1   Title: Evolutionary biogeography of the reef-building coral genus *Galaxea* across the Indo-

2   Pacific ocean

3   Authors: Patricia H. Wepfer[1,2*], Yuichi Nakajima[1], Makamas Sutthacheep[2], Veronica Z.

4   Radice[3], Zoe Richards[4], Put Ang[5], Tullia Terraneo[6,7], Mareike Sudek[8], Atsushi Fujimura[9],

5   Robert J. Toonen[10], Alexander S. Mikheyev[11], Evan P. Economo[2,†] & Satoshi Mitarai[1,†]

6

7   [†]S. Mitarai and E. Economo jointly supervised this research.

8   [1] Marine Biophysics Unit, Okinawa Institute of Science and Technology, 1919-1 Tancha,

9   Onna-son, Okinawa, 904-0495, Japan

10   [2] Biodiversity and Biocomplexity Unit, Okinawa Institute of Science and Technology, 1919-1

11   Tancha, Onna-son, Okinawa, 904-0495, Japan

12   [2] Marine Biodiversity Research Group, Department of Biology, Faculty of Science,

13   Ramkhamhaeng Unuversity, Huamark, Bangkok 10240, Thailand

14   [3] The University of Queensland, Australian Research Council Centre of Excellence for Coral

15   Reef Studies, School of Biological Sciences, St. Lucia QLD 4072 Australia

16   [4] Trace and Environmental DNA Laboratory, School of Molecular and Life Science, Curtin

17   University, Bentley, Western Australia, 6102 and Aquatic Zoology Department, Western

18   Australian Museum, Welshpool, Welstern Australia, 6106

19   [5] Marine Science Laboratory, Chinese University of Hong Kong, Shatin, NT, Hong Kong

20   SAR, The People's Republic of China

21   [6] Red Sea Research Center, Division of Biological and Environmental Science and

22   Engineering, King Abdullah University of Science and Technology, Thuwal 23955-6900,

23   Saudi Arabia

24   [7]ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville 4811,

25   QLD, Australia

26   [8] Coral Reef Advisory Group, Department of Marine and Wildlife Resources, Fatago 96799,

27   American Samoa

28   [9] University of Guam Marine Laboratory, UOG Station, Mangilao, Guam 96923, USA

1 [10] Hawai'i Institute of Marine Biology, University of Hawai'i at Mānoa, 46-007 Lilipuna

2 Road, Kāne'ohe HI 96744, USA

3 [11] Ecology and Evolution Unit, Okinawa Institute of Science and Technology, 1919-1

4 Tancha, Onna-son, Okinawa, 904-0495, Japan

5

6 * patwepfer@gmail.com

Abstract

8 Stony corals (Scleractinia) form the basis for some of the most diverse ecosytems on Earth,

9 but we have much to learn about their evolutionary history and systematic relationships. In

10 order to improve our understanding of species in corals we here investigated phylogenetic

11 relationships between morphologically defined species and genetic lineages in the genus

12 *Galaxea* (Euphyllidae) using a combined phylogenomic and phylogeographic approach.

13 Previous studies revealed the nominal species *G. fascicularis* included three genetically well-

14 differentiated lineages (L, S & L+) in the western Pacific, but their distribution and

15 relationship to other species in the genus was unknown. Based on genomic (RAD-seq) and

16 mitochondrial sequence data (non-coding region between *cytb* and *ND2*) we investigated

17 whether the morphological taxa represent genetically coherent entities and what is the

18 phylogenetic relationship and spatial distribution of the three lineages of *G. fascicularis*

19 throughout the observed species range. Using the RAD-seq data, we find that the genus

20 *Galaxea* is monophyletic and contains three distinct clades: an Indo-Pacific, a Pacific, and a

21 small clade restricted to the Chagos Archipelago. The three lineages of *G. fascicularis* were

22 associated with different RAD-seq clades, with the 'L' lineage showing some morphological

23 distinction from the other two lineages (larger more asymmetrical polyps). In addition to

24 these, three more genetic lineages in *G. fascicularis* may be distinguished – a Chagossian, an

25 Ogasawaran, and one from the Indian-Red Sea. Among nominal taxa for which we have

26 multiple samples, *G. horrescens* was the only monophyletic species. The mitochondrial non-

27 coding region is highly conserved apart of the length polymorphism used to define L, S & L+

28 lineages and lacks the power to distinguish morphological and genetic groups resolved with

29 genomic RAD-sequencing. The polyphyletic nature of most species warrants a careful

30 examination of the accepted taxonomy of this group with voucher collections and their

31 comparison to type specimens to resolve species boundaries. Further insight to the speciation

1 process in corals will require international cooperation for the sharing of specimens to

2 facilitate scientific discovery.

Graphical abstract

4
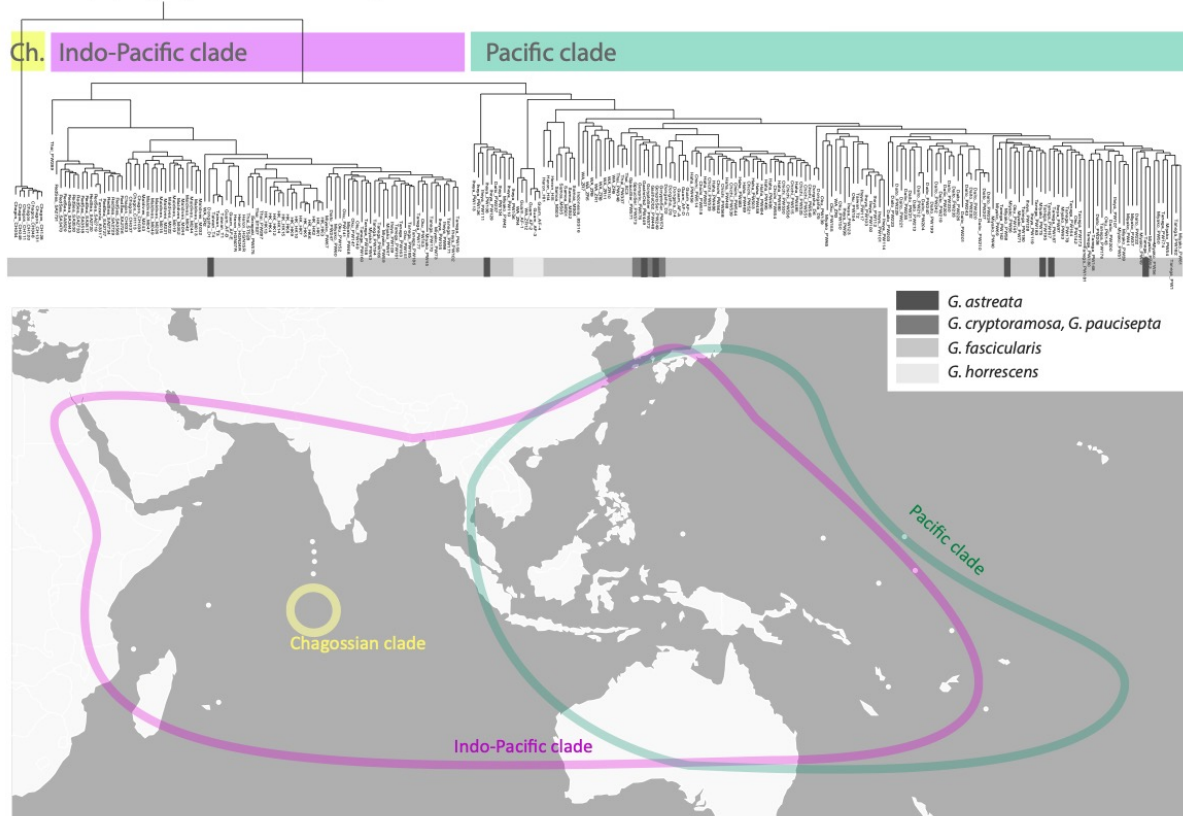
Keywords

6 RAD-seq, phylogenetics, Scleractinia, cryptic species, biogeography, mitochondrial
7 haplotype analysis, Indo-Pacific, Chagos, *Galaxea*

8

1

1. Introduction

Our understanding of scleractinian coral diversity and diversification processes is still underdeveloped despite their fundamental role in one of the world's most diverse ecosystems – coral reefs. Even on the family level the taxonomy and evolutionary history of the Scleractinia are not fully resolved (Romano and Cairns 2000, Fukami 2008, Kitahara et al. 2010) and less than half of all scleractinian species have been analyzed with modern phylogenetic methods. Traditional species delimitations based on macromorphological characters such as attributes of the corallite or colony growth form have been shown to differ from genetic classification and many taxonomic species may not represent evolutionary coherent entities, especially when comparing specimens across geographic regions (Fukami, 2008; Kitahara et al., 2010; Pinzon et al., 2013; Torres and Ravago-Gotanco, 2018). Furthermore, horizontal gene flow may be common in the Scleractinia (Mao et al., 2018; Veron, 1995; Willis et al., 2006), which further complicates the definition of species. Meaningful species delimitations are essential to understand evolution and diversification history and are crucial for the implementation of conservation measures for the protection of this threatened order (Ayre and Hughes, 2004). Using a phylogenomic and phylogeographic approach, we here attempt to shed light on the 'species problem' (Bernhard, 1902) in corals and holistically analyze the relationships between morphological, spatial, and genetic differentiation using the genus *Galaxea* Oken, 1815, as a model.

Galaxea is a small Indo-Pacific genus (ten described species, Table 1, (WoRMS, 2019)), and along with its phylogenetic sister *Simplastrea* Umbgrove, 1939, form the sister group to *Euphyllia* Dana, 1846 (Huang, 2012), although some uncertainties regarding the relationship to *Euphyllia* exist (Kitahara et al., 2016). The genus was recently reclassified from Oculinidae Gray, 1847, to Euphyllidae Veron, 2000 (Budd et al., 2012). The ten extant taxonomic species accepted to date are differentiated by colony branching patterns, the number of septa cycles, and corallite size (Veron and Stanfford-Smith, 2000). Among the ten taxonomic species the most common is *Galaxea fascicularis* L., 1767, distributed from the Red Sea to Micronesia, which is also the evolutionarily oldest species with a fossil record dated to the Oligocene (PBDB, 2018). Although *G. fascicularis* depends on its photosynthetic symbionts for nutrition (Radice et al., 2019), this species has been classified as 'stress-tolerant' because of its ability to increase particulate feeding when subject to elevated

1   seawater temperatures such as due to climate change (Ferrier-Pagès et al., 2010; Marshall and

2   Baird, 2000). The second most common taxon is *G. astreata* Lamarck, 1816, which

3   geographically overlaps with *G. fasciculari*s. The other eight species are much rarer and seem

4   to be restricted to South East Asia (Veron and Stafford-Smith 2000). The genetic coherence

5   of the taxonomic species and their phylogenetic relationships have never been investigated.

6   As in many other coral genera (e.g., *Stylophora* (Flot et al., 2011), *Acropora* (Ladner

7   and Palumbi, 2012), *Pocillopora* (Combosch et al., 2008), *Heliopora* (Yasuda et al., 2015;

8   Yasuda et al., 2014), and *Seriatopora* (Warner et al. 2015), there are morphologically

9   'cryptic' but genetically highly differentiated lineages within the taxon *G. fascicularis* (for

10   definition of 'cryptic' see (Bickford et al., 2007)). These lineages are relatively well-studied

11   in the Ryukyu Islands, Japan, where two distinct types of *G. fascicularis* had originally been

12   distinguished based on variation in the nematocyst anatomy (Hidaka 1992). They were later

13   found to differ in the length of a mitochondrial non-coding region by almost 300 bp

14   (intergenic region between *cytb* and *ND2*) (Watanabe et al. 2005), and microsatellite markers

15   revealed that the they were genomically well differentiated lineages (Abe et al., 2008;

16   Nakajima et al., 2015). Reproductive studies observed shifted spawning times in the field

17   (Heyward et al., 1987; Yamazato, 1988) and the lineages to rarely cross-fertilize under

18   laboratory conditions (Abe et al. 2008a). According to their mitochondrial length variation

19   the two lineages had been referred to as 'S' and 'L' for a short or a long intergenic region,

20   respectively. A third lineage from Japan 'L+' was found more recently, which has three more

21   base pairs than 'L' in the respective mitochondrial region and differs from both lineages in

22   the nuclear genome (Nakajima et al. 2016). The three lineages exist in sympatry on the coral

23   reefs in in the Ryukyu Islands (Hidaka 1992), indicating either sympatric ecological

24   segregation or neutral differentiation in an allopatric past, e.g. the currently observed

25   sympatry could be the result of a relatively recent breakdown of a dispersal barrier, such as

26   sea level rise after the Pleistocene (Bowen et al., 2013; Carpenter et al., 2010). However,

27   their geographic distributions elsewhere or potential microhabitat differentiation have not

28   been studied.

29   On the other hand of cryptic diversity, phenotypic plasticity of colony form within a

30   lineage can result in a single valid species encompassing multiple nominal taxa - e.g.,

31   *Pocillopora* (Johnston et al., 2017; Marti-Puig et al., 2014; Paz-Garcia et al., 2015),

32   *Stylophora* (Arrigoni et al., 2016), *Seriatopora* (Bongaerts et al., 2011), *Montipora* (Forsman

33   et al., 2010), *Porites* (Forsman et al., 2009). One way to identify cryptic species and

1 phenotypic plasticity is to examine phylogenetic relationships along a deeper time scale and

2 across a more inclusive phylogenetic group instead of a locally restricted or taxonomically

3 pre-selected number of species (Bickford et al., 2007). The extent of the genetic

4 differentiation between the lineages in *G. fascicularis* has never been compared to other

5 species within the genus and it is unclear how these lineages phylogenetically relate to each

6 other and other lineages across the taxonomic range of *Galaxea*.

7 Here we investigated the relationships between the taxonomic species and genetic

8 lineages in *Galaxea*. We specifically asked whether the nominal species based on gross

9 colony morphology represent biologically meaningful entities from a phylogenetic

10 perspective, and how the cryptic lineages in *G. fascicularis* are related and distributed across

11 the nominal species range. We gathered field and museum collections of the taxonomic

12 species *G. fascicularis, Galaxea horrescens* Dana*, 1846, Galaxea cryptoramosa* Fenner &

13 Veron, 2000*, G. astreata,* and *Galaxea paucisepta* Claerboudt, 1990, across the Indo-Pacific

14 and used restriction site-associated DNA (RAD) sequencing to obtain a thorough genomic

15 delineation. RAD-tag sequences are useful for both population genetics (Andrews et al.,

16 2016) and to address phylogenetic inference between recently diverged lineages (Cariou et

17 al., 2013; Emerson et al., 2010), and are therefore ideal for analyzing within-species

18 differentiation in *G. fascicularis,* as well as interspecific relationships in the genus *Galaxea*.

19 We investigated depth segregation and analyzed polyp size variation between the

20 morphologically cryptic lineages in *G. fascicularis* to see whether they could vary in their

21 habitat as a potential mechanism of ecological segregation (Prada and Hellberg, 2013;

22 Serrano et al., 2014). We further analyzed the characteristic mitochondrial non-coding

23 sequence between the genes *cytb* and *ND2* (Watanabe et al., 2005)*.* By geographically

24 mapping the distributions of the mitochondrial, morphological, and genomic diversity, we

25 finally discuss potential influences of biogeographic processes for the evolutionary history of

26 *Galaxea.*

2. Material and Methods

2.1. Specimen collection and identification

29 *Galaxea* specimens were collected from across the Indo-Pacific distribution range of the

30 genus (Fig. 4) and of six nominal species (*G. fascicularis* (589)*, G. astreata* (9)*, G.*

31 *cryptoramosa* (4)*, G. paucisepta* (2)*, G. horrescens* (6)*, G. longisepta* (1 museum specimen).

32 Field collections were gathered from the Red Sea (15, King Abduhllah University of Science

1 and Technology), Maldives (10), Chagos (10), the Great Barrier Reef (5, University of

2 Queensland), Western Australia (12, Curtin University), Thailand (205, Ramkamhaeng

3 University), Taiwan (6) and Dongsha (6, Academia Sinica), Japan (292), Hong Kong (13,

4 University of Hong Kong), American Samoa (5) (Coral Reef Advisory Group), and Guam (9,

5 University of Guam). To further increase geographic coverage, field collections were

6 complemented with museum specimens from the National Museum of Natural History,

7 Washington D. C. (Smithsonian Institution, 18), the Naturalis Biodiversity Center, Leiden

8 (2), and the Museum of the University of the Ryukyus Fujukan, Nishihara (16). All

9 specimens used, including their sampling location and available metadata, are listed in Tables

10 S1 and S2.

11   Species identification was performed analyzing field photographs (Fig. 1) and

12 remaining collection material when available (Table S1). In particular, following Veron,

13 (2000) and an unpublished taxonomic treatment given by van der Veer (2007), the following

14 morphological characters were considered to identify species: the number of septa cycles, the

15 size of polyps, and branching morphology. In *G. fascicularis*, primary and secondary septa

16 are similar or same in size so that the number of primary septa appears to be irregular or

17 extended. This feature can be observed through the coral tissue, which is why this taxonomic

18 species may be readily identified in the field or from field photographs, together with the

19 feature of massive and not branching colony morphology. Specimens with laminar growth

20 form and in which polyps had unequal septa cycles containing six uniform septa each and for

21 which polyps were > 3.5 mm in diameter were assigned to *G. astreata*. Specimens that were

22 similar to *G. astreata* but had smaller polyps (< 3.5 mm) with strictly 2 septal cycles, out of

23 which the second did often not reach the columnella, were assigned *G. paucisepta*. The

24 identification of these two species required the examination of polyp skeleton material at the

25 corallite level (Table S1). For further details regarding the identification of *G. astreata, G.*

26 *paucisepta* and *G. cryptoramosa* see Supplementary Information. Specimens that were thinly

27 branched and had small polyps shorter than the width of the branch they were sitting on were

28 assigned *G. horrescens*. Specimens that exhibited any form of irregular branching patterns

29 were assigned *G. cryptoramosa* following Van der Veer (2007). Identification of specimens

30 from Dongsha Atoll and Taiwan Island, and most from Western Australia were visually

31 confirmed by the sample providers from these areas (Allen Chen, Zoe Richards,

32 respectively).

1   Six outgroup specimens were also collected and added to the phylogenetic analysis

2   for rooting purposes and to test for the monophyly of *Galaxea*. Three species within the

3   Complex clade of the Scleractinia with two specimens each were chosen, including the

4   closely related *Euphyllia* c.f. *ancora* and *Pachyseris* c.f. *speciosa,* and two specimens of

5   *Acropora digitifera* (Huang 2012).

6   **2.2. DNA extraction**

7   Holobiont DNA from field collections were extracted using the DNeasy® Blood &

8   Tissue Kit (Qiagen, Hilden, Germany). The manufacturer's protocol was followed with

9   modifications of an extended initial incubation time for tissue lyses at 56ºC (4 –10 h), the

10   addition of 4ul of 100x RNAse A after lysis, and the application of a 1.5–2 times larger

11   volume of EtOH for denaturation for separating extensive amounts of mucus from the watery

12   phase.

13   Archival DNA was extracted from 40 museum samples that were satisfying the

14   criteria of having sufficient material, being of acceptable quality (i.e. without visible mold or

15   algal contamination), and not showing signs of chemical treatment for preservation purposes

16   (i.e. smell of xylene). Specimens from the National Museum of Natural History, Washington

17   D.C., and the Naturalis Biodiversity Center, Leiden, were extracted and treated in

18   collaboration with the ToBo laboratory at the Hawaii Institute of Marine Biology.

19   Particularly careful precautions against contaminations were taken for the extraction of

20   museum specimens as DNA is usually degraded and the yield is low, including sterilization

21   of tools with 10% bleach, 99% EtOH, and Bunsen burner in between processing of each

22   specimen, and autoclaving of tubes and tips. To remove potential surface contamination from

23   the dried specimens, skeleton pieces were soaked in 70% EtOH for 10 min to 1 h and air-

24   dried. The DNeasy® Blood & Tissue Kit extraction protocol was adjusted to a larger quantity

25   of extraction material of 0.2–1.3g per specimen and a longer denaturation incubation time of

26   18–22 h at 56 ºC with a larger amount of extraction buffer and Proteinase K (up to 10x

27   more). After this step, the manufacturer's protocol was followed. Based on yield and quality

28   of the DNA fragments (> 500 bp bands on Agarose gel after electrophoresis), 28 specimens

29   were chosen for sequence analysis.

1   **2.3. RAD-seq analysis**

2   *2.3.1 Library preparation and assembly*

3   A total of 293 specimens were genotyped using RAD-tag sequencing. We used a RAD-

4   protocol (Tin et al., 2014) that is designed for low quantities of degraded DNA and may

5   therefore be suited to process marine invertebrate DNA for which quantity and quality are

6   often low. It involves a single digestion with the restriction enzyme EcoR1 and produces

7   short final fragments of 35-50 bp. Libraries were single-end sequenced using the Illumina

8   HiSeq platform at the Okinawa Institute of Science and Technology. Raw reads were quality

9   filtered and trimmed using Trimmomatic v.0.35 (Bolger et al., 2014). Samples with non-

10   sufficient amplification (many of the museum specimens) were dropped, leaving 272

11   specimens for loci assembly. Raw sequences were submitted to Genbank (BioProject

12   PRJNA576132, BioSamples SAMN12925065-SAMN12925355).

13   Loci were assembled in ipyrad v.0.7.19 (Eaton 2015) based on partially assembled

14   *Galaxea* reference sequences provided by the ReFuGe2020 consortium (Liew et al., 2016;

15   Voolstra et al., 2015) and unassembled raw reads from a previous study (Nakajima et al.,

16   2015). Reads were filtered to be minimally 35 bp long to enter assembly analysis in ipyrad

17   and clustered using a threshold of 0.9. A minimum depth of 6 and maximum depth of 10,000

18   within samples were used for base calling. Only biallelic sites were considered. Maximally,

19   four uncalled bases (Ns) and eight heterozygotes in consensus sequences were accepted and a

20   locus was allowed to have maximally 10 SNPs and 8 indels. A locus needed to be represented

21   in at least three samples. It has been shown that the random loss of loci due to low sequence

22   coverage across specimens of hierarchical redundancy should not affect deeper phylogenetic

23   relationships (Eaton et al., 2017). Therefore, and based on experience with a similar RAD

24   approach in other phylogenetic projects (Fischer et al., 2015, Darwell et al., 2020), we

25   accepted this relatively low vertical coverage of three specimens per locus (and a high

26   gappyness of our data) in our main analysis, in order to reveal the maximal resolution of

27   deeper relationships between groups of hierarchically equal specimens. However, as a

28   supplemental analysis the effect of higher vertical coverage on phylogenetic relationships

29   was also evaluated by filtering the final dataset to loci contained in at least 27 (10%) or 68

30   (25%) specimens. Finally, loci were trimmed by 5 bp at the 5' end by 5 bp because these

31   contained too many inconsistent variable sites. From the mapping statistics, we then again

32   excluded *Galaxea* individuals that had less than 1000 loci (two individuals). The ipyrad read

1   statistics and phylip file may be retrieved from the supplement (Table S4 and supplementary

2   file 1).

3   In order to assess the risk of contamination of our RAD-reads and the reference

4   genome with symbiont DNA, the assembled loci were blasted against a custom

5   Symbiodiniaceae database using the same clustering threshold as ipyrad (90% identity). The

6   Symbiodiniaceae database was composed of published genomic and transcriptomic

7   sequences: genome clade B (Shoguchi et al., 2013), transcriptomes of subclades A, A3, B,

8   B1, C1 (Pinzón et al., 2015), and transcriptomes of clades C and D (Ladner et al., 2012). We

9   further screened our loci for bacterial, archaeal, or viral contamination using a *k*-mer based

10   identification approach in Kraken v.1 (Wood and Salzberg, 2014). We used the ready-built

11   KrakenMini_DB_8GB provided by the program developers.

12   *2.3.2. Phylogenetic inference and network construction*

13   From the filtered 272 specimens a phylogenetic tree was estimated based on the SNP phylip

14   output file with a maximum likelihood approach using ExaML v.3 (Kozlov et al., 2015).

15   Twenty random starting trees were generated using RAxML v.8.2.4 (Stamatakis, 2006) and

16   given as input to ExaML. ExaML was run under the PSR model to find the most likely tree.

17   The same was also done in our supplemental, more filtered datasets (loci coverage > 27, or >

18   68 specimens). Node supports were estimated for our main analysis (3 specimens) using

19   Bootstrap analysis with 468 iterations. Bootstrapped alignments were created in RAxML and

20   likelihood searches were performed in ExaML as described above. Convergence of the

21   Bootstrap replicates was confirmed in RAxML using the autoMRE option (converged after

22   450 replicates). Direct supports (frequencies) for bipartitions were drawn on the tree using

23   RAxML. In addition, we used Booster (Lemoine et al., 2018) to calculate branch supports.

24   Booster implements a newly developed method of gradual distance measurements between

25   branches and is thought to perform better for large datasets derived from next generation

26   sequencing than traditional Felsenstein-statistics, because the presence of a single uncertain

27   specimen results in the uncertainty of the whole clade (Lemoine et al., 2018). The booster

28   instability metric for each specimen is given in Table S3. Specimens with high booster

29   instability are considered to be of uncertain phylogenetic position (Lemoine et al. 2018) and

30   may explain low Felsenstein Bootstrap supports of clades that include them.

31   To further evaluate the evolutionary relationships in *Galaxea* and to detect the

32   potential existence of incompatible loci in the genomes, we computed a network using the

1 Neighbor-Net algorithm implemented in SplitsTree v.4.15.1 (Huson, 1998). We used the

2 Hamming method to calculate pairwise distances based on the same SNPs as for the

3 phylogeny.

4 *2.3.3. Admixture and DAPC analysis*

5 In order to investigate potential species delimitations the 264 *Galaxea* specimens (without

6 outgroup specimens) were further analyzed in their genetic structure, using Admixture v.1.2.

7 (Durand et al., 2011) and discriminant analysis of principal components (DAPC). For these

8 the ipyrad VCF output file was filtered using VCFtools (Danecek et al., 2011) to only include

9 SNPs that were present in at least 50% of all *Galaxea* individuals, which reduced the number

10 of sites to 2275. The VCF file was transformed into a Plink bed file using PLINK v.1.9

11 (Chang et al., 2015) before running Admixture with default settings for K=1, 2, 3, …, 21. K

12 with the smallest cross validation error (CV error) was inferred to determine the most likely

13 number of ancestral lineages (Fig. S3). The results were plotted as barplots in R v.3.4 (R

14 Core Team, 2015), once ordered by sites and cluster (Figs. 3) and once ordered by sites and

15 individuals (Fig. S4).

16 A series of DAPC was run using the R adegenet package (Jombart and Ahmed, 2011).

17 We first analyzed all *Galaxea* samples together. A principal component analysis (pca) was

18 performed using the grPca function. The best number of clusters *K* was found using the

19 function find.clusters based on the lowest value of the Bayesian Information Criterion (Fig.

20 S6). As several highly similar *K* values were found (*K*=6, 7, or 8), each of the two main

21 *Galaxea* clades from the phylogeny (Pacific and Indo-Pacific clade) were also analyzed

22 separately. The number of principal components to retain was chosen so that at least 80% of

23 the variation is retained, resulting in 100 principal components for the analysis on all

24 specimens and the one on the Pacific clade, and 60 principal components for the analysis on

25 the Indo-Pacific clade (Fig. S7). The clusters were then analyzed using DAPC and plotted as

26 scatterplots (Fig. 5). Cluster assignments of individuals for all five analyses are given in

27 Table S5.

28 **2.4. Mitochondrial haplotype analysis**

29 The *Galaxea* characteristic mitochondrial non-coding region between *cyt b* and *ND2* was

30 analyzed by Sanger sequencing using published primers and protocols (Nakajima et al.,

31 2016). Each polymerase chain reaction (PCR) contained 1µl of 8 µM primers, 2 µl MilliQ

32 water, 5 µl AmpliTaq Gold Master Mix (Applied Biosystems, Thermo Fisher Scientific), and

1 1µl holobiont DNA. PCR products of successfully amplified samples were purified using

2 ExoSAP-IT (Affymetrix, Santa Clara, CA, USA) and sent for single-end sequencing to

3 Macrogen Japan Corporation, except for 16 museum specimens from Fujukan, which were

4 sequenced in-house following (Nakajima et al., 2016). In total 135 specimens were

5 sequenced in their mitochondrial region (chromatograms may be viewed in Supplemental file

6 2).

7 DNA sequences were examined and processed using Geneious v. 9.1.2 (Biomatters

8 Ltd.). Low quality base calls at the ends and primer sequences were removed. Some

9 specimens, especially the museum specimens, showed signs for containing multiple

10 haplotypes, i.e. both the longer L and shorter S sequences, resulting in double peaks in the

11 DNA chromatograph (indicated in Table S1). For these specimens, only the dominant

12 sequences were taken if the peaks were an order of magnitude larger than those of the minor

13 background sequences, and if they were identical to a sequence in at least one of the other

14 specimens. Sequences of too low quality were removed entirely. The clean sequences were

15 aligned to each other and previously published haplotype sequences by Watanabe *et al.*

16 (2005) (LA-LE and SA-SC), and Nakajima *et al.* (2016) (Watanabe's LA and SA as "L1"

17 and "S1", respectively, "L2" for here called LF and L+, Genbank accession numbers

18 LC155810 - 3). Previously unknown sequences (LG – LP, SB) were submitted to Genbank

19 (accession numbers MK054259 - MK054269). TCS haplotype network topology was inferred

20 using TCS v. 1.21 (Clement et al., 2002) and the network was drawn using PopART (Leigh

21 and Bryant, 2015) and Adobe Illustrator. The geographic distribution of the haplotypes was

22 mapped using PopART and adjusted in color using Adobe Illustrator and GIMP software.

23 The Nexus input file for PopART and TCS are given in the Supplement (Supplementary file

24 3).

25 For 199 additional specimens we determined the mitochondrial main type S, L, or L+

26 by fragment length analysis following the procedures described in Nakaema and Hidaka

27 (2015). Specimens were assigned type L if they had a fragment size of 457 bp, S if their

28 fragment size was 167 bp, and L+ if their fragment size was 460 bp. Specimens with equally

29 abundant multiple fragment sizes were excluded from the analysis. In order to infer the

30 relationship between the mitochondrial and genomic differentiation, the mitochondrial types

31 were mapped onto the RAD-seq phylogeny where available.

2.5. Morphological and depth-differentiation between lineages of *G. fascicularis*

In order to infer any indications for potential ecological evolution between the three lineages 'L', 'S', 'L+' in *G. fascicularis*, we compared their depth distribution and skeletal corallite morphology of 334 specimens from two geographic regions (Thailand and Japan). For most specimens from Japan both RAD and mitochondrial data was available, for other specimens we only used mitochondrial fragment length data for lineage identification (Table S1).

We assessed depth distributions among the lineages based on depth recordings for 176 specimens from Japan and 157 specimens from Thailand using a diver computer. The reading was corrected for the tidal level of the sampling site at the time of collection to represent average depth. The distributions were visualized by boxplots for each sampling site and lineage separately using R.

We quantified polyp sizes in 157 specimens from the Ryukyu, Daito and Ogasawara Islands (Japan) based on size-standardized field photographs using Fiji v.2.0.0 (Schindelin et al. 2012). For each specimen, polyp maximal diameter, polyp minimal diameter, and distances between polyps were measured in 3–5 measurements from representative polyps of a colony and averaged within a specimen. Fractions of minimal and maximal diameters were calculated ('shape'), and relative distances between polyps were calculated as fractions of measured distance to the maximal diameter ('dist.rel'). A principal component analysis (PCA) was performed to depict morphological variation in two dimensions using the morph.pca function and plotted with the ggbiplot function in R. Variation in maximal polyp diameter between the lineages was additionally tested in a Kruskal-Wallis rank sum test, after confirming a non-normal distribution of this trait in a Shapiro-Wilk normality test. Skeletal features of the corallite (septa) were additionally examined based on 83 specimens with available material (Table S1), including the number of cycles and the number of septa in each cycle.

3. Results

3.1. RAD-seq phylogeny and network

The RAD-seq analysis revealed 214,705 loci and 456,846 unique patterns that were shared by 3 or more individuals. Most museum specimens (26) except for two from Tanzania and Indonesia had to be discarded due to insufficient data (less than 1000 loci). In addition, four field specimens from Miyako (1), Taiwan (2), Daito (1), and Chichi Island (1) were also removed due to insufficient data. Most loci were represented in less than a quarter of the

1   individuals, resulting in a relatively high 'gappyness' of our data (0.901, see full ipyard read

2   statistics in Table S4). In average, 14322 loci were covered in a *Galaxea* individual. This

3   number was smaller in the outgroup specimens (1072 loci), as well as in the Chagossian

4   small clade (9597 loci) and in *G. horrescens* (5873 loci, Table S4). Only 27 loci (0.01%)

5   mapped to sequences of Symbiodiniaceae and none of the loci were classified to be of

6   bacterial, archaeal, or viral origin by *k*-mer analysis. For the filtered datasets we retrieved

7   183,500 and 29,657 unique patterns that were present in at least 27 or 68 specimens,

8   respectively.

9   The node supports for deeper nodes and nodes clustering geographic locations were

10 high (Booster supports 0.8-1, Felsenstein bootstraps 60-100%). Supports were lower for

11 nested, terminal clades within geographic regions (Booster supports <0.8), especially

12 considering direct bootstrap frequencies (0- 20% for some nodes in the Pacific clade). Clades

13 with very low bootstrap values contained at least one specimen with a high Booster

14 instability score >1 (Fig. 2, Table S3). Specimens with high Booster instability scores tended

15 to have fewer loci (8769) than other specimens (14308) in average (Table S3), which may be

16 responsible for some of the instability.

17 The phylogeny confirmed that the genus *Galaxea* is monophyletic with respect to the

18 outgroup genera *Euphyllia, Pachyseris*, and *Acropora*. *Galaxea* largely clustered into three

19 well-supported main clades (Fig. 2): a small clade only represented in Chagos and sister to all

20 other specimens (hereafter referred to as 'Chagossian Clade'), an Indo-Pacific clade

21 containing specimens from the Red Sea, the Indian Ocean, and lineage 'S' from the central

22 Indo-Pacific (hereafter referred to as 'Indo-Pacific Clade'), and a Pacific clade comprising

23 lineage 'L+', lineage 'L' and all nominal species (hereafter referred to as 'Pacific clade', Fig.

24 2). This was consistent in our supplemental analysis with higher vertical coverage, except for

25 unstable placement of the outgroup *Acropora* (Figs. S1, S2), which is probably related to low

26 coverage recovered for that taxon.

27 Within these main clades, specimens cluster according to geographical closeness. In

28 the Indo-Pacific clade, specimens from the Red Sea formed a sister group to an Indian Ocean

29 clade (Chagos and Maldives) and a clade containing samples from Asia and Australia. An

30 exception in the geographical structuring represented a specimen from Thailand PW289,

31 which was sister to all other individuals in this clade, which was confirmed to be a hybrid of

32 the two main clades (Fig. 3), In the Pacific clade, the *G. fascicularis* 'L+' - lineage formed a

1 strongly supported clade sister to the taxonomic species *G. horrescens* and all other

2 specimens. The remaining specimens within the Pacific clade grouped to a south-eastern

3 Pacific subclade containing specimens from American Samoa and the Great Barrier Reef, and

4 two subclades containing Western Australian and Asian specimens, respectively. Finer

5 geographic resolution could not be obtained and clades according to sampling sites within

6 island archipelagos or regions remained unresolved. Notably, all islands of the Ryukyu

7 archipelago (Japan) were mixed in a Ryukyu clade, unless they belonged to clonal clusters

8 (for example specimens from Iheya in the Pacific clade, see Wepfer (2018)). In the

9 supplemental analysis using fewer loci the geographic clustering was less clear (Figs. S1, S2)

10 Taxonomically, *Galaxea fascicularis* was polyphyletic, occurring throughout the

11 phylogeny. *Galaxea astreata* was also polyphyletic represented across most parts of the

12 Pacific clade and with two specimens also in the Indo-Pacific clade. All other taxonomic

13 species (*G. horrescens, G. cryptoramosa, G. paucisepta*) were included in the Pacific clade.

14 *Galaxea horrescens* was monophyletic with specimens from two geographic regions (Guam

15 and Western Australia). *Galaxea astreata, G. paucisepta* and *G. cryptoramosa* from the

16 Ryukyu Islands were genomically undifferentiated from each other but clearly distinct from

17 *G. fascicularis* specimens in the Ryukyu Islands. The cryptic lineages in *G. fascicularis* were

18 split between the main clades; lineage 'S' was nested in the Indo-Pacific clade, and lineages

19 'L' and 'L+' were grouped in the Pacific clade.

20 The Neighbor-Net network (Fig. 3) was consistent with the phylogeny in

21 distinguishing a Pacific clade L, clade L+, an Indo-Pacific clade, and a Chagossian clade. *G.*

22 *horrescens* was placed closer to the outgroup *Euhphyllia* in the network instead of being

23 nested within the Pacific clade like in the phylogeny, The geographic grouping within the

24 main clades was also less clear than in the phylogeny, except for the somewhat distinct clades

25 defined by specimens from Ogasawara, the Indian Ocean and the Red Sea.

26 **3.2. Admixture and DAPC**

27 Based on 2275 filtered SNPs (filtered to be represented in 50% of all specimens), Admixture

28 determined K=7 as the most likely number of ancestral lineages in the genus *Galaxea* (Figs.

29 4, S3). K=3 and K=5 were the next most likely numbers of ancestral lineages in our data and

30 are shown to infer relatedness between the seven main lineages. Admixture agreed with the

31 phylogeny in identifying three main groups, a Chagossian group, an Indo-Pacific group

32 containing the Admixture lineages 'Indo-Pacific-a' from the Indian Ocean and '-b' from the

1   central Indo-Pacific, and a Pacific group containing four lineages ('Pacific-a' to '-d').

2   'Pacific-a' is spread across the entire Pacific and contains apart from *G. fascicularis* also *G.*

3   *astreata, G. paucisepta,* and *G. cryptoramosa* from Okinawa, while 'Pacific-b' and 'Pacific-

4   c' are local to the Ryukyu Islands and Ogasawara, respectively. 'Pacific-d' represented by *G.*

5   *horrescens* and *G. fascicularis* lineage 'L+' is a well-defined already at the K=5-level and is

6   a mix of the ancestral Chagossian and Pacific lineages at the K=3-level. One Thai specimen

7   (Thai_PW289) was mixed of ancestral lineages belonging to different phylogenetic clades,

8   containing 'Indo-Pacific-b' and 'Pacific-a' to equal parts (Fig. S4).

9   DAPC found *K*=6, 7, or 8 to be the most likely number of clusters in all *Galaxea,* and

10   *K*=4 to be the most likely number of cluster in the Pacific clade and Indo-Pacific clade (Fig.

11   S6, Table S5). The Chagossian clade clustered very distantly to all other clusters in the

12   analyses over all *Galaxea* ('Cluster Chagos', Fig. 5a). After that, all analyses agreed in a

13   cluster for lineage L+ ('Cluster L+) and one for the Ogasawaran specimens ('Cluster

14   Ogasawara', Fig. 5a, b). The analysis on the Pacific clade further distinguished a cluster for

15   *G. horrescens* ('Cluster Horrescens') from a large Asia-Pacific cluster ('Cluster L', Fig. 5b),

16   and the analysis on the Indo-Pacific clade distinguished a cluster for the Red Sea ('Cluster

17   Red Sea') from the Indian Ocean ('Cluster Indian') and separated a cluster from mainly Hong

18   Kong ('Cluster Hong Kong') from the other Asian and Australian specimens ('Cluster S',

19   Fig. 5c, Table S5).

20   **3.4. Mitochondrial haplotype diversity**

21   Across all locations and taxonomic species, 12 mitochondrial haplotypes were found, two

22   haplotypes that were 135 bp short (S subtypes) and 10 haplotypes that had the longer 467 bp

23   or 470 bp sequences (L subtypes: LA, LG-LP, L+). LA and SA were the most widely

24   distributed and most frequent types. Most of the other haplotypes differed only by a single

25   substitution from LA and were only represented in one specimen. These rare haplotypes were

26   collapsed into LA for simplicity (Fig. 6). The complete resolution of the haplotype network

27   including the haplotypes that were not found here but were reported previously (LA-LE, SA,

28   SB Watanabe et al., 2005), and (LF; Nakajima et al*.,* 2016) may be retrieved from the

29   supplement (Fig. S5). SB only differed by one bp from SA but was common in Taiwan and in

30   the Great Barrier Reef. LH only occurred in *G. horrescens*. The specimens from the

31   Chagossian RAD-seq clade could not be amplified in this marker (multiple bands of the PCR

32   product in gel electrophoresis).

1   The mitochondrial haplotypes mapped inconsistently to the RAD-seq phylogeny. The

2   Pacific clade contained mostly the longer L subtypes (LA-LP, except for LJ), but also

3   contained the shorter S type in some specimens from the Ryukyu and Daito Islands. The

4   Indo-Pacific clade contained both L and S subtypes (LA, LJ, SA, SB); the Pacific specimens

5   had haplotypes SA or SB, and most specimens from the Indian Ocean and the Red Sea were

6   associated with LA or LJ (Fig. 2, 5).

7   **3.5. Depth distribution and morphological variation between lineages in *G. fascicularis***

8   Within sampling sites, no obvious difference in depth distributions among the lineages 'S',

9   'L', and 'L+' in *G. fascicularis* were found (Fig. S7). However, in Thailand the relative

10   abundances of 'S' and 'L' lineages differed by sampling site and depth. Sites in Trat were

11   shallower (3.7 m mean depth) and had more of the S than L type, whereas sites in Chumphon

12   were deeper (5.6 m mean depth) and had more of the L type.

13   The first two dimensions of the PCA explained 84% of the total morphological

14   variation in polyp size, polyp shape, and distance between polyps (Fig. 7). The three lineages

15   'L', 'L+', and 'S' in *G. fascicularis* largely overlapped in the ordination space. However,

16   lineage 'L' may grow larger and more asymmetrical polyps (more ellipsoid than circular)

17   than the other lineages as shown by differences along the PC1 axis corresponding to polyp

18   size and differences along the PC2 axis corresponding to polyp shape (ratio between the

19   longer and shorter polyp diameter). The non-parametric Kruskal-Wallis test detected a

20   significant difference in maximal polyp diameters between genetic lineages (Chi-squared =

21   31.879, df = 2, *p*-value ≤ 0.001) but polyp shape did not have an effect. The number of septa

22   cycles or their sizes (hierarchy) septa were variable (2-3 cycles, sometimes with

23   indistinguishable hierarchy, each with 6-8 septa) but did not differ between the three lineages

24   (Table S1).

25 ## 4. Discussion

26   We investigated genetic differentiation in the genus *Galaxea* across its distribution range and

27   assessed morphological variation and depth distribution between the cryptic lineages 'L', 'S',

28   and 'L+' of the taxonomic species *G. fascicularis* in the Ryukyu Islands, Japan. We found

29   that *Galaxea* was monophyletic with respect to the outgroup specimens and clustered into

30   three distinct clades, an Indo-Pacific, a Pacific, and a small Chagossian clade sister to the two

31   other clades. The clades may be further split into seven to nine separate lineages (Table 2).

32   *Galaxea fascicularis* was spread across all clades and thus clearly paraphyletic, as was *G.*

1 *astreata. Galaxea horrescens* was the only monophyletic named species. The in *G.*

2 *fascicularis* commonly used mitochondrial marker (non-coding region between *cyt b* and

3 *ND2*) only partially matched the genomic divergence and underestimated the diversity in the

4 genus. The previously described cryptic lineages in *G. fascicularis* belonged to separate

5 clades and differed morphologically in that lineage 'L' tended to have larger polyps than the

6 other two lineages 'S' and 'L+'.

7 **4.1. Morphological diversity and taxonomic implications**

8 The most common and best known species *Galaxea fascicularis* is polyphyletic and may

9 under consideration of a phylogenetic species concept (Donoghue, 1985) not be a valid taxon

10 as it is described and applied today (van der Veer, 2007; Veron, 2000): it intermingles with

11 other nominal species throughout our phylogenetic tree (Fig. 2). There are at least four

12 (Chagossian, 'L+', 'S', and 'L') but up to eight genetic entities that have the morphology of

13 *G. fascicularis* (see discussion below, Table 2). Some of the entities ('L' *vs.* 'S' and 'L+')

14 tend to differ in polyp size and shape (Fig. 5), but no other morphological or environmental

15 characteristics analyzed here varied. At least two of the entities (lineages 'L' and 'S') also

16 seem to be reproductively isolated from each other (Abe et al. 2008a) and thus may satisfy

17 the biological species concept (Mayr, 1942). However, non-monophyletic species are

18 increasingly accepted (Carnicero et al., 2019), and a future taxonomic analysis should

19 evaluate reproductive isolation in depth, as well as consider more and other morphological

20 traits, perhaps such as those associated with the soft-tissue (Hidaka, 1992), to finally decide

21 whether the taxon *G. fascicularis* should be split into separate formal species.

22 　　　*Galaxea astreata, G. cryptoramosa,* and *G. paucisepta* were different from *G.*

23 *fascicularis* from the same location (Okinawa), but were molecularly undifferentiated from

24 each other and *G. fascicularis* from other parts of the Pacific, belonging to the same ancestral

25 lineage 'Pacific-a' (Figs. 2, 3, 4). How these specimens relate to other conspecific

26 individuals, for example from their respective type areas (all within Coral Triangle) and what

27 constitutes within versus between species differentiation for the genus remain to be

28 investigated in future studies. Particularly problematic is *G. astreata,* for which the name is

29 used inconsistently (Van der Veer, 2007; Veron & Stafford-Smith, 2000) and the original

30 description by Lamarck, 1816, is vague.

31 　　　Out of the six taxa examined, *Galaxea horrescens* was the best-defined species and

32 may be the only valid taxonomic species recognized in this genus under the phylogenetic

1   species concept (Donoghue, 1985). Specimens from multiple locations formed a

2   monophyletic clade based on the RAD data and were also distinct in their mitochondrial

3   haplotype (mostly LH). Uncertainty exists regarding its phylogenetic position: in contrast to

4   the phylogenetic tree, Admixture and DAPC analysis, the network places *G. horrescens*

5   separate from the Pacific clade and next to the outgroup species *Euphyllia* (Fig. 3). It is

6   possible that this pattern resulted from hybridization with *Euphyllia,* and future studies

7   should clarify this question including more outgroup species. In contrast to the other species

8   in *Galaxea*, *G. horrescens* occupies a different ecological niche given its branching growth

9   form and brooding reproductive mode (Fadlallah, 1983), favoring the hypothesis of being a

10 outgroup to the rest of *Galaxea*. However, regardless of its exact position in the phylogeny,

11 the validity of the species *G. horrescens* is well-supported by both life history and genetic

12 characters.

13   The morphological and taxonomic diversity in *Galaxea* was highest in the Pacific

14 clade, with most nominal species (all but *G. fascicularis* and *G. astreata*) represented only in

15 this clade (Fig. 2). The branching colony growth forms of *G. horrescens* and *G.*

16 *cryptoramosa* seem to have evolved independently based on the distant phylogenetic

17 placement of the two species on the tree (Fig. 2). However, more specimens of *G.*

18 *cryptoramosa* from other locations and representatives of the third branching species *G.*

19 *acrhelia* (although van Veeren (2007) synonymized this taxon with *G. cryptoramosa*) are

20 needed, to analyze the emergence of branching in *Galaxea*. It is possible that branching

21 morphology is not a good taxonomic character, since *G. cryptoramosa* intermingled with *G.*

22 *astreata* and *G. paucisepta* in the phylogeny (Fig. 2), similar to what was found for branching

23 and mounding morphologies in *Porites* (Forsman et al., 2017) or branching proportions of

24 species in *Oculina* (Eytan et al., 2009).

25   This first phylogenetic analysis of the genus *Galaxea* highlights a clear need for a

26 taxonomic revision. Future studies are needed to determine the taxonomic rank of the genetic

27 entities in *Galaxea* and decide whether this genus should be extended by several new species

28 or whether these entities can be absorbed into existing taxonomic names. The present study

29 was significantly limited by the absence of field collections from within the Coral Triangle,

30 where many of the accepted species' type areas are located (Veron, 2000). Museum

31 specimens were mostly not useful for phylogenetic analysis with RAD-seq here, and

32 unfortunately, fresh collections proved impossible due to the challenging legal and

33 administrative procedures to obtain samples from the involved countries needed to resolve

1 these questions. To better comprehend biological diversity, fully understand species

2 relationships, and to complete the taxonomic revision clearly needed within the genus

3 *Galaxea,* specimen sharing across political borders will be necessary.

4 **4.2. The cryptic lineages in *G. fascicularis***

5   The sympatric lineages 'L', 'S', 'L+' in the Ryukyu Islands (Hidaka, 1992; Nakajima

6 et al., 2016; Watanabe et al., 2005) belonged to separate phylogenetic clades that were

7 associated with different ocean basins (Figs. 2B, 5). Admixture analysis and DAPC further

8 distinguished multiple entities within these lineages: three ancestral lineages 'Pacific-a', '-b',

9 '-c', and two DAPC clusters 'L' and 'Ogasawara' within lineage 'L'; and clusters 'S' and

10 'Hong Kong' within lineage 'S' (Figs. 4, 5). Lineage 'S' was closely related to other lineages

11 of the morphology *G. fascicularis* in the Indian Ocean (Admixture lineage 'Indo-Pacific-a',

12 or DAPC clusters 'Indian' and 'Red Sea', Fig. 4c). DAPC generally distinguished more

13 entities than Admixture. Synthesizing all analyses and drawing the most parsimonious

14 conclusion under consideration also of the phylogeny (Fig. 2) and network (Fig. 3), *G.*

15 *fascicularis* may be split into six separate cryptic lineages: Chagossian, Pacific-L+, Pacific-L,

16 Ogasawara, Indo-Pacific-S, and Indian-Red Sea (Fig. 5, Table 2).

17   The distribution and abundance of lineages 'L' and 'S' are about equal in the Ryukyu

18 Islands, however, this varied across the Pacific. More isolated islands further to the East such

19 as Samoa and the Ogasawara Islands only harbored Pacific lineages but no Indo-Pacific S.

20 Lineages may also occurre at varying abundances between specific sampling sites (Fig. S6;

21 Nakajima et al. 2016), suggesting the influence of underlying environmental factors

22 influencing local distribution patterns. Although the three lineages could not be distinguished

23 in their depth occurrences (Fig. S6), there may be some unmeasured traits that will

24 potentially give more insight into ecological differentiations between lineages.

25   Their distribution pattern in the Pacific could apart from the biogeographic history of

26 *Galaxea* (see below), perhaps be explained by the ability of lineage 'L' to disperse farther

27 distances more easily. For example, lineage 'L' has been observed to often have a less dense

28 coenosteum than lineage 'S' (Hidaka 1992, Wewengkang et al. 2007). A softer coenosteum

29 may lead to more frequent colony fragmentation and dispersal by rafting (Thiel and Haye,

30 2006), which would allow lineage 'L' to disperse to remote places more easily. Although the

31 correlation between coenosteum density and lineage identity was not significant in a previous

32 study (Wewengkang et al., 2007), this aspect may hold more insights regarding the

1   morphological, life-history, and dispersal differentiation between the cryptic lineages, when

2   they will be reassessed in the light of the present findings regarding polyp size (Fig. 5) and

3   genetic distinction within lineage 'L' (Admixture lineages and DAPC clades, Figs. 3-5).

4   There are signs for some asymmetric gene-flow from Indo-Pacific lineage 'S' into the

5   Pacific clade, based on the many specimens of lineage 'L' containing mitochondrial

6   haplotype S, but not the other way around (Fig. 2). This mismatch of mitochondrial and

7   genomic, mostly nuclear data could indicate asymmetric introgression (Moore, 1995; van

8   Oppen et al., 2001), which would be consistent with a laboratory experiment finding higher

9   fertilization success between female 'S' and male 'L' than the other way around (Abe et al.,

10   2008). However, apart from one specimen in Thailand (PW289), we did not detect mixing of

11   the Pacific and Indo-Pacific genomes in the network and Admixture analysis (Figs. 3, 4),

12   Hybridization between coral species has often been suggested (Ladner and Palumbi, 2012)

13   (Combosch and Vollmer, 2015), however, in *Galaxea* the level of hybridization between the

14   lineages may be rather rare. An important factor that has potentially influenced the

15   maintenance of the genetic identity of lineages 'L' and 'S' in the Ryukyu Islands may be

16   shifted spawning times as observed in Okinawa (Heyward et al., 1987; Watanabe et al., 2005;

17   Yamazato, 1988). Future research using more loci and other methods, such as for example *D*-

18   statistics (Durand et al., 2011), may provide more insight into a hybridization history between

19   the lineages, but due to the 'gappy' nature of our RAD data could not be done here.

20   Lineage 'L+' was rare and could only be confirmed from a few locations in the

21   Ryukyu Islands in Japan. In the phylogenetic tree, it is located basal to all other lineages of

22   the Pacific clade (Fig. 2) and the network analysis (Fig. 3), as well as the Admixture analysis

23   (Fig. 4, K=3) indicated some shared genes with the Chagossian clade. Interestingly,

24   Admixture analysis grouped lineage L+ with *G. horrescens* (Fig. 3), and the mitochondrial

25   haplotype L+, which was otherwise private to lineage L+, was found in a museum specimen

26   of *G. horrescens* from Palau (Fig. 4). These findings suggest that lineage 'L+' may be a very

27   old lineage possibly sharing a common ancestor with the Chagossian clade, and a more recent

28   common ancestor with *G. horrescens*. The long inner branch (Fig. 2) and its marginal and

29   sparse geographic distribution favor the hypothesis of representing a relic of an ancient

30   lineage that is now perhaps being outcompeted by the other lineages in *Galaxea* in most parts

31   of the distribution range. Further experiments addressing differences in fitness and a timed

32   phylogeny are required to test this hypothesis.

4.3. Low mitochondrial diversity and a successful deletion

Haplotype LA (or a highly similar version) was present in both the Pacific as well as the Indo-Pacific clade, indicating that it may be the ancestral type of the two main extant clades in *Galaxea.* However, to find the genealogy of this mitochondrial sequence for the whole genus, future studies will need to investigate this sequence in the Chagossian clade. Across the entire distribution range, there were only a few differences in this region, apart from the characteristic 300 bp deletion in the S-haplotypes (SA, SB and SC). This mitochondrial region was much less differentiated than the rest of the genome (Figs. 2, 4) consistent with previous findings in corals (Shearer *et al.* 2002), and may by itself not be sufficient to distinguish lineages or taxonomic species in the genus *Galaxea.* However, the investigation of more complete mitochondrial data in *Galaxea* may find additional variation or a more suitable mitochondrial marker for representing the diversity in this genus (such as the ORF in *Pocillopora*, Johnston et al. 2017).

The characteristic deletion in haplotype S may have happened at once and shortly after the establishment of the Indo-Pacific clade in the Pacific since there are no intermediate lengths and all representatives of the Indo-Pacific clade in the Pacific (*G. fascicularis* lineage 'S') contain this deletion. As discussed above, haplotype S was also present in some individuals of the Pacific clade and Maldives, and some specimens from the Pacific clade and some museum specimens included both the longer and shorter haplotype. However, none of the lineage 'S'-specimens contained one of the longer L-haplotypes. One explanation for this asymmetry could be that the spread of the shorter mitochondrial sequence is under positive selection, favoring its successful establishment in the central Indo-Pacific and its introgression into other lineages across the *Galaxea* phylogeny. Cases of positive or negative selection for mtDNA types exist (Meiklejohn et al., 2007), and shorter sequences are generally faster and 'cheaper' to replicate, which by itself could be a reason for its positive selection (Selosse et al., 2001). It is possible that specimens containing both haplotypes represent a case of heteroplasy resulting from a past hybridization event, in which intracellular purifying selection has not yet fully eliminated the longer haplotypes (Birky, 2001).

4.5. Evolutionary biogeography of *Galaxea*

The geologic history and well-preserved fossil record in the Scleractinia (Keith et al., 2013) may give insight into explaining large-scale diversity patterns in *Galaxea.* The oldest fossils

of *Galaxea* are dated to 33.9–28 Ma (stem group age) and were found in Jamaica, Iran, and Florida (PBDB 2018), indicating a Tethyan origin of the genus in the Oligocene. Since the early Miocene 23–20 Ma and coinciding with the closing Tethys (Rögl, 1998), *Galaxea* has been extinct in the Atlantic and restricted to the Indo-Pacific, when also the first records of the taxon *Galaxea fascicularis* were found in Indonesia, Fiji, Iran, and Australia. A record from Hawaii from the mid Miocene indicates that the genus may have had its full contemporary range (or larger) by 11 Ma (since it is not present in Hawaii currently).

The existence of the clade in Chagos at the base of the tree (Fig. 2) supports a possible origin of *Galaxea* in the Western Indian Ocean. Although without a dated phylogeny and ancestral state analysis this is hypothetical, this ancient clade perhaps is a relic from the early, tectonically dynamic Miocene times (Rögl 1998). The Western Indian Ocean has been suggested to be important for the evolution of other corals, for example *Stylophora,* for which two of three lineages occurred in the Indian Ocean (Flot et al., 2011). In *Stylophora,* the Chagossian population is closely related to the ones in Madagascar and South-East Africa, and the population from the Red Sea is related to the one from Mid-Eastern Africa (Keshavmurthy et al., 2013). This is consistent with our results that the specimen from Tanzania clustered with specimens from the Red Sea (Figs. 2, 3, 4) and it is a common biogeographic pattern across many marine animals and plants (Costello et al., 2017). Further sampling is needed to confirm whether the Chagossian lineage occurs only in Chagos or also in other regions, too. There is also always the chance that we missed important taxa from under-sampled regions, e.g. the Coral Triangle, which would change the phylogenetic topology and position of this Chagossian clade. However, the presence of both the ancient lineage as well as the wide-spread Indo-Pacific lineage in Chagos in our study provides new insight into the genetic diversity in Chagos, which is typically grouped with either the Indo-Pacific and/or western Indian Ocean (Costello et al., 2017; Crandall et al., 2019; Kulbicki et al., 2013).

The divergence between the Indo-Pacific and Pacific clade may be a result of periods of restricted water flow between the Indian and Pacific Oceans, possibly during the late Miocene or Pliocene, as has been suggested in many marine animals (Bowen et al., 2013; Gaither et al., 2011). The South-East Asian region has been tectonically dynamic throughout the Cenozoic era (Hall and Holloway, 1998), which is regarded to be an important driver of allopatric speciation in many marine organisms (Carpenter et al., 2010). While the Australian plate was much further south during the late Oligocene and created a well-mixed Tethys sea,

1   it moved up and has restricted marine dispersal towards the late Miocene, perhaps at times

2   completely isolating the two basins in the Pliocene (Hall and Holloway 1998). The fossil

3   record of other morphologically defined taxonomic *Galaxea* species supports this

4   hypothetical divergence time between the two clades of sometime during the late Miocene, as

5   the morphological diversification in this genus, particularly colony branching, is a trait

6   associated with the Pacific clade (Fig. 2): the oldest record of *G. paucisepta* and *G. acrhelia*

7   (synonym *G. cryptoramosa*) appeared at 7 Ma in Indonesia, and *'Acrhelia horrescens'*

8   (synonym *G. horrescens*) was identified from 2.5 Ma in the Ryukyu Islands (PBDB, 2018).

9   Thus, the record of the branching *G. acrhelia* at 7 Ma suggests that the divergence between

10   the Indo-Pacific and Pacific clade could be at least 7 M years old. However, in the light of the

11   inconsistency of morphological and genetic variation in corals, the fossil record should be

12   used with caution. For example, it is possible that the branching fossil specimens belong to an

13   entirely different, now extinct clade of *Galaxea* and do not relate to the Pacific Clade of this

14   study.

15   Although we did not date our phylogenetic tree, the overlapping distributions of the

16   two genetically very distinct Pacific and Indo-Pacific clades in the western Pacific may

17   perhaps be explained by a relatively recent invasion of lineage 'S' from the Indo-Pacific

18   clade into the Pacific. Lineage 'S' may only be present in Asia and Australia because of a

19   lack of time to disperse to more isolated places like the Ogasawara Islands and American

20   Samoa in detectable quantities. Like numerous other marine organisms with an overlapping

21   distribution in the central Indo-Pacific, the pattern could be linked to sea level fluctuations in

22   the Pleistocene (Crandall et al., 2008; DeBoer et al., 2014; Jackson et al., 2014), for example,

23   higher sea levels after the last glacial maximum (~ 15 000 years) reconnected the two ocean

24   basins again through the Indonesian flow through (Hoeksema, 2007). However, there were

25   several climatic cycles since the beginning of the Pleistocene and the invasion of lineage 'S'

26   could also date back to much longer times (~2 Ma) (Bowen et al., 2013; Gaither et al., 2011).

27   In order to know the point in time lineage 'S' could have invaded the Pacific, as well as

28   divergence times between the major clades in *Galaxea,* a molecular clock analysis is needed.

29   Future studies may potentially use selected specimens from the present phylogeny and

30   analyze divergence times with a number of genes with well-known evolutionary rates.

31   As mentioned above, our study was limited by the lack any fresh collections suitable

32   for genomic analysis from the Coral Triangle, the center of the geographic distribution range

33   as well as the taxonomic diversity of *Galaxea.* Policies that have originally been created to

1 protect biodiversity and biological resources of nations from commercial exploitation are

2 now increasingly preventing research needed for conservation goals (Prathapan et al., 2018).

3 This was discussed for the Nagoya protocol (Prathapan et al., 2018) but also applies to the

4 CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora)

5 regulations. International and collaborative efforts to reveal true biodiversity patterns should

6 be facilitated rather than hindered, in order to understand and address problems affecting

7 biodiversity on a global level, such as global mass coral mortality due to climate change.

8 5. Conclusions

9 Based on a genus-wide sampling across the Indo-Pacific and using high-resolution genomic

10 markers it was possible to infer a geographically well-resolved phylogeny of the genus

11 *Galaxea.* These data also provide insights into the co-occurrence patterns and potential

12 emergence histories of morphologically and ecologically undifferentiated genetic lineages.

13 We showed that *Galaxea* is composed of three genetically highly divergent but

14 morphologically similar clades, that call into question the currently accepted taxonomy of

15 this genus. In particular, the genetic lineages of *G. fascicularis* were associated with different

16 clades making this nominal species clearly polyphyletic. In addition to the three previously

17 known lineages in *G. fascicularis* ('L', 'S', 'L+') this morphology contains three more

18 genetic entities from Ogasawara, the Indian Ocean, and Chagos (Table 2). The rooted

19 phylogeny suggests that morphological diversification in colony growth forms was associated

20 with the Pacific clade, which contained the taxonomic species *G. horrescens, G.*

21 *cryptoramosa,* and *G. paucisepta*. The *Galaxea*-characteristic mitochondrial non-coding

22 region was highly conserved across the Indo-Pacific with only a single variable nucleotide

23 (aside from the length polymorphism) that lacks the power to differentiate both

24 morphological and genomic diversity in this genus. In conclusion, we confirmed another case

25 of mismatching taxonomic and phylogenetic species identity in Scleractinia, and our results

26 indicate a more detailed taxonomic examination of *Galaxea* is clearly warranted. Whether

27 such study will be possible in the future depends on resolving the legal and administrative

28 procedures to obtain and exchange scientific samples across international boundaries. There

29 is still much to learn about the species problem in corals, and this research adds to a growing

30 number of studies that highlight the importance of a complete geographic sampling and the

31 necessity to investigate beyond nominal species boundaries for revealing true diversity

32 patterns and evolutionary history in corals. Such taxonomic and geographic sampling is made

1 exceedingly difficult, particularly in the Coral Triangle due to efforts to protect biological

2 resources from piracy.

Acknowledgements

17

References

19 Abe, M., Watanabe, T., Suzuki, Y., Hidaka, M., 2008. Genetic and morphological

20 differentiation in the hermatypic coral *Galaxea fascicularis* in Okinawa, Japan. Plankton

21 Benthos Research 3, 174-179.

22 Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G., Hohenlohe, P.A., 2016. Harnessing

23 the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet 17, 81-92.

24 Arrigoni, R., Benzoni, F., Terraneo, T.I., Caragnano, A., Berumen, M.L., 2016. Recent origin

25 and semi-permeable species boundaries in the scleractinian coral genus Stylophora from the

26 Red Sea. Sci Rep-Uk 6.

27 Ayre, D.J., Hughes, T.P., 2004. Climate change, genotypic diversity and gene flow in reef‐

28 building corals. Ecol Lett 7, 273-278.

29 Bernhard, H.M., 1902. The species problem in corals. Nature 65, 560.

1   Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K.,

2   Das, I., 2007. Cryptic species as a window on diversity and conservation. Trends Ecol Evol

3   22, 148-155.

4   Birky, C.W., 2001. The inheritance of genes in mitochondria and chloroplasts: Laws,

5   mechanisms, and models. Annu Rev Genet 35, 125-148.

6   Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina

7   sequence data. Bioinformatics 30, 2114-2120.

8   Bongaerts, P., Riginos, C., Hay, K.B., van Oppen, M.J.H., Hoegh-Guldberg, O., Dove, S.,

9   2011. Adaptive divergence in a scleractinian coral: physiological adaptation of Seriatopora

10   hystrix to shallow and deep reef habitats. Bmc Evol Biol 11.

11   Bowen, B.W., Rocha, L.A., Toonen, R.J., Karl, S.A., 2013. The origins of tropical marine

12   biodiversity. Trends Ecol Evol 28, 359-366.

13   Budd, A.F., Fukami, H., Smith, N.D., Knowlton, N., 2012. Taxonomic classification of the

14   reef coral family Mussidae (Cnidaria: Anthozoa: Scleractinia). Zool J Linn Soc-Lond 166,

15   465-529.

16   Cariou, M., Duret, L., Charlat, S., 2013. Is RAD-seq suitable for phylogenetic inference? An

17   in silico assessment and optimization. Ecol Evol 3, 846-852.

18   Carnicero, P., Schönswetter, P., Garcia-Jacas, N., Galbany-Casals, M., 2019. Is there a need

19   for accepting paraphyletic taxa? A case study in the Sardinian endemic Cymbalaria muelleri

20   (Plantaginaceae). Bot J Linn Soc 191, 325-338.

21   Carpenter, K.E., Barber, P.H., Crandall, E.D., Ablan-Lagman, M.C.A., Ambariyanto, Ngurah

22   Mahardika, G., Manjaji-Matsumoto, B.M., Juinio-Meñez, M.-A., Santos, M.D., Starger, C.J.,

23   Toha, A.H.A., 2010. Comparative phylogeography of the Coral Triangle and implications for

24   marine management. Journal of Marine Biology 2011.

25   Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015.

26   Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience

27   4.

28   Clement, M., Snell, Q., Walker, P., Posada, D., Crandall, K., 2002. TCS: Estimating gene

29   genealogies. Parallel and Distributed Processing Symposium, International Proceedings 2,

30   184.

1   Combosch, D.J., Guzman, H.M., Schuhmacher, H., Vollmer, S.V., 2008. Interspecific

2   hybridization and restricted trans-Pacific gene flow in the Tropical Eastern Pacific

3   Pocillopora. Mol Ecol 17, 1304-1312.

4   Combosch, D.J., Vollmer, S.V., 2015. Trans-Pacific RAD-Seq population genomics confirms

5   introgressive hybridization in Eastern Pacific Pocillopora corals. Mol Phylogenet Evol 88,

6   154-162.

7   Costello, M.J., Tsai, P., Wong, P.S., Cheung, A.K.L., Basher, Z., Chaudhary, C., 2017.

8   Marine biogeographic realms and species endemicity. Nat Commun 8, 1057.

9   Crandall, E., Riginos, C., Bird, C., Liggins, L., Treml, E., Beger, M., Barber, P., Connolly, S.,

10   Cowman, P., DiBattista, J., 2019. The molecular biogeography of the Indo-Pacific: testing

11   hypotheses with multispecies genetic patterns. Global Ecol Biogeogr.

12   Crandall, E.D., Jones, M.E., Munoz, M.M., Akinronbi, B., Erdmann, M.V., Barber, P.H.,

13   2008. Comparative phylogeography of two seastars and their ectosymbionts within the Coral

14   Triangle. Mol Ecol 17, 5276-5290.

15   Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker,

16   R.E., Lunter, G., Marth, G.T., Sherry, S.T., 2011. The variant call format and VCFtools.

17   Bioinformatics 27, 2156-2158.

18   Darwell, C.T., Fischer, G., Sarnat, E.M., Friedman, N.R., Liu, C., Baiao, G., Mikheyev, A.S.,

19   Economo, E.P., 2020. Genomic and phenomic analysis of island ant community assembly.

20   Mol Ecol 00, 1-17.

21   DeBoer, T.S., Naguit, M.R.A., Erdmann, M.V., Ablan-Lagman, M.C.A., Ambariyanto,

22   Carpenter, K.E., Toha, A.H.A., Barber, P.H., 2014. Concordance between phylogeographic

23   and biogeographic boundaries in the Coral Triangle: conservation implications based on

24   comparative analyses of multiple giant clam species. B Mar Sci 90, 277-300.

25   Donoghue, M.J., 1985. A Critique of the Biological Species Concept and Recommendations

26   for a Phylogenetic Alternative. Bryologist 88, 172-181.

27   Durand, E.Y., Patterson, D., Reich, D., Slatkin, M., 2011. Testing for ancient admixture

28   between closely related populations. Mol Biol Evol 28, 2239–2252.

Eaton, D.A.R., Spriggs, E.L., Park, B., Donoghue, M.J., 2017. Misconceptions on Missing Data in RAD-seq Phylogenetics with a Deep-scale Example from Flowering Plants. Syst Biol 66, 399-412.

Emerson, K.J., Merz, C.R., Catchen, J.M., Hohenlohe, P.A., Cresko, W.A., Bradshaw, W.E., Holzapfel, C.M., 2010. Resolving postglacial phylogeography using high-throughput sequencing. P Natl Acad Sci USA 107, 16196-16200.

Eytan, R.I., Hayes, M., ARBOUR‐REILY, P., Miller, M., Hellberg, M.E., 2009. Nuclear sequences reveal mid‐range isolation of an imperilled deep‐water coral population. Mol Ecol 18, 2375-2389.

Fadlallah, Y.H., 1983. Sexual reproduciton, development and larval biology in Scleractinian Corals. Coral Reefs 2, 129-150.

Ferrier-Pagès, C., Rottier, C., Beraud, E., Levy, O., 2010. Experimental assessment of the feeding effort of three scleractinian coral species during a thermal stress: Effect on the rates of photosynthesis. J Exp Mar Biol Ecol 390, 118-124.

Fischer, G., Azorsa, F., Garcia, F.H., Mikheyev, A.S., Economo, E.P., 2015. Two new phragmotic ant species from Africa: morphology and next-generation sequencing solve a caste association problem in the genus Carebara Westwood. Zookeys, 77-105.

Flot, J.-F., Blanchot, J., Charpy, L., Cruaud, C., Licuanan, W.Y., Nakano, Y., Payri, C., Tillier, S., 2011. Incongruence between morphotypes and genetically delimited species in the coral genus *Stylophora*: phenotypic plasticity, morphological convergence, morphological stasis or interspecific hybridization? Bmc Ecol 11, 22.

Forsman, Z.H., Barshis, D.J., Hunter, C.L., Toonen, R.J., 2009. Shape-shifting corals: Molecular markers show morphology is evolutionarily plastic in Porites. Bmc Evol Biol 9.

Forsman, Z.H., Concepcion, G.T., Haverkort, R.D., Shaw, R.W., Maragos, J.E., Toonen, R.J., 2010. Ecomorph or Endangered Coral? DNA and Microstructure Reveal Hawaiian Species Complexes: Montipora dilatata/flabellata/turgescens & M. patula/verrilli. Plos One 5.

Forsman, Z.H., Knapp, I.S.S., Tisthammer, K., Eaton, D.A.R., Belcaid, M., Toonen, R.J., 2017. Coral hybridization or phenotypic variation? Genomic data reveal gene flow between Porites lobata and P-Cornpressa. Mol Phylogenet Evol 111, 132-148.

1 Fukami, H., Chen, C. A., Budd, A. F., Collins, A., Wallace, C., Chuang, Y. Y., . . . Knowlton,
2 N. (2008). Mitochondrial and Nuclear Genes Suggest that Stony Corals Are Monophyletic
3 but Most Families of Stony Corals Are Not (Order Scleractinia, Class Anthozoa, Phylum
4 Cnidaria). *Plos One, 3*(9). doi:ARTN e3222

5 Gaither, M.R., Bowen, B.W., Bordenave, T.-R., Rocha, L.A., Newman, S.J., Gomez, J.A.,
6 van Herwerden, L., Craig, M.T., 2011. Phylogeography of the reef fish *Cephalopholis argus*
7 (Epinephelidae) indicates Pleistocene isolation across the Indo-Pacific Barrier with
8 contemporary overlap in the Coral Triangle. Bmc Evol Biol 11, 189.

9 Hall, R., Holloway, J.D., 1998. Biogeography and geological evolution of SE Asia.
10 Backhuys.

11 Heyward, A., Yamazato, K., Yeemin, T., Minei, M., 1987. Sexual reproduction of corals in
12 Okinawa. Galaxea 6, 331-343.

13 Hidaka, M., 1992. Use of nematocyst morphology for taxonomy of some related species of
14 scleractinian corals. Galaxea 11, 21-28.

15 Hoeksema, B.W., 2007. Delineation of the Indo-Malayan centre of maximum marine
16 biodiversity: the Coral Triangle. In: Renema, W. (Ed.), Biogoegraphy, time, and place:
17 distributions, barriers, and islands. Springer, Dordrecht, The Netherlands, pp. 117-178.

18 Huang, D.W., 2012. Threatened Reef Corals of the World. Plos One 7, e34459.

19 Huson, D.H., 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics
20 (Oxford, England) 14, 68-73.

21 Jackson, A.M., Ambariyanto, Erdmann, M.V., Toha, A.H.A., Stevens, L.A., Barber, P.H.,
22 2014. Phylogeography of commercial tuna and mackerel in the Indonesian Archipelago. B
23 Mar Sci 90, 471-492.

24 Johnston, E.C., Forsman, Z.H., Flot, J.F., Schmidt-Roach, S., Pinzon, J.H., Knapp, I.S.S.,
25 Toonen, R.J., 2017. A genomic glance through the fog of plasticity and diversification in
26 Pocillopora. Sci Rep-Uk 7.

27 Jombart, T., Ahmed, I., 2011. adegenet 1.3-1: new tools for the analysis of genome-wide
28 SNP data Bioinformatics.

1 Keith, S.A., Baird, A.H., Hughes, T.P., Madin, J.S., Connolly, S.R., 2013. Faunal breaks and
2 species composition of Indo-Pacific corals: the role of plate tectonics, environment and
3 habitat distribution. P Roy Soc B-Biol Sci 280, 20130818.

4 Keshavmurthy, S., Yang, S.Y., Alamaru, A., Chuang, Y.Y., Pichon, M., Obura, D., Fontana,
5 S., De Palmas, S., Stefani, F., Benzoni, F., MacDonald, A., Noreen, A.M.E., Chen, C.S.,
6 Wallace, C.C., Pillay, R.M., Denis, V., Amri, A.Y., Reimer, J.D., Mezaki, T., Sheppard, C.,
7 Loya, Y., Abelson, A., Mohammed, M.S., Baker, A.C., Mostafavi, P.G., Suharsono, B.A.,
8 Chen, C.A., 2013. DNA barcoding reveals the coral "laboratory-rat", *Stylophora pistillata*
9 encompasses multiple identities. Sci Rep-Uk 3, 1520.

10 Kitahara, M. V., Cairns, S. D., Stolarski, J., Blair, D., & Miller, D. J. (2010). A
11 comprehensive phylogenetic analysis of the Scleractinia (Cnidaria, Anthozoa) based on
12 mitochondrial CO1 sequence data. *Plos One, 5*(7), e11490.

13 Kitahara, M.V., Fukami, H., Benzoni, F., Huang, D., 2016. The New Systematics of
14 Scleractinia: Integrating Molecular and Morphological Evidence. In: Goffredo, S., Dubinsky,
15 Z. (Eds.), The Cnidaria, Past, Present and Future: The world of Medusa and her sisters.
16 Springer International Publishing, Cham, pp. 41-59.

17 Kozlov, A.M., Aberer, A.J., Stamatakis, A., 2015. ExaML version 3: a tool for phylogenomic
18 analyses on supercomputers. Bioinformatics 31, 2577-2579.

19 Kulbicki, M., Parravicini, V., Bellwood, D.R., Arias-Gonzàlez, E., Chabanet, P., Floeter,
20 S.R., Friedlander, A., McPherson, J., Myers, R.E., Vigliola, L., 2013. Global biogeography of
21 reef fishes: a hierarchical quantitative delineation of regions. Plos One 8, e81847.

22 Ladner, J.T., Barshis, D.J., Palumbi, S.R., 2012. Protein evolution in two co-occurring types
23 of Symbiodinium: an exploration into the genetic basis of thermal tolerance in Symbiodinium
24 clade D. Bmc Evol Biol 12, 217.

25 Ladner, J.T., Palumbi, S.R., 2012. Extensive sympatry, cryptic diversity and introgression
26 throughout the geographic distribution of two coral species complexes. Mol Ecol 21, 2224-
27 2238.

28 Leigh, J., Bryant, D., 2015. PopART: Full-feature software for haplotype network
29 construction. Methods Ecol Evol 6, 1110-1116.

1 Lemoine, F., Domelevo Entfellner, J.B., Wilkinson, E., Correia, D., Dávila Felipe, M., De
2 Oliveira, T., Gascuel, O., 2018. Renewing Felsenstein's phylogenetic bootstrap in the era of
3 big data. Nature 556, 452-456.

4 Liew, Y.J., Aranda, M., Voolstra, C.R., 2016. Reefgenomics.Org - a repository for marine
5 genomics data. Database 2016, baw152.

6 Mao, Y., Shinzato, C., Satoh, N., 2018. An ancient genome duplication in the speciose reef-
7 building coral genus, *Acropora*. bioRxiv, 366435.

8 Marshall, P.A., Baird, A.H., 2000. Bleaching of corals on the Great Barrier Reef: differential
9 susceptibilities among taxa. Coral Reefs 19, 155-163.

10 Marti-Puig, P., Forsman, Z.H., Haverkort-Yeh, R.D., Knapp, I.S.S., Maragos, J.E., Toonen,
11 R.J., 2014. Extreme phenotypic polymorphism in the coral genus Pocillopora; micro-
12 morphology corresponds to mitochondrial groups, while colony morphology does not. B Mar
13 Sci 90, 211-231.

14 Mayr, E., 1942. Systematics and the origin of species from the viewpoint of a zoologist.
15 Columbia University Press, New York.

16 Meiklejohn, C.D., Montooth, K.L., Rand, D.M., 2007. Positive and negative selection on the
17 mitochondrial genome. Trends Genet 23, 259-263.

18 Moore, W.S., 1995. Inferring Phylogenies from Mtdna Variation - Mitochondrial-Gene Trees
19 Versus Nuclear-Gene Trees. Evolution 49, 718-726.

20 Nakaema, S., Hidaka, M., 2015. GFP distribution and fluorescence intensity in *Galaxea*
21 *fascicularis*: developmental changes and maternal effects Platax 12, 1-9.

22 Nakajima, Y., Shinzato, C., Satoh, N., Mitarai, S., 2015. Novel polymorphic microsatellite
23 markers reveal genetic differentiation between two sympatric types of *Galaxea fascicularis*.
24 Plos One 10, e0130176.

25 Nakajima, Y., Zayasu, Y., Shinzato, C., Satoh, N., Mitarai, S., 2016. Genetic differentiation
26 and connectivity of morphological types of the broadcast-spawning coral *Galaxea*
27 *fascicularis* in the Nansei Islands, Japan. Ecol Evol 6, 1457-1469.

28 Paz-Garcia, D.A., Hellberg, M.E., Garcia-de-Leon, F.J., Balart, E.F., 2015. Switch between
29 Morphospecies of Pocillopora Corals. Am Nat 186, 434-440.

1   PBDB, 2018. The data were downloaded from the Paleobiology Database, using the taxon
2   name 'Galaxea'

3   Pinzón, J.H., Kamel, B., Burge, C.A., Harvell, C.D., Medina, M., Weil, E., Mydlarz, L.D.,
4   Burge, C.A., Harvell, C.D., Kamel, B., Medina, M., Mydlarz, L.D., 2015. Data from: Whole
5   transcriptome analysis reveals changes in expression of immune related genes during and
6   after bleaching in a reef-building coral. In: Dryad (Ed.).

7   Pinzon, J. H., Sampayo, E., Cox, E., Chauka, L. J., Chen, C. A., Voolstra, C. R., &
8   LaJeunesse, T. C. (2013). Blind to morphology: genetics identifies several widespread
9   ecologically common species and few endemics among Indo-Pacific cauliflower corals
10 (Pocillopora, Scleractinia). Journal of Biogeography, 40(8), 1595-1608.

11 Prada, C., Hellberg, M.E., 2013. Long prereproductive selection and divergence by depth in a
12 Caribbean candelabrum coral. P Natl Acad Sci USA 110, 3961-3966.

13 Prathapan, K.D., Pethiyagoda, R., Bawa, K.S., Raven, P.H., Rajan, P.D., Countries, C.-S.,
14 2018. When the cure kills-CBD limits biodiversity research National laws fearing biopiracy
15 squelch taxonomy studies. Science 360, 1405-1406.

16 R Core Team, 2015. R: A language and environment for statistical computing. R Foundation
17 for Statistical Computing, Vienna.

18 Radice, V., Hoegh-Guldberg, O., Fry, B., Fox, M., Dove, S., 2019. Upwelling as the major
19 source of nitrogen for shallow and deep reef-building corals across an oceanic atoll system.
20 Funct Ecol In press. DOI: 10.1111/1365-2435.13314.

21 Rögl, F., 1998. Palaeogeographic Considerations for Mediterranean and Paratethys Seaways
22 (Oligocene to Miocene). Annalen des Naturhistorischen Museums in Wien. Serie A für
23 Mineralogie und Petrographie, Geologie und Palāontologie, Anthropologie und Prāhistorie
24 99, 279-310.

25 Selosse, M.A., Albert, B.R., Godelle, B., 2001. Reducing the genome size of organelles
26 favours gene transfer to the nucleus. Trends Ecol Evol 16, 135-141.

27 Serrano, X., Baums, I.B., O'Reilly, K., Smith, T.B., Jones, R.J., Shearer, T.L., Nunes, F.L.D.,
28 Baker, A.C., 2014. Geographic differences in vertical connectivity in the Caribbean coral
29 Montastraea cavernosa despite high levels of horizontal connectivity at shallow depths. Mol
30 Ecol 23, 4226-4240.

1 Shoguchi, E., Shinzato, C., Kawashima, T., Gyoja, F., Mungpakdee, S., Koyanagi, R.,

2 Takeuchi, T., Hisata, K., Tanaka, M., Fujiwara, M., Hamada, M., Seidi, A., Fujie, M., Usami,

3 T., Goto, H., Yamasaki, S., Arakaki, N., Suzuki, Y., Sugano, S., Toyoda, A., Kuroki, Y.,

4 Fujiyama, A., Medina, M., Coffroth, M.A., Bhattacharya, D., Satoh, N., 2013. Draft

5 Assembly of the Symbiodinium minutum Nuclear Genome Reveals Dinoflagellate Gene

6 Structure. Curr Biol 23, 1399-1408.

7 Stamatakis, A., 2006. RAxML-VI-HPC: Maximum Likelihood-based phylogenetic analyses

8 with thousands of taxa and mixed models. Bioinformatics 22, 2688-2690.

9 Thiel, M., Haye, P.A., 2006. The ecology of rafting in the marine environment. III.

10 Biogeographical and evolutionary consequences. Oceanogr Mar Biol 44, 323-429.

11 Tin, M.M.Y., Economo, E.P., Mikheyev, A.S., 2014. Sequencing degraded DNA from non-

12 destructively sampled museum specimens for RAD-tagging and low-coverage shotgun

13 phylogenetics. Plos One 9, e96793.

14 Torres, A. F., & Ravago-Gotanco, R. (2018). Rarity of the "common" coral *Pocillopora*

15 damicornis in the western Philippine archipelago. *Coral Reefs, 37*(4), 1209-1216.

16 van der Veer, E., 2007. Towards a revision of the coral genus *Galaxea* (Scleractinia:

17 Oculinidae). Naturalis. University of Leiden, Leiden.

18 van Oppen, M.J.H., McDonald, B.J., Willis, B., Miller, D.J., 2001. The evolutionary history

19 of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear

20 marker: Reticulation, incomplete lineage sorting, or morphological convergence? Mol Biol

21 Evol 18, 1315-1329.

22 Veron, J.E.N., 1995. Corals in space and time : the biogeography and evolution of the

23 Scleractinia. Comstock/Cornell, Ithaca.

24 Veron, J.E.N., 2000. Corals of the world (No. C/593.6 V4). Australian Institute of Marine

25 Science, Townsville, Australia.

26 Voolstra, C., Miller, D., Ragan, M., Hoffmann, A., Hoegh-Guldberg, O., Bourne, D., Ball, E.,

27 Ying, H., Foret, S., Takahashi, S., Weynberg, K., van Oppen, M., Morrow, K., Chan, C.X.,

28 Rosic, N., Leggat, W., Sprungala, S., Imelfort, M., Tyson, G., Kassahn, K., Lundgren, P.,

29 Beeden, R., Ravasi, T., Berumen, M., Abel, E., Fyffe, T., 2015. The ReFuGe 2020

30 Consortium—using "omics" approaches to explore the adaptability and resilience of coral

31 holobionts to environmental change. Frontiers in Marine Science 2, 68.

Watanabe, T., Nishida, M., Watanabe, K., Wewengkang, D.S., Hidaka, M., 2005. Polymorphism in nucleotide sequence of mitochondrial intergenic region in scleractinian coral (*Galaxea fascicularis*). Mar Biotechnol 7, 33-39.

Wepfer, P.H., 2018. Spatial genetic structure in the coral genus *Galaxea* (Euphyllidae) and their associated Symbiodiniaceae communities. Okinawa Institute of Science and Technology, Onna, Okinawa, Japan.

Wewengkang, D.S., Watanabe, T., Hidaka, M., 2007. Studies on morphotypes of the coral *Galaxea fascicularis* from Okinawa: polyp color, nematocyst shape, and coenosteum density. Journal of the Japanese Coral Reef Society 9, 49-59.

Willis, B.L., van Oppen, M.J.H., Miller, D.J., Vollmer, S.V., Ayre, D.J., 2006. The role of hybridization in the evolution of reef corals. Annu Rev Ecol Evol S 37, 489-517.

Wood, D., Salzberg, S., 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biology 15.

WoRMS, E.B., 2019. World Register of Marine Species.

Yamazato, K., 1988. Reproductive mechanism of *Galaxea fascicularis*. Report for Grant-in-Aid on "Early phase of reproduction of marine resource organisms" from the Ministry of Education, Science and Culture, Japan, 148-155.

Yasuda, N., Taquet, C., Nagai, S., Fortes, M., Fan, T.-Y., Harii, S., Yoshida, T., Sito, Y., Nadaoka, K., 2015. Genetic diversity, paraphyly and incomplete lineage sorting of mtDNA, ITS2 and microsatellite flanking region in closely related *Heliopora species* (Octocorallia). Mol Phylogenet Evol 93, 161-171.

Yasuda, N., Taquet, C., Nagai, S., Fortes, M., Fan, T.-Y., Phongsuwan, N., Nadaoka, K., 2014. Genetic structure and cryptic speciation in the threatened reef-building coral *Heliopora coerulea* along Kuroshio Current. B Mar Sci 90, 233-255.

Tables and Figures

Table 1. Currently accepted species in *Galaxea* by the World Register of Marine Species (WoRMS, 2019) and distributions as given by Veron (2000). The age refers to the oldest fossil record listed the Paleo Biology Database (PBDB, 2018) where available. Abbreviations: na = not available, SE= South East

| Species	First description	Distribution	Abundance	Age [My]
G. fascicularis	Linnaeus, 1767. Syst. Nat. 12th ed. v. 1 (pt 2): 1278	Red Sea, Indo-Pacific	common	20.43
G. astreata	Lamarck, J.B.d.1816. Animaux sans Vertèbres: 227	Red Sea, Indo-Pacific	common	11.6
G. pauciradiata (synonym G. astreata*)*	Blainville, H. M. de 1830. Zoophytes. In: Dict. Sci. Nat.: 548, pls. 68	Red Sea	uncommon	5.5
G. horrescens	Dana, J.D. 1846. United States Exploring Expedition during the years 1838-1842. Zoophytes 7: 1-740.	Central Indo-Pacific	uncommon	2.5
G. alta	Nemenzo, F., 1979. The Philippine Journal of Science 108: 1-25.	SE Asia	na	na
G. negrensis	Nemenzo 1979. Hoeksema, B.; Cairns, S. (2018). World List of Scleractinia.	SE Asia	na	na
G. paucisepta	Claereboudt, M. 1990. Galaxea 9: 1-8.	SE Asia	rare	7.246
G. cryptoramosa	Veron, J.E.N. & Stafford-Smith, M. 2000. Corals of the World: 114	SE Asia	uncommon	na
G. longisepta	Veron, J.E.N. & Stafford-Smith, M. 2000. Corals of the World: 116	SE Asia	rare	na
G. acrhelia (synonym G. cryptoramosa*)*	Veron, J.E.N. & Stafford-Smith, M. 2000. Corals of the World: 115	SE Asia	uncommon	7.246

1   **Table 2.** Summary and synthesis of evolutionary entities in *Galaxea.* Based on genomic RAD-seq

2   data, seven lineages may be distinguished as a parsimonious conclusion from Admixture (Fig. 4) and

3   DAPC (Fig. 5), while accounting for distinguished groupings in Neighbor-Net (Fig. 3) and monophyly

4   in the phylogenetic tree (Fig. 2).

| Synthesis	Admixture	DAPC	Neighbor-Net group	Monophyly	Nominal species	Mt-hapl.
Lineage L	Pacific-a	Cluster L	Pacific	paraphyletic	*G. fasc., G. astr., G. pauc., G. cryptor.*	LA, LM, LN LL, LK, LI, LG
	Pacific-b					
Ogasawaran	Pacific-c	Cluster Ogasawara	(Ogasawara)	yes	*G. fasc.*	LA
Lineage L+	Pacific-d	Cluster L+	L+	yes	*G. fasc., G. astr*	L+
G. horresc.		Cluster Horresc.	G. horresc.	yes	*G. horresc.*	LH
Indian-Red Sea	Indo-Pacific a	Cluster Indian	Indo-Pacific (Red Sea)	paraphyletic	*G. fasc.*	LA, LJ, SA
		Cluster Red Sea				
Lineage S	Indo-Pacific-b	Cluster S	(S)	yes	*G. fasc., G. astr.*	SA, SB
		Cluster HK				
Chagossian	Chagossian	Cluster Chagos	Chagossian	yes	*G. fasc.*	n.a.

5

38
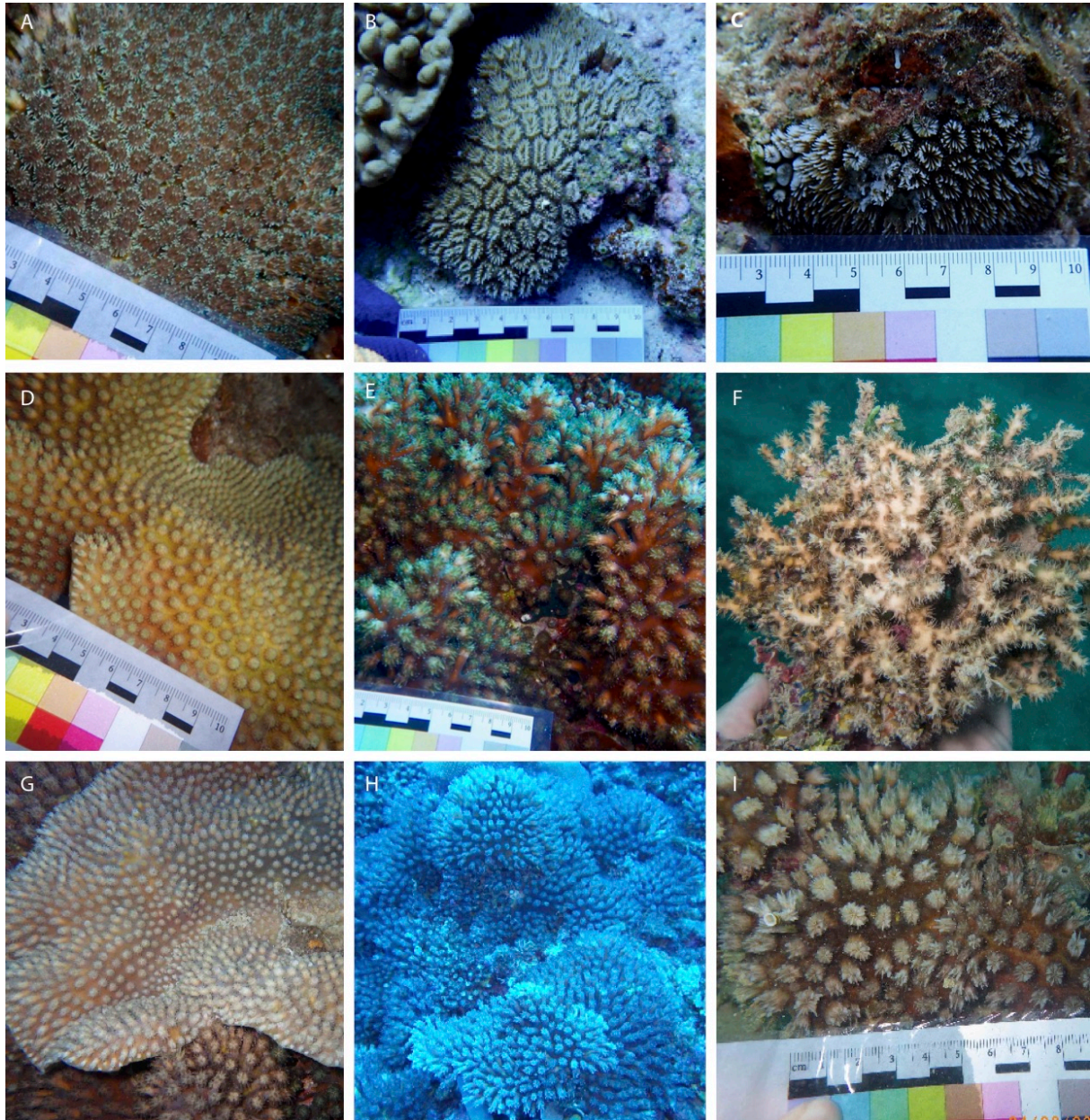
Figure 1. Example photographs of *Galaxea* specimens of different taxonomic species. A: *G. fascicularis,* lineage "S", PW575 from Seragaki; B: *G. fascicularis,* lineage "L", PW100 from Iheya, C: *G. fascicularis,* lineage "L+", PW42 from Miyako; D: *G. paucisepta,* PW571 from Seragaki; E: *G. cryptoramosa,* PW249 from Seragaki; F: *G. horrescens,* AF-3 from Guam; G: *G paucisepta* (overgrowing *G. astreata*)*,* PW573 from Seragaki; H: *G. astreata,* PW572 from Seragaki; I: *G. astreata,* PW448 from Motobu.
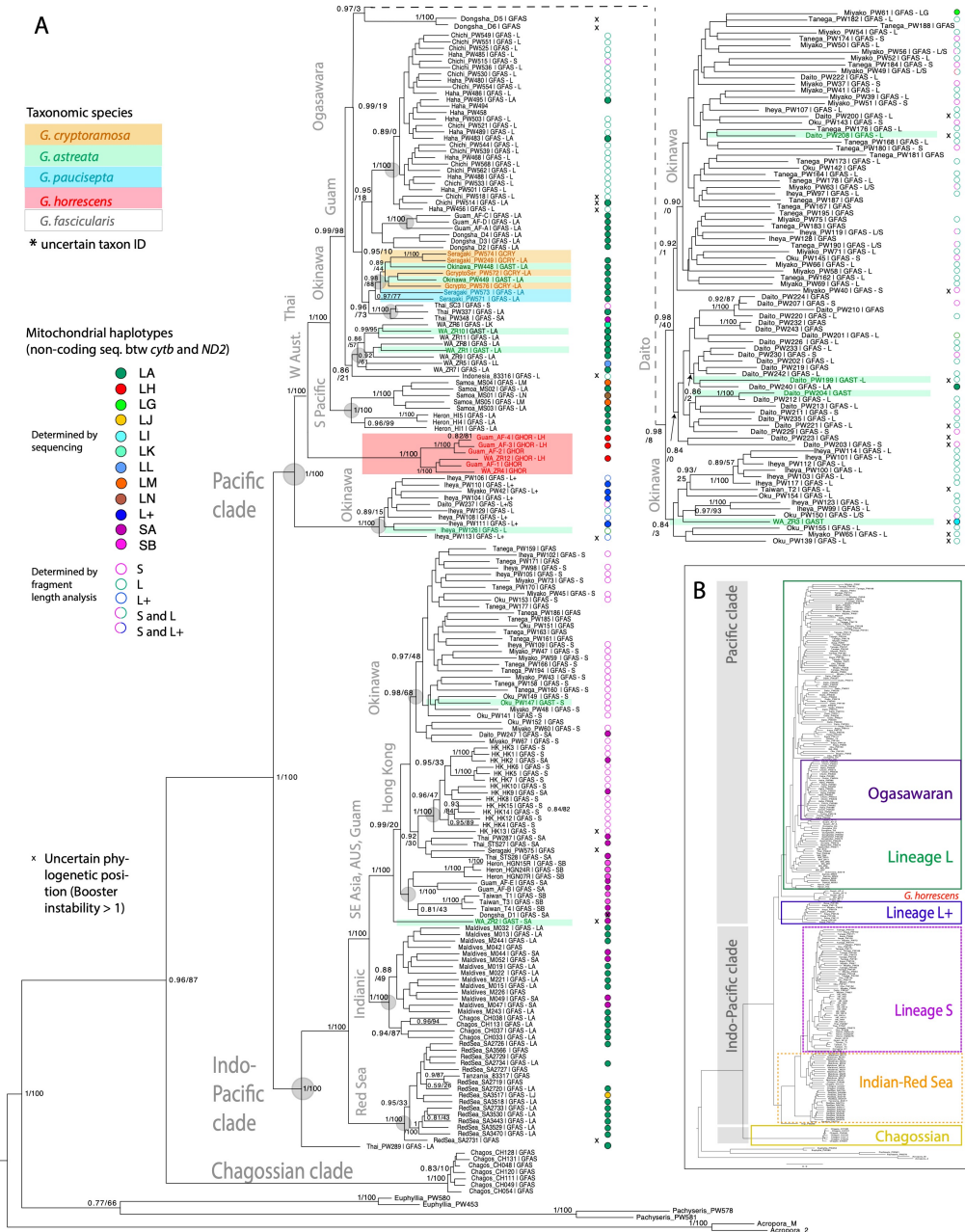
Figure 2. RAD-seq phylogeny of *Galaxea*. A: Tips are labeled according to their geographic origin, sample number, taxonomic species and mitochondrial haplotype. In addition, the taxonomic species are illustrated by tip color, and the mitochondrial haplotypes by circles next to the labels retrieved either by Sanger sequencing (filled circles) or fragment length analysis (empty circles). Crosses (x) mark specimens with booster instability >1 (Table S3). Node supports are given as Booster distances > 0.8 and direct bootstrap frequencies based on 468 bootstrap replicates. More details to sampling origin may be retrieved from Tables S1 and S2. B: The insertion shows the overview topology of the whole tree and the lineages in *Galaxea* synthesized in this study (Table 2).

1
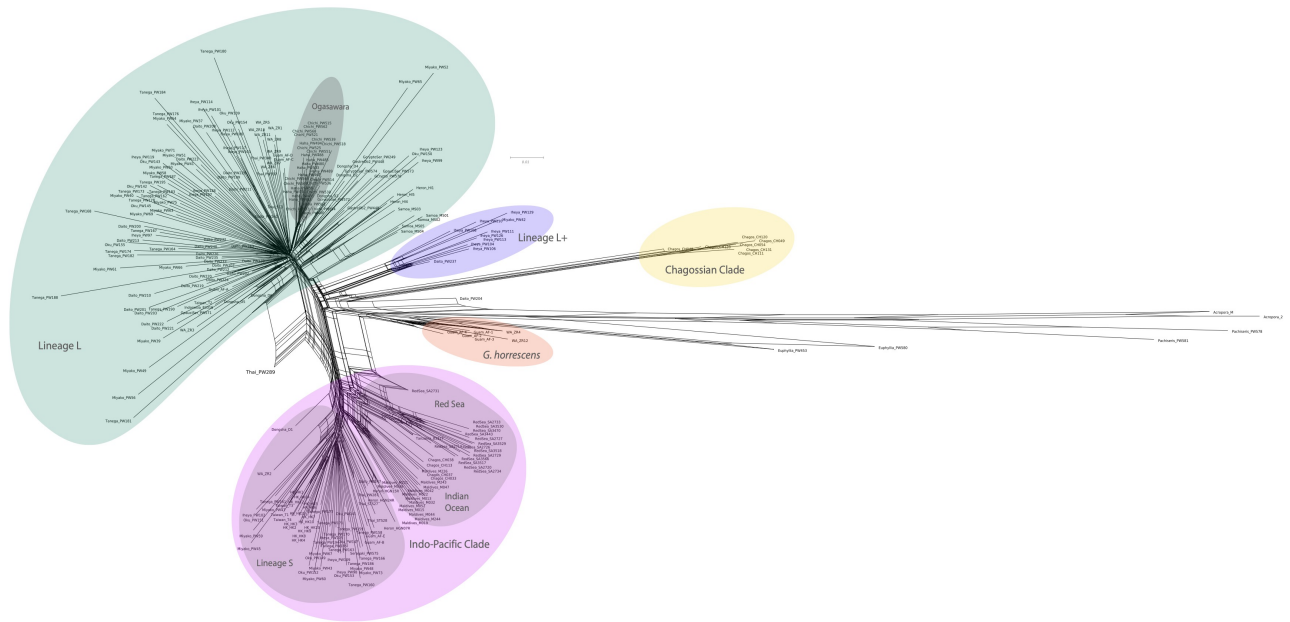
2

3 **Figure 3.** Neighbor-Net network implemented in SplitsTree. Specimens are grouped into five major

4 branches. In addition, the Ogasawaran specimens may be distinguished from the large Pacific group

5 (Lineage L). The distinction between Lineage S from the rest of the Indo-Pacific clade is less clear,

6 indicating incomplete lineage sorting or mixing. Hybridization between the main clades is only evident

7 in one specimen (Thai _PW289). In contrast to the phylogeny (Fig. 2), *G. horrescens* is placed closer

8 to the outgroup. The other nominal species (*G. astreata, G. paucisepta, G. cryptoramosa*) belonged to
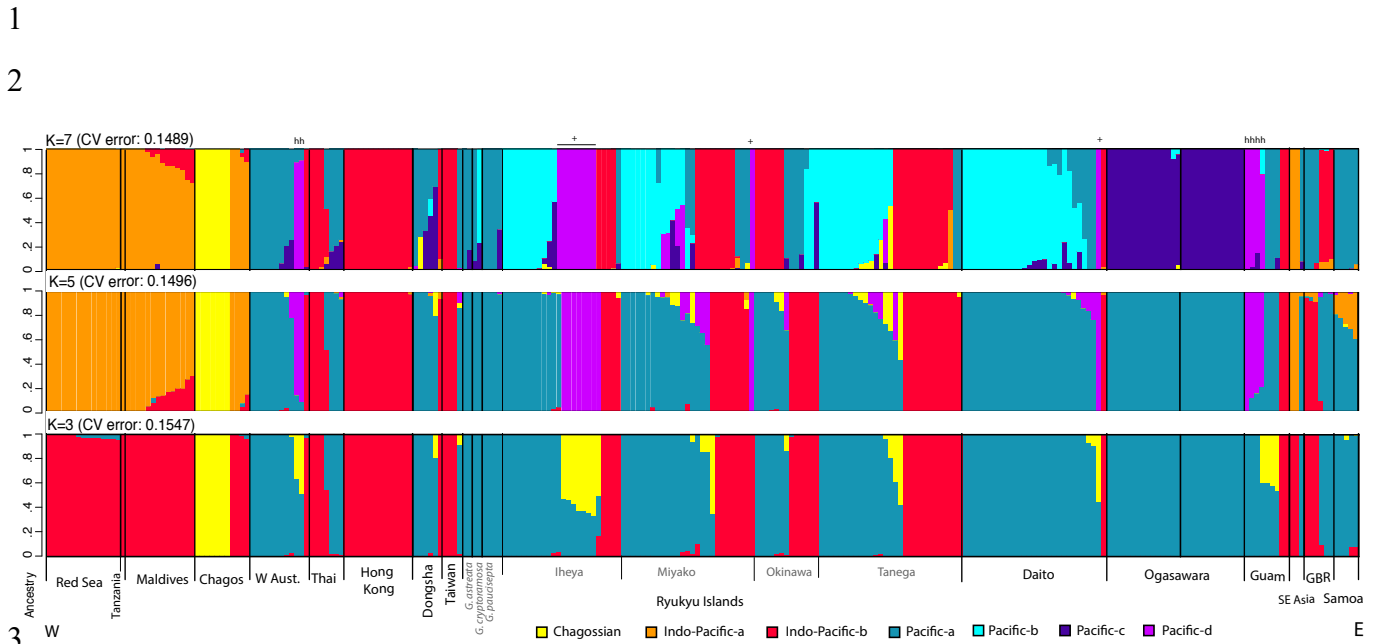
9 lineage L.

10

1

2

3

4 **Figure 4.** Admixture ancestral assignments for *Galaxea* specimens (in alphabetical order) based on

5 2275 genetic sites present in at least 50% of all individuals. K=7 was the most likely number of

6 ancestral lineages, but results for K= 3 and 5 (Fig. S3) are also shown to illuminate the relationships

7 between lineages. K=3 distinguishes the Pacific (blue, 'Pacific-a'), Indo-Pacific (red, 'Indo-Pacific-b'),

8 and Chagossian clade (yellow) identified in the phylogenetic analysis; from those, K=5 distinguishes

9 the lineages L+ (marked with '+') and *G. horrescens* (marked with 'h') from the Pacific clade (pink,

10 'Pacific-d') and a second Indian Ocean with Red Sea cluster from the 'Indo-Pacific' clade (orange,

11 'Indo-Pacific-a'); and K=7 splits an Ogasawaran lineage (dark blue, 'Pacific-c') and a lineage

12 represented in SW Japan (Ryukyu Islands and Daito, turquoise, 'Pacific-b'). Highest diversity of

13 ancestral lineages are in the Pacific, particularly in the Ryukyu Islands. Labels for each sample may
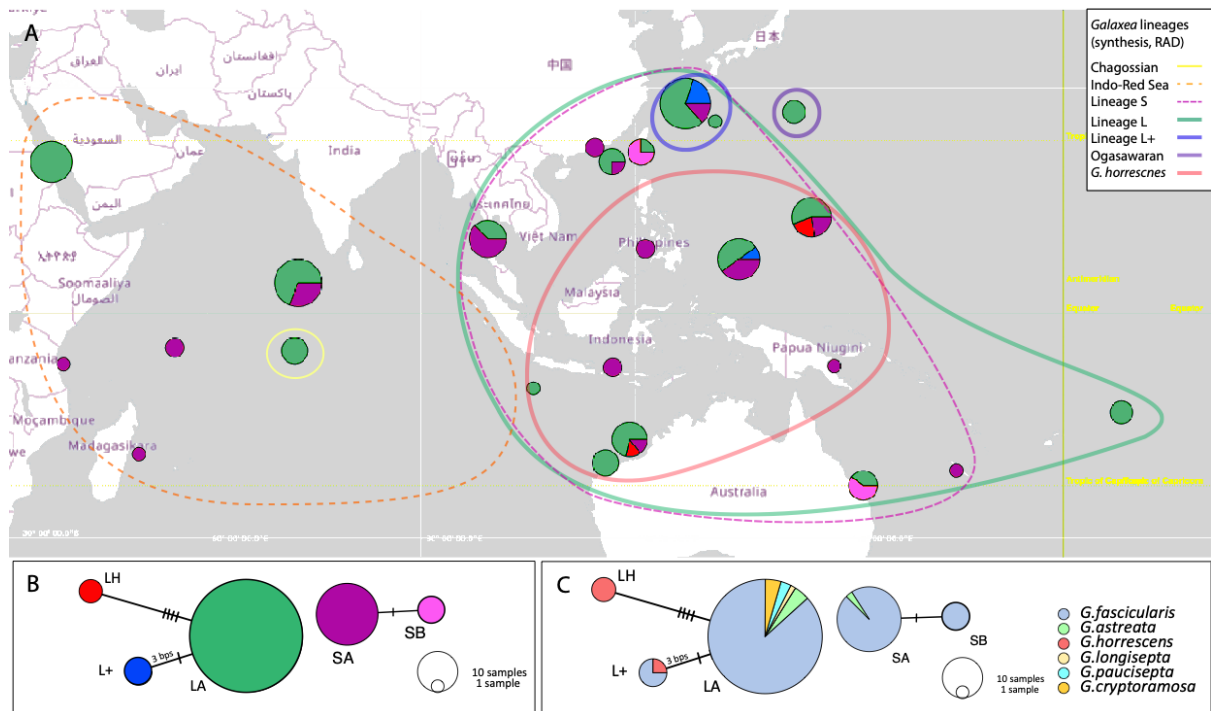
14 be retrieved from Fig. S4.

15

42
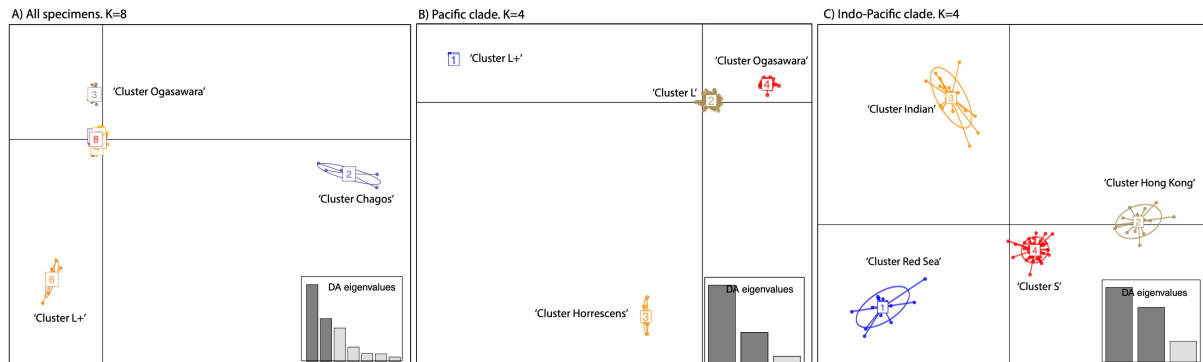
Figure 5. DAPC scatter plots for A) all *Galaxea* specimens, B) Pacific clade, C) Indo-Pacific clade. In summary, nine distinct DAPC-clusters are distinguished in *Galaxea*: one Chagossian, four Pacific, and four Indo-Pacific clusters.

1

2

3 **Figure 6.** A: Geographic distribution of RAD-lineages in *Galaxea* as synthesized in Table 2 and
4 mitochondrial haplotypes. The mitochondrial haplotypes refer to the non-coding region between the
5 genes *cytb* and *ND2.* B: Mitochondrial haplotype network of haplotypes that differed by more than one
6 substitution or were found in more than one specimen. C: Mitochondrial haplotype network by
7 taxonomic species. Most taxonomic species contain haplotype LA, except for *G. horrescens,* which
8 contained LH, and L+, and *G. astreata*, in which one specimen contained SA. LA is the most common
9 and most widely distributed mt-haplotype. The two networks differ by a deletion of 300 bp in SA and
10 SB.

11
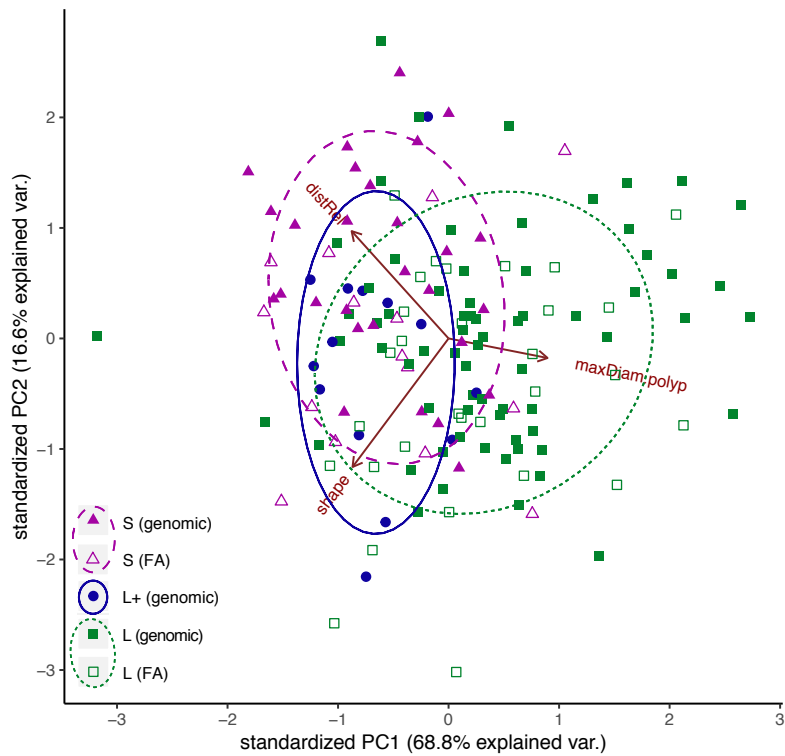
12

Figure 7. Morphology principal component analysis on cryptic lineages in *G. fascicularis* from Japan.
Lineage 'L' grows somewhat larger polyps than the other lineages. Specimens for which only
mitochondrial data from fragment length analysis (FA) was available are colored in a lighter shade.
Abbreviations: diam.max = maximal polyp diameter, dist.rel = space between polyps, shape = ratio of
shorter to longer polyp diameter.

Author statement

Patricia H. Wepfer: conceptualization, formal analysis, investigation, resources, writing-original, visualization, project administration, funding

Yuichi Nakajima: resources, funding, supervision

Makamas Sutthacheep: resources

Veronica Z. Radice: resources

Zoe Richards: resources

Put Ang: resources

Tullia Terraneo: resources

Mareike Sudek: resources

Atsushi Fujimura: resources

Robert J. Toonen: resources, writing-review, supervision

Alexander S. Mikheyev: methodology, funding, resources

Evan P. Economo: conceptualization, writing-review, resources, supervision

Satoshi Mitarai: resources, funding, supervision