TITLE:

A universal subcuticular bacterial symbiont of a coral predator, the crown-of-thorns starfish

RUNNING TITLE: Subcuticular bacterium of a coral predator

AUTHORS:

Naohisa WADA[1,2†], Hideaki YUASA[3†], Rei KAJITANI[3], Yasuhiro GOTOH[4], Yoshitoshi OGURA[4], Dai YOSHIMURA[3], Atsushi TOYODA[5], Sen-Lin TANG[2], Yukio HIGASHIMURA[1], Hugh SWEATMAN[6], Zac FORSMAN[7], Omri BRONSTEIN[8], Gal EYAL[9,10], Nalinee THONGTHAM[11], Takehiko ITOH[3*], Tetsuya HAYASHI[4*], Nina YASUDA[1*]

AFFILIATIONS:

1: Faculty of Agriculture, University of Miyazaki, 1-1 Gakuenkibanadai-Nishi, Miyazaki, Miyazaki 889-2192, Japan

2: Biodiversity Research Center, Academia Sinica, No.128, Sec 2, Academia Rd, Nangang, Taipei 11529 Taiwan

3: School of Life Science and Technology, Department of Life Science and Technology, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8550, Japan

4: Department of Bacteriology, Faculty of Medical Sciences, Kyushu University, 3-1-1 Maidashi Higashi-ku Fukuoka 812-8582, Japan

5: Center for Information Biology, National Institute of Genetics, Yata 1111, Mishima, Shizuoka 411-8540, Japan

6: Australian Institute of Marine Science, PMB No.3, Townsville, QLD 4810, Australia

7: Hawai‘i Institute of Marine Biology, School of Ocean & Earth Sciences & Technology, University of Hawai‘i at Mānoa, Coconut Island, Kāne‘ohe, HI, United States

8: School of Zoology, Faculty of Life Sciences & The Steinhardt Museum of Natural History, Tel Aviv University, Tel Aviv 6997801, Israel

9: ARC Centre of Excellence for Coral Reef Studies, School of Biological Sciences, The University of Queensland, St. Lucia, QLD 4072, Australia

10: The Mina & Everard Goodman Faculty of Life Sciences, Bar-Ilan University, Ramat Gan 5290002, Israel

11: Phuket Marine Biological Center, Box 60, Phuket 83000 Thailand

†EQUAL FIRST AUTHORS:

Naohisa Wada & Hideki Yuasa

*CORRESPONDING AUTHORS:

Tetsuya HAYASHI (thayash”at”bact.med.kyushu-u.ac.jp), Takehiko ITOH (Takehiko”at”bio.titech.ac.jp), and Nina YASUDA (nina27”at”cc.miyazaki-u.ac.jp)

Abstract

Background:

Population outbreaks of the crown-of-thorns starfish (*Acanthaster planci* sensu lato；COTS), a primary predator of reef-building corals in the Indo-Pacific Ocean, are major concern s in coral reefs management. While biological and ecological knowledge on COTS has been accumulated since the 1960s, little is known about their associated bacteria. The aim of this study was to provide fundamental information on dominant COTS-associated bacteria by multi-disciplinary approach.

Methods:

We first conducted 16S rRNA metabarcoding for bacterial profiles on different body parts of COTS, and obtained full-length 16S rRNA gene sequence of the single dominant bacterium for phylogenetic analysis. A total 205 COTS individuals from 17 locations throughout the Indo-Pacific Ocean was examined for the presence of COTS associated bacteria (named COTS27). The COTS27 localization was visualized by Fluorescence in situ hybridization (FISH) using a COTS27-specific probe. Furthermore, COTS27 chromosome genome was reconstructed from the hologenome sequence data of COTS.

Results:

We discovered that a single bacterium exists at high densities in the subcuticular space of COTS forming a biofilm-like structure over the entire body surface of COTS. COTS27 belongs to a clade that presumably represents a distinct order (so-called marine spirochetes) in the phylum *Spirochaetes* and is universally present in COTS throughout the Indo-Pacific Ocean. The reconstructed genome of COT27 includes some genetic traits that are probably linked to adaptation to marine environments and evolution as an extracellular endosymbiont in subcuticular spaces.

Conclusions:

COTS27 can be found in the allopatrically spectated 3 different COTS species, ranging from northern Red Sea to the central Pacific, implying symbiotic relationship started before speciation (about 2 million years ago). The universal and nearly mono-specific association of COTS27 with COTS potentially provides a useful model system to study symbiont-host interactions in marine invertebrates.

Keywords:

Crown-of-thorns starfish, Subcuticular bacteria, Marine spirochetes

Introduction

Coral reefs support almost one-third of the world's marine coastal species [1,2]. However, population outbreaks of a coral predator, the crown-of-thorns starfish (*Acanthaster planci* sensu lato; COTS), are a great threat to Indo-Pacific coral reef ecosystem integrity and biodiversity [3–5]. A 27-year study in the Great Barrier Reef concluded that COTS outbreaks and tropical cyclones were the main causes of the loss of reef-building corals [6]. While some aspects of the biology of COTS, such as their reproduction, larval ecology, phylogeography, and behaviour, have been studied intensively [5], little is known about their associated microbiota.

The bacterial symbionts of marine invertebrates have been shown to be important to their host organisms [7]. In echinoderms, bacterial communities may play a role in larval settlement [8], amino acid uptake on the integument [9], and digestive strategies in the gut [10,11], and these communities may even drive morphological variations [12]. Bacterial symbionts are prevalent on the body surfaces of echinoderms [13], with high host specificity [14,15]. Notably, extracellular endosymbionts known as subcuticular bacteria (SCB [16]) reside under the cuticular layer of echinoderm fauna from all five extant classes, and it has been postulated that these bacteria provide dissolved free amino acids to their echinoderm hosts [9,17]. To date, molecular genetic approaches targeting the 16S rRNA gene have revealed that several proteobacteria (*Alphaproteobacteria* and *Gammaproteobacteria*) are SCB that are distributed in the subcuticular space of two brittle star species [13,18], one holothurian species [19], and one asteroid species [19].

Despite their potential biological importance, studies of the bacteria associated with COTS have been mostly culture-based, and only two culture-independent studies have been published to date. Carrier et al. reported shifts in COTS larval microbiomes with diet [20]. Høj et al. found that adult COTS have tissue-specific bacterial communities, largely comprising four major bacterial groups: *Mollicutes* in male gonads, *Spirochaetales* in the body wall, *Hyphomonadaxeae* in the tube feet, and *Oceanospirillales* in all tissues [21]. Although these studies significantly increased our

understanding of the COTS microbiome, there is still a great lack of knowledge regarding COTS-associated bacteria, particularly SCB. In the current study, we aimed to obtain primary information on the indigenous bacteria of the body surface of COTS. A single dominant bacterium operational taxonomic unit (OTU) named COTS27 was initially identified by means of 16S rRNA metabarcoding screening; therefore, we engaged in a comprehensive investigation of its phylogenetic status. The indigeneity of COTS27 was examined in three *Acanthaster* species throughout the Indo-Pacific, and the localization of COTS27 in the body of COTS was investigated by fluorescence *in situ* hybridization (FISH) using the designed COTS27-specific probe. Furthermore, we reconstituted the chromosome sequence of COTS27 from the hologenome sequence data of a COTS to elucidate the genetic features of COTS27 and gain further insights into interactions between COTS and COTS27.

Results

Identification of a single OTU (COTS27) that dominates the body surface microbiota of COTS using 16S rRNA metabarcoding analysis

We used 16S rRNA metabarcoding to analyse the bacterial composition of the microbiota in eight different organs and body parts (tips and bases of aboral spines from both discs and arms, tips and bases of ambulacral spines, tube feet, and pyloric stomachs) of six COTS individuals (**Fig. 1**) that were collected in Miyazaki and Okinawa, Japan (three individuals from each location and triplicates of each body part). Seawater samples from the same locations were similarly analysed for their bacterial compositions (three samples from each location). After quality filtering, 1,427,570 and 108,334 sequences of bacterial origins were obtained from the COTS (n=130 for all body parts and organs in all replicates) and seawater (n=6) samples, respectively (averages of 10,981 and 18,056 sequences per sample, respectively; **Suppl. table S1**). From the abovementioned sequences, 671 bacterial OTUs were identified, 503 and 401 of which were found in the COTS and seawater samples, respectively. There were 233 OTUs that were common to both. The OTUs that were identified in the COTS and seawater samples represented 19 and 16 phyla, 34 and 24 classes, 93 and 68 orders, and 145 and 96 families, and 29 and 22 OTUs were unclassified, respectively.

In the six COTS individuals that were examined, a single unclassified OTU (OTU1) occupied 61.8% of the total sequences on average, predominately in most body parts and organs of both the

Okinawan and Miyazaki COTS populations (60.3% and 63.8% of the total sequences on average were assigned to OTU1 in the Okinawa and Miyazaki COTS collections, respectively; **Fig. 2**), despite more than 720 km separating these populations. The high abundance of OTU1 in all individuals was attributed to the surface body parts (68.8% and 79.1% of the sequences from all spine and tube foot samples, respectively; 8.0% of the sequences from the pyloric stomach samples) (**Fig. 2** and **Suppl. table S2**). OTU1 was abundant at both the oral (ambulacral spines and tube feet) and aboral (discs and arm spines) sides (**Suppl. fig. S1** and **Suppl. table S2**). The tips and bases of the spines showed roughly the same levels of OTU1 abundance (**Suppl. fig. S1** and **Suppl. table S2**). Five of the 88 spine samples that were examined (containing both tip and base) contained no or only a small abundance of OTU1 (**Suppl. fig. S1**); however, OTU1 was abundant in the other two DNA preparations of the triplicates from the same sample in all cases, suggesting that the exceptional data from the five preparations were due to some technical problems. Regarding the seawater samples, OTU1 was only detected in the Okinawan samples and in a small amount (0.026%; **Suppl. fig. S1**). The relatively abundant bacteria other than OTU1 are described in **Appendix 1**. In total, we identified 41 different OTUs, including OTU1, in all COTS individuals from the two locations, which may represent the core members of the bacterial community within COTS (**Suppl. fig. S2**). The core bacterial OTUs other than OTU1 occupied up to 18.4% (the abundance of each OTU was less than 3.5%) of the total reads from all COTS samples (**Suppl. fig. S2d**). These results indicate that a single bacterium (OTU1) predominantly colonizes the body surface of COTS. Hereafter, we refer to the bacterium corresponding to our OTU1 as COTS27.

Phylogeny of COTS27 from 16S rRNA gene sequences

To elucidate the phylogenetic status of COTS27, we determined the full-length 16S rRNA gene sequences of COTS27 in five tube foot samples obtained from Miyazaki (n=3) and Okinawa (n=2). The five sequences were largely identical (99.9-100% similarity), and there was a partial sequence overlap with the 16S rRNA gene sequence of a spirochete-like bacterium (GenBank accession No. PRJNA420398) that was a dominant bacterium on the body wall of COTS from the Great Barrier Reef [21]. Both the maximum likelihood (ML) and neighbour joining (NJ) phylogenetic trees (**Fig. 3a** and **Suppl. fig. S3**), which were based on full-length 16S rRNA gene sequences, showed that the 5 COTS27 sequences formed a distinct subclade within one of the three clades of the unclassified

spirochete cluster (named clade I; **Fig. 3a** and **Suppl. fig. S3**). All sequences in this unclassified spirochete cluster originated from marine environments and marine invertebrates (see **Appendix 2** for more details of clade I), apart from a single sequence obtained from a wetland soil sample (GenBank accession No. FQ660021.1). Hereafter, we refer to these spirochetes as "marine spirochetes", as referred to by Høj et al. (2018) [21]. These marine spirochetes formed a distinct cluster within the phylum *Spirochaetes*, with the order *Brachyspirales* being their closest relative (**Fig. 3a** and **Suppl. fig. S3**).

These results indicate that COTS27, together with many unclassified marine environmentally associated spirochetes, forms a distinct clade (marine spirochetes) within the phylum *Spirochaetes*. A single rRNA gene sequence (GenBank accession No. FN424158.1) was previously detected from the sponge *Chondrilla* and was proposed as a new lineage of spirochete related to *Brachyspiraceae* [22]. In our analysis, this sequence also belonged to the marine spirochetes (clade II; **Fig. 3a** and **Suppl. fig. S3**). Notably, the 16S rRNA gene sequences of the marine spirochetes, including COTS27, showed only a 76.3–78.1% identity to those of the order *Brachyspirales*, which is well below the proposed threshold for defining a novel order (82.0%) [23]. Thus, the marine spirochetes most likely represent a distinct order in the phylum *Spirochaetes*.

Universal association of COTS27 with COTS throughout the Indo-Pacific Ocean

The presence of COTS27 or COTS27-like bacteria in COTS individuals inhabiting various geographic regions was determined by PCR that was designed to specifically amplify a 261 bp fragment of the COTS27 16S rRNA gene. PCR products were obtained from all 195 COTS individuals that were collected at 15 locations throughout the Indo-Pacific Ocean and included three known species of COTS (**Figs. 1a** and **c**). Sequencing of the PCR products that were obtained from 53 randomly selected individuals confirmed the presence of COTS27 or very close relatives in these COTS. The ML and NJ trees based on these 261 bp sequences (**Suppl. figs. S4a** and **b**) revealed that all sequences formed a tight cluster together with the six COTS27 sequences obtained from the abovementioned phylogenetic analysis and with those obtained from the genome reconstruction described below. However, the sequences from the Israeli COTS population (Red Sea species) formed a separate clade from those of the other Indo-Pacific populations (northern Indian Ocean species or Pacific Ocean species). Among the northern Indo-Pacific species, only a single nucleotide

polymorphism (SNP) was detected in one sequence obtained from Japan (Wakayama C29 adult JPN; **Suppl. figs. S4a** and **b**). These results indicate the universal association of COTS27 with the Indo-Pacific COTS species.

Localization and biofilm-like structure formation of COTS27 in subcuticular spaces across the body surface of COTS

The localization of COTS27 in COTS was analysed with FISH using three probes: a general bacterial probe, EUB338mix, a COTS27-specific probe, COTSsymb, and a negative control probe, Non338. For this analysis, six different tissue types (aboral spines from both the discs and arms, tube feet, pyloric stomachs, pyloric caeca and gonads; see **Fig. 1**) were prepared from three individuals collected in Miyazaki, and three serial sections of each tissue type were analysed. COTS27 was consistently detected on the body surface. The localization pattern was similar in all individuals in the subcuticular spaces of the body surface (**Figs. 4** and **5**), as demonstrated by the binding of the general bacterial probe (see the binding signals of EUB338mix and COTSsymb at the COTS central disc spines in **Figs. 4a–b**). COTS27 was localized in the subcuticular spaces on both the aboral side (**Figs. 5a-d**; spines of the discs and arms, dermal papulae, and pedicellariae, see **Fig. 1b** and **Suppl. fig. S5 for** their anatomical locations and structures) and ambulacral side (**Figs. 5e**; the stems of the tube feet). No COTS27 or any other bacterial signals were detected in the pyloric caeca and gonads (**Fig. 5g–h**). Likewise, no COTS27 was found in the pyloric stomachs, although numerous cyanobacteria-like bacteria were detected (**Fig. 5f**). In the cross sections, COTS27 displayed continuous layer-like signals (**Figs. 4b, 5a-c** and **5e**), although a patchy distribution was also occasionally observed. Furthermore, three-dimensional (3D) images showed that COTS27 formed a biofilm-like structure along the ruffled structure of the epidermis of pedicellariae (**Fig. 5d**). These observations indicate that COTS27 is an SCB that covers nearly all the surface area (the epidermis) of COTS by forming a biofilm-like structure. COTS27 cells appear to have filamentous or long rod-like shapes (**Figs. 5c** and **e**), but different approaches are required to accurately determine their cell morphology.

Reconstruction of the COTS27 chromosome

A circular chromosome sequence of COTS27 was reconstructed using the hologenome sequence from a COTS sample collected in Miyazaki (**Fig. 6** and **Suppl. table S3**), with 90.66% completeness and 0.26% contamination as evaluated by CheckM [24]. Although 23 gaps remained in the final assembly, all were derived from tandem repeats in the genic regions, and the estimated gap sizes were less than 28 bp. The COTS27 chromosome was 2,684,921 bp in length, with a 39.6% average GC-content, and contained 1,650 protein-coding genes, three rRNA genes, and 35 tRNA genes. No transposable elements or prophages were detected. The 1,650 protein-coding genes included a giant gene (53,043 bp in length; COTS27_01023), but its function is currently unpredictable. Among the three rRNA genes, the 16S rRNA gene was located separately from the 23S and 5S rRNA genes. The 35 tRNA genes covered all 20 basic amino acids. Phylogenetic analysis using the sequences of 43 conserved marker genes with 5,656 reference bacterial and archaeal genomes placed COTS27 in the phylum *Spirochaetes* (**Fig. 3b**), supporting the results of the 16S rRNA sequence-based analysis (**Fig. 3a**).

Biological features of COTS27 inferred from the gene repertoire

In the Clusters of Orthologous Groups (COG) functional category-based principal component analysis of COTS27 using 716 high-quality *Spirochaetes* genomes obtained from the IMG database [25], COTS27 was placed in a distinct position with regard to all *Spirochaetes* (**Suppl. fig. S6**), indicating potential biological features unique to COTS27. Subsequently, we performed a Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis to obtain basic information on the biology of COTS27 (**Suppl. table S4**). Complete or near-complete biosynthesis pathways for 18 of the 20 basic amino acids were identified, excluding those for asparagine and aspartic acid. Although the guanine ribonucleotide biosynthesis pathway was not complete (one block missing), all other nucleotide biosynthesis pathways were detected. For vitamin and cofactor biosynthesis, the complete biosynthetic pathways of nicotinamide adenine dinucleotide (NAD), coenzyme A and riboflavin and the C1-unit interconversion were identified. The pathways for fatty acid biosynthesis, beta-oxidation, and phosphatidylethanolamine biosynthesis were also detected. The conservation of these metabolic pathways suggests that COTS27 is not strongly metabolically dependent on the host COTS.

For energy production, COTS27 contained the complete glycolysis pathway and TCA cycle. Genes for succinate dehydrogenase, cytochrome c oxidase, and F-type ATPase were also identified;

however, no genes for NADH dehydrogenase were detected. Instead, COTS27 had an operon encoding a sodium-pumping NADH:ubiquinone oxidoreductase (Na^+-NQR) (**Suppl. fig. S7**). Some bacteria living in Na^+-rich environments (e.g., marine or intercellular environments) have Na^+-NQR that oxidizes NADH to NAD^+ and pumps Na^+ out of cells, thus functioning in the respiratory chain and in the maintenance of intercellular homeostasis in Na^+-rich environments [26]. In line with these features of Na^+-NQR, out of the 716 high-quality *Spirochaetes* genomes (above), only one genome, which was also reconstituted from the metagenome sequences of a seawater sample (*Spirochaetaceae* bacterium NP120, IMG Genome ID: 2509276057), contained the Na^+-NQR operon. The acquisition of Na^+-NQR may represent one of the mechanisms for the adaptation of COT27 to marine environments.

Consistent with the general characteristics of *Spirochaetes*, which are basically gram-negative, helical or spiral-shaped, and motile, with periplasmic flagella [27], COTS27 contained sets of genes for the biosynthesis of DAP-type peptidoglycan, lipopolysaccharide and phosphatidylethanolamine. While a set of genes for flagellar biosynthesis was identified, no gene for chemotaxis, such as methyl-accepting chemotaxis proteins and chemotaxis-related signal transduction, was detected. This feature is very unusual in *Spirochaetes*; most of the high-quality *Spirochaetes* genomes mentioned above (>98%) contained gene sets for both flagellar biosynthesis and chemotaxis. The remaining genomes, such as those from the genus *Sphaerochaeta*, lack genes for both flagellar biosynthesis and chemotaxis, suggesting that chemotaxis genes have been selectively lost from the COTS27 genome. It has been proposed that the active migration and colonization by symbionts through motility and chemotaxis are indispensable for the acquisition of microbial partners by host organisms from environments [28]. However, the selected lost chemotaxis genes appear to represent a specific adaptation strategy of COT27 as an SCB. COTS27 may require flagella to spread and stably and widely colonize subcuticular spaces, but chemotaxis is no longer required after specially adapting to the subcuticular spaces of COTS.

Discussion

We identified a single bacterium that is universally present and numerically dominant in the subcuticular spaces of COTS, covers all body surfaces of COTS as an SCB by forming a biofilm-like structure, and belongs to the so-called marine spirochetes, which likely represent a previously

undefined order in the phylum *Spirochaetes*. The universal association of COTS27 with the Indo-Pacific COTS species suggests a long history of the COTS-COTS27 association. COTS are thought to have diverged allopatrically into four species during the Pliocene-Early Pleistocene (1.95–3.65 Myr ago) in the Indo-Pacific Ocean [29]. Therefore, the association of COTS27 with at least three of the four extant COTS species (data on the fourth species are currently not available) implies that the mutual relationship between COTS and COTS27 started prior to the Pliocene or Early Pleistocene eras. This hypothesis is supported by the finding that COTS27 from the northern Red Sea (forming a different cluster from the Indo-Pacific regions; **Suppl. fig. S4**) was notably different from COTS27 from other regions, although additional phylogeographic analyses of the Indo-Pacific COTS and COTS27 are required. Regarding the evolutionary and functional aspects of COTS27 and its association with COTS, we obtained two key findings from the genome sequence analysis: 1) the presence of Na^+-NQR and 2) the selective loss of chemotaxis genes. These findings are intriguing because these features are likely linked to the adaptation to marine environments and evolution as an extracellular endosymbiont in subcuticular space, respectively. Additional genome sequences of marine spirochetes are required to verify this hypothesis and elucidate the evolutionary and functional aspects of the COTS27-COTS association.

SCB are commonly found in echinoderm fauna [13,19] and have been classified into three morphotypes [13,15,16]. Among these, COTS27 most likely belongs to the SCB Type 2, which has a long spiral shape and is commonly found in all five echinoderm classes [15,30]. Jackson et al. suggested the presence of a highly dominant *Spirochaetae* in the hard tissues (including the body wall) of some starfish species in the United States and Australia [31]. Such a wide distribution of spirochetes or spirochete-like bacteria in echinoderms suggests that many echinoderms may have established symbiotic relationships with marine spirochetes that are similar to that between COTS27 and COTS. Further explorations of SCB in a wider range of echinoderms would provide more detailed insights into the association between echinoderm hosts and marine spirochetes.

The outer body surfaces of marine organisms often represent a highly active interface between an organism (host) and the surrounding marine environment regarding aspects such as light exposure, gas exchange, nutrient uptake and interactions with other fouling organisms, consumers, and pathogens [32]. It is plausible that SCB are also involved in these interactions. However, the physiological and potentially ecological roles of SCB largely remain unexplored. The universal and

nearly mono-specific association of COTS27 with COTS would be an ideal model system to further explore the roles of SCB as well as symbiont-host interactions in marine invertebrates. Moreover, COTS27 could be used as an environmental marker to monitor and/or predict population outbreaks of COTS.

Conclusions

Despite the fact that the 205 COTS individuals utilized in our current analyses were collected over a 13-year period (2004–2017) and from 17 different locations across the Indo-Pacific, the COTS27 community remained exceptionally ubiquitous both spatially and temporally. Besides, it is likely that COTS have hosted COTS27 on the surface of their bodies for more than 2 million years before allopatric speciation occurred during the Pliocene-Early Pleistocene. These indicate that COTS27 are likely an extracellular endosymbiotic bacteria strongly associated with COTS. Recognizing this tight holobiontic relationship and chromosome genome information of COTS27 presented here will significantly contribute to our understanding of COTS and its microbiology, and may also serve as a model system for studying endosymbionts in marine invertebrates at large.

Materials and Methods

Sample collection and preparation for DNA analyses and histology

We collected 205 individual COTS from 17 locations throughout the Indo-Pacific (**Fig. 1**, and **Suppl. table S5**). Triplicate sub-samples were taken from eight different body parts and organs from six individuals that were collected in Okinawa and Miyazaki (three from each location) in Japan, (**Fig. 1**). Seawater samples were also collected in triplicate from each of the latter two locations. Consequently, 130 DNA samples from the six COTS individuals and six seawater DNA samples were prepared and used for the metabarcoding analysis (**Fig. 1c** and **Suppl. materials and methods**). The tube foot DNA samples from five of the six individuals were used to determine the full-length 16S rRNA gene sequence of COTS27. DNA samples prepared from the tube feet of 195 individuals collected from 15 geographic locations were used to examine the presence of COTS27 in three species of COTS [33] (**Fig. 1a, c**). DNA from the tube feet of one individual collected in Miyazaki (Japan) was used for the COTS hologenome sequencing. The samples were stored in 100% ethanol before the DNA extraction, and DNA was extracted as previously described [34].

Samples of six different tissue types (**Fig. 1**) were prepared from the remaining three individuals collected in Miyazaki for the FISH analyses (see **Suppl. materials and methods** for the details).

16S rRNA metabarcoding

16S rRNA amplicon libraries (V4 region) were prepared as previously described [35,36] using the primers listed in **Suppl. table S6** and **Suppl. fig. S8** and subjected to paired-end (PE) sequencing (300 bpx2) on the Illumina MiSeq platform. The obtained PE sequences were merged using software USEARCH v8.1.1861 [37] and filtered using MOTHUR v.1.36.1 [38] with the following criteria: 1) read lengths between 200 and 305 bp; 2) read quality score > 27; and 3) homopolymer read length < 8 bp. From the high-quality merged read sequences, chimeric reads were eliminated using USEARCH (parameters: reference mode, RDP Gold database, and minimum division of five) in UCHIME [39]; then 2,100,477 merged sequences without any ambiguous bases were retrieved. Finally, these sequences were clustered into operational taxonomic units (OTUs) with a cut-off value of 97% identity using UPARSE [40]. The OTUs were assigned to known taxonomic groups by mapping onto the Silva SSU 132 database (https://www.arb-silva.de/) using MOTHUR with a cut-off value of 65. The OTUs that were assigned to unknown taxa were further processed with Silva SINA [41]; the minimum identity criterion to query the sequences was set to 0.70. A total of 1,535,904 sequences were assigned to bacteria, and others were assigned as eukaryotes, Archaea, chloroplasts, or unknown origins (**Suppl. table S1**).

Phylogenetic analysis of COTS27 using the full-length 16S rRNA gene sequence

The full-length 16S rRNA gene sequences of COTS27 were obtained from five COTS individuals using a COTS27-specific primer set that was designed in this study (**Suppl. materials and methods**, **Suppl. table S6** and **Suppl. fig. S8**). The sequences were used to reconstruct phylogenetic trees using the maximum likelihood (ML) and neighbour joining (NJ) methods (**Suppl. materials and methods**).

PCR screening and sequencing of COTS27

COTS individuals were screened for the presence of COTS27 by PCR using primers that were designed to specifically amplify a 261 bp fragment of the 16S rRNA gene (**Suppl. table S6** and **Suppl. fig. S8**). The PCR products obtained from 53 randomly selected samples from all COTS27-positive samples (n=195) were sequenced and used for phylogenetic reconstruction (**Suppl. materials and methods**).

Fluorescence in situ hybridization (FISH)

FISH was performed using a COTS27-specific oligonucleotide probe (COTSsymb; **Suppl. table S6** and **Suppl. fig. S8**), EUB338mix [42], and Non338 [43]. COTSsymb was designed using the ARB software package [44] with the SSURef_NR99_128 database from ARB SILVA (https://www.arb-silva.de/) as a reference. The specificity of COTSsymb was tested in the Silva SSU 131 database with the TestProbe function (https://www.arb-silva.de/). The FISH experiments were performed on three serial sections (thickness of 5 μm) from the six body parts using the three probes separately, as previously described [45]. Bacterial localization was observed using a confocal laser scanning microscope (LSM 550; Zeiss, Germany) with two channels for Cy3 fluorescence (excitation: 555 nm; emission: BP 490-635) and COTS autofluorescence (excitation: 488 nm; emission: non-filter). In addition, we reconstructed three-dimensional (3D) structures from thick sections (thickness; 50 μm) of the disc spines using the Z-stack function (interval of 0.2 μm) of the LSM770 confocal laser scanning microscope (Zeiss) with two channels for Cy3 fluorescence (excitation: 561 nm; emission: 561-641 nm) and COTS autofluorescence (excitation: 405 nm; emission: 410-552 nm) and the surface rendering function of Imaris software (BitplaneAG, USA). We also verified non-hybridization on an *E. coli* culture for a specific-COTS27 oligonucleotide probe (data not shown); basically, the true bacteria signals could be distinguished from the non-specific binding of the probe by applying the negative control treatment with the Non338 oligonucleotide probe on the other serial section.

Reconstitution of the COTS27 chromosome from the hologenome sequences of a COTS sample

Two PE libraries and six mate-pair libraries were prepared and sequenced using Illumina HiSeq 2500 sequencers (**Suppl. table S7**). *De novo* assembly was performed using Platanus v. 1.2.3 [46].

To identify the COTS27-derived sequences in the assembly, scaffolds with a high coverage depth (≥×200) were selected. The average coverage depth for all scaffolds was ×130. The longest scaffold, which was identified as the COTS27 chromosome, was closed by Sanger sequencing, and an alternative assembly was obtained using Platanus-allee v. 2.0.0 [47]. A circular view of the COTS27 chromosome was created using CGView Server [48] and manual processing. The completeness of the final assembly was evaluated using CheckM v. 1.0.11 [24].

Gene prediction and functional annotation

The protein-coding sequences (CDSs) were predicted by PROKKA v. 1.12 [49], followed by manual curation. For functional annotation, Clusters of Orthologous Groups (COG) were assigned by querying the CDSs against the Conserved Domain Database (CDD) with COG position-specific scoring matrices (PSSMs) using RPS-BLAST. Additionally, K numbers of Kyoto Encyclopedia of Genes and Genomes (KEGG) were assigned to each CDS; BlastKOALA [50] and KofamKOALA [51] were used to perform the searches in the KEGG GENES and KOfam databases, respectively.

Principal component analysis and phylogenetic analysis based on the genome sequences

Principal component analysis was performed based on the compositions of the COG functional categories. The genome sequences of the *Spirochaetes* bacteria were retrieved from the DOE-JGI IMG database, and 716 high-quality genomes (completeness >90% and contamination <5% as evaluated by CheckM) were retained (see **Suppl. materials and methods** for the details). Whole genome sequence-based phylogenetic analysis was performed using CheckM v. 1.0.11 [24] to obtain the ML tree of the COTS27 and 5,656 bacterial and archaeal genomes based on the sequences of 43 conserved marker genes. The tree was visualized using FigTree v. 1.4.3 (http://tree.bio.ed.ac.uk/software/figtree/).

List of abbreviations

CDD: Conserved Domain Database

CDSs: protein-coding sequences

COG: Clusters of Orthologous Groups

COTS: Crown-of-thorns starfish

FISH: Fluorescence *in situ* hybridization

KEGG: Kyoto Encyclopedia of Genes and Genomes

ML: Maximum likelihood

NJ: Neighbour joining

NAD: Nicotinamide adenine dinucleotide

Na+-NQ: Sodium-pumping NADH:ubiquinone oxidoreductase

OTU: Operational taxonomic unit

SCB: Subcuticular bacteria

• *Funding*

This study was supported by JSPS KAKENHI 221S0002, 16H06279 (PAGS), Grant-in-aid for Young Scientists (A) (17H04996) and (B) (25870563), the Program to Disseminate Tenure Tracking System in University of Miyazaki, the Environment Research and Technology Development Fund (4RF-1501), MOE, Japan, the postdoctoral fellows programme of Academia Sinica (2nd session, 2017), Grant-in-Aid for JSPS Fellows (18J23317), and the European Union's Horizon 2020 research and innovation programme under MSCA No. 796025.

• *Authors' contribution*

N.W., H.Y., T.I., T. H, and N.Y. conceived the research idea and designed the experiments. N.W., H.Y., Y. H, H. S, Z.F., O. B, G.E., N.T., and N.Y. conducted the field sampling. N.W., Y.G., and N.Y. performed the 16S rRNA metabarcoding. N.W. and N.Y. performed the phylogenetic analysis using the full-length 16S rRNA gene. N.W. H. S, Z.F., OB, N.T., and N.Y. performed the PCR screening and sequencing of the PCR products. N.W. conducted the FISH experiments. H.Y., R.K., D.Y., and A.T. conducted the sequencing and analysis of the COTS27 genome. N.W., H.Y., T.I., T. H, and N.Y. made major contributions to the manuscript writing and figure making. R.K., Y.G., Y.O., S.T., H.S., Z.F., O.B., and G.E. contributed to writing and editing the manuscript. All authors critically reviewed, revised and ultimately approved this final version.

• ***Acknowledgements***

We thank members of Yasuda Laboratory and Atsushi Toyoda Laboratory, J. Suzuki, H. Kagawa, S. Nagai, Y. R. Kubota, Y. Uchida, K. Sakai, T. Miki, C. Shinzato, L. Høj, and D. Bourne for their logistical and technical support. We also thank M. Timmers, O. Ben-Zvi, N. Phongsuwan, S. Ayalon, the Seaver Institute, and the Interuniversity Institute for Marine Sciences of Eilat for local sampling and analysis.

References

1. Spalding MD, Ravilious C, Green EP. World atlas of coral reefs. Prepared at the UNEP World Conservation Monitoring Centre. Univ Calif Berkeley EEUULinks. 2001;

2. Wilkinson C. Status of coral reefs of the world: 2004. Australian Institute of Marine Science; 2004.

3. Moran P, Bradbury R. The crown-of-thorns starfish controversy. Search. 1989;20:3–6.

4. Birkeland C, Lucas J. Acanthaster Planci: Major Management Problem of Coral Reefs. CRC Press; 1990.

5. Pratchett MS, Caballes CF, Sweatman JAR-P& HPA. Limits to Understanding and Managing Outbreaks of Crown- of- Thorns Starfish (*Acanthaster* spp.). Oceanogr Mar Biol. 2014;52:133–200.

6. De'ath G, Fabricius KE, Sweatman H, Puotinen M. The 27–year decline of coral cover on the Great Barrier Reef and its causes. Proc Natl Acad Sci. 2012;109:17995–9.

7. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, et al. Animals in a bacterial world, a new imperative for the life sciences. Proc Natl Acad Sci. 2013;110:3229–36.

8. Bryan PJ, Rittschof D, McClintock JB. Bioactivity of echinoderm ethanolic body-wall extracts: an assessment of marine bacterial attachment and macroinvertebrate larval settlement. J Exp Mar Biol Ecol. 1996;196:79–96.

9. Lesser MP, Walker CW. Comparative study of the uptake of dissolved amino acid in sympatric brittle stars with and without endosymbiotic bacteria. Comp Biochem Physiol Part B Comp Biochem. 1992;101:217–23.

10. Thorsen MS. Abundance and biomass of the gut-living microorganisms (bacteria, protozoa and fungi) in the irregular sea urchin *Echinocardium cordatum* (Spatangoida: Echinodermata). Mar Biol. 1999;133:353–60.

11. Thorsen MS, Wieland A, Ploug H, Kragelund C, Nielsen PH. Distribution, identity and activity of symbiotic bacteria in anoxic aggregates from the hindgut of the sea urchin *echinocardium cordatum*. Ophelia. 2003;57:1–12.

12. Balakirev ES, Pavlyuchkov VA, Ayala FJ. DNA variation and symbiotic associations in phenotypically diverse sea urchin *Strongylocentrotus intermedius*. Proc Natl Acad Sci. 2008;105:16218–23.

13. Burnett WJ, McKenzie JD. Subcuticular bacteria from the brittle star *Ophiactis balli* (Echinodermata: Ophiuroidea) represent a new lineage of extracellular marine symbionts in the alpha subdivision of the class *Proteobacteria*. Appl Environ Microbiol. 1997;63:1721–4.

14. McKenzie JD, Kelly MS. Comparative study of sub-cuticular bacteria in brittlestars (Echinodermata: Ophiuroidea). Mar Biol. 1994;120:65–80.

15. Kelly MS, McKenzie JD. Survey of the occurrence and morphology of sub-cuticular bacteria in shelf echinoderms from the north-east Atlantic Ocean. Mar Biol. 1995;123:741–56.

16. Holland ND, Nealson KH. The Fine Structure of the Echinoderm Cuticle and the Subcuticular Bacteria of Echinoderms. Acta Zool. 1978;59:169–85.

17. McKenzie JD, Black KD, Kelly MS, Newton LC, Handley LL, Scrimgeour CM, et al. Comparisons of fatty acid and stable isotope ratios in symbiotic and non-symbiotic brittlestars from Oban Bay, Scotland. J Mar Biol Assoc U K. 2000;80:311–20.

18. Morrow KM, Tedford AR, Pankey MS, Lesser MP. A member of the Roseobacter clade, *Octadecabacter* sp., is the dominant symbiont in the brittle star *Amphipholis squamata*. FEMS Microbiol Ecol. 2018;94:fly030.

19. Lawrence SA, O'Toole R, Taylor MW, Davy SK. Subcuticular Bacteria Associated With Two Common New Zealand Echinoderms: Characterization Using 16S rRNA Sequence Analysis and Fluorescence *in situ* Hybridization. Biol Bull. 2010;218:95–104.

20. Carrier TJ, Wolfe K, Lopez K, Gall M, Janies DA, Byrne M, et al. Diet-induced shifts in the crown-of-thorns (*Acanthaster* sp.) larval microbiome. Mar Biol. 2018;165:157.

21. Høj L, Levy N, Baillie BK, Clode PL, Strohmaier RC, Siboni N, et al. Crown-of-Thorns Sea Star *Acanthaster* cf. *solaris* Has Tissue-Characteristic Microbiomes with Potential Roles in Health and Reproduction. McBain AJ, editor. Appl Environ Microbiol. 2018;84:e00181-18.

22. Neulinger SC, Stöhr R, Thiel V, Schmaljohann R, Imhoff JF. New phylogenetic lineages of the <Emphasis Type="Italic">Spirochaetes</Emphasis> phylum associated with <Emphasis Type="Italic">Clathrina</Emphasis> species (Porifera). J Microbiol. 2010;48:411–8.

23. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nat Rev Microbiol. 2014;12:635–45.

24. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–55.

25. Chen I-MA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, et al. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. Nucleic Acids Res. 2019;47:D666–77.

26. Reyes-Prieto A, Barquera B, Juárez O. Origin and Evolution of the Sodium -Pumping NADH: Ubiquinone Oxidoreductase. PLOS ONE. 2014;9:e96696.

27. Paster BJ. Phylum XV. Spirochaetes Garrity and Holt 2001. In: Krieg NR, Staley JT, Brown DR, Hedlund BP, Paster BJ, Ward NL, et al., editors. Bergey's Manual® Syst Bacteriol Vol Four Bacteroidetes Spirochaetes Tenericutes Mollicutes Acidobacteria Fibrobacteres Fusobacteria Dictyoglomi Gemmatimonadetes Lentisphaerae Verrucomicrobia Chlamydiae Planctomycetes. New York, NY: Springer New York; 2010. p. 471–566.

28. Raina J-B, Fernandez V, Lambert B, Stocker R, Seymour JR. The role of microbial motility and chemotaxis in symbiosis. Nat Rev Microbiol. 2019;17:284.

29. Vogler C, Benzie J, Lessios H, Barber P, Wörheide G. A threat to coral reefs multiplied? Four species of crown-of-thorns starfish. Biol Lett. 2008;4:696–9.

30. Kelly MS, Barker MF, McKenzie JD, Powell J. The Incidence and Morphology of Subcuticular Bacteria in the Echinoderm Fauna of New Zealand. Biol Bull. 1995;189:91–105.

31. Jackson EW, Pepe-Ranney C, Debenport SJ, Buckley DH, Hewson I. The Microbial Landscape of Sea Stars and the Anatomical and Interspecies Variability of Their Microbiome. Front Microbiol. 2018;9:1829.

32. Wahl M, Goecke F, Labes A, Dobretsov S, Weinberger F. The Second Skin: Ecological Role of Epibiotic Biofilms on Marine Organisms. Front Microbiol. 2012;3.

33. Haszprunar G, Spies M. An integrative approach to the taxonomy of the crown-of-thorns starfish species group (*Asteroidea: Acanthaster*): A review of names and comparison to recent molecular data. Zootaxa. 2014;3841:271–84.

34. Yasuda N, Taquet C, Nagai S, Yoshida T, Adjeroud M. Genetic connectivity of the coral-eating sea star *Acanthaster planci* during the severe outbreak of 2006–2009 in the Society Islands, French Polynesia. Mar Ecol. 2015;36:668–78.

35. Apprill A, McNally S, Parsons R, Weber L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. Aquat Microb Ecol. 2015;75:129–37.

36. Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, et al. Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems. 2016;1:e00009-15.

37. Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 2010;26:2460–1.

38. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, et al. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. Appl Environ Microbiol. 2009;75:7537–41.

39. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics. 2011;27:2194–200.

40. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods. 2013;10:996–8.

41. Pruesse E, Peplies J, Glöckner FO. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. Bioinformatics. 2012;28:1823–9.

42. Daims H, Brühl A, Amann R, Schleifer KH, Wagner M. The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. Syst Appl Microbiol. 1999;22:434–44.

43. Wallner G, Amann R, Beisker W. Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. Cytometry. 1993;14:136–43.

44. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, et al. ARB: a software environment for sequence data. Nucleic Acids Res. 2004;32:1363–71.

45. Wada N, Pollock FJ, Willis BL, Ainsworth T, Mano N, Bourne DG. *In situ* visualization of bacterial populations in coral tissues: pitfalls and solutions. PeerJ. 2016;4:e2424.

46. Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, et al. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. Genome Res. 2014;24:1384–95.

47. Kajitani R, Yoshimura D, Okuno M, Minakuchi Y, Kagoshima H, Fujiyama A, et al. Platanus-allee is a de novo haplotype assembler enabling a comprehensive access to divergent heterozygous regions. Nat Commun. 2019;10:1702.

48. Stothard P, Wishart DS. Circular genome visualization and exploration using CGView. Bioinformatics. 2005;21:537–9.

49. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.

50. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J Mol Biol. 2016;428:726–31.

51. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics. 2019;btz859.

Figures

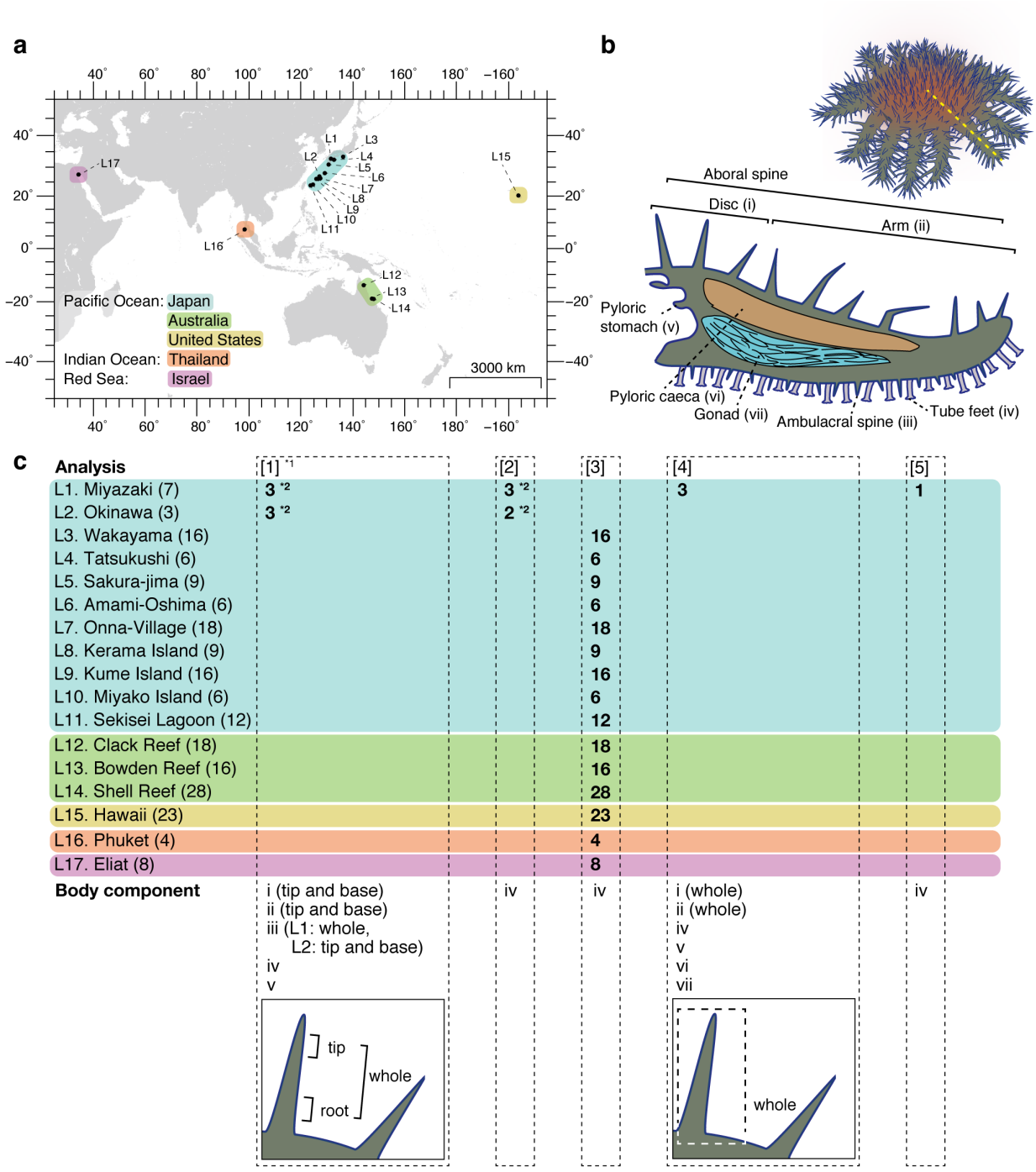

Fig. 1 Geographic and anatomical distributions of COTS individuals and the COTS body components analysed in this study.

The seventeen locations where the COTS individuals were collected (**a**) and the dissected body components of COTS for the analyses (**b**) are shown. The dashed yellow line (panel **b**) indicates the dissection line for the cross-sectional view. In panel (**c**), details of the samples used in each analysis are shown: [1] 16S rRNA metabarcoding, [2] phylogenetic analysis using the full-length 16S rRNA gene sequences, [3] PCR screening and sequencing of the

16S rRNA gene sequences of COTS27, [4] FISH analysis, and [5] hologenome sequencing analysis. *1: This analysis was performed in triplicate for each sample. *2: The same individuals were used in analyses [1] and [2].

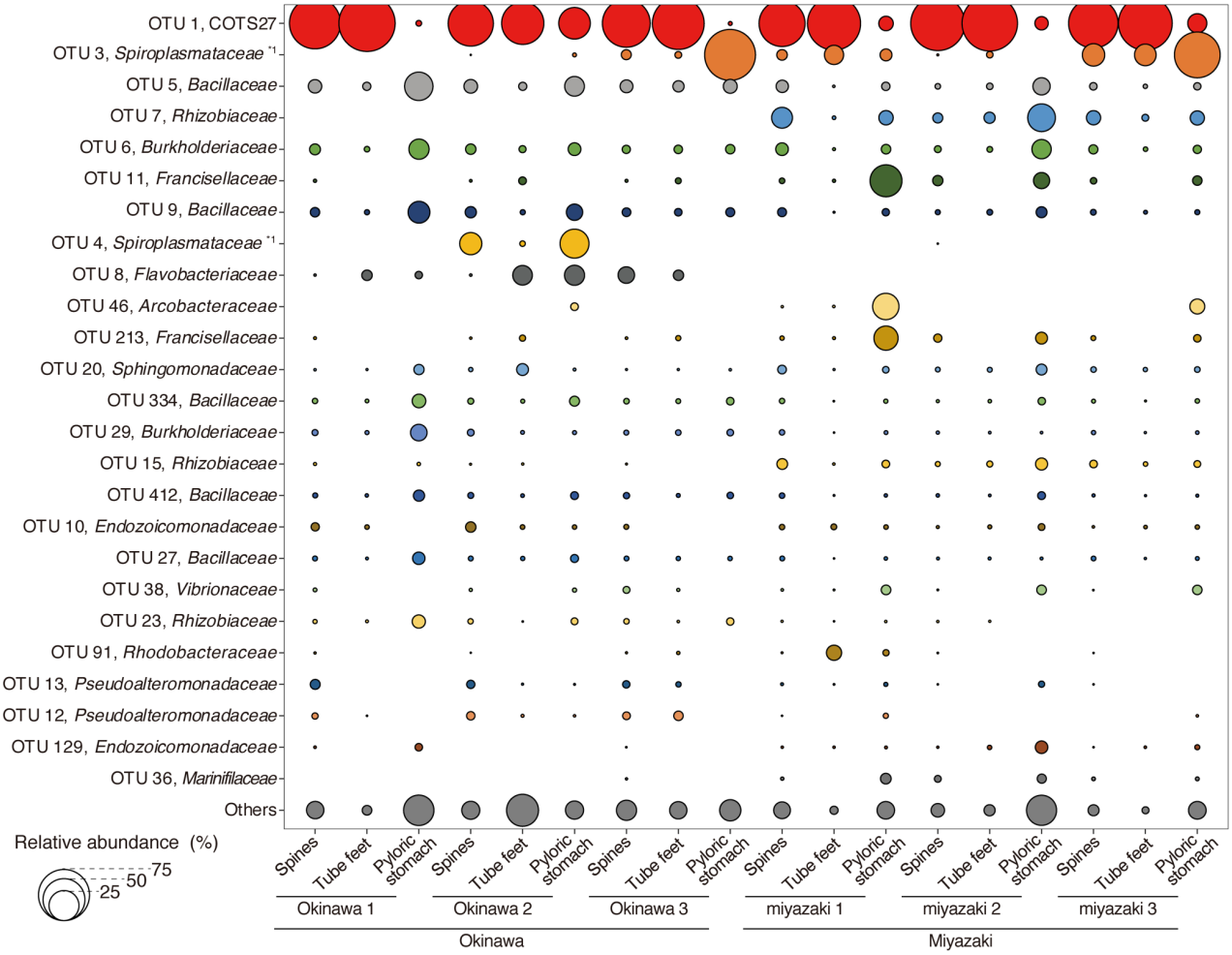

Fig. 2 The relative abundances of the 25 most abundant OTUs, including COTS27 (OTU1; red), in the total samples analysed in this study.

The bubble chart of the relative abundances was calculated from the merged replicates of each body part (spines, tube feet, and pyloric stomachs) in each COTS individual. The phylogenies of each OTU were determined based on the results (best hit) of BLAST searches against the NCBI nr/nt database. *1: The phylogenies of OTU3 and OTU4 were determined in the All-species Living Tree Project and RDP databases, respectively.

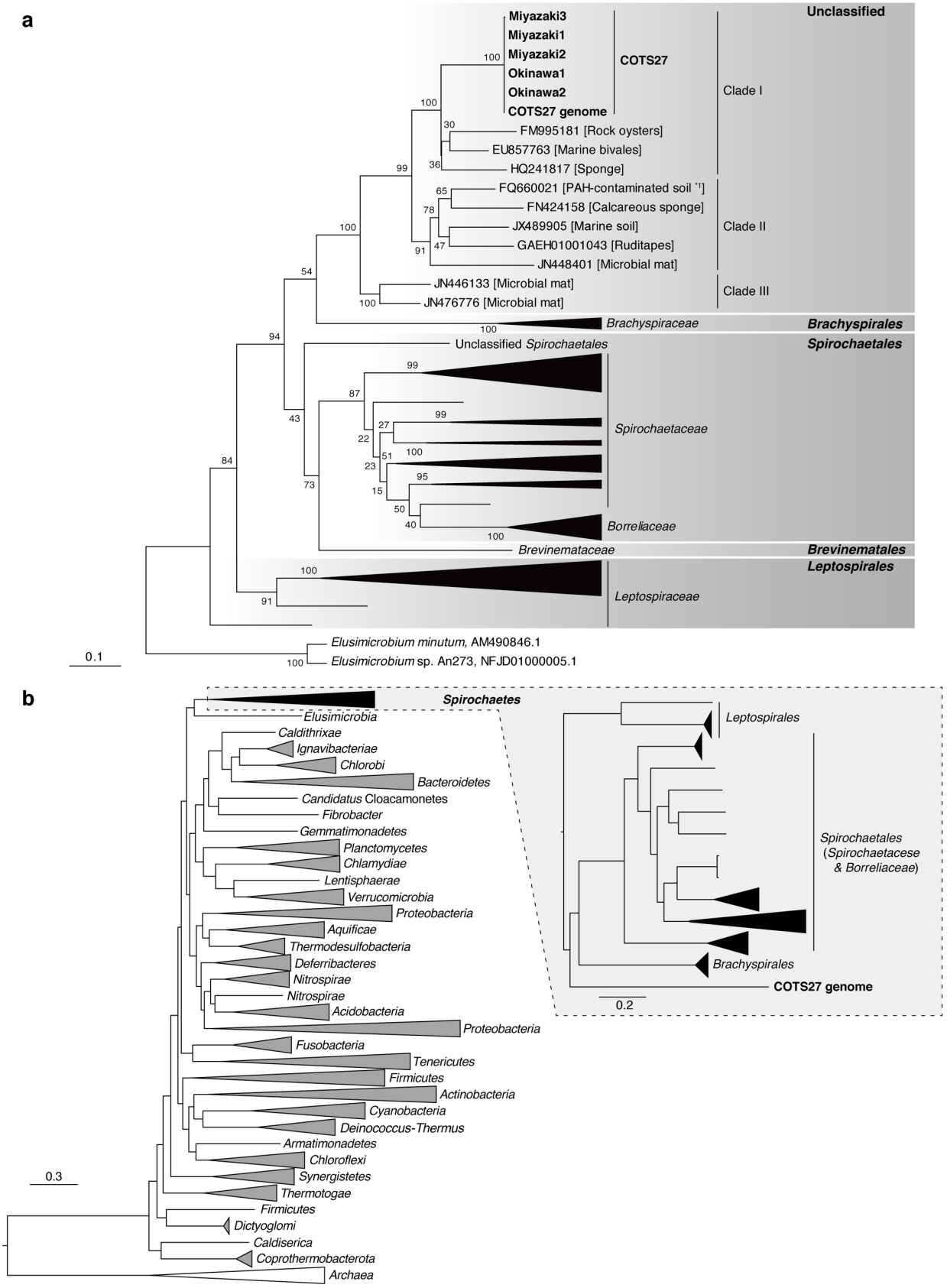

Fig. 3 The phylogenetic position of COTS27 in the phylum *Spirochaetes*.

Maximum likelihood (ML) trees were constructed based on the full-length 16S rRNA gene sequences (**a**) and the sequences of 43 conserved marker genes identified by CheckM (**b**). The bootstrap values in (**a**) were calculated by resampling 1,000 times. The scale bars indicate substitutions per site. *1: The gene with accession No. FQ660021.1 in panel (a) was obtained from a polycycle aromatic hydrocarbon (PAH)-contaminated soil sample in a mitigated wetland.

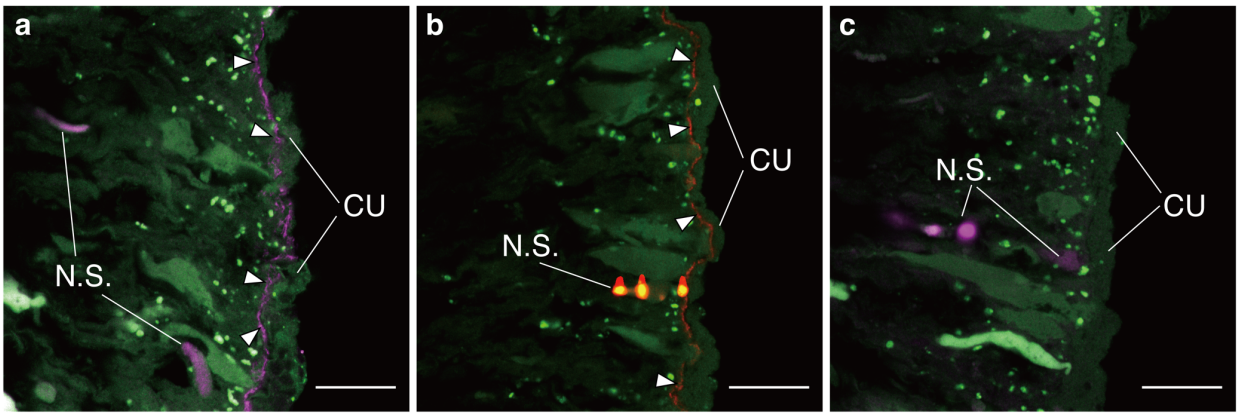

Fig. 4 FISH analysis of three serial sections of a COTS disc spine.

Each section was hybridized with the EUB338mix (**a,** purple; a general probe for bacteria), COTSsymb (**b**, red; a COTS27-specific probe), or Non338 (**c**, purple; a negative control to detect non-specific binding) probes. The probes were labelled with Cy3 in all panels and coloured with purple in panels **a** and **c** and with red in panel **b**. The green signals are tissue-derived autofluorescence. The arrowheads in panels **a** and **b** indicate layer-like signals from the general probe for bacteria (**a**) and the COTS27- specific probe (**b**). N.S. and CU indicate regions with non-specific binding and the outer cuticle complex, respectively. The scale bars represent 20 μm (**a-c**).

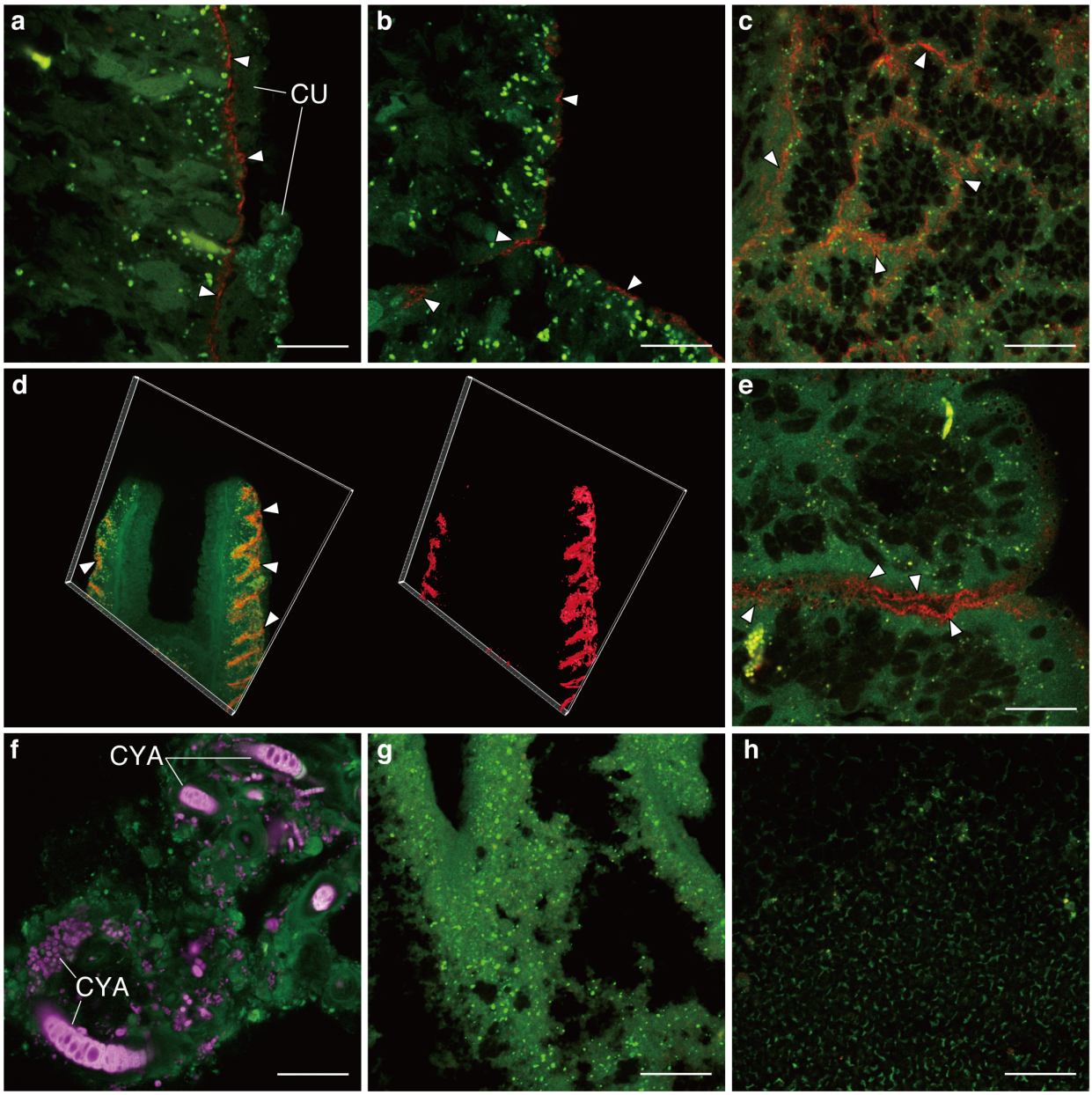

Fig. 5 Visualization of the COTS27 cells in different body parts of COTS using FISH.

COTS27 cells (red) residing in the subcuticular spaces of the body walls were detected with COTSsymb, a COTS27-specific probe, in the tips (**a**) and bases (**b**) of aboral spines on the discs and arms, respectively, dermal papula (**c**), pedicellariae on the aboral side (**d**; 3D image [left] and 3D rendering image [right]), and tube feet (**e**). Many non-COTS27 bacteria (purple) were detected in the pyloric stomachs (**f**) using the EUB338mix probe. No visible bacteria were detected in the pyloric caeca (**g**) and gonads (**h**) in this study (the images were obtained applying the COTS27-specific probe). The arrowheads indicate signals from COTS27. The green signals are tissue-derived autofluorescence. CU: outer cuticle complex; CYA: cyanobacteria-like cells. Scale bars (**a–c** and **e–h**) indicate 20 μm. The 3D image in panel (**d**) was taken with an original objective of x40.

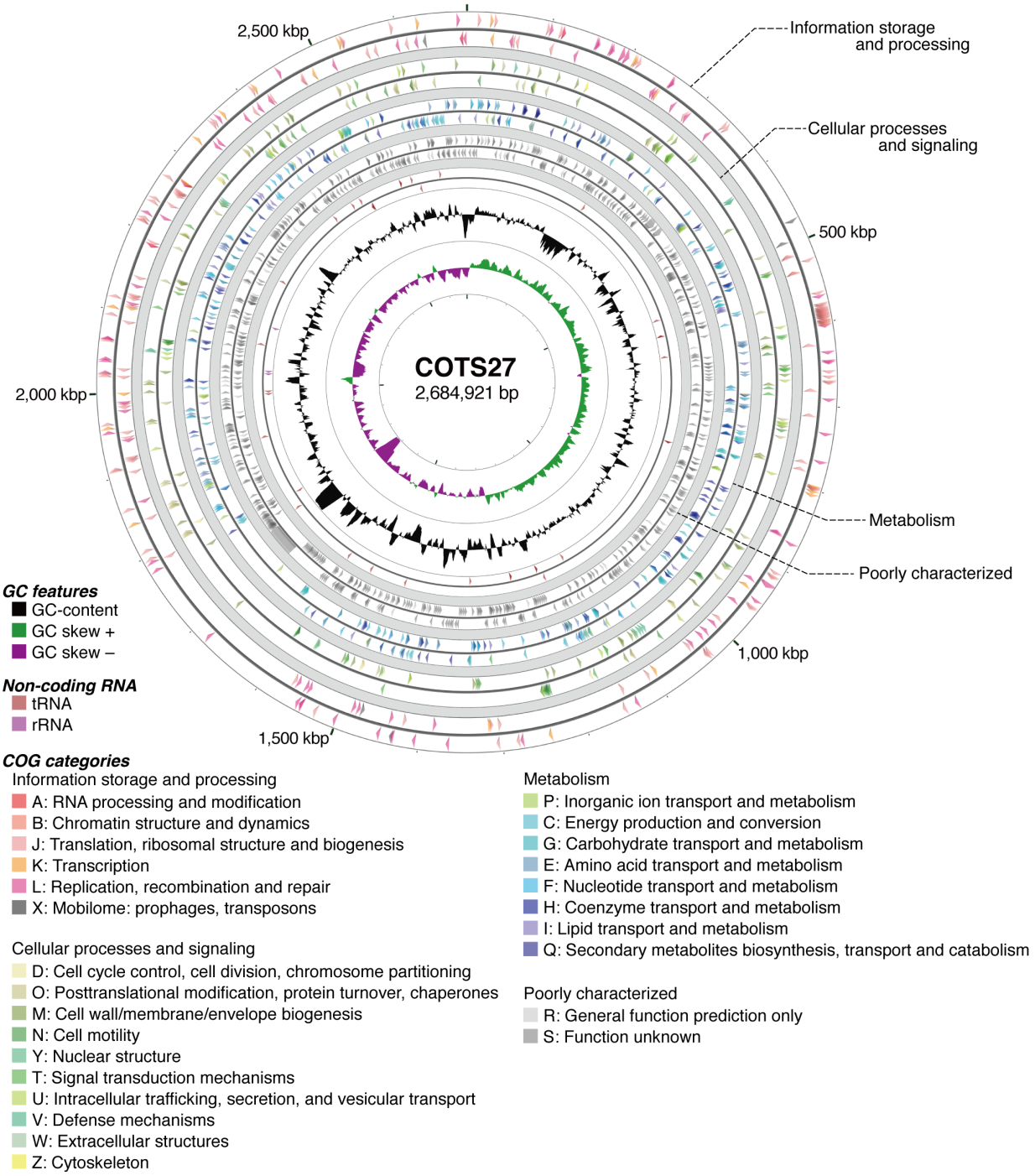

Fig. 6 Circular view of the COTS27 chromosome.

From the outside to the centre, each circle indicates forward strand CDSs; reverse strand CDSs; forward strand tRNA and rRNA genes; reverse strand tRNA and rRNA genes; GC-content; and GC skew. The CDSs were coloured according to the COG functional category of each CDS. The circular maps were created using CGView Server and the designations were then superposed manually.