Comparative genomics: Dominant coral-bacterium *Endozoicomonas acroporae* metabolizes dimethylsulfoniopropionate (DMSP)

Kshitij Tandon[1,2,3], Pei-Wen Chiang[1], Chih-Ying Lu[1], Naohisa Wada[1], Shan-Hua Yang[4], Ya-Fan Chan[1], Ping-Yun Chen[5], Hsiao-Yu Chang[5], Ming-Shean Chou[5], Wen-Ming Chen[6], Sen-Lin Tang[1,2#]

[1]Biodiversity Research Center, Academia Sinica, Taipei 115, Taiwan

[2]Bioinformatics Program, Institute of Information Science, Taiwan International Graduate Program, Academia Sinica, Taipei 115, Taiwan

[3]Institute of Cellular and Molecular Biology, National Tsing Hua University, Hsinchu 300, Taiwan

[4]Institute of Fisheries Science, National Taiwan University, Taipei 10617, Taiwan.

[5]Institute of Environmental Engineering, National Sun Yat-sen University, Kaohsiung 80424 Taiwan, ROC

[6]Laboratory of Microbiology, Department of Seafood Science, National Kaohsiung Marine University, No. 142, Hai-Chuan Rd, Nan-Tzu, Kaohsiung City 811, Taiwan, ROC.

Corresponding author

Corresponding author Email: sltang@gate.sinica.edu.tw

Abstract

Dominant coral-associated *Endozoicomonas* bacteria species are hypothesized to play a role in the coral-sulfur cycle by metabolizing Dimethylsulfoniopropionate (DMSP) into Dimethylsulfide (DMS); however, no sequenced genome to date harbors genes for this process. In this study, we assembled high-quality (>95% complete) genomes of strains of a recently added species *Endozoicomonas acroporae* (Acr-14^T, Acr-1 and Acr-5) isolated from the coral *Acropora muricata* and performed comparative genomic analysis on genus *Endozoicomonas*. We identified the first DMSP CoA-transferase/lyase—a *dddD* gene homolog found in all *E. acroporae* strains—and functionally characterized bacteria capable of metabolizing DMSP into DMS via the DddD cleavage pathway using RT-qPCR and gas chromatography (GC). Furthermore, we demonstrated that *E. acroporae* strains can use DMSP as the sole carbon source and have genes arranged in an operon-like manner to link DMSP metabolism to the central carbon cycle. This study confirms the role of *Endozoicomonas* in the coral sulfur cycle.

41 **Introduction**

42 Coral reefs are one of the most diverse ecosystems on Earth, with over 800 different coral
43 species known to date. Of these, corals in the genus *Acropora* are some of the most abundant
44 reef-building corals across the Indo-Pacific region [1]. These corals are also some of the
45 significant producers of dimethylsulfoniopropionate (DMSP) [2, 3]. DMSP is an organic sulfur
46 compound present in significant amounts in animals that harbor symbiotic algae such as
47 scleratinian corals and giant clams [2]. DMSP is present in coral tissues, mucus and
48 endosymbiotic dinoflagellates (Symbiodiniaceae) [4, 5]. In marine algae, DMSP acts as a
49 protectant from various stresses such as oxidative and osmotic stress [6]. Moreover, DMSP also
50 acts as an attractant for specific bacterial groups that have been reported to be part of coral-
51 associated bacterial communities and underpin coral health [7].

52 Once released from marine planktonic dinoflagellates, most of the DMSP is emanated to
53 surrounding water where it is readily available for microbial catabolic conversion as a source of
54 reduced carbon and sulfur [8, 9]. DMSP is a central molecule for marine sulfur cycling and is
55 degraded by bacteria via two pathways, a cleavage pathway and a demethylation pathway [9,
56 10]. The majority (~75%) of the DMSP is metabolized via the demethylation/demethiolation
57 pathway, producing methylmercaptopropionate [11, 12, 13]. The cleavage pathway, which
58 accounts for the remaining ~25%, produces dimethylsulfide (DMS)—a climate-active gas and
59 acrylic acid [13, 14, 15]. Genes and an enzyme (DddL) involved in the cleavage pathway has
60 been identified in *Sulfitobacter* [16]. Another gene, *dddD,* has been identified in *Marinomonas*
61 sp., which cleaves DMSP to produce DMS and 3-hydroxypropionate (3HP), not the usual
62 acrylate [13]. Moreover, DMSP metabolism yielding DMS and acrylate was recently also

63 identified in coccolithophore algae [17]. The DMS produced after cleavage of DMSP is then
64 released in the surrounding water [12].

65 Coral-associated bacterial communities are highly abundant, dynamic and diverse [18,
66 19, 20, 21]. These bacterial communities can be found in various niches associated with corals
67 like coral mucus [22], spaces in the skeleton [23, 24, 25], and within coral tissues [26, 27, 28].
68 Raina et al (2008) [8] confirmed that coral-associated bacteria have the potential to metabolize
69 organic sulfur compounds present in the coral tissues. They inferred that the majority of the
70 DMSP-degrading bacteria belongs to class *Gammaproteobacteria,* including *Alteromonas-,*
71 *Arhodomonas-, Idiomarina-, Pseudomonas-,* and *Spongiobacter (Endozoicomonas)*-related
72 organisms. Of these, *Arhodomonas-, Pseudomonas-,* and *Roseobacter*-related species harbor a
73 DMSP lyase—i.e. the *dddD* gene.

74 Of all the organisms in the tremendously diverse coral holobionts, species in the genus
75 *Endozoicomonas* (class: *Gammaproteobacteria*) have been studied widely for their genomics
76 and functional and ecological roles [29, 30]. *Endozoicomonas* species were found to be
77 abundant in the coral holobionts across the globe [31] and have been hypothesized to be
78 potential indicators of coral health [32]—they have high abundance in healthy corals and a
79 relatively low abundance in diseased or stressed corals [33, 34, 35]. Furthermore, nearly
80 complete and draft genomes of *Endozoicomonas* species isolated from corals and other marine
81 invertebrates [29, 34, 37] have been recently assembled. First, a nearly complete genome of
82 the *Endozoicomonas* isolate *Endozoicomonas montiporae* CL-33^T from the encrusting pore coral
83 *Montiporae aequituberculata* has provided insights into how this bacterium can interact with its
84 host from outside to inside of the coral cell with the help of potential effector proteins [30]. A

85 recent comparative genomics analysis also identified a high number of Type III secretion
86 system-related genes and that most gene ontology terms were associated with the generic
87 transport of molecules and those genomes of *Endozoicomonas* species show high plasticity
88 [30]. *Endozoicomonas* species have been hypothesized to play a role in the coral sulfur cycle by
89 effectively metabolizing DMSP into DMS [38, 39]. However, no study has confirmed the genus'
90 role and no sequenced genome has harbored any genes in this process [30]. Hence, the role of
91 this coral endosymbiont in the coral sulfur cycle remains elusive.

92 In this study, we assembled high quality and nearly complete draft genomes of newly
93 added *Endozoicomonas acroporae* strains and profiled its abundance in different coral species in
94 the Indo-pacific region. We identified and functionally characterized the first DMSP Co-A
95 transferase/lyase; dddD homolog in all *E. acroporae* strains. Furthermore, we provide conclusive
96 evidence that *E. acroporae* has a role in the coral sulfur cycle by effectively metabolizing DMSP
97 into DMS and is able to use DMSP as the sole carbon source for growth and survival.

98 **Materials and Methods**

99

100 **Culturing and whole-genome sequencing of *E. acroporae* Acr-1, Acr-5 and Acr-14^T**

101 Strains of *E. acroporae* Acr-1, Acr-5 and Acr-14^T were isolated from the coral *Acropora*
102 *muricata* from Kenting, off the southern coast of Taiwan, and cultured using a method described
103 previously [40]. Genomic DNA isolation was performed using the cetyltrimethylammonium
104 bromide (CTAB) method [41]; the quality of the isolated DNA was assessed using NanoDrop 1000
105 (Thermo Scientific, USA). High quality DNA was sent to the core sequencing facility at Biodiversity
106 Research Center, Academia Sinica (BRCAS), Taiwan for whole-genome sequencing on the Illumina

107 Miseq platform, with a TRUSeq DNA paired-end library generated to achieve an insert size of 500
108 bp.

109 **Genome assembly and annotation**

110 Reads obtained from Illumina MiSeq were quality-filtered and trimmed (Phred score ≥ 30)
111 using NGS QC toolkit v2.3.3 [42]. Quality-filtered and trimmed reads were *de novo* assembled
112 using CLC Genomics Workbench version 1.10.1 (Qiagen) with a bubble size of 40 and automatic
113 word size enabled. Minimum contig length was set to ≥500 bp (no scaffolding was performed).
114 Assembled genomes were quality checked for completeness, contamination, and heterogeneity
115 using CheckM [43]. Other *Endozoicomonas* species genomes were downloaded from the NCBI
116 genomes database (last accessed January 2018). Gene prediction was performed using Prokka
117 [44] on all genomes used in this study to avoid bias from different gene prediction methods.
118 Predicted genes were annotated by rapid annotation using the Subsystem Technology (RAST)
119 Server [45] with gene calls from Prokka preserved.

120 **Identification of genomic characteristics**

121 The *in-silico* DNA-DNA hybridization (DDH) percentages among the genomes of genus
122 *Endozoicomonas* were calculated using the genome-to-genome distance calculator from the
123 DSMZ server [46], and the Average Nucleotide Identity (ANI) (https://enve-
124 omics.ce.gatech.edu/ani/) calculator was used to calculate the ANI values. Further, Amino-Acid
125 Identity (AAI) among the genomes was calculated with CompareM
126 (https://github.com/dparks1134/CompareM). Clustered Regularly Interspaced Short
127 Palindromic Repeat (CRISPR) structures in all genomes were identified using Prokka; prophages

128 and phages within the genomes were identified using the PHAge Search Tool (PHAST) [47], which
129 classified the phages as intact, incomplete and questionable. Type III secretion system (T3SS)
130 proteins were identified by EffectiveT3 [48] using EffectiveDB with the animal classification
131 module and selective (0.9999) restriction value method enabled. Insertion Sequence (IS)
132 elements in the genomes of *E. acroporae* strains and *E. montiporae* were identified using the
133 ISfinder database (https://www.is.biotoul.fr) with blastn and an e-value threshold of 1e-5.

134 **16S rRNA gene phylogenetic analysis**

135 To determine robust phylogenetic relationships within the genus *Endozoicomonas*, all
136 available 16S rRNA sequences (80 in number) were downloaded from the NCBI taxonomy
137 database, for which host information was available to understand the distribution of
138 *Endozoicomonas* species in different marine invertebrates and identify the position of *E.*
139 *acroporae* strains within the genus *Endozoicomonas*. Sequences were aligned using MUSCLE [49]
140 in MEGA7 [50] with default parameters, and conserved regions were extracted manually.
141 Furthermore, the conserved regions were again aligned using the above method. The
142 phylogenetic tree was constructed using the maximum-likelihood method with Kimura 2-
143 parameter and Gamma distributed with invariant sites enabled in MEGA7 with 1000 bootstraps.

144 ***E. acroporae* distribution and abundance in different coral species from the Indo-Pacific region**

145 We analyzed microbial community data publically available from three different
146 studies—1) our laboratory's previous study [34]; 2) a study of coral-associated bacterial
147 communities in the Red Sea [51]; and 3) coral-associated bacterial community from reefs in the
148 east and west coast of Australia, including Ningaloo Reef, Lizard Island, reefs from the northern

149 sector of the Great Barrier Reef, and Lorde Howe Island [52]—to profile the abundance of *E.*
150 *acroporae* strains in different corals species from Penghu Archipelago, Taiwan, the Red Sea, Saudi
151 Arabia, and east and west Australia. Operational Taxonomic Unit (OTU) abundance profiles and
152 their representative sequences were obtained from supplementary materials in Shiu JH et al,
153 2017 [34], Ziegler M et al, 2016 [51], and Pollock FJ et al, 2018 [52]. We performed similarity
154 searches on all the OTU sequences in the three studies against an in-house 16S rRNA gene
155 sequence database of *Endozoicomonas* species with standalone blastn [53] and profiled the
156 relative abundance of *E. acroporae* strains at the three locations with e-value <1e-5 and identity
157 threshold ≥ 97%.

158 **Comparative genomics: Pan-genome analysis and core-genome phylogeny**

159 Pan-genome analysis was performed with Roary [54] using GFF files of all the genomes
160 obtained with Prokka**.** Core genome identification and alignment was performed on all genomes
161 using the parameters *-i 80, -e, –n*. A core genome-based phylogenetic tree was constructed using
162 Fasttree [55] with –nt, -gtr parameters and visualized in the GENETYX-Tree Program
163 (https://www.genetyx.co.jp/).

164 **Identification of stress response genes, *dddD CoA-transferase/lyase,* and DMSP metabolism**
165 **related operons**

166 The Stress Response Subsystem was analyzed for distribution of different categories of
167 stress responsive genes present in the genomes of *Endozoicomonas*. The Sulfur Metabolism
168 Subsystem in the RAST analysis annotated a *dddD* gene capable of metabolizing DMSP into DMS
169 within the "Sulfur Metabolism- no subcategory." Furthermore, the presence of domains in the

170 *dddD* gene and DMSP metabolism-related operon genes was determined with web-based
171 Conserved Domain Search Service [56,57], NCBI, with default parameters.

172 **DMSP degradation by strain Acr-14^T**

173 *Endozoicomonas acroporae* Acr-14^T was cultivated on the Modified Marine Broth
174 Version 4 (MMBV4) medium [29] with several modifications (Supplementary Table S1) at 25 °C
175 for 48 hours. Carbon sources, 0.1 % maltose, and 0.5 mM DMSP were added to the enrichment
176 culture of strain Acr-14^{T}and kept at 25 °C for 24 hours (OD_{600} of ~1.0 after incubation). The
177 enrichment culture was centrifuged at 2000 x *g* for 10 minutes, and the supernatant was
178 discarded. The enrichment culture pellet was washed with 1 ml fresh minimal medium
179 (Supplementary Table S2) twice to remove MMBV4 medium containing DMSP. Two
180 experimental groups were made for the test: (A) 1 ml of washed bacteria was resuspended in a
181 40 ml minimal medium containing 0.2% casamino acid and 1 mM DMSP; (B) 1 ml of washed
182 bacteria was resuspended in a 40 ml minimal medium containing 0.2% casamino acid without
183 DMSP. The cultures were sampled at 0, 16, 24 and 48 hours for RT-qPCR.

184 **Gene expression of *dddD* by RT-qPCR**

185 Total RNA was extracted using a TRI-reagent solution (Invitrogen, Carlsbad, CA, USA).
186 The culture samples (1-2 ml) were centrifuged for 1 min at 12000 x *g* and 4 °C following the
187 manufacturer's guidelines. The RNA pellet was air-dried and resuspended in nuclease-free
188 water. Residual DNA was removed using a TURBO DNA-free Kit (Invitrogen). RNA quality was
189 determined using NanoDrop ND-1000 UV-Vis Spectrophotometer (Nano-Drop Technologies).
190 RNA integrity was assessed by electrophoresis on 1% agarose-guanidine thiocyanate gel.

191 Complementary DNA (cDNA) was synthesized from purified RNA using the SuperScript IV First
192 Strand Synthesis System for RT-qPCR (Invitrogen) following the manufacturer's guidelines. For
193 cDNA, RT and non-RT samples were screened for residual DNA contamination using the
194 hypervariable V6V8 region of the bacterial 16S rRNA gene (U968F and U1391R).

195 A primer pair was made for the *dddD* gene based on the genome data of strain Acr-14^{T}
196 for Quantitative PCR using DNASTAR Lasergene [58] and Primer-BLAST tool on BLAST search
197 (NCBI): *dddD*-F (5'- ACCGCATCGCACCACTCAGG-3') and *dddD*-R (5'-GGCCCCGGTTGTTTCATCAT-
198 3'). The endogenous control was performed with *rpoD* gene, *rpoD*-F (5'-
199 AAGGCGGTGGACAAGTTCG-3') and *rpoD*-R (5'-GATGGTGCGGGCCTGGTCTG-3'). RT-qPCR assays
200 were carried out using Applied Biosystems QuantStudio^{TM} 5 Real-Time PCR System. The
201 standard cycling program consisted of cycles of UDG activation at 50 °C for 2 minutes, initial
202 denaturation-activation at 95 °C for 2 minutes and 40 cycles of denaturation at 95 °C for 10
203 seconds and annealing at 60 °C for 40 seconds using PowerUp SYBR Green Master Mix (Thermo
204 Fisher Scientific, USA). A dissociation step was performed to confirm the specificity of the
205 product and avoid the production of primer dimers. For all reactions, 10 ng of template DNA
206 was added to a reaction of 10 µl. The 10 µl reactions contained 5 µl of PowerUp SYBR Green
207 Master Mix (Thermo Fisher Scientific, USA), 3.4 µl of sterilized nuclease-free water, 0.3 µl each
208 of the forward and reverse primers (final conc. 0.3 µM), and 1µl of DNA template. Each sample
209 was performed in duplicate. The relative quantification of the expression ratio was calculated
210 by the comparative $2^{-\Delta\Delta C_T}$ method. Differences between treatments were statistically tested
211 using *t-test*.

212 **Quantification of released DMS**

213 To assay the functional activity of the DddD protein in *Endozoicomonas acroporae* strain
214 Acr-14^T, cultures were first grown in MMB medium with 0.5 mM DMSP for 24 hours. After that,
215 3 ml culture was collected, spun down, and washed twice with minimal medium and then
216 resuspended with 1 ml minimal medium. This 1 ml culture was then injected into sterile 60 ml
217 vials sealed with a rubber stoppers containing designated medium Treatment (a): 20 ml
218 minimal medium with 0.2 % casamino acid, and 1 mM DMSP; Treatment (c): 20 ml minimal
219 medium with 0.2 % casamino acid, acting as the negative control; and a culture-free Treatment
220 (b): 20 ml minimal medium with 0.2 % casamino acid, and 1 mM DMSP acted as the control.
221 Vials were then incubated at 25 °C, 200 rpm, in dark. After incubation for 0, 24 and 48 hours, 1
222 ml of the head space air sample was collected and injected into a gas chromatograph
223 (Shimadzu, GC-14B) fitted with a flame ionization detector (150 °C) and column (SGE, 60m ×
224 0.53mm ID BP624 × 3.0 µm) to detect the DMS concentration.

225 ***E. acroporae* growth on DMSP as the sole carbon source**

226 *E. acroporae* Acr-1, Acr-5, and Acr-14^T and *E. montiporae* CL-33^T strains were cultivated on
227 MMB medium (along with 0.1% maltose for *E. montiporae* CL-33^T) at 25 °C for 72 hours. The
228 carbon source, 0.1 mM DMSP, was added to the enrichment culture of *E. acroporae* strains and
229 kept at 25 °C for 24 hours (OD_{600} of ~0.3 after incubation). 0.1 mM DMSP and 0.1% maltose
230 were added to the enrichment culture of *E. montiporae* and kept at 25 °C for 24 hours (OD_{600} of
231 ~0.8 after incubation). The enrichment culture was centrifuged at 2000 x g for 10 minutes, and
232 the supernatant was discarded. The enrichment culture pellet was washed twice by 1 ml fresh
233 minimal medium. After resuspending with minimal medium, the enrichment cultures were
234 added to the treatments and adjusted to OD_{600} of 0.06. All the treatments were kept at 25 °C,

235 200 rpm, and the OD_{600} of each treatment was recorded after incubating for 24, 48, and 72
236 hours.

237 *E. acroporae* Acr-14^T, Acr-1, Acr-5, and *E. montiporae* CL-33^T were cultivated in minimal
238 medium with 0.2% casamino acid and 0.1 mM DMSP to test their ability in usage DMSP as the
239 sole carbon source. Treatments with different concentrations of DMSP were set to confirm
240 whether *E. montiporae* CL-33^T can grow with DMSP or not. In these treatments, *E. montiporae*
241 CL-33^T was cultivated in minimal medium with 0.2% casamino acid and 3 mM DMSP; 0.2%
242 casamino acid and 7 mM DMSP; 0.2% casamino acid and 1% maltose as a positive control.

243

244 **Data Availability**

245 Draft genomes of *E. acroporae* Acr-1 and *E. acroporae* Acr-5 have been deposited into GenBank
246 under accession IDs SAUT00000000 and SAUU00000000, respectively. *E. acroporae* Acr-14^T
247 genome has been previously made public under accession ID: PJPV00000000

248 **Results**

249 **16S rRNA gene phylogeny and *E. acroporae* abundance profiling in different coral species**
250 **from the Indo-Pacific region**

251 All three strains of *E. acroporae* had only one copy of 16S rRNA gene, compared to seven copies
252 in the *E. montiporae* CL-33^T genome (Supplementary Table S3). 16S rRNA gene based phylogeny
253 clustered sequences in a somewhat host-specific way (Supplementary Figure S1). We identified
254 separate clades for *E. montiporae* and *E. acroporae* (Supplementary Figure S1). The closest

255 relative of *E. acroporae* strains was a new species, *Endozoicomonas coralli*, within the same
256 clade, whose genome has not been sequenced yet. Another *Endozoicomonas* species was
257 *Endozoicomonas atrinae* WP70^{T}, whose genome has been sequenced.

258 Relative abundance profiles of *E. acroporae* strains were determined in different coral
259 species from three distinct geographical regions (Fig. 1). We determined that *E. acroporae* strains
260 were abundant in corals from the Red Sea, Saudi Arabia (15-22%); Penghu Archipelago, Taiwan
261 (>50%); and coral reefs in eastern (10-49%) and western (10-75%) Australia. Presence and
262 differential relative abundance of *E. acroporae* strains provide evidence that this species is one
263 of the abundant bacteria among coral holobionts spread across distinct geographic locations.

264

265 **Genome assembly features**

266 Whole-genome sequencing of *E. acroporae* isolates produced high-quality genome
267 assemblies with completeness >95% and contamination <3% (Supplementary Figure S2). The *E.*
268 *acroporae* Acr-1 genome was assembled into 299 contigs with a total size of 6,024,033 bp, and
269 we predicted 5,059 coding genes and 77 tRNAs. The *E. acroporae* Acr-5 genome was 6,034,674
270 bp in length with 295 contigs and 5,101 coding genes and 80 tRNAs. Genome assembly details of
271 *E. acroporae* Acr-14^{T} were 6.048 Mb long, 309 contigs, 5018 coding genes and 81 tRNAs; these
272 have also been reported in an earlier study (Tandon K et al, 2018). A list of all genomes used in
273 this study—along with their genome size (4.049-6.69 Mb) and host (coral, sea slug, comb pen
274 shell and sponge)—is shown in Table S4. All *E. acroporae* genomes (Acr-1, Acr-5 and Acr-14^{T}) had
275 similar genome sizes (6.024 -6.048 Mb) and GC contents (49.2, 49.3, and 49.3%, respectively).

276 **Genomic characteristics of genus *Endozoicomonas***

277 *Endozoicomonas* species have large genomes ranging from 4.049 Mb (*Endozoicomonas*
278 sp. AB1) to 6.69 Mb long (*E. elysicola* DSM22380). The average genome has ~5,100 protein coding
279 genes and a GC content of 47.6%. The average DDH, ANI and AAI values among the
280 *Endozoicomonas* species were ~46, ~73 and ~75%, respectively, which are at the lower end of
281 the 62-100% range of interspecies variation within a genus (Kim M et al, 2014), indicating high
282 genomic diversity (Fig. 2A, B, C).

283 A diverse array of IS elements were identified when comparing the genomes of *E.*
284 *acroporae* and *E. montiporae* species. For example, of the total 44 IS elements identified, 22 were
285 only present in *E. acroporae* genomes, while 18 were only present in *E. montiporae* genomes.
286 Relatively few IS elements (ISEc46, ISEcret1, ISEc42 and ISPa18) were present in both species.
287 Each genome also harbored some unique IS elements; among these, Acr-14^T had 6 unique IS
288 elements (Supplementary Figure S3).

289 *E. acroporae* genomes had more genes coding for T3SS with Acr-14^T (523), Acr-5(499) and
290 Acr-1 (499) than *E. monitporae* CL-33^T (249), *E. atrinae* WP70^T (382) or other *Endozoicomonas*
291 species (Supplementary Table S5). All the genomes encode complete pathways for the biogenesis
292 of most amino-acids. We further identified different phage insertions in all *Endozoicomonas*
293 genomes (Supplementary Table S6) and different CRISPR counts in the genomes of *E. acroporae*
294 Acr-14^T, Acr-5 and Acr-1 (Supplementary Table S3).

295 **Core genome phylogeny**

296 Core gene (n=308)-based phylogenetic analysis identified a somewhat host-specific
297 phylogeny and also hinted towards a high genomic divergence within the species of genus
298 *Endozoicomonas. Endozoicomonas* genomes isolated from the same host clustered very tightly
299 together, e.g. *E. acroporae* and *E. montiporae* with the coral host and *E. ascidiicola* and *E.*
300 *arenosclerae* with a sponge host (Fig. 2D). Moreover, *E. numazuensis* and *E. arenosclerae*
301 genomes shared a branch, both being isolated from the sponge. *E. acroporae* Acr-1 and Acr-5
302 clustered tightly while sharing a branch with *E. acroporae* Acr-14^T, which confirms that Acr-1 and
303 Acr-5 are closer than Acr-14^T; the former had ANI value of 99.81%.

304

305 **Analyzing RAST subsystems**

306 RAST subsystems; carbohydrates, protein metabolism, and amino acids and derivatives
307 had the highest number of genes, representing an average of 11.69, 10.85 and 13.29% of total
308 annotated genes, respectively (Supplementary Figure S4), in all *Endozoicomonas* genomes. We
309 focused our analysis on two other important subsystems, stress response and sulfur metabolism.
310 The highest number of genes were annotated for oxidative stress (39.16% of stress genes) in the
311 stress response subsystem, followed by unclassified (16%) and detoxification stress (12.53%)
312 (Supplementary Figure S5). Interestingly, in sulfur metabolism, we identified a *dddD* gene
313 homolog capable of metabolizing DMSP into DMS in *E. acroporae* Acr-14^T, Acr-5, and Acr-1
314 genomes. This gene was not present in other *Endozoicomonas* genomes (Supplementary Table
315 S4). The *dddD* gene present in *E. acroporae* genomes have two identical CaiB domains (position:
316 1-416, 439-821) belonging to the coenzyme-A transferase superfamily (Supplementary Fig S7).

317 **DMSP metabolism operon and links to the central carbon cycle**

318 We identified homologs of the DMSP transcriptional regulator (LysR family), a sulfur
319 transporter belonging to the betaine/carnitine/choline transporter (BCCT), and genes involved in
320 producing acetate from DMSP (Fig. 3A). We discerned a complete pathway with genes arranged
321 in a consecutive manner to form an operon that can yield acetate from DMSP metabolism via
322 three-step enzymatic reactions mediated by DddD, 3-hydroxypropionate dehydrogenase (EC
323 1.1.1.59), and malonate-semialdehyde dehydrogenase (EC 1.2.1.18) (Fig. 3B) in all the *E.*
324 *acroporae* genomes with significant gene and protein identity for all the three genes (>97%) and
325 proteins (>97%) (Fig. 3C). Identification of the DMSP cleavage pathway leading to central carbon
326 metabolism suggested that *E. acroporae* species can use DMSP as a carbon source.

327 ***dddD* gene activity and DMS quantification**

328 RT-qPCR was used to examine the expression of the *dddD* gene in *E. acroporae* Acr-14^{T}.
329 *dddD* gene expression increased with sampling time in the condition with 1 mM DMSP (Fig. 4A).
330 *dddD* gene expression in this condition was 42.77, 56.52 and 91.37 times higher than samples
331 without DMSP at 16, 24 and 48 hours, respectively (t-test, $p<0.05$), confirming that *dddD* was
332 active in *E. acroporae*. After confirming the *dddD* gene expression, we quantified the amount of
333 DMS released by *E. acroporae* when incubated in a DMSP-rich environment with a time series
334 detection (0, 24 and 48 hours). We identified that the DMS signal can only be detected in
335 Treatment (a). There was no DMS signal in the control groups (Treatments (b) and (c)) (Fig 4B). A
336 temporal increase in released DMS concentration in Treatment (a) confirmed the ability of *E.*
337 *acroporae* to metabolize DMSP into DMS.

338 **DMSP as the sole carbon source**

339 *E. acroporae* Acr-14^T, Acr-1, Acr-5 all showed the ability to use DMSP as the sole carbon
340 source. Acr-14^T had a mean OD_{600} value ~0.7 at the 72-hour time point, which was the highest
341 across all treatments (Fig. 4C). The mean OD_{600} value of Acr-1 and Acr-5 was ~0.45 and ~0.3,
342 respectively, after three days of incubation (Fig. 4C). In contrast, *E. montiporae* CL-33^T did not
343 show any sign of growth across all time points (Fig. 4C). We than tested whether *E. montiporae*
344 CL-33^T could grow under different concentrations of DMSP, equivalent to the same carbon
345 concentration of maltose that can support *E. montiporae* growth (Supplementary Figure S6). The
346 result showed that the *E. montiporae* CL-33^T OD_{600} value did not increase across all treatments
347 0.1, 3, and 7 mM of DMSP after three days of incubation, which further confirmed that *E.*
348 *montiporae* CL-33^T cannot use DMSP as its sole carbon source, but also requires maltose as the
349 carbon source for growth (Supplementary Figure S6).

350 **Discussion**

351 In this study, we performed genomic (and comparative genomic) and functional assays,
352 and abundance profiling of strains from a newly added species *E. acroporae* in the exclusive
353 marine bacterial genus *Endozoicomonas.* We assembled high quality, nearly complete genomes
354 of *E. acroporae* strains and identified for the first time the *dddD* gene responsible for the
355 previously hypothesized ability of *Endozoicomonas* species to metabolize DMSP. Furthermore,
356 we characterized the functional activity of the identified gene, the ability of *E. acroporae* strains
357 to grow on DMSP as their sole carbon source. This is the first study to establish the role of an
358 *Endozoicomonas* species in the coral sulfur cycle with genomic and functional evidence.

359 *Endozoicomonas* species have been suggested to play a role in the coral sulfur cycle via
360 DMSP metabolism [8,59]. However, no gene related to DMSP metabolism was found in
361 previously sequenced *Endozoicomonas* genomes (Supplementary Table S4) [30]. Furthermore,
362 using enrichment cultures from coral mucus, tissue and skeleton, certain bacterial species have
363 been reported that can degrade DMSP, including *Spongiobacter (Endozoicomonas)*
364 *nickelotolerans,* but no gene was identified and it was hypothesized to use a different pathway
365 for DMSP degradation [8]. In this study, we investigated the sulfur metabolism pathway in high
366 quality and nearly complete draft genomes (Supplementary Figure S2) of *E. acroporae* strains and
367 identified a *dddD* gene, which is a Co-A transferase/lyase that cleaves DMSP to volatile DMS [9].
368 Among the eight identified DMSP lyases in different bacterial species, only *dddD* produces 3-HP
369 and DMS, whereas others form acrylate and DMS [60]. 3-HP is not the usual product of other
370 DMSP lyases, which directly produce acrylate after the cleavage due to their unusual organization
371 of class III Co-A transferase domains [61]. Co-A transferase domains present in the DddD protein
372 are similar to CaiB (Supplementary Figure S7), which is a homodimer capable of adding a Co-A to
373 carnitine [13]. RT-qPCR-based temporal analysis of *E. acroporae* strains in a DMSP-rich
374 environment showed that *dddD* activity increased significantly (Fig. 4A); moreover, the GC
375 analysis also showed that high amounts of DMS were released, providing the evidence that *dddD*
376 is functionally active (Fig. 4B).

377 Genomic analysis also identified genes arranged in consecutive order to form an operon
378 with DMSP transport, metabolism, and transcriptional regulator genes (Fig. 3A) in all the strains
379 of *E. acroporae*. *dddT* is a transporter that imports molecules like DMSP [13, 62], and thus has
380 been suggested to be an importer of DMSP. *dddR* is a transcription regulator with the ability to

381 activate the expression of the *dddD* gene in response to DMSP [13]. The functions of *dddB* and
382 *dddC* have not been characterized biochemically, but they are hypothesized to be involved in
383 oxido-reductive function based on their sequences similarity with other oxido-reductases,
384 thereby modifying DMSP either before or after the addition of acyl CoA moiety by *dddD* [13].
385 Interestingly, the arrangement of genes around *dddD* has been associated to be of the "Pick 'n
386 Mix" form in different $DddD^+$ bacteria, like *Marinomonas sp.* MWYL1, *Sagittus* sp. E37, and *B.*
387 *cepacia* AMMD [62]; a similar pattern was observed in *E. acroporae* strains as well (Fig. 3A).

388 The ability of *E. acroporae* strains to use DMSP as their sole carbon source (Fig. 4C)
389 compared to *E. montiporae*—which is unable to utilize DMSP, even at higher concentrations
390 (Supplementary Figure S6)—is further evidence that *E. acroporae* strains can metabolize DMSP
391 and use it for growth and survival. This finding indicates that the DMSP metabolism is directly
392 linked to the central carbon cycle, as we predicted in the genomic analysis (Fig. 3B). Furthermore,
393 it also confirmed the hypothesis that marine bacteria that harbor *dddD* have the ability to grow
394 on DMSP as the sole carbon source [63].

395 Bacterial genome size can reflect that bacteria's evolutionary dependency on its host
396 when they are engaged in obligate symbiosis [64]. A striking feature of obligate symbionts is that
397 they have smaller genome sizes than facultative symbionts [65]. However, genomes of
398 *Endozoicomonas* species are relatively long, with a size range of 4.09 Mb to 6.69 Mb; this includes
399 *E. acroporae* strains, which have an average genome length of ~6.03 Mb, suggesting that they
400 might have free living stage. The possibility that *Endozoicomonas* species have a free living stage
401 is further supported by their large number of coding genes (Avg. ~5,000 proteins). These features
402 suggest that genome streamlining [66], a notable feature of symbiotic bacteria, does not

403 predominantly occur in genus *Endozoicomonas*. Moreover, a diverse array of phages
404 (Supplementary Table S6) and IS elements (Supplementary Figure S3) provides clues about the
405 infection and colonization histories of different marine hosts, aided by frequent divergence
406 events. Recent studies by Ding JY et al. (2016) [29] reported a high proportion of repeat
407 sequences and also a diverse variety of IS in the genome of *E. montiporae,* which may help the
408 bacterium adapt to the host and also identify the *N-deglycosylation* enzyme that helps penetrate
409 the mucus layer of the host. These findings suggest that *Endozoicomonas* can transition between
410 different symbiotic lifestyles.

411 We identified a high number of secretory genes (Type III secretion system, T3SS) in the all
412 the genomes of *E. acroporae* (Supplementary Table S5) that may help transport organic
413 maromolecules between the host and symbiont. The T3SS system helps bacteria interact with
414 their host, and a complete gene set for the assembly of this vital secretory system was observed
415 in all the genomes used in this study. *E. acroporae* has secretory genes that regulate host
416 mechanisms, similar to those of *E. montiporae*, and can increase the chances that bacteria survive
417 in the host while improving the host fitness. Additionally, other genes found in *E. acroporae* may
418 also provide clues about the survival strategy of *E. acroporae* in hosts. For example, identification
419 of a catalase gene in *E. acroporae* strains—along with phosphoenolpyruvate synthase and 7,8-
420 dihydro-8-oxoguanine triphosphatase as secretory genes from T3SS—can help the bacterium
421 survive by scavenging H_2O_2, modulating the gluconeogeneis and confer resistance from oxidative
422 stresses to the host respectively. Up-regulating genes involved in gluconeogenesis have been
423 proposed as a response to stress-induced starvation in corals [67].

424 Phylogenetic analysis based on a core-genome (308 genes) clustered the species in a
425 somewhat host-specific manner, as it has done in previous studies [30, 31]. However, it was
426 interesting to see that, even when the host type is the same (i.e. stony coral), *E. montiporae* and
427 *E. acroporae* did not share any branch, and their strains cluster tightly within their clade, which
428 suggests that the co-diversification between host and symbiont is complex in nature (Fig. 2D).
429 Moreover, a highly reduced core genome along with similar ANI, DDH and AAI at the lower end
430 of inter species range give clues about the interspecies diversity within genus *Endozoicomonas*
431 (Fig. 2A, B, C).

432 With the high abundance of *Endozoicomonas* species reported in coral holobionts across
433 the globe and very few culturable isolates known to date, it has been hypothesized that
434 *Endozoicomonas* is either species specific or can have a large host range. In our study, we
435 identified as being *E. acroporae* prevalent in diverse coral species from geographically distant
436 locations in the Indo-Pacific region. Relative abundance varies among coral species (Fig. 1). The
437 presence of *E. acroporae* in different coral species supports the idea that this new species has a
438 large host range in the genus *Endozoicomonas;* it also highlights the importance of studying this
439 genus in greater detail. Identification of other isolates that can be cultured will enhance our
440 understanding of this diverse marine genus. This further points to the need to advance our
441 culturing techniques and combine the sequencing approaches with culturomics.

442 *Endozoicomonas* species have only been isolated from marine invertebrates; hence it is
443 reasonable to believe that *Endozoicomonas* and their hosts experiences ever-present oxidative
444 stress [69] in the marine environment. Identification of a high percentage of oxidative stress
445 response genes in all the genomes of *Endozoicomonas* provides clues for understanding its

446 adaptation to the marine environment and potential ability to mitigate oxidative stress.
447 Moreover, earlier studies have reported that DMSP, which has antioxidant properties, and its
448 breakdown products are functionally important in the coral stress response [69]. In our study,
449 we also identified 7,8-dihydro-8-oxoguanine triphosphatase as a T3SS effector protein, which
450 might provide resistance against mitochondrial dysfunction and spontaneous mutagenesis [29,
451 70] to the coral host during thermal stress, which is also an indication of a putative symbiotic
452 association between E. *acroporae* and *A. muricata.*

453 **Conclusion**

454 *Endozoicomonas* species have been long hypothesized to play a role in the coral sulfur cycle, and
455 this study provides genomic and functional evidence for the first time to support this hypothesis.
456 *E. acroporae* strains can not only metabolize DMSP to produce DMS, but also use it as the sole
457 carbon source for growth and survival. In our study we also identified the first DMSP-related
458 operon in *E. acroporae*, which links DMSP metabolism to the central carbon cycle. Furthermore,
459 the presence of stress responsive genes at higher proportions gives clues about how this genus
460 adapted in marine environments. Since very few species in this diverse marine genus have been
461 cultured to date, we believe that there are still many other functional aspects, and more focus
462 on this genus in regards to coral reef health can provide better insights in the near future.

463 **Acknowledgements**

464 This study was supported by funding from Academia Sinica. K.T would like to acknowledge the
465 Taiwan International Graduate Program (TIGP) for its fellowship towards his graduate studies.

The authors would like to acknowledge support from Dr. Shu-Fen Chiou during the GC experiment. We would like to thank Noah Last of Third Draft Editing for his English language editing.

Author contribution

K.T and S.L.T conceived the idea for this study. K.T assembled the genomes, performed the bioinformatics analysis, and wrote the manuscript. P.W.C cultured the strains and performed the RT-qPCR analysis. C.Y.L and Y.F.C performed GC experiment and analysis. C.Y.L performed the sole carbon test and analysis. N.W and S.Y helped write the manuscript and modify the illustrations. P.Y.C, H.Y.C, and M.S.C helped with the GC experiments and provided the instruments for conducting the experiment. W.M.C provided the cultures. S.L.T supervised the overall study. All authors read and approved the manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. Veron, JEN. *Corals of the World* Vol.1-3 (Australian Institute of Marine Sciences,2000).
2. Broadent AD, Jones GB, and Jones RJ. DMSP in Corals and Benthic Algae from the Great Barrier Reef. Estuar. Coast. Shelf Sci. 2002; 55:547-555**.**
3. Raina JB, Tapiolas DM, Foret S, Lutz A, Abrego D, Ceh J *et al.* DMSP biosynthesis by and animal and its role in coral thermal stress response. Nature. 2013; 502:677-680.
4. Broadent AD and Jones GB. DMS and DMSP in mucus ropes, coral mucus, surface films and sediment pore waters from coral reefs in the Great Barrier Reef. Mar Freshwater Res. 2004; 55:849-855
5. LaJeunesse TC, Parkinson JE, Gabrielson PW, Jeong HJ, Reimer JD, Voolstra CR, et al. Systematic revision of symbiodiniaceae highlights the antiquity and diversity of coral endosymbionts. *Curr Bio.* 2018; 28:2570-80.

6. Sunda W, Kieber DJ Kiene, R.P. and Huntsman, S. An antioxidant function for DMSP and DMS in marine algae. *Nature.* 2002; 418:317-320.
7. Raina JB, Dinsdale EA, Willis BL, Bourne DG. Do the organic sulfur compounds DMSP and DMS drive coral microbial associations? *Trends Microbiol*. 2010; 18:101-108.
8. Raina JB. Tapiolas D, Willis D, Bourne DG. Coral-associated bacteria and their role in the biogeochemical cycling of sulfur. *Appl Environ Microbiol*. 2009; 75:3492.
9. Reisch CR, Moran MA, Whitman WB. Bacterial catabolism of Dimethylsulfoniopropionate (DMSP). Front Microbiol. 2011; 2:172.
10. Curson ARJ, Todd JB, Sullivan MJ, Johnston AWB. Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. Nat Rev Microbiol. 2011; 9:849-859.
11. Howard, EC. *et al.* Bacterial taxa that limit sulfur flux from the ocean. Science. 2006; 314:649-652.
12. Kiene RP., Linn LJ., and Burton JA. New and important roles for DMSP in marine microbial communities. *J. Sea Res.* 2000; 43:209-224.
13. Todd JD, Rogers R, Li YG, Wexler M, Bond PL, Sun L. Structural and regulatory genes required to make the gas dimethylsulfide in bacteria. Science. 2007; 315:666-669.
14. Moran MA. *et al.* Genome sequence of *Silibacter pomeroyi* reveals adaptation to the marine environment. *Nature*. 2004; 432:910-913.
15. Yoch DC. Dimethylsulfoniopropionate: its source, role in the marine food web, and biological degradation of dimethylsulfide. *Appl Environ Microbiol*. 2002; 68:5804-5815.
16. Curson ARJ, Rogers R, Todd JD, Brearely CA, and Johnston AWB. Molecular genetic analysis of a dimethylsulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine α-proteobacteria and *Rhodobacter spharoides. Environ. Microbiol.* 2008: 10;757-767.
17. Alcolombri U. *et al.* Indentification of the algal dimethyl sulfide releasing enzyme: a missing link in the marine sulfur cycle. *Science.* 2015: 348:1466-1469.
18. Kooperman N, Ben-Dov E, Kramarsky-Winter E, Barak Z, Kushmaro A. Coral mucus-associated bacterial communities from natural and aquarium environments. *FEMS Microbiol Lett.* 2007: 276;106-113.
19. Rohwer F, Seguritan V, Azam F and Knowlton N. Diversity and distribution of coral-associated bacteria. *Mar. Ecol. Prog. Ser.* 2002: 243;1-10.
20. Yang SH, Tseng CH, Huang CR, Chen CP, Tandon K, Lee STM, *et al.* Long-Term Survey Is Necessary to Reveal Various Shifts of Microbial Compositions in Corals. *Front. Microbiol.* 2017: 8;1094.
21. Zanevald JR, Burkepile DE, Shantz AA, Pritchard CE, McMinds R, Payet JP, *et al.* Overfishing and nutrient pollution interact with temperature to disrupt coral reefs down to microbial scales. *Nat. Commun.* 2016: 7;11833.
22. Lee STM, Davy S, Tang SL and Kench PS. Mucus sugar content shapes the bacterial community structure in thermally stressed *Acropora muricata. Fron. Microbiol.* 2016: 7;1371.

532 23. Ferrer LM and Szmant AM. Nutrient regeneration by the endolithic community in coral
533 skeletons. In: Proceedings of the 6th International Coral Reef Symposium, Vol. 1; 1988.
534 p.1-4.
535 24. Yang SH, Lee STM, Huang CR, Tseng CH, Chiang PW, Chen CP. *et al.* Prevalence of
536 potential nitrogen-fixing, green sulfur bacteria in the skeleton of reef-building coral
537 *Isopora palifera. Limnol Oceanogr.* 2016: 61;1078-86.
538 25. Yang SH, Tandon K, Lu CY, Wada N, Shih CJ, Hsiao SSY. *et al.* Metagenomic, phylogenetic,
539 and functional characterization of predominant endolithic green sulfur bacteria in the
540 coral *Isopora palifera*. *Microbiome.* 2019; 7:3.
541 26. Reshef L, Koren O, Loya Y, Zilber-Rosenberg I and Rosenberg E. The coral probiotic
542 hypothesis. *Environ. Microbiol.* 2006: 8:2068-2073.
543 27. Ritchie KB. Regulation of microbial populations by coral surface mucus and mucus-
544 associated bacteria. *Mar. Ecol. Prog. Ser.* 2006; 322:1-14.
545 28. Rosenberg E, Koren O, Reshef L, Efrony R, and Zilber-Rosenberg I. The role of
546 microorganisms in coral health, disease and evolution. *Nat Rev Microbiol.* 2007: 5;355-
547 362.
548 29. Ding JY, Shiu JH, Chen WM, Chiang YR and Tang SL. Genomic insight into the host-
549 endosymbiont relationship of *Endozoicomonas montiporae* CL-33(T) with its coral host.
550 *Front. Microbiol.* 2016: 7;251.
551 30. Neave MJ, Michell CT, A Apprill and Voolstra CR. *Endozoicomonas genomes* reveal
552 functional adaptation and plasticity in bacterial strains symbiotically associated with
553 diverse marine hosts. *Sci. Rep.* 2017: 7;40579.
554 31. Neave MJ, Apprill A, Ferrier-Pagès C and Voolstra CR. Diversity and function of prevalent
555 symbiotic marine bacteria in the genus *Endozoicomonas*. *Appl. Microbiol. Biotechnol.* 2016:
556 100;8315-8324.
557 32. Bourne DG, Iida Y, Uthicke S and Smith-Keune C. Changes in coral-associated microbial
558 communities during a bleaching event. *ISME J.* 2008: 2;350-363.
559 33. Vezzulli L, Pezzati E, Huete-Stauffer C, Pruzzo C, Cerrano C. 16S rDNA pyrosequencing of
560 the Mediterranean Gorgonian *Paramuricea clavata* reveals a link among alterations in
561 bacterial holobiont members, anthropogenic influence and disease outbreaks. *PLoS*
562 *One.* 2013: 8;6745.
563 34. Shiu JH, Keshavmurthy S, Chiang PW, Chen HJ, Lou SP, Tseng CH. *et al.* Dynamics of
564 coral-associated bacterial communities acclimated to temperature stress based on
565 recent thermal history. *Sci. Rep.* 2017: 7;14933.
566 35. O'Brien P, Smith HA, Fallon S, Fabricus K, Willis BT, Morrow KM. *et al.* Elevated CO_2 Has
567 Little Influence on the Bacterial Communities Associated with the pH-Tolerant Coral,
568 Massive *Porites* spp. *Front. Microbiol.* 2018: 9;2621.
569 36. Neave MJ, Michell CT, Apprill A, Voolstra CR. Whole-genome sequences of three
570 symbiotic *Endozoicomonas* strains. *Genome Announc.* 2014: 2;e00802-e00814.

571 37. Tandon K, Chiang PW, Chen WM, Tang SL. Draft Genome Sequence of *Endozoicomonas*
572 *acroporae* Strain Acr-14^{T}, Isolated from *Acropora* Coral. *Genome Announc.* 2018: 6(6):
573 e0156-17.
574 38. Ransome E, Rowley SJ, Thomas S, Tait K, Munn CB. Disturbance to conserved bacterial
575 communities in the coral water gorgonian coral *Eunicella verrucosa. FEMS Micorbiol.*
576 *Ecol.* 2014: 90;404-416.
577 39. Pike RE, Haltli R, Kerr RG. Description of *Endozoicomonas euniceicola* sp. nov. and
578 *Endozoicomonas gorgoniicola* sp. nov., bacteria isolated from the octocoral *Eunicea*
579 *fusca* and *Plexaura* sp., and an emended description of the genus *Endozoicomonas. Int J*
580 *Syst Evol Microbiol*. 2014: 63;4294-302.
581 40. Sheu SY, Lin KR, Hsu MY, Sheu DS, Tang SL. et al. *Endozoicomonas acroporae* sp. nov.
582 isolated from *Acropora* coral. *Int J Syst Evol Micorbiol.* 2017: 67;3791-3797.
583 41. Wilson K. Preparation of genomic DNA from bacteria. *Current Protocols in Molecular*
584 *Biology*. 2001: 56;241-245.
585 42. Patel RK and Jain M. NGS QC toolkit: a toolkit for quality control of next generation
586 sequencing data. *PLoS One.* 2012: 7; e30619.
587 43. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing
588 quality of microbial genomes recovered from isolates, single cells, and metagenomes.
589 *Genome Res.* 2015: 25;1043-1055.
590 44. Seeman T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics.* 2014:
591 20;2068-2069.
592 45. Aziz RK, Bartels D, Best DD, DeJogh M, Disz T, Edwards RA. *et al.* The RAST Server: rapid
593 annotations using subsystems technology. *BMC Genomics*. 2008: 9;75.
594 46. Meier-Kolthof JP, Klenk HP, and Goker M. Taxonomic use of DNA G+C content and DNA-
595 DNA hybridization in the genomic age. *Int. J. Syst. Microbiol.* 2014: 64;325-356.
596 47. Zhou Y, Liang Y, Lynch KH, Denis JJ, and Wishart DS. PHAST: a fast phage search tool.
597 *Nucleic Acids Res.* 2011: 37; W347-W352.
598 48. Eichinger V, Nussbaumer T, Platzer A, Jehl MA, Arnold R, Rattei T. *et al.* EffectiveDB-
599 updates and novel feature annotation of bacterial secreted proteins and Type III, IV, VI
600 secretion systems. *Nucleic Acids Res.* 2016: 44; D669-D674.
601 49. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high
602 throughput. *Nucleic Acids Res.* 2004: 32;1792-1797.
603 50. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis
604 Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* 2016: 33(7);180-1874.
605 51. Ziegler M, Roik A, Porter A, Zubier K, Mudarris SM, Ormond R. *et al.* Coral microbial
606 community dynamics in response to anthropogenic impacts near a major city in the
607 central Red Sea. *Mar. Pollut. Bull.* 2016: 105;629-640.
608 52. Pollock FJ, McMinds R, Smith S, Bourne DG, Willis BL, Medina M. *et al*. Coral-associated
609 bacteria demonstrate phylosymbiosis and cophylogeny. *Nat. Commun.* 2018: 9;4921.
610 53. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J.*
611 *Mol. Biol.*1990: 215;403-410.

612 54. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG. *et al.* Roay: rapid large-
613 scale prokaryote pan genome analysis. *Bioinformatics.* 2015: 31(22); 3691-3693.
614 55. Price MN, Dehal PS, Arkin AP. Fasttree: computing large minimum-evolution trees with
615 profiles instead of a distance matrix. *Mol Biol Evol.* 2009: 26;1641-1650.
616 56. Marchler-Bauer A, Bryant SH. CD-Search: protein domain annotations on the fly. *Nucleic*
617 *Acids Res.* 2005: 32; W327-W331.
618 57. Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY. *et al.*
619 CDD/SPARCLE: functional classification of proteins via subfamily domain architectures.
620 *Nucleic Acids Res.* 2017: 45; D200-D203.
621 58. Burland TG. DNASTAR's Lasergene sequence analysis software. *Methods Mol. Biol.* 2000:
622 132;71-91.
623 59. Bourne DG, Morrow KM and Webster NS. Insights into the coral microbiome:
624 underpinning the health and resilience of reef ecosystems. *Annu. Rev. Microbiol.* 2016:
625 70;317-340.
626 60. Bullock HA, Luo HW and Whitman WB. Evolution of dimethylsulfoniopropionate
627 metabolism in marine phytoplankton and bacteria. *Front. Microbiol.* 2017: 8;637.
628 61. Yoch DC. Dimethylsulfoniopropionate: its sources, role in the marine food web, and
629 biological degradation to dimethylsulfide. *Appl. Environ Microbiol.* 2002: 68; 5804-5815.
630 62. Johnston AWB, Todd JD, Sun L, Nikolaidou-Katsaridou MN, Curson ARJ, Rogers R.
631 Molecular diversity of bacterial production of the climate active gas, dimethylsulphide, a
632 molecule that impinges on local and global symbioses. *J. Exp. Bot.* 2008: 59;1059-1067.
633 63. Todd JD, Curson ARJ, Nikolaidou-Katsaridou N, Brealey CA, Watmough NJ, Chan Y. *et al.*
634 Molecular dissection of bacterial acrylate catabolism- unexpected links with
635 dimethylsulfoniopropionate catabolism and dimethylsulfide production. *Environ.*
636 *Microbiol.* 2010: 12;327-343.
637 64. Moran NA, McCutcheon JP and Nakabachi A. Genomics and evolution of heritable
638 bacterial symbionts. *Annu. Rev. Genet.* 2008: 42;165-190.
639 65. McCutcheon JP and Moran NA. Extreme genome reduction in symbiotic bacteria. *Nature*
640 *Rev. Microbiol.* 2011: 10;13-36.
641 66. Kwan JC, Donia MS, Han AW, Hirose E, Haygoog MG, Schmidt EW. Genome streamlining
642 and chemical defense in a coral reef symbiosis. *Proc. Natl. Acad. Sci.* USA 2012:
643 109;0655-20660.
644 67. Kenkel C, Meyer E and Matz M. Gene expression under chronic heat stress in
645 populations of the mustard hill coral (*Porites asteroids)* from different thermal
646 environments. *Mol. Ecol.* 2013: 22; 4322-4324.
647 68. Downs CA, Fauth JE, Halas JC, Dustan P, Bemiss J, Woodley CM. Oxidative stress and
648 seasonal coral bleaching. *Free Radical Biol. Med.* 2002: 33; 533-543.
649 69. Arias-González JE, Johnson C, Seymour RM, Perez P, Aliño P (2011) Scaling up models of the
650 dynamics of coral reef ecosystems: An approach for science-based management of global
651 change Coral Reefs: An Ecosystem in Transition. Springer, pp373–388.

652 70. Salvioli A, Ghignone S, Novero M, Navazio L, Venice F, Bagnaresi P *et al.* Symbiosis with an
653 endobacterium increases the fitness of mycorrhizal fungus, raising its bioenergetics potential.
654 *ISME J.* 2016: 10;130-144.

655

656

657 **Figure Legends**

658 **Fig. 1.** Distribution and abundance profiles of *E. acroporae* strains in the Red Sea, Saudi Arabia;
659 Penghu Archipelago, Taiwan; and eastern and western Australian coral reefs. The green color
660 shows relative abundance of *E. acroporae* strains.

661 **Fig. 2. Genomics characteristics of genus *Endozoicomonas.*** Heatmaps based on **A)** Average
662 Amino-acid Identity (AAI), **B)** Average Nucleotide Identity (ANI)**, C)** DNA-DNA Hybridization (DDH)
663 for all the sequenced genomes from genus *Endozoicomonas.* **D)** Core-genome (308)-based
664 phylogenetic tree.

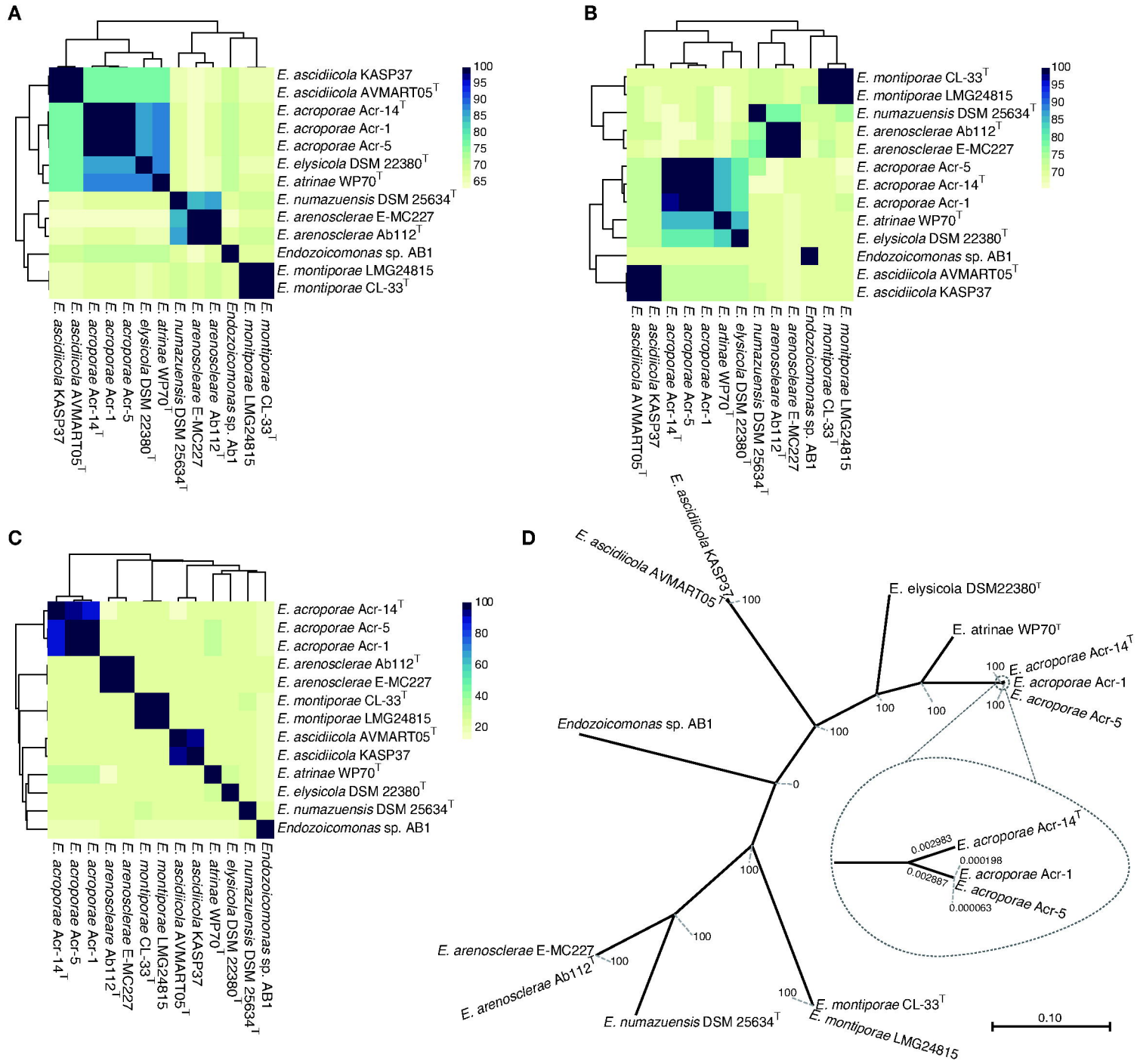
665 **Fig. 3. 3A)** DMSP metabolism operon, with genes arranged in an operon manner with in all *E.*
666 *acroporae* strains, with a similar arrangement in *Marinomonas* sp. MWYL1, with different
667 putative DMSP transporters (*BetS* and *BetL*). **3B)** Metabolic pathway represented by the genes,
668 linking DMSP metabolism to the central carbon cycle, producing acetate from DMSP with the
669 release of DMS. **3C)** Heatmaps based on gene/protein identity for the three enzymes used in the
670 metabolic cycle, DddD, 3-hydroxypropionate dehydrogenase (EC 1.1.1.59), and malonate-
671 semialdehyde dehydrogenase (EC 1.2.1.18).

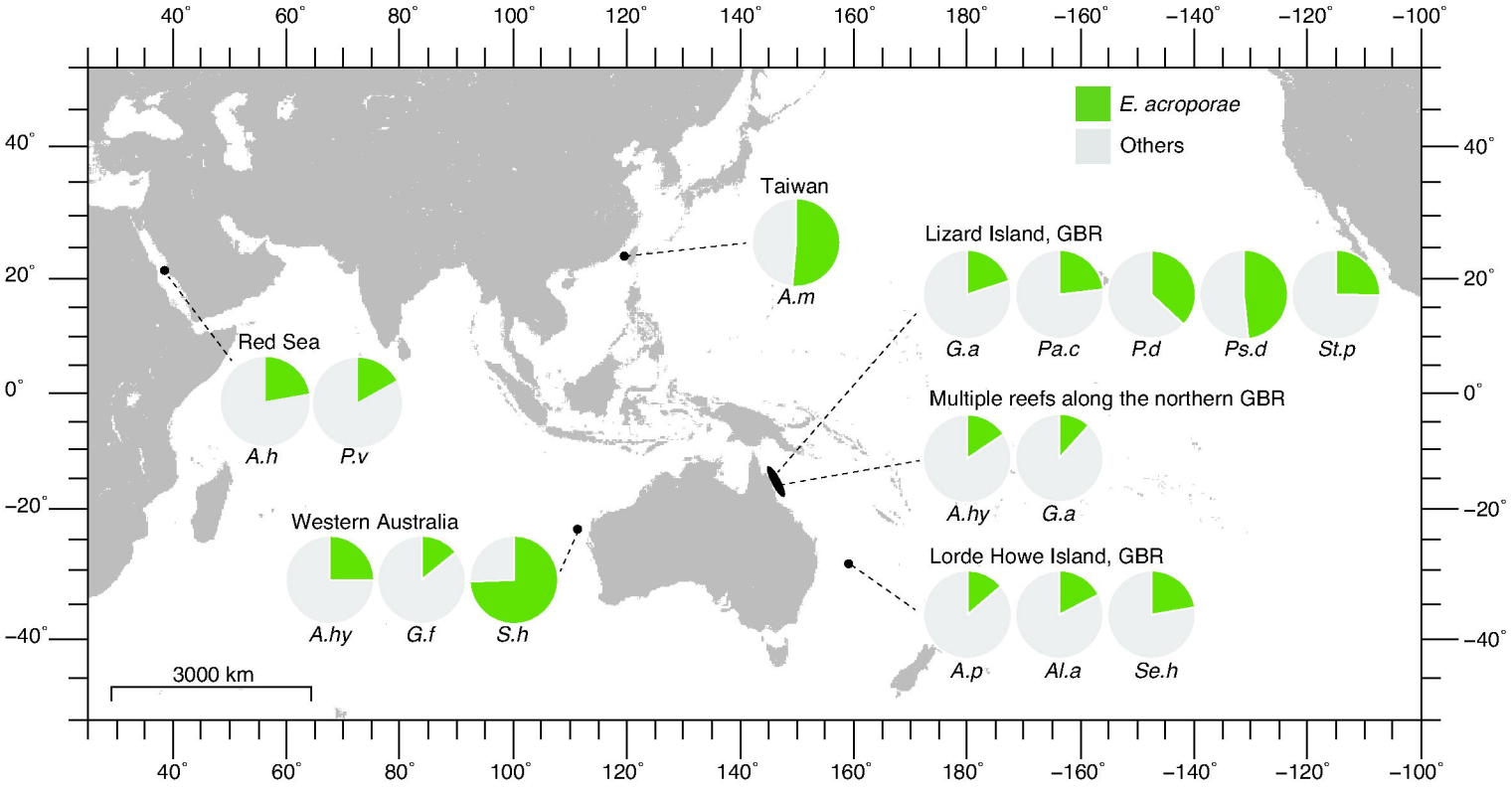
672 **Fig. 4. 4A)** RT-qPCR of the *dddD* gene from *E. acroporae* Acr-14^{T} to measure temporal expression
673 change in the presence of 1 mM DMSP. **4B)** GC-based quantification of DMS released from *E.*

674 *acroporae* Acr-14^{T} cultures in the presence of DMSP (1mM). **4C)** Sole Carbon Test; *E. acroporae*

675 strains can use DMSP as the sole carbon source for their growth and survival, whereas *E.*

676 *montiporae* CL-33^{T} cannot grow on DMSP as the sole carbon source.


A
E. ascidiicola KASP37
E. ascidiicola AVMART05T
E. acroporae Acr-14T
E. acroporae Acr-1
E. acroporae Acr-5
E. elysicola DSM 22380T
E. atrinae WP70T
E. numazuensis DSM 25634T
E. arenosclerae E-MC227
E. arenosclerae Ab112T
Endozoicomonas sp. AB1
E. montiporae LMG24815
E. montiporae CL-33T
100
95
90
85
80
75
70
65
B
E. montiporae CL-33T
E. montiporae LMG24815
E. numazuensis DSM 25634T
E. arenosclerae Ab112T
E. arenosclerae E-MC227
E. acroporae Acr-5
E. acroporae Acr-14T
E. acroporae Acr-1
E. atrinae WP70T
E. elysicola DSM 22380T
Endozoicomonas sp. AB1
E. ascidiicola AVMART05T
E. ascidiicola KASP37
100
95
90
85
80
75
70
C
E. acroporae Acr-14T
E. acroporae Acr-5
E. acroporae Acr-1
E. arenosclerae Ab112T
E. arenosclerae E-MC227
E. montiporae CL-33T
E. montiporae LMG24815
E. ascidiicola AVMART05T
E. ascidiicola KASP37
E. atrinae WP70T
E. elysicola DSM 22380T
E. numazuensis DSM 25634T
Endozoicomonas sp. AB1
100
80
60
40
20
D
E. ascidiicola AVMART05T
E. ascidiicola KASP37
E. elysicola DSM22380T
E. atrinae WP70T
E. acroporae Acr-14T
E. acroporae Acr-1
E. acroporae Acr-5
Endozoicomonas sp. AB1
E. arenosclerae E-MC227
E. arenosclerae Ab112T
E. numazuensis DSM 25634T
E. montiporae CL-33T
E. montiporae LMG24815
0.002983
0.000198
0.002887
0.000063
0.10



Acropora
- *A. hemprichii* (*A.h*)
- *A. hyacinthus* (*A.hy*)
- *A. muricata* (*A.m*)
- *A. pallifera* (*A.p*)

Pocillopora
- *P. verrucosa* (*P.v*)
- *P. damicornis* (*P.d*)

Galaxea
- *G. fascicularis* (*G.f*)
- *G. astreata* (*G.a*)

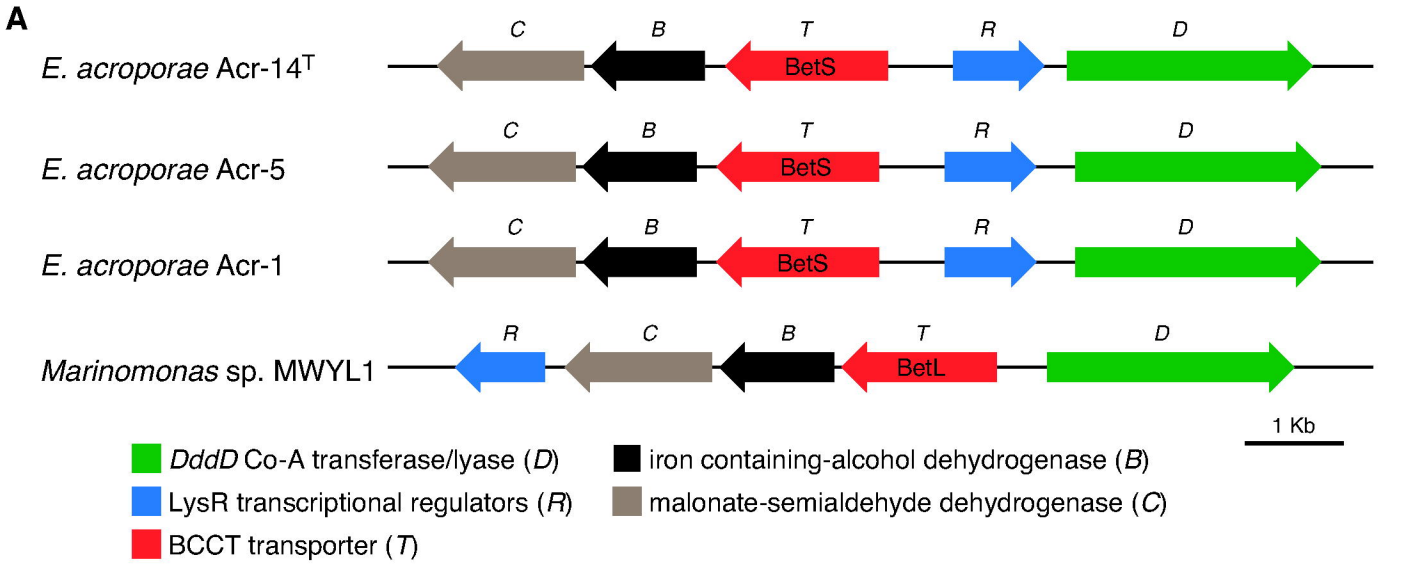
Playthoa
- *P. caesia* (*Pa.c*)

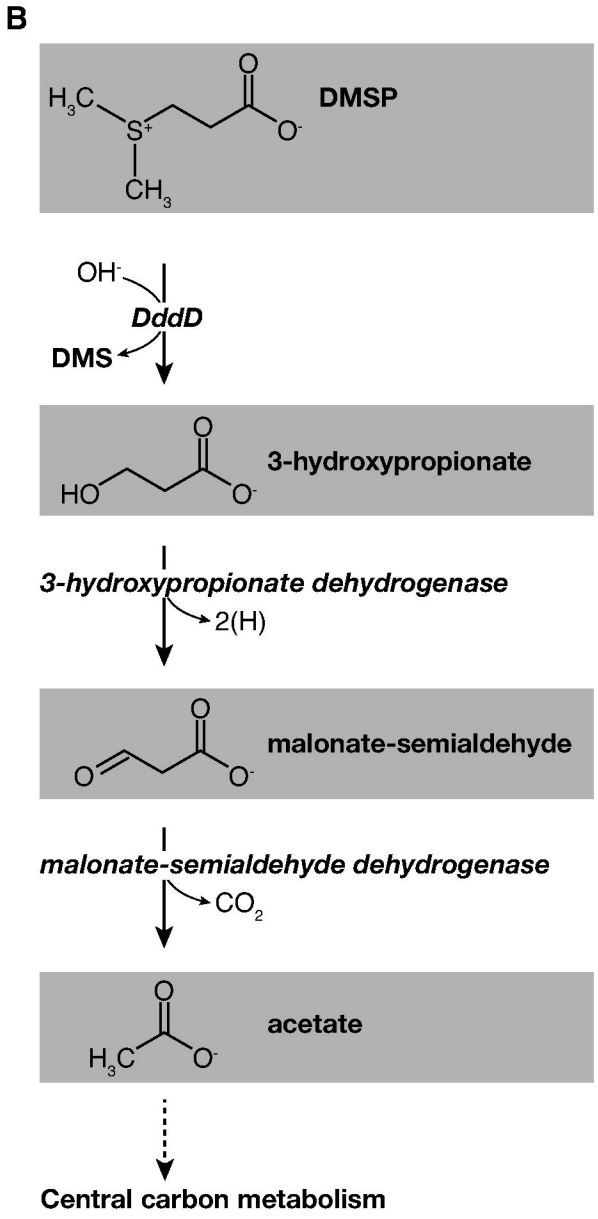
Psammacora
- *P. digitata* (*Ps.d*)

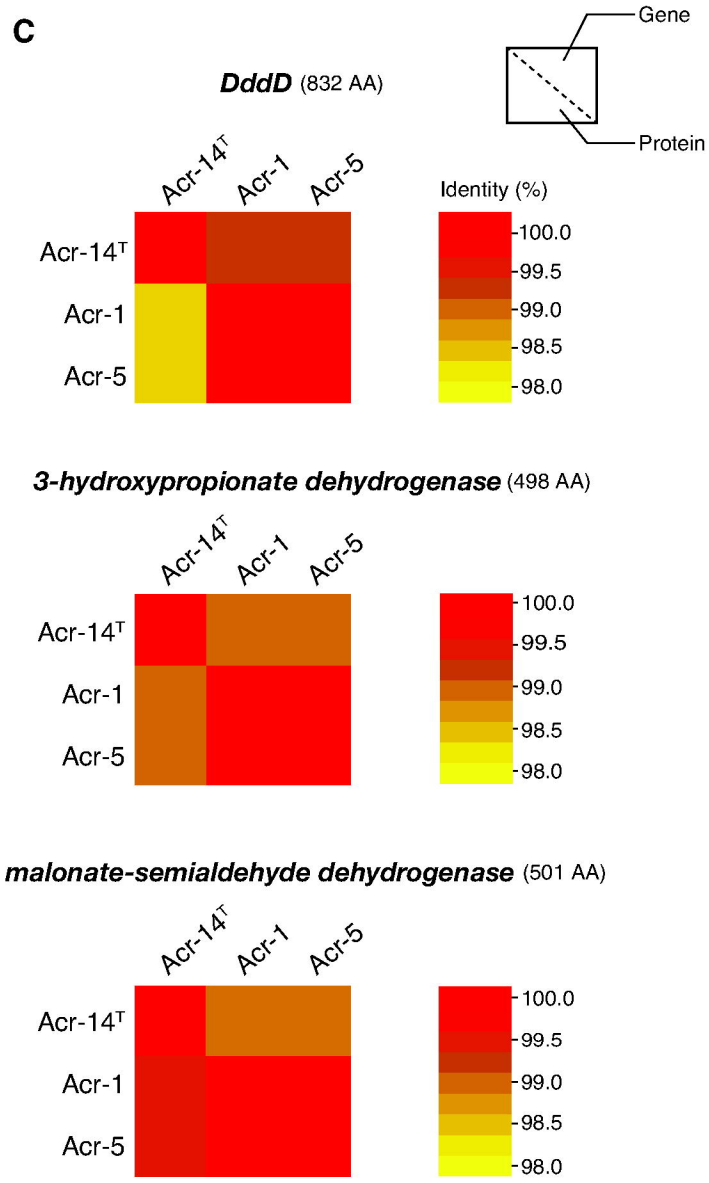
Seriatopora
- *S. hystrix* (*Se.h*)

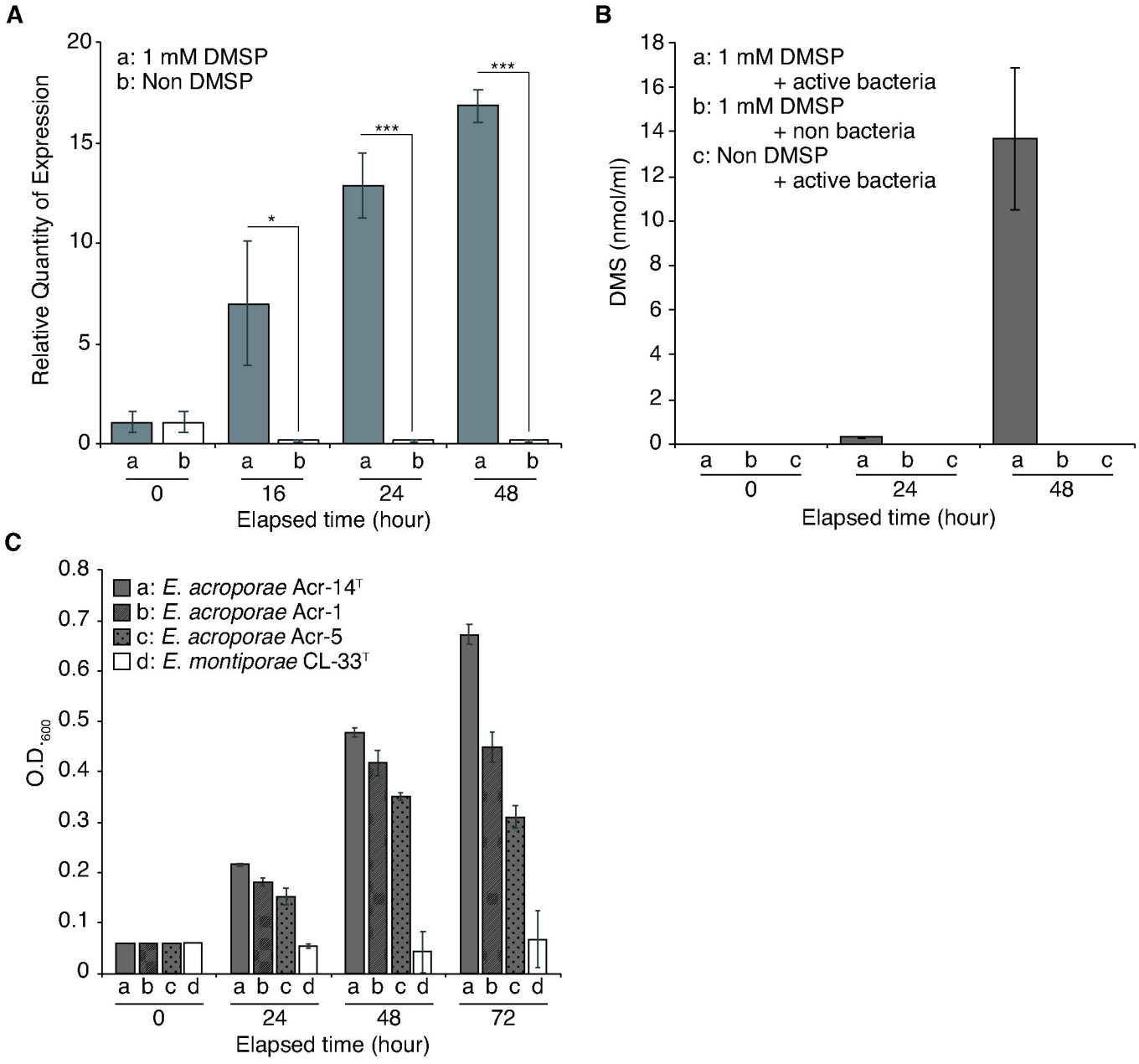
Stylophora
- *S. pistiallata* (*St.p*)

Alveopora
- *A. allingi* (*Al.a*)


A
E. acroporae Acr-14T
E. acroporae Acr-5
E. acroporae Acr-1
Marinomonas sp. MWYL1
C
B
T
R
D
BetS
BetL
1 Kb
DddD Co-A transferase/lyase (D)
LysR transcriptional regulators (R)
BCCT transporter (T)
iron containing-alcohol dehydrogenase (B)
malonate-semialdehyde dehydrogenase (C)


B
DMSP
OH-
DddD
DMS
3-hydroxypropionate
3-hydroxypropionate dehydrogenase
2(H)
malonate-semialdehyde
malonate-semialdehyde dehydrogenase
CO2
acetate
Central carbon metabolism


C
Gene
Protein
DddD (832 AA)
Identity (%)
Acr-14T
Acr-1
Acr-5
100.0
99.5
99.0
98.5
98.0
3-hydroxypropionate dehydrogenase (498 AA)
malonate-semialdehyde dehydrogenase (501 AA)


A
a: 1 mM DMSP
b: Non DMSP
Relative Quantity of Expression
20
15
10
5
0
*
***
***
a b a b a b a b
0 16 24 48
Elapsed time (hour)
B
a: 1 mM DMSP
+ active bacteria
b: 1 mM DMSP
+ non bacteria
c: Non DMSP
+ active bacteria
DMS (nmol/ml)
18
16
14
12
10
8
6
4
2
0
a b c a b c a b c
0 24 48
Elapsed time (hour)
C
a: E. acroporae Acr-14T
b: E. acroporae Acr-1
c: E. acroporae Acr-5
d: E. montiporae CL-33T
O.D.600
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
a b c d a b c d a b c d a b c d
0 24 48 72
Elapsed time (hour)