

# Lawrence Berkeley National Laboratory

LBL Publications

## Title

Metagenomics-resolved genomics provides novel insights into chitin turnover, metabolic specialization, and niche partitioning in the octocoral microbiome

## Permalink

<https://escholarship.org/uc/item/5hw1c3zx>

## Journal

Microbiome, 10(1)

## ISSN

2049-2618

## Authors

Keller-Costa, Tina

Kozma, Lydia

Silva, Sandra G

et al.

## Publication Date

2022

## DOI

10.1186/s40168-022-01343-7

## Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

RESEARCH

Open Access



# Metagenomics-resolved genomics provides novel insights into chitin turnover, metabolic specialization, and niche partitioning in the octocoral microbiome

Tina Keller-Costa<sup>1,2\*</sup>, Lydia Kozma<sup>1,3</sup>, Sandra G. Silva<sup>1,2</sup>, Rodolfo Toscan<sup>4</sup>, Jorge Gonçalves<sup>5</sup>, Asunción Lago-Lestón<sup>6</sup>, Nikos C. Kyrpides<sup>7</sup>, Ulisses Nunes da Rocha<sup>4</sup> and Rodrigo Costa<sup>1,2,5\*</sup>

## Abstract

**Background:** The role of bacterial symbionts that populate octocorals (Cnidaria, Octocorallia) is still poorly understood. To shed light on their metabolic capacities, we examined 66 high-quality metagenome-assembled genomes (MAGs) spanning 30 prokaryotic species, retrieved from microbial metagenomes of three octocoral species and seawater.

**Results:** Symbionts of healthy octocorals were affiliated with the taxa *Endozoicomonadaceae*, *Candidatus Thioglobaceae*, *Metamycoplasmataceae*, unclassified *Pseudomonadales*, *Rhodobacteraceae*, unclassified *Alphaproteobacteria* and *Ca. Rhabdochlamydiaceae*. Phylogenomics inference revealed that the *Endozoicomonadaceae* symbionts uncovered here represent two species of a novel genus unique to temperate octocorals, here denoted *Ca. Gorgonimonas eunicellae* and *Ca. Gorgonimonas leptogorgiae*. Their genomes revealed metabolic capacities to thrive under suboxic conditions and high gene copy numbers of serine-threonine protein kinases, type 3-secretion system, type-4 pili, and ankyrin-repeat proteins, suggesting excellent capabilities to colonize, aggregate, and persist inside their host. Contrarily, MAGs obtained from seawater frequently lacked symbiosis-related genes. All *Endozoicomonadaceae* symbionts harbored endo-chitinase and chitin-binding protein-encoding genes, indicating that they can hydrolyze the most abundant polysaccharide in the oceans. Other symbionts, including *Metamycoplasmataceae* and *Ca. Thioglobaceae*, may assimilate the smaller chitin oligosaccharides resulting from chitin breakdown and engage in chitin deacetylation, respectively, suggesting possibilities for substrate cross-feeding and a role for the coral microbiome in overall chitin turnover. We also observed sharp differences in secondary metabolite production potential between symbiotic lineages. Specific *Proteobacteria* taxa may specialize in chemical defense and guard other symbionts, including *Endozoicomonadaceae*, which lack such capacity.

**Conclusion:** This is the first study to recover MAGs from dominant symbionts of octocorals, including those of so-far unculturable *Endozoicomonadaceae*, *Ca. Thioglobaceae* and *Metamycoplasmataceae* symbionts. We identify a thus-far unanticipated, global role for *Endozoicomonadaceae* symbionts of corals in the processing of chitin, the most

\*Correspondence: tinakellercosta@tecnico.ulisboa.pt; rodrigoscosta@tecnico.ulisboa.pt

<sup>1</sup> Institute for Bioengineering and Biosciences, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

abundant natural polysaccharide in the oceans and major component of the natural zoo- and phytoplankton feed of octocorals. We conclude that niche partitioning, metabolic specialization, and adaptation to low oxygen conditions among prokaryotic symbionts likely contribute to the plasticity and adaptability of the octocoral holobiont in changing marine environments. These findings bear implications not only for our understanding of symbiotic relationships in the marine realm but also for the functioning of benthic ecosystems at large.

**Keywords:** Symbiosis, MAGs, Gorgonians, Facultative anaerobes, Chitinases, Secondary metabolism, *Endozoicomonadaceae*, *Thioglobaceae*, *Mollicutes*

## Background

Octocorals (Octocorallia, Anthozoa, Cnidaria) are an integral part of benthic marine ecosystems, increasing habitat complexity and biodiversity where they abound [1, 2]. Octocorals differ from scleractinian corals (Hexacorallia) by the eightfold symmetry of their polyps and usually lack a stony skeleton, though their tissue is supported by small, internal skeletal elements (calcite sclerites) [3]. They amount to over 3500 species [4], possessing worldwide distribution from polar over temperate to tropical regions and from shallow waters to the deep sea [2, 5]. Octocorals suspension feed on large quantities of debris and phyto- and zooplankton [6]. They thus play paramount roles in coastal food chains, helping to regulate primary and secondary production [6, 7]. They also form associations with various microorganisms, including micro-eukaryotes, bacteria, archaea, and viruses [reviewed in 7].

Over the past two decades, heat waves and infectious diseases have led to significant mortalities in Mediterranean and northeast Atlantic octocoral populations [8–12]. These mortality events can alter critical ecosystem processes and result in biodiversity loss in the benthos of temperate zones [9]. In other parts of the world, however, octocorals continued to thrive, while scleractinian corals faced a rapid decline from climate change. Studies in the Red Sea, the Caribbean, and the Pacific and Indian Oceans showed reef community shifts from scleractinian corals towards octocorals, increasing octocoral cover from less than 10% to nearly 50% in some regions [reviewed in 7]. It is yet to be determined why some octocorals resist while others are affected by climate change scenarios — but their associated microbiomes likely play a fundamental role in this response.

Owing to advances in high-throughput sequencing and analysis of phylogenetic marker genes (e.g., the 16S rRNA gene), our understanding of the taxonomic composition, diversity, host specificity, geographic variability, and seasonal stability of the octocoral microbiome has increased considerably in the past few years [7, 13–17]. The prokaryotic assemblages of corals, and octocorals, are frequently dominated by *Endozoicomonadaceae* and other *Gammaproteobacteria* phylotypes, in addition to

members of *Alphaproteobacteria*, *Mollicutes*, *Flavobacteriia*, *Actinobacteria* and *Spirochaetes* [7, 15, 18, 19]. *Endozoicomonadaceae* symbionts can make up to 96% of a coral prokaryotic community and are generally considered indicators of coral health [7, 15, 20]. Despite their ubiquity and abundance in corals worldwide, less than ten coral-associated *Endozoicomonas* strains (comprising five formally described *Endozoicomonas* species) exist in culture or had their genomes sequenced, as shown by a meta-analysis of over 3050 coral bacterial isolates from more than 80 coral species [21]. Among these, only two cultured *Endozoicomonas* species, *E. euniceicola* and *E. gorgoniicola*, derive from octocorals [22], and no genome sequence of an octocoral-derived *Endozoicomonadaceae* symbiont is yet publicly available. Moreover, recent studies have suggested that cultured *Endozoicomonas* isolates are phylogenetically distant from the dominant, so-far uncultured *Endozoicomonadaceae* phylotypes populating temperate octocorals [13, 15, 19]. Therefore, we still know little about the role of this “core” bacterial family in octocorals and in the functioning of benthic ecosystems at large.

Despite our current, improved view of prokaryotic diversity in corals, knowledge of the functional attributes of coral-associated prokaryotes remains limited, especially regarding the symbionts of octocorals. Collectively, recent omics studies indicate the provision of vitamins and amino acids and participation in nutrient cycling and dimethylsulfoniopropionate metabolism as putative roles of bacterial symbionts of scleractinian corals [21, 23–26]. Comparative genomics of *Alphaproteobacteria*, *Vibrio*, and *Aquimarina* isolates from octocorals further revealed the presence of various biosynthetic gene clusters (SM-BGCs) coding for polyketide, terpene, and antimicrobial peptide production, suggesting an involvement of octocoral bacteria in chemical defense [21, 27–29]. Our recent shotgun metagenomics survey showed that high abundances of eukaryotic-like proteins, exo- and endonucleases, phage lysogenization regulators, and micronutrient acquisition-related genes distinguish the prokaryotic communities of healthy from necrotic octocoral tissue, likely contributing to the stability of the symbiotic microbiome [15]. However, the functional

contributions and mechanisms of interaction of microbial symbionts within the octocoral holobiont remain largely unknown.

In this study, we examine metagenome-assembled genomes (MAGs) from 17 shotgun-sequenced microbial metagenomes of three octocoral species (healthy and necrotic *Eunicella gazella* tissue, healthy *Eunicella verrucosa*, and *Leptogorgia sarmentosa*) and surrounding seawater to connect microbial taxonomy with function in a cultivation-independent fashion. We shed light on the likely roles and mechanisms of interaction of dominant bacterial taxa in octocorals. This study is the first to retrieve and compare multiple draft genomes from uncultured symbionts of octocorals, addressing the hypothesis of microbial niche partitioning in the octocoral holobiont.

## Methods

### Metagenome samples

The 17 microbial metagenomes used for the binning of MAGs in this study are publicly available (PRJEB13222) and have been described in our earlier study [15]. Briefly, the samples were collected by scuba diving at 17-m depth on June 17, 2014, in the Atlantic Ocean off the coast of Faro, Algarve, and Portugal (“Pedra da Greta”: Latitude 36° 58′ 47.2 N, Longitude 7° 59′ 20.8 W). Branches (10–20 cm each) of 10 colonies from three octocoral (Alcyonacea, Gorgoniidae) species, *Eunicella gazella* ( $N=6$ : 3 × healthy, 3 × necrotic; EG15H, EG15N; EG16H, EG16N; EG18H, EG18N), *Eunicella verrucosa* ( $N=4$ ; EV01–EV04), and *Leptogorgia sarmentosa* ( $N=3$ ; LS06–LS08), were sampled. From *E. gazella*, branches of both healthy (H) and necrotic (N) tissue were sampled from three independent colonies, each of which displaying both conditions to directly compare the octocoral-associated microbial community in healthy versus necrotic states. The necrotic *E. gazella* tissue was characterized by a change of tissue color (from white to brown) and integrity which indicates necrosis and ultimately leads to coenenchyme detachment and loss. Octocoral samples were placed, in situ, separately in Ziploc® bags containing natural seawater. In addition, replicate samples of surrounding seawater ( $N=4$ ; SW01–SW04; ca 2 L each) were collected ca. 1 m above the corals in separate Ziploc® bags. All samples were transported to the laboratory in a cooling box within 1.5 h post sampling and immediately processed upon arrival. Octocoral branches were washed with artificial seawater to remove exogenous microorganisms and aseptically cut into smaller pieces. The soft tissue (coenenchyme and polyps) was then separated from the inner gorgonin skeleton with a scalpel and homogenized in sterile  $\text{Ca}^{2+}$  and  $\text{Mg}^{2+}$ -free artificial seawater (CMFASW: 27 g L<sup>-1</sup> NaCl, 1 g L<sup>-1</sup> NaSO<sub>4</sub>, 0.8 g

L<sup>-1</sup> KCl and 0.18 g L<sup>-1</sup> NaHCO<sub>3</sub>, 1 g of soft tissue per 9 mL CMFASW w/v) using a sterile mortar and pestle. The obtained cell homogenates were then subjected to differential centrifugation as described in [19] to retrieve microbial cell pellets. Each seawater sample was filtered through a sterile 0.22 μm nitrocellulose membrane filter (Millipore, Billerica, MA, USA; 47 mm) using a vacuum pump. Microbial pellets and seawater filters were stored at –80 °C until metagenomic DNA extraction with the UltraClean® Soil DNA isolation kit (MO BIO, Carlsbad, CA, USA) according to the manufacturer’s instructions. Equivalent amounts of DNA per sample biotope (i.e., healthy, and necrotic octocoral tissue, seawater) were sent for next-generation shotgun sequencing on an Illumina HiSeq 2500 device at MR DNA (Shallowater, TX, USA). DNA libraries were prepared for sequencing using the Nextera DNA Sample preparation kit (Illumina) after the manufacturer’s instructions and sequenced paired end with sequence depth calibrated at c. 20 million 101 bp reads per sample.

### Binning of metagenome assembled genomes (MAGs)

Metagenome reads obtained for each sample were assembled using the MetaWRAP v1.0.5 pipeline [30] and encompassed reads quality control with the MetaWRAP galore module, followed by assembly with the meta-SPAdes module 3.13.0 [31]. Eukaryotic contigs were then filtered out of the resulting assemblies using EukRep v.0.6.6 [32], generating individual “prokaryotic-enriched” assembly files per sample [15]. Metagenomic binning was performed in this study on the “prokaryotic-enriched” assemblies using MetaBAT 2 [33], Maxbin 2.0 [34], and CONCOCT [35] within the MetaWRAP binning module. Binning\_refiner [36] was used to refine bins and produce a superior bin set within MetaWRAP. Genome completeness, contamination, and strain heterogeneity scores were calculated using CheckM v1.0.11 [37] with default parameters. Following the approach of Parks et al. [38], an overall quality score was calculated for each MAG (completeness — 5 × contamination), and MAGs possessing a quality score above 50% were kept for analysis, resulting in a final dataset of 66 MAGs. We then categorized the 66 MAGs into “high-quality MAGs” (when completeness was above 90% and contamination below 5%) and “medium quality MAGs” (completeness above 50% and contamination below 10%), according to MIMAG guidelines [39]. The Microbial Genome Atlas (MiGA) [40] was used to obtain genome metrics such as number of contigs, genome size, GC content, and N50 values. MiGA (accessed on 25th of November 2021) was also used to identify the closest relative (i.e., genome of a type strain) of each MAG, based on whole-genome average amino acid identity values (AAI%).

### Taxonomic assignment and species-level similarity of MAGs

The Genome Taxonomy Database Toolkit (GTDB-Tk) [41] v.1.5.0 (release 06-RS202) was used to perform taxonomy assignment of all MAGs obtained in this study. High taxonomic rank assignments were afterwards manually curated to comply with the list of prokaryotic names with standing in nomenclature (LPSN) [42, 43] as deemed necessary. FastANI [44] was used to compute whole-genome average nucleotide identity (ANI) percent values in a pairwise fashion whenever two or more MAGs shared the same taxonomic assignment. MAGs which shared  $\geq 95\%$  ANI were considered to belong to the same species [44–46].

### Phylogenomics of the *Endozoicomonadaceae* and *Ca. Thioglobaceae* families

Owing to their high frequency across healthy octocoral samples and presumed taxonomic distinctiveness and novelty, we thoroughly explored the phylogenomic relatedness of the 11 *Endozoicomonadaceae* and six *Ca. Thioglobaceae* MAGs of this study with their closest relatives. Two phylogenomic trees were created, one for each family. Details on the genomes used for tree construction are provided in Additional file 1. Both trees were constructed with the SpeciesTreeBuilder v.01.0 application of the DOE Systems Biology Knowledgebase (KBase) [47] using the function “Insert Set of Genomes into Species Tree,” after annotating all isolate genomes, MAGs and SAGs with Prokka [48]. SpeciesTreeBuilder uses the FastTree2 algorithm [49] to infer maximum-likelihood (ML) phylogenies for large alignments. Alignments were based on a set of 49 core genes defined by Clusters of Orthologous Groups of proteins (COG) families. Graphical visualization and editing of the trees were made in iTOL v4 (Interactive Tree Of Life) [50].

### Functional annotation of MAGs

Functional annotation of MAGs encompassed the generation of COG profiles, metabolic pathway reconstruction, and genome mining for secondary metabolite biosynthetic gene clusters (SM-BGCs). COG annotation was performed for all MAGs using our in-house, automated genome annotation pipeline MeLanGE as documented on GitHub (<https://sandragodinhosilva.github.io/MeLanGE>). Briefly, all MAGs (contig fasta files) were first annotated with Prokka v1.14.6 [48] to obtain GenBank (gbk) format and amino acid fasta files. Thereafter, proteins were queried against the COGs database implemented within NCBI's Conserved Domain Database (CDD) through Reversed Position Specific Blast (RPS-BLAST) from the BLAST + suite (v2.9.0), and the best hit

per ORF, above the cutoff of  $E 1e-5$ , was selected. MAGs were further annotated with the RAST server version 2.0 [51, 52] using the RASTtk annotation scheme with default settings and the “build metabolic models” setting activated. Using the “KEGG metabolic analysis” tool within RAST, KEGG metabolic maps were constructed for selected metabolic pathways and individual MAGs (not to model metabolic interactions between species). Complete chitinase coding sequences (CDS) present on the 11 *Endozoicomonadaceae* MAGs were retrieved from RAST and characterized as described in Additional file 1. Identification of SM-BGCs across all MAGs was performed using antiSMASH v5.0 [53] with default parameters and extra features “all on”. The degree of novelty of SM-BGCs was assessed through matches with the Minimum Information about a Biosynthetic Gene cluster (MIBiG) database [54, 55].

### Data analyses and statistics

Multivariate analysis of COG-based functional profiles was carried out on Hellinger-transformed data (i.e., square root of the relative abundance of each COG entry on a MAG). Euclidean distances were then calculated from COG abundance distributions across MAGs, and a principal components analysis (PCA) was performed using PAST v3.25 [56]. One-way permutational analysis of variance (PERMANOVA) was used to test for overall differences in functional profiles among MAGs belonging to different taxonomic orders. To determine COG functions that contributed most to the dissimilarity between MAGs at the order level, similarity percentage analysis (SIMPER) was performed in PAST v 3.25 [84]. The top ten most differentiating COG functions were then plotted as vectors on the PCA diagram to explore relationships between gene functions and symbiont taxonomy. Moreover, Welch's unequal variances *t*-tests (one sided) were conducted within STAMP v2.0.953 [57] to identify COG entries that were significantly enriched on MAGs affiliated with the *Endozoicomonadaceae* ( $N=11$ ) and *Ca. Thioglobaceae* ( $N=6$ ) families. Multiple test correction was performed with the Benjamini–Hochberg method to decrease false discovery rates, and COG entries representing fivefold (*Ca. Thioglobaceae* MAGs) and tenfold (*Endozoicomonadaceae* MAGs) enrichments were selected for further analysis.

## Results

### Dataset overview

The final dataset analyzed in this study comprised 66 MAGs, 65 of which derived from *Bacteria* and one from *Archaea* (Table S1, Additional file 2). Twenty-five MAGs were obtained from healthy octocoral samples

(all species together), 14 MAGs from necrotic *E. gazella* tissue, and 27 MAGs, including the archaeal one, from seawater. Of the 66 MAGs, 30 and 36 were of high and medium quality, respectively. Genome completeness ranged from 55.03 to 99.86% with an average of 83.57% across the 66 MAGs, while contamination ranged from 0 to 4.64% with an average of 1.16% (Table 1; Table S2, Additional file 2). Average genome size and GC content

ranged from only 0.63 Mb and 22.3% in *Metamycoplasmataceae* to 4.2 Mb and 59.6% in *Ca. Inquilinaceae* (*Alphaproteobacteria*), both symbionts of healthy octocorals (Table 1, Table S1, Additional file 2). The mean and median genome size of the 66 MAGs was 2.38 Mb and 2.46 Mb, respectively, whereby genome size was more related with MAG taxonomy than with origin (Table 1; Table S2, Additional file 2).

**Table 1** Genome features of the MAGs from the 30 prokaryotic species analyzed in this study

Class	Species-level taxonomic affiliation of MAGs	N <sup>a</sup> MAGs	Genome size Mb	GC content %	N <sup>o</sup> of coding Sequences	Complete-ness %	Contami-nation %
<i>Mollicutes</i>	<i>Metamycoplasmataceae</i> (DT-68 sp.)	3	0.63 ± 0.13	22.33 ± 0.72	685 ± 63	84.07 ± 17.26	0.08 ± 0.14
<i>Chlamydia</i>	<i>Ca. Rhabdochlamydiaceae</i> (SZUA-160 sp.)	1	0.88	44.38	1015	60.07	0.00
<i>Gammaproteobacteria</i>	<i>Ca. Gorgonimonas eunicellae</i>	7	2.79 ± 0.13	29.68 ± 0.05	2177 ± 121	73.17 ± 1.38	1.52 ± 0.96
	<i>Ca. Gorgonimonas leptogorgiae</i>	4	3.15 ± 0.60	29.77 ± 0.29	2603 ± 355	70.20 ± 4.15	1.06 ± 0.40
	<i>Ca. Thiocorallibacter gorgonii</i>	2	1.41 ± 0.16	29.19 ± 0.17	2612 ± 6	90.95 ± 6.72	1.24 ± 0.35
	<i>Ca. Microaerophilica antagostica</i>	4	1.85 ± 0.12	37.47 ± 0.07	1763 ± 101	93.05 ± 2.33	0.82 ± 0.50
	<i>Pseudomonadales</i> (DT-91 sp.)	2	3.29 ± 0.52	40.98 ± 0.09	3505 ± 240	77.02 ± 18.70	1.35 ± 1.90
	<i>Cardiobacteriales</i> sp.	2	2.50 ± 0.20	42.72 ± 0.05	2560 ± 145	85.70 ± 6.34	0.57 ± 0.00
<i>Alphaproteobacteria</i>	<i>Luminiphilus</i> sp009886815	3	3.19 ± 0.48	52.79 ± 0.19	3307 ± 580	85.67 ± 10.49	3.39 ± 1.10
	<i>Ruegeria</i> sp. (EG16H_Bin1)	1	2.49	56.18	3405	55.03	0.38
	<i>Ruegeria</i> sp900313035	1	4.16	56.55	4564	87.01	1.38
	<i>Yoonia</i> sp. (EG16N_Bin1)	1	3.62	56.98	3750	99.39	0.38
	<i>Yoonia</i> sp. (EG16N_Bin5)	1	2.69	53.45	2878	95.52	1.22
	<i>Yoonia</i> sp. (EV01H_Bin2)	1	3.05	57.15	3798	83.28	1.47
	<i>Aliiroseovarius</i> sp. (EG18N_Bin4)	1	2.06	51.31	2606	70.61	1.13
	<i>Planktomarina</i> sp002683685	4	2.28 ± 0.26	54.87 ± 0.07	2392 ± 105	93.47 ± 7.83	0.87 ± 0.70
	<i>Amylibacter</i> sp900197625	4	1.70 ± 0.27	37.10 ± 0.70	1855 ± 183	75.16 ± 14.49	1.54 ± 1.37
	<i>Rhodobacteraceae</i> (LGRT01 sp001642945)	3	2.91 ± 0.05	51.36 ± 0.05	2903 ± 27	96.36 ± 0.81	0.36 ± 0.19
	<i>Ca. Parvibaculales</i> (UBA8337 sp900197605)	2	1.35 ± 0.03	54.77 ± 0.02	1646 ± 37	80.11 ± 3.25	1.74 ± 0.00
	<i>Lentilitoribacter</i> sp900537175	1	3.63	44.39	3580	97.27	0.52
	Unclassified <i>Rhizobiaceae</i> (EG18N_Bin5)	1	2.79	58.16	2656	98.96	0.08
	Unclassified <i>Rhizobiaceae</i> (EG18N_Bin7)	1	2.58	41.22	2613	93.88	0.2
	<i>Ca. Inquilinaceae</i> (JAAAOG01 sp.)	1	4.23	59.61	5446	68.01	0.44
	Unclassified <i>Alphaproteobacteria</i>	2	1.67 ± 0.20	50.03 ± 0.16	1668 ± 224	89.25 ± 4.56	0.81 ± 0.38
<i>Flavobacteriia</i>	<i>Aquimarina</i> sp. (EG16N_Bin3)	1	3.83	32.22	3501	93.00	0.67
	<i>Flavobacteriaceae</i> MS024-2A sp002167945	1	1.49	37.84	1570	66.19	2.73
	<i>Flavobacteriaceae</i> (MS024-2A sp002292265)	1	1.39	41.03	1492	74.42	2.26
	<i>Schleiferiaceae</i> (UBA10364 sp002387615)	4	1.83 ± 0.07	44.29 ± 0.04	1796 ± 71	95.44 ± 2.09	0.62 ± 0.34
<i>Verrucomicrobiae</i>	<i>Akkermansiaceae</i> (UBA985 sp003527555)	5	2.57 ± 0.79	48.44 ± 0.71	2410 ± 376	88.65 ± 18.1	1.13 ± 0.39
<i>Poseidoniia</i>	<i>Ca. Poseidoniaceae</i> (MGIIa-L1 sp.)	1	1.56	52.10	1772	61.82	1.60

<sup>a</sup> For species with more than one MAG, average values and standard deviations are shown

### Taxonomic affiliation of MAGs

The 66 MAGs belonged to six phyla, seven classes, 15 orders, 16 families, and at least 30 species as defined by 95% ANI thresholds (Fig. 1a–c; Table S1, Additional file 2). The 21 *Gammaproteobacteria* MAGs retrieved from octocorals all represented so-far uncultured and unclassified lineages, with 11 MAGs affiliating with the *Endozoicomnadaceae* family (Fig. 1b). Of these, nine derived from the microbiomes of healthy octocoral tissue and two from necrotic tissue. Six octocoral-derived MAGs affiliated with the family Ca. *Thioglobaceae*, five of them obtained from the microbiomes of healthy tissue. Moreover, two *Gammaproteobacteria* MAGs from healthy *L. sarmentosa* were affiliated with candidate taxon DT-91 of the order *Pseudomonadales*, while two MAGs from necrotic *E. gazella* were identified as unclassified *Cardiobacterales*. Contrarily, the three *Gammaproteobacteria* MAGs found in seawater belonged to the genus *Luminiphilus* (*Haliaceae*, *Cellvibrionales*). Notably, four MAGs from healthy octocoral tissue were affiliated with intracellular bacterial symbionts [58] of the families *Metamycoplasmataceae* and Ca. *Rhabdochlamydiaceae*. The five *Alphaproteobacteria* MAGs from healthy octocoral tissue were either unclassified at order level ( $N=2$ ) or belonged to the candidate family *Inquilinaceae* and the *Rhodobacteraceae* genera *Ruegeria* and *Yoonia* (Fig. 1c). Among the MAGs obtained from necrotic *E. gazella* tissue was also one affiliating with the genus *Aquimarina* (*Bacteroidetes*), a taxon frequently cultured from octocorals [19, 28], as well as three MAGs affiliating with the *Rhizobiaceae* family, one of them identified as *Lentilitoribacter*. Overall, little overlap was observed at species level between MAGs reconstructed from healthy octocoral tissue versus necrotic tissue versus seawater. Of the 11 bacterial species recovered from healthy octocorals, only two species, the here proposed Ca. *Gorgonimonas eunicellae* and Ca. *Thiocorallibacter gorgonii*, were also recovered from necrotic samples. However, their average genome coverage, a proxy for relative abundance, was much higher in healthy than necrotic *E. gazella* tissue (nearly sevenfold difference in coverage for Ca. *G. eunicellae* and nearly twofold difference in coverage for Ca. *T. gorgonii* in healthy vs. necrotic samples; Fig. 1c).

### Phylogenomics of *Endozoicomnadaceae* and Ca. *Thioglobaceae* MAGs

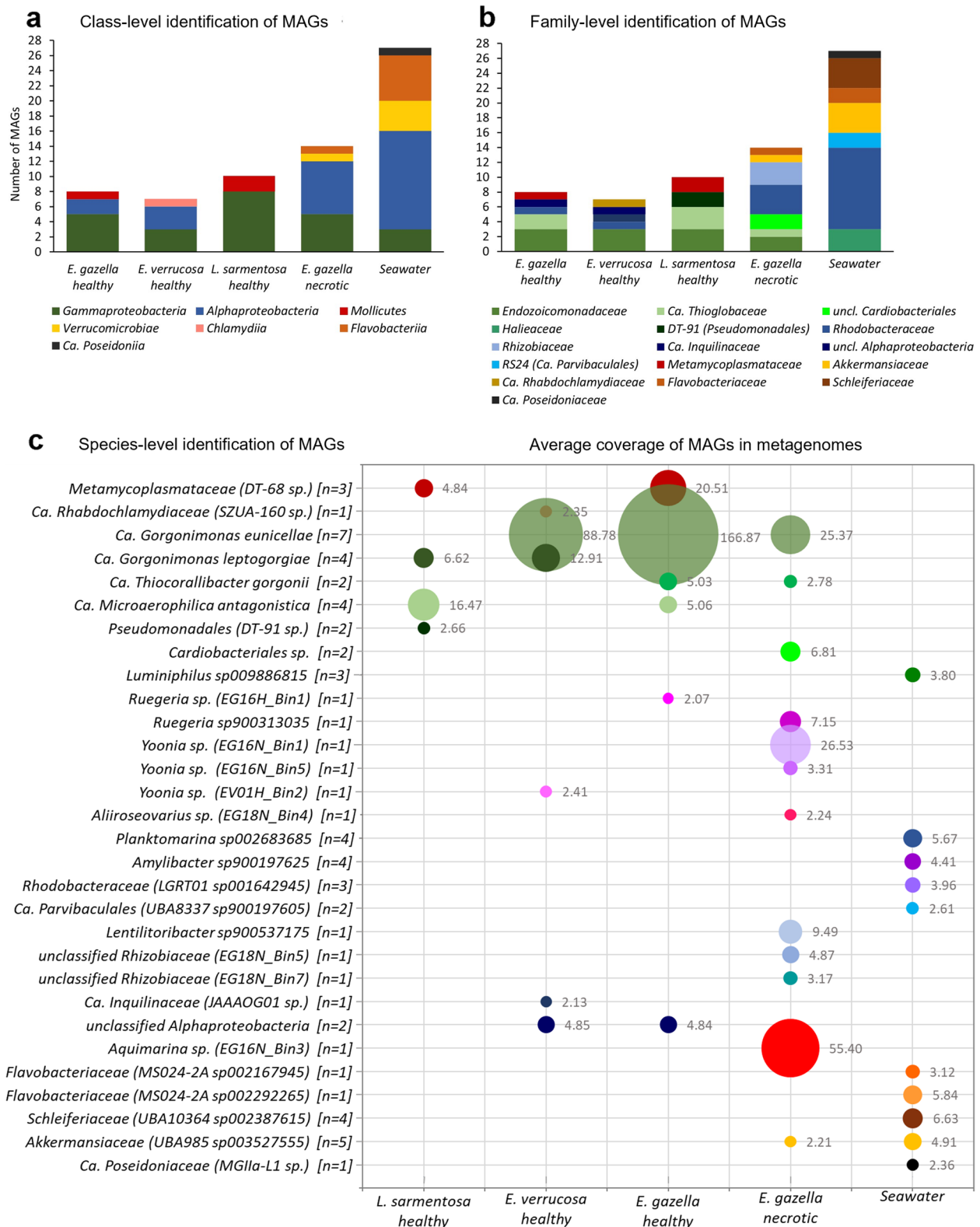
Phylogenomics analysis of the *Endozoicomnadaceae* family comprised 29 publicly available type genomes, MAGs, and SAGs plus the 11 octocoral-derived MAGs retrieved in this study (Fig. 2). The latter formed their own, well supported and deeply branching clade, separate from the genomes of all formally described genera

with cultured representatives (i.e., *Kistimonas*, *Parendo- zoicomonas*, and *Endozoicomonas*). This clade comprised two subclusters, each representing a novel species, sharing ~89.8% ANI between them. Subcluster 1 contained seven MAGs, all obtained from *Eunicella* hosts. Subcluster 2 comprised four MAGs, three from *L. sarmentosa* and one from *E. gazella*. The closest type strain genome to the 11 *Endozoicomnadaceae* MAGs was *Endozoicomonas atrinae* GCA\_001647025T, which shared only 52–53% AAI with them (Table S1, Additional file 2), well below the 65% threshold considered by MiGA for same-genus classification. This indicates that the 11 *Endozoicomnadaceae* MAGs represent two distinct species, forming a novel yet uncultured genus unique to temperate octocorals. We propose the names *Candidatus Gorgonimonas eunicellae* (corresponding to subcluster 1) and Ca. *Gorgonimonas leptogorgiae* (corresponding to subcluster 2) for the two species.

Phylogenomics inference of the Ca. *Thioglobaceae* family showed that the six octocoral-derived *Thioglobaceae* MAGs formed two separate clusters, representing distinct, novel species as judged by ANI values way below 80% compared with the remainder genomes of the family (Fig. S1, Additional file 1). The first clade comprised two *E. gazella*-derived MAGs which formed a subcluster within other Ca. *Thioglobaceae* clusters of the genera *Thioglobus* and *Thiomultimodus*. The second clade was composed of the other four *Thioglobaceae* MAGs, derived from healthy *L. sarmentosa* and *E. gazella* tissue, which formed a well-supported, deeply branching phylogenomic node on their own, sharing only 46% AAI with genomes of their closest type strains, namely *Sulfurivirga caldicuralii* GCA\_900141795T and *Thiohalobacter thiocyanaticus* GCA\_003932505T (Table S1, Additional file 2). This indicates that the six octocoral-derived Ca. *Thioglobaceae* MAGs not only represent two distinct species but most likely two distinct, novel genera, here proposed Ca. *Thiocorallibacter gorgonii* and Ca. *Microaerophilica antagonistica* (Fig. S1, Additional file 1), which so far lack cultured representatives.

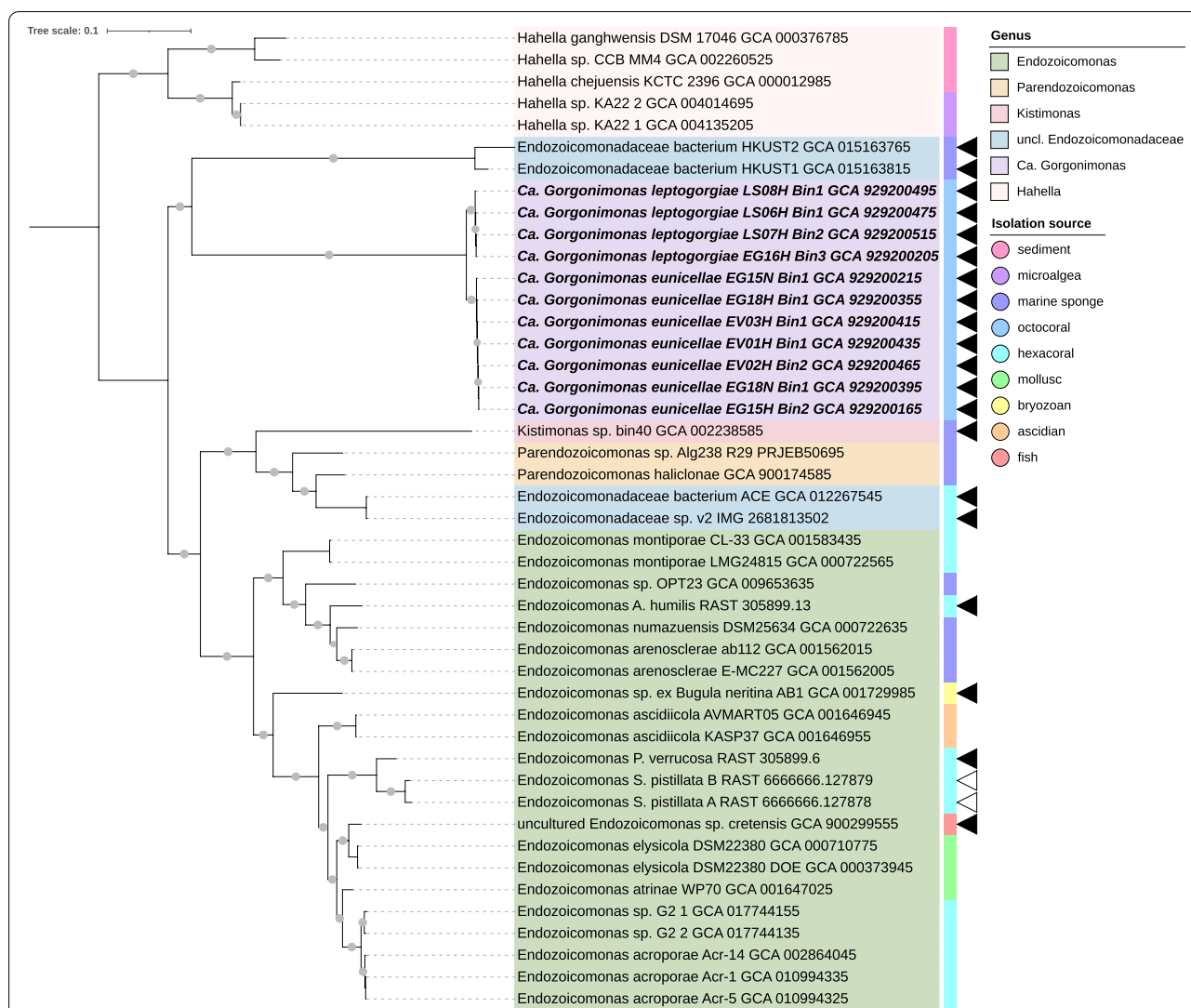
### Functional profiling of MAGs from octocoral and seawater microbiomes

Multivariate analysis based on COG functional profiles showed that the MAGs grouped primarily according to their (order level) taxonomic affiliations (PERMANOVA,  $F=9.869$ ,  $P=0.0001$ ) (Fig. 3). The 11 *Endozoicomnadaceae* MAGs formed a very tight cluster, much distant from all other MAGs. Such separate clustering was mostly determined by the high copy number of ankyrin repeat motifs (COG0666) and serine/threonine protein kinase-encoding genes (COG0515) on the



**Fig. 1** Taxonomic affiliation at **a** class, **b** family, and **c** species level of the 66 metagenome-assembled genomes (MAGs) obtained from the microbial metagenomes of healthy and necrotic *Eunicella gazella* tissue, healthy *E. verrucosa*, and *Leptogorgia sarmentosa*, and seawater. Average genome coverage values (inferred by mapping unassembled reads against MAG contigs) represent a relative abundance estimate of the prokaryotes in each microhabitat (**c**)



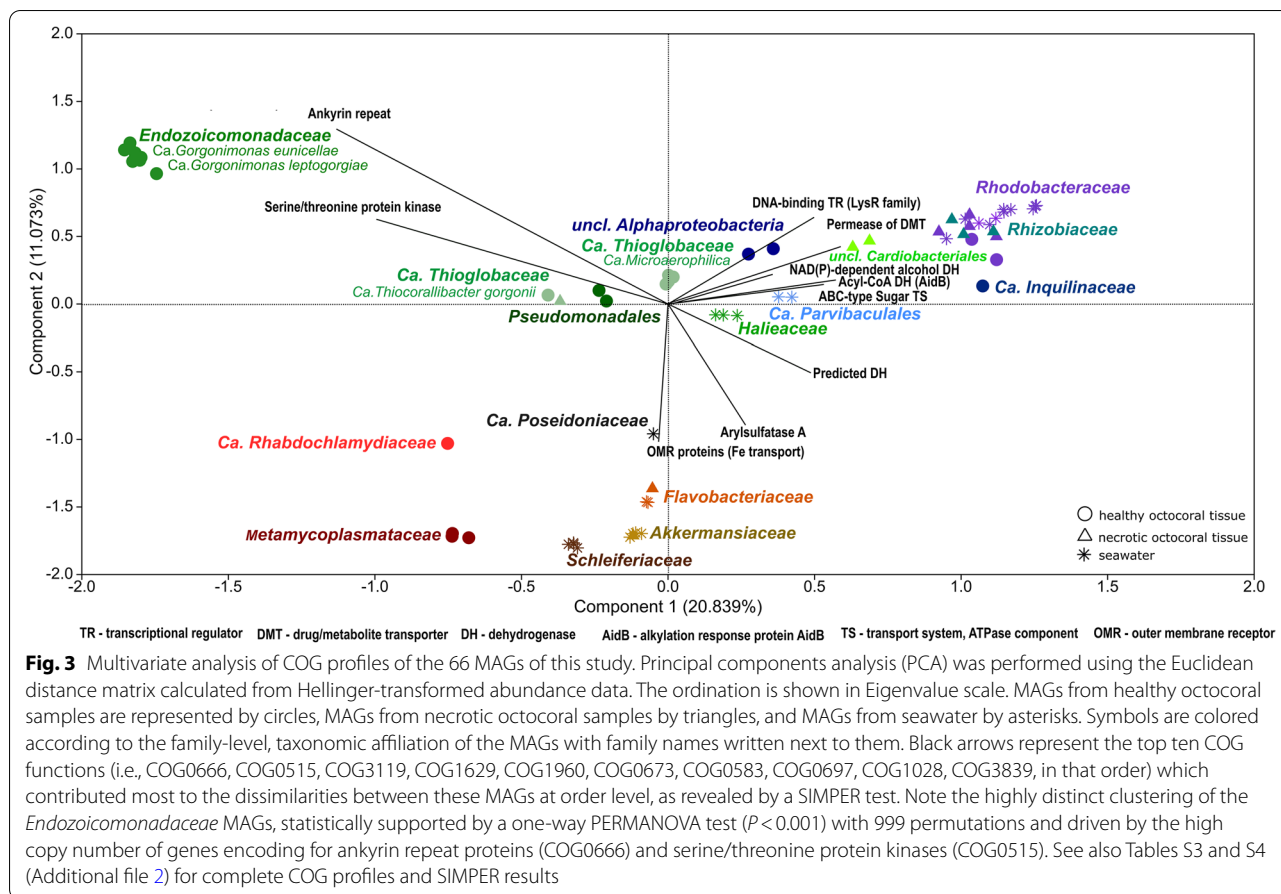


**Fig. 2** Phylogenomic analysis of the *Endozoicomonadaceae* family using the speciesTreeBuilder v.01.0. Evolutionary history was inferred by using a maximum likelihood method (FastTree 2) based on alignment similarity of a set of 49 cores, universal genes defined by Clusters of Orthologous Groups of proteins (COG) gene families. Gray dots on the branches indicate bootstrap support of > 70%. Black triangles indicate a MAG, white triangles a SAG, and the remaining genomes were from isolates. The 11 *Endozoicomonadaceae* MAGs of this study are highlighted in bold-italics. All other *Endozoicomonadaceae* genomes ( $n = 29$ ) were publicly available on RAST, IMG, or NCBI. Assembly accession numbers are given next to the strain names. Five *Hahella* spp. genomes of the closely related *Hahellaceae* family were used as outgroup to root the tree. The colored bar next to the tree shows the isolation source of the genomes. Note that all *Endozoicomonadaceae* genomes (including MAGs and SAGs) are derived from the microbiomes of marine animals, mainly marine invertebrates. The tree is drawn to scale and was style edited in iTOL

*Endozoicomonadaceae* MAGs. SIMPER analysis showed that these two COGs were indeed the functions that contributed most to the dissimilarity between all MAGs at order level (Tables S3, S4, Additional file 2).

The six *Ca. Thioglobaceae* MAGs formed two, well-separated clusters in the ordination space, one comprising the two *Ca. Thiocorallibacter gorgonii* MAGs and another one with the four *Ca. Microaerophilica antagonistica* MAGs, congruent with our phylogenomic

assessment. The positioning of *Alphaproteobacteria* MAGs in the PCA diagram was influenced, among others, by the presence and abundance of genes encoding for LysR family transcriptional regulator (COG0583), ABC sugar transport system (COG3839), Acyl-CoA and NAD(P)-dependent alcohol dehydrogenases (COG1960 and COG1028), and drug metabolite transporters (COG0697).



**Functional features enriched in *Endozoicomonadaceae* symbionts of octocorals**

The 11 *Endozoicomonadaceae* MAGs were drastically enriched in COGs related to eukaryotic-like proteins (Fig. 4), mainly ankyrin repeats ( $q < 0.0001$ ; Welch’s test) and, to a lesser extent, WD40 repeats and tetratricopeptide repeats (Table S5, Additional file 2, Fig. S2, Additional file 1). These MAGs also displayed high abundance of COG entries related to the type 3 secretion system ( $q < 0.0001$ ; Welch’s test), serine/threonine protein kinases, serine protease inhibitors ( $q < 0.0001$ ; Welch’s test), and the membrane-anchored periplasmic protein YejM. Other typical features of all 11 *Endozoicomonadaceae* MAGs were several COG entries associated with type 4-pilus (tfp) production (Figs. 4 and 5), and the consistent presence of COG3206 encoding a protein involved

in exopolysaccharide (EPS) biosynthesis, which was not observed to such extent in any of the other 55 MAGs investigated here. The two *Endozoicomonadaceae* species identified in this study were distinguished by a consistent presence of serine/threonine phosphatase encoding genes on the four MAGs of *Ca. Gorgonimonas leptogorgiae* which were absent on the seven MAGs of *Ca. G. eunicellae*.

The *Endozoicomonadaceae* MAGs show capacity for pyruvate metabolism and to convert acetyl-CoA to acetate via acetyl phosphate, through a phosphate acetyl transferase (EC 2.3.1.8) and an acetate kinase (EC 2.7.2.1), respectively, a process able to generate ATP independently from aerobic conditions (Fig. S3, Additional file 1). The consistent presence of pyruvate formate-lyase-activating enzyme encoding genes (COG1189) on the 11

(See figure on next page.)

**Fig. 4** Functional analysis of the 66 MAGs according to their COG profiles. Values for each entry represent the percentage of COGs assigned to a given function relative to the total number of COGs annotated per MAG. The color code from blue over yellow to dark red reflects an increase in the percentage of COGs related to each function (blue = 0% (minimum value); yellow = 50% percentile; dark red = 3.66% (maximum value). When functions were represented by more than one COG entry across the data set, the coding sequence (CDS) counts of these functionally belonging COGs were summed, and the number of COG entries that contributed to each function is given in brackets behind the COG description. Identification of COGs that contributed to this figure can be found in Tables S5A and S5B, Additional file 2

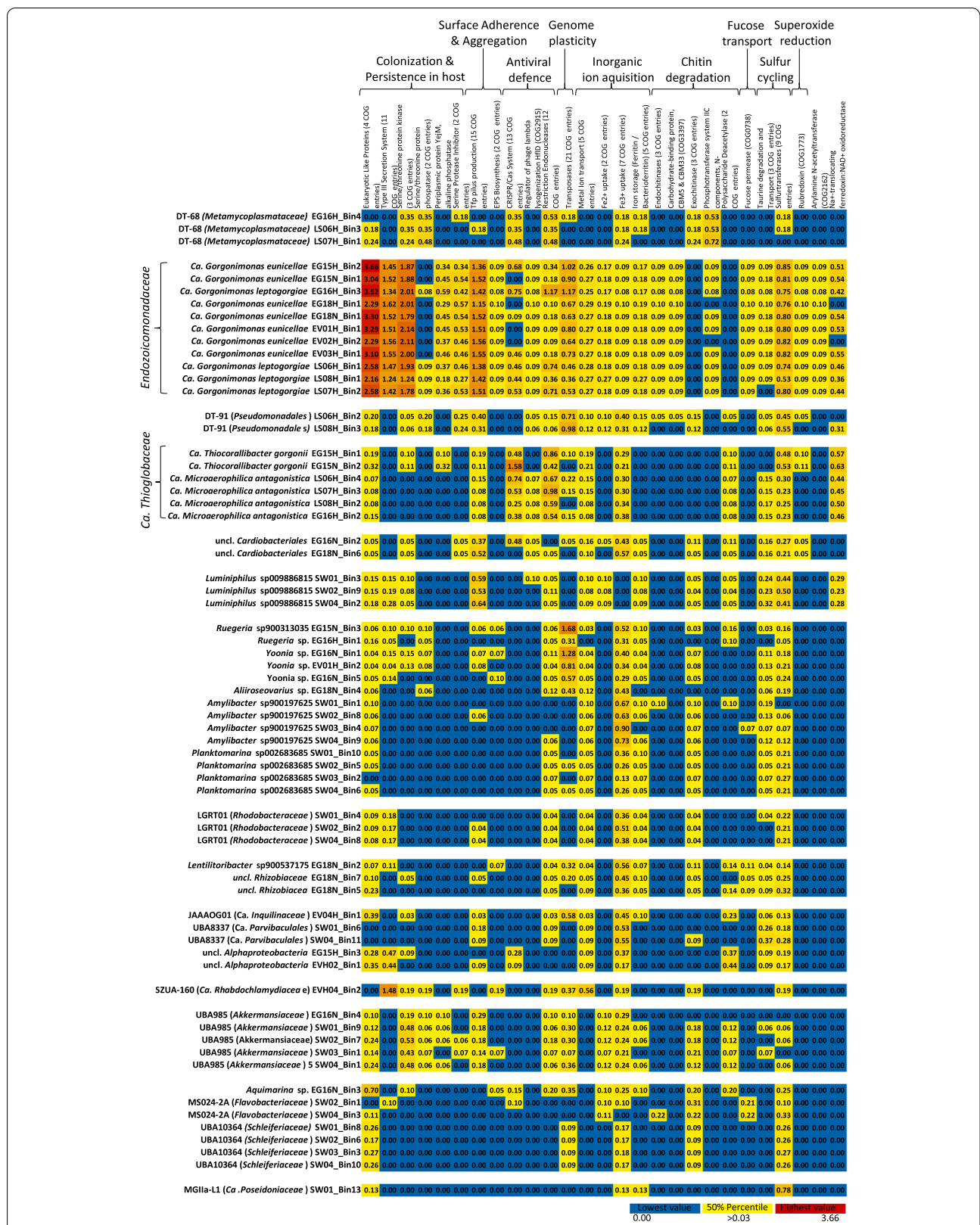


Fig. 4 (See legend on previous page.)

*Endozoicomonadaceae* genomes (Fig. S2, Additional file 1, Welch's test  $q < 0.0001$ ) further suggests that these symbionts may supply the citric acid cycle with acetyl-CoA from pyruvate during anaerobic glycolysis. Indication for adaptation of *Endozoicomonadaceae* symbionts to suboxic conditions was also found through the consistent presence of *feoA/B* genes (COG1918, COG0370) encoding for ferrous iron ( $\text{Fe}^{2+}$ ) uptake systems, distinguishing this taxon from the other 55 MAGs of this study (Welch's *t*-test,  $q < 0.0001$ ). In this regard, *Endozoicomonadaceae* MAGs were also distinguished by the presence of rubredoxin encoding genes (Figs. 4, 5 and 6; Fig. S2, Additional file 1), a class of iron-containing proteins which play an important role in superoxide reduction that can be found in several anaerobic and sulfate-reducing bacteria.

All *Endozoicomonadaceae* MAGs possessed a gene encoding an endo-chitinase (COG3469) involved in the extracellular breakdown of chitin polymers (Figs. 4, 5 and 6). We found a high degree of novelty within these 11 endo-chitinase encoding genes, as they clustered into two distinct groups and possessed less than 50% amino acid sequence similarity to publicly available endo-chitinases (Blastp search). Remarkably, these groups mirror the phylogenomic relatedness of the *Endozoicomonadaceae* MAGs, representing two endo-chitinase gene clades, each from *Ca.* species *Gorgonimonas leptogorgiae* and *G. eunicellae* (Fig. S4, Additional file 1). Protein family (Pfam) analysis confirmed the presence of a GH18 domain with an active site on all 11 genes. These genes were all complete (start and stop codon present) and carried a signal peptide sequence, indicating the protein can be excreted from the cell (Table S6A, Additional file 2). We also screened all publicly available genomes from cultured and uncultured *Endozoicomonadaceae* representatives and detected endo-chitinases on 32 out of 42 *Endozoicomonadaceae* genomes (Table S6B, Additional file 2). Moreover, genes for exo-chitinases, the enzymes that cleave the smaller chito-oligomer products into mono-sugars, were present in *Metamycoplasmataceae*, *Ca. Rhabdochlamydiaceae*, and *Pseudomonadales* symbionts of octocorals, while polysaccharide/chitin deacetylase encoding genes, which lead to the production of chitosan, were found in *Ca. Thioglobaceae* and several *Alphaproteobacteria* symbionts.

#### Metabolic inference of *Ca. Thioglobaceae* symbionts of octocorals

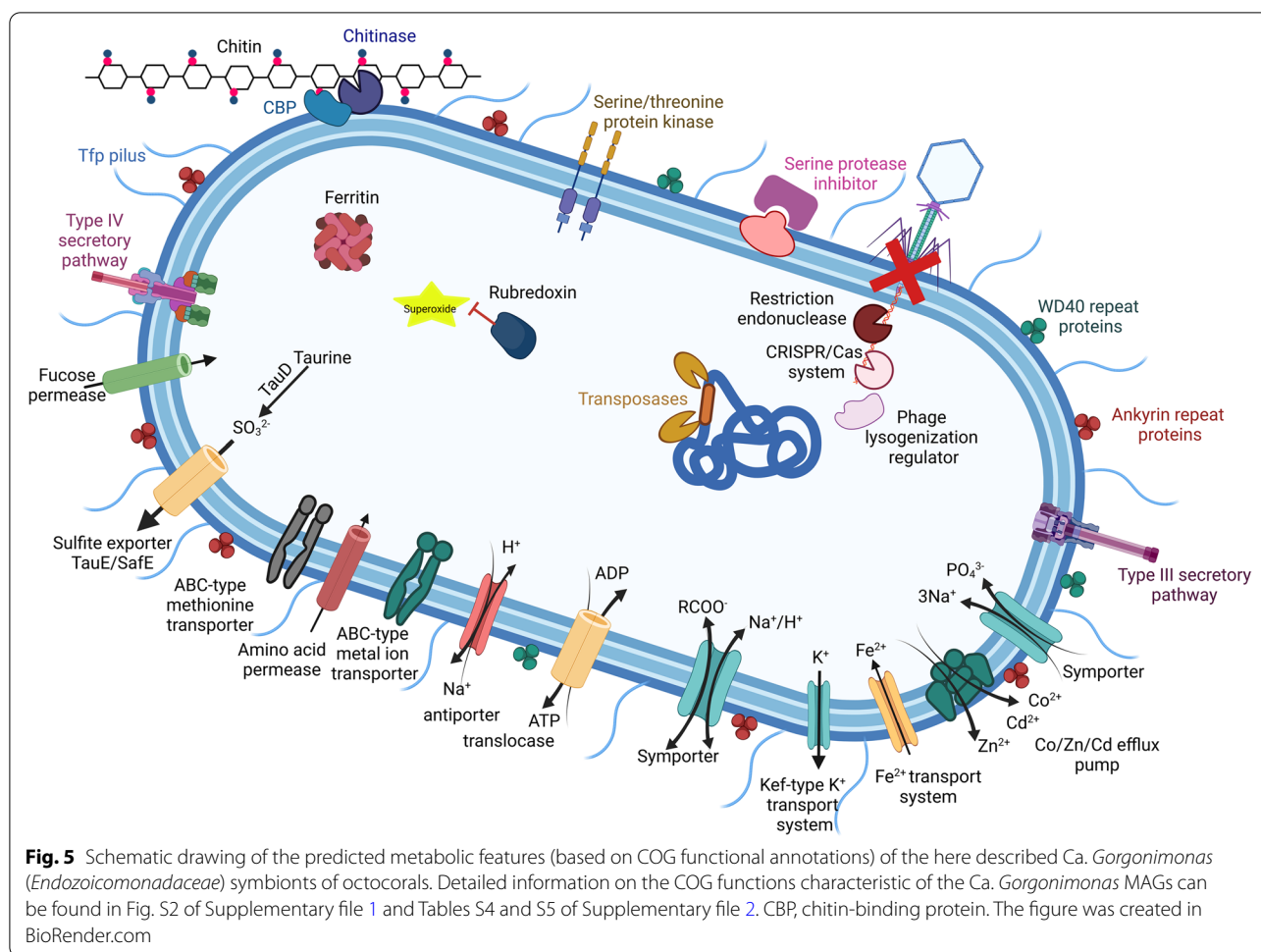
The *Ca. Thioglobaceae* MAGs were significantly enriched in CRISPR/Cas system-associated endonuclease Cas2 (COG1343) (Fig. S5, Additional file 1). Several other CRISPR/Cas protein-encoding genes were found on all *Ca. Thioglobaceae*, all *Metamycoplasmataceae*, and many *Endozoicomonadaceae*

MAGs (Fig. 4; Table S5, Additional file 2). *Ca. Thioglobaceae* MAGs were further significantly enriched in  $\text{Na}^+$ -translocating ferredoxin:NAD<sup>+</sup> oxidoreductase (Rnf complex) encoding genes (Fig. S5, Additional file 1), which were also present in great abundance in nine of the 11 *Endozoicomonadaceae* MAGs. An assimilatory sulfite reductase (EC 1.8.1.2) was also found on the *Ca. Thioglobaceae* MAGs, pointing towards a role in sulfur cycling in the octocoral holobiont. The four *Ca. Microaerophilica antagonistica* MAGs encoded for versatile taurine utilization pathways and its metabolism to aminoacetaldehyde and sulfite, via taurine deoxygenase (TauD; EC 1.14.11.17), or to sulfoacetaldehyde via taurine-pyruvate-aminotransferase (EC 2.6.1.77) (Fig. S6, Additional file 1).

All *Ca. Thioglobaceae* MAGs showed an extensive genetic repertoire for ammonium assimilation and transformation of inorganic nitrogen into amino acids, possessing genes coding for glutamine and asparagine synthetase (EC 6.3.1.2; EC 6.3.5.4), L-asparaginase (EC 3.5.1.1), glutamine amidotransferase chain of NAD synthetase (EC 6.3.1.5), aminoethyltransferase (EC 2.1.2.10), and glutamate synthase (EC 1.4.1.13; EC 1.4.7.1). The latter was significantly enriched in this symbiotic family compared to the other 60 MAGs obtained in this study (one-sided Welch's test,  $p > 0.001$ , Fig. S5, Additional file 1). Finally, *Ca. Thiocorallibacter gorgonii* MAG EG15H\_bin1 stood out as the only MAG harboring ribulose 1,5-bisphosphate carboxylase — RuBisCo (EC 4.1.1.39, COG4451, COG1850) and several other genes involved in the reductive dicarboxylate cycle (Fig. S7, Additional file 1), suggesting a chemoautotrophic lifestyle and “dark carbon fixation” ability of this octocoral symbiont.

#### Genomic features of prokaryotes from necrotic octocoral tissue and surrounding seawater

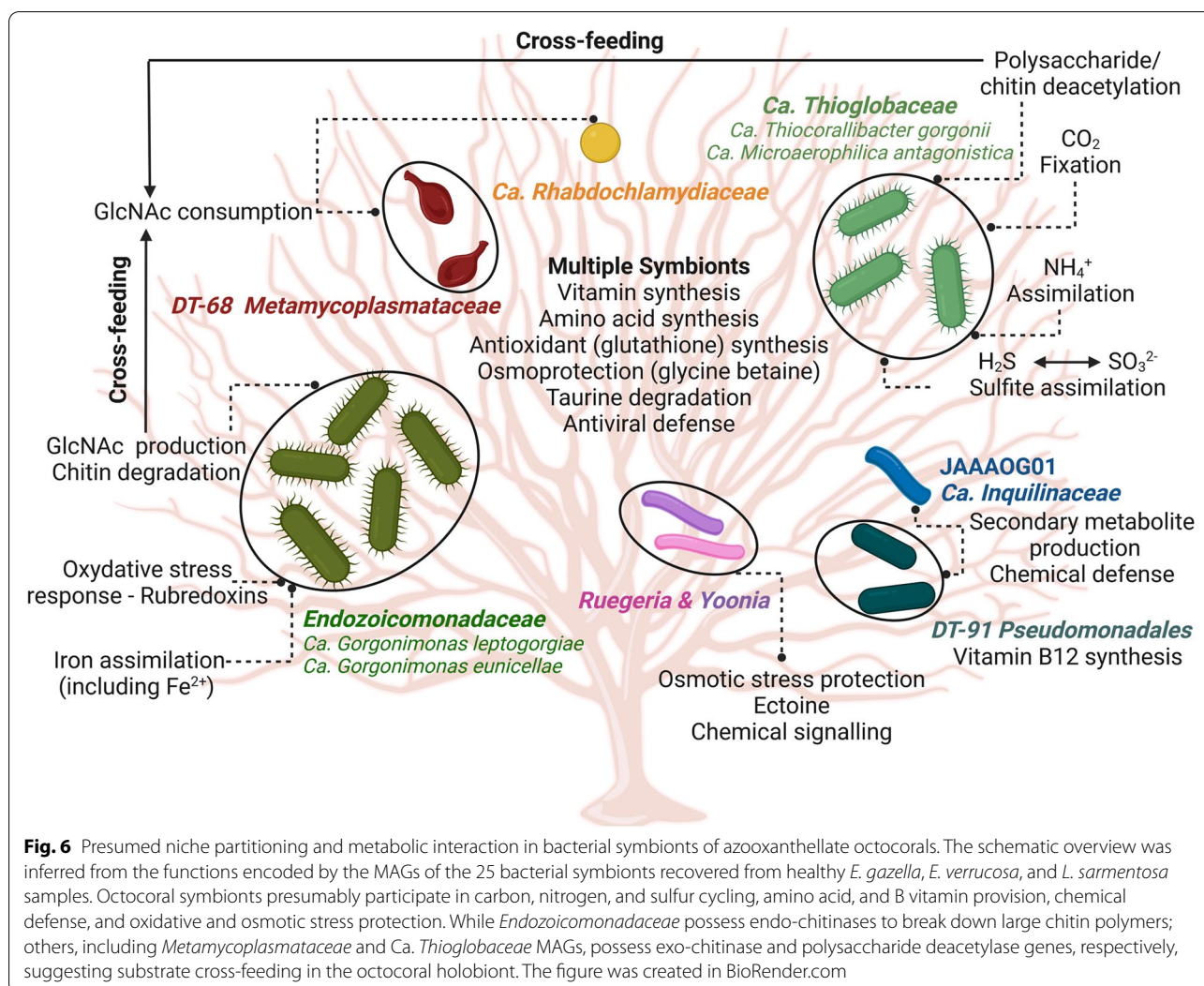
Alongside several *Alphaproteobacteria* MAGs of the *Rhodobacteraceae* and *Rhizobiaceae* families, we recovered an *Aquimarina* sp. (*Flavobacteriaceae*, *Bacteroidetes*) MAG with high genome coverage, suggesting high abundance of this taxon in necrotic *E. gazella* tissue. The *Aquimarina* MAG not only harbored a variety of SM-BGCs typical for the genus [59, 60] but also a high number of eukaryotic-like proteins, particularly leucine-rich and tetratricopeptide repeats, likely facilitating its association with eukaryotic hosts. This MAG further harbored genes involved in chitin deacetylation and chitin oligomer utilization as well as several transposase encoding genes (Fig. 4), suggesting an ability of quick genome rearrangement to adapt to different environmental settings. However, the *Aquimarina* MAG distinguished itself from the dominant symbionts



obtained from healthy octocoral tissue by the presence of a virulence factor (COG2996) and nitric oxide reductase and arginase genes, indicating this species is equipped to attack the host's immune system and to display opportunistic-to-pathogenic behavior, as discussed elsewhere [15, 60]. Underpinning their free-living lifestyle, the *Rhodobacteraceae* MAGs recovered from seawater are distinguished by the absence of serine-threonine protein kinases and serine protease inhibitors, and lack of the CRISPR/CAS system which we found enriched in many of the *Gammaproteobacteria* symbionts recovered from healthy octocorals. Likewise, the archaeal (*Ca. Poseidoniaceae*) MAG from surrounding seawater also lacked the aforementioned genes and other typical symbiosis features that would indicate host preference (Fig. 4), consistent with the notion of *Ca. Poseidoniaceae* spp. as primarily planktonic archaea [61]. It harbored, however, genes involved in the biosynthesis of four (i.e., B2, B3, B7, B9) of eight B vitamins and of some, mostly “nonessential” amino acids, e.g., alanine, serine, and glycine.

### B-vitamin and amino acid biosynthesis capacities in octocoral symbionts

All *Gammaproteobacteria* species obtained from healthy octocoral tissue possessed genes for the synthesis of at least six of eight B vitamins. For example, the *Endozoicomonadaceae* MAGs harbored genes for vitamin B1 (thiamine), B2 (riboflavin), B3 (niacin/NAD), B6 (pyridoxine), B7 (biotin), and B9 (folate) biosynthesis but lacked genes for B5 (pantothenic acid/CoA) and B12 (cobalamin) synthesis. The *Pseudomonadales* MAGs from healthy octocoral, however, possessed genes for the synthesis of all eight B vitamins (Table S7, Additional file 2). On the contrary, the *Metamycoplasmataceae* and *Rhabdochlamydiaceae* symbionts completely lacked B vitamin biosynthesis capacities, indicating genome streamlining congruent with their very small genome sizes. Moreover, both taxa possessed very limited coding potential for amino acid biosynthesis, harboring genes only for the synthesis of two and five of the 20 most common proteinogenic amino acids, respectively. The *Gammaproteobacteria* species from healthy octocorals, however,



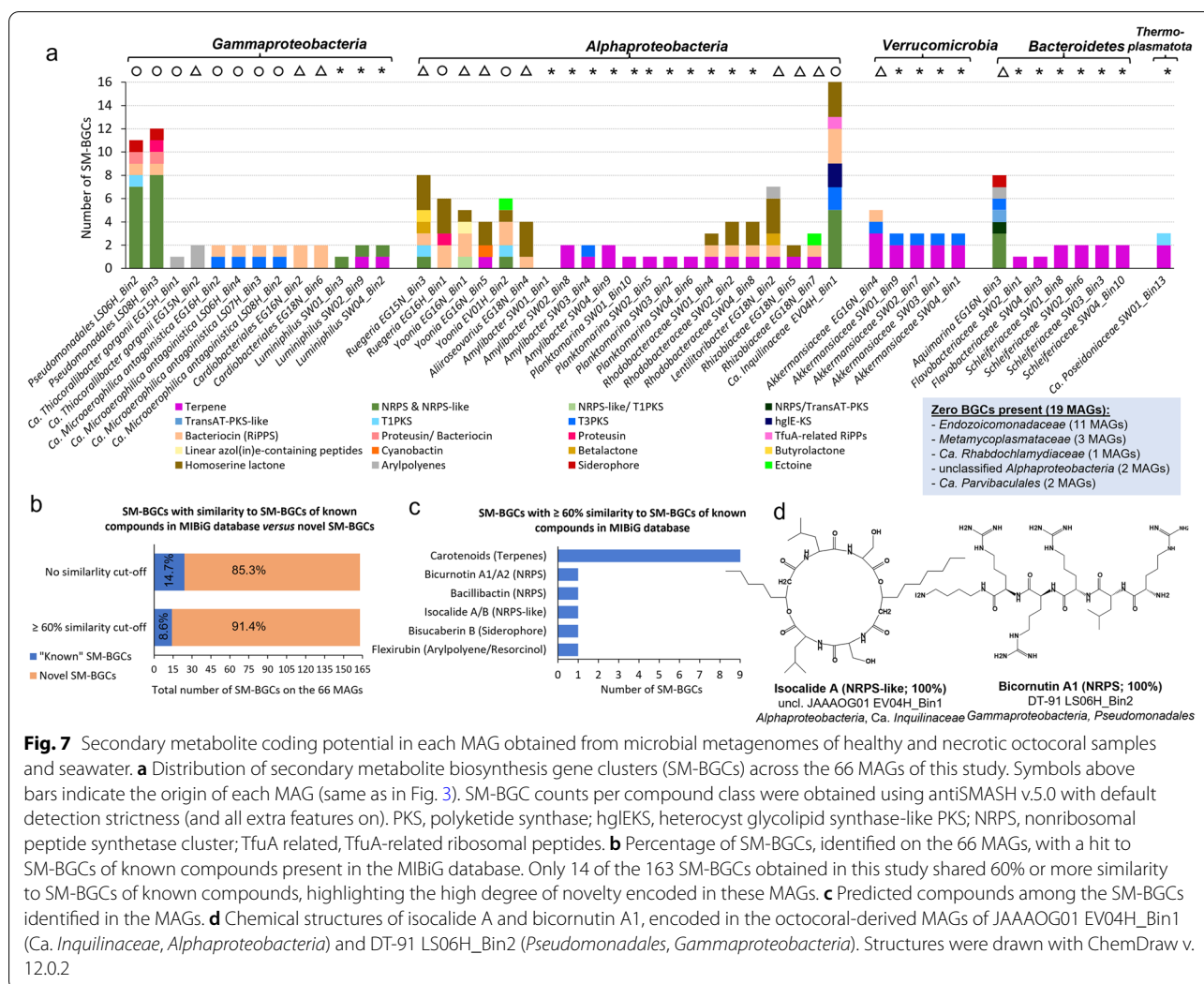
collectively harbored genes for the synthesis of 19 out of 20 proteinogenic amino acids, among them all nine “essential” amino acids, including the energetically costly methionine, histidine, and tryptophan (Table S7, Additional file 2). While the *Gammaproteobacteria* symbionts lacked the capacity to synthesize arginine, we identified arginine biosynthesis genes in an unclassified *Alphaproteobacteria* symbiont retrieved from healthy octocoral tissue. The *Alphaproteobacteria* species found in healthy octocorals were in general also well equipped to synthesize most B vitamins, including B12, and “essential” amino acids.

#### Secondary metabolite biosynthetic potential of octocoral symbionts

Genome mining with antiSMASH revealed that 46 of 66 MAGs analyzed in this study harbored between one and 16 (*Ca. Inquilinaceae* EV04\_Bin1) SM-BGCs, while 20 MAGs

lacked SM-BGCs, among them all *Endozoicomonadaceae*, *Metamycoplasmataceae*, *Ca. Rhabdochlamydiaceae*, and unclassified *Alphaproteobacteria* MAGs from octocorals (Fig. 7a; Table S8, Additional file 2). Notably, the two *Ca. Thioglobaceae* species presented distinct secondary metabolite coding potential. While *Ca. Thiocorallibacter gorgonii* MAGs harbored one or two arylpolyene cluster(s) (which may function as antioxidants), *Ca. Microaerophilica antagonistica* MAGs harbored one T3PKS and one bacteriocin/RiPP SM-BGC each, which may indicate antagonistic potential. The two *Pseudomonadales* MAGs from healthy *L. sarmentosa* samples showed rich SM-BGC profiles, with 7–8 NRPS and 2–3 bacteriocin/RiPP/protetoxin clusters, plus a siderophore SM-BGC (Figs. 6 and 7a).

Only 24 out of 163 SM-BGCs detected across all MAGs showed some homology with SM-BGCs encoding known compounds present in the MIBiG database, with 14 SM-BGCs sharing a similarity of 60–100% (Fig. 7b; Table S9,



Additional file 2). One of the NRPS clusters of *Pseudomonadales* MAG LS06H\_Bin2 showed 100% similarity to the NRPS cluster of the antimicrobial peptide bicornutin A [62] (Fig. 7c and d). Another NRPS cluster with 100% similarity to a SM-BGC of a known compound, namely the antibiotic and cyclic depsipeptide isocalide A [63], was identified on the *Ca. Inquilinaceae* MAG from healthy *E. verrucosa*, which was also the MAG with the richest SM-BGC profile across the entire dataset.

### Discussion

This is the first functional genomics study of uncultivated symbionts of octocorals. Notably, MAGs of dominant *Endozoicomonadaceae* symbionts were retrieved from all octocoral species and from nine out of 10 healthy coral specimens analyzed, strengthening the current understanding of these symbionts as indicators of coral health. Extensive amino acid and B vitamin biosynthesis capacities were common traits of

the dominant *Gammaproteobacteria* fraction (*Endozoicomonadaceae*, *Ca. Thioglobaceae*, and *Pseudomonadales* MAGs) characteristic of healthy octocorals (summarized in Fig. 6). This outcome corroborates recent research on scleractinian coral symbionts, which suggested amino acid and vitamin biosynthesis among the key traits for holobiont functioning [23, 26]. Our metagenomic data suggest that cross-feeding and niche partitioning, particularly of chitin and its breakdown products, B vitamins, and amino acids, may be mechanisms to promote coexistence and energy conservation among symbionts of azooxanthellate octocorals, which deserves further investigation and experimental validation in the future. Niche partitioning is proposed by sharp differences in secondary metabolite biosynthesis potential between symbiotic lineages, whereby *Endozoicomonadaceae*, *Metamycoplasmataceae* and *Chlamydiales* symbionts did not possess a single SM-BGC, indicating they might not

engage in chemical defense. Instead, *Pseudomonadales* and specific *Alphaproteobacteria* symbionts of healthy octocorals are likely to fulfill this task. They possessed the richest secondary metabolism across all 66 MAGs examined, including the genomic blueprint for the biosynthesis of antibiotics such as isocalide A [63]. Below, we highlight key genomic features and putative mechanisms of interaction among octocoral symbionts, focusing on the *Endozoicomonadaceae-Thioglobaceae-Metamycoplasmataceae* triad.

#### Discovery of novel *Endozoicomonadaceae* symbionts highly adapted to life in octocorals

This study uncovered two species of a novel *Endozoicomonadaceae* genus so far unique to temperate octocorals, here denoted Ca. *Gorgonimonas eunicellae* and Ca. *Gorgonimonas leptogorgiae*. These *Gorgonimonas* MAGs showed many traits that facilitate colonization of and persistence inside the host animal, including high abundances of ankyrin repeats. Studies on marine sponges showed that bacteria expressing ankyrin genes avoid phagocytosis by sponge amoebocytes, thus residing inside the sponge by evading its immune system [64]. A similar mechanism likely exists in coral-associated bacteria. Some bacteriophages also carry genes for ankyrin biosynthesis and possibly even transfer this information across different microbial symbionts [65]. Our earlier research demonstrated that ankyrin repeats are enriched in the microbial metagenomes of healthy octocorals [15], and this study now attributes this trait to specific *Endozoicomonadaceae* symbionts. A high copy number of ankyrin repeats was recently found among some cultured *Endozoicomonas* strains [21, 26]. Moreover, *Endozoicomonas marisrubri* 6c was shown to upregulate ankyrin repeat expression upon exposure to scleractinian coral extract [26]. This study is the first to show the enrichment of this trait in dominant, deeply branching, and uncultivated *Endozoicomonadaceae* lineages within the octocoral microbiome, further underpinning the importance of ankyrin repeats in coral symbioses. *Gorgonimonas* MAGs also possessed a high abundance of genes related to the type 3 secretion system, serine/threonine protein kinases, serine protease inhibitors, and the protein YejM. These features enable colonization and interference with host defense mechanisms among bacterial pathogens [66–68]. There is, however, growing evidence that beneficial bacteria use similar mechanisms to suppress immune responses and facilitate host colonization [69]. Indeed, parasitic and mutualistic symbioses often share similar evolutionary histories. Thus, host-microbe interactions may rather represent a flexible gradient, known as the “parasite-mutualist continuum,” than a static binary system, and

can shift and evolve depending on ecological context [70]. To accurately discern the beneficial-to-parasitic behavior of coral symbionts in a context-dependent manner, future studies need to combine multi-omics analyses with controlled in vivo mesocosm experiments.

All *Gorgonimonas* MAGs showed great capacities to produce tfp pili, bacterial surface appendages required for adherence to solid surfaces or cells, including host cells [71]. Tfp pili are also involved in bacterial aggregation (agglutination) and flagellum-independent movement on solid surfaces via gliding or twitching motility, coupled with chemotaxis [71]. CARD-FISH analysis demonstrated that *Endozoicomonas* bacteria indeed aggregate in dense, cyst-like structures inside the coral endoderm [72]. The here identified tfp pili likely provide the mechanisms for *Endozoicomonadaceae* symbionts to form such aggregations and attach to and move inside their host.

#### Ca. *Thioglobaceae* genomes code for a sophisticated antiviral defense system and chemoautotrophic metabolism

The Ca. *Thioglobaceae* MAGs were significantly enriched in CRISPR/Cas protein-associated genes, also found in all *Metamycoplasmataceae* (DT-68) and many *Endozoicomonadaceae* MAGs. These proteins are prokaryotic defense mechanisms against bacterial viruses and foreign DNA known to be enriched in the microbiomes of marine sponges [65, 73] and octocorals [15]. A recent study of the *Thioglobaceae* pan-genome showed that CRISPR/Cas systems are relatively enriched in mussel, sponge, and scleractinian coral symbionts compared to free-living *Thioglobaceae* lineages [74]. The Ca. *Thioglobaceae* and *Endozoicomonadaceae* MAGs of this study exhibited intraspecific variability in CRISPR/Cas gene numbers, leading to a flexible pan-immune system within host-associated populations, as suggested elsewhere [75].

We found one Ca. *Thioglobaceae* MAG to harbor RuBisCo and other genes for light-independent carbon fixation and chemoautotrophic lifestyle — as known for some *Thioglobaceae* symbionts from deep-sea sponges and hydrothermal vent bivalves [76]. It could represent an alternative carbon fixation route in azooxanthellate octocorals that lack photosynthetic Symbiodinaceae and other algae symbionts. It also indicates that the ecological range of chemoautotrophic, symbiotic *Thioglobaceae* extends beyond hydrothermal vent and deep-sea ecosystems well into mesophotic zones and above.

Beyond their putative role in carbon supply, most Ca. *Thioglobaceae* MAGs possessed the genetic repertoire for ammonium assimilation, sulfur cycling, and taurine processing within the octocoral holobiont.



Taurine is a sulfur-containing amino acid widely found in animal tissue, and several studies suggested a role of sponge and coral symbionts in sulfur cycling via taurine metabolism [21, 23, 73, 77]. However, as observed in this study, taurine utilization is a widespread trait among marine host-associated and free-living, seawater-derived *Alpha*- and *Gammaproteobacteria* and not restricted to obligate symbiotic clades.

#### ***Endozoicomonadaceae*, *Ca. Thioglobaceae*, and *Metamycoplasmataceae* symbionts may adopt a facultative anaerobic lifestyle**

We found genome-based evidence for *Endozoicomonadaceae*, *Ca. Thioglobaceae*, and *Metamycoplasmataceae* symbionts to possess a facultative anaerobic metabolism and the ability to thrive in low oxygen conditions. *Endozoicomonadaceae* symbionts may generate energy without oxygen via the pyruvate metabolism and with the help of acetate kinases and pyruvate formate lyases [78, 79]. All *Ca. Thioglobaceae* and most *Endozoicomonadaceae* MAGs also harbored Na<sup>+</sup>-translocating ferredoxin:NAD<sup>+</sup> oxidoreductase (Rnf complex) encoding genes. The Rnf complex is involved in anaerobic respiration and alternative routes of energy generation by translocating sodium across the cell membrane, a common feature of anaerobic (acetate forming) acetogens (e.g., *Acetobacterium woodii*) [80]. Recent studies demonstrated that the Rnf complex is not limited to acetogens and instead can be found in several Gram-negative, facultative anaerobic bacteria [81–83]. The *Metamycoplasmataceae* (DT-68) and some *Ca. Thioglobaceae* MAGs further possessed lactate dehydrogenase encoding genes to ferment pyruvate into lactate and thus recycle NADH + H<sup>+</sup>, for glycolysis to continue under anaerobic conditions. Facultative anaerobe symbionts and anaerobic metabolism seem to play a crucial role in nutrient cycling in marine sponges [84]. Few cultured *Endozoicomonas* strains from octocorals are classified as facultative anaerobes, such as type strains *Endozoicomonas euni-ceicola* and *E. gorgoniicola* [22], indicating that octocoral holobionts might benefit from the facultative anaerobic lifestyles of their symbionts. The consistent presence of genes coding for ferrous iron (Fe<sup>2+</sup>) uptake systems on all *Endozoicomonadaceae* MAGs provides further support for this hypothesis, since Fe<sup>2+</sup> is usually available in anaerobic environments while ferric iron (Fe<sup>3+</sup>) is dominant under oxygenated conditions. Organisms frequently exposed to oxygen limitation in their natural habitat rely on ferrous iron uptake mechanisms [85]. The 11 *Endozoicomonadaceae* MAGs were also enriched in ferritin encoding genes for iron storage, likely supporting growth of these symbionts during iron starvation [86], a situation

they may frequently encounter in the host animal. Iron is indeed a limiting micronutrient in some marine environments and was shown to restrain primary productivity in corals [87, 88]. Together, such metabolic capacities could enable these symbionts not to drain their obligate aerobic host from oxygen but to provide it with iron whenever environmental oxygen is limited. Another distinguishing feature of the *Endozoicomonadaceae* symbionts of this study was the presence of rubredoxins, a class of iron-containing proteins found in several anaerobic and sulfate-reducing bacteria. Rubredoxins act as electron carriers in many biochemical processes and can play a crucial role in reducing reactive oxygen species [89–91]. Both effective oxidative stress response and iron-sequestration mechanisms were recently proposed as key beneficial traits of coral probiotics [92].

#### **Chitin breakdown and cross-feeding possibilities in the octocoral holobiont revealed**

The distribution of endo-chitinase (*Endozoicomonadaceae*), exo-chitinase (*Metamycoplasmataceae* and several other symbiont taxa), and chitin deacetylase (*Ca. Thioglobaceae* and *Alphaproteobacteria* spp.) encoding genes across octocoral symbionts points towards their likely role in chitin processing and carbon and nitrogen turnover in their host. This indicates possible mechanisms of substrate cross-feeding [93] between symbiotic partners. Cross-feeding cascades among microbes may promote the evolution of small genomes as here observed, for example, in *Metamycoplasmataceae* (0.63 Mb) and *Ca. Rhabdochlamydiaceae* (1 Mb) symbionts. These symbionts may have evolved as auxotrophs, relying on the acquisition of certain nutrients and vitamins from other microbial partners of the coral holobiont. This may also reduce competition and enable coexistence of multiple host-associated taxa through functional specialization. Exo-chitinase (EC 3.2.1.52) activity was measured earlier in crude extracts of the octocoral *Gorgonia ventalina* [94]. Moreover, chitinolytic activity was reported in seven scleractinian coral species, and chitinase-like genes were identified in the genome of *Acropora digitifera* [95]. Raimundo and colleagues reported that the abundance of endo-chitinase and chitin-binding-protein encoding genes in healthy octocoral tissue levels up with those from surrounding environments [96]. Taken together, these observations suggest that chitinases are widely distributed among corals, many of which feed on chitin-rich phyto- and zooplankton [6]. These enzymes may also play a role in the animals' defense against fungal infections [7], whereby chitinases from bacterial symbionts may further strengthen the coral immune system.

## Conclusion

This study examines the first microbial MAGs ever retrieved from octocorals, revealing that uncultured *Endozoicomonadaceae*, *Ca. Thioglobaceae*, and *Mycoplasmoidales* symbionts are seemingly well adapted to life in low oxygen conditions. Moreover, we identified a thus-far unanticipated, global role for *Endozoicomonadaceae* symbionts of corals in chitin processing and C and N cycling across benthic ecosystems. Other symbionts such as *Ca. Thioglobaceae* were well equipped to control bacteriophage attacks and to adopt a chemoautotrophic lifestyle. We conclude that the prokaryotic symbiome of octocorals participates in carbon (particularly chitin), nitrogen and sulfur cycling, amino acid and B vitamin provision, chemical defense, and oxidative and osmotic stress protection. Traits are not shared equally among all members of the symbiotic community; instead, niche partitioning and metabolic specialization among taxonomically unique symbionts may contribute to efficient functioning of and co-existence in the healthy octocoral holobiont.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-022-01343-7>.

**Additional file 1: Figure S1.** Phylogenomic analysis of the *Ca. Thioglobaceae* family using SpeciesTreeBuilder v.01.0. **Figure S2.** COG functions ( $N=76$ ) significantly enriched ( $q$ -value  $< 0.05$ ) in the octocoral-derived *Ca. Gorgonimonas* (*Endozoicomonadaceae*) MAGs ( $N= 11$ , green), compared with all other MAGs ( $N=55$ ) of this dataset. **Figure S3.** KEGG metabolic pathway map of the *Ca. Gorgonimonas* (*Endozoicomonadaceae*) MAGs of this study. **Figure S4.** Phylogeny of 101 full-length, bacterial endo-chitinase (EC 3.2.1.14) encoding genes, including the 11 endo-chitinase (GH18-family) genes from *Endozoicomonadaceae* MAGs (grey shadings) of this study. **Figure S5.** COG functions ( $N=34$ ) significantly enriched ( $p$ -value  $< 0.05$ ) in the *Ca. Thioglobaceae* MAGs ( $N= 6$ , dark green), compared with all other MAGs ( $N=60$ ) of this dataset. An one-sided Welch's t-test for unequal variances was performed in STAMP v.2.1.3. Multiple test correction was performed using the Benjamini-Hochberg correction (FDR). To further limit the displayed number of significant entries, an effect size filter was also applied, setting the "ratio of proportions" to 5.00. **Figure S6.** KEGG metabolic pathway map of the *Ca. Thioglobaceae* MAGs of this study. The metabolic map features glycolysis, pyruvate, nitrogen, sulfur, and taurine metabolism. EC numbers of enzymes catalyzing the reactions are given in rectangular boxes. EC numbers highlighted in green represent enzymes encoded on *Ca. Thioglobaceae* MAGs. Beige boxes indicate connections to other metabolic pathways active in these MAGs. **Figure S7.** KEGG carbon fixation map of *Ca. Thiocorallibacter gorgonii* MAG EG15H\_Bin1 (*Ca. Thioglobaceae*). EC numbers of enzymes catalyzing the reactions are given in rectangular boxes. EC numbers highlighted in green represent enzymes encoded on this MAG. Beige boxes indicate connections to other metabolic pathways active in these MAGs. Orange dots highlight the substrates (Ribulose-1,5-bisphosphate,  $\text{CO}_2$ , and  $\text{H}_2\text{O}$ ) and product (2 Glycerate-3-phosphate,  $2 \text{H}^+$ ) of Rubisco (EC 4.1.1.39).

**Additional file 2: Table S1.** General genomic features, taxonomic classification and genome assembly accession numbers of each of the 66 metagenome-assembled genomes (MAGs) obtained from octocoral- and seawater-derived microbial metagenomes. **Table S2.** Overview of the number and quality of MAGs obtained from octocoral- and seawater-derived microbial metagenome samples. **Table S3.** Clusters

of Orthologous Groups of proteins (COGs) annotation of the 66 MAGs analysed in this study. **Table S4.** Results of the SIMPER test performed on COG profiles (Hellinger-transformed abundances; Euclidean distances) of the 66 MAGs grouped at order level. **Table S5A.** Absolute abundances (counts) of COG functions shown in Figure 4 that distinguished the 66 MAGs the most (based on SIMPER and Welch's tests). **Table S5B.** Relative abundances of COG functions shown in Figure 4 that distinguished the 66 MAGs the most (based on SIMPER and Welch's tests). **Table S6A.** Features of the endo-chitinase (EC 3.2.1.14) genes found on the 11 *Endozoicomonadaceae* MAGs of this study. **Table S6B.** Genes involved in chitin degradation present on the 11 *Endozoicomonadaceae* MAGs of this study and other, publicly available *Endozoicomonadaceae* genomes. **Table S7.** Amino acid ( $n= 20$ ) and B vitamin ( $n= 8$ ) biosynthesis capacities (based on genomic evidence) of the 11 bacterial species recovered from the microbiomes of healthy octocoral tissue. **Table S8.** Secondary metabolite biosynthetic gene clusters (SM-BGCs) present on the 66 MAGs, annotated using antiSMASH bacterial version 5.0. **Table S9.** List of SM-BGCs with some level of homology to MIBiG database entries.

## Acknowledgements

Not applicable

## Authors' contributions

TKC and RC designed the study. RC, TKC, ALL, JMSG, NK, and UNR provided resources and materials. TKC, RT, and SGS processed data. TKC and LK analyzed data and prepared figures. TKC wrote the first manuscript draft. RC revised the first manuscript draft. All authors revised and improved the manuscript prior to submission. The authors read and approved the final manuscript.

## Funding

This work was supported by the Portuguese Foundation for Science and Technology (FCT), through the research projects EXPL/MAR-EST/1664/2013 and EXPL/BIA-MIC/0286/2021, and by the Fundo Azul program of Direção-Geral de Política do Mar (DPGM; Ministry of the Sea, Portugal) through grant FA\_05\_2017\_032. We further acknowledge national funds from the FCT, in the scope of the projects UIDB/04565/2020 and UIDP/04565/2020 of iBB and the project LA/P/0140/2020 of i4HB. Computational support was received from INCD funded by FCT and FEDER under the project 01/SAICT/2016 n° 022153. TKC is the recipient of an investigator contract (CEECIND/00788/2017) and SGS of a PhD scholarship (PD/BD/143029/2018) conceded by FCT. LK acknowledges a scholarship from the Swiss-European Mobility Programme (SEMP). UNR and RT were financed by the Helmholtz Association (VH-NG-1248 Micro 'Big Data'). The work conducted by the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under contract no. DE-AC02-05CH11231.

## Availability of data and materials

The 66 MAGs are available in the European Nucleotide Archive (ENA) under the study accession number PRJEB50578 and sample accession numbers ERS10420767 — ERS10420805 (from octocorals) and ERS10422230 — ERS10422256 (from seawater). Assembly accession numbers for each MAG can be found in Table S1 of Additional file 2. All raw metagenome data are deposited under the study accession number PRJEB13222.

## Declarations

### Ethics approval and consent to participate

This article does neither contain any studies with human participants nor vertebrate animals or cephalopods performed by any of the authors. This study was exempt from ethical approval procedures according to the current Portuguese legislation.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interests.

**Author details**

<sup>1</sup>Institute for Bioengineering and Biosciences, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal. <sup>2</sup>Associate Laboratory i4HB - Institute for Health and Bioeconomy, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal. <sup>3</sup>École Polytechnique Fédérale de Lausanne, Écublens, Switzerland. <sup>4</sup>Helmholtz Centre for Environmental Research, Leipzig, Germany. <sup>5</sup>Centro de Ciências Do Mar, Universidade Do Algarve, Faro, Portugal. <sup>6</sup>Centro de Investigación Científica Y de Educación Superior de Ensenada, Ensenada, Mexico. <sup>7</sup>Department of Energy, Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.

Received: 6 May 2022 Accepted: 3 August 2022

Published online: 22 September 2022

**References**

- Curdia J, Monteiro P, Afonso CML, Santos MN, Cunha MR, Goncalves JMS. Spatial and depth-associated distribution patterns of shallow gorgonians in the Algarve coast (Portugal, NE Atlantic). *Helgol Mar Res*. 2013;67(3):521–34.
- McFadden CS, Sánchez JA, France SC. Molecular phylogenetic insights into the evolution of Octocorallia: a review. *Integr Comp Biol*. 2010;50(3):389–410.
- Fabricius KE, Alderslade P. Soft corals and sea fans: a comprehensive guide to the tropical shallow water genera of the central-west Pacific, the Indian Ocean and the Red Sea. Townsville: Australian Institute of Marine Science; 2001.
- Williams GC, Cairns SD. Systematic list of valid octocoral genera: <http://researcharchive.calacademy.org/research/izg/OCTOCLASS.htm>. 2019, 2019(20.09.2019).
- Pérez CD, Neves B, Cordeiro RTS, Williams GC, Cairns SD. Diversity and Distribution of Octocorallia. In: The Cnidaria, past, present and future: the world of Medusa and her sisters. Edited by Goffredo S, Dubinsky Z: Springer International Publishing; 2016. p. 109–123.
- Gili JM, Coma R. Benthic suspension feeders: their paramount role in littoral marine food webs. *Trends Ecol Evol*. 1998;13(8):316–21.
- van de Water JAJM, Allemand D, Ferrier-Pagès C. Host-microbe interactions in octocoral holobionts - recent advances and perspectives. *Microbiome*. 2018;6:64.
- Cerrano C, Bavestrello G, Bianchi CN, Cattaneo-vietti R, Bava S, Morganti C, Morri C, Picco P, Sara G, Schiaparelli S, et al. A catastrophic mass-mortality episode of gorgonians and other organisms in the Ligurian Sea (northwestern Mediterranean), summer 1999. *Ecol Lett*. 2002;3(4):284–93.
- Ponti M, Perlini RA, Ventra V, Grech D, Abbiati M, Cerrano C. Ecological shifts in Mediterranean coralligenous assemblages related to gorgonian forest loss. *PLoS ONE*. 2014;9(7):e102782.
- Hall-Spencer JM, Pike J, Munn CB. Diseases affect cold-water corals too: *Eunicella verrucosa* (Cnidaria : Gorgonacea) necrosis in SW England. *Dis Aquat Org*. 2007;76(2):87–97.
- Vezzulli L, Previati M, Pruzzo C, Marchese A, Bourne DG, Cerrano C. VibrioSea-Consortium: *Vibrio* infections triggering mass mortality events in a warming Mediterranean Sea. *Environ Microbiol*. 2010;12(7):2007–19.
- Turicchia E, Abbiati M, Sweet M, Ponti M. Mass mortality hits gorgonian forests at Montecristo Island. *Dis Aquat Organ*. 2018;131:79–85.
- La Riviere M, Garrabou J, Bally M. Evidence for host specificity among dominant bacterial symbionts in temperate gorgonian corals. *Coral Reefs*. 2015;34(4):1087–98.
- La Riviere M, Roumagnac M, Garrabou J, Bally M. Transient shifts in bacterial communities associated with the temperate gorgonian *Paramuricea clavata* in the northwestern Mediterranean Sea. *Plos One*. 2013;8(2):e57385.
- Keller-Costa T, Lago-Leston A, Saraiva JP, Toscan R, Silva SG, Gonçalves J, Cox CJ, Kyrpidis NC, Nunes da Rocha U, Costa R. Metagenomic insights into the taxonomy, function and dysbiosis of prokaryotic communities in octocorals. *Microbiome*. 2021;9:72.
- van de Water JAJM, Voolstra CR, Rottier C, Cocito S, Peirano A, Allemand D, Ferrier-Pagès C. Seasonal stability in the microbiomes of temperate gorgonians and the red coral *Corallium rubrum* across the Mediterranean Sea. *Microbial Ecology*. 2018;75(1):274–88.
- Haydon TD, Suggett DJ, Siboni N, Kahlke T, Camp EF, Seymour JR. Temporal variation in the microbiome of tropical and temperate octocorals. *Microbial Ecology*. 2022;83(4):1073–87.
- van de Water JAJM, Melkonian R, Junca H, Voolstra CR, Reynaud S, Allemand D, Ferrier-Pagès C. *Spirochaetes* dominate the microbial community associated with the red coral *Corallium rubrum* on a broad geographic scale. *Sci Rep*. 2016;6:27277.
- Keller-Costa T, Eriksson D, Gonçalves JMS, Gomes NCM, Lago-Leston A, Costa R. The gorgonian coral *Eunicella labiata* hosts a distinct prokaryotic consortium amenable to cultivation. *FEMS Microbiol Ecol*. 2017;93(12):1–19.
- Bourne DG, Morrow KM, Webster NS. Insights into the coral microbiome: underpinning the health and resilience of reef ecosystems. *Annu Rev Microbiol*. 2016;70:317–40.
- Sweet M, Vilella H, Keller-Costa T, Costa R, Romano S, Bourne DG, Cardenas A, Huggett M, Kerwin A, Kuek F, et al. Insights into the cultured bacterial fraction of corals. *mSystems*. 2021;6(3):e0124920.
- Pike RE, Haltli B, Kerr RG. Description of *Endozoicomonas euniceicola* sp nov and *Endozoicomonas gorgoniicola* sp nov., bacteria isolated from the octocorals *Eunicea fusca* and *Plexaura* sp., and an emended description of the genus *Endozoicomonas*. *Int J Syst Evol Microbiol*. 2013;63:4294–302.
- Robbins SJ, Singleton CM, Chan CX, Messer LF, Geers AU, Ying H, Baker A, Bell SC, Morrow KM, Ragan MA, et al. A genomic view of the reef-building coral *Porites lutea* and its microbial symbionts. *Nat Microbiol*. 2019;4(12):2090–100.
- Neave MJ, Michell CT, Apprill A, Voolstra CR. *Endozoicomonas* genomes reveal functional adaptation and plasticity in bacterial strains symbiotically associated with diverse marine hosts. *Sci Rep*. 2017;7:40579.
- Tandon K, Lu C-Y, Chiang P-W, Wada N, Yang S-H, Chan Y-F, Chen P-Y, Chang H-Y, Chiou Y-J, Chou M-S, et al. Comparative genomics: dominant coral-bacterium *Endozoicomonas acroporae* metabolizes dimethylsulfoxide (DMSO). *ISME J*. 2020;14(5):1290–303.
- Pogoreutz C, Oakley CA, Räderer N, Cárdenas A, Perna G, Xiang N, Peng L, Davy SK, Ngugi DK, Voolstra CR. Coral holobiont cues prime *Endozoicomonas* for a symbiotic lifestyle. *ISME J*. 2022;16(8):1883–95.
- Raimundo I, Silva SG, Costa R, Keller-Costa T. Bioactive secondary metabolites from octocoral-associated microbes—new chances for blue growth. *Mar Drugs*. 2018;16(12):485.
- Keller-Costa T, Silva R, Lago-Lestón A, Costa R. Genomic insights into *Aquimarina* sp. EL33, a bacterial symbiont of the gorgonian coral *Eunicella labiata*. *Genome Announc*. 2016;4(4):e00855-00816.
- Couceiro JF, Keller-Costa T, Marques M, Woyke T, Kyrpidis NC, Whitman WB, Costa R. The *Roseibium album* (*Labrenzia alba*) genome possesses multiple symbiosis factors possibly underlying host-microbe relationships in the marine benthos. *Microbiol Resour Announc*. 2021;10(34):e0032021.
- Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*. 2018;6(1):158.
- Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*. 2017;27:824–34.
- West PT, Probst AJ, Grigoriev IV, Thomas BC, Banfield JF. Genome-reconstruction for eukaryotes from complex natural microbial communities. *Genome Res*. 2018;28(4):569–80.
- Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7:e7359.
- Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 2016;32(4):605–7.
- Alneberg J, Bjarnason BS, de Bruijn I, Schirmer M, Quick J, Ijaz UZ, Lahti L, Loman NJ, Andersson AF, Quince C. Binning metagenomic contigs by coverage and composition. *Nat Methods*. 2014;11(11):1144–6.
- Song WZ, Thomas T. Binning\_refiner: improving genome bins through the combination of different binning programs. *Bioinformatics*. 2017;33(12):1873–5.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25(7):1043–55.

38. Parks DH, Rinke C, Chuvochina M, Chaumeil P-A, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol.* 2017;2(11):1533–42.
39. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, Schulz F, Jarett J, Rivers AR, Eloie-Fadrosch EA, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol.* 2017;35(8):725–31.
40. Rodriguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Konstantinidis KT. The Microbial Genomes Atlas (MiGA) webservice: taxonomic and gene diversity analysis of Archaea and bacteria at the whole genome level. *Nucleic Acids Res.* 2018;46(W1):W282–8.
41. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics.* 2019;36(6):1925–7.
42. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol.* 2020;70(11):5607–12.
43. Parte AC. LPSN—list of prokaryotic names with standing in nomenclature. *Nucleic Acids Res.* 2013;42(D1):D613–6.
44. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun.* 2018;9(1):5114.
45. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol.* 2007;57(Pt 1):81–91.
46. Konstantinidis KT, Tiedje JM. Genomic insights that advance the species definition for prokaryotes. *Proc Natl Acad Sci USA.* 2005;102(7):2567–72.
47. Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, et al. KBase: the United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol.* 2018;36(7):566–9.
48. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30(14):2068–9.
49. Price MN, Dehal PS, Arkin AP. FastTree 2 - approximately maximum-likelihood trees for large alignments. *PLoS ONE.* 2010;5(3):e9490–e9490.
50. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47(W1):W256–9.
51. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, et al. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 2014;42:D206–14.
52. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics.* 2008;9:75.
53. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research.* 2019;47(W1):W81–7.
54. Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hoof JJJ, van Santen JA, Tracanna V, Suarez Duran HG, Pascal Andreu V, et al. MIBiG 2.0: a repository for biosynthetic gene clusters of known function. *Nucleic Acids Res.* 2020;48(D1):D454–D458.
55. Medema MH, Kottmann R, Yilmaz P, Cummings M, Biggins JB, Blin K, de Bruijn I, Chooi YH, Claesen J, Coates RC, et al. Minimum Information about a Biosynthetic Gene cluster. *Nat Chem Biol.* 2015;11:625.
56. Hammer Ø, Harper DAT, Ryan PD. PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica.* 2001;4(1):9pp.
57. Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics.* 2014;30(21):3123–4.
58. Maire J, Blackall LL, van Oppen MJH. Intracellular bacterial symbionts in corals: Challenges and future directions. *Microorganisms.* 2021;9(11):2209.
59. Silva SG, Keller-Costa T, Bloom J, Costa R. Comparative genomics reveals complex natural product biosynthesis potential and carbon metabolism across host-associated and free-living *Aquimarina* (*Bacteroidetes*) species. *Environ Microbiol.* 2019;21:4002–19.
60. Silva SG, Paula P, da Silva JP, Mil-Homens D, Teixeira MC, Fialho AM, Costa R, Keller-Costa T. Insights into the antimicrobial activities and metabolomes of *Aquimarina* (*Flavobacteriaceae*, *Bacteroidetes*) species from the rare marine biosphere. *Mar Drugs.* 2022;20(7):423.
61. Rinke C, Rubino F, Messer LF, Youssef N, Parks DH, Chuvochina M, Brown M, Jeffries T, Tyson GW, Seymour JR, et al. A phylogenomic and ecological analysis of the globally abundant Marine Group II archaea (*Ca. Poseidoniales* ord. nov.). *ISME Journal.* 2019;13(3):663–75.
62. Böszörményi E, Ersek T, Fodor A, Fodor AM, Földes LS, Hevesi M, Hogan JS, Katona Z, Klein MG, Kormány A, et al. Isolation and activity of *Xenorhabdus* antimicrobial compounds against the plant pathogens *Erwinia amylovora* and *Phytophthora nicotianae*. *J Appl Microbiol.* 2009;107(3):746–59.
63. Dose B, Niehs SP, Scherlach K, Flórez LV, Kaltenpoth M, Hertweck C. Unexpected bacterial origin of the antibiotic icosalide: two-tailed depeptide assembly in multifarious *Burkholderia* symbionts. *ACS Chem Biol.* 2018;13(9):2414–20.
64. Nguyen M, Liu M, Thomas S. Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis. *Mol Ecol.* 2014;23(6):1635–45.
65. Jahn MT, Arkhipova K, Markert SM, Stigloher C, Lachnit T, Pita L, Kupczok A, Ribes M, Stengel ST, Rosenstiel P, et al. A phage protein aids bacterial symbionts in Eukaryote immune evasion. *Cell Host Microbe.* 2019;26(4):542–550.e545.
66. Coburn B, Sekirov I, Finlay BB. Type III secretion systems and disease. *Clin Microbiol Rev.* 2007;20(4):535–49.
67. Canova MJ, Molle V. Bacterial serine/threonine protein kinases in host-pathogen interactions. *J Biol Chem.* 2014;289(14):9473–9.
68. Gabale U, Peña Palomino PA, Kim H, Chen W, Ressler S. The essential inner membrane protein YejM is a metalloenzyme. *Sci Rep.* 2020;10(1):17794.
69. Stringlis IA, Zamioudis C, Berendsen RL, Bakker PAHM, Pieterse CMJ. Type III secretion system of beneficial rhizobacteria *Pseudomonas simiae* WCS417 and *Pseudomonas defensor* WCS374. *Front Microbiol.* 2019;10:1631.
70. Drew GC, Stevens EJ, King KC. Microbial evolution and transitions along the parasite–mutualist continuum. *Nat Rev Microbiol.* 2021;19(10):623–38.
71. Shi W, Sun H. Type IV pilus-dependent motility and its possible role in bacterial pathogenesis. *Infect Immun.* 2002;70(1):1–4.
72. Neave MJ, Apprill A, Ferrier-Pages C, Voolstra CR. Diversity and function of prevalent symbiotic marine bacteria in the genus *Endozoicomonas*. *Appl Microbiol Biotechnol.* 2016;100(19):8315–24.
73. Karimi E, Ramos M, Gonçalves JMS, Xavier JR, Reis MP, Costa R. Comparative metagenomics reveals the distinctive adaptive features of the *Spongia officinalis* endosymbiotic consortium. *Front Microbiol.* 2017;8:2499. <https://doi.org/10.3389/fmicb.2017.02499>.
74. Ansoerge R, Romano S, Sayavedra L, Rubin-Blum M, Gruber-Vodicka H, Scilipoti S, Molari M, Dubilier N, Petersen J. The hidden pangenome: comparative genomics reveals pervasive diversity in symbiotic and free-living sulfur-oxidizing bacteria. *bioRxiv.* 2020;12.11.421487. <https://doi.org/10.1101/2020.12.11.421487>.
75. Bernheim A, Sorek R. The pan-immune system of bacteria: antiviral defence as a community resource. *Nat Rev Microbiol.* 2020;18(2):113–9.
76. Morris RM, Spietz RL. The physiology and biogeochemistry of SUP05. *Ann Rev Mar Sci.* 2022;14(1):261–75.
77. Engelberts JP, Robbins SJ, de Goeij JM, Aranda M, Bell SC, Webster NS. Characterization of a sponge microbiome using an integrative genome-centric approach. *ISME J.* 2020;14(5):1100–10.
78. Thauer RK, Jungermann K, Decker K. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol Rev.* 1977;41(1):100–80.
79. Crain AV, Broderick JB. Pyruvate formate-lyase and its activation by pyruvate formate-lyase activating enzyme. *J Biol Chem.* 2014;289(9):5723–9.
80. Biegel E, Müller V. Bacterial Na<sup>+</sup>-translocating ferredoxin:NAD<sup>+</sup>-oxidoreductase. *Proc Natl Acad Sci.* 2010;107(42):18138–42.
81. Kuhns M, Trifunović D, Huber H, Müller V. The Rnf complex is a Na<sup>+</sup>-coupled respiratory enzyme in a fermenting bacterium, *Thermotoga maritima*. *Commun Biol.* 2020;3(1):431.
82. Biegel E, Schmidt S, González JM, Müller V. Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. *Cell Mol Life Sci.* 2011;68(4):613–34.

83. Hreha TN, Mezić KG, Herce HD, Duffy EB, Bourges A, Pryshchep S, Juarez O, Barquera B. Complete topology of the RNF complex from *Vibrio cholerae*. *Biochemistry*. 2015;54(15):2443–55.
84. Schuster A, Strehlow BW, Eckford-Soper L, McAllen R, Canfield DE, Hallam SJ. Effects of seasonal anoxia on the microbial community structure in demosponges in a marine lake in Lough Hyne, Ireland. *mSphere*. 2021;6(1):e00991–e00920.
85. Lau CKY, Krewulak KD, Vogel HJ. Bacterial ferrous iron transport: the Feo system. *FEMS Microbiol Rev*. 2015;40(2):273–98.
86. Andrews SC. Iron storage in bacteria. *Adv Microb Physiol*. 1998;40:281–351.
87. Ferrier-Pagès C, Schoelzke V, Jaubert J, Muscatine L, Hoegh-Guldberg O. Response of a scleractinian coral, *Stylophora pistillata*, to iron and nitrate enrichment. *J Exp Mar Biol Ecol*. 2001;259(2):249–61.
88. Entsch B, Sim RG, Hatcher BG. Indications from photosynthetic components that iron is a limiting nutrient in primary producers on coral reefs. *Mar Biol*. 1983;73(1):17–30.
89. Chen L, Liu MY, LeGall J, Fareleira P, Santos H, Xavier AV. Rubredoxin oxidase, a new flavo-hemo-protein, is the site of oxygen reduction to water by the “strict anaerobe” *Desulfovibrio gigas*. *Biochem Biophys Res Commun*. 1993;193(1):100–5.
90. Gardner AM, Helmick RA, Gardner PR. Flavorubredoxin, an inducible catalyst for nitric oxide reduction and detoxification in *Escherichia coli*. *J Biol Chem*. 2002;277(10):8172–7.
91. LeGall J, Liu MY, Gomes CM, Braga V, Pacheco I, Regalla M, Xavier AV, Teixeira M. Characterisation of a new rubredoxin isolated from *Desulfovibrio desulfuricans* 27774: definition of a new family of rubredoxins. *FEBS Lett*. 1998;429(3):295–8.
92. Peixoto RS, Sweet M, Villela HDM, Cardoso P, Thomas T, Voolstra CR, Høj L, Bourne DG. Coral probiotics: Premise, promise, prospects. *Ann Rev Anim Biosci*. 2021;9(1):265–88.
93. Smith NW, Shorten PR, Altermann E, Roy NC, McNabb WC. The classification and evolution of bacterial cross-feeding. *Front Ecol Evol*. 2019;7:153.
94. Douglas NL, Mullen KM, Talmage SC, Harvell CD. Exploring the role of chitinolytic enzymes in the sea fan coral *Gorgonia ventalina*. *Marine Biol*. 2007;150(6):1137–44.
95. Yoshioka Y, Tanabe T, Iguchi A. The presence of genes encoding enzymes that digest carbohydrates in coral genomes and analysis of their activities. *PeerJ*. 2017;5:e4087.
96. Raimundo I, Silva R, Meunier L, Valente S, Keller-Costa T, Costa R. Functional metagenomics reveals differential chitin degradation and utilization features across free-living and host-associated marine microbiomes. *Microbiome*. 2021;9:43.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

