Title: *Endozoicomonas*-chlamydiae interactions in cell-associated microbial aggregates of the coral *Pocillopora acuta*

Justin Maire[1,*], Kshitij Tandon[1], Astrid Collingro[2], Allison van de Meene[1], Katarina Damjanovic[3], Cecilie Ravn Gøtze[1,3], Sophie Stephenson[3], Gayle K. Philip[4], Matthias Horn[2], Neal E. Cantin[3], Linda L. Blackall[1], Madeleine J. H. van Oppen[1,3]

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

[2] Centre for Microbiology and Environmental Systems Science, University of Vienna, Vienna, 1030, Austria.

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

[4] Melbourne Bioinformatics, The University of Melbourne, Parkville, VIC, 3010, Australia.

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

Abstract

Corals are associated with a variety of bacteria, which occur in the surface mucus layer, gastrovascular cavity, skeleton, and tissues. Some tissue-associated bacteria form clusters, termed cell-associated microbial aggregates (CAMAs), which are poorly studied. Here, we provide a comprehensive characterization of CAMAs in the coral *Pocillopora acuta*. Combining imaging techniques, laser capture microdissection, and amplicon and metagenome sequencing we show that CAMAs: (i) are located in the tentacle tips and may be intracellular; (ii) contain *Endozoicomonas, Kistimonas* (both Gammaproteobacteria), and *Simkania* (Chlamydiota) bacteria; (iii) *Endozoicomonas* may provide vitamins to its host and use secretion systems and/or pili for colonization and aggregation; (iv) *Endozoicomonas* and *Simkania* occur in distinct, but adjacent, CAMAs; (v) *Simkania* may rely on acetate and heme provided by neighboring *Endozoicomonas*. Our study provides detailed insight into coral endosymbionts, which will guide the assessment of their suitability for probiotic approaches to mitigate coral bleaching.

Keywords: microbiome, *Endozoicomonas*, CAMA, endosymbiont, coral, chlamydiae

Introduction

Coral reefs are among the most important ecosystems on the planet because they constitute hotspots for biodiversity (*1*), and provide many goods and services, including coastal protection and important industries such as tourism and fisheries (*2*). Scleractinian corals are foundation members of coral reefs, as their skeleton forms the reefs' three-dimensional structure that provides habitats for many reef-dwelling organisms and because they are the main primary producers on the reef. Corals rely on a wide range of microorganisms for their survival, including protists, bacteria, archaea, fungi and viruses (*3–5*). Photosynthetic dinoflagellates of the Symbiodiniaceae family are by far the most studied coral-associated microorganisms. They are vital symbionts because they provide their hosts with most of their energy through the translocation of photosynthates (*6*). Some coral-associated bacteria also have roles that are critical to the coral host, such as protection against pathogens (*7*), and cycling of nitrogen and sulfur (*8–10*).

Bacteria can colonize all microhabitats in corals and are most abundant and diverse in the mucus and skeleton (*3*, *4*). Tissue-associated bacteria are less diverse (*9*, *11*) and some are known to form large, dense clusters termed cell-associated microbial aggregates (CAMAs). Initially described in the early 1980s, they were thought to be linked to coral disease in *Acropora palmata* (*12*, *13*). CAMAs have since been described in numerous cnidarian species, including hard corals, soft corals, and anemones, regardless of disease status, and are particularly common in the genera *Acropora, Pocillopora, Porites, Platygyra* and *Stylophora* (*14–23*). Apart from location and distribution within the coral host, knowledge about CAMAs is scarce, with a handful of studies applying taxa-specific histological techniques to identify bacteria residing in CAMAs (*16*, *19*, *23*, *24*), and only one recent study applying culture-independent genomics to assess their functional potential (*21*). These studies have shown that bacteria residing in CAMAs often belong to the *Endozoicomonas* genus, a widespread coral symbiont (*24*). Several genomes of *Endozoicomonas* cultured from corals have been sequenced (*10*, *25*, *26*), but the location of these strains within the coral host and their ability to form CAMAs was not assessed. Hence, a holistic understanding of coral-associated CAMAs is still lacking, despite the huge progress made in coral microbiome research in the past couple of decades.

Here, we provide a detailed characterization of CAMAs in the coral *Pocillopora acuta*, including their location, subcellular structure, transmission, community composition, and functional potential (Figure S1, Table S1). *Pocillopora acuta* is an asexual brooder and sexual broadcast spawner coral (*27*). It acquires bacteria both vertically, from the mother colony, and horizontally, from the environment (*28*). Associated bacterial communities are often dominated by *Endozoicomonas* (*28*, *29*), and CAMAs have been detected in asexually produced larvae, suggesting their vertical transmission in this species (*28*). Using a combination of fluorescence *in situ* hybridization (FISH), confocal laser scanning microscopy (CLSM), and transmission electron microscopy (TEM), we show that CAMAs are located in the epidermal layer in tentacle tips of *P. acuta* polyps and are likely intracellular. CAMAs were precisely excised with laser capture microdissection (LCM) and community composition analysis showed they contain bacteria of the families Endozoicomonadaceae and Simkaniaceae, a member of the Chlamydiota phylum. Metagenomic analyses suggest Endozoicomonadaceae may be beneficial to its coral host, but also to Simkaniaceae present in adjacent CAMAs.

Results and discussion

CAMAs are located in *Pocillopora acuta*'s tentacles

Pocillopora acuta colonies were sampled from the field, reared in captivity over a three-year period, and asexually produced two successive generations of corals (Figure S2A). The two latter generations were used for this study and are hereafter referred to as generations F1 and F2. To investigate the location of CAMAs within *P. acuta* polyps, we first conducted whole-mount FISH with a universal bacterial probe (EUB338-mix). CLSM images showed that CAMAs are present at the tip of the tentacles (Figure 1A-C), but not anywhere else in the polyp. This was the case for all three genotypes we sampled (Figure S2A), C2_12, R2_8, and F1_6. Nematocysts, present in the tentacles, displayed significant non-specific binding, as evidenced by the simultaneous use of an antisense probe (Figure 1A). FISH and CLSM performed on polyp sections yielded the same pattern (Figure 1D-E). Analysis of adult offspring (F2 generation) samples also revealed a multitude of CAMAs at the tip of the tentacles (Figure 1F-G). *In situ* location in other cnidarians has failed to show any consistency across species: CAMAs have been found in the tentacles of *Exaiptasia diaphana* (*18*, *20*) and *Stylophora pistillata* (*21*), the mesenteries of *Acropora formosa* and *A. palmata* (*12*, *14*), and throughout all tissue types in *Acropora hyacinthus* (*22*). We observed between 0 and 10 CAMAs per polyp in *P. acuta*, and they measured between 20 and 80 µm in diameter, which is similar to other coral species (*21*, *23*). CAMAs were exclusively seen in the epidermis and are thus not in contact with Symbiodiniaceae cells (Figure 1E), which exist in the gastroderm. Therefore, CAMA bacteria are more likely to interact with the coral animal than with the Symbiodiniaceae in *P. acuta*. CAMAs have only been detected in the epidermis in *E. diaphana* (*18*, *20*) and *S. pistillata* (*19*), while they colonize the gastrodermis in *Pocillopora verrucosa* (*19*), *S. pistillata* (*14*, *15*, *19*), *Acropora aspera* (*15*) and *A. formosa* (*14*). This variability in location suggests that CAMAs may play distinct roles across coral holobionts. For example, gastrodermal CAMAs may interact with Symbiodiniaceae, while CAMAs in tentacles may bear defensive functions. Additionally, distinct infection mechanisms may exist and allow different bacterial taxa to colonize and aggregate in different tissues, *e.g.* tentacles may be a privileged site of contact between the coral and environmental bacteria.

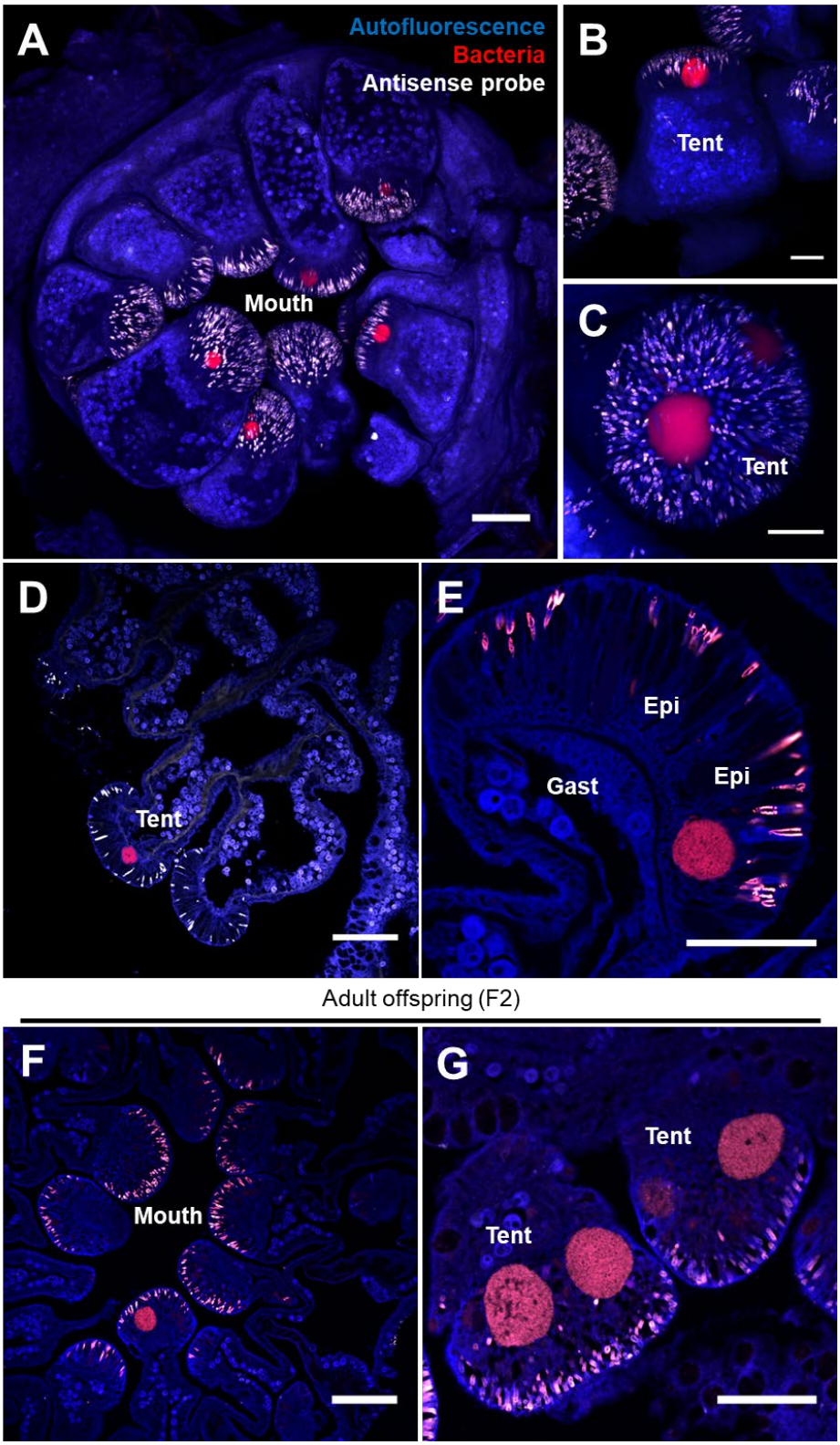

Figure 1: **CAMAs are located in the epidermis of *P. acuta*'s tentacles.** A-C: CAMA location by whole-mount FISH and CLSM in a whole adult polyp (A, maximum projection of a z-stack) and tentacles (B, C). D-G: CAMA location by FISH on sectioned adult polyps. Genotypes: R2_8 (A, B); F1_6 (C, E, F); C2_12 (D, G). Generation: F1 (A-E), F2 (F, G). Blue: autofluorescence; red: EUB338-mix probe (all bacteria); white: non-EUB probe (negative control). Gast: Gastrodermis; Epi: Epidermis; Tent: Tentacle. Scale bars: 100 µm for A, D, F; 50 µm for B, C, E, G.

While CAMAs are consistently found within cnidarian tissues, whether they are intra- or extracellular varies; they are intracellular in *E. diaphana* (*20*), extracellular in *A. palmata* (*13*), and both intra- and extracellular in *Porites compressa* (*23*). TEM images revealed that CAMAs are densely packed with bacteria (Figure 2A-B), which all show similar morphology and a highly visible nucleoid region (Figure 2C-D). TEM images in a previous study also showed the presence of nucleoid regions in CAMA bacteria (*20*). We observed a membrane encasing the CAMAs (Figures 2C-D and S3). This membrane resembled other coral cell membranes (Figure 2C, S3A and S3F; see for example the membrane of a cnidocyte, surrounding a nematocyst in Figure S3E) and was composed of lipid bilayers (Figure S3D. This suggests that CAMAs are intracellular in *P. acuta*. Because no eukaryotic organelle was observed inside CAMAs, it is possible that CAMAs are encased inside a vacuole, inside a coral cell. This is supported by the presence of eukaryotic content near CAMAs (*e.g.,* Golgi apparatus in Figure 2D), and the presence of a second membrane seemingly surrounding the CAMA membrane (Figures 2C, S3A, S3B, S3C, and S3F), and sometimes fusing with the CAMA membrane (Figures S3A and S3C). Alternatively, CAMAs may be extracellular, *i.e.* in between coral cells, and contained in a membrane of prokaryotic origin. DAPI staining combined with FISH showed the presence of one or more nuclei within CAMAs (Figure 2E-F), which is consistent with an intracellular nature. We also observed CAMAs that seemed to be disintegrating, with bacteria of a similar morphology seen outside the CAMAs (Figure 2F). This suggests that CAMAs are dynamic within the coral holobiont and may be able to form and deform after colonization of the animal. The mechanisms by which these bacteria aggregate, and whether the process is controlled by the host, bacteria, or both, require further investigation.

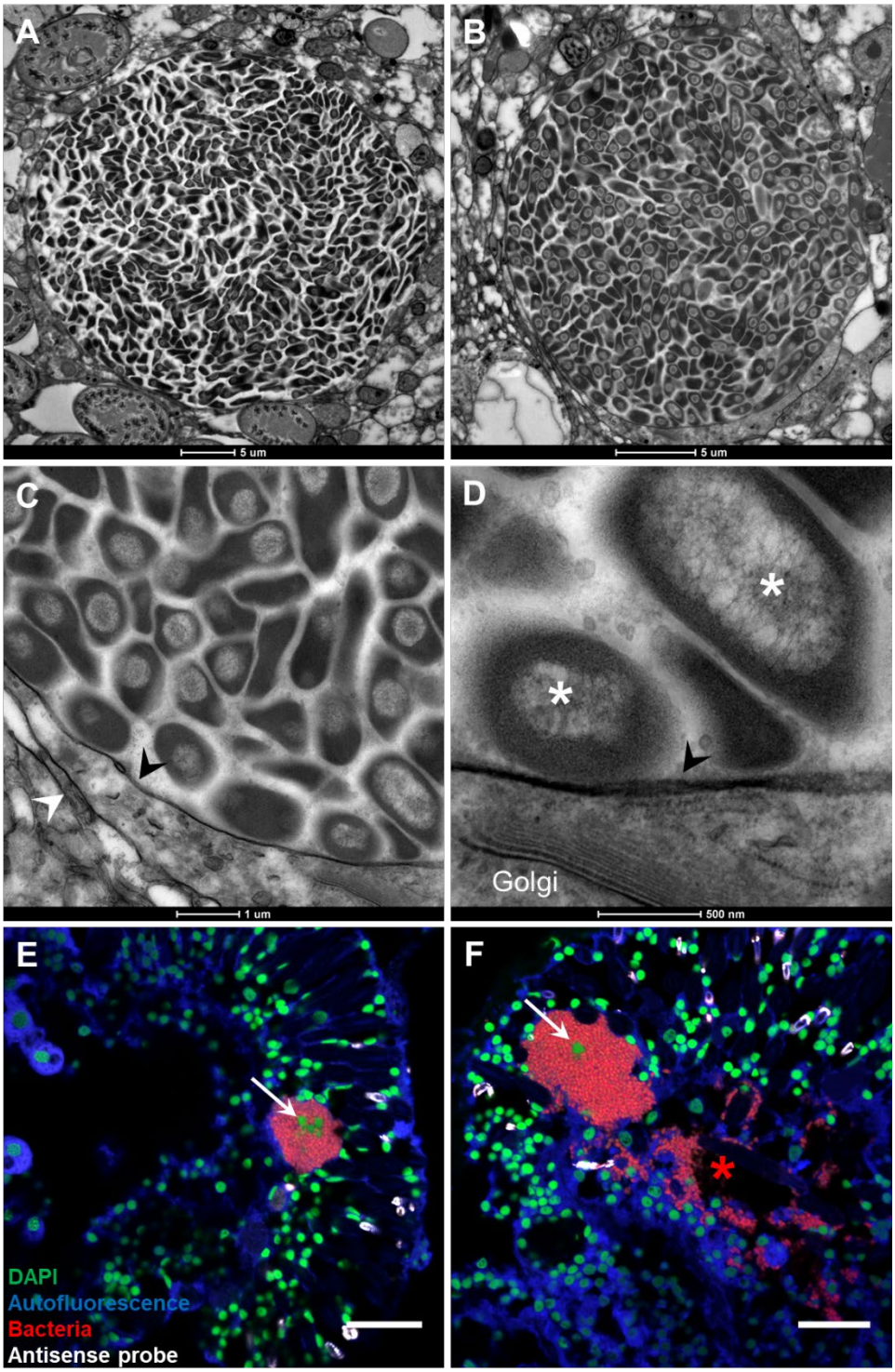

Figure 2: CAMAs may be intracellular. A-D: Subcellular structure of CAMAs observed in TEM. Black arrowheads point at a possible membrane surrounding the CAMAs. White arrowheads point at possible coral cell membranes. White asterisks show dense nucleoid regions of bacteria. E, F: Combined FISH and DAPI staining observed by CLSM on sectioned adult polyps. White arrows point at host nuclei inside CAMAs. The red asterisk highlights a zone where bacteria are outside a CAMA. All images are from F1 generation samples. Genotype: C2_12 (A-E); F1_6 (F). Blue: autofluorescence; red: EUB338-mix probe (all bacteria); white: non-EUB probe (negative control), green: DAPI. Scale bars: 20 µm for E, F.

152

153 **Endozoicomonadaceae and Simkaniaceae are the main CAMA members**

154 Using laser capture microdissection (LCM), we specifically sampled CAMAs from two
155 successive generations (F1 and F2, Figure S2A) of adult coral colonies belonging to the
156 single F1_6 genotype (Figure S4A), that were raised entirely under controlled experimental
157 conditions. 16S rRNA gene metabarcoding was performed to investigate the taxonomy of the
158 bacteria contained in CAMAs. Sequencing statistics for this and the following metabarcoding
159 experiments are summarized in Table S2. Fourteen ASVs were detected, although only five
160 had a relative abundance above 0.1% in at least one sample (Table S3A). Four of those were
161 assigned to the *Endozoicomonas* genus, were present in all samples, and amounted to more
162 than 95% of all reads (Figure 3A). ASVs 01, 02, 03 showed >99% identity with each other
163 (Figure S5), suggesting that they might belong to one single *Endozoicomonas* strain, while
164 ASV04 was more distant (~96% identity to ASVs 01, 02, and 03) and is likely a separate
165 strain. *Endozoicomonas* is widespread in marine invertebrates, including corals (*19*, *24*), and
166 sometimes dominates assemblages of coral-associated bacterial communities (*16*, *21*, *28*, *30*).

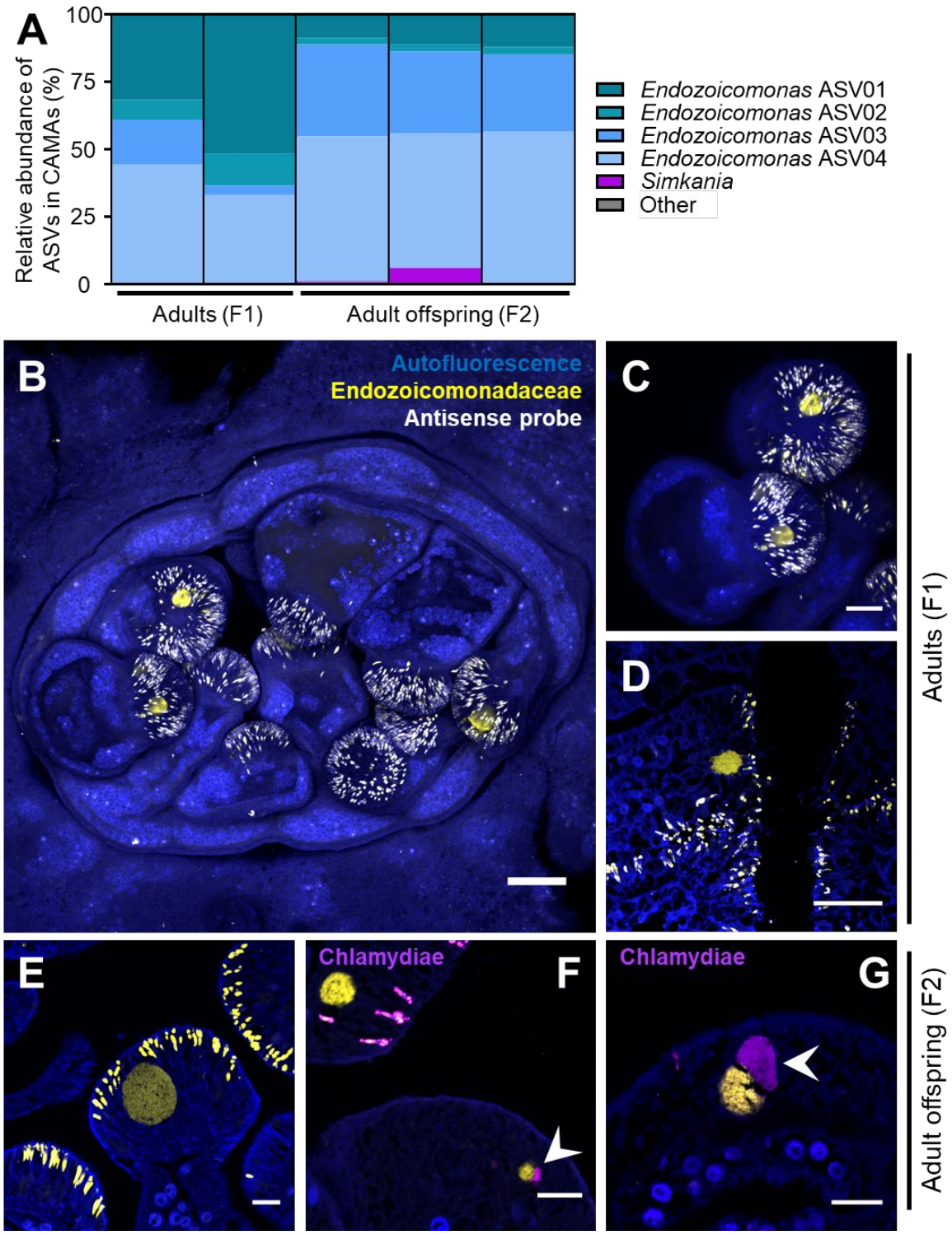

Figure 3: *Endozoicomonas* and *Simkania* are the two major taxa in CAMAs of the F1_6 genotype. A: Relative abundance of bacterial ASVs in CAMAs of adult polyps (F1_6 genotype, F1 and F2 generation) isolated by LCM. Each bar is a single sample. ASVs classified as 'others' represent less than 0.1% of the communities and cannot be seen in the figure. Detailed composition is available in Table S3. B-C: Location of *Endozoicomonas* in CAMAs by whole-mount FISH in a whole adult polyp (B, maximum projection of a z-stack) and tentacle (C). D-G: Location of *Endozoicomonas* and *Simkania* in CAMAs by FISH on sectioned adult polyps. White arrowheads point at *Simkania* CAMAs (purple), adjacent to *Endozoicomonas* CAMAs (yellow). All photos are from the F1_6 genotype. Generation: F1

(B-D), F2 (E-G). Blue: autofluorescence; yellow: End663 probe (Endozoicomonadaceae); magenta: Chls523 probe (chlamydiae). Scale bars: 100 µm for B; 50 µm for C, D; 25 µm for E, F.

FISH using a previously published *Endozoicomonas* probe (*16*) confirmed that *Endozoicomonas* are present in the CAMAs (Figure 3B-E). Simultaneous use of the *Endozoicomonas* probe with a universal bacterial probe showed complete co-localization (Figure S6A), confirming that *Endozoicomonas* are the main, and likely only, bacteria making up these CAMAs. FISH using the same probe has previously shown the presence of *Endozoicomonas* in CAMAs of *S. pistillata* and *P. verrucosa* (*16*, *19*, *24*), suggesting that *Endozoicomonas* commonly forms CAMAs in a range of coral species. Endozoicomonadaceae also form aggregates in the tissues of ascidians and fish (*31*, *32*), and inside the nuclei of mussel cells (*33*), suggesting that a conserved mechanism across this bacterial family may allow them to form aggregates inside animal hosts. Recently, Wada et al. (2022) also used LCM to sample single CAMAs and confirmed with 16S rRNA gene metabarcoding the presence of several *Endozoicomonas* strains within individual CAMAs of *S. pistillata* (*21*). Because we pooled many CAMAs in our samples and used a genus-wide *Endozoicomonas* FISH probe, we cannot say whether several strains co-occur in a single CAMA in *P. acuta*. The co-occurrence of two or more *Endozoicomonas* strains in CAMAs is puzzling and suggests they might show a degree of metabolic complementation or dependence on one another. It also raises the question of strain compatibility and whether all *Endozoicomonas* strains can co-exist within a CAMA.

The last ASV with a relative abundance over 0.5% was assigned to the *Simkania* genus (Figure 3A) and was present in just two samples, both belonging to the F2 generation. The Simkaniaceae family belongs to the Chlamydiota phylum, which are obligate intracellular bacteria infecting a vast array of animals and thriving as symbionts of protists in diverse environments (*34*). Simkaniaceae ASVs are often detected at high relative abundances in coral samples (*21*, *30*, *35*), as well as in Symbiodiniaceae cultures (*36*, *37*), although their roles, as well as the cells they infect, remain uninvestigated. In the present study, FISH using a previously published chlamydiae probe (*38*), simultaneously with the *Endozoicomonas* probe, showed *Simkania* inclusions are distinct from, but always adjacent to *Endozoicomonas* CAMAs in *P. acuta* (Figure 3F-G). Use of the chlamydiae probe combined with the universal bacterial probe confirmed *Simkania* occur alone within their inclusions (Figure S6B). While this is the first evidence of *Simkania* inclusions in corals, use of the Gimenez stain has previously suggested the presence of *Chlamydia*- or *Rickettsia*-like bacteria inclusions in *Acropora muricata* (*23*). Because chlamydiae are intracellular bacteria, the presence of *Simkania* inclusions reinforces the hypothesis that CAMAs are intracellular. Chlamydiae are known to co-occur with other bacteria in *Acanthamoeba* cells (*39*), although whether *Simkania* and *Endozoicomonas* share the same *P. acuta* cells is unclear. The close spatial proximity between them suggests the different bacterial species are interacting within the coral holobiont. Interestingly, both *Endozoicomonas* and chlamydiae are fish pathogens, forming cysts in the gills and skin and negatively affecting breathing (*40*). Such cysts strongly resemble CAMAs, although there is no record of *Endozoicomonas* and chlamydiae infecting the same fish hosts.

Horizontal (*Endozoicomonas*) and vertical (*Kistimonas*) transmission of CAMA bacteria

223 The presence of CAMAs and of all four *Endozoicomonas* ASVs in both the F1 and F2
224 generations shows that this symbiotic association persists across generations and may be
225 vertically transmitted to asexually produced larvae. However, we were not able to observe
226 any CAMAs in freshly released F2 larvae (Figure S7). Because we may miss CAMAs when
227 only looking at sections, we further conducted 16S rRNA gene metabarcoding on whole
228 larvae of the F1_6 genotype to assess the presence of *Endozoicomonas* (Table S2). Of the 179
229 ASVs detected in these samples, none were assigned to the *Endozoicomonas* genus or to the
230 Endozoicomonadaceae family (Table S4), suggesting the larvae do not contain CAMAs or
231 *Endozoicomonas*. This contrasts with previous data showing the presence of CAMAs and of
232 *Endozoicomonas* in larvae of distinct *P. acuta* genotypes sampled from Orpheus Island (*28*).

233 To further investigate the vertical transmission in the genotypes sampled by Damjanovic et
234 al. (*28*) from Orpheus Island, we processed stored larval samples from that study belonging to
235 the two genotypes OI2 and OI3 (Figure S2B, Table S1). FISH confirmed the presence of
236 CAMAs in the larvae (Figure 4A-B), which also hybridized to the *Endozoicomonas* FISH
237 probe (Figure 4C-D). We sampled CAMAs from both genotypes using LCM (Figures S1 and
238 S4B) and examined their bacterial composition with 16S rRNA gene metabarcoding (Tables
239 S2 and S3B). In the OI2 genotype, two ASVs assigned to *Kistimonas* made up more than
240 80% of the reads (Figure 4E). Only one nucleotide differs between the two ASVs, suggesting
241 they may be a single *Kistimonas* strain. These two ASVs were also present in the OI3
242 genotype at a lower relative abundance (Figure 4E). Other ASVs in these samples may be
243 contaminants that were too abundant to be removed bioinformatically as initial biomass was
244 low in these samples. *Kistimonas* belongs to the Endozoicomonadaceae family, is
245 phylogenetically close to *Endozoicomonas*, and has previously been isolated from marine
246 invertebrates (*41*). Intriguingly, *Kistimonas* were not detected by Damjanovic et al. (2020),
247 although the larvae were collected at the same time and from the same parent colonies.
248 Therefore, we re-analyzed data from Damjanovic et al (*28*), comprised of whole larvae,
249 recruits and adults, using an updated SILVA database (v138 vs v132). Among the seven
250 *Endozoicomonas* ASVs previously reported, three were assigned to *Kistimonas* in our new
251 analysis (Figure 4F), two of which perfectly matched the two ASVs found in our excised
252 CAMA samples (ASVs 09 and 10). This could be due to the relatively recent description of
253 *Kistimonas*, which might not have been included in v132 of the SILVA database.
254 Importantly, only *Kistimonas,* and no *Endozoicomonas*, were detected in whole larvae and
255 recruits, while both genera are present in the later life stages (Figure 4F).

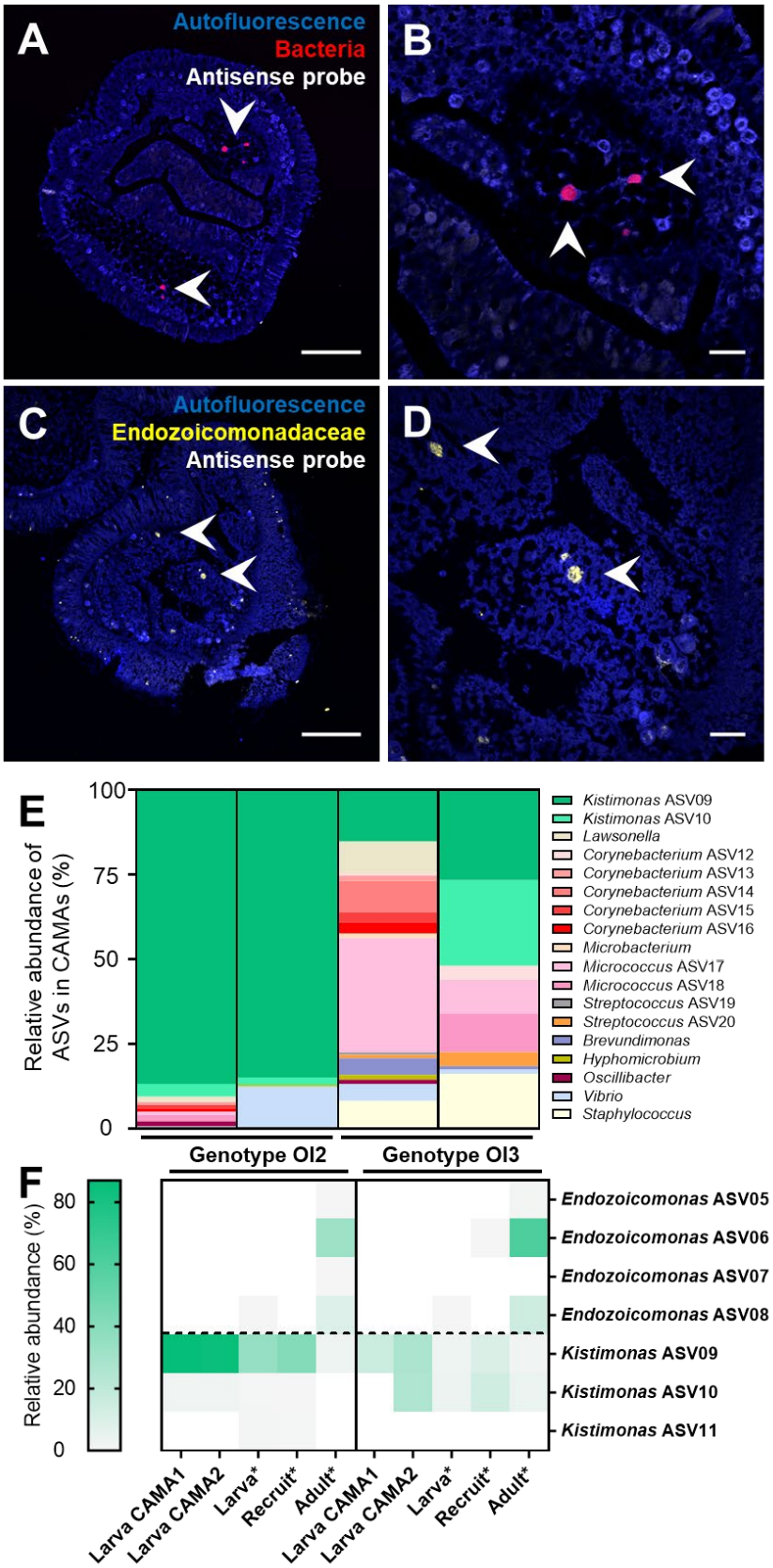

Figure 4: *Kistimonas* is the major CAMA taxon in larvae of the OI2 and OI3 genotype and is vertically transmitted. A-D: CAMA location by FISH on sectioned larvae. Arrowheads point at CAMAs. Blue: autofluorescence; yellow: End663 probe (Endozoicomonadaceae); red: EUB338-mix probe (all bacteria). Scale bars: 100 µm for A, C;

20 µm for B, D. E: Relative abundance of bacterial ASVs in CAMAs of larvae (OI2 and OI3 genotypes) isolated by LCM. Each bar is a single sample. F: Relative abundance of *Kistimonas* and *Endozoicomonas* ASVs in CAMAs, whole larvae, recruits, and adults of the OI2 and OI3 genotypes. Columns with an asterisk (Larva, Recruit, Adult) contain data re-analyzed from Damjanovic et al. (*28*).

Taken together, these data show that *Kistimonas* are vertically transmitted in genotypes OI2 and OI3, while *Endozoicomonas* are acquired horizontally in genotypes F1_6, OI2 and OI3, likely at the recruit or adult stage. Whether the vertical transmission of *Kistimonas* involves CAMAs or individual bacteria able to establish a new population remains unknown.

Interestingly, the *Simkania* ASV detected in the F2 adult samples of the F1_6 genotype was detected in every whole larval sample of the same genotype, with relative abundances ranging from 0.2 to 0.9% (Table S4). This suggests that *Simkania* is efficiently transmitted to asexually produced larvae in the F1_6 genotype, and that CAMAs containing *Simkania* may have been missed when sampling the F1 adult samples. *Endozoicomonas* is likely to have an environmental life stage, while *Kistimonas* and *Simkania* may be obligate symbionts. This is expected for *Simkania*, as all described chlamydiae depend on a eukaryotic host for propagation, whereas *Kistimonas*, just like *Endozoicomonas*, has been cultured *in vitro* from several marine hosts (*41*). These different transmission modes are likely to have impacted genome evolution and host-microbe interactions.

CAMA bacteria belong to undescribed *Endozoicomonas* and *Simkania* species

To assemble the genomes and investigate the functional potential of uncultured CAMA bacteria, metabarcoding samples were used for shotgun sequencing and metagenomic analyses. Insufficient material was available of the OI2 and OI3 genotype larval samples to conduct metagenomics, thus only F1_6 samples were processed. Three metagenome-assembled genomes (MAGs) were recovered, one from the F1 generation sample, and two from the F2 generation samples. MAG data are detailed in Table 1. One MAG from each generation was assigned to *Endozoicomonas* (Pac_F1 and Pac_F2a), while the second MAG from the F2 generation sample was assigned to *Simkania* (Pac_F2b), consistently with our metabarcoding results. Pac_F1 and Pac_F2a genome sizes are similar with other *Endozoicomonas* genomes (*10*, *25*, *26*), while Pac_F2b is significantly smaller, which is consistent its classification as a member of the intracellular chlamydiae (*42*). Pac_F2b's genome size is like other animal-associated chlamydiae (1-1.5 Mb), whereas protist-associated chlamydiae usually possess larger genomes (2-3 Mb) (*42*, *43*). This may be due to the higher stability of a multicellular organism and relaxed selection pressure on more bacterial genes leading to deeper genomic erosion (*44*).

299 **Table 1: Summary of the three metagenome-assembled genomes recovered in CAMAs**
300 **sampled from adults of the F1_6 genotype.**

MAG	Pac_F1	Pac_F2a	Pac_F2b
Genus assigned by GTDB-Tk	*Endozoicomonas*	*Endozoicomonas*	*Simkania*
Size (bp)	6,938,003	5,907,264	1,247,175
Coverage (mean value)	2,063.5x	3,539x	500.7x
Completeness (%)	94.25	88.76	88.02
Contamination (%)	1.65	2.48	0.21
G+C content (%)	49	51.3	43.9
N50	6,009	6,281	41,614
L50	314	248	12
Number of contigs	1,770	1,443	47
Number of coding sequences (CDS)	5,947	5,256	1,074
Number of rRNAs	10	5	1
Number of tRNAs	41	50	34

301

302 Because *Endozoicomonas* are so widespread in marine invertebrates, and particularly corals
303 (*24*), we assessed their taxonomic placement. A full length 16S rRNA gene (1,536 bp) was
304 recovered from Pac_F2a and two non-overlapping partial sequences were retrieved from
305 Pac_F1 (401 bp and 1,129 bp respectively). ASV01 and ASV04 had 100% identity with
306 Pac_F1 (1,129 bp fragment) and Pac_F2a, respectively. A phylogenetic tree based on more
307 than 1,000 Endozoicomonadaceae 16S rRNA gene sequences shows our two strains cluster
308 together with other *Endozoicomonas* strains isolated from *P. damicornis* (*45*) and *S. pistillata*
309 (*16*, *21*) corals (Figure S8). This indicates a certain degree of host specialization in the
310 Pocilloporidae family (*21*). However, this contrasts with other animal-*Endozoicomonas*
311 associations, with host phylogenies rarely matching symbiont phylogenies (*21*, *25*, *26*), which
312 is further evidence that *Endozoicomonas* possess an environmental life stage, are not obligate
313 symbionts, and are not undergoing genome streamlining. Pac_F1 and Pac_F2a form a well-
314 supported clade together with *E. acroporae* (isolated from *Acropora* sp. corals), and to a
315 lesser extent *E. atrinae,* and *E. elysicola*, although they likely represent a new species (~95-
316 97% identity with *E. acroporae*, *E. atrinae,* and *E. elysicola*). A phylogenetic tree based on
317 120 bacterial gene markers from 15 *Endozoicomonas* genomes confirmed this trend (Figure
318 5, Table S5A). Two clades were observed, with coral-associated *Endozoicomonas* present in
319 both. Pac_F1 and Pac_F2a clustered together and were closest to two *Endozoicomonas*
320 genomes recently sequenced from *S. pistillata* CAMAs (*21*). Average nucleotide identity
321 (ANI) and average amino acid identity (AAI) showed strong similarity between Pac_F1 and
322 Pac_F2a (98% and 96%, respectively) (Figure S9A-B). The highest values were obtained
323 with *Stylophora*- (ANI 83%, AAI 78%) and *Acropora*-associated (ANI 80%, AAI 70%)
324 *Endozoicomonas*. Overall, this data suggests that Pac_F1 and Pac_F2a are two different
325 strains that belong to a novel, undescribed *Endozoicomonas* species, which is closely related
326 to other Pocilloporidae-associated *Endozoicomonas*.

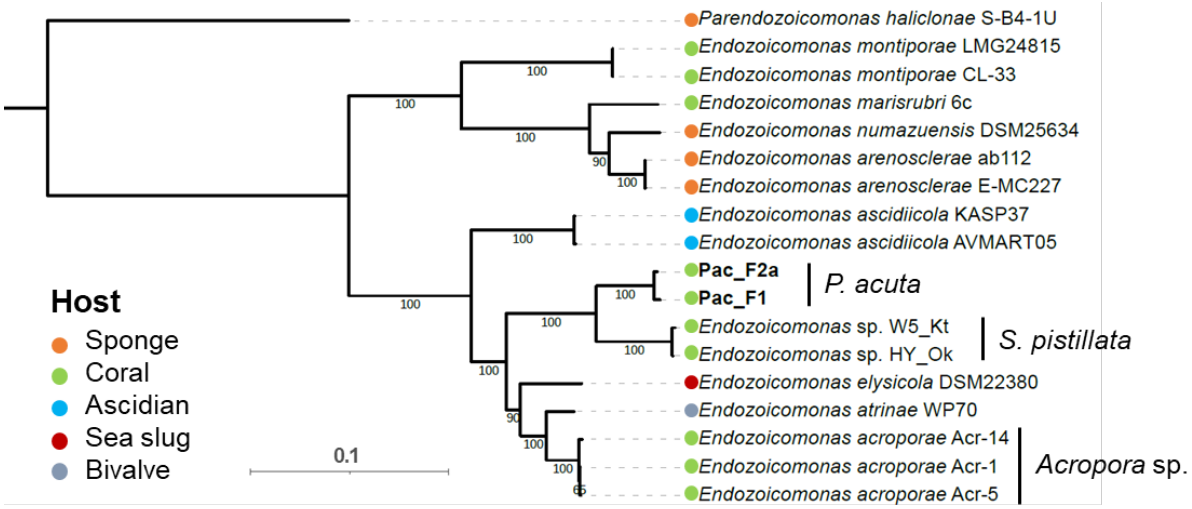

Figure 5: Two MAGs sequenced from CAMAs belong to new species of *Endozoicomonas.* Maximum-likelihood phylogeny of *Endozoicomonas* based on 120 marker genes, 15 *Endozoicomonas* genomes, and one *Parendozoicomonas* genome (outgroup). Bootstrap support values based on 1000 replications are provided. Additional data on the reference genomes is available in Table S5A.

Only a partial 16S rRNA sequence (662 bp) was retrieved for Pac_F2b, so we created a phylogenetic tree based on 15 conserved marker genes, along with other chlamydiae and Simkaniaceae genomes (Table S5B). As suggested by the GTDB-tk classification, Pac_F2b falls in the same clade as *Simkania negevensis* and is closest to a MAG isolated from the coral *Cyphastrea* sp. sampled in subtropical Lord Howe Island (Australia) (Figure 6 and S10). AAI was highest with the MAG isolated from *Cyphastrea* sp. (75%) and *S. negevensis* (64%) (Figure S11). ANI values were too low (<75%) to be reliable. This suggests that Pac_F2b in an undescribed species that belongs to the *Simkania* genus.

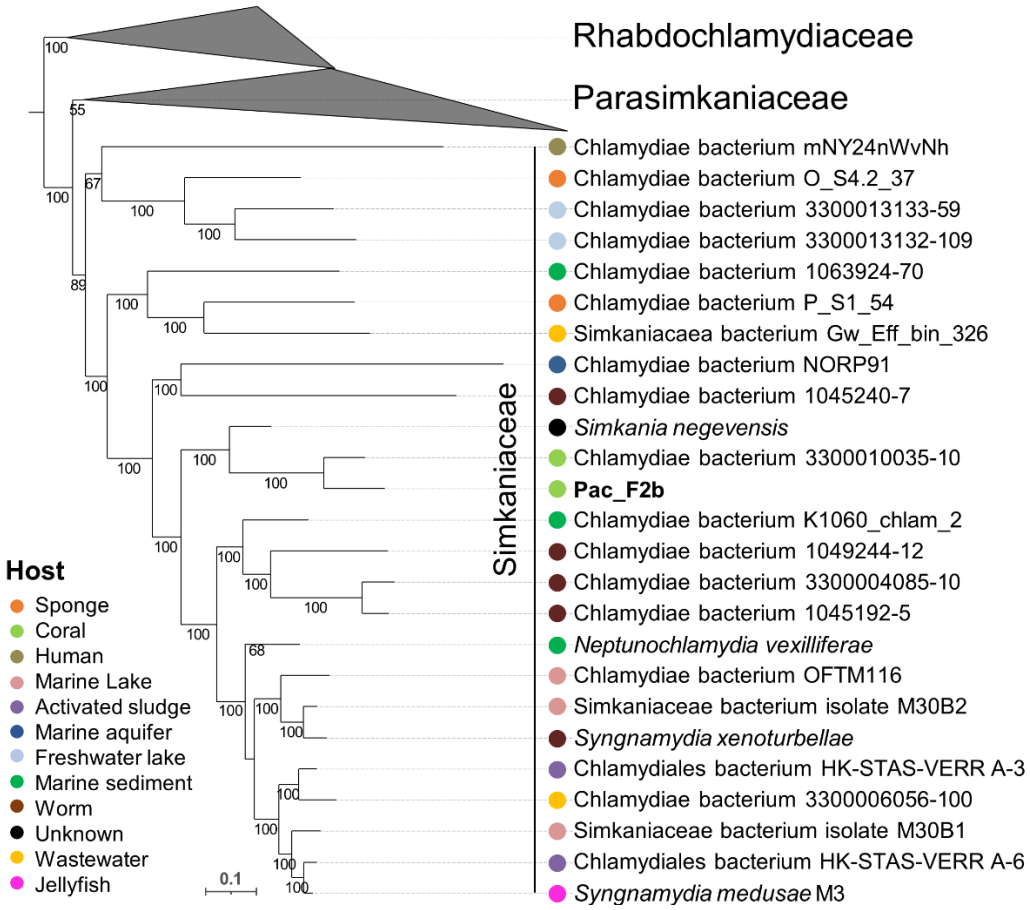

Figure 6: One MAG sequenced from CAMAs belongs to a new species of *Simkania.* Chlamydial maximum-likelihood phylogeny based on 15 conserved non-supervised orthologous groups (NOGs) in 24 Simkaniaceae, 8 Parasimkaniaceae, and 21 Rhabdochlamydiaceae (outgroup) genomes. Bootstrap support values based on 1000 replications are provided. Additional data on the reference genomes is available in Table S5B. A tree containing the entire chlamydial phylogeny is available in Figure S10.

Endozoicomonas may provide vitamins to their coral host and protect it against pathogens

The functional potential of the three MAGs recovered from CAMAs of *P. acuta* was investigated through KEGG, RAST and InterProScan analyses and is summarized in Figure 7. The completeness of selected KEGG pathways is provided in Figure S12. Detailed Prokka and eggNOG-mapper annotations are provided in Dataset S1. We first focused on genes for host-symbiont interactions in the two *Endozoicomonas* MAGs, Pac_F1 and Pac_F2a (Figure 7A). First, a total of nine secondary metabolites were bioinformatically predicted across the two MAGs (Table S6). Only one showed similarity with a known biosynthetic gene cluster, which encodes the antibacterial protein rhizomide (*46*). The predicted secondary metabolites were also annotated by Prokka and were identified as antimicrobial proteins, such as gramicidin, tyrocidine, and microcin. Additionally, several near-complete secretion systems were detected, including a type II, type III, and type VI (T2SS, T3SS, and T6SS, respectively). Note that a T6SS was only detected in Pac_F1. All three secretion systems may be involved in host infection and/or host cell entry, although their effectors remain unknown. While T2SSs and T3SSs are routinely observed in *Endozoicomonas* (*10*, *21*, *25*), this is only

366 the second report of a T6SS in a coral-associated *Endozoicomonas* (*21*). These secretion
367 systems, along with the secondary metabolites, may play a role in host colonization,
368 outcompeting other bacteria present in the coral holobiont, as well as preventing other
369 bacteria from invading CAMAs and promoting the maintenance of these structures. High
370 numbers of genes encoding eukaryotic-like proteins were also detected (Table S7), including
371 more than 100 ankyrin-repeat proteins in each MAG. These are also found in other
372 *Endozoicomonas* (*10*, *21*, *25*), are often effectors of secretion systems, and hypothesized to
373 facilitate host-bacteria interactions by mediating bacterial protein-eukaryotic host protein
374 interactions (*47*). Finally, both MAGs possess the complete machinery to synthesize type IV
375 pili (T4P). While never discussed in earlier publications, the pathway is also complete in
376 other *Endozoicomonas* genomes (Table S8). However, the gene encoding the actual pilus,
377 *pilA*, is missing in *Parendozoicomonas haliclonae*. T4P are involved in bacterial motility and
378 may play a role in host colonization, but they also support the formation of biofilms and
379 aggregates (*48*). Specifically, *pilY1* encodes an adhesin that may be involved in host
380 recognition. Therefore, we hypothesize that T4P are involved in the formation and
381 maintenance of *Endozoicomonas* CAMAs. As *E. acroporae* forms CAMA-like structures in
382 culture when incubated with dimethylsulfoniopropionate (DMSP) (*10*), it is possible that
383 metabolic signals, such as DMSP, regulate the *in hospite* expression of T4P synthesis genes
384 and in turn the formation of CAMAs.

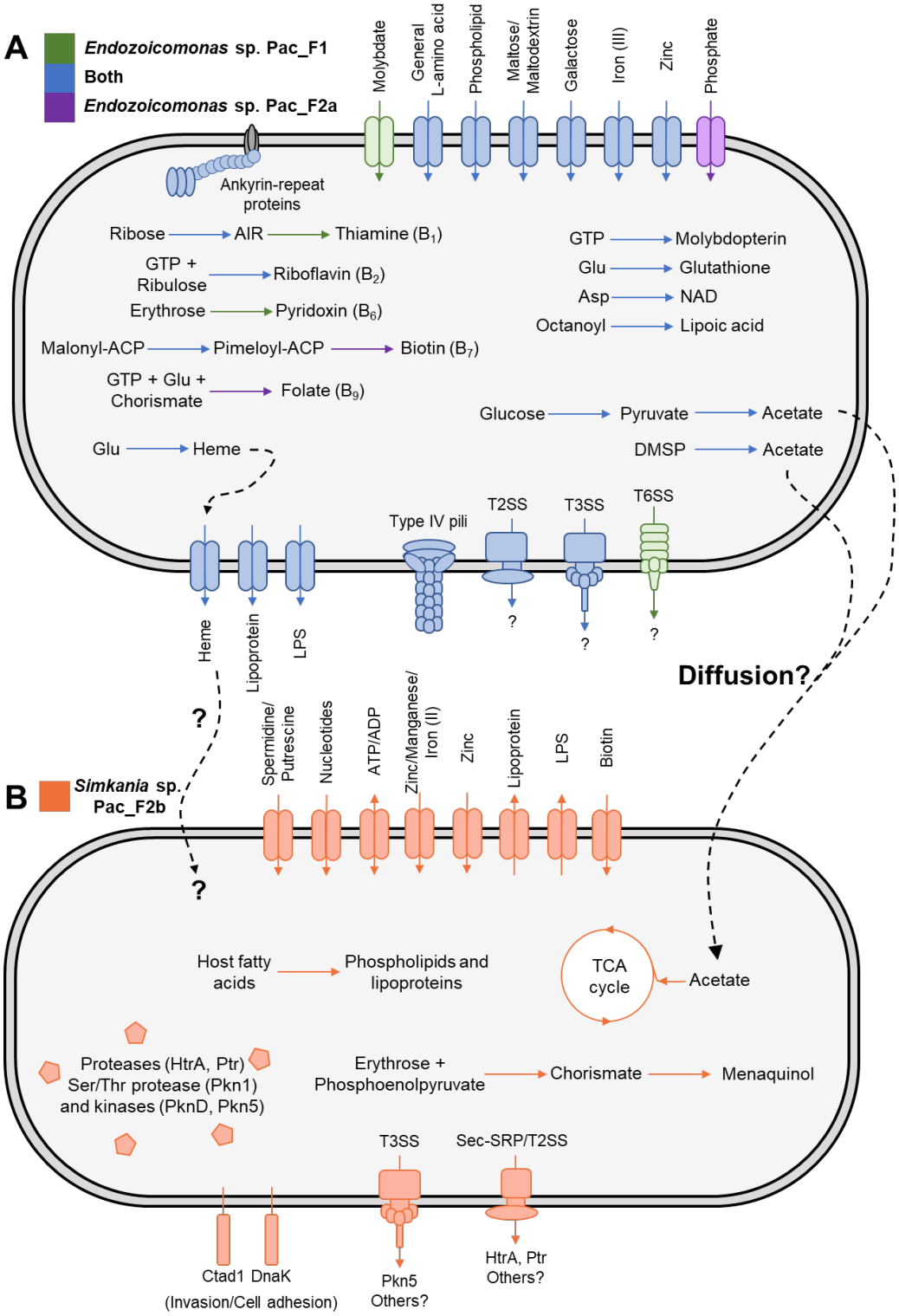

Figure 7: Overview of the genomic potential of the three MAGs recovered from *P. acuta* CAMAs, *Endozoicomonas* sp. Pac_F1 (A) and Pac_F2a (A), and *Simkania* sp. Pac_F2b (B). The pathways represented here are either complete or only lack one gene (Figure S12). Dashed arrows represent hypothetical interactions between *Endozoicomonas* and *Simkania* (Figure S13). Sec-SRP: Sec translocase and signal recognition particle pathway; T2SS: type II secretion system; T3SS: type III secretion system; T6SS: type VI secretion system; TCA: tricarboxylic acid; AIR: aminoimidazole ribotide; DMSP: dimethylsulfonioproprionate; LPS: lipopolysaccharide; NAD: nicotinamide adenine dinucleotide.

394 Both *Endozoicomonas* MAGs encode pathways for the synthesis of several B vitamins,
395 including thiamine (B_1), riboflavin (B_2), pyridoxin (B_6), biotin (B_7) and folate (B_9). Both
396 corals and Symbiodiniaceae cannot produce B vitamins and are dependent on external
397 sources. B vitamin provisioning from bacteria has been shown in hematophagous insects (*49*)
398 and marine algae (*50*), and B vitamin synthesis pathways have been detected in other
399 *Endozoicomonas* (*21*, *25*, *26*). Further, thiamine and pyridoxin synthesis was upregulated in
400 *E. marisrubri* incubated with coral host extracts (*25*). It is therefore likely that
401 *Endozoicomonas* provide B vitamins to its coral host and/or Symbiodiniaceae, although how
402 these vitamins are exported remains elusive. As in other *Endozoicomonas* (*21*, *26*), we
403 identified a wide range of ABC transporters, notably for zinc, iron, amino acids, maltose,
404 galactose, or heme. *Endozoicomonas* MAGs recently obtained from CAMAs in *S. pistillata*
405 showed the presence of genes involved in phosphate transport, metabolism, and storage, as
406 well as a phosphotransferase system, which led to the hypothesis that CAMAs may play a
407 role in phosphate cycling and balance in the coral holobiont (*21*). Here, we only found genes
408 encoding an ABC transporter for phosphate and a cellobiose phosphotransferase system, but
409 no other genes that may suggest a role in phosphate metabolism. Finally, we also detected
410 several pathways for the synthesis of molecules involved in antioxidant processes, including
411 lipoic acid, heme (incorporated into hemoproteins, such as catalases and peroxidases), and
412 glutathione. Antioxidant synthesis is a trait of high interest in the design of bacterial
413 probiotics for the mitigation of coral bleaching (*51*, *52*), as the overproduction of reactive
414 oxygen species by Symbiodiniaceae is thought to be the main driver of coral bleaching (*53*).
415 Thus, *Endozoicomonas* is likely to be a beneficial symbiont for its host, potentially providing
416 vitamins, antioxidants, and antimicrobial molecules. As the relative abundance of
417 *Endozoicomonas* often diminishes during thermal stress (*54*), the loss of its associated
418 benefits may exacerbate the negative effects of marine heatwaves, including the increase in
419 abundance of opportunistic pathogens.

420

421 ***Simkania* may benefit from *Endozoicomonas* for growth and survival**

422 The *Simkania* MAG Pac_F2b exhibited reduced metabolic abilities (Figure 6B), which is
423 consistent with other chlamydiae and their intracellular lifestyle (*42*). It possesses all the
424 genes for glycolysis, the tricarboxylic acid (TCA) cycle, and the pentose phosphate pathway.
425 Additionally, Pac_F2b encodes the genes for the transformation of acetate into acetyl-CoA,
426 potentially enabling this *Simkania* strain to use acetate as a substrate for the TCA cycle and as
427 a carbon source (Figure S12). In a similar way, sponge-associated chlamydiae can use acetoin
428 as an energy source (*55*). Interestingly, both *Endozoicomonas* MAGs can produce acetate via
429 two pathways (Figure S13): (i) through the conversion of acetyl-coA, produced by pyruvate
430 oxidation, into acetate; (ii) through the conversion of DMSP into methylmercaptopropionate
431 by DmdA, and subsequently into acetate, although not all the genes in this pathway were
432 found. Because acetate can freely diffuse across membranes, we hypothesize that excess
433 acetate produced by *Endozoicomonas* present in neighboring CAMAs can be acquired by
434 *Simkania* for its own growth. Like most chlamydiae, Pac_F2b lacks the genes for *de novo*
435 nucleotide biosynthesis and can only synthesize few amino acids (alanine, aspartate, and
436 glutamate), but possesses a complete shikimate pathway. Other hallmark chlamydial genes
437 were detected, including a T3SS, adhesins, ATP/ADP translocases, and virulence effectors
438 such as proteases and kinases (Figure 7B). An orthogroup analysis revealed that 518 of
439 Pac_F2b's genes (almost 50%) were present in more than 90% of all publicly available
440 chlamydial genomes and 204 genes were shared by two other *Simkania* genomes (Dataset
441 S2). This suggests that Pac_F2b's overall genomic content is similar to that of other

chlamydiae. The only vitamin/cofactor Pac_F2b can produce is menaquinol, which is common in chlamydiae (*55*) and plays a role in bacterial growth in *C. trachomatis* (*58*). Pac_F2b may import other vitamins from host cells or adjacent *Endozoicomonas* through vitamin transporters such as BioY, present in Pac_F2b. Unlike most other chlamydiae (*55*), it does not encode the heme biosynthesis pathway. Although the genome sequence of the Pac_F2b MAG is not complete, it is highly unlikely that the full heme pathway is present. None of the heme biosynthesis genes could be detected in the closest relative, the *Cyphastrea*-associated Chlamydiae bacterium 3300010035-10, either and these genes are scattered at six loci throughout the chromosome in *Simkania negevensis*. Heme is essential for many cellular processes, including antioxidant functions, as well as infectivity and virulence in many pathogens. For example, heme production is essential for the intracellular survival of *Brucella abortus* (*59*). Heme-deficient strains of *Staphylococcus aureus* show reduced tissue colonization in mice compared to wild-type strains (*60*), although they persist more easily intracellularly, perhaps because of lower toxin production (*61*). The specific role of heme in chlamydiae remains unknown, although mutations in the *hemG* gene, involved in heme synthesis, were associated with increased infectivity in *C. trachomatis* (*62*). In *P. acuta*, *Simkania* may acquire heme from the host, or from neighboring *Endozoicomonas*, which possess the complete heme biosynthesis pathway and an ABC transporter for heme. However, no heme import system was detected in Pac_F2b and how heme would cross CAMA or *Simkania*'s membranes remains unclear. Alternatively, it is possible that the lack of heme synthesis in Pac_F2b makes this strain less virulent and able to persist in inclusions in *P. acuta*, similarly to *S. aureus*.

In summary, it is possible that *Simkania* is reliant on *Endozoicomonas* for acetate, which *Simkania* can use to stimulate its growth, as well as heme and vitamins. Interactions between bacteria within a holobiont are known in insects where endosymbiotic bacteria share a single cell, *e.g. Buchnera/Serratia* in the aphid *Cinara cedri* (*63*) or *Tremblaya/Moranella* in the mealybug *Planococcus citri* (*64*) (note that in the latter, *Moranella* resides within *Tremblaya* cells). In both cases, the synthesis of some amino acids requires a patchwork of genes present in one symbiont or the other. The potential metabolic interactions we uncovered in *P. acuta* may explain the spatial proximity of the two types of CAMAs, and why we did not observe any *Simkania* CAMAs on their own. Nevertheless, the presence of *Simkania* without *Endozoicomonas* in larvae samples of the F1_6 genotype (Table S4) suggests that potential *Endozoicomonas*-*Simkania* interactions are not necessary for *Simkania* survival. However, *Simkania* was in low relative abundance in these samples (< 1%) and was not observed forming CAMAs (Figure S6). This suggests that either its growth is impeded, or that it is instead dependent on host-produced acetate and/or heme. In the latter hypothesis, *Simkania* would not be restricted to aggregate around *Endozoicomonas* and may be able to source host compounds in different coral cell types, hence the absence of CAMAs in these samples. Simkaniaceae and Endozoicomonadaceae co-occur in corals of the *Acropora, Stylophora* and *Pocillopora* genera (*21*, *30*, *35*, *65*). These corals cover the complex and robust clades within the Scleractinia. The interactions suggested by our genomic data may thus be widespread in coral holobionts and should be studied further.

Suitability for the development of probiotic treatments

We provided here a comprehensive overview of the biology of CAMAs in the coral *P. acuta*. CAMAs are located at the tip of the tentacles, contain members of the Endozoicomonadaceae and Simkaniaceae families, exhibit mixed-mode transmission, and genomic data suggests they bear a wide range of functions. This further establishes *Endozoicomonas* as a widespread

490 mutualistic symbiont of corals and we provide additional evidence that *Endozoicomonas*
491 symbionts vary in their genetic potential and life strategies. This is also the first evidence,
492 both spatial and genomic, for CAMAs constituted by different microbes suggesting a
493 relevance for interactions between CAMAs for CAMA and host functioning.

494 Our study provides crucial insights into the biology of these endosymbionts, which may be
495 used for the development of probiotics for the mitigation of coral bleaching. Specifically,
496 several aspects suggest the strong suitability of *Endozoicomonas*. First, they may have several
497 beneficial functions, including an antagonistic effect towards other bacteria, and the
498 production of B vitamins and antioxidants. Second, our data strongly suggests that
499 *Endozoicomonas* is intracellular, and shows efficient spreading from colony to colony, as all
500 individuals from two successive generations were infected. This stability is essential if
501 probiotics are to lastingly colonize coral hosts and successfully spread through coral
502 populations, without having to repeatedly inoculate probiotics (*4*, *51*). However, as they are
503 not located in the gastrodermis in *P. acuta*, it is unclear whether any beneficial function
504 would affect Symbiodiniaceae (*e.g.*, through the transfer of antioxidants to the gastrodermis)
505 and thus be relevant for coral bleaching. *Kistimonas*, through its vertical transmission, would
506 also be a candidate of choice and needs to be investigated further. Future work should focus
507 on culturing these strains. *Kistimonas* may be more challenging to culture if vertical
508 transmission has resulted in some genome degeneration and higher host dependence than
509 *Endozoicomonas*. Cultured isolates will allow us to perform targeted analyses to confirm
510 their functional potential, as well as reinoculation experiments to trial their potential
511 efficiency as probiotics.

512

513 **Material and methods**

514 **Coral husbandry and aquarium conditions**

515 For the F1_6, C2_12 and R2_8 genotypes, 3 parent colonies were collected from 3 reefs in
516 July 2017, Feather reef (colony ID: 202B F1_6; Lat: -17.51854 Long: 146.38940), Coates
517 reef (colony ID: 307A C2_12; Lat: -17.18850 Long: 146.37180) and Rib reef (colony ID:
518 211D R2_8; Lat: -18.47122 Long: 146.87360), along the central Great Barrier Reef in
519 Australia (Table S1, Figure S2A). The colonies were brought back to the National Sea
520 Simulator at the Australian Institute of Marine Science (AIMS) in Townsville, QLD,
521 Australia. Parent colonies were kept in aquaria, running an annual temperature profile based
522 on historical SST records (1985-2012) for the mid-shelf Central GBR obtained from National
523 Oceanic and Atmospheric Administration (NOAA) Coral Reef Watch (CRW) daily global
524 5km (0.05 degree) satellite product v3.1 (Figure S14). The annual temperature profile
525 targeted an annual average temperature for the central GBR of 26.14 ± 0.23°C (± 1 daily SD);
526 an annual range from 23.16 to 29.88°C, with a summer heat stress accumulation equivalent to
527 0.95°C-weeks (NOAA CRW v3.1). Corals were maintained in a semi-recirculating system
528 with a daily turnover of ~300%, and circulation within tanks was achieved with two pumps
529 (Maxspect 350 series) generating gyres at opposite ends of the tanks; partial CO_2 pressure
530 was maintained at a daily average value 464 ± 40.94 ppm (± 1 daily SD) allowing for diurnal
531 variability driven by natural photosynthesis and respiration cycles; daily light cycles were set
532 to match local conditions with daily and seasonal variations in solar intensity and lunar cycles
533 to maintain synchronicity of planulation patterns for *P. Acuta* colonies. From July to October
534 2017, fully developed F1 larvae, released asexually by the parent colonies, were collected
535 with flowthrough collection devices on the outflow of isolation tanks and settled separately
536 under constant temperature-controlled flow-through seawater onto pre-conditioned aragonite

537 settlement plugs. From September 2018 to February 2019, F2 larvae, generated asexually by
538 the F1 colonies (12-16 months old at that point), were collected with flow-through collection
539 devices on the outflow of isolation tanks and settled separately under constant temperature-
540 controlled flow-through seawater onto pre-conditioned aragonite settlement plugs.
541 Throughout their life, both generations were kept in the same mesocosm systems under the
542 same annual conditions as parent colonies.

543 For the OI2 and OI3 genotypes, parent colonies were collected from Orpheus Island, in the
544 central Great Barrier Reef in Australia (Table S1, Figure S2B), and maintained and sampled
545 as previously described (*28*).

546

547 **Sampling**

548 For adult colonies, small coral fragments were snapped off the colonies with forceps, rinsed
549 with 0.22 µm FSW and placed in cryovials. Forceps were rinsed in 80% ethanol between
550 sampling each colony. Two to three fragments per colony were placed in one cryovial. Three
551 colonies were sampled per genotype. For FISH and LCM, branches were fixed for 24 hrs in
552 4% paraformaldehyde (PFA) in 0.22 µm-filtered seawater (FSW), rinsed twice in FSW, and
553 stored in 50% ethanol-PBS at -20°C. For transmission electron microscopy (TEM), branches
554 were fixed for 24 hrs in 4% PFA + 0.1% glutaraldehyde in FSW, rinsed once in FSW, once in
555 PBS 1X, and stored in PBS 1X at 4°C. Following fixation, coral branches were decalcified in
556 EDTA 10%. EDTA was renewed every two days, and samples were kept at 4°C on a rotating
557 wheel, until there was no skeleton left (around two weeks). Samples were then rinsed in PBS
558 1X, and stored at 4°C in PBS 1X.

559 Larvae were sampled as previously described (*28*). Briefly, before planulation, colonies were
560 maintained in individual acrylic aquaria that received indirect natural sunlight and coarsely
561 filtered seawater. A filter was fitted at each outlet of the acrylic tank to collect released
562 planulae. Each larva was washed with FSW transferred from the filter into an Eppendorf tube
563 (FISH) or cryovial (metabarcoding) using a sterile pipette tip and as much water as possible
564 was removed without disturbing the larvae. For FISH and LCM, larvae were fixed for 24 hrs
565 in 4% PFA in FSW, rinsed twice in FSW, and stored in 50% ethanol-PBS at -20°C. For 16S
566 rRNA gene metabarcoding, larvae were snap frozen and stored at -80°C. For FISH, around
567 10 larvae per genotype were fixed. For metabarcoding, larvae were pooled (three per tube),
568 immediately snap-frozen following sampling and kept at -80°C.

569

570 **FISH on slides**

571 Sample processing, sectioning of larvae and decalcification of branches were performed by
572 the Melbourne Histology Platform (University of Melbourne). First, samples were dehydrated
573 in 50% ethanol for 1 hr, 70% ethanol for 1 hr, 90% ethanol for 45 min, 100% ethanol for 1 hr
574 (three times), 50% ethanol and 50% xylene for 45 min, xylene for 45 min (twice), and
575 paraffin for 50 min (three times). Samples were subsequently embedded in solid paraffin, and
576 3-µm sections were cut using a Microm HM 325 rotary microtome and plain microscopic
577 slides (Menzel, Germany). Each slide had four serial sections. FISH was then performed as
578 previously described (*28*), except final probe concentration during hybridization was 5 ng/µL.
579 See Table S9 for probe sequences and formamide concentrations (*16*, *38*, *66*, *67*). Slides were
580 mounted in CitiFluor^{TM} CFM3 mounting medium (proSciTech, Australia) containing 3 µg/uL
581 4',6-diamidino-2-phenylindole dihydrochloride (DAPI; Merck, Germany) to stain nucleic

582 acids, covered with a coverslip and sealed with clear nail polish. Slides were kept at 4°C until
583 observation.

584

585 **Whole-mount FISH**

586 Single polyps were dissected in PBS 1X using Dumont tweezers under a dissecting
587 microscope. Tissue was cleared of autofluorescence by incubating polyps in 50% methanol
588 for 10 min, 75% methanol for 10 min, 90% methanol for 10 min, 100% methanol for 10 min,
589 90% methanol + 0.1% triton for 10 min, 75% methanol + 0.1% triton for 10 min, 50%
590 methanol + 0.1% triton for 10 min, PBS 1X + 0.1% triton for 10 min, and stored at 4°C in
591 PBS 1X.

592 The whole-mount FISH protocol was adapted from previously published protocols (*28*, *68*).
593 Samples were first permeabilized in 70% acetic acid for 2 min and twice rinsed in PBS 1X
594 for 5 min. Samples were then incubated in PBS 1X + 0.1% triton for 10 min, PBS 1X for 5
595 min, pepsin at 0.2 mg/mL in HCl 0.01 M at 37°C for 30 min, PBS 1X for 5 min, HCl 0.2 M
596 for 12 min, and Tris-HCl 20 mM for 10 min. Hybridization was performed in the dark for 3
597 hrs at 46°C in 2 mL of hybridization buffer (0.9 M NaCl, 20 mM Tris-HCl pH 7.2, 0.01%
598 SDS, 5 ng/mL of probe; see Table S9 for formamide concentration). A negative control with
599 no probe was included. Samples were then washed for 15 min at 48°C in washing buffer (20
600 mM Tris-HCl pH 7.2, 0.01% SDS; see Table S9 for NaCl concentration). Samples were then
601 rinsed twice in ice-cold milliQ water and deposited onto 8-well coverslip-bottom slides (ibidi,
602 USA). A drop of milliQ water was deposited onto each polyp to avoid drying. Slides were
603 kept at 4°C until observation.

604

605 **Confocal laser scanning microscopy**

606 Slides were observed on a Nikon A1R confocal laser scanning microscope (Nikon, Japan)
607 with the NIS325 Element software. Virtual band mode was used to acquire variable emission
608 bandwidth to tailor acquisition for specific fluorophores. The fluorophores Cy3 and Atto550
609 were excited using the 561 nm laser line, Atto647 using the 640 nm laser line, DAPI using
610 the 405 nm laser line, and the coral autofluorescence using the 488 nm laser line with a
611 detection range of 570-620 nm for Cy3 and Atto550, 660-710 nm for Atto647, 500-550 nm
612 for coral autofluorescence and 420-480 nm for DAPI. For three-dimensional reconstructions
613 of Z-stacks (Figures 1A and 3B), sections were acquired using Z steps of 3 µM with the 10X
614 objective and 1.1 µM with the 20X objective. Nd2 files were processed using ImageJ. Linear
615 adjustments of brightness and contrast were performed when necessary and applied to the
616 entire image and to each channel independently. Channels were then merged together. Z-
617 stacks were projected in two-dimensional images using the 'Max Intensity' projection type.

618

619 **Transmission electron microscopy**

620 Following decalcification, single polyps were dissected in PBS 1X using Dumont tweezers
621 under a dissecting microscope. Samples were washed in distilled water 3 x 10 min followed
622 by post-fixation with 1% osmium tetroxide in ddH_2O for 1 hr. Samples were then dehydrated
623 in an acetone series of 10, 20, 40, 60, 80 and 100% acetone. Two further exchanges of 100%
624 acetone were undertaken before infiltrating and embedding the samples in Spurr's resin. Thin
625 (100 nm) sections were cut on a Leica UC7 ultramicrotome and collected on copper slot

626 grids. The sections were post-stained with 1% uranyl acetate in water for 10 min and 3% lead
627 citrate for 2 min, and imaged on an FEI Tecnai Spirit transmission electron microscopy
628 equipped with an Eagle CCD camera (Thermofisher). CAMAs were only detected in one
629 polyp of the C2_12 genotype.

630

631 **Laser capture microdissection of CAMAs**

632 Genotypes F1_6 (adults, generations F1 and F2) and O2 (larvae) were chosen to perform
633 laser capture microdissection (LCM) to extract CAMAs. For each group, three replicates
634 containing either one branch or three larvae were processed. For each replicate, 10 slides each
635 containing 8 sections were processed (~80 sections per replicate in total). FISH was
636 performed as described above, using the EUBmix-338-atto550 probe. Slides were not
637 mounted and left without coverslip. To check that FISH had worked and to take photos
638 comparing before and after LCM, one slide per group was covered with a coverslip (without
639 any mounting medium) and sealed with clear nail polish for CLSM imaging as described
640 above. The nail polish was then carefully dissolved by applying small amounts of acetone
641 with a cotton swab and the coverslip was removed. All slides were processed by LCM the
642 next day to avoid any loss of signal. LCM was performed on a PALM Laser Microdissector
643 (Zeiss, Germany), using the PALM Robo 3.2 software, and a 20X air objective. CAMAs
644 were detected using a TRITC filter (excitation 550 ± 12.5 nm; emission 605 ± 35 nm).
645 Individual CAMAs were captured using a UV laser (355 nm) into the caps of 200 µL
646 AdhesiveCap Clear tubes (Zeiss, Germany). These caps are filled with clear adhesive
647 material that ensures sample retention. CAMAs from all slides from one replicate were all
648 collected in one tube. For each replicate, tissue areas without CAMAs were also separately
649 captured as a control. DNA extraction was then conducted within the cap, using the
650 Arcturus® PicoPure® DNA Extraction Kit (Applied Biosystems, USA). 155 µL of
651 Reconstitution Buffer into one vial of Proteinase K to dissolve the enzyme. 10 µL of the
652 Proteinase K solution was then added to each cap and incubated at 65°C for 18 hrs. Tubes
653 were then centrifuged for 5 min at 10,000 × *g* to bring samples down into the tubes and stored
654 at -20°C. Two to three caps containing no sample, but that were open in the LCM facility to
655 capture air contamination, were also included as extraction blanks.

656

657 **DNA extractions of whole larvae**

658 Six replicates each containing 3 larvae were processed. DNA extractions were performed
659 using a salting-out method as previously described (*69*). Three blank DNA extractions were
660 conducted as negative controls. A mock community, ZymoBIOMICS Microbial Community
661 DNA Standard (Zymo Research), was included to check sequencing and processing quality.

662

663 **16S rRNA gene metabarcoding**

664 Hypervariable regions V5-V6 of the 16S rRNA genes were amplified using the primer set
665 784F (5′ GTGACCTATGAACTCAGGAGTCAGGATTAGATACCCTGGTA 3′) and 1061R
666 (5′ CTGAGACTTGCACATCGCAGCCRRCACGAGCTGACGAC 3′). Adapters were
667 attached to the primers and are shown as underlined. Bacterial 16S rRNA genes were PCR-
668 amplified on a SimpliAmp Thermal Cycler (Applied Biosystems, ThermoFisher Scientific).
669 Each reaction contained 1 µL of DNA template, 1.5 µL of forward primer (10 µM stock), 1.5
670 µL of reverse primer (10 µM stock), 7.5 µL of 2x QIAGEN Multiplex PCR Master Mix

671 (Qiagen, Germany) and 3.5 µL of nuclease-free water (Thermofisher), with a total volume of
672 15 µL per reaction. Three triplicate PCRs were conducted for each sample and three no
673 template PCRs were conducted as negative controls. PCR conditions for the 16S rRNA genes
674 were as follows: initial denaturation at 95ºC for 3 min, then 18 cycles of: denaturation at
675 95ºC for 15 s, annealing at 55ºC for 30 s, and extension at 72ºC for 30 s; with a final
676 extension at 72ºC for 7 minutes. Samples were then held at 4ºC. Following PCR, triplicates
677 were pooled, resulting in 45 µL per sample. Metabarcoding library preparation was
678 conducted as previously described (*70*) and sequencing was performed at the Walter and
679 Eliza Hall Institute (WEHI) in Melbourne, Australia on one MiSeq V3 system (Illumina) with
680 2x300bp paired-end reads.
681
682
683 **Bacterial 16S rRNA gene analysis**

684 QIIME2 v 2020.11 (*71*) was used for processing 16S rRNA gene sequences. The plugin
685 demux (*71*) was used to create an interactive plot to visualize the data and assess the quality,
686 for demultiplexing and quality filtering of raw sequences. The plugin cutadapt (*72*) was used
687 to remove the primers and MiSeq adapters. Plugin DADA2 (*73*) was used for denoising and
688 chimera checking, trimming, dereplication, generation of a feature table, joining of paired-
689 end reads, and correcting sequencing errors and removing low quality reads (Q-score < 30).
690 Summary statistics were obtained using the feature-table to ensure processing was successful.
691 Taxonomy was assigned by training a naive Bayes classifier with the feature-classifier plugin
692 (*71*), based on a 99% similarity to the V5-V6 region of the 16S rRNA gene in the SILVA 138
693 database to match the 784F/1061R primer pair used (*74*). Mitochondria and chloroplast reads
694 were filtered out. Analyses and graphs were performed using Rstudio version 2022.02.2 and
695 the phyloseq package (*75*). Metadata file, taxonomy table, phylogenetic tree and ASV table
696 were imported into R to create a phyloseq object. Contaminant ASVs, arising from kit
697 reagents and sample manipulation, were identified using the package decontam (*76*). The
698 function 'isNotContaminant' was used as it is more stringent and more adequate for low-
699 biomass samples. One replicate for each the Adult F1 condition, larvae OI2 and larvae OI3
700 were removed because of heavy contamination. It is worth noting that bacteria belonging to
701 the *Brachybacterium* genus accounted for more than 90% of the contamination (Table S2),
702 which were contaminants of the Qiagen PCR kit.

703

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

705 To reach sufficient quantities for shotgun sequencing, DNA samples obtained through LCM
706 needed to be amplified. Whole-genome amplification was carried out using a SeqPlex DNA
707 Amplification Kit (Sigma-Aldrich, USA), which is specifically designed for low-quantity,
708 fragmented samples. The same DNA samples that were used for 16S rRNA gene
709 metabarcoding were used and, for each sample, 1.5 µL of each of the three replicates were
710 pooled before amplification. This was only carried out using the F1_6 genotype samples
711 (both generations F1 and F2), as there was not enough material left from the O2 and O3
712 genotype samples. Amplification was performed according to the manufacturer's
713 instructions, using 29 cycles in an end-point PCR. Intermediate DNA concentrations were
714 measured using a Qubit Fluorometer and a Qubit high-sensitivity dsDNA quantification
715 assay. Primer removal was also carried out using the SeqPlex DNA Amplification kit. Final
716 DNA quality and quantities were checked on a TapeStation (Agilent, USA). 16S rRNA gene
717 metabarcoding of the amplified samples showed similar results to the non-amplified samples,
718 confirming that amplification bias was minimal (Table S3A).

719 Samples were then sent to the Ramaciotti Centre for Genomics (Sydney, Australia) for
720 sequencing. Library preparation was performed with a NextFLEX Rapid DNA Seq 2.0 Prep
721 kit and sequenced on a NovaSeq 6000 SP 2x100 bp Flowcell Illumina platform.

722

723 **Data analysis of metagenomics**

724 Each sample was run on two different lanes, so reads from each lane were first merged,
725 according to their direction (i.e., Forward reads from both lanes were merged, and Reverse
726 reads from both lanes were merged). Adapters were removed from the raw data using
727 Cutadapt v4.0 (*72*) with default settings. Quality control was performed using FastQC
728 v0.11.9 (*77*) and low-quality sequences (phred score < 30) were trimmed using Trimmomatic
729 v0.36 (*78*) (LEADING:30 HEADCROP:10 MINLEN:70). High-quality reads were mapped
730 against a *Pocillopora acuta* draft genome (*79*) using Bowtie2 v2.4.5 with default parameters
731 (*80*) to remove host related reads from the metagenome. Mapped reads were then removed
732 using Samtools v1.11 (*81*). Only paired-end host-removed reads were used for metagenome
733 assembly. Metagenome assembly was carried out using MEGAHIT v1.2.9 (*82*) with a
734 minimum contig length of 1000 bp and the following k-mers: 21, 33, 55, 77. Trimmed reads
735 were then mapped back to the assembly using Samtools v1.11 (*81*). Assembled contigs were
736 binned using the binning (with metabat2, concoct and maxbin2 tools) and bin_refinement
737 (with >70% completeness and <10% contamination as cut-off parameters) modules of
738 MetaWRAP v1.3.2 (*85*). Bin quality and taxonomy were assessed using CheckM2 v0.1.3
739 (*86*) and GTDB-Tk v2.1.0 (*87*), respectively. These bins were reassembled to further improve
740 the contiguity and bin completeness and contamination stats using the reassemble_bins
741 module SPAdes implemented in MetaWRAP v1.3.2. The final taxonomy of individual
742 contigs on a per bin level was assessed using CAT/BAT v5.2.3 (*84*), and any contig
743 belonging to a different phylum than the taxonomy assigned by GTDB-Tk was manually
744 removed. Bin coverage was obtained using CoverM v0.6.1
745 (https://github.com/wwood/CoverM) using the "genome" option.

746

747 **Genome annotation**

748 Gene prediction and annotation of bins was performed using Prokka v1.14.6 with default
749 settings (*88*), RAST v2.0 (*89*), KEGG-mapper Reconstruct (*90*), eggNOG-mapper v2.1.6
750 with a cut-off e-value of 1.10^{-5} (*91*) with the eggNOG database v. 5.0 (*92*), and InterProScan
751 v5.55 with Pfam domain annotations (*93*). For Pac_F2b (*Simkania* MAG), Prokka annotation
752 was performed using the "--protein" option with a well-annotated set of chlamydial genes as
753 initial reference. Secondary metabolites were predicted using antiSMASH v6.1.1 (*94*). For
754 the annotation of DMSP demethylases (DmdABC), conserved domains were double-checked
755 with web-based CD-search using e-value cut-off 1e-5 (*95*). Average Nucleotide Identities
756 (ANI) and Average Amino acid Identities (AAI) of all three MAGs were calculated using a
757 genome-based matrix calculator (*96*).

758

759 **Phylogenetic analyses**

760 16S rRNA gene sequences were recovered from MAGs using barrnap as implemented in
761 Prokka. The 16S rRNA gene phylogenetic tree of the Endozoicomonadaceae family was
762 constructed using full-length 16S rRNA gene sequences from the SILVA v132 database. For
763 Pac_F2a, a full-length 16S rRNA gene (1536 bp) was recovered and used for the

764 phylogenetic tree construction. For Pac_F1, two fragments were recovered, of 1129 bp and
765 401 bp respectively and only the 1129-bp fragment was used. A MAFFT alignment was
766 created using Geneious Prime v2019.1.3. This alignment was used to generate a maximum
767 likelihood phylogenetic tree with 1000 ultrafast bootstraps using IQ-Tree v2.2.0.3 (*97*) with
768 the best model TIM3e+R6, selected by ModelFinder wrapped in IQ-tree (*98*). Sequences
769 belonging to the *Zookishella* genus were chosen as an outgroup. For the *Endozoicomonas*
770 whole-genome phylogenetic tree, GTDB-Tk v2.1.0 (*87*) was used to create an alignment
771 based on 120 gene markers. Nucleotide sequences were chosen because all the genomes
772 belong to the same genus. The genome of *P. haliclonae* was selected as an outgroup. The
773 reference genomes are listed in Table S5. A maximum likelihood phylogenetic tree was
774 constructed based on the alignment with IQ-Tree v2.1.2 (*97*) using the best model LG+F+R3,
775 selected by ModelFinder wrapped in IQ-tree (*98*), and 1000 ultrafast bootstrap replicates
776 (*99*).

777 For comparative genome analysis of the Pac_F2b MAG, a dataset of high quality chlamydial
778 genomes based on Dharamshi et al. (*55*) was used. This dataset was complemented by
779 additional chlamydial genomes available on NCBI (last accessed on 01/09/2022). Only
780 genomes with a completeness >70% (determined with CheckM2 v0.1.3 (*86*)) and average
781 nucleotide identities <95% (determined with FastANI v1.33 (*100*)) were considered for the
782 final dataset which included 139 chlamydial genomes (Table S5B). For further comparative
783 analysis of the chlamydial genomes, all encoded protein sequences were clustered into
784 orthologous groups (OGs) with OrthoFinder v2.5.2 with default parameters (*101*). In order to
785 obtain the phylogenetic affiliation of the Pac_F2b MAG, a set of concatenated protein
786 sequences from 15 non-supervised orthologous groups (NOGs) was used (*55*) (Table S10). It
787 has been shown previously that the usage of these 15 NOGs retrieves the same topology for
788 chlamydial phylogeny as the application of larger protein sets (*55*). Thus, the dataset from
789 Dharamshi et al. (*55*) including a large outgroup of other members of the Planctomycetes,
790 Verrucomicobia, Chlamydiae (PVC) superphylum was expanded by the Pac_F2b MAG and
791 the additional genomes from NCBI. Proteins of these genomes belonging to the 15 NOGs
792 were aligned to the existing alignments from Dharamshi et al. (*55*) with MAFFT v7.490 "--
793 add" (*102*). The resulting single protein alignments were subsequently trimmed with BMGE
794 v1.12 (*103*) (entropy 0.6) and concatenated. Maximum likelihood phylogeny was inferred
795 with IQ-TREE v2.1.2 (*97*) with 1000 ultrafast bootstrap replicates (*99*) and 1000 replicates of
796 the SH-like approximate likelihood ratio test (*104*) under the LG+C60+F+R4 model.

797

798 **References**

799 1. R. Fisher, R. A. O'Leary, S. Low-Choy, K. Mengersen, N. Knowlton, R. E. Brainard, M. J. Caley,
800 Species richness on coral reefs and the pursuit of convergent global estimates. *Current Biology*.
801 **25**, 500–505 (2015).

802 2. D. L. Santavy, C. L. Horstmann, L. M. Sharpe, S. H. Yee, P. Ringold, What is it about coral reefs?
803 Translation of ecosystem goods and services relevant to people and their well-being.
804 *Ecosphere*. **12**, e03639 (2021).

805 3. D. G. Bourne, K. M. Morrow, N. S. Webster, Insights into the coral microbiome: Underpinning
806 the health and resilience of reef ecosystems. *Annual Review of Microbiology*. **70**, 317–340
807 (2016).

4. M. J. H. van Oppen, L. L. Blackall, Coral microbiome dynamics, functions and design in a changing world. *Nature Reviews Microbiology*. **17**, 557 (2019).

5. J. Maire, P. Buerger, W. Y. Chan, P. Deore, A. M. Dungan, M. R. Nitschke, M. J. H. van Oppen, Effects of Ocean Warming on the Underexplored Members of the Coral Microbiome. *Integrative and Comparative Biology* (2022), doi:10.1093/ICB/ICAC005.

6. M. R. Nitschke, S. L. Rosset, C. A. Oakley, S. G. Gardner, E. F. Camp, D. J. Suggett, S. K. Davy, "The diversity and ecology of Symbiodiniaceae: A traits-based review" in *Advances in Marine Biology* (Academic Press, 2022; https://www.sciencedirect.com/science/article/pii/S0065288122000141).

7. K. B. Ritchie, Regulation of microbial populations by coral surface mucus and mucus-associated bacteria. *Marine Ecology Progress Series*. **322**, 1–14 (2006).

8. J. Ceh, M. R. Kilburn, J. B. Cliff, J.-B. Raina, M. van Keulen, D. G. Bourne, Nutrient cycling in early coral life stages: Pocillopora damicornis larvae provide their algal symbiont (Symbiodinium) with nitrogen acquired from bacterial associates. *Ecology and Evolution*. **3**, 2393–2400 (2013).

9. M. P. Lesser, C. H. Mazel, M. Y. Gorbunov, P. G. Falkowski, Discovery of symbiotic nitrogen-fixing cyanobacteria in corals. *Science (New York, N.Y.)*. **305**, 997–1000 (2004).

10. K. Tandon, C.-Y. Lu, P.-W. Chiang, N. Wada, S.-H. Yang, Y.-F. Chan, P.-Y. Chen, H.-Y. Chang, Y.-J. Chiou, M.-S. Chou, W.-M. Chen, S.-L. Tang, Comparative genomics: Dominant coral-bacterium Endozoicomonas acroporae metabolizes dimethylsulfoniopropionate (DMSP). *The ISME Journal*. **14**, 1290–1303 (2020).

11. J. Maire, L. L. Blackall, M. J. H. van Oppen, Intracellular Bacterial Symbionts in Corals: Challenges and Future Directions. *Microorganisms*. **9**, 2209 (2021).

12. E. C. Peters, A survey of cellular reactions to environmental stress and disease in Caribbean scleractinian corals. *Helgoländer Meeresuntersuchungen*. **37**, 113–137 (1984).

13. E. C. Peters, J. J. Oprandy, P. P. Yevich, Possible causal agent of “white band disease” in caribbean acroporid corals. *Journal of Invertebrate Pathology*. **41**, 394–396 (1983).

14. T. D. Ainsworth, M. Fine, L. L. Blackall, O. Hoegh-Guldberg, Fluorescence in situ hybridization and spectral imaging of coral-associated bacterial communities. *Applied and Environmental Microbiology*. **72**, 3016–3020 (2006).

15. T. D. Ainsworth, O. Hoegh-Guldberg, Bacterial communities closely associated with coral tissues vary under experimental and natural reef conditions and thermal stress. *Aquatic Biology*. **4**, 289–296 (2009).

16. T. Bayer, M. J. Neave, A. Alsheikh-Hussain, M. Aranda, L. K. Yum, T. Mincer, K. Hughen, A. Apprill, C. R. Voolstra, The microbiome of the red sea coral stylophora pistillata is dominated by tissue-associated endozoicomonas bacteria. *Applied and Environmental Microbiology*. **79**, 4759–4762 (2013).

17. M. La Rivière, M. Garel, M. Bally, Localization of endobacteria in the gastrodermis of a Mediterranean gorgonian coral, Paramuricea clavata, using fluorescence in situ hybridization. *Mar Biol*. **163**, 206 (2016).

18. M. J. McKinstry, G. B. Chapman, D. M. Spoon, E. C. Peters, The Occurrence of Bacterial Colonies in the Epidermis of the Tentacles of the Sea Anemone Aiptasia pallida (Anthozoa: Actinaria). *Transactions of the American Microscopical Society*. **108**, 239 (1989).

19. M. J. Neave, R. Rachmawati, L. Xun, C. T. Michell, D. G. Bourne, A. Apprill, C. R. Voolstra, Differential specificity between closely related corals and abundant Endozoicomonas endosymbionts across global scales. *ISME Journal*. **11**, 186–200 (2017).

20. E. E. Palincsar, W. R. Jones, J. S. Palincsar, M. A. Glogowski, J. L. Mastro, Bacterial aggregates within the epidermis of the sea anemone Aiptasia pallida. *The Biological Bulletin*. **177**, 130–140 (1989).

21. N. Wada, M.-T. Hsu, K. Tandon, S. S.-Y. Hsiao, H.-J. Chen, Y.-H. Chen, P.-W. Chiang, S.-P. Yu, C.-Y. Lu, Y.-J. Chiou, Y.-C. Tu, X. Tian, B.-C. Chen, D.-C. Lee, H. Yamashiro, D. G. Bourne, S.-L. Tang, High-resolution spatial and genomic characterization of coral-associated microbial aggregates in the coral Stylophora pistillata. *Science Advances*. **8**, eabo2431 (2022).

22. N. Wada, M. Ishimochi, T. Matsui, F. J. Pollock, S.-L. Tang, T. D. Ainsworth, B. L. Willis, N. Mano, D. G. Bourne, Characterization of coral-associated microbial aggregates (CAMAs) within tissues of the coral Acropora hyacinthus. *Scientific reports*. **9**, 14662 (2019).

23. T. Work, G. Aeby, Microbial aggregates within tissues infect a diversity of corals throughout the Indo-Pacific. *Marine Ecology Progress Series*. **500**, 1–9 (2014).

24. M. J. Neave, A. Apprill, C. Ferrier-Pagès, C. R. Voolstra, Diversity and function of prevalent symbiotic marine bacteria in the genus Endozoicomonas. *Applied Microbiology and Biotechnology*. **100**, 8315–8324 (2016).

25. C. Pogoreutz, C. A. Oakley, N. Rädecker, A. Cárdenas, G. Perna, N. Xiang, L. Peng, S. K. Davy, D. K. Ngugi, C. R. Voolstra, Coral holobiont cues prime Endozoicomonas for a symbiotic lifestyle. *The ISME Journal 2022*, 1–13 (2022).

26. M. J. Neave, C. T. Michell, A. Apprill, C. R. Voolstra, Endozoicomonas genomes reveal functional adaptation and plasticity in bacterial strains symbiotically associated with diverse marine hosts. *Scientific Reports*. **7**, 1–12 (2017).

27. H. A. Smith, A. Moya, N. E. Cantin, M. J. H. van Oppen, G. Torda, Observations of Simultaneous Sperm Release and Larval Planulation Suggest Reproductive Assurance in the Coral Pocillopora acuta. *Frontiers in Marine Science*. **6**, 362 (2019).

28. K. Damjanovic, P. Menéndez, L. L. Blackall, M. J. H. van Oppen, Mixed-mode bacterial transmission in the common brooding coral Pocillopora acuta. *Environmental Microbiology*. **22**, 397–412 (2020).

29. T. D. Haydon, J. R. Seymour, J. B. Raina, J. Edmondson, N. Siboni, J. L. Matthews, E. F. Camp, D. J. Suggett, Rapid Shifts in Bacterial Communities and Homogeneity of Symbiodiniaceae in Colonies of Pocillopora acuta Transplanted Between Reef and Mangrove Environments. *Frontiers in Microbiology*. **12** (2021), doi:10.3389/FMICB.2021.756091/FULL.

30. K. Damjanovic, L. L. Blackall, L. M. Peplow, M. J. H. van Oppen, Assessment of bacterial community composition within and among Acropora loripes colonies in the wild and in captivity. *Coral Reefs*. **39**, 1245 (2020).

31. L. Schreiber, K. U. Kjeldsen, P. Funch, J. Jensen, M. Obst, S. López-Legentil, A. Schramm, Endozoicomonas Are Specific, Facultative Symbionts of Sea Squirts. *Frontiers in Microbiology*. **7** (2016) (available at https://www.frontiersin.org/articles/10.3389/fmicb.2016.01042).

32. P. Katharios, H. M. B. Seth-Smith, A. Fehr, J. M. Mateos, W. Qi, D. Richter, L. Nufer, M. Ruetten, M. Guevara Soto, U. Ziegler, N. R. Thomson, R. Schlapbach, L. Vaughan, Environmental marine pathogen isolation using mesocosm culture of sharpsnout seabream: striking genomic and morphological features of novel Endozoicomonas sp. *Sci Rep*. **5**, 17609 (2015).

33. F. U. Zielinski, A. Pernthaler, S. Duperron, L. Raggi, O. Giere, C. Borowski, N. Dubilier, Widespread occurrence of an intranuclear bacterial parasite in vent and seep bathymodiolin mussels. *Environmental Microbiology*. **11**, 1150–1167 (2009).

34. A. Collingro, S. Köstlbacher, M. Horn, Chlamydiae in the environment. *Trends in Microbiology*. **28**, 877 (2020).

35. W. Y. Chan, L. M. Peplow, P. Menéndez, A. A. Hoffmann, M. J. H. van Oppen, The roles of age, parentage and environment on bacterial and algal endosymbiont communities in Acropora corals. *Molecular Ecology*. **28**, 3830 (2019).

36. P. Buerger, R. T. Vanstone, J. Maire, M. J. H. van Oppen, Long-Term Heat Selection of the Coral Endosymbiont Cladocopium C1acro (Symbiodiniaceae) Stabilizes Associated Bacterial Communities. *International Journal of Molecular Sciences*. **23**, 4913 (2022).

37. J. Maire, S. K. Girvan, S. E. Barkla, A. Perez-Gonzalez, D. J. Suggett, L. L. Blackall, M. J. H. Van Oppen, Intracellular bacteria are common and taxonomically diverse in cultured and in hospite algal endosymbionts of coral reefs. *The ISME Journal*. **15**, 2028–2042 (2021).

38. S. Poppert, A. Essig, R. Marre, M. Wagner, M. Horn, Detection and differentiation of chlamydiae by fluorescence in situ hybridization. *Applied and Environmental Microbiology*. **68**, 4081–4089 (2002).

39. I. Lagkouvardos, J. Shen, M. Horn, Improved axenization method reveals complexity of symbiotic associations between bacteria and acanthamoebae. *Environmental Microbiology Reports*. **6**, 383–388 (2014).

40. M. I. Blandford, A. Taylor-Brown, T. A. Schlacher, B. Nowak, A. Polkinghorne, Epitheliocystis in fish: An emerging aquaculture disease with a global impact. *Transboundary and Emerging Diseases*. **65**, 1436–1446 (2018).

41. S. P. Glaeser, P. Kämpfer, "Kistimonas" in *Bergey's Manual of Systematics of Archaea and Bacteria*, A. Oren, Ed. (Wiley, 2020; https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118960608.gbm01541), pp. 1–7.

42. S. Köstlbacher, A. Collingro, T. Halter, F. Schulz, S. P. Jungbluth, M. Horn, Pangenomics reveals alternative environmental lifestyles among chlamydiae. *Nature Communications*. **12**, 4021 (2021).

43. A. Collingro, P. Tischler, T. Weinmaier, T. Penz, E. Heinz, R. C. Brunham, T. D. Read, P. M. Bavoil, K. Sachse, S. Kahane, M. G. Friedman, T. Rattei, G. S. A. Myers, M. Horn, Unity in Variety—The Pan-Genome of the Chlamydiae. *Mol Biol Evol*. **28**, 3253–3270 (2011).

926 44. J. Dharamasi, S. Köstlbacher, M. E. Schön, A. Collingro, T. J. G. Ettema, M. Horn, Gene gain
927 facilitated endosymbiotic evolution of Chlamydiae. *Nature Microbiology* (In press).

928 45. D. G. Bourne, C. B. Munn, Diversity of bacteria associated with the coral Pocillopora damicornis
929 from the Great Barrier Reef. *Environ Microbiol*. **7**, 1162–1174 (2005).

930 46. X. Wang, H. Zhou, H. Chen, X. Jing, W. Zheng, R. Li, T. Sun, J. Liu, J. Fu, L. Huo, Y. Li, Y. Shen, X.
931 Ding, R. Müller, X. Bian, Y. Zhang, Discovery of recombinases enables genome mining of cryptic
932 biosynthetic gene clusters in Burkholderiales species. *Proc Natl Acad Sci U S A*. **115**, E4255–
933 E4263 (2018).

934 47. J. E. Martyn, L. Gomez-Valero, C. Buchrieser, The evolution and role of eukaryotic-like domains
935 in environmental intracellular bacteria: the battle with a eukaryotic cell. *FEMS Microbiology*
936 *Reviews*. **46**, fuac012 (2022).

937 48. L. Craig, K. T. Forest, B. Maier, Type IV pili: dynamics, biophysics and functional consequences.
938 *Nat Rev Microbiol*. **17**, 429–440 (2019).

939 49. V. Michalkova, J. B. Benoit, B. L. Weiss, G. M. Attardo, S. Aksoy, Vitamin B6 generated by
940 obligate symbionts is critical for maintaining proline homeostasis and fecundity in tsetse flies.
941 *Applied and environmental microbiology*. **80**, 5844–53 (2014).

942 50. M. T. Croft, A. D. Lawrence, E. Raux-Deery, M. J. Warren, A. G. Smith, Algae acquire vitamin
943 B12 through a symbiotic relationship with bacteria. *Nature*. **438**, 90–93 (2005).

944 51. J. Maire, M. J. H. van Oppen, A role for bacterial experimental evolution in coral bleaching
945 mitigation? *Trends in Microbiology*. **30**, 217–228 (2022).

946 52. R. S. Peixoto, P. M. Rosado, D. C. de A. Leite, A. S. Rosado, D. G. Bourne, Beneficial
947 Microorganisms for Corals (BMC): Proposed mechanisms for coral health and resilience.
948 *Frontiers in Microbiology*. **8**, 341 (2017).

949 53. M. Szabó, A. W. D. Larkum, I. Vass, "A Review: The role of reactive oxygen species in mass coral
950 bleaching" in *Advances in Photosynthesis and Respiration* (Springer, 2020;
951 https://doi.org/10.1007/978-3-030-33397-3_17), vol. 45, pp. 459–488.

952 54. J. M. McDevitt-Irwin, J. K. Baum, M. Garren, R. L. Vega Thurber, Responses of coral-associated
953 bacterial communities to local and global stressors. *Frontiers in Marine Science*. **4**, 262 (2017).

954 55. J. E. Dharamshi, N. Gaarslev, K. Steffen, T. Martin, D. Sipkema, T. J. G. Ettema, Genomic
955 diversity and biosynthetic capabilities of sponge-associated chlamydiae. *ISME J*, 1–16 (2022).

956 56. J. Yao, V. J. Dodson, M. W. Frank, C. O. Rock, Chlamydia trachomatis Scavenges Host Fatty
957 Acids for Phospholipid Synthesis via an Acyl-Acyl Carrier Protein Synthetase. *J Biol Chem*. **290**,
958 22163–22173 (2015).

959 57. C. J. Rosario, B. R. Hanson, M. Tan, The Transcriptional Repressor EUO Regulates Both Subsets
960 of Chlamydia Late Genes. *Mol Microbiol*. **94**, 888–897 (2014).

961 58. B. M. Dudiak, T. M. Nguyen, D. Needham, T. C. Outlaw, D. G. McCafferty, Inhibition of the
962 futalosine pathway for menaquinone biosynthesis suppresses Chlamydia trachomatis infection.
963 *FEBS Letters*. **595**, 2995–3005 (2021).

964 59. M. Almirón, M. Martínez, N. Sanjuan, R. A. Ugalde, Ferrochelatase is present in Brucella
965 abortus and is critical for its intracellular survival and virulence. *Infect Immun*. **69**, 6225–6230
966 (2001).

967 60. N. D. Hammer, M. L. Reniere, J. E. Cassat, Y. Zhang, A. O. Hirsch, M. Indriati Hood, E. P. Skaar,
968 Two heme-dependent terminal oxidases power Staphylococcus aureus organ-specific
969 colonization of the vertebrate host. *mBio*. **4**, e00241-13 (2013).

970 61. C. von Eiff, C. Heilmann, R. A. Proctor, C. Woltz, G. Peters, F. Götz, A site-directed
971 Staphylococcus aureus hemB mutant is a small-colony variant which persists intracellularly. *J*
972 *Bacteriol*. **179**, 4706–4712 (1997).

973 62. P. Engström, B. D. Nguyen, J. Normark, I. Nilsson, R. J. Bastidas, Å. Gylfe, M. Elofsson, K. A.
974 Fields, R. H. Valdivia, H. Wolf-Watz, S. Bergström, Mutations in hemG Mediate Resistance to
975 Salicylidene Acylhydrazides, Demonstrating a Novel Link between Protoporphyrinogen Oxidase
976 (HemG) and Chlamydia trachomatis Infectivity. *J Bacteriol*. **195**, 4221–4230 (2013).

977 63. M. J. Gosalbes, A. Lamelas, A. Moya, A. Latorre, The Striking Case of Tryptophan Provision in
978 the Cedar Aphid Cinara cedri. *J Bacteriol*. **190**, 6026–6029 (2008).

979 64. J. P. McCutcheon, C. D. von Dohlen, An Interdependent Metabolic Patchwork in the Nested
980 Symbiosis of Mealybugs. *Current Biology*. **21**, 1366–1372 (2011).

981 65. M. Ziegler, C. G. B. Grupstra, M. M. Barreto, M. Eaton, J. BaOmar, K. Zubier, A. Al-Sofyani, A. J.
982 Turki, R. Ormond, C. R. Voolstra, Coral bacterial community structure responds to
983 environmental change in a host-specific manner. *Nature Communications*. **10**, 3092 (2019).

984 66. H. Daims, A. Brühl, R. Amann, K. H. Schleifer, M. Wagner, The domain-specific probe EUB338 is
985 insufficient for the detection of all bacteria: Development and evaluation of a more
986 comprehensive probe set. *Systematic and Applied Microbiology*. **22**, 434–444 (1999).

987 67. G. Wallner, R. Amann, W. Beisker, Optimizing fluorescent in situ hybridization with rRNA-
988 targeted oligonucleotide probes for flow cytometric identification of microorganisms.
989 *Cytometry*. **14**, 136–143 (1993).

990 68. J. Maire, N. Parisot, M. Galvao Ferrarini, A. Vallier, B. Gillet, S. Hughes, S. Balmand, C. Vincent-
991 Monégat, A. Zaidman-Rémy, A. Heddi, Spatial and morphological reorganization of
992 endosymbiosis during metamorphosis accommodates adult metabolic requirements in a
993 weevil. *Proceedings of the National Academy of Sciences*. **117**, 19347–19358 (2020).

994 69. L. M. Hartman, M. J. H. van Oppen, L. L. Blackall, Microbiota characterization of Exaiptasia
995 diaphana from the Great Barrier Reef. *Animal Microbiome*. **2**, 10 (2020).

996 70. J. Maire, L. L. Blackall, M. J. H. van Oppen, Microbiome characterization of defensive tissues in
997 the model anemone Exaiptasia diaphana. *BMC Microbiology*. **21**, 152 (2021).

998 71. E. Bolyen, J. R. Rideout, M. R. Dillon, N. A. Bokulich, C. C. Abnet, G. A. Al-Ghalith, H. Alexander,
999 E. J. Alm, M. Arumugam, F. Asnicar, Y. Bai, J. E. Bisanz, K. Bittinger, A. Brejnrod, C. J. Brislawn, C.
1000 T. Brown, B. J. Callahan, A. M. Caraballo-Rodríguez, J. Chase, E. K. Cope, R. Da Silva, C. Diener,
1001 P. C. Dorrestein, G. M. Douglas, D. M. Durall, C. Duvallet, C. F. Edwardson, M. Ernst, M. Estaki,
1002 J. Fouquier, J. M. Gauglitz, S. M. Gibbons, D. L. Gibson, A. Gonzalez, K. Gorlick, J. Guo, B.
1003 Hillmann, S. Holmes, H. Holste, C. Huttenhower, G. A. Huttley, S. Janssen, A. K. Jarmusch, L.
1004 Jiang, B. D. Kaehler, K. Bin Kang, C. R. Keefe, P. Keim, S. T. Kelley, D. Knights, I. Koester, T.

1005 Kosciolek, J. Kreps, M. G. I. Langille, J. Lee, R. Ley, Y. X. Liu, E. Loftfield, C. Lozupone, M. Maher,
1006 C. Marotz, B. D. Martin, D. McDonald, L. J. McIver, A. V. Melnik, J. L. Metcalf, S. C. Morgan, J. T.
1007 Morton, A. T. Naimey, J. A. Navas-Molina, L. F. Nothias, S. B. Orchanian, T. Pearson, S. L.
1008 Peoples, D. Petras, M. L. Preuss, E. Pruesse, L. B. Rasmussen, A. Rivers, M. S. Robeson, P.
1009 Rosenthal, N. Segata, M. Shaffer, A. Shiffer, R. Sinha, S. J. Song, J. R. Spear, A. D. Swafford, L. R.
1010 Thompson, P. J. Torres, P. Trinh, A. Tripathi, P. J. Turnbaugh, S. Ul-Hasan, J. J. J. van der Hooft,
1011 F. Vargas, Y. Vázquez-Baeza, E. Vogtmann, M. von Hippel, W. Walters, Y. Wan, M. Wang, J.
1012 Warren, K. C. Weber, C. H. D. Williamson, A. D. Willis, Z. Z. Xu, J. R. Zaneveld, Y. Zhang, Q. Zhu,
1013 R. Knight, J. G. Caporaso, Reproducible, interactive, scalable and extensible microbiome data
1014 science using QIIME 2. *Nature Biotechnology*. **37**, 852–857 (2019).

1015 72. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads.
1016 *EMBnet.journal*. **17**, 10 (2011).

1017 73. B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, S. P. Holmes, DADA2:
1018 High-resolution sample inference from Illumina amplicon data. *Nature Methods*. **13**, 581–583
1019 (2016).

1020 74. N. A. Bokulich, B. D. Kaehler, J. R. Rideout, M. Dillon, E. Bolyen, R. Knight, G. A. Huttley, J.
1021 Gregory Caporaso, Optimizing taxonomic classification of marker-gene amplicon sequences
1022 with QIIME 2's q2-feature-classifier plugin. *Microbiome*. **6**, 90 (2018).

1023 75. P. J. McMurdie, S. Holmes, phyloseq: An R package for reproducible interactive analysis and
1024 graphics of microbiome census data. *PLoS ONE*. **8**, e61217 (2013).

1025 76. N. M. Davis, Di. M. Proctor, S. P. Holmes, D. A. Relman, B. J. Callahan, Simple statistical
1026 identification and removal of contaminant sequences in marker-gene and metagenomics data.
1027 *Microbiome*. **6**, 226 (2018).

1028 77. S. Andrews, FASTQC. A quality control tool for high throughput sequence data | BibSonomy,
1029 (available at https://www.bibsonomy.org/bibtex/f230a919c34360709aa298734d63dca3).

1030 78. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data.
1031 *Bioinformatics (Oxford, England)*. **30**, 2114–2120 (2014).

1032 79. J. Vidal-Dupiol, C. Chaparro, M. Pratlong, P. Pontarotti, C. Grunau, G. Mitta, Sequencing, de
1033 novo assembly and annotation of the genome of the scleractinian coral, Pocillopora acuta.
1034 *bioRxiv*, 698688 (2020).

1035 80. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature methods*. **9**,
1036 357–9 (2012).

1037 81. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin,
1038 1000 Genome Project Data Processing Subgroup, The Sequence Alignment/Map format and
1039 SAMtools. *Bioinformatics*. **25**, 2078–2079 (2009).

1040 82. D. Li, C. M. Liu, R. Luo, K. Sadakane, T. W. Lam, MEGAHIT: an ultra-fast single-node solution for
1041 large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*
1042 *(Oxford, England)*. **31**, 1674–1676 (2015).

1043 83. A. Gurevich, V. Saveliev, N. Vyahhi, G. Tesler, QUAST: Quality assessment tool for genome
1044 assemblies. *Bioinformatics*. **29**, 1072–1075 (2013).

1045 84. F. A. B. Von Meijenfeldt, K. Arkhipova, D. D. Cambuy, F. H. Coutinho, B. E. Dutilh, Robust
1046 taxonomic classification of uncharted microbial sequences and bins with CAT and BAT. *Genome*
1047 *Biology*. **20**, 1–14 (2019).

1048 85. G. V. Uritskiy, J. Diruggiero, J. Taylor, MetaWRAP - A flexible pipeline for genome-resolved
1049 metagenomic data analysis. *Microbiome*. **6**, 158 (2018).

1050 86. A. Chklovski, D. H. Parks, B. J. Woodcroft, G. W. Tyson, CheckM2: a rapid, scalable and accurate
1051 tool for assessing microbial genome quality using machine learning (2022), p.
1052 2022.07.11.499243, , doi:10.1101/2022.07.11.499243.

1053 87. P. A. Chaumeil, A. J. Mussig, P. Hugenholtz, D. H. Parks, GTDB-Tk: a toolkit to classify genomes
1054 with the Genome Taxonomy Database. *Bioinformatics*. **36**, 1925–1927 (2020).

1055 88. T. Seemann, Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. **30**, 2068–2069
1056 (2014).

1057 89. R. K. Aziz, D. Bartels, A. A. Best, M. DeJongh, T. Disz, R. A. Edwards, K. Formsma, S. Gerdes, E.
1058 M. Glass, M. Kubal, F. Meyer, G. J. Olsen, R. Olson, A. L. Osterman, R. A. Overbeek, L. K. McNeil,
1059 D. Paarmann, T. Paczian, B. Parrello, G. D. Pusch, C. Reich, R. Stevens, O. Vassieva, V. Vonstein,
1060 A. Wilke, O. Zagnitko, The RAST Server: rapid annotations using subsystems technology. *BMC*
1061 *Genomics*. **9**, 75 (2008).

1062 90. M. Kanehisa, Y. Sato, KEGG Mapper for inferring cellular functions from protein sequences.
1063 *Protein Sci*. **29**, 28–35 (2020).

1064 91. C. P. Cantalapiedra, A. Hernández-Plaza, I. Letunic, P. Bork, J. Huerta-Cepas, eggNOG-mapper
1065 v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic
1066 Scale. *Mol Biol Evol*. **38**, 5825–5829 (2021).

1067 92. J. Huerta-Cepas, D. Szklarczyk, D. Heller, A. Hernández-Plaza, S. K. Forslund, H. Cook, D. R.
1068 Mende, I. Letunic, T. Rattei, L. J. Jensen, C. von Mering, P. Bork, eggNOG 5.0: a hierarchical,
1069 functionally and phylogenetically annotated orthology resource based on 5090 organisms and
1070 2502 viruses. *Nucleic Acids Research*. **47**, D309–D314 (2019).

1071 93. P. Jones, D. Binns, H. Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A.
1072 Mitchell, G. Nuka, S. Pesseat, A. F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S. Y. Yong, R.
1073 Lopez, S. Hunter, InterProScan 5: genome-scale protein function classification. *Bioinformatics*.
1074 **30**, 1236–1240 (2014).

1075 94. K. Blin, S. Shaw, A. M. Kloosterman, Z. Charlop-Powers, G. P. van Wezel, M. H. Medema, T.
1076 Weber, antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids*
1077 *Research*. **49**, W29–W35 (2021).

1078 95. A. Marchler-Bauer, S. H. Bryant, CD-Search: protein domain annotations on the fly. *Nucleic*
1079 *Acids Research*. **32**, W327–W331 (2004).

1080 96. L. M. Rodriguez-R, K. T. Konstantinidis, The enveomics collection: a toolbox for specialized
1081 analyses of microbial genomes and metagenomes (2016),
1082 doi:10.7287/PEERJ.PREPRINTS.1900V1.

1083 97. B. Q. Minh, H. A. Schmidt, O. Chernomor, D. Schrempf, M. D. Woodhams, A. von Haeseler, R.
1084 Lanfear, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the
1085 Genomic Era. *Molecular Biology and Evolution*. **37**, 1530–1534 (2020).

1086 98. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. von Haeseler, L. S. Jermiin, ModelFinder: fast
1087 model selection for accurate phylogenetic estimates. *Nat Methods*. **14**, 587–589 (2017).

1088 99. D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, L. S. Vinh, UFBoot2: Improving the
1089 Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*. **35**, 518–522 (2018).

1090 100. C. Jain, L. M. Rodriguez-R, A. M. Phillippy, K. T. Konstantinidis, S. Aluru, High throughput ANI
1091 analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*. **9**, 5114
1092 (2018).

1093 101. D. M. Emms, S. Kelly, OrthoFinder: phylogenetic orthology inference for comparative
1094 genomics. *Genome Biology*. **20**, 238 (2019).

1095 102. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7:
1096 Improvements in performance and usability. *Molecular Biology and Evolution*. **30**, 772–780
1097 (2013).

1098 103. A. Criscuolo, S. Gribaldo, BMGE (Block Mapping and Gathering with Entropy): a new software
1099 for selection of phylogenetic informative regions from multiple sequence alignments. *BMC*
1100 *Evolutionary Biology*. **10**, 210 (2010).

1101 104. S. Guindon, J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, O. Gascuel, New Algorithms
1102 and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of
1103 PhyML 3.0. *Systematic Biology*. **59**, 307–321 (2010).

1104

1105 **Acknowledgments**

1106 We thank the Biosciences Microscopy Unit (University of Melbourne) and the Biological
1107 Optical Microscopy Platform (University of Melbourne) for the use of their confocal
1108 microscopes and laser capture dissection, and particularly Dr Gabriela Segal, Dr Shane
1109 Cheung and Dr Sam Mills for their valuable assistance. We thank the Melbourne Histology
1110 Platform (University of Melbourne), Laura Leone, and Lisa Foster for their assistance with
1111 sample sectioning. We also thank the Melbourne Research Cloud and the Life Science
1112 Compute Cluster (LiSC; http://cube.univie.ac.at/lisc) for providing the high-performance
1113 computing instances needed for this work. We are grateful to Dr Sanjida Topa and Ms Isini
1114 Buthgamuwa for their efforts in trying to culture *Endozoicomonas*.

1115 **Funding:** Australian Research Council Laureate Fellowship FL180100036 (MJHvO); Paul
1116 G. Allen Family Foundation; Austrian Science Fund grant no. P32112 (AC)

1117 **Author contributions:** Conceptualization: JM, LB, MJHvO; Formal analysis: JM, KT, AC,
1118 GP, MH; Funding acquisition: MJHvO, NC; Investigation: JM, AvdM, KD, CRG, SS, NC;
1119 Methodology: JM; Supervision: LB, MJHvO; Visualization: JM, KT, AC; Writing - Original
1120 Draft Preparation: JM, MJHvO. All authors have read and edited the manuscript.

1121 **Competing interests:** The authors declare that they have no competing interests.

1122 **Data availability:** Raw data are available under NCBI BioProject IDs PRJNA891910 (Adult
1123 CAMAs, MiSeq and NovaSeq raw data, MAGs), PRJNA891898 (Larva CAMAs, MiSeq raw
1124 data), PRJNA891892 (Whole larva microbiome, MiSeq raw data).