data

MDPI

Data Descriptor

Genome Analysis of the Marine Bacterium *Labrenzia* sp. Strain 011, a Potential Protective Agent of Mollusks

Jamshid Amiri Moghaddam [1], Antonio Dávila-Céspedes [1], Mohammad Alanjary [2], Jochen Blom [3], Gabriele M. König [1,4,*] and Till F. Schäberle [4,5,6,*]

1 Institute for Pharmaceutical Biology, University of Bonn, 53115 Bonn, Germany; jamirimoghaddam@uni-bonn.de (J.A.M.); antonio.dc@uni-bonn.de (A.D.-C.)
2 Bioinformatics Group, Wageningen University, 6708PB Wageningen, The Netherlands; mohammad.alanjary@wur.nl
3 Bioinformatics and Systems Biology, Justus-Liebig-University Giessen, 35390 Giessen, Germany; jochen.blom@computational.bio.uni-giessen.de
4 German Center for Infection Research (DZIF), Partner Site Cologne/Bonn, 50935 Cologne, Germany
5 Institute for Insect Biotechnology, Justus-Liebig-University Giessen, 35390 Giessen, Germany
6 Department of Bioresources, Fraunhofer Institute for Molecular Biology and Applied Ecology, 35394 Giessen, Germany
* Correspondence: g.koenig@uni-bonn.de (G.M.K.); till.f.schaeberle@agrar.uni-giessen.de (T.F.S.)

Received: 18 January 2019; Accepted: 13 February 2019; Published: 20 February 2019

Abstract: The marine bacterium *Labrenzia* sp. strain 011 was isolated from the coastal sediment of Kronsgaard, Germany. The *Labrenzia* species are suggested to be protective agents of mollusks. *Labrenzia* sp. strain 011 produces specialized metabolites, which showed activity against a range of microorganisms, thereunder strong inhibitory effects against *Pseudoroseovarius crassostreae* DSM 16,950 (genus *Roseovarius*), the causative agent of oyster disease. The genome of *Labrenzia* sp. strain 011 was sequenced and assembled into 65 contigs, has a size of 5.1 Mbp, and a G+C content of 61.6%. A comparative genome analysis defined *Labrenzia* sp. strain 011 as a distinct new species within the genus *Labrenzia*, whereby 44% of the genome was contributed to the *Labrenzia* core genome. The genomic data provided here is expected to contribute to a deeper understanding of the mollusk-protective role of *Labrenzia* spp.

Dataset: This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. QCYM00000000. The version described in this paper is the first version, QCYM01000000 (https://www.ncbi.nlm.nih.gov/nuccore/QCYM01000000).

Keywords: *Labrenzia*; draft genome; comparative genomics; antimicrobial; oyster disease; *Roseovarius crassostreae*

1. Summary

Bacteria of the genus *Labrenzia* colonize surfaces, such as oyster shells, and may produce antibacterial compounds, which inhibit the growth of other bacteria [1–4]. *Labrenzia* sp. strain 011 showed activity

against the oyster pathogen *Roseovarius crassostreae* [1]. *R. crassostreae* has an adverse effect on natural oyster populations and on oyster farming operations [5]. In addition, strains of the genus *Labrenzia*, which produce compounds showing antimicrobial activity, were associated with soft corals and the marine sponge *Erylus discophorus* [6,7]. Moreover, an analysis of the available genome of *Labrenzia* sp. strain EL143 showed many genes that are linked to the symbiotic relationship with sessile hosts, genes that can be linked to resistance mechanisms against antibiotics and toxic compounds, and genes corresponding to a strong dehalogenation potential [8]. This can be regarded as a requirement for filter-feeding organisms that are exposed to halogenated substances in their environment, and might use bacterial symbionts with dehalogenase activity for detoxification and nutrition [9]. These reports reflect the importance of *Labrenzia* species and their potential for the protection of marine bivalves and for biotechnological applications. Therefore, the genome of *Labrenzia* sp. strain 011 will enable the identification of biosynthetic gene clusters corresponding to protective compounds. The data shown here can be useful for research groups working on natural product discovery, by enabling further genome-mining approaches.

2. Data Description

The draft genome sequence of *Labrenzia* sp. strain 011 consists of 65 contigs (>1000 bp) with 5,102,962 bp in length, and a G+C content of 61.6%. There were 4812 coding sequences (CDSs) that were predicted (this number includes proteins annotated as hypothetical), of which 2280 CDSs (48%) were categorized in 473 different subsystems with identified functional roles.

A phylogenetic tree of all of the *Labrenzia* strains with the available genomes based on the core genomes alignment revealed *Labrenzia* sp. strain OB1 and *L. marina* DSM 17,023, which were isolated from coastal seawater in La Jolla, CA, USA, and South Korea, respectively, as the most closely related strains to *Labrenzia* sp. strain 011 (Figure 1).

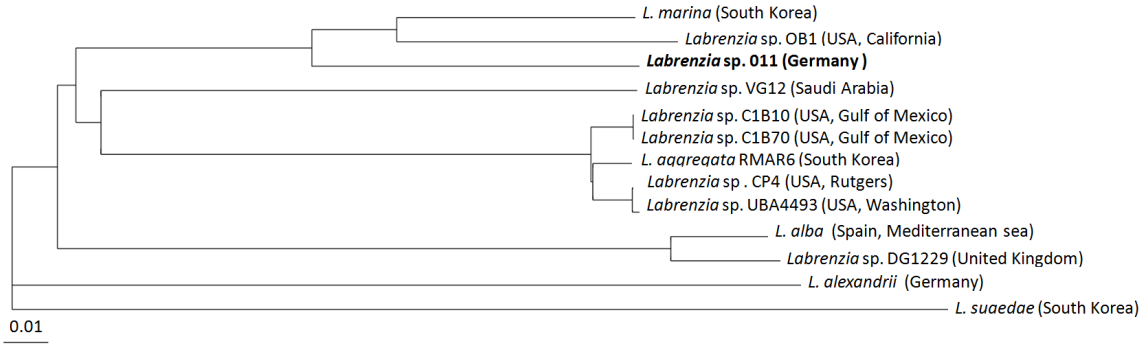

Figure 1. Phylogenetic tree of selected *Labrenzia* strains with available genomes. The tree was build out of a core of 2131 genes per genome. The geographic origins of the strains are given in parentheses. The tree was calculated with 100 iterations. All branches have 100/100 bootstrap support, except the branch between *L. aggregate* RMAR6 and *Labrenzia* sp. UBA4493/*Labrenzia* sp. CP4, which is 61/100.

In order to obtain further insight into the degree of similarity between the analyzed genomes, the numbers of the core genes and of the singletons were determined. There were 2131 CDS that contributed to the core genome of the *Labrenzia* strains, equivalent to ~44% of the *Labrenzia* sp. strain 011 genome (Figure 2A). To identify the actual core genome of a species, it is possible to use an approximate approach by extrapolating the number of core genes for an infinite number of genomes [10]. Using this methodology, it was calculated that the core genome will be around 2113 CDS, based on a decay function ($2929.005 \times \exp(-x/3.229) + 2112.783$, see Figure 2B). The pan genome increases with every additional *Labrenzia* strain, indicating an open pan genome of *Labrenzia* (Heaps' law function: $5736.13 \times x^{0.462}$, see Figure 2C).

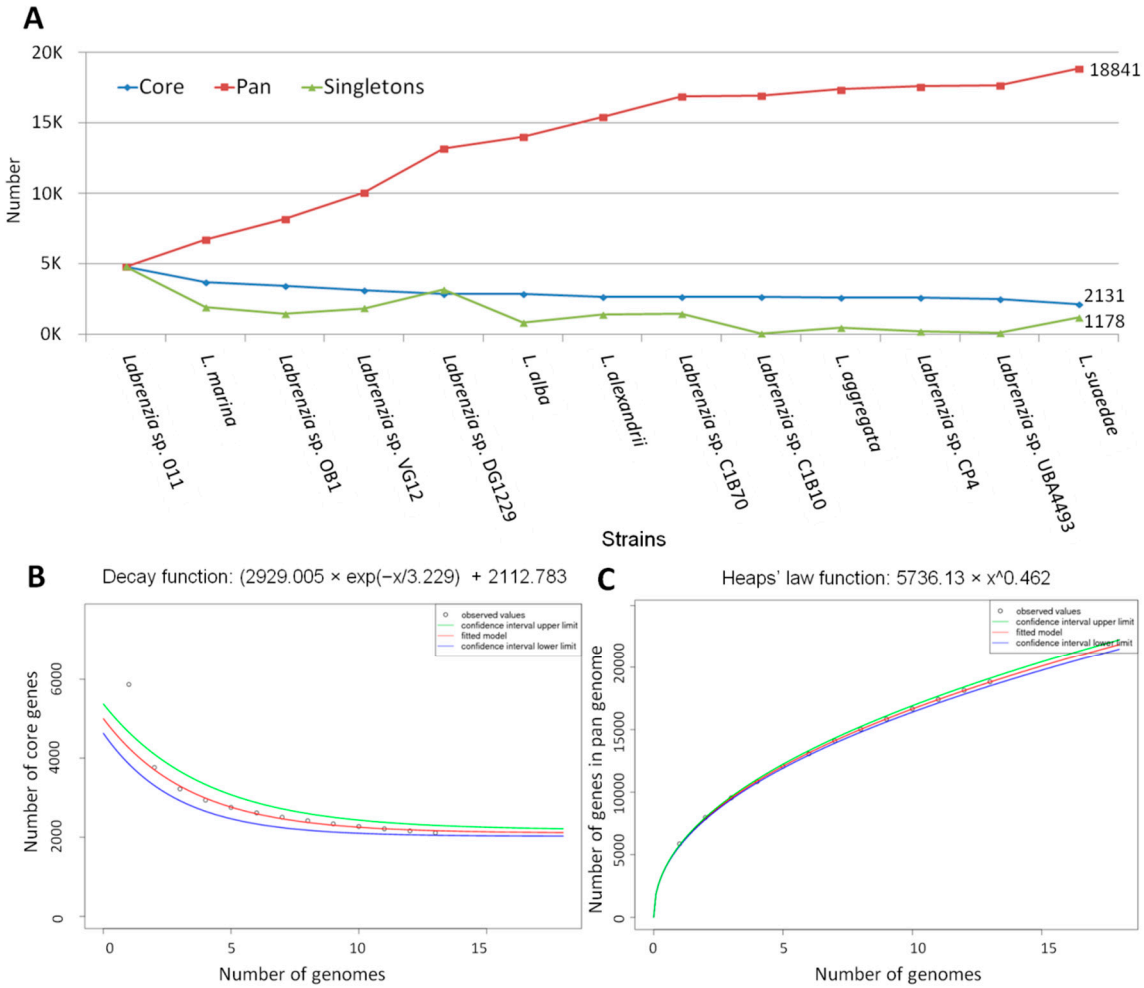

Figure 2. (**A**) Core vs. pan genome plot of the genomes. (**B**) Core genome development plot. (**C**) Pan genome development plot.

The average nucleotide identity (ANI) values between *Labrenzia* sp. strain 011 and all of the analyzed *Labrenzia* strains was between 73.55% to 84.85% in the pair-wise sequence comparisons (Figure 3). This puts the strain only into distant relation to other strains, as values smaller than 80–85% ANI must be regarded as distantly related [11]. The in-silico DNA–DNA hybridization (isDDH) values between *Labrenzia* sp. strain 011 and the other *Labrenzia* strains was between 22.7% to 33.1%, whereby the highest values were obtained for *Labrenzia* sp. strain OB1 and *L. marina* DSM 17023, verifying the phylogenetic relationship between these two and strain 011. Furthermore, differences in the G+C content between *Labrenzia* sp. strain 011 and other *Labrenzia* strains were between 1.32–5.38%, which supports the species delineation (Table 1). Therefore, the in silico parameters (ANI $\geq$ 96%, isDDH $\geq$ 70%, and difference in G+C content of $\leq$ 1%) [11–13] define *Labrenzia* sp. strain 011 as a distinct new species of the genus *Labrenzia* (Figure 3, Table 1). Instead, CP4, UBA4493, C1B70, and C1B10 seem to be strains closely related to *L. aggregata* RMAR6, with ANI values between 97–100% (Figure 3).

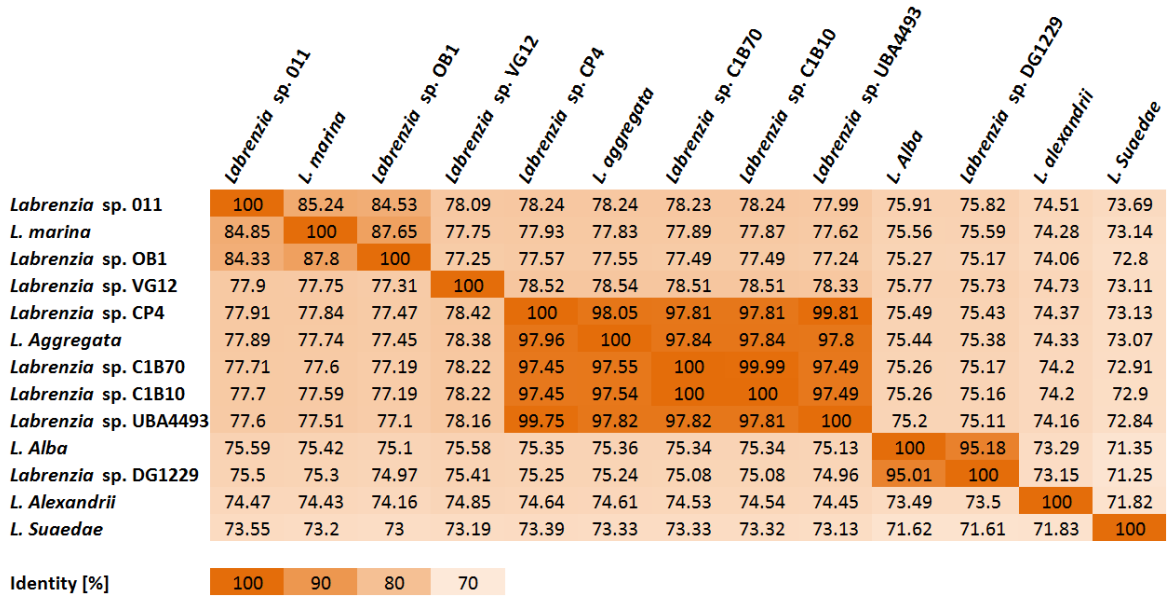

Figure 3. Average nucleotide identity (ANI) heat map of the selected *Labrenzia* strains.

Table 1. In silico DNA–DNA hybridization (isDDH) and G+C difference of *Labrenzia* sp. strain 011 vs. other *Labrenzia* strains.

***Labrenzia* sp. Strain 011 vs.**	**isDDH%**	**G+C Difference%**
Labrenzia sp. strain OB1	33.1	2.20
Labrenzia marina	30.2	1.41
Labrenzia sp. strain C1B70	26.7	2.71
Labrenzia sp. strain C1B10	26.7	2.71
Labrenzia sp. strain CP4	26.6	2.52
Labrenzia sp. strain VG12	26.5	1.62
Labrenzia aggregata	26.5	2.57
Labrenzia sp. strainUBA4493	26.3	2.56
Labrenzia sp. strain DG1229	24.8	5.38
Labrenzia alba	24.4	5.25
Labrenzia alexandrii	23.5	5.26
Labrenzia suaedae	22.7	1.32

The genome of *Labrenzia* sp. strain 011 carries genes related to nitrogen metabolism and denitrification (56 CDSs), polyhydroxybutyrate metabolism (32 CDSs), and many genes that are related to stress response, for example, heat and cold shock (169 CDSs) (Figure 4). *Labrenzia* sp. strain 011 belongs to the family of *Rhodobacteraceae*, which is a sister family of the *Rhizobiales*. The latter fix nitrogen in plant roots [14]. This data may explain the denitrification ability of the oyster microbiome, which is dominated by *Rhodobacteraceae* [15]. The bacteria of this family are surface colonizers and are known for the production of compounds with antibacterial activity, which prohibit the growth of other bacteria; thereby, shaping the microbiome [15,16].

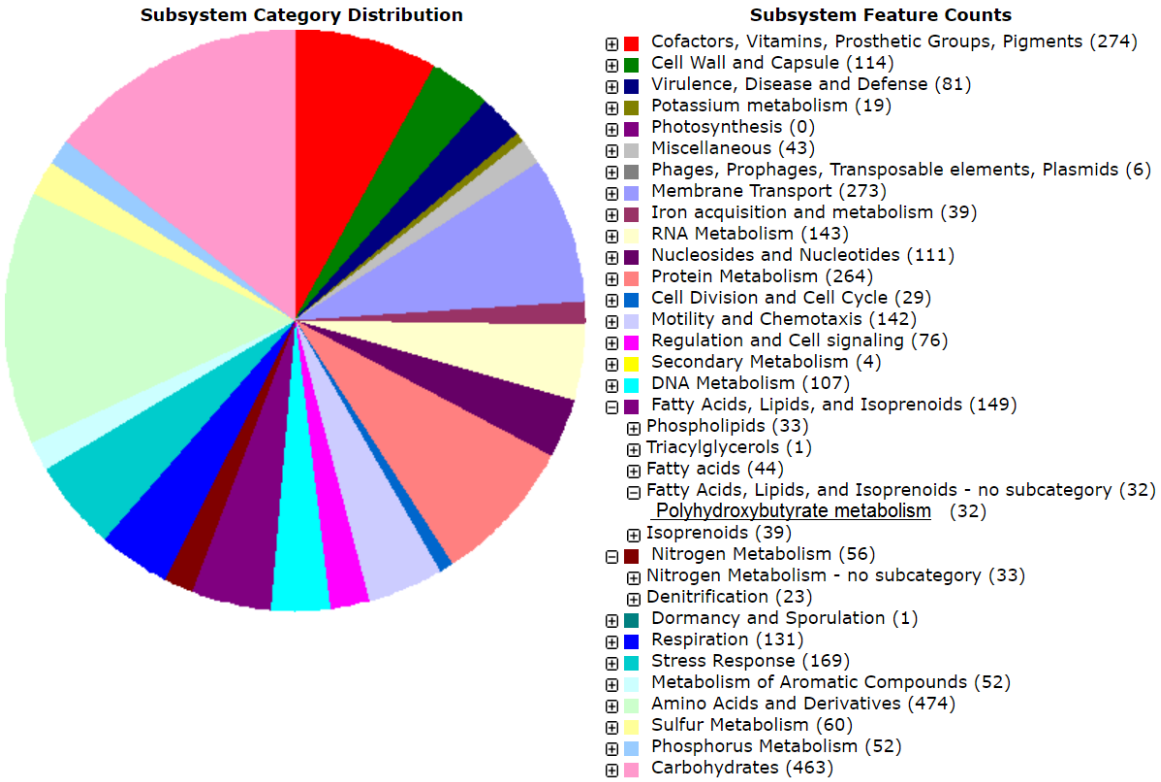

Figure 4. Subsystem category distribution and feature counts in the genome of *Labrenzia* sp. strain 011.

In total, 11 biosynthetic gene clusters (BGCs) were identified (5.3% of the genome), including one type-I polyketide synthase, one terpene, one bacteriocin, four fatty acids, and four saccharide BGCs (Figure 5). Additionally, 23 putative gene clusters were identified using the cluster finder algorithm (3.8% of the genome), thereunder three BGCs for cyclopropane fatty acid synthases (Figure 5).

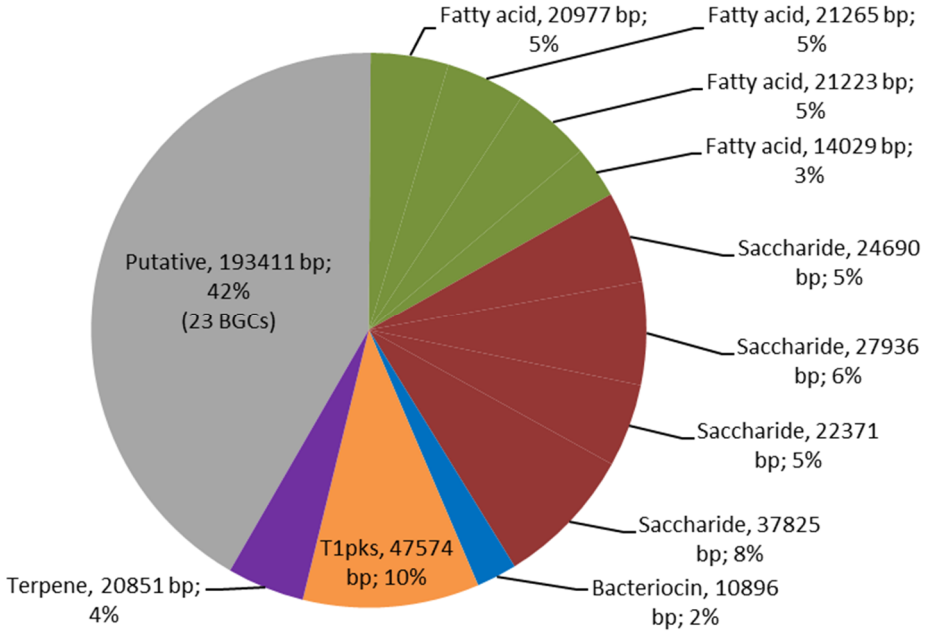

Figure 5. Distribution of the biosynthetic gene clusters (BGCs) in the genome of *Labrenzia* sp. strain 011. In total, 463,048 bp (equal to 9.1% of the genome) were identified. The identified regions and percentages of the total are given.

3. Methods

3.1. Sequencing and Assembly

The marine bacterium *Labrenzia* sp. strain 011 was isolated from sediment from the coastal area of Kronsgaard, Germany. The phenotypic appearance of its colonies is creamy yellow on Difco™ marine agar 2216 (Table 2). The genomic DNA isolation of *Labrenzia* sp. strain 011 was performed as described before [17]. In brief, a one-week culture in a marine broth liquid medium was used to harvest the cell pellets. Therefrom, the DNA was isolated using the GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich). Illumina shotgun paired-end sequencing libraries were generated and sequenced on a MiSeq instrument (Illumina, San Diego, CA, USA). Quality filtering using Trimmomatic version 0.36(6) resulted in 495,158 paired-end reads for *Labrenzia* sp. strain 011. The paired-end reads were combined using the Spades assembler v3.10, yielding initial sequence scaffolds [18]. Scaffolds smaller than 1 kb were filtered and 65 contigs remained as determined with Quast [19]. The genome completeness was estimated using CheckM [20] and the genus level marker genes, resulting in a value of 83.2%.

Table 2. Features of *Labrenzia* sp. strain 011, and MIGS mandatory information.

Items	Description
Investigation type	Bacteria
Strain	*Labrenzia* sp. 011
Gram stain	Negative
Cell shape	Rod
Pigmentation	Creamy yellow
Temperature optimum	30 °C
Latitude and longitude	54.731111 N 9.964167 E
Geographic location name	Kronsgaard, Germany
Collection date	15-Aug-2012
Environmental biome	M arine biome (ENVO:00000447)
Environmental feature	Sea coast (ENVO:00000303)
Environmental material	Marine sediment (ENVO_03000033)
Environmental package	Surface sediment
Relationship to oxygen	Aerobe
Number of replicons	1
Sequencing method	Illumina

3.2. Genome Annotation and Comparison

The coding sequences (CDS) of the genome were determined using the RAST prokaryotic genome annotation server [21]. The annotated GenBank file was uploaded to the EDGAR 2.2 genomic pipeline [22] for phylogeny and genome comparison. For this analysis, all of the available genome sequences of the *Labrenzia* strains were used (accession numbers in parentheses), as follows: *L. alexandrii* DFL-11^T (ACCU00000000), *L. aggregata* RMAR6-6 chromosome (CP019630), *L. suaedae* DSM 22153^T (FRBW00000000), *L. alba* CECT 5095^T (CXWE00000000), *Labrenzia* sp. strain CP4 (CP011927), *Labrenzia* sp. strain VG12 (CP022529), *Labrenzia* sp. strain DG1229 (AYYG00000000), *Labrenzia* sp. strain C1B10 (AXBY00000000), *Labrenzia* sp. strain C1B70 (AXCE00000000), *Labrenzia* sp. strain OB1 (JSEP00000000), *L. marina* DSM 17023^T (PPCN00000000), and *Labrenzia* sp. strain UBA4493 (DGNL00000000). For the in silico comparison of the strains, the average nucleotide identity (ANI) matrix of all of the conserved genes in the core genome was computed by the BLAST algorithm using JSpeciesWS [23], and was visualized as heat map. The in-silico DNA–DNA hybridization (isDDH) was performed based on the identities/high-scoring segment pairs (HSP) length formula using the DSMZ genome to the genome distance calculator (GGDC) service tool [12]. The biosynthetic gene clusters (BGCs) for the specialized metabolites were identified using antiSMASH v4 [24], and default parameters and the incorporation of the ClusterFinder algorithm were applied.

Author Contributions: Methodology, J.A.M.; software, J.A.M., M.A., and J.B.; validation, T.F.S.; formal analysis, J.A.M.; investigation, J.A.M.; resources, G.M.K. and T.F.S.; data curation, J.A.M.; writing—original draft preparation, J.A.M. and T.F.S.; writing—review and editing, J.A.M., M.A., A.D.-C., J.B., G.M.K., and T.F.S.; visualization, J.A.M.; supervision, G.M.K. and T.F.S.; project administration, G.M.K. and T.F.S.; funding acquisition, G.M.K. and T.F.S.

Funding: This research was funded by the German Centre for Infection Research (DZIF) through grant TTU09.811, and by the German Federal Ministry of Education and Research (BMBF) through grant 16GW0117K. J.A.M. was funded by the Ministry of Science, Research, and Technology, Iran. The funders had no role in the study design, data collection, and interpretation, or the decision to submit the work for publication.

Acknowledgments: We thank Mila Goralski for technical assistance. We also thank Göttingen Genomics Laboratory (G2L) for the sequencing of *Labrenzia* sp. strain 011.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Amiri Moghaddam, J.; Dávila-Céspedes, A.; Kehraus, S.; Crüsemann, M.; Köse, M.; Müller, C.E.; König, G.M. Cyclopropane-Containing Fatty Acids from the Marine Bacterium *Labrenzia* sp. 011 with Antimicrobial and GPR84 Activity. *Mar. Drugs* **2018**, *16*, 369. [CrossRef] [PubMed]
2. Boettcher, K.J.; Barber, B.J.; Singer, J.T. Additional evidence that juvenile oyster disease is caused by a member of the Roseobacter group and colonization of nonaffected animals by *Stappia stellulata*-like strains. *Appl. Environ. Microbiol.* **2000**, *66*, 3924–3930. [CrossRef] [PubMed]
3. Maloy, A.P.; Ford, S.E.; Karney, R.C.; Boettcher, K.J. Roseovarius crassostreae, the etiological agent of Juvenile Oyster Disease (now to be known as Roseovarius Oyster Disease) in *Crassostrea virginica*. *Aquaculture* **2007**, *269*, 71–83. [CrossRef]
4. Pujalte, M.J.; Carmen Macián, M.; Arahal, D.R.; Garay, E. *Stappia alba* sp. nov., isolated from Mediterranean oysters. *Syst. Appl. Microbiol.* **2005**, *28*, 672–678. [CrossRef]
5. Méndez-Vilas, A. *Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology: Bacteria in Molluscs: Good and Bad Guys*; Formatex Research Center: Badajoz, Spain, 2010; pp. 136–147. ISBN 978-84-614-6194-3.
6. Chen, Y.-H.; Kuo, J.; Sung, P.-J.; Chang, Y.-C.; Lu, M.-C.; Wong, T.-Y.; Liu, J.-K.; Weng, C.-F.; Twan, W.-H.; Kuo, F.-W. Isolation of marine bacteria with antimicrobial activities from cultured and field-collected soft corals. *World J. Microbiol. Biotechnol.* **2012**, *28*, 3269–3279. [CrossRef]
7. Graca, A.P.; Bondoso, J.; Gaspar, H.; Xavier, J.R.; Monteiro, M.C.; de La Cruz, M.; Oves-Costales, D.; Vicente, F.; Lage, O.M. Antimicrobial activity of heterotrophic bacterial communities from the marine sponge *Erylus discophorus* (Astrophorida, Geodiidae). *PLoS ONE* **2013**, *8*, e78992. [CrossRef] [PubMed]
8. Rodrigues, G.N.; Lago-Lestón, A.; Costa, R.; Keller-Costa, T. Draft Genome Sequence of Labrenzia sp. Strain EL143, a Coral-Associated Alphaproteobacterium with Versatile Symbiotic Living Capability and Strong Halogen Degradation Potential. *Genome Announc.* **2018**, *6*, e00132-18. [CrossRef]
9. Novak, H.R.; Sayer, C.; Isupov, M.N.; Paszkiewicz, K.; Gotz, D.; Spragg, A.M.; Littlechild, J.A. Marine Rhodobacteraceae L-haloacid dehalogenase contains a novel His/Glu dyad that could activate the catalytic water. *FEBS J.* **2013**, *280*, 1664–1680. [CrossRef]
10. Tettelin, H.; Masignani, V.; Cieslewicz, M.J.; Donati, C.; Medini, D.; Ward, N.L.; Angiuoli, S.V.; Crabtree, J.; Jones, A.L.; Durkin, A.S.; et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial "pan-genome". *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 13950–13955. [CrossRef]
11. Goris, J.; Konstantinidis, K.T.; Klappenbach, J.A.; Coenye, T.; Vandamme, P.; Tiedje, J.M. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* **2007**, *57*, 81–91. [CrossRef]
12. Meier-Kolthoff, J.P.; Klenk, H.-P.; Göker, M. Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. *Int. J. Syst. Evol. Microbiol.* **2014**, *64*, 352–356. [CrossRef] [PubMed]
13. Colston, S.M.; Fullmer, M.S.; Beka, L.; Lamy, B.; Gogarten, J.P.; Graf, J. Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using Aeromonas as a test case. *mBio* **2014**, *5*, e02136. [CrossRef] [PubMed]
14. Williams, K.P.; Sobral, B.W.; Dickerman, A.W. A robust species tree for the alphaproteobacteria. *J. Bacteriol.* **2007**, *189*, 4578–4586. [CrossRef] [PubMed]

15. Arfken, A.; Song, B.; Bowman, J.S.; Piehler, M. Denitrification potential of the eastern oyster microbiome using a 16S rRNA gene based metabolic inference approach. *PLoS ONE* **2017**, *12*, e0185071. [CrossRef] [PubMed]
16. Dang, H.; Li, T.; Chen, M.; Huang, G. Cross-ocean distribution of Rhodobacterales bacteria as primary surface colonizers in temperate coastal marine waters. *Appl. Environ. Microbiol.* **2008**, *74*, 52–60. [CrossRef] [PubMed]
17. Harms, H.; Poehlein, A.; Thürmer, A.; König, G.M.; Schäberle, T.F. Draft Genome Sequence of Zobellia sp. Strain OII3, Isolated from the Coastal Zone of the Baltic Sea. *Genome Announc.* **2017**, *5*, e00737-17. [CrossRef] [PubMed]
18. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol. A J. Comput. Mol. Cell Biol.* **2012**, *19*, 455–477. [CrossRef]
19. Gurevich, A.; Saveliev, V.; Vyahhi, N.; Tesler, G. QUAST: Quality assessment tool for genome assemblies. *Bioinformatics* **2013**, *29*, 1072–1075. [CrossRef]
20. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [CrossRef]
21. Aziz, R.K.; Bartels, D.; Best, A.A.; DeJongh, M.; Disz, T.; Edwards, R.A.; Formsma, K.; Gerdes, S.; Glass, E.M.; Kubal, M.; et al. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genom.* **2008**, *9*, 75. [CrossRef]
22. Blom, J.; Kreis, J.; Spänig, S.; Juhre, T.; Bertelli, C.; Ernst, C.; Goesmann, A. EDGAR 2.0: An enhanced software platform for comparative gene content analyses. *Nucleic Acids Res.* **2016**, *44*, W22–W28. [CrossRef] [PubMed]
23. Richter, M.; Rosselló-Móra, R.; Oliver Glöckner, F.; Peplies, J. JSpeciesWS: A web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* **2016**, *32*, 929–931. [CrossRef] [PubMed]
24. Blin, K.; Wolf, T.; Chevrette, M.G.; Lu, X.; Schwalen, C.J.; Kautsar, S.A.; Suarez Duran, H.G.; de Los Santos, E.L.C.; Kim, H.U.; Nave, M.; et al. antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **2017**, *45*, W36–W41. [CrossRef] [PubMed]