



# Genome Sequencing Suggests Diverse Secondary Metabolism in Coral-Associated *Aquimarina megaterium*

Joana Fernandes Couceiro,<sup>a,b,c</sup> Tina Keller-Costa,<sup>a,b,c</sup> Nikos C. Kyrpides,<sup>d</sup>  Tanja Woyke,<sup>d</sup>  William B. Whitman,<sup>e</sup>  Rodrigo Costa<sup>a,b,c,f</sup>

<sup>a</sup>Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

<sup>b</sup>4HB—Institute for Health and Bioeconomy, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

<sup>c</sup>Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

<sup>d</sup>Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California, USA

<sup>e</sup>Department of Microbiology, University of Georgia, Athens, Georgia, USA

<sup>f</sup>Centro de Ciências do Mar (CCMAR), Universidade do Algarve, Faro, Portugal

**ABSTRACT** We report here the genome sequences of three *Aquimarina megaterium* strains isolated from the octocoral *Eunicella labiata*. We reveal a coding potential for versatile carbon metabolism and biosynthesis of natural products belonging to the polyketide, nonribosomal peptide, and terpene compound classes.

Members of the genus *Aquimarina* (*Bacteroidetes*, *Flavobacteriaceae*) play important roles in carbon and nitrogen cycling in marine environments (1). They are also a rich reservoir of secondary metabolites with promising antimicrobial activities (1–3). Many *Aquimarina* species live in association with eukaryotic hosts, such as marine sponges (4), octocorals (5, 6), clams (7), or algae (8), although some strains are emerging pathogens, causing epizootic shell disease in crustaceans (9). *Aquimarina megaterium* is a strictly aerobic, nonflagellated bacterium that was first isolated from surface seawater in the South Pacific (10). Only one *A. megaterium* genome, that of planktonic type strain XH134, is currently available. Here, we report the genomes of three octocoral-associated *A. megaterium* strains isolated from *Eunicella labiata* in the northeast Atlantic.

The strains were isolated from one *E. labiata* specimen collected off the coast of Faro, Portugal (6). Host-derived microbial cell suspensions were plated on half-strength marine agar and incubated at 18°C for 1 week, and single colonies were streaked until purity on the same medium (6). *Aquimarina* isolates were identified by Sanger sequencing of 16S rRNA genes amplified from genomic DNA extracted from pure cultures freshly grown in marine broth, using the Wizard genomic DNA purification kit (Promega, USA) (6). The same DNA extracts were used for genome sequencing at the Joint Genome Institute (JGI) as part of the Genomic Encyclopedia of Type Strains Phase IV project. Default parameters were used for all software unless otherwise specified. Genome libraries (300 bp) were prepared with the KAPA HyperPrep kit (Kapa Biosystems) and sequenced using the Illumina NovaSeq 6000 platform (S4 flow cell). Raw reads were quality filtered per JGI standard operating practice (SOP) protocol 1061 using BBTtools v.38.86 (<http://bbtools.jgi.doe.gov>). Filtered reads were assembled into contigs using SPAdes v.3.14.1 (11) with 25, 55, and 95 k-mers. Organism and project metadata were deposited in the Genomes OnLine database (12), and contigs were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP v.6.2) (13) and DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4) (14). Results were submitted to the Integrated Microbial Genomes and Microbiomes system (IMG/M) for comparative analysis (15). Genome completeness and contamination were assessed with the Microbial Genomes Atlas (16). AntiSMASH v6.0 (17) was used to identify secondary metabolite biosynthetic gene clusters (SM-BGCs).

The general features of the genomes are provided in Table 1. Average nucleotide identities (ANIs) obtained on IMG/M (15) for octocoral-derived strains EL32, EL35, and

**Editor** Frank J. Stewart, Montana State University

**Copyright** © 2022 Couceiro et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Rodrigo Costa, [rodrigoscosta@tecnico.ulisboa.pt](mailto:rodrigoscosta@tecnico.ulisboa.pt).

The authors declare no conflict of interest.

**Received** 17 June 2022

**Accepted** 4 October 2022

**Published** 19 October 2022

**TABLE 1** General features of the *Aquimarina megaterium* strain genomes reported in this study

Strain <sup>a</sup> IMG ID <sup>b</sup>	Size (Mb)	GC (%)	Coverage (N <sub>50</sub> ) <sup>c</sup>	No. of		Length (bp)	Estimate (%) of:		Genes	CDSs	Count by <sup>d</sup> :				Accession no. <sup>e</sup>		
				contigs	Read		Completeness	Contamination			RNA	rRNA	tRNA	ncRNA		Pfam	
EL_32	2880550035	6.06	33.00	249.6 ×	44 (334,002)	22,409,030	151	97.2	0.9	5,331† 5,213*	5,266† 5,166*	46† 47*	3† 3*	40† 40*	3† 4*	3,842†	JADOU0000000000, GCF_015752005.1, SRR13202134
EL_35	2880555367	6.06	32.99	248.9 ×	46 (325,984)	27,572,010	151	97.2	0.9	5,330† 5,215*	5,265† 5,168*	46† 47*	3† 3*	40† 40*	3† 4*	3,843†	JADOUH0000000000, GCF_015751855.1, SRR13202145
EL_43	2880762256	6.05	33.00	249.7 ×	48 (326,032)	28,518,382	151	97.2	0.9	5,321† 5,203*	5,256† 5,156*	46† 47*	3† 3*	40† 40*	3† 4*	3,839†	JADOU0000000000, GCF_015752065.1, SRR13202146

<sup>a</sup> All strains reported in this study have been isolated from the octocoral host *Eunicella labiata*.

<sup>b</sup> Unique genome identifier on the Integrated Microbiomes and Metagenomes (IMG/M) portal.

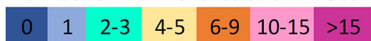
<sup>c</sup> Contig N<sub>50</sub> metrics are provided in numbers of base pairs.

<sup>d</sup> Annotation was performed using the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4) (†) and the NCBI Prokaryotic Genome Annotation Pipeline (Pgap v.6.2) (\*).

<sup>e</sup> Accession numbers for GenBank, the assembly, and raw reads are listed in that order.

Pfam ID (A)	Description	EL32	EL35	EL43
pfam00263	Bacterial type II and III secretion system protein	1	1	1
pfam00437	Type II/IV secretion system protein	1	1	1
pfam00482	Type II secretion system (T2SS), protein F	1	1	1
pfam04717	Type VI secretion system/phage-baseplate injector OB domain	1	1	1
pfam10614	Type VIII secretion system (T8SS), CsgF protein	1	1	1
pfam11751	Type IX secretion system membrane protein PorP/SprF	12	12	12
pfam13585	CHU_C Type IX secretion signal domain	17	17	17
pfam19572	Type IX secretion system protein PorV	1	1	1
Pfam ID (B)	Description	EL32	EL35	EL43
pfam00150	Cellulase (glycosyl hydrolase family 5)	2	2	2
pfam00182	Chitinase class I	1	1	1
pfam01832	Mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase	1	1	1
pfam02927	Cellulase N-terminal ig-like domain	1	1	1
pfam05089	Alpha-N-acetylglucosaminidase (NAGLU) tim-barrel domain	1	1	1
pfam10566	Glycoside hydrolase 97	2	2	2
pfam12971	Alpha-N-acetylglucosaminidase (NAGLU) N-terminal domain	1	1	1
pfam12972	Alpha-N-acetylglucosaminidase (NAGLU) C-terminal domain	1	1	1
pfam14508	Glycosyl-hydrolase 97 N-terminal	2	2	2
pfam14509	Glycosyl-hydrolase 97 C-terminal, oligomerisation	2	2	2
pfam16317	Glycosyl hydrolase family 99	1	1	1
pfam17678	Glycosyl hydrolase family 92 N-terminal domain	6	6	6
SM-BCGs (C)	Description	EL32	EL35	EL43
Arylpolyene-Resorcinol	Hybrid aryl polyene-resorcinol cluster [ <b>flexirubin: 75%</b> ]	1	1	1
Siderophore	Siderophore cluster [ <b>bisucaberin B: 100%</b> ]	1	1	1
RRE-containing-RiPP-like	Hybrid RRE-element containing and RiPP-like cluster	1	1	1
RiPP-like	Ribosomally synthesized & post-translationally modified peptides	2	2	2
Terpene	Terpene cluster [ <b>(1) carotenoid: 25%</b> ]	2	2	2
T3PKS	Type III polyketide synthase cluster	1	1	1
NRPS	Non-ribosomal peptide synthase cluster [ <b>(1) tyrocidine: 12%</b> ]	5	5	5
transAT-PKS-PKS-like	Hybrid transAT polyketide synthase cluster and other types of PKS cluster [ <b>macrobrevin: 26%</b> ]	1	1	1
transAT-PKS-like	Trans-AT PKS fragment, with trans-AT domain not found	2	2	2

**Number of Pfam entries or SM-BCGs:**



**FIG 1** Presence and abundance of select functional features encoded by the *Aquimarina megaterium* genomes reported in this study. Pfam entries involved in protein secretion systems (A) and carbohydrate metabolism (B) are shown, along with SM-BCGs (C) identified with AntiSMASH v.6.0.

EL43 and *A. megaterium* type strain XH134 were 99.06% or above, ascertaining their same-species status. Likewise, digital DNA-DNA hybridization (dDDH) probabilities (18) calculated for all strain combinations were above 97%.

All three genomes contain genes encoding several glycoside hydrolases, feature cellulase-, chitinase-, and *N*-acetylglucosaminidase-encoding genes, and contain multiple domains related to the type IX secretion system involved in gliding motility and chitinase export (19) (Fig. 1), underpinning a versatile carbon metabolism (1). Moreover, all strains possess the potential to synthesize terpenes, polyketides, nonribosomal and ribosomal peptides, flexirubin-like pigments, and siderophores (Fig. 1), congruent with the notion of *Aquimarina* species as reservoirs of chemical diversity of biotechnological interest (1, 3).

**Data availability.** The genome sequences of the three *Aquimarina megaterium* strains have been deposited in GenBank/NCBI by the JGI. GenBank accession numbers are listed in Table 1.

## ACKNOWLEDGMENTS

This study was financed by the Fundo Azul program of “Direção-Geral de Política do Mar” (DGPM, Ministry of the Sea, Portugal) through grant FA\_05\_2017\_032. Sequencing, assembly, and annotation of the three *Aquimarina megaterium* genomes were performed at the Joint Genome Institute (JGI) within the framework of the Genomic Encyclopedia of Type Strains Phase IV project (KMG-IV): “sequencing the most valuable type-strain genomes for metagenomic binning, comparative biology, and taxonomic classification.” The work (proposal 10.46936/10.25585/60001079) conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under contract no. DE-AC02-05CH11231. Further support was provided to the Institute of Bioengineering and Biosciences by “Programa Operacional Regional de Lisboa” (project no. 007317) and Fundação para a Ciência e a Tecnologia (FCT) and the European Regional Development Fund (ERDF) through grant UIDB/04565/2020. J.F.C. is the recipient of a Ph.D. scholarship conceded by FCT in the framework of the MIT Portugal Ph.D. program (PRT/BD/153503/2021). T.K.C. is the recipient of a Research Scientist contract conceded by FCT (CEECIND/00788/2017).

## REFERENCES

- Silva SG, Blom J, Keller-Costa T, Costa R. 2019. Comparative genomics reveals complex natural product biosynthesis capacities and carbon metabolism across host-associated and free-living *Aquimarina* (*Bacteroidetes*, *Flavobacteriaceae*) species. *Environ Microbiol* 21:4002–4019. <https://doi.org/10.1111/1462-2920.14747>.
- Esteves ALS, Hardoim CCP, Xavier JR, Gonçalves JMS, Costa R. 2013. Molecular richness and biotechnological potential of bacteria cultured from *Irciniidae* sponges in the north-east Atlantic. *FEMS Microbiol Ecol* 85:519–536. <https://doi.org/10.1111/1574-6941.12140>.
- Dieterich CL, Probst SI, Ueoka R, Sandu I, Schäfle D, Molin MD, Minas HA, Costa R, Oxenius A, Sander P, Piel J. 2022. Aquimarinins, peptide antibiotics with amino-modified C-termini from a sponge-derived *Aquimarina* sp. bacterium. *Angew Chem Int Ed Engl* 61:e202115802. <https://doi.org/10.1002/anie.202115802>.
- Yoon BJ, You HS, Lee DH, Oh DC. 2011. *Aquimarina spongiae* sp. nov., isolated from marine sponge *Halichondria oshoro*. *Int J Syst Evol Microbiol* 61:417–421. <https://doi.org/10.1099/ijs.0.022046-0>.
- Keller-Costa T, Silva R, Lago-Lestón A, Costa R. 2016. Genomic insights into *Aquimarina* sp. strain EL33, a bacterial symbiont of the gorgonian coral *Eunicella labiata*. *Genome Announc* 4:e00855-16. <https://doi.org/10.1128/genomeA.00855-16>.
- Keller-Costa T, Eriksson D, Gonçalves JMS, Gomes NCM, Lago-Leston A, Costa R. 2017. The gorgonian coral *Eunicella labiata* hosts a distinct prokaryotic consortium amenable to cultivation. *FEMS Microbiol Ecol* 93:fx143. <https://doi.org/10.1093/femsec/fox143>.
- Park SC, Choe HN, Baik KS, Seong CN. 2012. *Aquimarina mytili* sp. nov., isolated from the gut microflora of a mussel, *Mytilus coruscus*, and emended description of *Aquimarina macrocephali*. *Int J Syst Evol Microbiol* 62:1974–1979. <https://doi.org/10.1099/ijs.0.032904-0>.
- Lin B, Lu G, Zheng Y, Xie W, Li S, Hu Z. 2012. *Aquimarina agarilytica* sp. nov., an agarolytic species isolated from a red alga. *Int J Syst Evol Microbiol* 62:869–873. <https://doi.org/10.1099/ijs.0.027136-0>.
- Quinn RA, Metzler A, Smolowitz RM, Tlusty M, Chistoserdov AY. 2012. Exposures of *Homarus americanus* shell to three bacteria isolated from naturally occurring epizootic shell disease lesions. *J Shell Res* 31:485–493. <https://doi.org/10.2983/035.031.0208>.
- Yu T, Zhang Z, Fan X, Shi X, Zhang XH. 2014. *Aquimarina megaterium* sp. nov., isolated from seawater. *Int J Syst Evol Microbiol* 64:122–127. <https://doi.org/10.1099/ijs.0.055517-0>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Mukherjee S, Stamatis D, Bertsch J, Ovchinnikova G, Sundaramurthi JC, Lee J, Kandimalla M, Chen IMA, Kyrpides NC, Reddy TBK. 2021. Genomes OnLine Database (GOLD) v.8: overview and updates. *Nucleic Acids Res* 49:D723–D733. <https://doi.org/10.1093/nar/gkaa983>.
- Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Huntemann M, Ivanova NN, Mavromatis K, James Tripp H, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen IMA, Pati A, Nielsen T, Markowitz VM, Kyrpides NC. 2015. The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4). *Stand Genomic Sci* 10:86. <https://doi.org/10.1186/s40793-015-0077-y>.
- Chen IMA, Chu K, Palaniappan K, Ratner A, Huang J, Huntemann M, Hajek

- P, Ritter S, Varghese N, Seshadri R, Roux S, Woyke T, Elie-Fadrosh EA, Ivanova NN, Kyrpides NC. 2021. The IMG/M data management and analysis system v.6.0: new tools and advanced capabilities. *Nucleic Acids Res* 49:D751–D763. <https://doi.org/10.1093/nar/gkaa939>.
16. Rodriguez-R LM, Gunturu S, Harvey WT, Rosselló-Mora R, Tiedje JM, Cole JR, Konstantinidis KT. 2018. The Microbial Genomes Atlas (MiGA) web-server: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic Acids Res* 46:W282–W288. <https://doi.org/10.1093/nar/gky467>.
17. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, Medema MH, Weber T. 2019. AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 47:W81–W87. <https://doi.org/10.1093/nar/gkz310>.
18. Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 50:D801–D807. <https://doi.org/10.1093/nar/gkab902>.
19. Lauber F, Deme JC, Lea SM, Berks BC. 2018. Type 9 secretion system structures reveal a new protein transport mechanism. *Nature* 564:77–82. <https://doi.org/10.1038/s41586-018-0693-y>.