A novel, tissue-associated and vertically transmitted bacterial symbiont in the coral *Pocillopora acuta*

Justin Maire[1,*], Sarah Jane Tsang Min Ching[1], Katarina Damjanovic[1,2], Hannah E. Epstein[2,3,#], Louise M. Judd[4], Linda L. Blackall[1], Madeleine J. H. van Oppen[1,2]

[1] School of BioSciences, The University of Melbourne, Parkville, 3010, VIC, Australia.

[2] Australian Institute of Marine Science, PMB No 3, Townsville, 4810, QLD, Australia.

[3] ARC Centre of Excellence for Coral Reef Studies, James Cook University, Townsville, 4811, QLD, Australia

[4] Doherty Applied Microbial Genomics, Department of Microbiology and Immunology, The University of Melbourne at the Peter Doherty Institute for Infection and Immunity, Parkville, 3010, VIC, Australia

[#] Present address: School of Life Sciences, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, United Kingdom

*Corresponding author: justin.maire@unimelb.edu.au

Abstract

Coral microhabitats are colonized by a myriad of microorganisms, including diverse bacteria which are essential for host functioning and survival. However, the location, transmission, and functions of individual bacterial species living inside the coral tissues remain poorly studied. Here, we show that a previously undescribed bacterial symbiont of the coral

Pocillopora acuta forms cell-associated microbial aggregates (CAMAs) within the mesenterial filaments. CAMAs were found in both adults and larval offspring, providing evidence of vertical transmission. *In situ* laser capture microdissection of CAMAs followed by 16S rRNA gene metabarcoding and shotgun metagenomics produced a near complete metagenome-assembled genome. We subsequently cultured the CAMA bacteria from *P. acuta* colonies, and sequenced and assembled their genomes. Phylogenetic analyses showed that the CAMA bacteria belong to an undescribed Endozoicomonadaceae genus and species, which we propose to name *Sororendozoicomonas aggregata* gen. nov sp. nov. Metabolic pathway reconstruction from its genome sequence suggests this species can synthesize most amino acids, several B vitamins, and antioxidants, which may be beneficial to its coral hosts. This study provides detailed insights into a new member of the widespread Endozoicomonadaceae family, thereby improving our understanding of coral holobiont functioning. Vertically transmitted, tissue-associated bacteria, such as *S. aggregata* may be key candidates for the development of microbiome manipulation approaches with long-term positive effects on the coral host.

Key words (3-10)

Endozoicomonas, endozoicomonadaceae, cnidarian, symbiosis, microbiome, genomics

Competing interests statement

The authors declare they have no competing interests.

45 **Main text**

46 Corals associate with diverse bacteria with essential functions, such as protection against
47 pathogens and nutrient cycling [1]. Among these bacteria, a small portion resides within the
48 coral tissues [2], sometimes forming cell-associated microbial aggregates (CAMAs) [3–6].
49 However, the identification and characterization of CAMA-forming bacteria in corals
50 remains in its infancy. The most common CAMA bacteria belong to the *Endozoicomonas*
51 genus [3, 4, 7], widespread coral symbionts generally considered to be mutualistic through
52 sulfur cycling and B vitamin synthesis [3, 7–10]. *Simkania* sp. (Chlamydiota) was recently
53 shown to also form CAMAs in *Pocillopora acuta* from Feather Reef (Great Barrier Reef
54 [GBR], Australia; Figure S1) [3], although its function remains unclear.

55 Here, we characterized CAMAs present in *P. acuta* colonies from Orpheus Island, also in the
56 GBR. *P. acuta* is an asexual brooder, producing fully formed, genetically identical larvae, and
57 a sexual broadcast spawner [11]. The microbiome of Orpheus Island colonies is dominated by
58 *Endozoicomonas* in early and adult life stages [6, 12], and CAMAs were detected in larvae,
59 indicating possible vertical transmission [6]. Using samples from these two studies (Figure
60 S2A-B), we verified by fluorescence *in situ* hybridization (FISH) that both adults (X7 colony
61 [12]) and larvae (OI2 and OI3 colonies [6]) possessed CAMAs (Figure 1A-E). In adult
62 polyps, CAMAs localized to the mesenterial filaments (Figure 1A-C). In larvae, CAMAs
63 were found in the gastrodermis (Figure 1D-E).

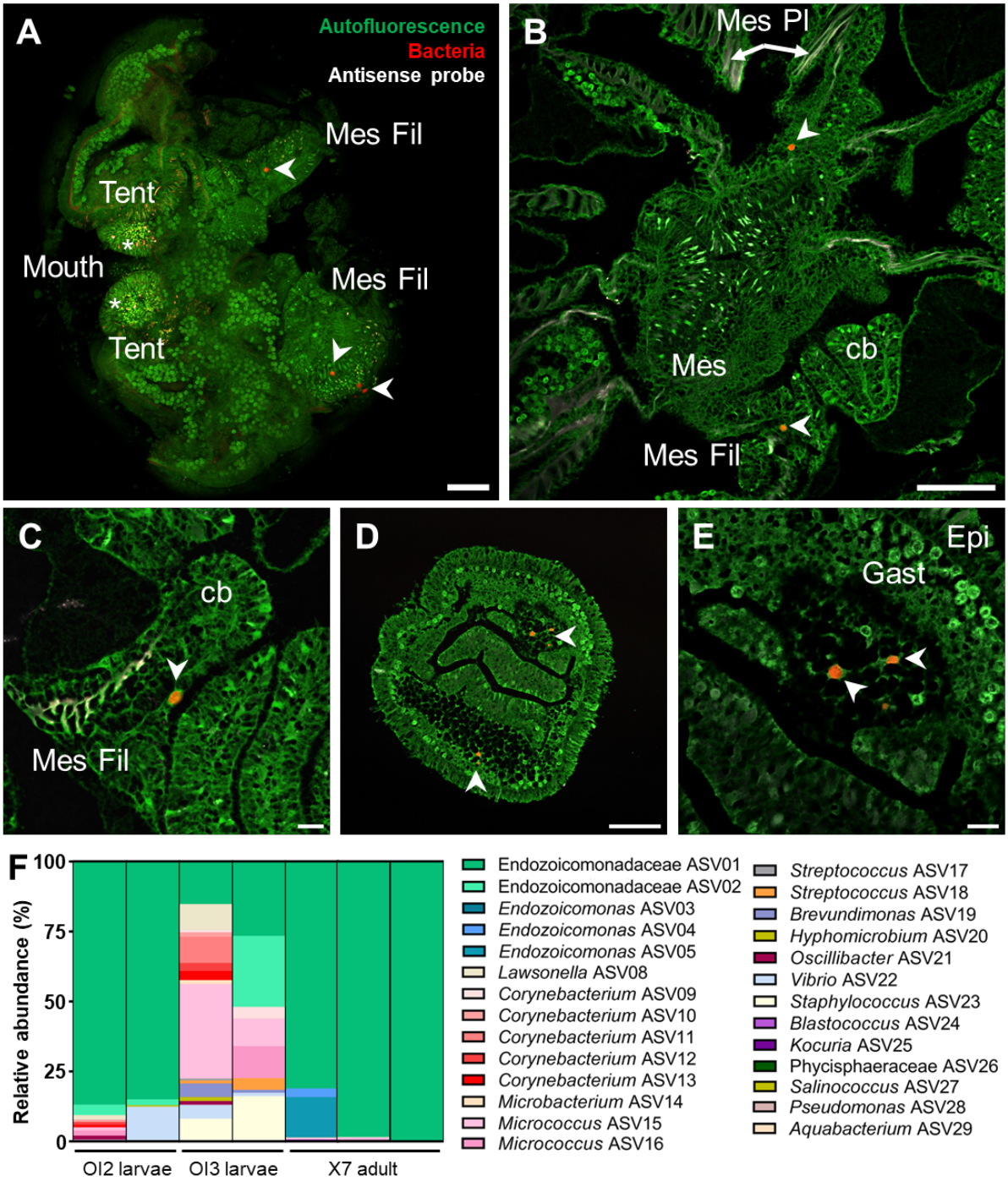

Figure 1: Location and identity of CAMAs in Orpheus Island *Pocillopora acuta*. A: CAMAs in an adult polyp (X7 colony) by whole-mount FISH. B-E: FISH images showing the location of CAMAs in sectioned adult polyps (B, C; X7 colony) and sectioned larvae (D, E; OI2 colony). White arrowheads point at CAMAs. Green: autofluorescence; red: EUB338-mix probe (all bacteria); white: non-EUB probe (negative control). Note the non-specific

binding in tentacles (red and white signal overlapping, white asterisks), often localized in nematocysts. Tent: tentacle; Mes Fil: mesenterial filaments; Mes: mesenteries; Mes Pl: mesogleal plates; cb: cnidoglandular band; Gast: gastrodermis; Epi: epidermis. Scale bars: 100 µm for A, B, D; 20 µm for C and E. F: Relative abundance of bacterial ASVs in CAMAs of larvae (OI2 and OI3 colonies) and adult branches (X7 colony) isolated by LCM. Each bar represents one replicate (either three pooled larvae or one adult branch). Detailed composition is available in Table S2.

To taxonomically identify the bacteria present in the CAMAs, we excised CAMAs using laser capture microdissection (LCM) , followed by 16S rRNA gene metabarcoding (Figure S3, Table S1). Two ASVs assigned to *Kistimonas* (Endozoicomonadaceae ASVs 01 and 02) made up most of the reads in OI2 larvae and X7 branches (Figure 1F, Table S2). While the samples were collected at the same site, they originated from different host colonies and were collected several months apart, highlighting the temporal stability of this association among individuals and through ontogeny. *Kistimonas* is part of the Endozoicomonadaceae family along with *Endozoicomonas*. Because of taxonomic discrepancies discussed below, we refer to these two ASVs as Endozoicomonadaceae ASVs 01 and 02. The two ASVs were also present in OI3 larvae at a lower relative abundance. In addition, three *Endozoicomonas* ASVs were found in X7 and were identical to *Endozoicomonas* present in *P. acuta* from Feather Reef [3]. FISH using an Endozoicomonadaceae probe provided additional support for this taxonomy in both adult and larval CAMAs (Figure S4).

While *Kistimonas* was not reported in the original studies [6, 12], a re-analysis of both previous studies reassigned three *Endozoicomonas* ASVs to *Kistimonas* (Figure S5, Table S3), two of which were 100% identical to Endozoicomonadaceae ASVs 01 and 02 found in

our excised CAMA samples. Virtually no *Endozoicomonas* strains (only one ASV at 0.02-0.04% relative abundance) were detected in whole larvae or recruits in our re-analysis of the earlier datasets (Figure S5), suggesting that *Endozoicomonas* are acquired horizontally in Orpheus Island *P. acuta*, as observed for Feather Reef *P. acuta* [3]. Conversely, Endozoicomonadaceae ASVs 01 and 02 were present at all life stages (*i.e.* larvae, recruits and adults of the same colonies, Figure S5) and formed CAMAs, confirming their vertical transmission. Whether whole CAMAs or individual bacteria are transmitted from parent colony to asexually produced larvae remains unknown. Strong evidence for vertical transmission of bacteria in corals has only been reported twice before [13, 14].

Shotgun sequencing of the X7 CAMA samples yielded one metagenome-assembled genome (MAG) (Pac_X7, Table S4), the 16S rRNA gene sequence of which was identical to Endozoicomonadaceae ASV01. An additional *P. acuta* colony (P3, Figure S2C) was collected from Orpheus Island in June 2023, from which we cultured ten bacterial isolates with 16S rRNA gene sequences identical to Pac_X7, one of which was whole-genome sequenced (Pac_P3-11-1, Table S4). Both assemblies had 99.8% Average Nucleotide Identity (ANI) and 99.6% Average Amino acid Identity (AAI) (Figure S6), confirming that the Pac_P3-11-1 isolate is indeed a CAMA bacterium. Phylogenetic analysis of 120 marker genes placed both genomes separately from all described Endozoicomonadaceae genera (Figure 2A, Table S5), along with Endozoicomonadaceae SCSIO12664 (99.6% 16S rRNA gene sequence identity, 99.2% ANI, 98.8% AAI), isolated from *Pocillopora damicornis* in the South China Sea [15]. This was supported by 16S rRNA gene phylogeny (Figure S7). Together with low AAI (highest score of 60.1% with *E. elysicola*, Figure S6) and low 16S rRNA gene identity (highest score of 93.97% with *E. gorgoniicola*), these data support the placement of Pac_P3-11-1 and Endozoicomonadaceae SCSIO12664 in a new genus and species, which we propose to name *Sororendozoicomonas aggregata*, gen. nov., sp. nov. This likely explains the

discrepancies observed with metabarcoding-based taxonomy, as taxonomy would be assigned to a closely related taxon rather than an undescribed genus.

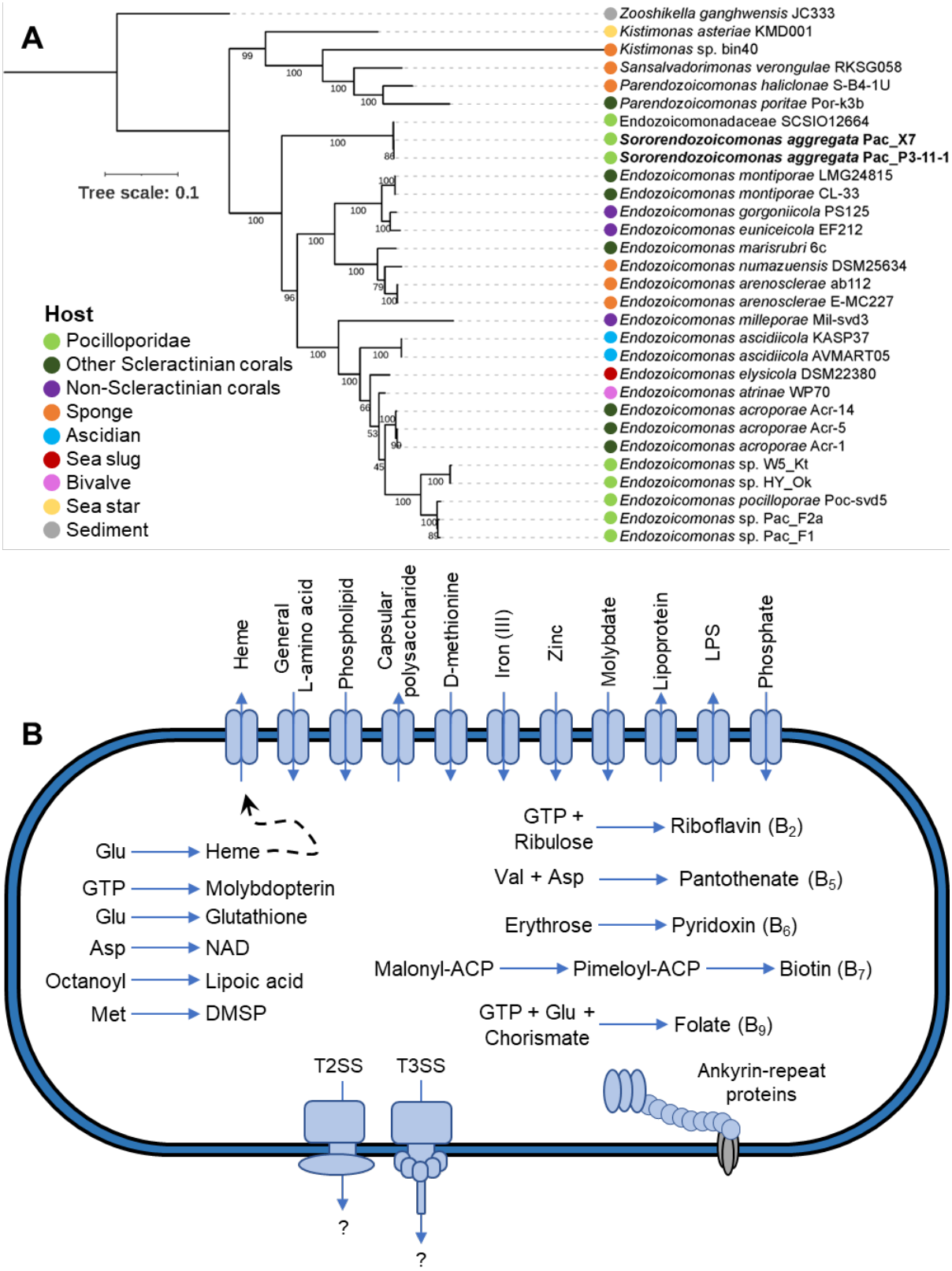

Figure 2: Taxonomy and functional potential of *Sororendozoicomonas aggregata*, CAMA member of Orpheus Island *Pocillopora acuta*. A: Maximum likelihood phylogeny of Endozoicomonadaceae based on 120 marker genes and 28 Endozoicomonadaceae genomes in addition to the two genomes recovered in this study. *Zooshikella* was chosen as the outgroup. Bootstrap support values based on 1000 replications are provided. Additional data on the reference genomes are available in Table S5. B: Overview of the metabolic potential from the genome sequence of Pac_P3-11-1 recovered from *P. acuta* CAMAs. The pathways represented here are at least 80% complete (Fig. S7). Dashed arrows represent hypotheticals. T2SS: type II secretion system; T3SS: type III secretion system; DMSP: dimethylsulfoniopropionate; LPS: lipopolysaccharide; NAD: nicotinamide adenine dinucleotide.

Genome annotation (Figures 2B and S8, Table S6) revealed Pac_P3-11-1 encodes several genes indicative of host-symbiont interactions, including type II and type III secretion systems, 35 eukaryotic-like proteins (Table S7), and three secondary metabolites with putative antimicrobial activity (Table S8). This gene repertoire is similar to that of other coral-associated Endozoicomonadaceae [3, 4, 8–10] and may regulate host colonization, aggregation, and vertical transmission. Additionally, we recovered pathways for the biosynthesis of all amino acids (except phenylalanine, tyrosine, and arginine), as well as riboflavin (vitamin B_2), pantothenate (B_5), pyridoxin (B_6), biotin (B_7), and folate (B_9), which could assist with coral metabolism [3, 4, 9, 10, 16]. Finally, Pac_P3-11-1 shows potential for the scavenging of reactive oxygen species, which are believed to drive coral bleaching when present in excess [17], as it possesses pathways for the synthesis of the antioxidants heme, lipoic acid, and glutathione, and the *dsyB* gene, essential for the synthesis of dimethylsulfonioproprionate [18].

147 In conclusion, we described and cultured the first coral tissue-associated, vertically
148 transmitted member of the Endozoicomonadaceae family. *Sororendozoicomonas* was not
149 found in other GBR *P. acuta* populations that harbor *Endozoicomonas* [3], suggesting
150 geographical location may impact Endozoicomonadaceae associations. Additionally,
151 *Endozoicomonas* CAMAs were previously found in pocilloporid tentacles [3, 4], contrasting
152 with *Sororendozoicomonas*' location in the mesenterial filaments.

153 As climate change increasingly triggers coral bleaching and mortality, novel methods are
154 urgently needed to increase coral climate resilience, including the manipulation of coral-
155 associated microorganisms. Endozoicomonadaceae are great candidates for coral microbiome
156 manipulation approaches [19, 20], due to their prevalence, tissue location (*i.e.* more stably
157 associated than bacteria residing in the mucus or skeleton), and putative benefits. The vertical
158 transmission and cultivability of *S. aggregata* make it an ideal candidate, ensuring its stability
159 within the coral holobiont across generations and thereby the long-term sustainability of
160 conservation/restoration approaches based on microbiome manipulation.

161

162 **Description of *Sororendozoicomonas* gen. nov.**

163 *Sororendozoicomonas* (So.ror.en.do.zo.i.co.mo'nas. L. fem. n. soror, sister; *Endozoicomonas*,
164 taxonomic name of a bacterial genus; *Sororendozoicomonas* referring to the close
165 relationship to the bacterial genus *Endozoicomonas*). The name was registered through
166 SeqCode [21].

167 The genus *Sororendozoicomonas* represents a distinct lineage within the
168 Endozoicomonadaceae family, sister to *Endozoicomonas*, supported by concatenated marker
169 gene phylogenies (Figures 2A) as well as 16S rRNA gene phylogeny (Figure S7) in the
170 present study. *Sororendozoicomonas* show similar functional and metabolic potential to

171 *Endozoiocomonas* species (Figure 2B). The new genus contains two strains isolated from
172 *Pocillopora* corals (Pac-P3-11-1, from the present study, and SCSIO12664 [15]) and
173 potentially additional strains isolated from Red Sea *Stylophora pistillata* [22] (only 16S
174 rRNA gene sequences are available - KC669256.1, KC669230.1, KC669217.1,
175 KC669131.1).

176

177 **Description of *Sororendozoicomonas aggregata* sp. nov.**

178 *Sororendozoicomonas aggregata* (ag.gre.ga'ta. L. fem. part. adj. *aggregata,* joined together,
179 referring to the ability to form aggregates within coral tissues). The name was registered
180 through SeqCode [21]. The new species contains two strains isolated from *Pocillopora* corals
181 (Pac-P3-11-1, from the present study, and SCSIO12664 [15]).

182 Type strain: Pac_P3-11-1, isolated from tissues of the coral *Pocillopora acuta*, collected from
183 Orpheus Island, Australia (18°36'16.26" S, 146°29'24.27" E). Taxonomic assignment based
184 on concatenated marker gene phylogenies (Figure 2A) as well as 16S rRNA gene phylogeny
185 (Figure S7). Grows on Marine Agar 2216 (Difco) between 23°C and 28°C, forming circular,
186 white, raised, sticky colonies with entire margins.

187

188 **Acknowledgements**

189 This research was supported by the Australian Research Council (FL180100036 to MJHvO;
190 DP160101468 to MJHvO and LLB), the Native Australian Animals Trust (to JM), Paul G.
191 Allen Philanthropies (to MJHvO), and AIMS@JCU, the joint venture between James Cook
192 University and the Australian Institute of Marine Science, as part of H.E.E's PhD research.
193 H.E.E acknowledges receipt of an AIMS@JCU Postgraduate Scholarship. We thank: Corinne

E. Allen for the collection of coral colonies from Orpheus Island (permit G22/46479.1); the Biosciences Microscopy Unit (University of Melbourne) and the Biological Optical Microscopy Platform (University of Melbourne) for the use of their confocal microscopes and laser capture dissection, and particularly Dr Gabriela Segal, Dr Shane Cheung and Dr Sam Mills for their valuable assistance; the Melbourne Histology Platform (University of Melbourne), Laura Leone, and Lisa Foster for their assistance with sample sectioning; Dr Kshitij Tandon and Dr Gayle K. Philip for assistance with genome analyses; the Melbourne Research Cloud, the University of Melbourne's Research Computing Services and the Petascale Campus Initiative for providing the high-performance computing instances needed for this work.

Data availability statement

Raw data are available under NCBI BioProject IDs PRJNA891898 (larvae CAMAs, MiSeq raw data) and PRJNA974967 (adult CAMAs, MiSeq and NovaSeq raw data, Pac_X7 MAG assembly; Pac_P3-11-1 NextSeq raw reads and assembly). The complete 16S rRNA gene sequence of Pac_P3-11-1 is available under the accession number OR505856.

Author contribution

Conceptualization: JM, MvO, LB; Investigation: JM, SJTMC, KD, HE, LJ; Methodology: JM; Resources: KD, HE; Formal analysis: JM; Visualization: JM; Funding acquisition: MvO, LB, JM; Writing - Original Draft Preparation: JM, MvO. All authors reviewed and edited the final manuscript.

Material and methods

Coral larvae collection (OI2 and OI3 colonies)

For OI2 and OI3, parent colonies were collected in 2017 from Orpheus Island (Little Pioneer Bay; 18°36'3.6"S 146°29'20.4"E), in the central Great Barrier Reef in Australia (Figure S1 and S2A), and maintained and sampled as part of a previous experiment [6]. Briefly, before planulation, colonies were maintained in individual acrylic aquaria that received indirect natural sunlight and 0.4 µm-filtered seawater. A filter was fitted at each outlet of the acrylic tank to collect released planulae. Larvae were washed with 0.22-µm filtered seawater (FSW), transferred from the filter into a microcentrifuge tube using a sterile pipette tip, and as much water as possible was removed from the tube without disturbing the larvae. Larvae were fixed for 10 hrs at 4°C in 4% paraformaldehyde (PFA) prepared in FSW, rinsed twice in FSW, and stored in 50% ethanol-PBS at -20°C. Around 10 larvae per parent colony were fixed.

Adult coral collection (X7 and P3 colonies)

For the X7 colony (Figure S2B), coral fragments were sampled from the field in 2016 from the same site at Orpheus Island as part of a previous experiment (Figure S1) [12]. From this study, only samples collected in November 2016 were processed here. Four small coral fragments (2-5 cm in length) were snapped off the colonies with forceps, fixed for 24 hrs in 4% PFA in FSW, rinsed twice in FSW, and stored in 50% ethanol-PBS at -20°C. Following fixation, coral branches were decalcified in EDTA 10%. EDTA was renewed every two days, and samples were kept at 4°C on a rotating wheel, until there was no skeleton left (around two weeks). Samples were then rinsed in PBS 1X, and stored at 4°C in PBS 1X.

For the P3 colony (Figure S2C), coral colonies were collected at 1-4 m depth with a chisel and hammer from Orpheus Island (Little Pioneer Bay, 18°36'16.26" S, 146°29'24.27" E,

241 Collection permit G22/46479.1) in June 2023. Four colonies were brought back to the
242 Australian Institute of Marine Science for overnight holding before being shipped to the
243 University of Melbourne the next day. On arrival, the corals were placed in 130 L
244 recirculating system tanks and kept there overnight for 10 h at 25 °C in 35 p.p.t. reconstituted
245 sea water (Red Sea Salt™, R11065, Red Sea, USA).

246

247 **Fluorescence *in situ* Hybridization (sections)**

248 Sample processing, embedding, and sectioning (3-µm thickness) were performed by the
249 Melbourne Histology Platform (University of Melbourne) as previously described [3].
250 Fluorescence *in situ* hybridization (FISH) was then performed as previously described [6],
251 but final probe concentration during hybridization was 5 ng/µL. See Table S9 for probe
252 sequences, fluorophores, and formamide concentrations. Slides were mounted in CitiFluor™
253 CFM3 mounting medium (proSciTech, Australia), covered with a coverslip and sealed with
254 clear nail polish. Slides were kept at 4°C until observation.

255

256 **Fluorescence *in situ* Hybridization (whole-mount)**

257 Single polyps of the X7 colony were dissected in PBS 1X using Dumont tweezers under a
258 dissecting microscope. Samples were cleared of autofluorescence by incubating polyps in
259 methanol-PBS 1X (50:50 v/v) for 10 min, methanol-PBS 1X (75:25 v/v) for 10 min,
260 methanol-PBS 1X (90:10 v/v) for 10 min, 100% methanol for 10 min, methanol-PBS 1X +
261 Triton X-100 0.2% (90:10 v/v) for 10 min, methanol-PBS 1X + Triton X-100 0.2% (75:25
262 v/v) for 10 min, methanol-PBS 1X + Triton X-100 0.2% (50:50 v/v) for 10 min, PBS 1X +
263 0.2% Triton X-100 for 10 min, and stored at 4°C in PBS 1X. FISH was performed on whole
264 polyps as previously described [3], with the probes shown in Table S9**.** Before observation,

265 single polyps were cut in half to expose the gastric cavity and mesenteries, and deposited
266 onto 8-well coverslip-bottom slides (ibidi, USA) with a drop of milliQ water to avoid drying.

267

268 **Confocal Laser Scanning Microscopy**

269 Slides were observed on a Nikon A1R confocal laser scanning microscope (Nikon, Japan)
270 with the NIS325 Element software. Virtual band mode was used to acquire variable emission
271 bandwidth to tailor acquisition for specific fluorophores. The fluorophores Atto550 were
272 excited using the 561 nm laser line, Atto647 using the 640 nm laser line, and the coral
273 autofluorescence using the 488 nm laser line with a detection range of 570-620 nm for
274 Atto550, 660-710 nm for Atto647, and 500-550 nm for coral autofluorescence. For three-
275 dimensional reconstructions of Z-stacks (for whole polyps), sections were acquired using Z
276 steps of 3.2 µM with the 10X objective. Nd2 files were processed using ImageJ. Z-stacks
277 were projected in two-dimensional images using the 'Max Intensity' projection type. Linear
278 adjustments of brightness and contrast were performed when necessary and applied to the
279 entire image and to each channel independently. Channels were then given artificial colors
280 (see figure legends) and merged.

281

282 **Laser Capture Microdissection of CAMAs and DNA extraction**

283 Laser capture microdissection (LCM) of CAMAs in OI2, OI3, and X7 samples was
284 performed as previously described [3]. For OI2 and OI3 larvae, three replicates each
285 containing three larvae were processed. For each replicate, ten slides each containing eight
286 sections were processed (~80 3-µm sections per replicate in total). For X7 branches, three
287 replicates each containing one branch were processed. For each replicate, 15 slides each
288 containing four sections were processed (~60 3-µm sections per replicate in total). Tissue

289 areas without CAMAs were also separately captured as a negative control. DNA extraction
290 was then conducted as previously described [3], using the Arcturus® PicoPure® DNA
291 Extraction Kit (Applied Biosystems, USA). Three caps containing no captured tissue, but that
292 were open in the LCM facility to capture air contamination, were also included as extraction
293 blanks.

294

295 **16S rRNA gene metabarcoding**

296 Hypervariable regions V5-V6 of the 16S rRNA genes were amplified using the primer set
297 784F (5′ GTGACCTATGAACTCAGGAGTCAGGATTAGATACCCTGGTA 3′) and 1061R
298 (5′ CTGAGACTTGCACATCGCAGCCRRCACGAGCTGACGAC 3′). Adapters were
299 attached to the primers and are underlined. Bacterial 16S rRNA genes were PCR-amplified
300 on a SimpliAmp Thermal Cycler (Applied Biosystems, ThermoFisher Scientific). Each
301 reaction contained 1 µL of DNA template, 1.5 µL of forward primer (10 µM stock), 1.5 µL of
302 reverse primer (10 µM stock), 7.5 µL of 2x QIAGEN Multiplex PCR Master Mix (Qiagen,
303 Germany) and 3.5 µL of nuclease-free water (Thermofisher), with a total volume of 15 µL per
304 reaction. Three triplicate PCRs were conducted for each sample and three no-template PCRs
305 were conducted as negative controls. PCR conditions for the 16S rRNA genes were as
306 follows: initial denaturation at 95°C for 3 min, then 18 cycles of: denaturation at 95°C for 15
307 s, annealing at 55°C for 30 s, and extension at 72°C for 30 s; with a final extension at 72°C for
308 7 minutes. Samples were then held at 4°C. Following PCR, triplicates were pooled, resulting
309 in 45 µL per sample. Metabarcoding library preparation was conducted as previously
310 described [23] and sequencing was performed at the Walter and Eliza Hall Institute (WEHI)
311 in Melbourne, Australia on one MiSeq V3 system (Illumina) with 2x300bp paired-end reads.
312

313 **16S rRNA gene metabarcoding analysis**

314 QIIME2 v 2021.8 [24] was used for processing 16S rRNA gene sequences. The plugin demux
315 [24] was used to create an interactive plot to visualize the data and assess the quality, for
316 demultiplexing and quality filtering of raw sequences. The plugin cutadapt [25] was used to
317 remove the primers and MiSeq adapters. DADA2 [26] was used for denoising and chimera
318 checking, trimming, dereplication, generation of a feature table, joining of paired-end reads,
319 correcting sequencing errors, and removing low quality reads (Q-score < 30). Summary
320 statistics were obtained using the feature-table to ensure processing was successful.
321 Taxonomy was assigned by training a naive Bayes classifier with the feature-classifier plugin
322 [24], based on a 99% similarity to the V5-V6 region of the 16S rRNA gene in the SILVA 138
323 database to match the 784F/1061R primer pair used [27]. Mitochondria and chloroplast reads
324 were filtered out. Analyses were performed using Rstudio version 2022.02.2 and the phyloseq
325 package [28]. Metadata file, taxonomy table, phylogenetic tree and ASV table were imported
326 into R to create a phyloseq object. Contaminant ASVs, arising from kit reagents and sample
327 manipulation, were identified using the package decontam [29]. The function
328 'isNotContaminant' was used as it is more stringent and more adequate for low-biomass
329 samples. One replicate for each OI2 and OI3 were removed because of heavy contamination.
330 It is worth noting that bacteria belonging to the *Brachybacterium* genus accounted for more
331 than 90% of the contamination (Table S1), which were contaminants of the Qiagen PCR kit
332 previously observed [3]. For the re-analysis of previously acquired data, raw reads were
333 downloaded from SRA (SRP150755 for X7, and SRP187380 for OI2 and OI3). Data was re-
334 analyzed as described above. Plots were generated using GraphdPad Prism 9.

335

336 **Genome amplification of LCM samples and shotgun sequencing**

337 To reach sufficient quantities for shotgun sequencing, DNA samples obtained through LCM
338 were amplified as previously described [3], using a SeqPlex DNA Amplification Kit (Sigma-
339 Aldrich, USA), which is specifically designed for low-quantity, fragmented DNA. The same
340 DNA samples that were used for 16S rRNA gene metabarcoding were used and, for each
341 sample, 1 µL (OI2 and OI3) or 3 µL (X7) of each of the three replicates were pooled before
342 amplification. Only the X7 sample yielded sufficient DNA for sequencing. Samples were
343 then sent to the Australian Genomic Research Facility (Melbourne, Australia) for sequencing.
344 Library preparation was performed with an IDT xGen cfDNA & FFPE DNA Library Prep kit
345 (Integrated DNA Technologies, USA) and samples were sequenced on a NovaSeq 6000 S4
346 2×150 bp Flowcell Illumina platform.

347

348 **Metagenomic data analysis**

349 Quality control was performed using FastQC v0.11.9 [30]. Adapters, low-quality sequences
350 (phred score < 30), as well as the first and last 10 bp were trimmed using Trim Galore v0.6.2
351 [31]. High-quality reads were mapped against a *P. acuta* draft genome [32] using Bowtie2
352 v2.4.2 with default parameters [33] to remove host-related reads from the data. Mapped reads
353 were then removed using Samtools v1.11 [34]. Only paired-end host-removed reads were
354 used for metagenome assembly. Metagenome assembly was carried out using MEGAHIT
355 v1.2.9 [35] with a minimum contig length of 1000 bp and the following k-mers: 21, 33, 55,
356 77, 99. Contigs from kmer 99 were used for all downstream processing. Contig taxonomy
357 was assessed using CAT/BAT v5.2.3 [36], and contigs assigned to Eukarya, Archea, or Virus
358 were removed. Assembled contigs were binned using the binning (with metabat2, concoct
359 and maxbin2 tools) and bin_refinement (with >70% completeness and <10% contamination
360 as cut-off parameters) modules of MetaWRAP v1.3.2 [38]. Bin quality and taxonomy were

361 assessed using CheckM v1.2.2 [39] and GTDB-Tk v2.3.0 [40], respectively. These bins were
362 reassembled to further improve the contiguity and bin completeness and contamination stats
363 using the reassemble_bins module implemented in MetaWRAP v1.3.2. The final taxonomy of
364 individual contigs on a per bin level was assessed using CAT/BAT v5.2.3 [36], and any contig
365 belonging to a different phylum than the taxonomy assigned by GTDB-Tk was manually
366 removed. Bin coverage was obtained using CoverM v0.6.1
367 (https://github.com/wwood/CoverM) using the "genome" option. The full sequence of the
368 16S rRNA gene (1569 bp) was obtained using Barrnap v0.9 [41].

369

370 **Pure culturing of bacteria from *Pocillopora acuta***

371 Using clean bone cutters, fragments between 3 and 7 cm in length were sampled from the P3
372 colony, rinsed in FSW, and placed in sterile zip-lock bags containing FSW. Using a water
373 flosser (Waterpik® Water Flosser) fitted with a clean nozzle, coral tissue was blasted off the
374 skeleton. The resulting slurry was collected in a sterile zip-lock bag and transferred to 50 mL
375 Falcon tubes. The tubes were centrifuged at 3750 × *g* for 10 min (Allegra X-12R, Beckman
376 Coulter, USA) to obtain coral tissue pellets. The pellets were resuspended in 1 mL FSW and
377 transferred to a single 1.5 mL tube per colony and centrifuged at 5000 × *g* for 10 min to
378 remove any residual mucus from the sample. The tissue pellets were resuspended in 1 ml of
379 FSW using a tissue lyser (Tissue-Lyser II, Qiagen, Australia) and the resuspension was
380 homogenized using sterile glass homogenizers for 30 s. The resulting tissue homogenates
381 were each serially diluted from 10^{-1} to 10^{-7} and plated in triplicates onto Marine Agar 2216
382 (MA, BD Difco). Following one week of incubation at 23 °C, individual bacterial colonies
383 were selected and streaked onto fresh MA plates. Purified bacterial isolates were obtained

384 after a further 2 rounds of sub-culturing onto fresh MA plates following a weekly incubation
385 period at 25 °C and 26 °C, respectively.

386

387 **16S rRNA gene sequencing of cultured bacteria**

388 Individual, pure, freshly grown bacterial colonies were suspended in 20 µL sterile Milli-Q®
389 water, incubated for 10 min at 95°C then used as templates in colony PCRs. PCR
390 amplification of the bacterial 16S rRNA gene was with primers 27F and 1492R [42]. The
391 PCR was performed with 20 µl Mango Mix™ (Bioline, UK), 0.25 µM of each primer and
392 2 µl of DNA template in a final volume of 40 µl with nuclease free water (Ambion, Thermo
393 Fisher Scientific Inc., TX, USA). The thermal cycling protocol was as follows: 94°C for
394 5 min; 30 cycles of 94°C for 1 min, 50°C for 45 s and 72°C for 90 s; and a final extension of
395 10 min at 72°C. Amplicons were purified and sequenced on an ABI sequencing instrument by
396 Macrogen (Seoul, South Korea). Trimmed high-quality read data from each isolate were used
397 for presumptive identification by querying the 16S rRNA gene sequences via Blastn and
398 compared to other 16S rRNA gene sequences obtained in this study.

399

400 **DNA extraction and whole-genome sequencing**

401 The P3-11-1 isolate was selected for whole-genome sequencing. A single colony was picked
402 with an inoculation loop and DNA extraction was performed on the QIAsymphony using the
403 DSP Virus/Pathogen Mini Kit (Qiagen). Library preparation performed using Nextera XT
404 (Illumina Inc.) according to manufacturer's instructions. Whole-genome sequencing was
405 performed on NextSeq 500/550 with a 150bp PE kit.

406

407 **Genome assembly, taxonomy, and annotation**

408 Raw reads were trimmed and quality-filtered using Trimmomatic v0.36 [43]
409 (HEADCROP:10 LEADING:5 TRAILING:5 SLIDINGWINDOW:4:28 MINLEN:30). The
410 quality of raw reads before and after trimming was checked with FASTQC v0.11.9 [30].
411 Trimmed and quality-filtered reads were de novo assembled into contigs using SPAdes
412 v3.15.5 [44] with 21, 33, 55, 77 and 99 k-mers, and the option "--careful" was applied to
413 minimise the number of mismatches and short indels. All contigs with a length of <1000 bp
414 were removed using BBMap v38.96 [45]. Subsequently, the levels of completeness and
415 contamination of assembled genomes were assessed using CheckM v1.2.2 using the
416 "lineage_wf" workflow [39]. Coverage was obtained using CoverM v0.6.1
417 (https://github.com/wwood/CoverM) using the "genome" option.

418 Taxonomic assignment of all assembled genomes was carried out using GTDB-Tk v2.3.0
419 [40] using the "classify_wf" workflow. GTDB-Tk assigned the taxonomy of the genomes
420 based on 120 bacterial marker genes. Based on the GTDB-Tk alignment of the two genomes
421 from this study and 28 additional genomes (Table S5), a phylogenetic tree was built in IQ-
422 Tree v2.2.2.3 [46] using the best model Q.insect+F+R4, selected by ModelFinder wrapped in
423 IQ-tree [47], and 1000 ultrafast bootstrap replicates [48]. The tree was visualized in iTOL v6
424 [49]. *Zookishella* was chosen as an outgroup. Average Nucleotide Identities (ANI) and
425 Average Amino acid Identities (AAI) were calculated using a genome-based matrix calculator
426 [50] and plotted in R using the pheatmap package [51]. Gene prediction was performed in
427 Bakta v1.7.0 [52], KEGG-mapper Reconstruct [53], and InterProScan v5.55 with Pfam
428 domain annotations [54]. Pfam annotations were used to look for eukaryotic-like proteins
429 (ankyrin-repeat domains, WD40 domains, tetratricopeptide repeat) and the *dsyB* gene
430 (PF00891). Secondary metabolites were predicted using antiSMASH v7.0.0 [55].

431 The 16S rRNA gene phylogenetic tree of the Endozoicomonadaceae family was constructed
432 using 38 published, full-length 16S rRNA gene sequences from the Endozoicomonadaceae
433 family. The full-length 16S rRNA gene (1572 bp) obtained from the P3-11-1 assembly was
434 used for the phylogenetic tree construction. A MAFFT alignment was created using Geneious
435 Prime v2019.1.3. The full alignment was stripped of columns containing 99% or more gaps.
436 This alignment was used to generate a maximum likelihood phylogenetic tree with 1000
437 ultrafast bootstraps using IQ-Tree v2.2.2.3 [46] with the best model TIM3+F+I+G4, selected
438 by ModelFinder wrapped in IQ-tree [47]. Sequences belonging to the *Zookishella* genus were
439 chosen as an outgroup.

440

441 **References**

442 1. Bourne DG, Morrow KM, Webster NS. Insights into the coral microbiome: Underpinning the
443 health and resilience of reef ecosystems. *Annual Review of Microbiology* 2016; **70**: 317–340.
444 2. Maire J, Blackall LL, van Oppen MJH. Intracellular Bacterial Symbionts in Corals: Challenges and
445 Future Directions. *Microorganisms* 2021; **9**: 2209.
446 3. Maire J, Tandon K, Collingro A, van de Meene A, Damjanovic K, Gotze CR, et al. Colocalization
447 and potential interactions of Endozoicomonas and chlamydiae in microbial aggregates of the
448 coral Pocillopora acuta. *Science Advances* 2023; **9**: eadg0773.
449 4. Wada N, Hsu M-T, Tandon K, Hsiao SS-Y, Chen H-J, Chen Y-H, et al. High-resolution spatial and
450 genomic characterization of coral-associated microbial aggregates in the coral Stylophora
451 pistillata. *Science Advances* 2022; **8**: eabo2431.
452 5. Work T, Aeby G. Microbial aggregates within tissues infect a diversity of corals throughout the
453 Indo-Pacific. *Marine Ecology Progress Series* 2014; **500**: 1–9.
454 6. Damjanovic K, Menéndez P, Blackall LL, van Oppen MJH. Mixed-mode bacterial transmission in
455 the common brooding coral Pocillopora acuta. *Environmental Microbiology* 2020; **22**: 397–412.

7. Neave MJ, Apprill A, Ferrier-Pagès C, Voolstra CR. Diversity and function of prevalent symbiotic marine bacteria in the genus Endozoicomonas. *Applied Microbiology and Biotechnology* 2016; **100**: 8315–8324.

8. Tandon K, Lu C-Y, Chiang P-W, Wada N, Yang S-H, Chan Y-F, et al. Comparative genomics: Dominant coral-bacterium Endozoicomonas acroporae metabolizes dimethylsulfoniopropionate (DMSP). *The ISME Journal* 2020; **14**: 1290–1303.

9. Pogoreutz C, Oakley CA, Rädecker N, Cárdenas A, Perna G, Xiang N, et al. Coral holobiont cues prime Endozoicomonas for a symbiotic lifestyle. *The ISME Journal 2022* 2022; 1–13.

10. Hochart C, Paoli L, Ruscheweyh H-J, Salazar G, Boissin E, Romac S, et al. Ecology of Endozoicomonadaceae in three coral genera across the Pacific Ocean. *Nat Commun* 2023; **14**: 3037.

11. Smith HA, Moya A, Cantin NE, van Oppen MJH, Torda G. Observations of Simultaneous Sperm Release and Larval Planulation Suggest Reproductive Assurance in the Coral Pocillopora acuta. *Frontiers in Marine Science* 2019; **6**: 362.

12. Epstein HE, Torda G, van Oppen MJH. Relative stability of the Pocillopora acuta microbiome throughout a thermal stress event. *Coral Reefs* 2019; **38**: 373.

13. Sharp KH, Distel D, Paul VJ. Diversity and dynamics of bacterial communities in early life stages of the Caribbean coral Porites astreoides. *The ISME journal* 2012; **6**: 790–801.

14. Ceh J, van Keulen M, Bourne DG. Intergenerational transfer of specific bacteria in corals and possible implications for offspring fitness. *Microb Ecol* 2013; **65**: 227–231.

15. Li J, Zou Y, Yang J, Li Q, Bourne DG, Sweet M, et al. Cultured Bacteria Provide Insight into the Functional Potential of the Coral-Associated Microbiome. *mSystems* 2022; **7**: e00327-22.

16. Neave MJ, Michell CT, Apprill A, Voolstra CR. Endozoicomonas genomes reveal functional adaptation and plasticity in bacterial strains symbiotically associated with diverse marine hosts. *Scientific Reports* 2017; **7**: 1–12.

481 17. Szabó M, Larkum AWD, Vass I. A Review: The role of reactive oxygen species in mass coral
482 bleaching. *Advances in Photosynthesis and Respiration*. 2020. Springer, pp 459–488.
483 18. Curson ARJ, Liu J, Bermejo Martínez A, Green RT, Chan Y, Carrión O, et al.
484 Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene
485 in this process. *Nature Microbiology* 2017; **2**: 17009.
486 19. Doering T, Maire J, Oppen MJH van, Blackall LL. Advancing coral microbiome manipulation to
487 build long-term climate resilience. *Microbiol Aust* 2023; **44**: 36–40.
488 20. Maire J, van Oppen MJH. A role for bacterial experimental evolution in coral bleaching
489 mitigation? *Trends in Microbiology* 2022; **30**: 217–228.
490 21. Hedlund BP, Chuvochina M, Hugenholtz P, Konstantinidis KT, Murray AE, Palmer M, et al.
491 SeqCode: a nomenclatural code for prokaryotes described from sequence data. *Nat Microbiol*
492 2022; **7**: 1702–1708.
493 22. Bayer T, Neave MJ, Alsheikh-Hussain A, Aranda M, Yum LK, Mincer T, et al. The microbiome of
494 the red sea coral stylophora pistillata is dominated by tissue-associated endozoicomonas
495 bacteria. *Applied and Environmental Microbiology* 2013; **79**: 4759–4762.
496 23. Maire J, Blackall LL, van Oppen MJH. Microbiome characterization of defensive tissues in the
497 model anemone Exaiptasia diaphana. *BMC Microbiology* 2021; **21**: 152.
498 24. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible,
499 interactive, scalable and extensible microbiome data science using QIIME 2. *Nature*
500 *Biotechnology* 2019; **37**: 852–857.
501 25. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
502 *EMBnet.journal* 2011; **17**: 10.
503 26. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution
504 sample inference from Illumina amplicon data. *Nature Methods* 2016; **13**: 581–583.

505 27. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic
506 classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin.
507 *Microbiome* 2018; **6**: 90.
508 28. McMurdie PJ, Holmes S. phyloseq: An R package for reproducible interactive analysis and
509 graphics of microbiome census data. *PLoS ONE* 2013; **8**: e61217.
510 29. Davis NM, Proctor DiM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and
511 removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 2018;
512 **6**: 226.
513 30. Andrews S. FASTQC. A quality control tool for high throughput sequence data.
514 https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 17 Oct 2022.
515 31. Krueger F, James F, Ewels P, Afyounian E, Weinstein M, Schuster-Boeckler B, et al. TrimGalore:
516 v0.6.10 - add default decompression path. 2023. Zenodo.
517 32. Vidal-Dupiol J, Chaparro C, Pratlong M, Pontarotti P, Grunau C, Mitta G. Sequencing, de novo
518 assembly and annotation of the genome of the scleractinian coral, Pocillopora acuta. *bioRxiv*
519 2020; 698688.
520 33. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature methods* 2012; **9**:
521 357–9.
522 34. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map
523 format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.
524 35. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large
525 and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics (Oxford,*
526 *England)* 2015; **31**: 1674–1676.
527 36. Von Meijenfeldt FAB, Arkhipova K, Cambuy DD, Coutinho FH, Dutilh BE. Robust taxonomic
528 classification of uncharted microbial sequences and bins with CAT and BAT. *Genome Biology*
529 2019; **20**: 1–14.

530 37. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: Quality assessment tool for genome
531 assemblies. *Bioinformatics* 2013; **29**: 1072–1075.
532 38. Uritskiy GV, Diruggiero J, Taylor J. MetaWRAP - A flexible pipeline for genome-resolved
533 metagenomic data analysis. *Microbiome* 2018; **6**: 158.
534 39. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of
535 microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*
536 2015; **25**: 1043–1055.
537 40. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with
538 the Genome Taxonomy Database. *Bioinformatics* 2020; **36**: 1925–1927.
539 41. Seemann T. barrnap 0.9: rapid ribosomal RNA prediction.
540 https://github.com/tseemann/barrnap. .
541 42. Lane D. 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds). *Nucleic Acid*
542 *techniques in bacterial systematics*. 1991. John Wiley & Sons, New York, pp 115–47.
543 43. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
544 *Bioinformatics (Oxford, England)* 2014; **30**: 2114–2120.
545 44. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new
546 genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;
547 **19**: 455–477.
548 45. Bushnell B, Rood J, Singer E. BBMerge – Accurate paired shotgun read merging via overlap.
549 *PLOS ONE* 2017; **12**: e0185056.
550 46. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-
551 TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era.
552 *Molecular Biology and Evolution* 2020; **37**: 1530–1534.
553 47. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model
554 selection for accurate phylogenetic estimates. *Nat Methods* 2017; **14**: 587–589.

555 48. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast
556 Bootstrap Approximation. *Molecular Biology and Evolution* 2018; **35**: 518–522.
557 49. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display
558 and annotation. *Nucleic Acids Research* 2021; **49**: W293–W296.
559 50. Rodriguez-R LM, Konstantinidis KT. The enveomics collection: a toolbox for specialized analyses
560 of microbial genomes and metagenomes. 2016.
561 51. Kolde R. Pheatmap: pretty heatmaps. 2017.
562 52. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. Bakta: rapid and
563 standardized annotation of bacterial genomes via alignment-free sequence identification.
564 *Microb Genom* 2021; **7**: 000685.
565 53. Kanehisa M, Sato Y. KEGG Mapper for inferring cellular functions from protein sequences.
566 *Protein Sci* 2020; **29**: 28–35.
567 54. Jones P, Binns D, Chang HY, Fraser M, Li W, McAnulla C, et al. InterProScan 5: genome-scale
568 protein function classification. *Bioinformatics* 2014; **30**: 1236–1240.
569 55. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, et al. antiSMASH 7.0: new and
570 improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic*
571 *Acids Research* 2023; **51**: W46–W50.
572