.io: Interactive viewing and compar- ison of
711 large phylogenetic trees on the web. *Molecular Biology and Evolution* **33:** 2163–2166
712
713 **Rocha LA, Craig MT, Bowen BW. 2007.** Phylogeography and the conservation of coral reef
714 fishes. *Coral Reefs* **26:** 501-512
715

716 **Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L,**
717 **Huelsenbeck, JP. 2012.** MrBayes 3.2: efficient Bayesian phylogenetic inference and model
718 choice across a large model space. *Systematic Biology* **61:** 539-542
719
720 **Sbrocco EJ, Barber PH. 2013.** MARSPEC: ocean climate layers for marine spatial ecology.
721 *Ecology* **94:** 979-979
722
723 **Schorr G, Holstein N, Pearman PB, Guisan A, Kadereit JW. 2012.** Integrating species
724 distribution models (SDMs) and phylogeography for two species of Alpine Primula. *Ecology and*
725 *Evolution* **2:** 1260-1277
726
727 **Siddall M, Rohling EJ, Almogi-Labin A, Hemleben C, Meischner D, Schmelzer I, Smeed**
728 **DA. 2003.** Sea-level fluctuations during the last glacial cycle. *Nature* **423:** 853
729
730 **Simmonds SE, Chou V, Cheng SH, Rachmawati R, Calumpong HP, Mahardika GN,**
731 **Barber PH. 2018.** Evidence of host-associated divergence from coral-eating snails (genus
732 *Coralliophila*) in the Coral Triangle. *Coral Reefs* **37:** 355-371
733 **Stamatakis A. 2014.** RAxML version 8: a tool for phylogenetic analysis and post-analysis of
734 large phylogenies. *Bioinformatics* **30:** 1312-1313
735
736 **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: molecular
737 evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30:** 2725-2729
738

739 **Tang QY, Liu SQ, Yu D, Liu HZ, Danley PD. 2012.** Mitochondrial capture and incomplete
740 lineage sorting in the diversification of balitorine loaches (Cypriniformes, Balitoridae) revealed
741 by mitochondrial and nuclear genes. *Zoologica Scripta* **41:** 233-247
742
743 **Thompson JD, Higgins DG, Gibson TJ. 1994.** CLUSTAL W: improving the sensitivity of
744 progressive multiple sequence alignment through sequence weighting, position-specific gap
745 penalties and weight matrix choice. *Nucleic Acids Research* **22:** 4673-4680
746
747 **Tyberghein L, Verbruggen H, Pauly K, Troupin C, Mineur F, De Clerck O. 2012.** Bio-
748 ORACLE: A global environmental dataset for marine species distribution modelling. *Global*
749 *Ecology and Biogeography* **21:** 272–281
750
751 **Van Der Meer MH, Jones GP, Hobbs J-PA, van Herwerden L. 2012.** Historic hybridization
752 and introgression between two iconic Australian anemone fish and contemporary patterns of
753 population connectivity. *Ecology and Evolution* **2:** 1592–1604
754
755 **Victor BC. 2015.** How many coral reef fish species are there? Cryptic diversity and the new
756 molecular taxonomy. Ecology of Fishes on Coral Reefs. Cambridge University Press,
757 Cambridge, United Kingdom, 76-87
758
759 **Voris HK. 2000.** Maps of Pleistocene sea levels in Southeast Asia: shorelines, river systems and
760 time durations. *Journal of Biogeography* **27:** 1153−1167

761 **Wabnitz C. 2003.** From ocean to aquarium: the global trade in marine ornamental species (No.
762 17). UNEP/Earthprint
763
764 **Waldrop E, Hobbs JPA, Randall, JE, DiBattista JD, Rocha LA, Kosaki RK, Berumen ML,**
765 **Bowen BW. 2016.** Phylogeography, population structure and evolution of coral-eating
766 butterflyfishes (Family Chaetodontidae, genus *Chaetodon*, subgenus Corallochaetodon). *Journal*
767 *of Biogeography* **43:** 1116-1129
768
769 **Warren DL, Glor RE, Turelli M. 2008.** Environmental niche equivalency versus conservatism:
770 quantitative approaches to niche evolution. *Evolution* **62:** 2868-2883
771
772 **Westneat MW, Alfaro ME. 2005.** Phylogenetic relationships and evolutionary history of the
773 reef fish family Labridae. *Molecular phylogenetics and evolution* **36:** 370-390
774
775 **Whitney JL, Donahue MJ, Karl SA. 2018.** Niche divergence along a fine-scale ecological
776 gradient in sympatric color morphs of a coral reef fish. *Ecosphere* **9:** e02015
777
778 **Wiens JJ. 2004.** Speciation and ecology revisited: phylogenetic niche conservatism and the
779 origin of species. *Evolution* **58:** 193–197
780
781 **Wiens JJ, Graham CH. 2005.** Niche conservatism: integrating evolution, ecology, and
782 conservation biology. *Annual Review of Ecology, Evolution, and Systematics* **36:** 519-539
783

784 **Wisz MS, Hijmans RJ, Li J, Peterson AT, Graham CH, Guisan A. 2008.** Effects of sample
785 size on the performance of species distribution models. *Diversity and Distributions* **14:** 763-773
786

Figure 1

Map of sampling locations. Abbreviations of the locations are given in Table 1, and the number in the parenthesis is the sample size

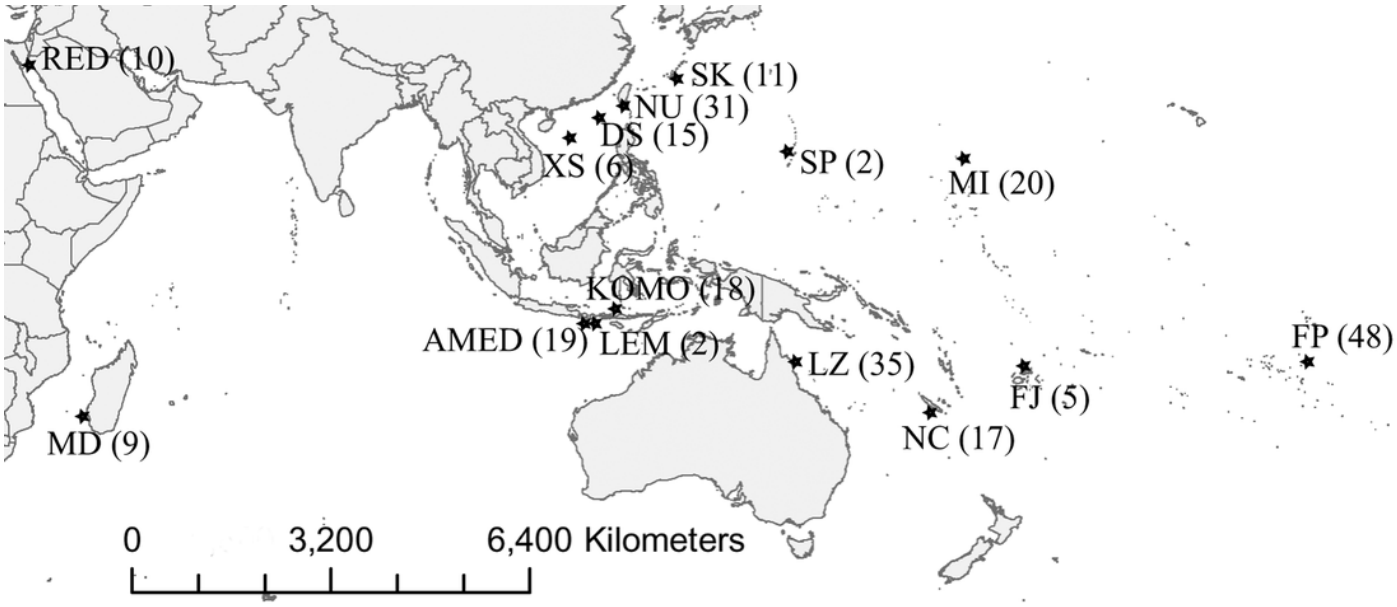
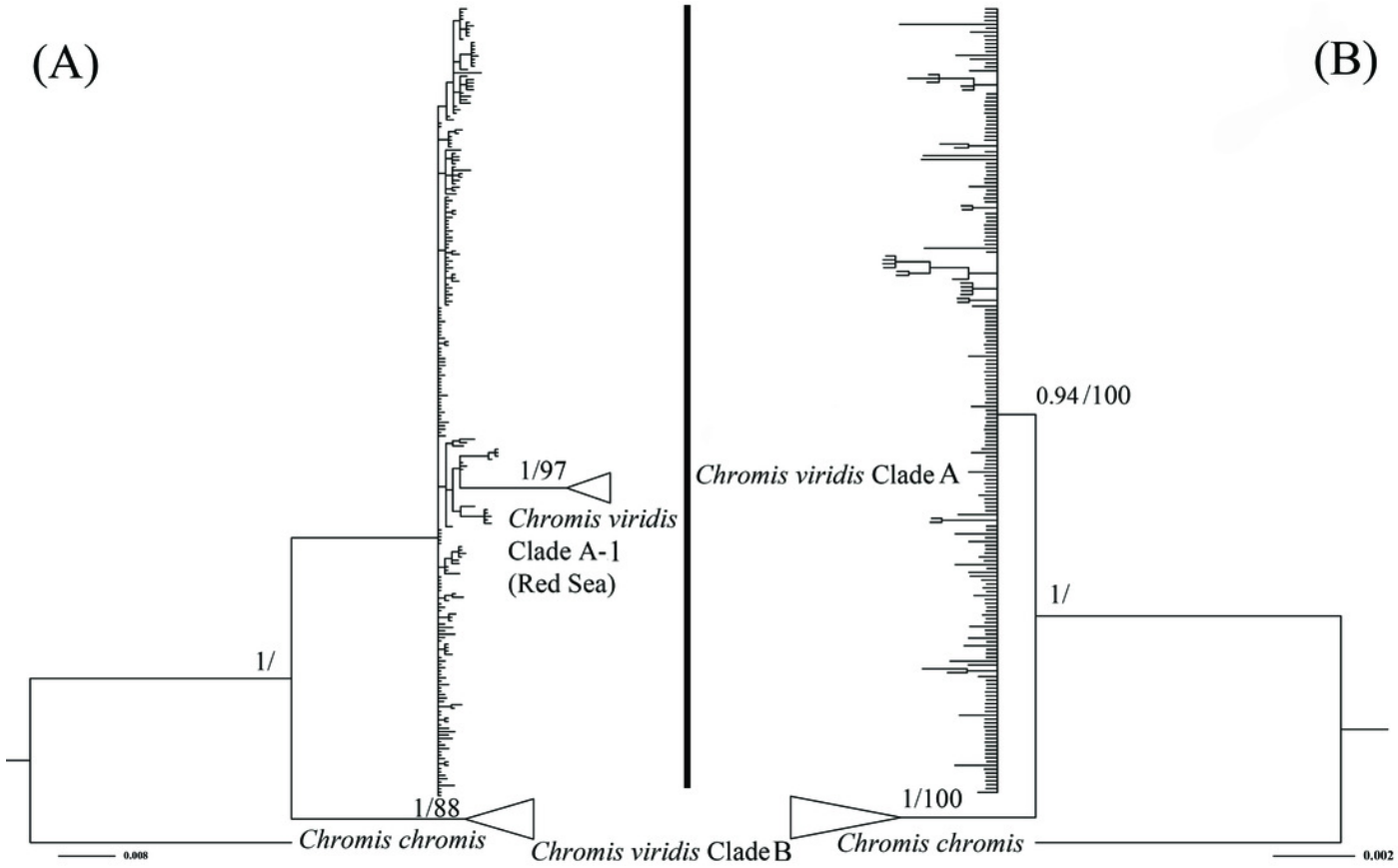


Figure 2

Bayesian phylogenetic tree (A, B) and corresponding haplotype network (C, D) based on two genetic markers including *Cytb* and Rag2.

Nodes are presented only for those with bootstrap scores >85% majority rule for maximum likelihood and >90% majority probabilities for Bayesian probability values (BI/ML). For the Haplotype network, different colors indicate different clades (e.g., white = *Chromis viridis* Clade A, gray = *Chromis viridis* Clade B, and black = Red Sea), mutation step larger than 20 were denoted by hatch marks with number of mutation steps


(A)
(B)
1/97
Chromis viridis
Clade A-1
(Red Sea)
1/
1/88
Chromis chromis
0.008
Chromis viridis Clade B
Chromis viridis Clade A
0.94/100
1/
1/100
Chromis chromis
0.002

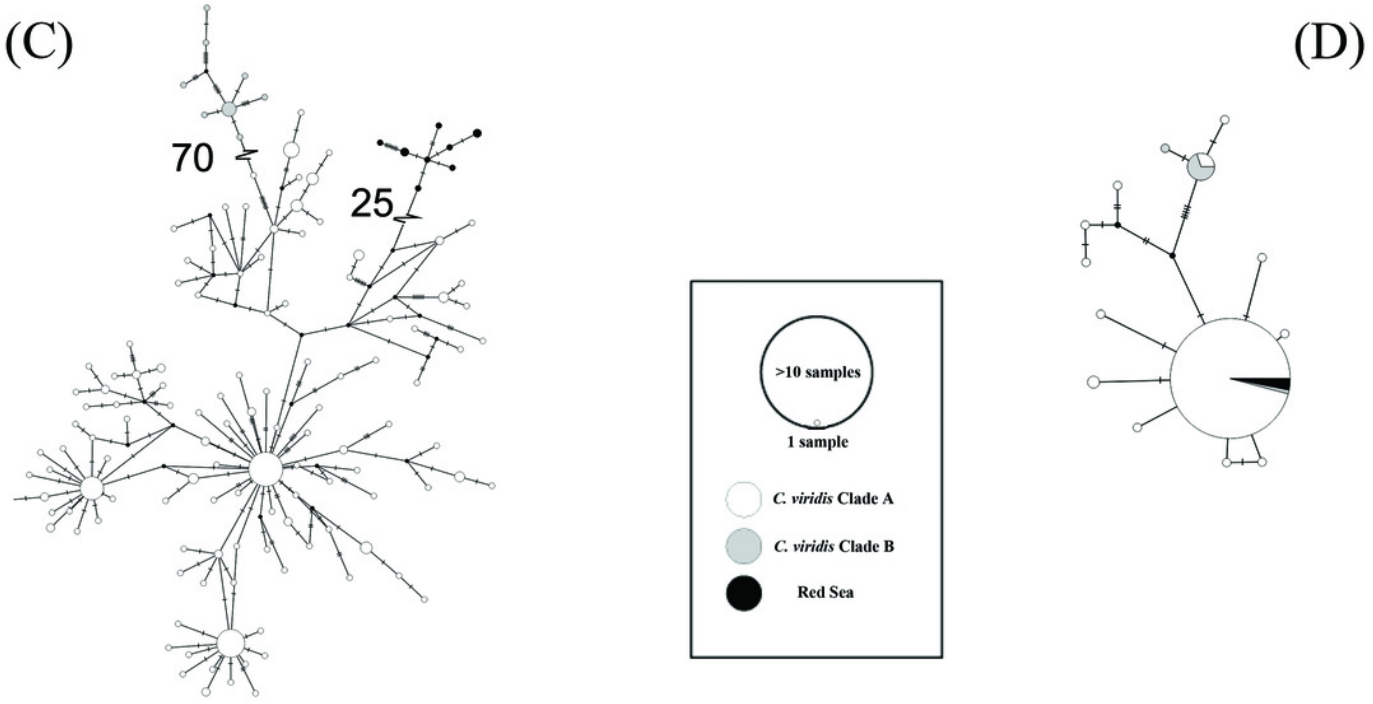

(C)
70
25
(D)
>10 samples
1 sample
C. viridis Clade A
C. viridis Clade B
Red Sea

Figure 3

Time-tree of cryptic lineages among *Chromis viridis* obtained from BEAST.

With a 3.1 MYA time constraint on the node between *Chromis atrilobata* and *Chromis multilineata* (Quenouille et al., 2004). Horizontal grey bars at nodes indicate 95% posterior probability densities (HPD) intervals of age.

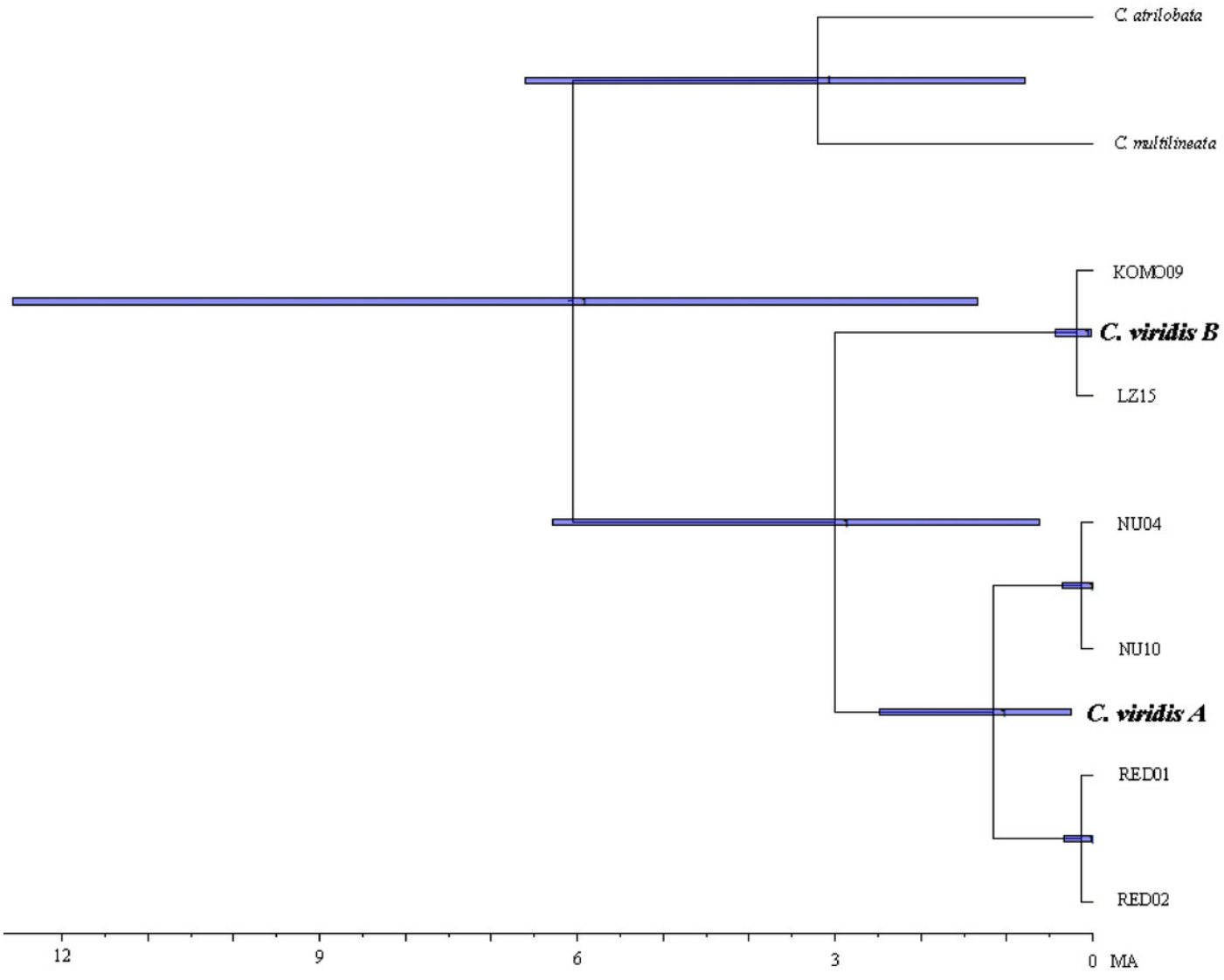


Figure 4

Non-metric multidimensional scaling (NMDS) plot of the 15 sampling sites.

The two NMDS axes (NMDS1 and NMDS2) represent environmental gradients defined by the 11 environmental factors examined across the sites. The convex hulls for the sampling sites where *C. viridis* A and *C. viridis* B were found are shown in red and blue, respectively. Please see the legend of Fig.1 for the abbreviation of the site names.

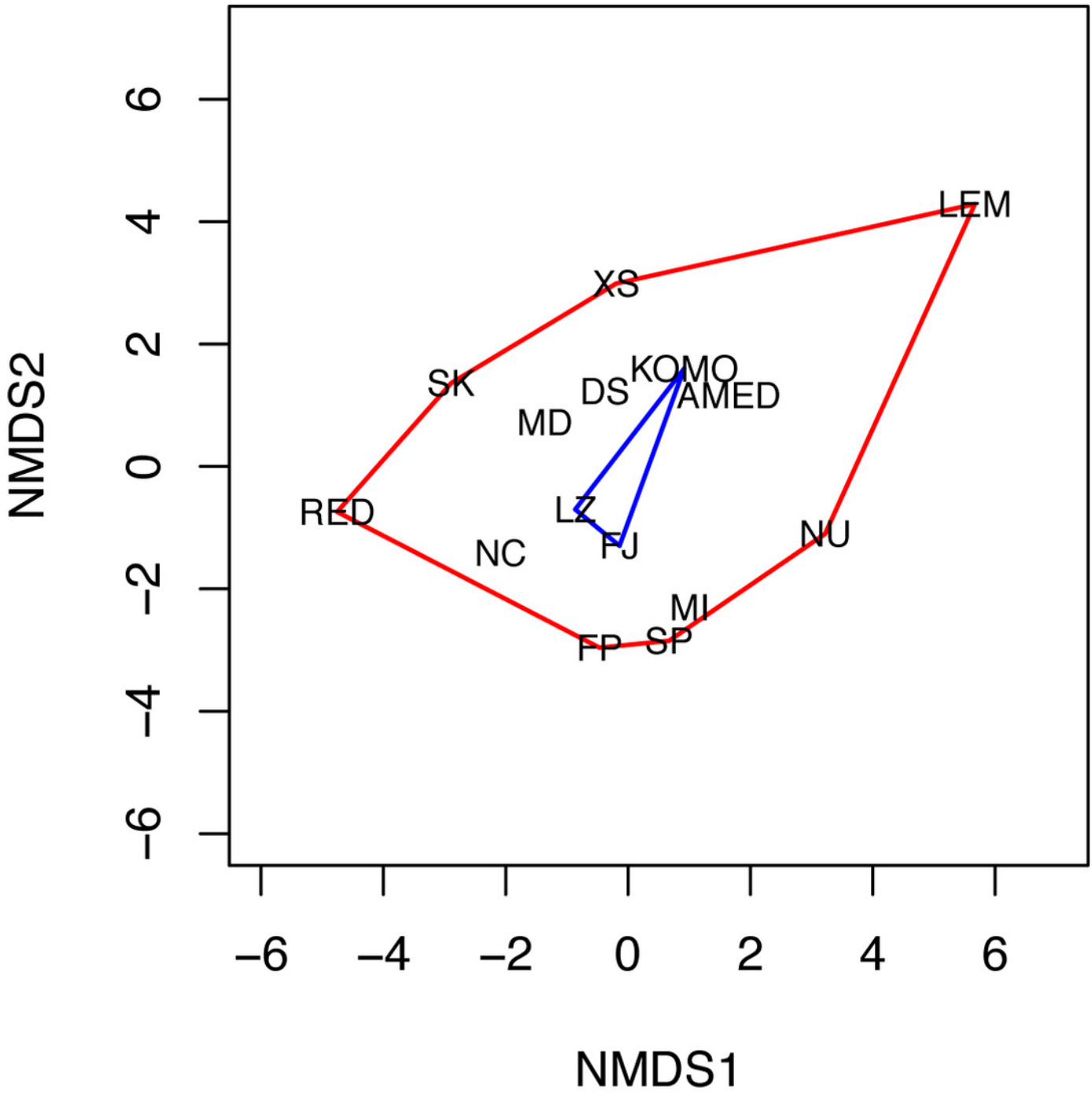

NMDS2
6
4
2
0
−2
−4
−6
NMDS1
−6
−4
−2
0
2
4
6
LEM
XS
SK
KOMO
DS
AMED
MD
RED
LZ
FJ
NC
NU
MI
SP
FP

Table 1(on next page)

Sampling locations and diversity indices based on *Cytb* sequences (911bp) in 15 populations of *Chromis viridis* from the Indo-Pacific.

Abbreviations are as follows: N = sample size, nh = number of haplotype, h = haplotype diversity, and π = nucleotide diversity.

Table 1. Sampling locations and diversity indices based on *Cytb* sequences (911bp) in 15 populations of *Chromis viridis* from the Indo-Pacific. Abbreviations are as follows: N = sample size, nh = number of haplotype, h = haplotype diversity, and π = nucleotide diversity.

		Cytb				Rag2			
Abb.	Location	N	nh	h ± SD	π ± SD	n	n*h*	*h* ± SD	π ± SD
RED	Eliat, Israel	10	8	0.956 ± 0.059	0.004 ± 0.002	6	6	1.000 ± 0.096	0.003 ± 0.002
MD	Toliara, Madagascar	9	6	0.889 ± 0.091	0.007 ± 0.004	9	9	1.000 ± 0.052	0.005 ± 0.003
AMED	Amed, Bali, Indonesia	19	14	0.953 ± 0.036	0.007 ± 0.004	18	17	0.994 ± 0.021	0.004 ± 0.002
LEM	Nusa Lembongan, Bali, Indonesia	2	2	1.000 ± 0.500	0.003 ± 0.004	2	2	1.000 ± 0.500	0.006 ± 0.006
KOMO	Komodo, Indonesia	18	11	0.882 ± 0.064	0.014 ± 0.007	18	18	1.000 ± 0.019	0.005 ± 0.003
XS	Xisha, China	6	6	1.000 ± 0.096	0.006 ± 0.004	8	8	1.000 ± 0.063	0.004 ± 0.003
DS	Dongsha, Taiwan	15	11	0.933 ± 0.054	0.005 ± 0.003	15	15	1.000 ± 0.024	0.005 ± 0.003
NU	NPP Ⅲ Inlet, Taiwan	31	24	0.972 ± 0.020	0.005 ± 0.003	23	23	1.000 ± 0.013	0.004 ± 0.002
SK	SesokoIsland, Japan	11	9	0.946 ± 0.066	0.005 ± 0.003	11	11	1.000 ± 0.034	0.005 ± 0.003
SP	Saipan, USA	2	1	0.000 ± 0.000	0.000 ± 0.000	2	2	1.000 ± 0.500	0.000 ± 0.000
MI	Marshall Island, R.O. Marshall Islands	20	15	0.942 ± 0.043	0.005 ± 0.003	13	13	1.000 ± 0.030	0.004 ± 0.003
NC	New Caledonia, French	17	16	0.993 ± 0.023	0.006 ± 0.003	15	15	1.000 ± 0.024	0.004 ± 0.002
LZ	Lizard Island, Australia	35	26	0.968 ± 0.020	0.039 ± 0.019	35	35	0.998 ± 0.007	0.009 ± 0.005
FJ	Fiji	5	5	1.000 ± 0.127	0.054 ± 0.033	4	4	1.000 ± 0.177	0.015 ± 0.011
FP	Moorea Island, French Polynesia	48	21	0.7846 ± 0.0612	0.0028 ± 0.0016	39	35	0.9933 ± 0.0080	0.0042 ± 0.0025

Table 2(on next page)

Mann-Whitney U tests for the environmental differences between presence locations of clade A and those of clade B, and between presence and absence locations of clade A.

Table 2. Mann-Whitney U tests for the environmental differences between presence locations of clade A and those of clade B, and between presence and absence locations of clade A.

Environmental Factor	A presence (Mean ± SD)	B presence (Mean ± SD)	A absence (Mean ± SD)	A presence vs. B presence		A presence vs. A absence	
				U	P-value	U	P-value
Temperature (°C)	27.44±0.86	27.15±1.35	27.08±1.47	20.5	0.86	20	0.84
Salinity (PSS)	34.31±0.86	34.55±1.43	34.61±1.57	22.5	1	18	1
Current velocity (m^{-1})	0.09±0.06	0.13±0.19	0.14±0.21	23.5	0.95	17	0.95
Nitrate ($mol \cdot m^{-3}$)	0.007±0.009	0.11±0.31	0.13±0.34	25.5	0.77	15	0.73
Phosphate ($mol \cdot m^{-3}$)	0.25±0.05	0.23±0.06	0.22±0.06	18.5	0.68	22	0.63
Silicate ($mol \cdot m^{-3}$)	3.57±1.99	3.74±1.85	3.79±1.90	23.5	0.95	17	0.95
Dissolved molecular oxygen ($mol \cdot m^{-3}$)	203.15±1.50	203.65±3.34	203.78±3.70	22.5	1	18	1
Iron ($umol \cdot m^{-3}$)	0.001±0.0003	0.0007±0.0005	0.0006±0.0005	10.5	0.17	30	0.1
Chlorophyll ($mg \cdot m^{-3}$)	0.14±0.04	0.15±0.12	0.15±0.13	18.5	0.68	22	0.63
Phytoplankton ($umol \cdot m^{-3}$)	1.19±0.16	1.14±0.48	1.13±0.54	15.5	0.44	25	0.36
Primary productivity ($g \cdot m^{-3} \cdot day^{-1}$)	0.006±0.003	0.006±0.008	0.007±0.009	17.5	0.59	23	0.54