





The *Roseibium album* (*Labrenzia alba*) Genome Possesses Multiple Symbiosis Factors Possibly Underpinning Host-Microbe Relationships in the Marine Benthos

Joana Fernandes Couceiro,^{a,b} Tina Keller-Costa,^{a,b} Matilde Marques,^{a,b} Nikos C. Kyrpides,^{c,d}  Tanja Woyke,^{c,d} William B. Whitman,^e  Rodrigo Costa^{a,b,c,d,f}

^aInstitute for Bioengineering and Biosciences, Department of Bioengineering, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal

^bInstitute for Health and Bioeconomy Associate Laboratory, Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal

^cDepartment of Energy, Joint Genome Institute, Berkeley, California, USA

^dLawrence Berkeley National Laboratory, Berkeley, California, USA

^eDepartment of Microbiology, University of Georgia, Athens, Georgia, USA

^fCentro de Ciências do Mar, University of Algarve, Faro, Portugal

ABSTRACT Here, we announce the genomes of eight *Roseibium album* (synonym *Labrenzia alba*) strains that were obtained from the octocoral *Eunicella labiata*. Genome annotation revealed multiple symbiosis factors common to all genomes, such as eukaryotic-like repeat protein- and multidrug resistance-encoding genes, which likely underpin symbiotic relationships with marine invertebrate hosts.

Roseibium album (1) (homotypic synonyms, *Stappia alba* [2] and *Labrenzia alba* [3]) is a Gram-negative, marine alphaproteobacterium (order *Hyphomicrobiales*, family *Stappiaceae*) that has frequently been isolated from sessile, filter-feeding invertebrates such as sponges, corals, and oysters (2, 4–6). To date, however, only a few genomes of this species are available. To illuminate the putative roles of *R. album* in association with marine animals, here we report the genomes of eight *R. album* strains that were isolated from the octocoral *Eunicella labiata* (Table 1), and we present symbiosis factors and environmental resistance traits common to all genomes (Fig. 1).

The eight *R. album* strains were isolated from *E. labiata* specimens that had been collected off the coast of Faro, Portugal, after plating of host-derived homogenates on half-strength marine agar medium followed by incubation at 18°C for 4 weeks (4). Genomic DNA was extracted from pure cultures using the Wizard genomic DNA purification kit (Promega, Madison, WI, USA) as described previously (4). The isolates were sequenced at the Joint Genome Institute (JGI) as part of the Genomic Encyclopedia of Type Strains Phase IV (KMG-V) 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 BBTools v38.86 (<http://bbtools.jgi.doe.gov>). Filtered reads were assembled into contigs using SPAdes v3.14.1 (7) with 25, 55, and 95 k-mers, and contigs were annotated using the Integrated Microbial Genomes and Microbiomes (IMG/M) pipeline v5.0.17 (8). We report the Clusters of Orthologous Groups of proteins (COG) profiles obtained for all eight strains using the IMG/M platform (9). Average nucleotide identity (ANI) values obtained on the IMG/M platform (9) for all pairwise combinations among the strains were above 98.3%, supporting the same-species status of the strains. Genome completeness and contamination scores were assessed with the Microbial Genomes Atlas (MiGA) (10) (Table 1). AntiSMASH v6.0 (11) was used to identify secondary metabolite biosynthetic gene clusters (SM-BGCs).

Citation Couceiro JF, Keller-Costa T, Marques M, Kyrpides NC, Woyke T, Whitman WB, Costa R. 2021. The *Roseibium album* (*Labrenzia alba*) genome possesses multiple symbiosis factors possibly underpinning host-microbe relationships in the marine benthos. *Microbiol Resour Announc* 10:e00320-21. <https://doi.org/10.1128/MRA.00320-21>.

Editor J. Cameron Thrash, University of Southern California

Copyright © 2021 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.

Received 30 March 2021

Accepted 19 July 2021

Published 26 August 2021

TABLE 1 General features of the *Roseibium album* (*Labrenzia alba*) genomes reported in this study

Genome ^a	IMG/M identifier ^b	Genome size (Mb)	GC content (%)	Genome coverage (x)	No. of contigs	Contig N ₅₀ (bp)	No. of reads	Read length (bp)	Completeness (%)	Contamination (%)	No. of coding sequences	No. of RNAs	No. of genes in COGs	GenBank accession no.	Assembly accession no.	SRA accession no.
<i>Labrenzia alba</i> EL_13	2882936267	7.00	56.44	211.9	19	668,536	19,078,598	151	98.1	2.8	6,475	59	5,303	JADOU0000000000	GCA_015751945.1	SRR13202144
<i>Labrenzia alba</i> EL_126	2880767578	7.14	56.34	210.4	76	352,437	31,261,630	151	98.1	2.8	6,732	61	5,352	JADOUN0000000000	GCA_015752395.1	SRR13202553
<i>Labrenzia alba</i> EL_132	2880560698	7.06	56.44	215.3	51	400,118	33,793,138	151	98.1	2.8	6,629	56	5,334	JADOU0000000000	GCA_015752105.1	SRR13202552
<i>Labrenzia alba</i> EL_142	2880567409	7.29	56.31	207.9	54	363,626	31,961,096	151	98.1	2.8	6,808	66	5,460	JADOU0000000000	GCA_015752355.1	SRR13202549
<i>Labrenzia alba</i> EL_159	2880774395	6.9	56.44	215.3	18	899,415	32,269,048	151	98.1	2.8	6,384	62	5,230	JADOU0000000000	GCA_015752425.1	SRR13202554
<i>Labrenzia alba</i> EL_162	2880574310	6.9	56.44	214.6	18	899,415	20,727,438	151	98.1	2.8	6,387	62	5,230	JADOU0000000000	GCA_015752365.1	SRR13202550
<i>Labrenzia alba</i> EL_195	2880580784	7.13	56.36	211.6	31	464,795	36,215,300	151	98.1	2.8	6,593	59	5,356	JADOU0000000000	GCA_015752055.1	SRR13202548
<i>Labrenzia alba</i> EL_208	2880780866	7.06	56.44	214.8	49	402,705	22,234,602	151	98.1	2.8	6,638	56	5,334	JADOU0000000000	GCA_015752025.1	SRR13202551

^a All strains reported in this study were isolated from the octocoral host *Eumicella labiata*.

^b Unique genome identifier at the IMG/M portal.

COG ID	Description	EL13	EL126	EL132	EL142	EL159	EL162	EL195	EL208
COG0457	Tetratricopeptide (TPR) repeat	14	18	15	15	15	15	16	15
COG0790	TPR repeat	12	9	9	11	11	11	11	9
COG0666	Ankyrin repeat	1	1	1	1	1	1	1	1
COG2319	WD40 repeat	6	4	5	4	4	4	4	5
COG1566	Multidrug resistance efflux pump	7	8	7	8	6	6	6	7
COG4300	Cadmium resistance protein CadD, predicted permease	1	1	1	1	1	1	1	1
COG3793	Tellurite resistance protein	2	3	3	2	3	3	3	3
COG4103	Tellurite resistance protein B (TerB) family	2	2	2	2	2	2	2	2
COG1275	Tellurite resistance protein TehA and related permeases	0	0	0	0	0	0	0	0
COG0726	Peptidoglycan/xylan/chitin deacetylase, PgdA/CDA1 family	3	2	2	4	3	3	3	2
COG2861	YibQ, putative polysaccharide deacetylase 2 family	1	1	1	1	1	1	1	1
COG3525	N-acetyl-beta-hexosaminidase	1	1	1	1	1	1	1	1
COG0363	Glucosamine-6-phosphate isomerase/deaminase	2	1	2	1	1	1	2	2
COG2175	Taurine dioxygenase, alpha-ketoglutarate-dependent	3	2	3	3	3	3	3	3
COG4521	ABC-type taurine transport system, periplasmic comp.	1	1	1	1	1	1	1	1
COG4525	ABC-type taurine transport system, ATPase component	1	1	1	1	1	1	1	1
COG3119	Arylsulfatase A or related enzyme	10	9	9	9	9	9	10	9
SM-BGCs	Description								
transAT-PKS	TransAT polyketide synthase cluster	1	1	1	1	1	1	1	1
T3PKS	Type III polyketide synthase cluster	1	1	1	1	1	1	1	1
NRPS	Non-ribosomal peptide synthetase cluster	1	1	1	1	1	1	1	1
RiPP-like	Ribosomally synthesized & post-translationally modified peptides	2	2	2	1	1	1	1	2
TfuA-related	Thioamidite cluster	1	1	1	2	1	1	1	1
Thiopeptide	Thiopeptide cluster	1	1	1	2	1	1	1	1
Lanthipeptide	Lanthipeptide class II cluster	1	1	1	0	0	0	0	1
Terpene	Terpene cluster	1	1	1	1	1	1	1	1
Betalactone	Betalactone cluster	1	1	1	1	1	1	1	1
Hserlactone	Homoserine lactone cluster	0	1	1	1	0	0	0	1

Number of COG entries or SM-BGCs:
0 1 2-3 4-6 7-10 11-15 16-20

FIG 1 Selected COG functions characteristic of the *Roseibium album* (*Labrenzia alba*) genomes described here, as well as SM-BGCs present in all genomes. Values for each entry represent the numbers of coding sequences assigned to COG functions per genome (top) and the numbers of SM-BGCs coding for major compound classes (such as polyketides and terpenes) per genome (bottom).

Analysis of the *R. album* genomes revealed the presence of various COGs important for the establishment of symbiotic relationships, including eukaryotic-like WD40, ankyrin, and tetratricopeptide repeats (Fig. 1). COG functions related to drug (e.g., COG1566) and heavy metal resistance were common to all eight genomes, including traits specific for tellurite resistance that are usually encoded on plasmids (12) (Fig. 1). The eight strains harbor genes for the utilization of chitin, a trait that was previously reported for sponge and coral symbionts (13), as revealed by the presence of polysaccharide deacetylases and exochitinases. Other features included a coding potential for a possible role in sulfur cycling, e.g., through the catabolism of the sulfolipid cerebroside 3-sulfate and of the amino-sulfonic acid taurine, two compounds that are widely synthesized in animal tissue and utilized by bacterial symbionts (14). Moreover, all *R. album* genomes possess the potential to produce a diverse range of secondary metabolites. Indeed, we found a variety of SM-BGCs encoding terpenes, *trans*-AT-polyketide synthases and type III polyketide synthases, nonribosomal peptide synthetases, and several ribosomally synthesized peptides in this genome pool (Fig. 1).

Data availability. The genome sequences of the eight *R. album* (*L. alba*) strains have been deposited in GenBank by the JGI. 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 eight *Roseibium album*/*Labrenzia alba* genomes were performed at the JGI within the framework of the KMG-V. The work conducted by the JGI, a U.S. Department of Energy (DOE) Office of Science User Facility,

is supported by the Office of Science of the DOE under contract DE-AC02-05CH11231. Further support was provided to the Institute for Bioengineering and Biosciences by Programa Operacional Regional de Lisboa (project 007317), the Fundação para a Ciência e a Tecnologia (FCT), and the European Regional Development Fund (ERDF) through grant UIDB/04565/2020. T.K.-C. is the recipient of a research scientist contract conceded by FCT (grant CEECIND/00788/2017).

REFERENCES

- Hördt A, López MG, Meier-Kolthoff JP, Schleuning M, Weinhold LM, Tindall BJ, Gronow S, Kyrpides NC, Woyke T, Göker M. 2020. Analysis of 1,000+ type-strain genomes substantially improves taxonomic classification of *Alphaproteobacteria*. *Front Microbiol* 11:468. <https://doi.org/10.3389/fmicb.2020.00468>.
- Pujalte MJ, Macián MC, Arahal DR, Garay E. 2005. *Stappia alba* sp. nov., isolated from Mediterranean oysters. *Syst Appl Microbiol* 28:672–678. <https://doi.org/10.1016/j.syapm.2005.05.010>.
- Biebl H, Pukall R, Lünsdorf H, Schulz S, Allgaier M, Tindall BJ, Wagner-Döbler I. 2007. Description of *Labrenzia alexandrii* gen. nov., sp. nov., a novel alphaproteobacterium containing bacteriochlorophyll a, and a proposal reclassification of *Stappia aggregata* as *Labrenzia aggregata* comb. nov., and of *Stappia alba* as *Labrenzia alba* comb. nov., and emended descriptions of the genera *Pannonibacter*, *Stappia* and *Roseibium*, and of the species *Roseibium denhamense* and *Roseibium hamelinense*. *Int J Syst Evol Microbiol* 57:1095–1107. <https://doi.org/10.1099/ijs.0.64821-0>.
- Keller-Costa T, Eriksson D, Gonçalves JMS, Gomes NCM, Lago-Lestón A, Costa R. 2017. The gorgonian coral *Eunicella labiata* hosts a distinct prokaryotic consortium amenable to cultivation. *FEMS Microbiol Ecol* 93: fix143. <https://doi.org/10.1093/femsec/fix143>.
- Rodrigues GN, Lago-Lestón A, Costa R, Keller-Costa T. 2018. Draft genome sequence of *Labrenzia* sp. strain EL143, a coral-associated alphaproteobacterium with versatile symbiotic living capability and strong halogen degradation potential. *Genome Announc* 6:e00132-18. <https://doi.org/10.1128/genomeA.00132-18>.
- Karimi E, Keller-Costa T, Slaby BM, Cox CJ, da Rocha UN, Hentschel U, Costa R. 2019. Genomic blueprints of sponge-prokaryote symbiosis are shared by low abundant and cultivatable *Alphaproteobacteria*. *Sci Rep* 9:1999. <https://doi.org/10.1038/s41598-019-38737-x>.
- 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>.
- Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen IM, 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 IA, Chu K, Palaniappan K, Ratner A, Huang J, Huntemann M, Hajek P, Ritter S, Varghese N, Seshadri R, Roux S, Woyke T, Eloe-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>.
- 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>.
- 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>.
- Chasteen TG, Fuentes DE, Tantaleán JC, Vásquez CC. 2009. Tellurite: history, oxidative stress, and molecular mechanisms of resistance. *FEMS Microbiol Rev* 33:820–832. <https://doi.org/10.1111/j.1574-6976.2009.00177.x>.
- Raimundo I, Silva R, Meunier L, Valente SM, Lago-Lestón A, Keller-Costa T, Costa R. 2021. Functional metagenomics reveals differential chitin degradation and utilization features across free-living and host-associated marine microbiomes. *Microbiome* 9:43. <https://doi.org/10.1186/s40168-020-00970-2>.
- Karimi E, Slaby BM, Soares AR, Blom J, Hentschel U, Costa R. 2018. Metagenomic binning reveals versatile nutrient cycling and distinct adaptive features in alphaproteobacterial symbionts of marine sponges. *FEMS Microbiol Ecol* 94:fy074. <https://doi.org/10.1093/femsec/fy074>.