



# **MOLECULAR ECOLOGY AND GENETIC DIVERSITY OF THE ROSEOBACTER CLADE**

EDITED BY: Rolf Daniel, Meinhard Simon and Bernd Wemheuer  
PUBLISHED IN: *Frontiers in Microbiology*



# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2018 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-538-6

DOI 10.3389/978-2-88945-538-6

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# MOLECULAR ECOLOGY AND GENETIC DIVERSITY OF THE ROSEOBACTER CLADE

Topic Editors:

**Rolf Daniel**, University of Göttingen, Germany

**Meinhard Simon**, University of Oldenburg, Germany

**Bernd Wemheuer**, University of Göttingen, Germany; The University of New South Wales, Australia

Marine bacteria and archaea are key players in the biogeochemical cycling of nitrogen, carbon, and other elements. One important lineage of marine bacteria is the Roseobacter group. Members of this clade are the most abundant bacteria in marine ecosystems constituting up to 25% of the marine bacterioplankton. They have been detected in various marine habitats from coastal regions to deep-sea sediments and from polar regions to tropical latitudes. These bacteria are physiologically and genetically very versatile. Utilization of several organic and inorganic compounds, sulfur oxidation, aerobic anoxygenic photosynthesis, carbon monoxide oxidation, DMSP demethylation, and production of secondary metabolites are some of the important functional traits found in this clade.

Moreover, several isolates are available allowing in-depth analysis of physiological and genetic characteristics. Although the Roseobacter group has been intensively studied in recent years, our understanding of its ecological contributions and the evolutionary processes shaping the genomes of this clade is still rather limited.

**Citation:** Daniel, R., Simon, M., Wemheuer, B., eds. (2018). Molecular Ecology and Genetic Diversity of the Roseobacter Clade. Lausanne: Frontiers Media.  
doi: 10.3389/978-2-88945-538-6

# Table of Contents

- 04 Editorial: Molecular Ecology and Genetic Diversity of the Roseobacter Clade**  
Rolf Daniel, Meinhard Simon and Bernd Wemheuer
- 07 Composition of Total and Cell-Proliferating Bacterioplankton Community in Early Summer in the North Sea – Roseobacters are the Most Active Component**  
Insa Bakenhus, Leon Dlugosch, Sara Billerbeck, Helge-Ansgar Giebel, Felix Milke and Meinhard Simon
- 21 Adaptation of Surface-Associated Bacteria to the Open Ocean: A Genomically Distinct Subpopulation of *Phaeobacter gallaeciensis* Colonizes Pacific Mesozooplankton**  
Heike M. Freese, Anika Methner and Jörg Overmann
- 33 The Biogeographical Distribution of Benthic Roseobacter Group Members Along a Pacific Transect is Structured by Nutrient Availability Within the Sediments and Primary Production in Different Oceanic Provinces**  
Marion Pohlner, Julius Degenhardt, Avril J. E. von Hoyningen-Huene, Bernd Wemheuer, Nora Erlmann, Bernhard Schnetger, Thomas H. Badewien and Bert Engelen
- 47 Complementary Metaproteomic Approaches to Assess the Bacterioplankton Response Toward a Phytoplankton Spring Bloom in the Southern North Sea**  
Lars Wöhlbrand, Bernd Wemheuer, Christoph Feenders, Hanna S. Ruppertsberg, Christina Hinrichs, Bernd Blasius, Rolf Daniel and Ralf Rabus
- 60 Exoproteome Analysis of the Seaweed Pathogen *Nautella Italica* R11 Reveals Temperature-Dependent Regulation of RTX-Like Proteins**  
Melissa Gardiner, Adam M. Bournazos, Claudia Maturana-Martinez, Ling Zhong and Suhelen Egan
- 69 Evolution of Dimethylsulfoniopropionate Metabolism in Marine Phytoplankton and Bacteria**  
Hannah A. Bullock, Haiwei Luo and William B. Whitman
- 86 *FnrL* and Three *Dnr* Regulators are Used for the Metabolic Adaptation to Low Oxygen Tension in *Dinoroseobacter Shiba***  
Matthias Ebert, Sebastian Laaß, Andrea Thürmer, Louisa Roselius, Denitsa Eckweiler, Rolf Daniel, Elisabeth Härtig and Dieter Jahn
- 105 Large Diversity and Original Structures of Acyl-Homoserine Lactones in Strain MOLA 401, a Marine Rhodobacteraceae Bacterium**  
Margot Doberva, Didier Stien, Jonathan Sorres, Nathalie Hue, Sophie Sanchez-Ferandin, Véronique Eparvier, Yoan Ferandin, Philippe Lebaron and Raphaël Lami
- 115 The Composite 259-kb Plasmid of *Martelella Mediterranea* DSM 17316<sup>T</sup>—A Natural Replicon With Functional RepABC Modules From Rhodobacteraceae and Rhizobiaceae**  
Pascal Bartling, Henner Brinkmann, Boyke Bunk, Jörg Overmann, Markus Göker and Jörn Petersen
- 131 Plasmid Transfer in the Ocean – A Case Study From the Roseobacter Group**  
Jörn Petersen and Irene Wagner-Döbler





# Editorial: Molecular Ecology and Genetic Diversity of the *Roseobacter* Clade

Rolf Daniel<sup>1</sup>, Meinhard Simon<sup>2</sup> and Bernd Wemheuer<sup>1,3,4\*</sup>

<sup>1</sup> Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, University of Göttingen, Göttingen, Germany, <sup>2</sup> Biology of Geological Processes, Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany, <sup>3</sup> Centre for Marine Bio-Innovation, The University of New South Wales, Sydney, NSW, Australia, <sup>4</sup> School of Biological Earth and Environmental Sciences, The University of New South Wales, Sydney, NSW, Australia

**Keywords:** *Roseobacter* group, microbial ecology, microbial evolution, molecular ecology, microbial diversity

## Editorial on the Research Topic

### Molecular Ecology and Genetic Diversity of the *Roseobacter* Clade

The *Roseobacter* clade, more recently referred to as *Roseobacter* group, is a paraphyletic group within the *Rhodobacteraceae* (Alphaproteobacteria) (Simon et al., 2017). It is one of the most widely distributed and abundant bacterial groups in the marine ecosystem constituting up to 30% of bacterial communities in pelagic environments. *Roseobacter* group members inhabit a great variety of marine habitats and niches. They exhibit a free-living or surface-associated lifestyle and even occur in oxic and anoxic sediments (Luo and Moran, 2014). They are physiologically and genetically very versatile. Some of the important functional traits found in the *Roseobacter* group are the utilization of various organic and inorganic compounds including the catabolism of dimethylsulfoniopropionate (DMSP), energy acquisition by sulfur oxidation, aerobic anoxygenic photosynthesis and carbon monoxide oxidation, and the production of secondary metabolites (Buchan et al., 2005; Wagner-Döbler and Biebl, 2006; Brinkhoff et al., 2008; Todd et al., 2012).

Although various aspects of the *Roseobacter* group have been studied in recent years (e.g., Luo and Moran, 2014; Wemheuer et al., 2014, 2017; Gram et al., 2015; Voget et al., 2015; Lutz et al., 2016; Zhang et al., 2016), our knowledge about its ecological significance and the evolutionary processes shaping the genomes of this group is still limited. The 10 publications presented in this research topic “Molecular Ecology and Genetic Diversity of the *Roseobacter* Clade” highlight new and interesting findings on the evolution, biodiversity, and functions of the *Roseobacter* group in the marine environment. Contributions include original research, a perspective, and a comprehensive review.

In three contributions, culture-independent approaches are employed to assess the abundance and distribution of *Roseobacter* group members in marine pelagic systems (Bakenhus et al.; Freese et al.) and Pacific sediments (Pohlner et al.). Bakenhus et al. highlight the major role of several pelagic members of the *Roseobacter* group in processing phytoplankton-derived organic matter, although this group constituted only a minor proportion of the total bacterioplankton community. Freese et al. show that a previously unknown, distinct group of *Phaeobacter gallaeciensis* possess a limited number of group-specific genes, which may be relevant for its association with mesozooplankton and for its colonization in marine pelagic systems.

As most studies on the abundance and diversity of the *Roseobacter* group were conducted on pelagic samples (e.g., Giebel et al., 2011; Wemheuer et al., 2015; Billerbeck et al., 2016), the

## OPEN ACCESS

### Edited by:

Alison Buchan,  
University of Tennessee, Knoxville,  
United States

### Reviewed by:

Jose M. Gonzalez,  
Universidad de La Laguna, Spain

### \*Correspondence:

Bernd Wemheuer  
bwemheu@gwdg.de

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 21 March 2018

**Accepted:** 16 May 2018

**Published:** 01 June 2018

### Citation:

Daniel R, Simon M and Wemheuer B  
(2018) Editorial: Molecular Ecology  
and Genetic Diversity of the  
*Roseobacter* Clade.  
Front. Microbiol. 9:1185.  
doi: 10.3389/fmicb.2018.01185

distribution and function of this group in sediments is less understood (but see Kanukollu et al., 2015). In their contribution, Pohlner et al. demonstrate that different oligo- and ultraoligotrophic oceanic provinces in the subtropics and tropics of the Pacific were characterized by specific sediment communities and *Roseobacter* group members, distinct from those of the more productive temperate and subarctic regions. *Roseobacter*-affiliated OTUs were dominated by uncultured members, demonstrating the need to obtain cultured *Roseobacter* representatives from sediments to link community structures to specific metabolic processes at the seafloor.

Aside from community patterns, the functional response of the ambient bacterial community toward a *Phaeocystis globosa* bloom in the southern North Sea was studied using metaproteomic approaches (Wöhlbrand et al.) This study highlights the application of different sample preparation techniques and mass spectrometric methods for a comprehensive characterization of marine bacterioplankton responses to changing environmental conditions. The comprehensive approach verified previous metaproteomic studies of marine bacterioplankton (e.g., Sowell et al., 2011; Teeling et al., 2012; Georges et al., 2014), but also revealed new insights into carbon and nitrogen metabolism.

Gardiner et al. demonstrate for the first time temperature-dependent regulation of the RTX-like proteins in the important seaweed pathogen *Nautella italica* R11 and thus provides the basis for future functional studies on the temperature-dependent manner of secreted proteins and their role in pathogenicity and/or environmental persistence of *N. italica* R11. This is of crucial importance as increasing ocean temperatures associated with climate change are predicted to cause greater host stress and more extensive disease events in macroalgae.

Two studies focused on adaptations to environmental properties of the *Roseobacter* group (Bullock et al.; Ebert et al.) Ebert et al. describe for the first time a regulatory network solely composed of four Crp/Fnr-family regulators responsible for the metabolic adaptation to low oxygen tension observed in the marine bacterium *Dinoroseobacter shibae*. Bullock et al. review the evolution of DMSP metabolism in marine phytoplankton and bacteria, thereby illustrating that the enzymes of DMSP demethylation and cleavage pathways are examples of the various processes of enzyme adaptation and evolution, which occurred within the *Roseobacter* group in the last 250 million years.

*N*-acyl-homoserine lactones (AHLs) constitute the major class of semiochemicals in quorum sensing (QS) systems (Williams, 2007; Pappenfort and Bassler, 2016). Complex mixtures of AHLs have been found for the several members of the *Roseobacter* clade (Wagner-Döbler et al., 2005). In their contribution, Doberva et al. discover an unsuspected capacity of the marine *Rhodobacteraceae* strain MOLA 401 to synthesize 20 different putative AHLs by a combination of biosensor-based screening and liquid chromatography coupled to mass spectrometry and nuclear magnetic resonance. The authors conclude that the higher diversity of signaling molecules, unusual for a single strain, shows new molecular adaptations of QS systems to planktonic life.

Horizontal gene transfer (HGT) is an important driver of bacterial diversification and the evolution of prokaryotic genomes (Polz et al., 2013; Sun et al., 2015). Two articles in this research topic highlight the importance of HGT in the *Roseobacter* group. Bartling et al. identified a *Roseobacter*-specific RepABC-type operon in the draft genome of the marine rhizobium *Marteella mediterranea* DSM 17316<sup>T</sup>, whereas Petersen and Wagner -Döbler provide the first evidence for conjugational plasmid transfer across biogeographical and phylogenetic barriers in the *Rhodobacteraceae*.

In summary, the articles presented in this research topic demonstrate the benefits of using multidisciplinary approaches to analyze and deepen our knowledge of the ecological significance, functions, and the evolutionary processes shaping the genomic basis and responses of the *Roseobacter* group to environmental conditions. Moreover, many challenges and questions were identified that remain to be addressed. We thank all the participating authors for their contributions, which we believe will be the basis for future investigations into the function, evolution, and diversity of the fascinating *Roseobacter* group.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

## FUNDING

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) within the Collaborative Research Center Roseobacter (TRR51).

## REFERENCES

- Billerebeck, S., Wemheuer, B., Voget, S., Poehlein, A., Giebel, H.-A., Brinkhoff, T., et al. (2016). Biogeography and environmental genomics of the *Roseobacter* group affiliated pelagic CHAB-I-5 lineage. *Nat. Microbiol.* 1:16063. doi: 10.1038/nmicrobiol.2016.63
- Brinkhoff, T., Giebel, H. A., and Simon, M. (2008). Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch. Microbiol.* 189, 531–539. doi: 10.1007/s00203-008-0353-y
- Buchan, A., González, J. M., and Moran, M. A. (2005). Overview of the marine roseobacter lineage. *Appl. Environ. Microbiol.* 71, 5665–5677. doi: 10.1128/AEM.71.10.5665-5677.2005
- Georges, A. A., El-Swaiss, H., Craig, S. E., Li, W. K. W., and Walsh, D. A. (2014). Metaproteomic analysis of a winter to spring succession in coastal northwest Atlantic Ocean microbial plankton. *ISME J.* 8, 1301–1313. doi: 10.1038/ismej.2013.234
- Giebel, H. A., Kahlhoefer, D., Lemke, A., Thole, S., Gahl-Janssen, R., Simon, M., et al. (2011). Distribution of *Roseobacter* RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J.* 5, 8–19. doi: 10.1038/ismej.2010.87
- Gram, L., Rasmussen, B. B., Wemheuer, B., Bernbom, N., Ng, Y. Y., Porsby, C. H., et al. (2015). *Phaeobacter inhibens* from the *Roseobacter* clade has an environmental niche as a surface colonizer in harbors. *Syst. Appl. Microbiol.* 38, 483–493. doi: 10.1016/j.syapm.2015.07.006

- Kanukollu, S., Wemheuer, B., Herber, J., Billerbeck, S., Lucas, J., Daniel, R., et al. (2015). Distinct compositions of free-living, particle-associated and benthic communities of the *Roseobacter* group in the North Sea. *FEMS Microbiol. Ecol.* 92:145. doi: 10.1093/femsec/fiv145
- Luo, H., and Moran, M. A. (2014). Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol. Mol. Biol. Rev.* 78, 573–587. doi: 10.1128/MMBR.00020-14
- Lutz, C., Thomas, T., Steinberg, P., Kjelleberg, S., and Egan, S. (2016). Effect of interspecific competition on trait variation in *Phaeobacter* inhibens biofilms. *Environ. Microbiol.* 18, 1635–1645. doi: 10.1111/1462-2920.13253
- Papenfort, K., and Bassler, B. L. (2016). Quorum sensing signal-response systems in Gram-negative bacteria. *Nat. Rev. Microbiol.* 14, 576–588. doi: 10.1038/nrmicro.2016.89
- Polz, M. F., Alm, E. J., and Hanage, W. P. (2013). Horizontal gene transfer and the evolution of bacterial and archaeal population structure. *Trends Genet.* 29, 170–175. doi: 10.1016/j.tig.2012.12.006
- Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., et al. (2017). Phylogenomics of *Rhodobacteraceae* reveals evolutionary adaptation to marine and non-marine habitats. *ISME J.* 11, 1483–1499. doi: 10.1038/ismej.2016.198
- Sowell, S. M., Abraham, P. E., Shah, M., Verberkmoes, N. C., Smith, D. P., Barofsky, D. F., et al. (2011). Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *ISME J.* 5, 856–865. doi: 10.1038/ismej.2010.168
- Sun, B.-F., Li, T., Xiao, J.-H., Jia, L.-Y., Liu, L., Zhang, P., et al. (2015). Horizontal functional gene transfer from bacteria to fishes. *Sci. Rep.* 5:18676. doi: 10.1038/srep18676
- Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336, 608–611. doi: 10.1126/science.1218344
- Todd, J. D., Kirkwood, M., Newton-Payne, S., and Johnston, A. W. B. (2012). DddW, a third DMSP lyase in a model *Roseobacter* marine bacterium, Ruegeria pomeroyi DSS-3. *ISME J.* 6, 223–226. doi: 10.1038/ismej.2011.79
- Voget, S., Wemheuer, B., Brinkhoff, T., Vollmers, J., Dietrich, S., Giebel, H.-A., et al. (2015). Adaptation of an abundant *Roseobacter* RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. *ISME J.* 9, 371–384. doi: 10.1038/ismej.2014.134
- Wagner-Döbler, I., and Biebl, H. (2006). Environmental biology of the marine *Roseobacter* lineage. *Annu. Rev. Microbiol.* 60, 255–280. doi: 10.1146/annurev.micro.60.080805.142115
- Wagner-Döbler, I., Thiel, V., Eberl, L., Allgaier, M., Bodor, A., Meyer, S., Ebner, S., et al. (2005). Discovery of complex mixtures of novel long-chain quorum sensing signals in free-living and host-associated marine alphaproteobacteria. *Chembiochem* 6, 2195–2206. doi: 10.1002/cbic.200500189
- Wemheuer, B., Güllert, S., Billerbeck, S., Giebel, H.-A., Voget, S., Simon, M., et al. (2014). Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol. Ecol.* 87, 378–389. doi: 10.1111/1574-6941.12230
- Wemheuer, B., Wemheuer, F., Hollensteiner, J., Meyer, F.-D., Voget, S., and Daniel, R. (2015). The green impact: bacterioplankton response towards a phytoplankton spring bloom in the southern North Sea assessed by comparative metagenomic and metatranscriptomic approaches. *Front. Microbiol.* 6:805. doi: 10.3389/fmicb.2015.00805
- Wemheuer, B., Wemheuer, F., Meier, D., Billerbeck, S., Giebel, H.-A., Simon, M., et al. (2017). Linking compositional and functional predictions to decipher the biogeochemical significance in DFAA turnover of abundant bacterioplankton lineages in the North Sea. *Microorganisms* 5:E68. doi: 10.3390/microorganisms5040068
- Williams, P. (2007). Quorum sensing, communication and cross-kingdom signalling in the bacterial world. *Microbiology* 153, 3923–3938. doi: 10.1099/mic.0.2007/012856-0
- Zhang, Y., Sun, Y., Jiao, N., Stepanauskas, R., and Luo, H. (2016). Ecological genomics of the uncultivated marine *Roseobacter* lineage CHAB-I-5. *Appl. Environ. Microbiol.* 82, 2100–2111. doi: 10.1128/AEM.03678-15

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Daniel, Simon and Wemheuer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Composition of Total and Cell-Proliferating Bacterioplankton Community in Early Summer in the North Sea – Roseobacters Are the Most Active Component

Insa Bakenhus<sup>†</sup>, Leon Dlugosch<sup>†</sup>, Sara Billerbeck, Helge-Ansgar Giebel, Felix Milke and Meinhard Simon<sup>\*</sup>

## OPEN ACCESS

### Edited by:

Marcelino T. Suzuki,  
Sorbonne Universités, Université  
Pierre et Marie Curie (UPMC), CNRS,  
France

### Reviewed by:

Christian Jeannot,  
Station Biologique de Roscoff, France  
Klaus Jürgens,  
Leibniz Institute for Baltic Sea  
Research (LG), Germany

### \*Correspondence:

Meinhard Simon  
m.simon@icbm.de

<sup>†</sup> These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 22 March 2017

**Accepted:** 31 August 2017

**Published:** 13 September 2017

### Citation:

Bakenhus I, Dlugosch L,  
Billerbeck S, Giebel H-A, Milke F and  
Simon M (2017) Composition of Total  
and Cell-Proliferating  
Bacterioplankton Community in Early  
Summer in the North Sea –  
Roseobacters Are the Most Active  
Component. *Front. Microbiol.* 8:1771.  
doi: 10.3389/fmicb.2017.01771

*Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany*

Heterotrophic bacterioplankton communities play an important role in organic matter processing in the oceans worldwide. In order to investigate the significance of distinct phylogenetic bacterial groups it is not only important to assess their quantitative abundance but also their growth dynamics in relation to the entire bacterioplankton. Therefore bacterial abundance, biomass production and the composition of the entire and cell-proliferating bacterioplankton community were assessed in North Sea surface waters between the German Bight and 58°N in early summer by applying catalyzed reporter deposition (CARD-FISH) and bromodeoxyuridine fluorescence *in situ* hybridization (BrdU-FISH). *Bacteroidetes* and the *Roseobacter* group dominated the cell-proliferating fraction with 10–55 and 8–31% of total BrdU-positive cells, respectively. While *Bacteroidetes* also showed high abundances in the total bacterial fraction, roseobacters constituted only 1–9% of all cells. Despite abundances of up to 55% of total bacterial cells, the SAR11 clade constituted <6% of BrdU-positive cells. *Gammaproteobacteria* accounted for 2–16% of the total and 2–13% of the cell-proliferating cells. Within the two most active groups, BrdU-positive cells made up 28% of *Bacteroidetes* as an overall mean and 36% of roseobacters. Estimated mean growth rates of *Bacteroidetes* and the *Roseobacter* group were 1.2 and 1.5 day<sup>-1</sup>, respectively, and much higher than bulk growth rates of the bacterioplankton whereas those of the SAR11 clade and *Gammaproteobacteria* were 0.04 and 0.21 day<sup>-1</sup>, respectively, and much lower than bulk growth rates. Only numbers of total and cell-proliferating roseobacters but not those of *Bacteroidetes* and the other groups were significantly correlated to chlorophyll fluorescence and bacterioplankton biomass production. The *Roseobacter* group, besides *Bacteroidetes*, appeared to be a major player in processing phytoplankton derived organic matter despite its low partitioning in the total bacterioplankton community.

**Keywords:** bacteria, community composition, CARD-FISH, BrdU-FISH, *Roseobacter*, North Sea



## INTRODUCTION

Heterotrophic bacterioplankton communities play an important role in the cycling of carbon, nitrogen and other nutrients in the oceans worldwide. About half of the phytoplankton primary production, supplied in the form of dissolved (DOM) and particulate organic matter (POM), is degraded and mineralized by heterotrophic bacteria (Azam and Malfatti, 2007). The metabolic pathways to break down the complex DOM and POM are diverse and carried out by a multitude of members of the bacterioplankton, distinct in their growth, substrate and environmental requirements (Teeling et al., 2012; Lucas et al., 2015; Sunagawa et al., 2015; Milici et al., 2016; Osterholz et al., 2016). Over the last decade, our insight into the community structure and functional diversity of the bacterioplankton has improved greatly with the establishment of culture-independent methods and in particular the application of fluorescence *in situ* hybridization (FISH) and next generation sequencing technologies. In oceanic environments, *Alpha*- and *Gammaproteobacteria* as well as *Flavobacteria* and *Sphingobacteria* of the *Bacteroidetes* phylum constitute the major fractions of bacterioplankton communities but other phylogenetic lineages, such as *Actinobacteria* and *Planctomycetes*, contribute as well (Schattenhofer et al., 2009; Wietz et al., 2010; Teeling et al., 2012; Buchan et al., 2014; Sunagawa et al., 2015; Milici et al., 2016). Even though the quantitative contribution to the community provides a first view into the significance of a given bacterial taxonomic group, it does not provide clear-cut information on its functional significance and role in organic matter processing. This type of information may be provided by assessing the composition of the metabolically active members, relative to the total bacterioplankton community on the basis of (1) the 16S rRNA and its gene or its expression patterns (Campbell et al., 2011; Gifford et al., 2014), (2) metatranscriptomic analyses (Ottesen et al., 2011; Varaljay et al., 2015), (3) FISH-based activity measurements coupled to microautoradiography (MAR-FISH; Cottrell and Kirchman, 2000; Malmstrom et al., 2007) or finally (4) bromodeoxyuridine (BrdU-FISH; Pernthaler et al., 2002b; Tada et al., 2011). Whereas the former provide detailed information on expression patterns of functional genes with a high phylogenetic resolution only MAR-FISH and BrdU-FISH yield quantitative data on the numeric abundance of metabolically active, i.e., protein synthesizing or DNA proliferating bacterial taxa. Whereas MAR-FISH applies tritiated substrates including leucine and thymidine to trace protein synthesizing and DNA proliferating cells BrdU-FISH applies a thymidine analog for tracing DNA proliferating cells. A compilation of many studies showed that on average 40% of total cells are detected as metabolically active by MAR-FISH (del Giorgio and Gasol, 2008). Similar data are not yet available for BrdU-FISH as only few studies applied this approach (Pernthaler et al., 2002b; Tada et al., 2011, 2013). Proportions of BrdU-active cells in these studies range between 5 and 37%.

In a finer scale the active fraction of a given phylogenetic bacterial group may vary greatly, from <10% to >40% as demonstrated by studies in the Arctic and Atlantic Ocean

(Malmstrom et al., 2007; Alonso-Sáez et al., 2012), the Southern Ocean (Straza et al., 2010; Tada et al., 2013) and the western North Pacific (Tada et al., 2011). *Bacteroidetes* and SAR11, often quantitatively dominating bacterioplankton communities, usually constitute only relatively low fractions of the active cells, whereas the *Roseobacter* group, constituting much lower fractions of bacterioplankton communities, often exhibit high fractions, i.e., >50%, of active cells (Malmstrom et al., 2007; Tada et al., 2011, 2013; Alonso-Sáez et al., 2012). While this observation was in particular oceanic environments, it may suggest more broadly, that the *Roseobacter* group contributes relatively more than the other bacterial groups to organic matter processing but is more susceptible to mortality including grazing and viral lysis.

The North Sea is a coastal sea with pronounced on-off shore gradients of inorganic nutrients, DOM, phytoplankton biomass and bacterioplankton growth and community composition (McQuatters-Gollop et al., 2007; Giebel et al., 2011; Osterholz et al., 2016). Bacterioplankton community dynamics have recently been studied extensively in the German Bight of the North Sea (Teeling et al., 2012; Wemheuer et al., 2014; Lucas et al., 2015; Voget et al., 2015). Despite these extensive studies still little is known about the relative and absolute significance of these phylogenetic groups for bacterial biomass production and cell proliferation. Therefore the aim of this study was to assess cell proliferation of these bacterial groups by BrdU-FISH together with bacterial biomass production and relevant microbial and geochemical variables in early summer between the German Bight and the northern North Sea at 58°N.

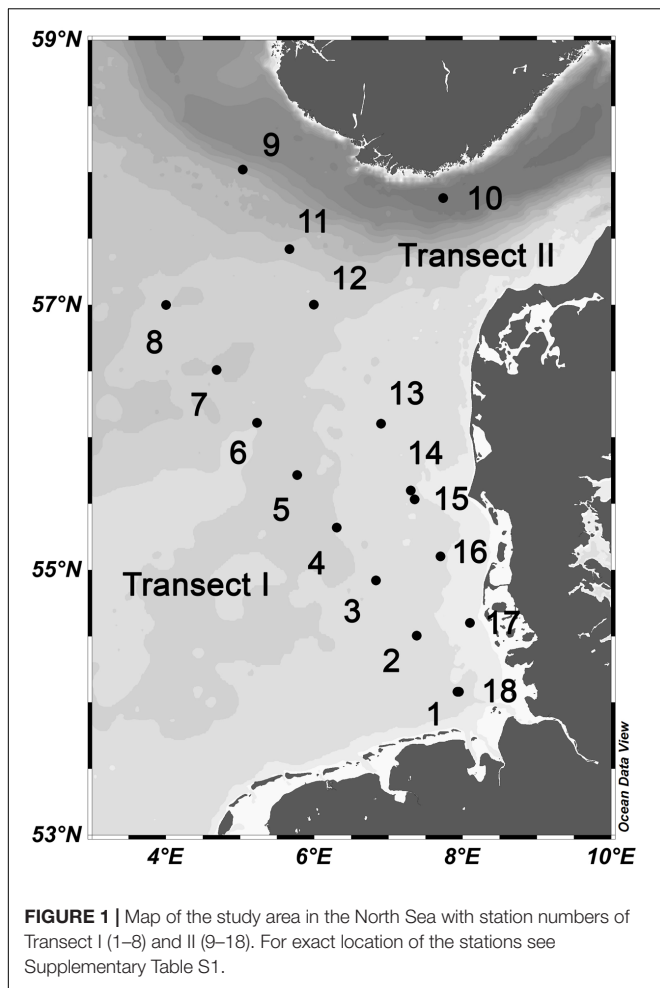
## MATERIALS AND METHODS

### Study Area and Sampling

Surface samples were collected during a cruise with RV Heincke in the North Sea at 3 m depth between the German Bight at ~54°N and 58°N west of Norway from 23 May to 7 June 2014 (**Figure 1** and Supplementary Table S1). On the way north a more westerly transect with eight stations (Transect I) was sampled from 23 to 26 May whereas on the way south a more easterly and coastal transect (Transect II) with nine stations and an additional station in the Norwegian Trench (station 10) was visited between 27 May and 5 June. Depth profiles of conductivity, temperature and chlorophyll *a* (Chl*a*) fluorescence were measured at each station using a CTD sensor (OTS 1500, Meereselektronik, Kiel, Germany) and a Wetlabs ECO FL fluorometer (Philomath, OR, United States). Samples were collected by 5-L Niskin bottles mounted on a rosette (Hydrobios, Kiel, Germany). Subsamples were withdrawn from the bottles immediately after retrieval and further processed for various measurements (see below).

### Biogeochemical and Microbial Variables

Subsamples for the analysis of particulate organic carbon (POC) and nitrogen (PON) were filtered onto precombusted (2 h, 450°C) and preweighed GF/F filters (Whatman). Filters were rinsed with distilled water to remove salt and kept frozen at -20°C until analysis as described in Lunau et al.



(2006). Subsamples for Chl $a$  and phaeopigments were filtered onto GF/F filters (Whatman, 47 mm diameter), immediately wrapped into aluminum foil and kept frozen at  $-20^{\circ}\text{C}$  until further spectrophotometric analysis in the lab. The spectrophotometric analysis was performed as described in Giebel et al. (2011). Numbers of prokaryotic cells, referred as bacterioplankton hereafter, were determined by flow cytometry as described in Osterholz et al. (2015). The cells of the bacterioplankton community were discriminated in high nucleic acid (HNA) and low nucleic acid (LNA) content cells by their distinct fluorescence yield and each subpopulation was delineated via manual gating in a plot of green fluorescence (FL1,  $533 \pm 15 \text{ nm}$ ) vs. red fluorescence (FL3,  $> 670 \text{ nm}$ ).

Rates of bacterioplankton biomass production (BP) were determined by the incorporation of  $^{14}\text{C}$ -leucine. Briefly, triplicate 5-mL subsamples and a formalin-killed control were incubated with  $^{14}\text{C}$ -leucine ( $10.8 \text{ GBq mmol}^{-1}$ , Hartmann Analytic, Germany) at a final concentration of 20 nM in the dark at *in situ* temperature for 1 h and further processed as described (Lunau et al., 2006). Biomass production was calculated using a conversion factor of  $3.05 \text{ kg C (mol leucine)}^{-1}$  according to Simon and Azam (1989).

Bacterioplankton community growth rates ( $\mu$ ;  $\text{day}^{-1}$ ) were calculated as  $\mu = \ln(B1) - \ln(B0)$ , where  $B0$  and  $B1$  ( $B0 + BP$ ) are bacterioplankton biomass at  $T0$  and 1 h later. Bacterioplankton biomass was calculated from bacterial cell numbers, assuming a carbon content of  $20 \times 10^{-15} \text{ g C per cell}$  (Simon and Azam, 1989) and BP is bacterioplankton biomass production as outlined above.

## CARD- and BrdU-FISH

Seawater samples of 55 mL were transferred to dark bottles and incubated with 5-bromo-2'-deoxyuridine (BrdU) (Sigma-Aldrich, Germany; 20  $\mu\text{M}$  final concentration) and thymidine (Sigma-Aldrich, Germany; 33 nM final concentration). A formaldehyde-fixed sample (2% final concentration) served as control. After 4 h incubation at *in situ* temperature samples were fixed with formaldehyde at a final concentration of 2% (v/v) for 1 h at room temperature, filtered onto 0.2  $\mu\text{m}$  polycarbonate filters (Whatman) and stored at  $-20^{\circ}\text{C}$  until further analysis in the lab.

Bacterioplankton community composition was analyzed by catalyzed reporter deposition fluorescence *in situ* hybridization (CARD-FISH) using horseradish peroxidase-labeled oligonucleotide probes specific for *Bacteroidetes*, the *Bacteroidetes* subgroup *Polaribacter*, *Gammaproteobacteria*, the SAR11 clade, the *Roseobacter* group and the *Roseobacter* subgroup RCA (*Roseobacter* clade affiliated) cluster (Table 1). Analysis was carried out according to Pernthaler et al. (2002a), but hybridization and amplification in glass humidity chambers according to Bennke et al. (2013) with a hybridization time of 2 h at  $46^{\circ}\text{C}$ . Hybridization with the probe SAR11-441R was carried out using a 45% [vol/vol] formamide hybridization buffer at  $35^{\circ}\text{C}$  overnight. Oligonucleotide probes as well as unlabeled competitor (Manz et al., 1992) and helper probes (Fuchs et al., 2000) targeting the 16S rRNA of the RCA cluster were newly designed by evaluating probes for the RCA cluster with increased group coverage compared to probe RCA826 (Selje et al., 2004) by using the probe design tool in ARB (Ludwig et al., 2004). Analyses of hybridization efficiency via mathFISH<sup>1</sup> (Supplementary Table S2; Yilmaz et al., 2011) as well as hybridization with RCA isolate RCA23 (*Planktomarina temperata*, GenBank Accession Number GQ369962, Supplementary Figure S1) resulted in best efficiency and specificity for probe RCA996 covering 91% of the sequences of the RCA cluster (Silva SSU Ref NR 128 database, Supplementary Figure S2). For determining optimal stringency conditions of probe RCA996 a series of hybridizations at formamide concentrations from 0 to 70% were carried out against two RCA isolates, RCA23 and LE17 (*Roseobacter* sp., GenBank Accession Number GQ468665). A *Paracoccus* strain (*Paracoccus* sp. GWS-BW-H72M, GenBank Accession Number AY515424) served as negative control. Optimal formamide concentration of 35% was defined as the highest concentration before decreasing signal intensity.

In the results, proportions of the *Roseobacter* group and of *Bacteroidetes* are presented as those of the total minus the

<sup>1</sup><http://mathfish.cee.wisc.edu/index.html>

**TABLE 1** | Probes, their target group, sequence data and formamide (FA) concentration used for this study.

Probe	Target group	Sequence 5'–3'	FA (%)	Reference
CF319a	<i>Bacteroidetes</i>	TGG TCC GTGTCT CAG TAC	35	Manz et al., 1996
POL740	<i>Polaribacter</i>	CCC TCA GCG TCA GTA CAT ACG T	35	Malmstrom et al., 2007
ROS536	<i>Roseobacter</i>	CAA CGC TAA CCC CCT CCG	35	Brinkmeyer et al., 2000
	Competitor	CAA CGC TAG CCC CCT CAG		
RCA996	RCA cluster	TCT CTG GTA GTA GCA CAG GAT	35	This study
	Competitor	CTC TGG GAG TAG CAC AGG		
	Helper978	GTC AAG GGT TGG TAA GGT		
	Helper1017	CCC GAA GGG AAC GTA CCA		
Gam42a	<i>Gammaproteobacteria</i>	GCC TTC CCA CAT CGT TT	35	Manz et al., 1992
	Competitor Bet42a	GCC TTC CCA CTT CGT TT		
SAR11-441R	SAR11-clade	TAC AGT CAT TTT CTT CCC CGA C	45	Morris et al., 2002
NON338	Negative control	ACT CCT ACG GGA GGC AGC	35	Wallner et al., 1993

RCA cluster and minus the *Polaribacter* cluster, respectively. The latter subgroups are given as separate data such that the total of *Roseobacter*+RCA and *Bacteroidetes*+*Polaribacter* represent proportions of the total *Roseobacter* group and the *Bacteroidetes* phylum.

BrdU-FISH was performed according to Pernthaler et al. (2002b) but with the following modifications: The enzymatic permeabilization time was extended to 45 min. The digestion of intercellular DNA during the antibody-reaction was carried out separately by incubation of the filter sections for 30 min at 37°C in digestion buffer (50 mM Tris-HCl, 5 mM MgCl<sub>2</sub>) and using the restriction enzymes ExoIII (53.3 U/mL) and HaeIII (20 U/mL). The antibody reaction was shortened to 2 h and the washing step after the BrdU-amplification carried out in 1x PBS buffer. Proportions of the *Roseobacter*+RCA group and the *Bacteroidetes*+*Polaribacter* group are shown in the same way as the CARD-FISH data (see above).

Microscopic images of the filter sections were acquired semi-automatically using the epifluorescence microscope AxioImager.Z2m including the software package AxioVisionVs 40 V4.8.2.0 (Carl Zeiss, Jena, Germany). Relative abundances of the phylogenetic groups were determined using the automated image analysis software ACMEtool3 (©M. Zeder<sup>2</sup>). Absolute numbers of bacteria detected by CARD-FISH and BrdU-FISH, respectively, were calculated from the relative abundances determined by CARD-FISH and BrdU-FISH, respectively, and the total bacterial numbers obtained by flow cytometry. Negative control counts (hybridization with HRP-Non338) were always below 1% of DAPI-stained cells.

## Statistical Analysis

Pearson's correlation analysis (95% confidence interval) was performed to determine the correlation between the abundance of CARD-FISH targeted and BrdU-positive bacterial groups and environmental, microbial and biogeochemical variables.

<sup>2</sup><http://www.technobiology.ch>

## RESULTS

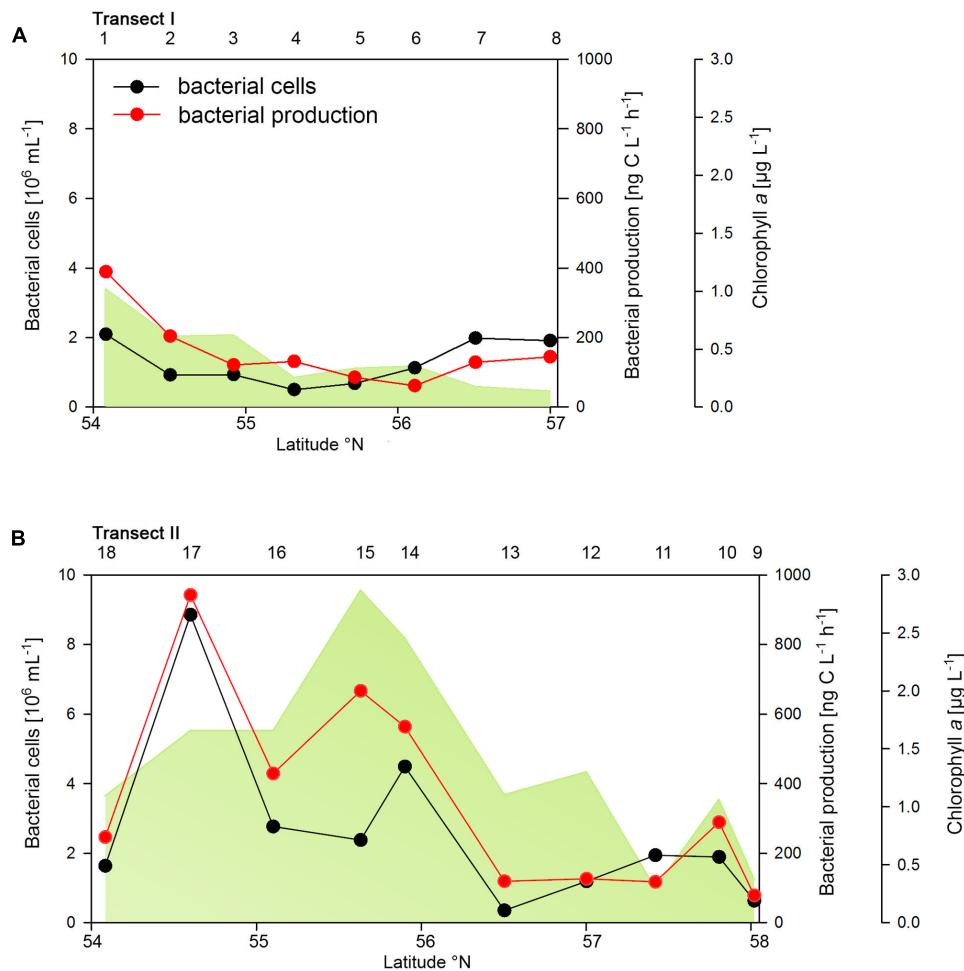
### Environmental and Biogeochemical Characteristics

Water temperature ranged from 11 to 16°C with lower values in the northern North Sea (Supplementary Table S1). Salinity varied between 24.5 and 35.1 with the lowest salinities at stations 9 and 10 close to the Norwegian coast affected by Baltic Sea water (Supplementary Table S1). Salinities in the German Bight ranged between 29.9 and 32.8 and further north and offshore between 33.0 and 35.1.

The two transects differed markedly in their biogeochemical and microbial properties. Transect I, located in the more off shore westerly region, exhibited generally lower concentrations of POC, PON, Chla and bacterioplankton cell numbers and rates of BP than Transect II which stretched over the easterly coastal regions (Figure 2 and Supplementary Table S1). On Transect I Chla concentrations did not exceed 1.0 µg L<sup>-1</sup> and gradually decreased toward the northern region to <0.2 µg Chla L<sup>-1</sup> (Figure 2A). In contrast, on Transect II concentrations exceeded 1.0 µg Chla L<sup>-1</sup> at 8 of the 10 stations with maxima of 2.5 and 2.8 µg Chla L<sup>-1</sup> between 55° and 56°N (Figure 2B). POC and PON also exhibited higher concentrations on Transect II than on Transect I and covaried with Chla (Supplementary Table S1).

### Bacterioplankton Abundance and Growth

Bacterioplankton cell numbers on Transect I ranged between 0.5 × 10<sup>6</sup> and 2.1 × 10<sup>6</sup> mL<sup>-1</sup> with lowest values in the central part of the transect (Figure 2A). At the four southern stations at least 72% of the cells were HNA cells whereas at stations 5 to 7 the fraction of HNA cells was reduced to between 55 and 37% (Figure 3A). On Transect II bacterioplankton abundance varied greatly with highest cell numbers of 8.9 × 10<sup>6</sup> and 4.5 × 10<sup>6</sup> mL<sup>-1</sup> at 54.60° and 55.60°N, respectively, and values of <2.0 × 10<sup>6</sup> mL<sup>-1</sup> at stations further north and at the southernmost station (Figure 2B). South of 56°N HNA cells dominated by 73–75% whereas at 56.50°N and further north HNA constituted only 39–54% (Figure 3B). Bacterioplankton



**FIGURE 2 |** Bacterial cell numbers, biomass production and chlorophyll a concentrations measured spectrophotometrically of Transects I (A) and II (B) during early summer in the North Sea between 54° and 58°N. The numbers above the panels indicate station numbers.

biomass production generally covaried with bacterioplankton cell numbers with rates not exceeding  $400 \text{ ng C L}^{-1} \text{ h}^{-1}$  on Transect I, gradually decreasing toward the northern region, and rates of up to  $942 \text{ ng C L}^{-1} \text{ h}^{-1}$  in the southern part of Transect II (Figure 2). Both variables in the entire data set were highly significantly linearly correlated ( $r^2 = 0.77$ ,  $p < 0.001$ ). Bacterioplankton community growth rates ranged between 0.06 and 0.4 per day (Figure 4). Highest values occurred on Transect I at stations 2 and 4 and on Transect II at stations 13 and 15. Growth rates did not covary with other microbial or biogeochemical variables.

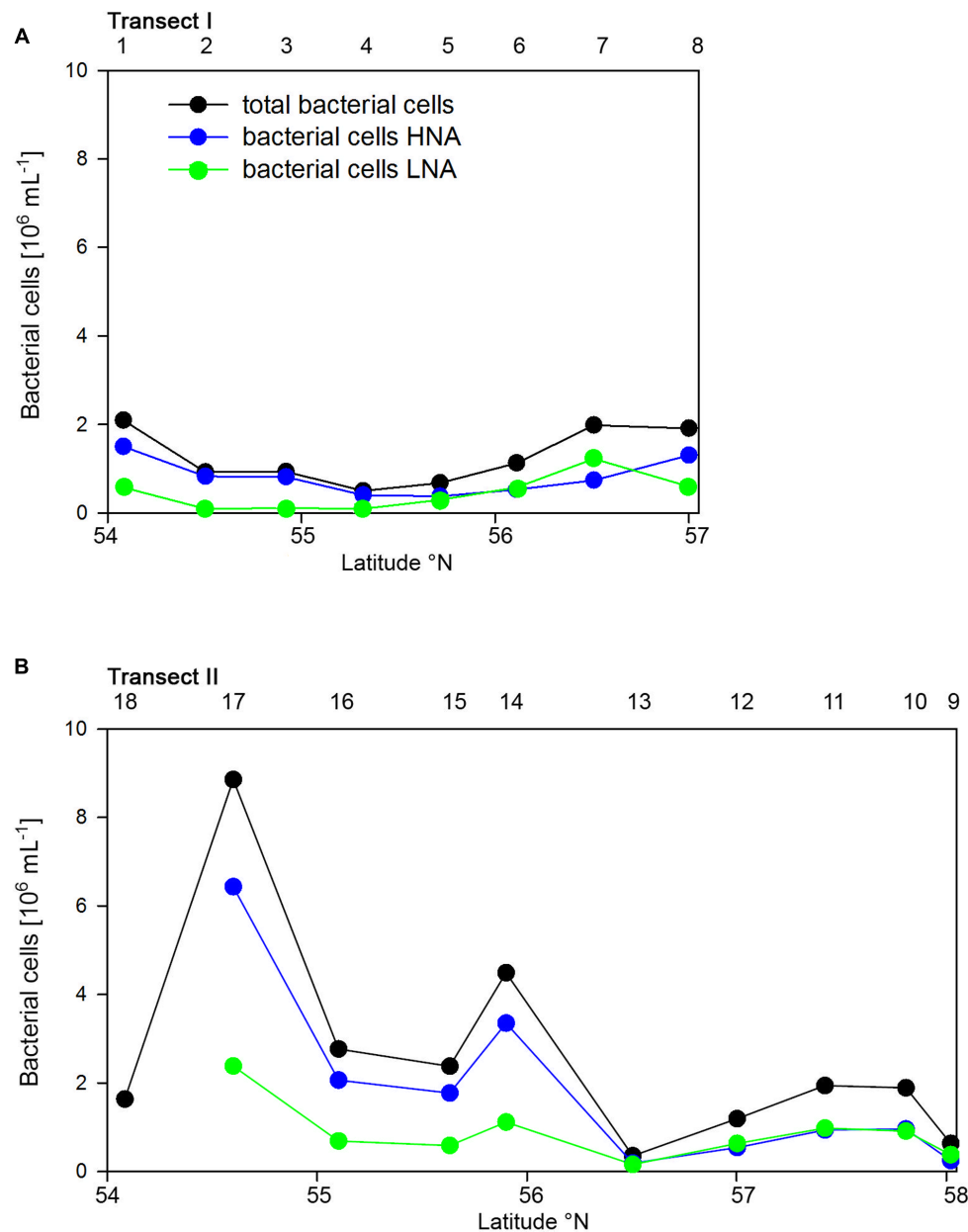
## Bacterioplankton Community Structure

Abundances of the bacterioplankton phylogenetic groups detected by CARD-FISH generally covaried with total bacterioplankton cells as shown by highly significant linear correlations of all phylogenetic groups (mean  $r^2 = 0.73$ , range 0.62–0.90;  $p < 0.001$ ) except the *Polaribacter* cluster (Figure 5). On Transect I 42 to 84% of total cells were identified by the probes applied and on Transect II 37 to 64% (Figures 5A,C).

SAR11 was the most abundant group and dominated the bacterioplankton communities except at three stations on Transects I and II, respectively (Figures 5A,C). This group constituted 4.5 to 43.8% of total bacterioplankton cells on Transect I and 13.4 to 34.8% on Transect II (Figures 5B,D). The second most abundant group was *Bacteroidetes* including the *Polaribacter* cluster, which comprised 5.7 to 36.2% on Transect I and 3.6 to 25.9% on Transect II (Figures 5B,D). Eight to 49% of this phylogenetic group belonged to the *Polaribacter* cluster (Figures 5B,D). *Gammaproteobacteria* constituted 1.9 to 16.4% and 2.5 to 12.8% of total bacterioplankton cells on Transects I and II, respectively (Figures 5B,D) and the *Roseobacter* group 2 to 10.4% and 1.3 to 11.3% on Transects I and II, respectively. The *Roseobacter* RCA cluster accounted for 14 to 80% of this group (Figures 5B,D).

Numbers of DNA-proliferating, i.e., BrdU-positive cells exhibited generally similar patterns as total bacterioplankton cell numbers on both transects (Figures 6A,C). Their proportion ranged from 5 to 22% of total cells with an overall mean of 11%. Forty four to 75% (mean =  $62 \pm 9\%$ ) of total BrdU-positive cells





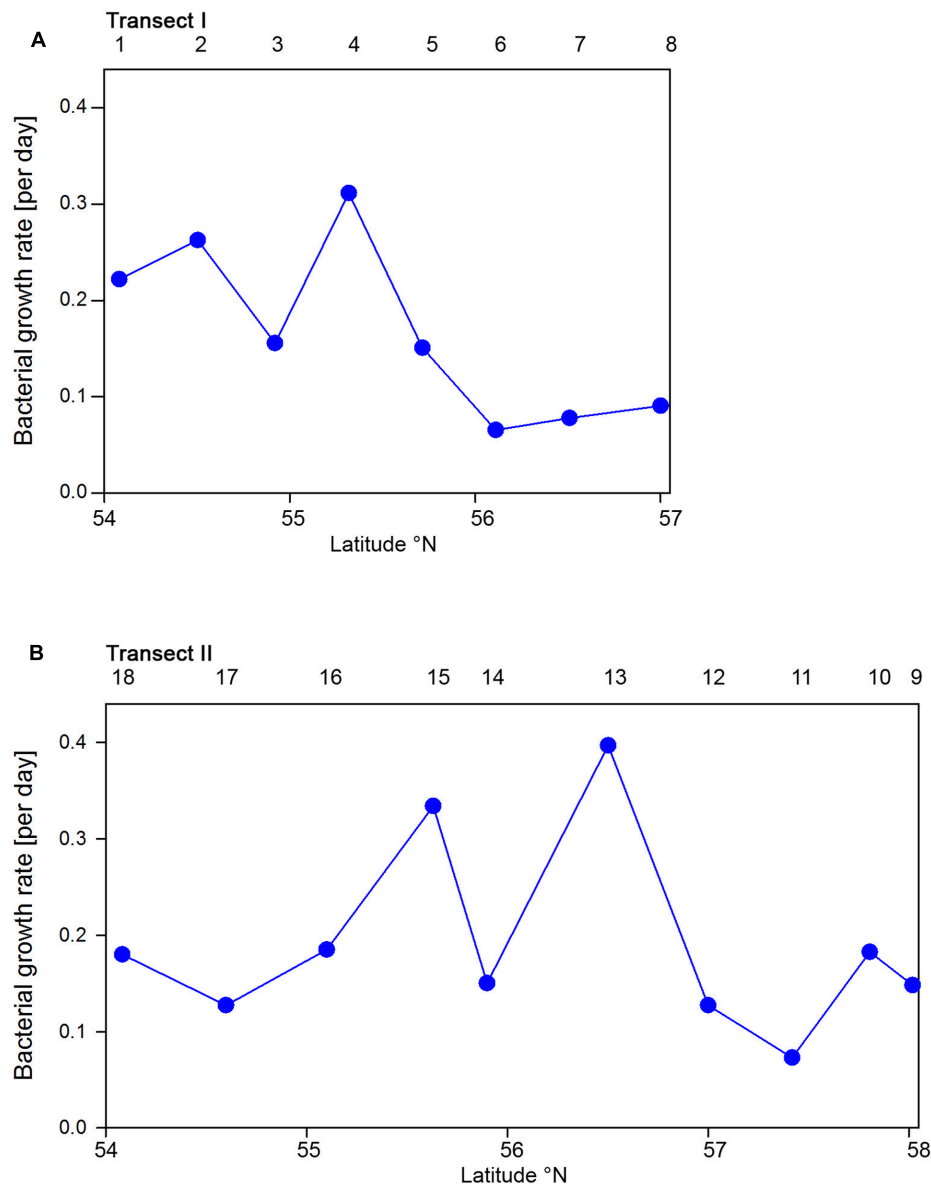
**FIGURE 3 |** Numbers of total, high (HNA) and low nucleic acids (LNA) containing bacterial cell of Transects I (A) and II (B) during early summer in the North Sea between 54° and 58°N. The numbers above the panels indicate station numbers.

could be assigned to cells detected by CARD-FISH and therefore to a bacterial phylogenetic group.

*Bacteroidetes* and the *Roseobacter* group dominated the DNA-proliferating cells at all except two stations and constituted 10–52% (mean =  $39 \pm 10\%$ ) and 8–31% (mean =  $17 \pm 6\%$ ) of total BrdU-positive cells, respectively (Figures 6, 7). The SAR11 clade always contributed <6% and as an overall mean only 1.5% to the DNA-proliferating cells (Figures 6B,D). Also *Gammaproteobacteria* exhibited low proportions of DNA-proliferating cells with an overall mean of 5%. Only at station 14 on Transect II this group constituted a larger fraction

(28%; Figure 6D). The *Roseobacter* group with its RCA cluster and *Bacteroidetes* with its *Polaribacter* cluster were clearly overrepresented among the BrdU-positive cells whereas SAR11 and *Gammaproteobacteria* were greatly underrepresented among those cells (Figure 7).

The proportion of DNA-proliferating cells within each phylogenetic group reflects the relative cell proliferation and thus growth activity of the respective group (Pernthaler et al., 2002b). Generally this feature was in line with the proportion of the given group of total BrdU-positive cells (Figures 7, 8). Within *Bacteroidetes* 28% were BrdU-positive as an overall mean



**FIGURE 4 |** Bacterioplankton bulk growth rates of Transects I (A) and II (B) during early summer in the North Sea between 54° and 58°N. The numbers above the panels indicate station numbers.

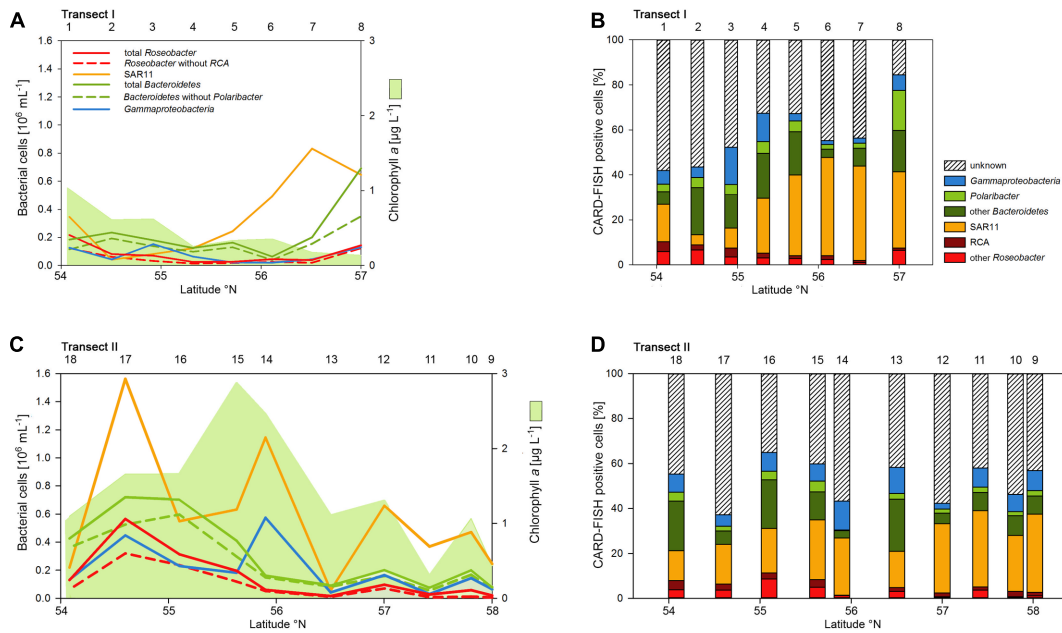
and 21% within its subcluster *Polaribacter* (Figure 8). The *Roseobacter* group harbored even higher proportions of BrdU-positive cells with a mean of 36 and 32% of its RCA cluster (Figure 8). In contrast, within the SAR11 clade <4% and as a mean 1% of the cells were BrdU-positive. Similarly, on average, only 5% of *Gammaproteobacteria* were BrdU-positive (Figure 8).

## Correlation Analysis

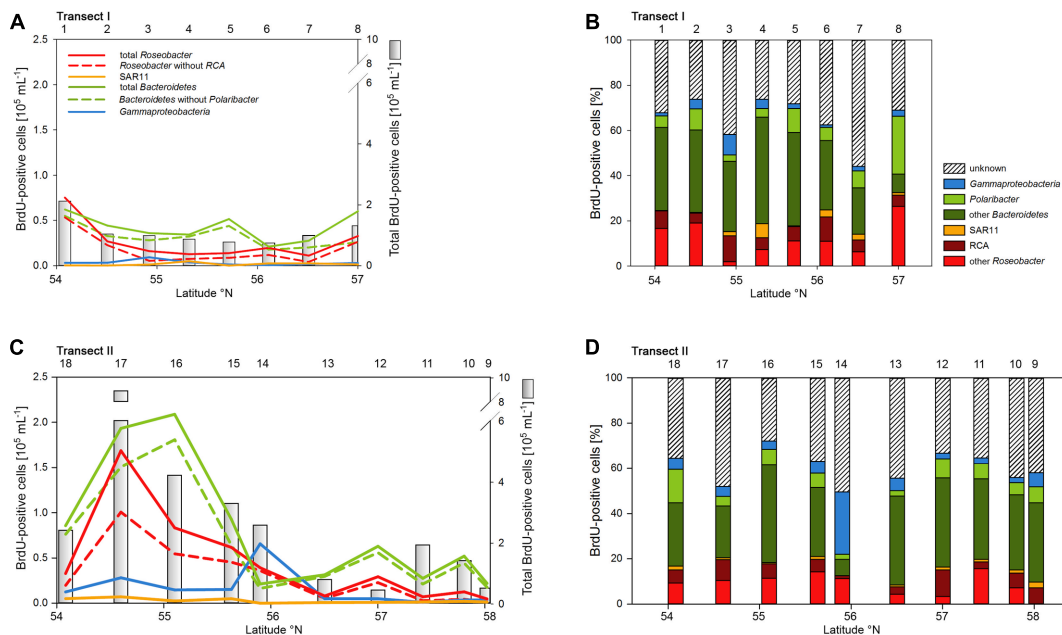
In order to obtain a more refined insight into possible factors controlling the dynamics of the different bacterial groups we carried out a Pearson correlation analysis of their absolute numbers based on the CARD- and BrdU-FISH data vs. the assessed hydrographic, biogeochemical and microbial variables.

Regarding the CARD-FISH data *Gammaproteobacteria* exhibited highly significant ( $p < 0.001$ ) correlations with nine of the 10 variables tested (Figure 9). Correlations with an  $r > 0.7$  existed for phaeopigment concentrations, BP and HNA cells. The *Roseobacter* group and its RCA cluster were highly significantly correlated with concentrations of Chla, BP, HNA, and LNA cells and  $r$  was larger than 0.7 in three of the four correlations. SAR11 was highly significantly correlated to BP and HNA and LNA cells with an  $r = 0.94$  for the latter. *Bacteroidetes* were highly significantly correlated to Chla fluorescence, BP and HNA cells.

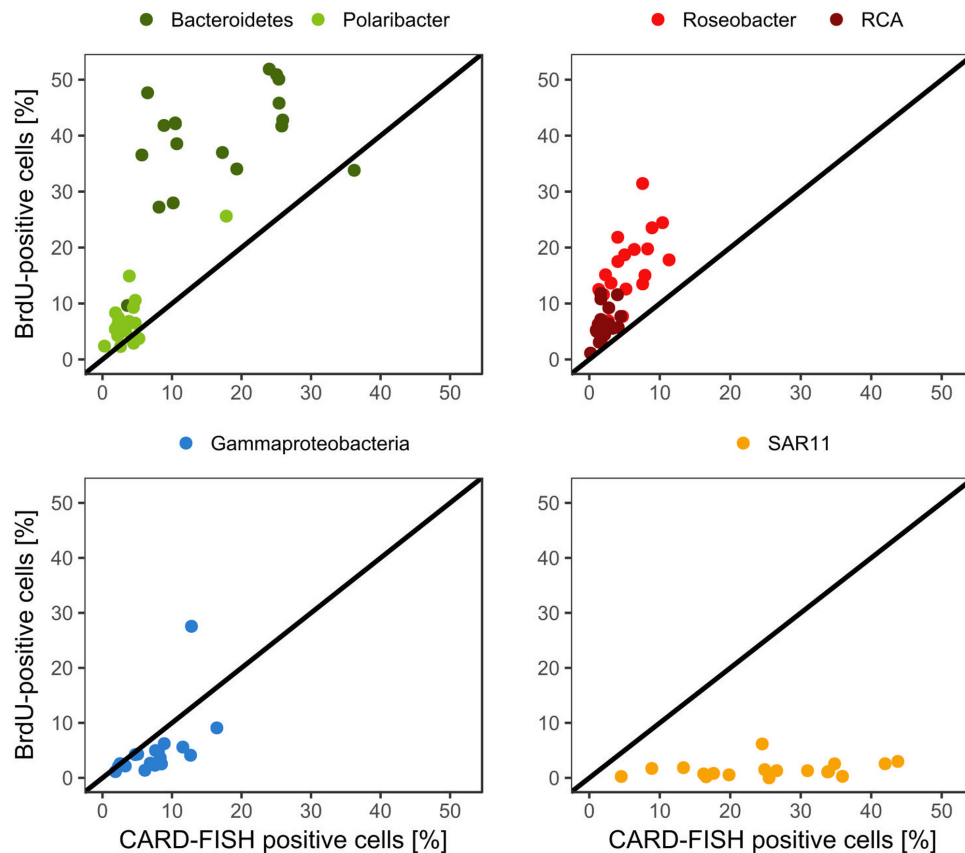
The correlation analysis of the absolute numbers of BrdU-positive cells with the same variables yielded quite different results (Figure 9). *Gammaproteobacteria* were also significantly



**FIGURE 5 |** Cell numbers of the SAR11 clade, the *Roseobacter* group and its RCA cluster, *Bacteroidetes* and its *Polaribacter* cluster and of *Gammaproteobacteria* assessed by CARD-FISH and chlorophyll *a* concentrations measured spectrophotometrically of Transects I (upper panel) and II (lower panel) during early summer in the North Sea between 54° and 58°N. The red bars represent the *Roseobacter* group without the RCA cluster and the dark green bars *Bacteroidetes* without the *Polaribacter* cluster. The numbers above the panels indicate station numbers. (A,C) Show absolute cell numbers and chlorophyll *a* and (B,D) percentages of DAPI-stained cells detected by a group-specific CARD-FISH probe.



**FIGURE 6 |** Cell numbers of the SAR11 clade, the *Roseobacter* group and its RCA cluster, *Bacteroidetes* and its *Polaribacter* cluster and of *Gammaproteobacteria* assessed by BrdU-FISH and numbers of total BrdU-positive cells of Transects I (upper panel) and II (lower panel) during early summer in the North Sea between 54° and 58°N. The red bars represent the *Roseobacter* group without the RCA cluster and the dark green bars *Bacteroidetes* without the *Polaribacter* cluster. The numbers above the panels indicate station numbers. (A,C) Show absolute cell numbers and (B,D) percentages of BrdU-FISH cells detected by a group-specific BrdU-FISH probe.



**FIGURE 7 |** Relative abundances of the *Roseobacter* group and its RCA cluster, *Bacteroidetes* and its *Polaribacter* cluster, the SAR11 clade and of *Gammaproteobacteria* assessed by CARD-FISH and BrdU-FISH during early summer in the North Sea between 54° and 58°N. The red dots represent the *Roseobacter* group without the RCA cluster and the dark green dots *Bacteroidetes* without the *Polaribacter* cluster. The diagonal line indicates equal percentages of the cells assessed by both methods.

( $p < 0.01$ ) correlated to concentrations of Chla and highly significantly ( $p < 0.001$ ) to phaeopigments, POC and PON but not to the other variables of the correlation of the CARD-FISH data. The BrdU-positive cells of the *Roseobacter* group and its RCA cluster exhibited rather similar highly significant correlations as the CARD-FISH data of this group with the strongest correlations to BP and HNA cells. *Bacteroidetes* did not exhibit any significant correlation and SAR11 only one to LNA cells.

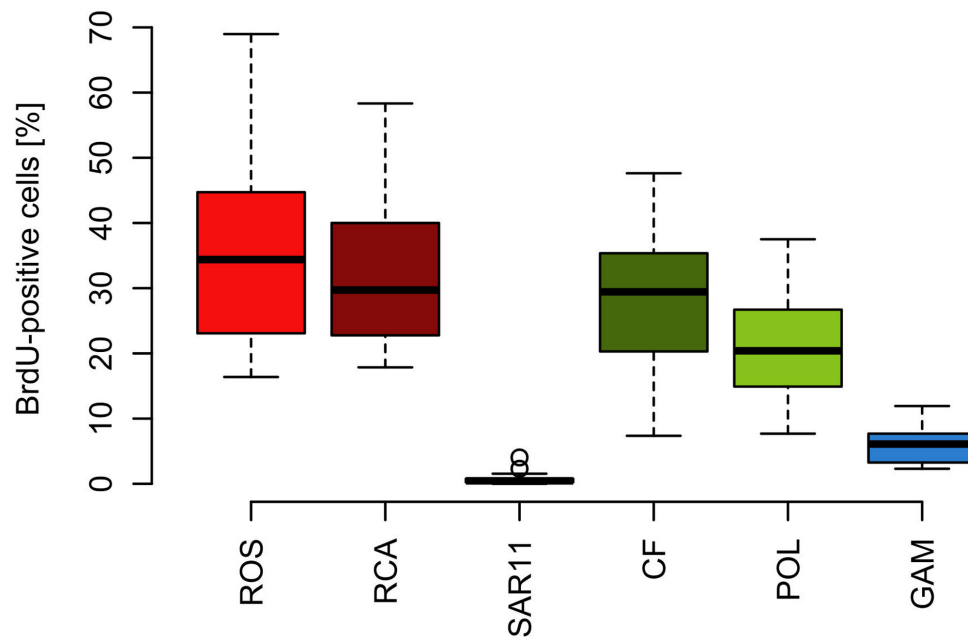
## DISCUSSION

The combined analysis of the bacterioplankton communities by CARD-FISH, BrdU-FISH and by their biomass production on the background of relevant biogeochemical and phytoplankton-related variables allowed us to obtain a rather detailed insight into the growth dynamics of the major bacterioplankton groups and how they were involved in organic matter processing in the North Sea in early summer. On the two transects we encountered different situations which reflected the trophic states of the coastal, nutrient richer, and the off shore, more

oligotrophic region of the North Sea (McQuatters-Gollop et al., 2007). Transect I in the more oligotrophic region exhibited generally lower concentrations of POC, Chla, bacterioplankton numbers and BP than Transect II in the more coastal region with particularly high values in the southern part. As bacterioplankton abundance and biomass production closely covaried and community bulk growth rates did not show pronounced differences between both transects this indicates that the general growth control, mainly by substrate supply and losses by grazing and viral lysis (Longnecker et al., 2010; Winter et al., 2010; Chow et al., 2014), was fairly similar.

The SAR11 clade dominated the bacterioplankton community at the great majority of the stations and *Bacteroidetes* was the second most abundant phylogenetic group, even dominating at a few stations on both transects in the southern part. *Gammaproteobacteria* and the *Roseobacter* group constituted lower proportions than the former two groups. This partitioning of the various phylogenetic groups in structuring the bacterioplankton communities is in line with many previous reports from the North Sea and other marine pelagic systems in the temperate zone (Schattenhofer et al., 2009; Tada et al., 2011, 2013; Teeling et al., 2012; Wemheuer





**FIGURE 8 |** Box-Whisker plots of the percentage of BrdU-positive cells within the *Roseobacter* group and its RCA cluster, the SAR11 clade, *Bacteroidetes* (CF) and its *Polaribacter* cluster and *Gammaproteobacteria* assessed by CARD-FISH during early summer in the North Sea between 54° and 58°N. Given are means, the 25 and 75% percentiles and the error bars of 5 and 95%. The two circles of the SAR11 clade are values above the 95% percentile.

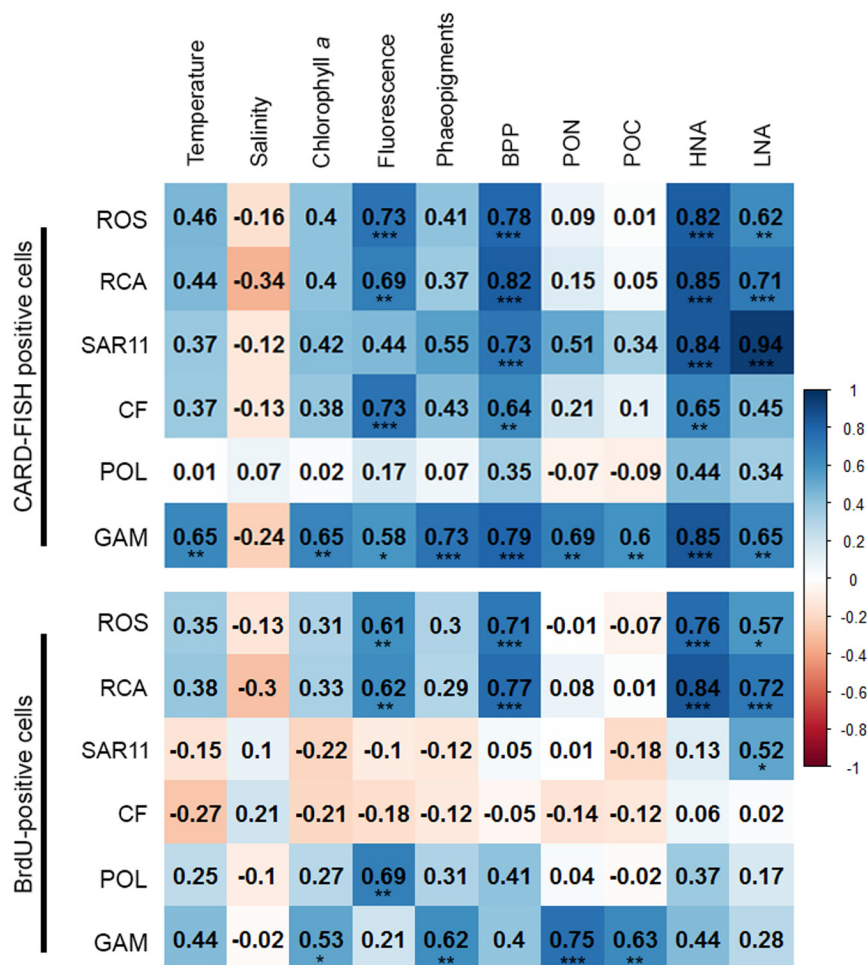
et al., 2014; Lucas et al., 2015). The fact that *Bacteroidetes*, the *Roseobacter* group and *Gammaproteobacteria* exhibited close correlations with Chl *a* fluorescence, BP, HNA and LNA cells suggests that their abundance and growth dynamics were controlled by phytoplankton biomass-related dynamics. As *Gammaproteobacteria* exhibited also close correlations to temperature, PON and POC this phylogenetic subclass appeared to respond to dynamics of non-phytoplankton related biogeochemical variables as well. In contrast, the SAR11 clade was correlated only to BP and HNA and LNA cells, obviously reflecting its general dominance in the bacterioplankton community without a distinct response to phytoplankton-related variables.

Our CARD-FISH analysis detected between 37 and 84% of total bacterial cells and as an overall mean 50% leaving room for undetected phylogenetic groups or bacterial cells. It must be kept in mind that the probes we applied may not have covered all phylogenetic bacterial groups known to be present in the North Sea such as the SAR86 and SAR116 clades (Teeling et al., 2012; Wemheuer et al., 2014; Lucas et al., 2015). On the other hand it may also be a result of methodological constraints of our CARD-FISH analysis by small cells below the cut off of our image analysis system.

The BrdU-positive cells constituted as a mean 11% of total bacterioplankton cells. This percentages is in the same range as reported from a study carried out between subtropical and Antarctic waters (Tada et al., 2013) and somewhat lower than reported from the western North Pacific (Tada et al., 2011). Since during the incubation assay BrdU accumulates in the cells over time the proportion of BrdU-positive cells is a function

of the incubation time and yields information on the relative cell-proliferation rate of the various phylogenetic groups. We chose an incubation time of 4 h whereas in the two studies cited bacterioplankton communities were incubated for 10 h. This implies that in our study the proportions of the BrdU-positive cells presumably were underestimated as compared to the two other studies. A longer incubation time may not have reached proportions of active bacterial cells as high as by MAR-FISH (~40%; del Giorgio and Gasol, 2008) but on average presumably >20%.

One potential constraint of the BrdU method is that not all bacterial taxa may be able to take up this thymidine analog as already stated by Urbach et al. (1999). It has been shown that several bacterial isolates from soil and of marine origin affiliated to *Actinobacteria*, *Bacteroidetes*, *Gamma*- and *Alphaproteobacteria* were unable to take up BrdU and that individual uptake rates during exponential growth varied up to tenfold (Hamasaki et al., 2007; Hellman et al., 2011). On the other hand these and other studies found that the great majority of marine isolates, affiliated to *Actinobacteria*, *Gammaproteobacteria*, *Roseobacter*, and *Bacteroidetes*, is capable of taking up BrdU (Pernthaler et al., 2002b; Hamasaki et al., 2007; Mou et al., 2007). Also the major pelagic marine lineage of *Alphaproteobacteria*, the SAR11 clade, was shown to take up BrdU (Tada et al., 2011, 2013). These authors reported that BrdU-positive SAR11 cells constituted 2 to 33% of total bacterial cells detected by CARD-FISH. Hence, these data suggest, in accordance with a previous statement (Pernthaler et al., 2002b), that BrdU does not show any indication of toxicity on uptake by the great majority of marine bacteria. In the present study



**FIGURE 9 |** Matrix of the Pearson correlation coefficients ( $r$ ) between absolute numbers of the various bacterial groups assessed by CARD-FISH (**upper**) and BrdU-FISH (**lower**) and hydrographic, microbial and biogeochemical variable assessed during early summer in the North Sea between 54° and 58°N. Color code: blue – positive correlations; red – negative correlations. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

we detected the major and well known bacterial phylogenetic groups as BrdU-positive and thus assume that in our samples no group was specifically affected by a potential toxicity of BrdU. The low fractions of the SAR11 clade in BrdU-positive cells, i.e., a relatively low cell proliferation and physiological activity is in line with the low gene expression levels reported for the southern North Sea as compared to *Roseobacter* RCA populations (Voget et al., 2015).

In total, more than 60% of the BrdU-positive cells were covered by the probes we applied. However, the BrdU-FISH analysis showed that the composition of the DNA-proliferating cells was very different than that of the total bacterioplankton community. *Bacteroidetes* and the *Roseobacter* group dominated among DNA-proliferating cells and the SAR11 clade and *Gammaproteobacteria*, except at one station, constituted only minor proportions. These data are in line with previous reports from the western North Pacific, the subtropical Indic and Southern Ocean (Tada et al., 2011, 2013). Interestingly, the phylogenetic groups represented among BrdU-positive cells were

strikingly differently correlated to the set of biological and hydrographic variables. Only the *Roseobacter* group and its RCA cluster exhibited significant and close correlations to BP, HNA, and LNA cells and Chla fluorescence and *Gammaproteobacteria* to phaeopigments, POC and PON. In contrast *Bacteroidetes* did not exhibit any significant correlation to biogeochemical variables and the SAR11 clade only to LNA cells. These results suggest that the growth dynamics of the *Roseobacter* group were much more closely controlled by phytoplankton-related processes than those of the other phylogenetic groups assessed, despite its rather small proportion of the total bacterioplankton community.

Based on the fact that the proportions of cells of a *Roseobacter* and *Alteromonas* isolate increased over time preceding cell proliferation Pernthaler et al. (2002b) suggested that BrdU uptake may be a sensitive measure of the *in situ* growth potential of the targeted phylogenetic groups. Therefore we used the BrdU uptake data to make an estimate of the potential growth rate of the *Roseobacter* group and the other phylogenetic lineages

assessed. The *Roseobacter* group and its RCA cluster harbored the largest fraction of BrdU-positive cells indicating that this group was the most active, i.e., cell-proliferating component of the bacterioplankton community. Assuming a linear accumulation of BrdU by the cells (Pernthaler et al., 2002b) and taking the mean proportion of 36% of BrdU-positive cells within the *Roseobacter* group and of 32% within its RCA cluster suggests a mean growth rate of the *Roseobacter* group and its RCA cluster of 1.5 and 1.3 per day, respectively. A similar calculation yields growth rates of *Gammaproteobacteria* and the SAR11 clade of 0.21 and 0.04 per day, respectively, and of *Bacteroidetes* of 1.2 per day. The growth rates of the *Roseobacter* group and its RCA cluster and of *Bacteroidetes* are far higher than the measured bulk growth rates of the entire bacterioplankton communities, whereas those of *Gammaproteobacteria* were in the range of these bulk growth rates and those of the SAR11 clade far below these rates (Figure 4). The Box-Whisker plot shows that the mean, median and range of the proportion of BrdU-positive cells within *Bacteroidetes* were lower than that of the *Roseobacter* group and its RCA cluster (Figure 8). This indicates that cell proliferation and thus growth was generally higher in the latter group and emphasizes that the *Roseobacter* group was the most active player of the bacterioplankton community in processing phytoplankton-derived organic matter despite its relatively low proportion of the total bacterioplankton community. The RCA cluster constituted at least 15 and in 65% of the samples more than 50% of the cells of the *Roseobacter* group detected by CARD-FISH as well as BrdU. These data further emphasize the great significance of this cluster as an active component of the bacterioplankton and add to previous reports in the North Sea and other pelagic marine systems (West et al., 2008; Giebel et al., 2009, 2011; Teeling et al., 2012; Wemheuer et al., 2014; Voget et al., 2015). The low proportions of the *Roseobacter* group in the CARD-FISH data may be a result of a high top-down control by grazing and viral lysis, the major mortality factors of pelagic bacteria, of the rapidly growing cells of this group (Longnecker et al., 2010; Winter et al., 2010; Chow et al., 2014).

*Bacteroidetes* and its *Polaribacter* cluster harbored almost as many BrdU-positive cells within their groups as the *Roseobacter* group but constituted a much higher fraction of total CARD-FISH and BrdU-positive cells. However, the BrdU-positive cells of this group did not exhibit a positive correlation to fluorescence or concentrations of Chl *a* or BP and only its *Polaribacter* cluster to Chl *a* concentrations. The proportions of BrdU-positive cells of *Bacteroidetes* within this group and in particular of the *Polaribacter* cluster were lower than those in the *Roseobacter* group. Estimated growth rates of *Bacteroidetes* were also slightly lower than those of the *Roseobacter* group. These findings suggest that *Bacteroidetes*, represented presumably mainly by its *Polaribacter* cluster and other *Flavobacteriaceae* (Teeling et al., 2012), were generally intensely involved in organic matter processing, but not specifically in relation to the phytoplankton blooms we encountered in early summer. Nonetheless, our findings corroborate generally the notion that *Bacteroidetes* and the *Roseobacter* group are the key phylogenetic groups involved in organic matter processing during phytoplankton blooms in pelagic marine systems in temperate to polar regions (Buchan

et al., 2014). The *Roseobacter* group, despite its lower share of the total bacterioplankton community, appears to be the more active component and prone to higher mortality losses than *Bacteroidetes*.

As a conclusion the application of CARD-FISH, BrdU-FISH and assessing BP provided a detailed insight into the partitioning of the various bacterial phylogenetic groups in the total and cell-proliferating bacterioplankton communities and allowed a refined insight into the growth dynamics of distinct phylogenetic groups in the North Sea in early summer. More than a decade ago this approach has been first tested in the North Sea (Pernthaler et al., 2002b) but so far not applied to a detailed study in this dynamic coastal sea. SAR11 dominated the total but constituted only minor proportions of the cell-proliferating bacterioplankton and exhibited low growth rates. In contrast the *Roseobacter* group constituted only minor proportions of the total but, together with *Bacteroidetes*, constituted a major proportion of the cell-proliferating bacterioplankton, exhibited the highest growth rates and the closest correlation to Chl *a* fluorescence and BP. Hence, this bacterial group appeared to be a major player in processing phytoplankton derived organic matter despite its low partitioning in the total bacterioplankton community. It seemed to be controlled top down by grazing and viral lysis but the significance of these controlling factors still needs to be shown.

## AUTHOR CONTRIBUTIONS

SB, IB, and FM carried out the field work during the cruise. IB designed the RCA probe. IB and LD performed the CARD-FISH and BrdU-FISH analyses. H-AG carried out the flow cytometric analyses of the bacterial community, FM analyzed the biogeochemical data. LD carried out the statistical data evaluation. IB, LD, and MS wrote the manuscript and all authors commented on the manuscript. MS designed the cruise as the responsible PI.

## FUNDING

This work was supported by Deutsche Forschungsgemeinschaft within the Collaborative Research Center *Roseobacter* (TRR51).

## ACKNOWLEDGMENTS

This work was carried out with RV Heincke as cruise HE-425 (AWI-HE425\_00). We thank the crew of the ship for their support during the cruise, Christine Beardsley for advice in the BrdU-FISH method and two reviewers for constructive suggestions on an earlier version of this publication.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01771/full#supplementary-material>

## REFERENCES

- Alonso-Sáez, L., Sánchez, O., and Gasol, J. M. (2012). Bacterial uptake of low molecular weight organics in the subtropical Atlantic: are major phylogenetic groups functionally different? *Limnol. Oceanogr.* 57, 798–808. doi: 10.4319/lo.2012.57.3.079E
- Azam, F., and Malfatti, F. (2007). Microbial structuring of marine ecosystems. *Nat. Rev. Microbiol.* 5, 782–791. doi: 10.1038/nrmicro1747
- Bennke, C. M., Neu, T. R., Fuchs, B. M., and Amann, R. (2013). Mapping glycoconjugate-mediated interactions of marine Bacteroidetes with diatoms. *Syst. Appl. Microbiol.* 36, 417–425. doi: 10.1016/j.syapm.2013.05.002
- Brinkmeyer, R., Rappé, M., Gallacher, S., and Medlin, L. (2000). Development of clade (*Roseobacter* and *Alteromonas*)- and taxon-specific oligonucleotide probes to study interactions between toxic dinoflagellates and their associated bacteria. *Eur. J. Phycol.* 35, 315–329. doi: 10.1017/S096702620000295X
- Buchan, A., LeCleir, G. R., Gulvik, C. A., and Gonzalez, J. M. (2014). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat. Rev. Microbiol.* 12, 686–698. doi: 10.1038/nrmicro3326
- Campbell, B. J., Yu, L., Heidelberg, J. F., and Kirchman, D. L. (2011). Activity of abundant and rare bacteria in a coastal ocean. *Proc. Natl. Acad. Sci. U.S.A.* 108, 12776–12781. doi: 10.1073/pnas.1101405108
- Chow, C.-E. T., Kim, D. Y., Sachdeva, R., Caron, D. A., and Fuhrman, J. A. (2014). Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J.* 8, 816–829. doi: 10.1038/ismej.2013.199
- Cottrell, M. T., and Kirchman, D. L. (2000). Natural assemblages of marine proteobacteria and members of the *Cytophaga-Flavobacter* cluster consuming low- and high-molecular-weight dissolved organic matter. *Appl. Environ. Microbiol.* 66, 1692–1697. doi: 10.1128/AEM.66.4.1692-1697.2000
- del Giorgio, P. A., and Gasol, J. M. (2008). “Physiological structure and single-cell activity in marian bacterioplankton,” in *Microbial Ecology of the Oceans*, 2nd Edn, ed. D. L. Kirchman (Hoboken NJ: Wiley), 243–298.
- Fuchs, B. M., Glöckner, F. O., Wulf, J., and Amann, R. (2000). Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes. *Appl. Environ. Microbiol.* 66, 3603–3607. doi: 10.1128/aem.66.8.3603-3607.2000
- Giebel, H.-A., Brinkhoff, T., Zwisler, W., Selje, N., and Simon, M. (2009). Distribution of *Roseobacter* RCA and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environ. Microbiol.* 11, 2164–2178. doi: 10.1111/j.1462-2920.2009.01942.x
- Giebel, H.-A., Kalhoefer, D., Lemke, A., Thole, S., Gahl-Janssen, R., Simon, M., et al. (2011). Distribution of *Roseobacter* RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J.* 5, 8–19. doi: 10.1038/ismej.2010.87
- Gifford, S. M., Sharma, S., and Moran, M. A. (2014). Linking activity and function to ecosystem dynamics in a coastal bacterioplankton community. *Front. Microbiol.* 5:185. doi: 10.3389/fmicb.2014.00185
- Hamasaki, K., Taniguchi, A., Tada, Y., Long, R. A., and Azam, F. (2007). Actively growing bacteria in the inland sea of Japan, identified by combined bromodeoxyuridine immunocapture and denaturing gradient gel electrophoresis. *Appl. Environ. Microbiol.* 73, 2787–2798. doi: 10.1128/AEM.02111-06
- Hellman, M., Berg, J., Brandt, K. K., and Hallin, S. (2011). Survey of bromodeoxyuridine uptake among environmental bacteria and variation in uptake rates in a taxonomically diverse set of bacterial isolates. *J. Microbiol. Methods* 86, 376–378. doi: 10.1016/j.mimet.2011.05.020
- Longnecker, K., Wilson, M. J., Sherr, E. B., and Sherr, B. F. (2010). Effect of top-down control on cell-specific activity and diversity of active marine bacterioplankton. *Aquat. Microb. Ecol.* 58, 153–165. doi: 10.3354/ame01366
- Lucas, J., Wichels, A., Teeling, H., Chafee, M., Scharfe, M., and Gerdt, G. (2015). Annual dynamics of North Sea bacterioplankton: seasonal variability superimposes short-term variation. *FEMS Microbiol. Ecol.* 91:fiv099. doi: 10.1093/femsec/fiv099
- Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371. doi: 10.1093/nar/gkh293
- Lunau, M., Lemke, A., Dellwig, O., and Simon, M. (2006). Physical and biogeochemical controls of microaggregate dynamics in a tidally affected coastal ecosystem. *Limnol. Oceanogr.* 51, 847–859. doi: 10.4319/lo.2006.51.2.0847
- Malmstrom, R. R., Straza, T. R. A., Cottrell, M. T., and Kirchman, D. L. (2007). Diversity, abundance, and biomass production of bacterial groups in the western Arctic Ocean. *Aquat. Microb. Ecol.* 47, 45–55. doi: 10.3354/ame047045
- Manz, W., Amann, R., Ludwig, W., Vancanneyt, M., and Schleifer, K.-H. (1996). Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. *Microbiology* 142, 1097–1106. doi: 10.1099/13500872-142-5-1097
- Manz, W., Amann, R., Ludwig, W., Wagner, M., and Schleifer, K.-H. (1992). Phylogenetic oligodeoxynucleotide probes for the major subclasses of *Proteobacteria*: problems and solutions. *Syst. Appl. Microbiol.* 15, 593–600. doi: 10.1016/S0723-2020(11)80121-9
- McQuatters-Gollop, A., Raitos, D. E., Edwards, M., Pradhan, Y., Mee, L. D., Lavender, S. J., et al. (2007). A long-term chlorophyll dataset reveals regime shift in North Sea phytoplankton biomass unconnected to nutrient levels. *Limnol. Oceanogr.* 52, 635–648. doi: 10.4319/lo.2007.52.2.0635
- Milici, M., Tomasch, J., Wos-Oxley, M. L., Wang, H., Jáuregui, R., Camarinha-Silva, A., et al. (2016). Low diversity of planktonic bacteria in the tropical ocean. *Sci. Rep.* 6:19054. doi: 10.1038/srep19054
- Morris, R. M., Rappé, M. S., Connon, S. A., Vergin, K. L., Siebold, W. A., Carlson, C. A., et al. (2002). SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* 420, 806–810. doi: 10.1038/nature01240
- Mou, X., Hodson, R. E., and Moran, M. A. (2007). Bacterioplankton assemblages transforming dissolved organic compounds in coastal seawater. *Environ. Microbiol.* 9, 2025–2037. doi: 10.1111/j.1462-2920.2007.01318.x
- Osterholz, H., Niggemann, J., Giebel, H.-A., Simon, M., and Dittmar, T. (2015). Inefficient microbial production of refractory dissolved organic matter in the ocean. *Nat. Commun.* 6:7422. doi: 10.1038/ncomms8422
- Osterholz, H., Singer, G., Wemheuer, B., Daniel, R., Simon, M., Niggemann, J., et al. (2016). Deciphering associations between dissolved organic molecules and bacterial communities in a pelagic marine system. *ISME J.* 10, 1717–1730. doi: 10.1038/ismej.2015.231
- Ottesen, E. A., Marin, R. III, Preston, C. M., Young, C. R., Ryan, J. P., Scholin, C. A., et al. (2011). Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. *ISME J.* 5, 1881–1895. doi: 10.1038/ismej.2011.70
- Pernthaler, A., Pernthaler, J., and Amann, R. (2002a). Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl. Environ. Microbiol.* 68, 3094–3101. doi: 10.1128/aem.68.6.3094-3101.2002
- Pernthaler, A., Pernthaler, J., Schattnerhofer, M., and Amann, R. (2002b). Identification of DNA-synthesizing bacterial cells in coastal North Sea plankton. *Appl. Environ. Microbiol.* 68, 5728–5736. doi: 10.1128/aem.68.11.5728-5736.2002
- Schattnerhofer, M., Fuchs, B. M., Amann, R., Zubkov, M. V., Tarran, G. A., and Pernthaler, J. (2009). Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ. Microbiol.* 11, 2078–2093. doi: 10.1111/j.1462-2920.2009.01929.x
- Selje, N., Simon, M., and Brinkhoff, T. (2004). A newly discovered *Roseobacter* cluster in temperate and polar oceans. *Nature* 427, 445–448. doi: 10.1038/nature02272
- Simon, M., and Azam, F. (1989). Protein content and protein synthesis rates of planktonic marine bacteria. *Mar. Ecol. Prog. Ser.* 51, 201–203. doi: 10.3354/meps051201
- Straza, T. R. A., Ducklow, H. W., Murray, A. E., and Kirchman, D. L. (2010). Abundance and single-cell activity of bacterial groups in Antarctic coastal waters. *Limnol. Oceanogr.* 55, 2526–2536. doi: 10.4319/lo.2010.55.6.2526
- Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., et al. (2015). Structure and function of the global ocean microbiome. *Science* 348:1261359. doi: 10.1126/science.1261359
- Tada, Y., Makabe, R., Kasamatsu-Takazawa, N., Taniguchi, A., and Hamasaki, K. (2013). Growth and distribution patterns of *Roseobacter*/*Rhodobacter*, SAR11, and Bacteroidetes lineages in the Southern Ocean. *Pol. Biol.* 36, 691–704. doi: 10.1007/s00300-013-1294-8



- Tada, Y., Taniguchi, A., Nagao, I., Miki, T., Uematsu, M., Tsuda, A., et al. (2011). Differing growth responses of major phylogenetic groups of marine bacteria to natural phytoplankton blooms in the western North Pacific Ocean. *Appl. Environ. Microbiol.* 77, 4055–4065. doi: 10.1128/aem.02952-10
- Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336, 608–611. doi: 10.1126/science.1218344
- Urbach, E., Vergin, K. I., and Giovannoni, S. J. (1999). Immunochemical detection and isolation of DNA from metabolically active bacteria. *Appl. Environ. Microbiol.* 65, 1207–1213.
- Varaljay, V. A., Robidart, J., Preston, C. M., Gifford, S. M., Durham, B. P., Burns, A. S., et al. (2015). Single-taxon field measurements of bacterial gene regulation controlling DMSP fate. *ISME J.* 9, 1677–1686. doi: 10.1038/ismej.2015.23
- Voget, S., Wemheuer, B., Brinkhoff, T., Vollmers, J., Dietrich, S., Giebel, H.-A., et al. (2015). Adaptation of an abundant *Roseobacter* RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. *ISME J.* 9, 371–384. doi: 10.1038/ismej.2014.134
- Wallner, G., Amann, R., and Beisker, W. (1993). Optimizing fluorescent in situ hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* 14, 136–143. doi: 10.1002/cyto.990140205
- Wemheuer, B., Güllert, S., Billerbeck, S., Giebel, H.-A., Voget, S., Simon, M., et al. (2014). Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol. Ecol.* 87, 378–389. doi: 10.1111/1574-6941.12230
- West, N. J., Obernosterer, I., Zemb, O., and Lebaron, P. (2008). Major differences of bacterial diversity and activity inside and outside of a natural iron-fertilized phytoplankton bloom in the Southern Ocean. *Environ. Microbiol.* 10, 738–756. doi: 10.1111/j.1462-2920.2007.01497.x
- Wietz, M., Gram, L., Jørgensen, B., and Schramm, A. (2010). Latitudinal patterns in the abundance of major marine bacterioplankton groups. *Aquat. Microb. Ecol.* 61, 179–189. doi: 10.3354/ame01443
- Winter, C., Bouvier, T., Weinbauer, M. G., and Thingstad, T. F. (2010). Trade-offs between competition and defense specialists among unicellular planktonic organisms: the “Killing the Winner” hypothesis revisited. *Microbiol. Mol. Biol. Rev.* 74, 42–57. doi: 10.1128/mmbr.00034-09
- Yilmaz, L. S., Parnerkar, S., and Noguera, D. R. (2011). MathFISH, a web tool that uses thermodynamics-based mathematical models for in silico evaluation of oligonucleotide probes for fluorescence in situ hybridization. *Appl. Environ. Microbiol.* 77, 1118–1122. doi: 10.1128/AEM.01733-10

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Bakenhus, Dlugosch, Billerbeck, Giebel, Milke and Simon. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Adaptation of Surface-Associated Bacteria to the Open Ocean: A Genomically Distinct Subpopulation of *Phaeobacter gallaeciensis* Colonizes Pacific Mesozooplankton

Heike M. Freese\*, Anika Methner and Jörg Overmann

Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany

## OPEN ACCESS

### Edited by:

Bernd Wemheuer,  
University of New South Wales,  
Australia

### Reviewed by:

James T. Hollibaugh,  
University of Georgia, United States  
Julia Grosse,  
GEOMAR Helmholtz Centre for Ocean  
Research Kiel (HZ), Germany

### \*Correspondence:

Heike M. Freese  
heike.freese@dsMZ.de

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 28 April 2017

**Accepted:** 16 August 2017

**Published:** 31 August 2017

### Citation:

Freese HM, Methner A and  
Overmann J (2017) Adaptation  
of Surface-Associated Bacteria to the  
Open Ocean: A Genomically Distinct  
Subpopulation of *Phaeobacter*  
*gallaeciensis* Colonizes Pacific  
Mesozooplankton.  
Front. Microbiol. 8:1659.  
doi: 10.3389/fmicb.2017.01659

The marine *Roseobacter* group encompasses numerous species which occupy a large variety of ecological niches. However, members of the genus *Phaeobacter* are specifically adapted to a surface-associated lifestyle and have so far been found nearly exclusively in disjunct, man-made environments including shellfish and fish aquacultures, as well as harbors. Therefore, the possible natural habitats, dispersal and evolution of *Phaeobacter* spp. have largely remained obscure. Applying a high-throughput cultivation strategy along a longitudinal Pacific transect, the present study revealed for the first time a widespread natural occurrence of *Phaeobacter* in the marine pelagial. These bacteria were found to be specifically associated to mesoplankton where they constitute a small but detectable proportion of the bacterial community. The 16S rRNA gene sequences of 18 isolated strains were identical to that of *Phaeobacter gallaeciensis* DSM26640<sup>T</sup> but sequences of internal transcribed spacer and selected genomes revealed that the strains form a distinct clade within *P. gallaeciensis*. The genomes of the Pacific and the aquaculture strains were highly conserved and had a fraction of the core genome of 89.6%, 80 synteny breakpoints, and differed 2.2% in their nucleotide sequences. Diversification likely occurred through neutral mutations. However, the Pacific strains exclusively contained two active Type I restriction modification systems which is commensurate with a reduced acquisition of mobile elements in the Pacific clade. The Pacific clade of *P. gallaeciensis* also acquired a second, homolog phosphonate transport system compared to all other *P. gallaeciensis*. Our data indicate that a previously unknown, distinct clade of *P. gallaeciensis* acquired a limited number of clade-specific genes that were relevant for its association with mesozooplankton and for colonization of the marine pelagial. The divergence of the Pacific clade most likely was driven by the adaptation to this novel ecological niche rather than by geographic isolation.

**Keywords:** *Phaeobacter gallaeciensis*, zooplankton, attached bacteria, bacterial adaptation, genome evolution, high-throughput cultivation

## INTRODUCTION

Bacteria of the genus *Phaeobacter* belong to the widespread and often abundant marine *Roseobacter* group (Buchan et al., 2005; Wagner-Döbler and Biebl, 2006; Simon et al., 2017). They are adapted to a surface associated lifestyle and are characterized by a high versatility of metabolic pathways (Martens et al., 2006; Dickschat et al., 2010; Newton et al., 2010; Thole et al., 2012). The cells attach tightly to abiotic and biotic surfaces, outcompeting other bacteria and can even invade established epiphytic communities (Rao et al., 2005, 2010; Frank et al., 2015). Colonization is accompanied by the production of the antibiotic tropodithietic acid (TDA) which inhibits a variety of bacteria including pathogenic vibrios (Brinkhoff et al., 2004; Porsby et al., 2008; Prado et al., 2009). The inhibitory activity may also involve antibacterial peptides encoded by hybrid non-ribosomal peptide synthetase/polyketide synthases (Ruiz-Ponte et al., 1999; Martens et al., 2007). Due to their antibiotic activity, members of the genus *Phaeobacter* exert a beneficial effect on different fish and shellfish and may act as probiotic in aquacultures (D'Alvise et al., 2012; Kesarcodi-Watson et al., 2012; Karim et al., 2013). However, *Phaeobacter* cells associated with senescent algae switch from a mutualistic to a parasitic lifestyle and produce algaecides ("roseobacticide") leading to the lysis of algal cells (Seyedsayamdost et al., 2011b).

Despite their high metabolic versatility and contrary to other representatives of the *Roseobacter* group (Billerbeck et al., 2016; Sonnenschein et al., 2017), *Phaeobacter gallaeciensis*, *P. inhibens*, and *P. porticola* have so far been detected exclusively in different aquacultures (Porsby et al., 2008; Prado et al., 2009; Balcazar et al., 2010; Xue et al., 2016) or on sessile invertebrates and abiotic surface in harbors (Gram et al., 2015; Breider et al., 2017). In addition, only a few solitary isolates from macroalgae, seaweed, and intertidal mudflats at coastal shores of Australia, France and Germany have been reported (Rao et al., 2005; Martens et al., 2006; Penesyan et al., 2009; Doghri et al., 2015).

Two strains of *P. inhibens* which had been isolated from an aquaculture at the Atlantic Coast of north western Spain (strain DSM 17395) and from a macroalga near Sydney, Australia (strain 2.10, DSM 24588), respectively, had identical 16S rRNA sequences, a genome sequence identity of 97%, a very high percentage of shared genes (88–93%), and a high synteny of genomes and plasmids (Thole et al., 2012). Considering that *Phaeobacter* spp. were mostly detected on surfaces in man-made environments, the presence of highly similar genotypes in locations as far as 18,000 km apart suggest that rapid means of dispersal exist for these bacteria. Also, populations of *Phaeobacter* spp. might exist in other marine environments which so far have escaped detection by established molecular or cultivation-based approaches. Accordingly, *Phaeobacter* may also occur associated to phytoplankton or zooplankton in the open ocean. We hypothesize that attachment to and transport by planktonic eukaryotes link the *Phaeobacter* populations in distant coastal environments.

In metagenomics databases from the oceans like the Global Ocean Sampling, *Phaeobacter* spp. are not detectable. Although, short 16S rRNA gene sequence fragments related to *Phaeobacter*

spp. are present in the Tara Ocean 16S rRNA sequence database<sup>1</sup>, they could not be unambiguously classified, as these short read sequences are also identical to several other *Roseobacter* group genera like *Sulfitobacter* and *Ruegeria* which are known to be globally distributed (Sonnenschein et al., 2017). Different *Phaeobacter* species and related genera have very low 16S rRNA gene sequence differences. In combination with their low abundance this currently precludes a reliable detection of these bacteria by culture-independent methods suitable for large sample sets. We therefore conducted a systematic study to detect *Phaeobacter* in the open Pacific Ocean along a longitudinal transect from New Zealand to Alaska employing advanced high-throughput enrichment methods with improved growth media. Subsequent comparative genomics of selected *Phaeobacter* isolates revealed a distinct subpopulation of *P. gallaeciensis* with particular adaptations. These findings contribute to our understanding of the evolution and dispersal of marine surface-associated bacteria.

## MATERIALS AND METHODS

### Sampling and Sample Preparation

Plankton samples were collected during cruise SO248 of the RV Sonne in May 2016 at 12 open water stations in the Pacific along a longitudinal transect between Auckland (New Zealand) and Dutch Harbor (Alaska) (Table 1). For sampling, a 3.2 m long Bongo net with an opening diameter of 61 cm and mesh sizes of 100 or 300  $\mu\text{m}$  was employed. Vertical tows were conducted from 150 m water depth to the surface with a constant velocity of 0.2  $\text{m s}^{-1}$  and horizontal hauls were done at the water surface for 45 min at a speed of 1.5–2 knots. Two ml biovolume of concentrated plankton was homogenized with a glass tissue grinder (Kimble Chase Gerresheimer, United States) and diluted with artificial sea water (ASW; Bruns et al., 2003) supplemented with 10 mM HEPES, pH 7.6. Samples from the supernatant of the plankton sample were used for parallel inoculations to investigate the bacteria which were not firmly attached to the zooplankton. A workflow scheme of the combined methods is depicted in Supplementary Figure S1.

### Enrichment, Isolation, and Screening Strategies

High-throughput liquid serial dilutions were performed on board in 96 deepwell plates (2 ml MASTERBLOCK®, Greiner Bio-One GmbH) with 48–128 inoculated wells per dilution ( $10^{-4}$  to  $10^{-10}$ ) and sample. Two different media were successfully employed to grow *Phaeobacter*. Medium HD contained some complex organic substrates at low concentrations and based on ASW supplemented with 0.6 mM glucose, 0.25  $\text{g l}^{-1}$  yeast extract, 0.5  $\text{g l}^{-1}$  peptone, 20  $\text{mg l}^{-1}$  cycloheximide, 0.2 mM ferric citrate, 1  $\text{ml l}^{-1}$  vitamin solution (Karsten and Drake, 1995), and 1  $\text{ml l}^{-1}$  trace element solution SL 10 (Widdel et al., 1983). Medium AM was more alkaline (pH 9) and based on ASW after Martens et al. (2006) but without silicate and with 10 mM CHES, 0.2 mM

<sup>1</sup><http://ocean-microbiome.embl.de/companion.html>

**TABLE 1** | Characteristics of samples and sampling depths. Temperature and salinity were recorded by a CTD system and taken from Badewien et al. (2016).

Station	Date time UTC	Local time	Latitude	Longitude	Temperature (°C)		Salinity (PSU)		Samples *	Cultivation
					Surface	150 m	Surface	150 m		
SO248_1	02.05.2016 20:51	08:51	29° 59.999' S	177° 0.002' E	23.5	17.9	35.9	35.8	100 and 300 $\mu$ m vertical, supernatant	High-throughput liquid, agar plates
SO248_4	08.05.2016 00:47	13:47	10° 20.001' S	176° 29.933' W	30.1	23.7	34.7	36.2	300 $\mu$ m horizontal	Agar plates
SO248_6	12.05.2016 05:19	18:19	0° 0.005' N	180° 0.000' E	29.5	19.7	35.2	35.3	300 $\mu$ m horizontal	Agar plates
SO248_8	16.05.2016 00:34	13:34	11° 0.008' N	179° 0.331' E	27.8	16.2	34.6	34.7	300 $\mu$ m horizontal	Agar plates
SO248_10	18.05.2016 23:52	13:52	22° 0.047' N	178° 18.985' E	25.2	18.0	35.5	35.0	300 $\mu$ m horizontal	Agar plates
SO248_12	21.05.2016 20:08	10:08	33° 59.912' N	177° 20.018' E	15.8	15.0	34.8	34.7	300 $\mu$ m horizontal, supernatant	High-throughput liquid
SO248_13	23.05.2016 08:18	22:18	39° 58.226' N	177° 19.611' E	11.9	9.3	34.3	34.2	100 $\mu$ m vertical, 300 $\mu$ m horizontal, supernatant	High-throughput liquid
SO248_14	24.05.2016 14:09	04:09	45° 0.002' N	178° 44.990' E	5.8	5.4	33.3	33.5	100 $\mu$ m vertical, supernatant	High-throughput liquid
SO248_16	26.05.2016 23:31	13:31	49° 59.462' N	179° 31.444' E	4.8	3.6	33.1	33.9	300 $\mu$ m vertical, supernatant	High-throughput liquid
SO248_17	29.05.2016 03:30	17:30	54° 0.081' N	179° 34.889' E	4.6	3.3	33	33.5	300 $\mu$ m horizontal, supernatant	High-throughput liquid
SO248_18	30.05.2016 04:50	19:50	56° 59.189' N	179° 37.476' E	4.2	3.8	33.2	33.5	100 $\mu$ m vertical, supernatant	High-throughput liquid
SO248_19	31.05.2016 01:37	16:37	58° 54.472' N	178° 55.749' W	5.4	3.7	32.9	33.2	100 $\mu$ m vertical, supernatant	High-throughput liquid

\* Underlined samples yielded *Phaeobacter*.

ferric citrate, 1 ml l<sup>-1</sup> selenite-tungstate solution (Tschech and Pfennig, 1984), supplemented with 0.5 mM glucose, 0.5 mM D-alanine, and 0.2 mM Tween® 80. To test isolation success on solidified media, plankton homogenate was streaked onto Difco marine broth Bacto agar (BD) in dilutions from 10<sup>-4</sup> to 10<sup>-8</sup> in duplicates. Deepwell plates and agar plates were incubated at 15°C while one parallel of each dilution of solid medium was incubated at *in situ* temperature for at least 10 days to test for temperature effects.

To detect *Phaeobacter* and closely related genera, cultures were screened first on board and later in our home lab via PCR using the specific forward primer PHA-16S-129f (5'-AAC GTG CCC TTC TCT AAG G-3'; Gram et al., 2015) and the universal reverse primer 907r (5'-CCG TCA ATT CMT TTG AGT TT-3'; Lane, 1991) (details see Supplementary Material). Positive cultures were sequenced and identified according to the NCBI type material database (Federhen, 2015). Most probable numbers (MPNs) of *Phaeobacter* were calculated after Jarvis et al. (2010). Strains of *P. gallaeciensis* were isolated from positive enrichments by streaking on solid media.

## Sequencing of the 16S rRNA Gene and the Internal Transcribed Spacer Region

To sequence the complete 16S rRNA gene and the downstream internal transcribed spacer (ITS) region, isolates were grown in marine broth, harvested by centrifugation and DNA was extracted by the Qiagen Blood & Tissue Kit. Amplification products were generated with primers 27f (5'-AGA GTT TGA TCM TGG CTC AG-3'; Lane, 1991) and 23S-130r (5'-GGG TTG CCC CAT TCR G-3'; Fisher and Triplett, 1999) and sequenced by Sanger sequencing. ITS sequences were aligned using ClustalW implemented in MEGA6 (Tamura et al., 2013) and manually curated. The phylogenetic tree was calculated employing the best fit option of the Maximum Likelihood method (Kimura 2-parameter model with gamma distribution, complete gap deletion) in MEGA6.

## Genome Sequencing, Assembly, and Annotation

Based on the results of the phylogenetic analysis, two strains originating from distant sampling locations were chosen for genome analysis. Genomic DNA was extracted with the JETFLX Genomic DNA Purification Kit (Genomed). SMRT sequencing was carried out on the PacBio RSII (Pacific Biosciences, Menlo Park, CA, United States) using the P6 chemistry (details see Supplementary Material). PacBio reads were assembled *de novo* in the SMRT Portal 2.3.0 and were corrected by paired-end Illumina reads, which were sequenced on the MiSeq (PE150), using the Burrows-Wheeler Aligner (Li and Durbin, 2009) and the CLC Genomics Workbench 7.0.1. The final assembly was circularized and adjusted to the replication system as start point<sup>2</sup>. Genome sequences were automatically annotated using Prokka 1.8 (Seemann, 2014) and were deposited in the NCBI GenBank (accession numbers: CP021040–CP021052). Information for

<sup>2</sup><https://github.com/boykebunk/genomefinish>



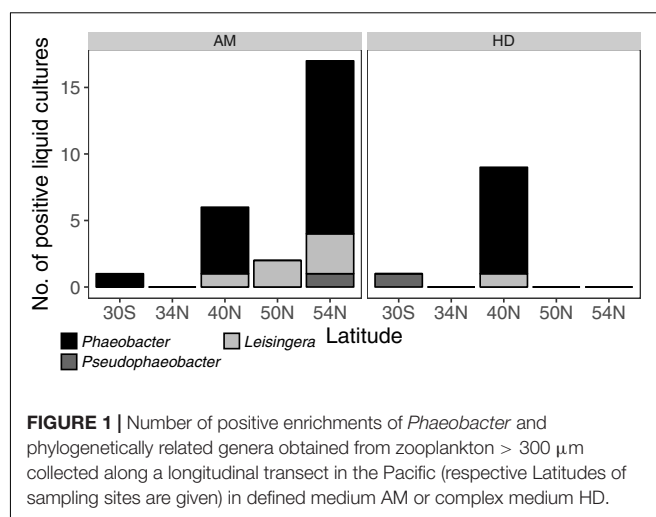
all other *Phaeobacter* strains which were used for genomic comparisons are given in Supplementary Table S1.

## Genome Comparisons

Polymorphic genome sites in the *Phaeobacter* genus were extracted from a core genome alignment using Parsnp and Gingr (Treangen et al., 2014). A phylogenetic network was calculated from the resulting matrix which contained 68,858 characters with the NeighborNet algorithm in SplitsTree 4.13.1 (Huson and Bryant, 2006). Phylogenetic distances of *P. gallaeciensis* were inferred from pairwise comparisons of complete genome sequences via the GGDC 2.1 web service (formula 2; Meier-Kolthoff et al., 2013). A BIONJ tree was calculated with the R package ape and rooted at midpoint (Gascuel, 1997; Paradis et al., 2004). Whole chromosome alignments of *P. gallaeciensis* were obtained with Mauve (Darling et al., 2010), sequence similarity between genomes estimated by BLAST (Altschul et al., 1990) and plotted together with the R package genoPlotR (Guy et al., 2010). Nucleotide diversity between Pacific and the other strains was calculated for all aligned ortholog genes along the genome with the R package PopGenome (Pfeifer et al., 2014). Prophages were identified with PHASTER (Arndt et al., 2016) and classified by Virfam (Lopes et al., 2014). Genomic islands were predicted with the IslandViewer web server (Dhillon et al., 2013). Mobile elements of the other *P. gallaeciensis* had been identified previously (Freese et al., submitted) and were included in the analysis. Orthologs of *P. gallaeciensis* were inferred via proteinortho (Lechner et al., 2011) and the core genome was defined as orthologs present in all strains. Subsequently, clade-specific orthologs were identified and their KEGG orthologies (ko identifiers) were determined using KAAS (Moriya et al., 2007) using a threshold of 30%. Methylation motifs were detected using the SMRT Portal with the RS Modification and Motif Analysis and were identified using REBASE (Roberts et al., 2015). Methylase sequence homologs for the hits were identified using BLAST.

## RESULTS

For the first time, *Phaeobacter* spp. were successfully enriched from the open ocean by employing our high-throughput liquid cultivation. Enrichments were obtained from very distantly located samples taken in the South Pacific (30°S), North Pacific (40°N), and in the Bering Sea (54°N) (Figure 1) where surface water temperatures differed between 23.5 and 4.6°C (Table 1). The observations that representatives of the phylogenetically related genera *Pseudophaeobacter* and *Leisingera* were also enriched and that total MPNs exceeded those of *Phaeobacter* by 1–2 orders of magnitude (Table 2) indicate that both types of liquid media were not highly selective for *Phaeobacter*. Cultivation attempts on marine broth agar failed for samples from the open Pacific (Table 1) even when liquid enrichments from the same station (SO248\_1) yielded *Phaeobacter*. Most notably, *Phaeobacter* were exclusively enriched from size fractions >300 µm which consists mostly of larger zooplankton organisms. In contrast neither the supernatant nor the size



**FIGURE 1 |** Number of positive enrichments of *Phaeobacter* and phylogenetically related genera obtained from zooplankton > 300 µm collected along a longitudinal transect in the Pacific (respective Latitudes of sampling sites are given) in defined medium AM or complex medium HD.

fraction >100 µm yielded enrichments of *Phaeobacter* even at stations SO248\_1 and SO248\_13 where both size fractions were tested in parallel.

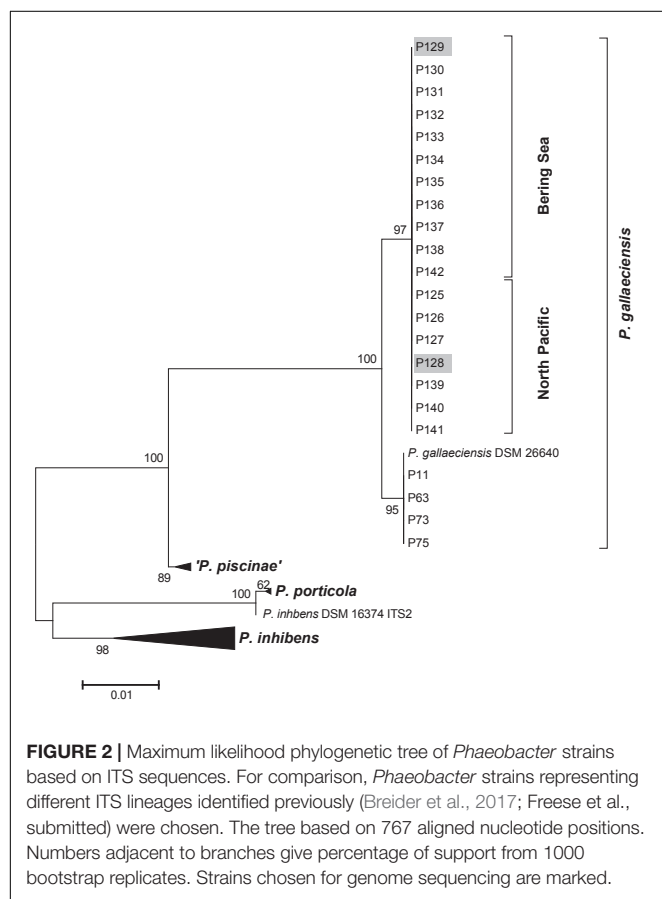
The specific MPNs calculated from the results of the liquid dilution series revealed highest *Phaeobacter* abundances of  $7.2 \cdot 10^4$  to  $1.2 \cdot 10^5$  cells (ml plankton biovolume)<sup>-1</sup> in the North Pacific (Table 2). In the warmer Southern Pacific, the *Phaeobacter* abundance was 2 orders of magnitude lower since only one well of the liquid dilution series proved to be positive (Figure 1).

A total of 18 strains were isolated from the samples collected in the North Pacific and the Bering Sea. The 16S rRNA gene sequences of all isolates were identical to the sequence of the *P. gallaeciensis* type strain DSM 26640<sup>T</sup> and to all other isolates of *P. gallaeciensis* which are available to date. The ITS sequences of all Pacific strains were also identical but clustered separately from the other *P. gallaeciensis* sequences due to five polymorphic sites (Figure 2). Based on the results of the ITS analysis, two strains isolated from locations 1500 km apart in the North Pacific (strain P128) and the Bering Sea (strain P129) and which had been enriched in two different media were chosen for genome sequencing. Their high quality, closed genome sequences were generated and compared to the five existing genome sequences of *P. gallaeciensis* strains originating from aquacultures in Spain and France (Supplementary Table S1).

The characteristics of the *P. gallaeciensis* genomes are summarized in Table 3. They share a large core genome (median, 89.6%), a high nucleotide similarity (median whole genome nucleotide similarity estimated as digital DNA:DNA hybridisation, 100%) and are highly syntenous with not more than 80 breakpoints (Figure 3). A considerable number of synteny breaks (31.5%) resulted from small sequence length variations (median 36 bp) in intergenic regions and 39.7% of the breaks occurred due to differences in gene content. The two isolates from the open Pacific were nearly identical in their genome sequence and differed by just 17 SNPs. Of these SNPs, 6 were located in intergenic regions whereas 11 were located in single core orthologs (nine non-synonymous, two synonymous). The two genomes of the Pacific strains differed by only 15

**TABLE 2** | Most probable numbers (MPNs) of *Phaeobacter* per ml plankton biovolume.

Station	Medium	<i>Phaeobacter</i>			Total numbers		
		MPN ml <sup>-1</sup>	95% confidence limits		MPN ml <sup>-1</sup>	95% confidence limits	
30°S	HD	0			1.5·10 <sup>6</sup>	1.1·10 <sup>6</sup>	1.9·10 <sup>6</sup>
30°S	AM	1.3·10 <sup>3</sup>	1.7·10 <sup>2</sup>	9.5·10 <sup>3</sup>	7.4·10 <sup>5</sup>	5.9·10 <sup>5</sup>	9.3·10 <sup>5</sup>
34°N	HD	0			6.4·10 <sup>5</sup>	4.4·10 <sup>5</sup>	9.3·10 <sup>5</sup>
34°N	AM	0			2.0·10 <sup>5</sup>	1.5·10 <sup>5</sup>	2.5·10 <sup>5</sup>
40°N	HD	1.2·10 <sup>5</sup>	5.7·10 <sup>4</sup>	2.3·10 <sup>5</sup>	1.1·10 <sup>7</sup>	8.7·10 <sup>6</sup>	1.4·10 <sup>7</sup>
40°N	AM	7.2·10 <sup>4</sup>	2.9·10 <sup>4</sup>	1.8·10 <sup>5</sup>	6.3·10 <sup>6</sup>	4.8·10 <sup>6</sup>	8.1·10 <sup>6</sup>
50°N	HD	0			6.0·10 <sup>6</sup>	4.9·10 <sup>6</sup>	7.4·10 <sup>6</sup>
50°N	AM	0			6.5·10 <sup>5</sup>	5.5·10 <sup>5</sup>	7.7·10 <sup>5</sup>
54°N	HD	0			3.4·10 <sup>6</sup>	2.7·10 <sup>6</sup>	4.4·10 <sup>6</sup>
54°N	AM	1.7·10 <sup>4</sup>	1.0·10 <sup>4</sup>	3.0·10 <sup>4</sup>	2.3·10 <sup>5</sup>	1.9·10 <sup>5</sup>	2.9·10 <sup>5</sup>



genes which are located on a small, 15-kb plasmid present in P129 but not P128 (labeled f in **Figure 3**). This plasmid carried a replication system and also a restriction system. Among the total of seven *P. gallaeciensis* strains, 84,374 SNPs were detected which were nearly homogeneously distributed over the whole genome (**Supplementary Figure S2**). The Pacific strains differed by 83,953 SNPs from the five aquaculture *P. gallaeciensis* strains whereas much fewer differences were detected among the latter (405 SNPs). Phylogenetic network analysis confirmed that the Pacific isolates constitute a separate clade within

*P. gallaeciensis* with no indication of recombination between the clades (**Figure 4**). Between the two clades, the average nucleotide similarity of homolog blocks still amounted to 97.8%. A higher sequence divergence was only detected for the RepABC-5 plasmid which also differed in gene content (**Figure 3**; plasmid labeled 'b' or 'x' have only 94.7% sequence similarity).

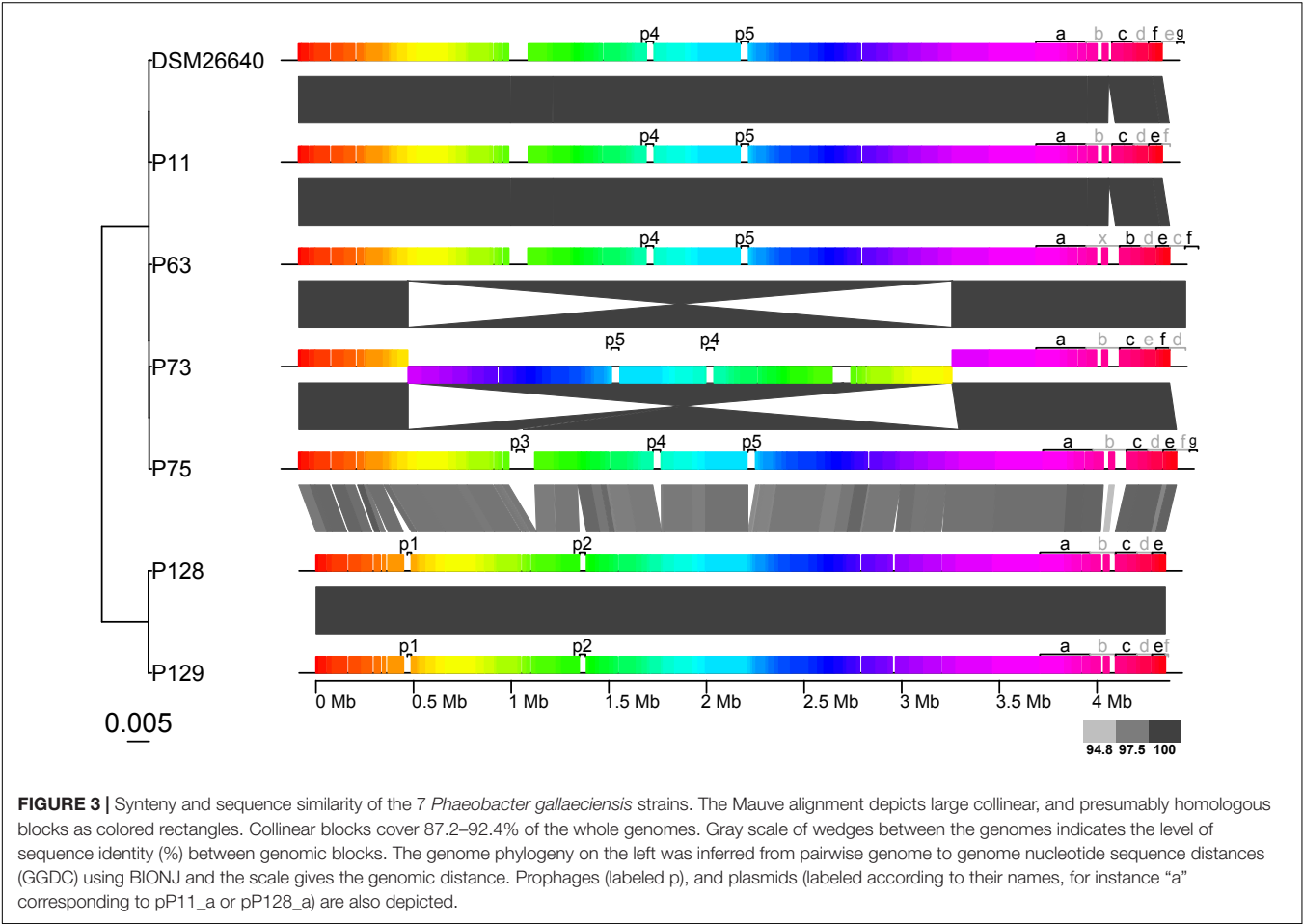
*Phaeobacter gallaeciensis* strains from the Pacific contained significantly fewer mobile genetic elements than the five strains from aquacultures (*t*-test  $p < 0.002$ ). Plasmids, prophages, and genomic islands together amounted 10.6% (464 kb) and 14.4% (653 kb) of the genomes, respectively (**Figure 5**). Significant differences were also apparent when only the mobile elements located in the chromosomes are considered (216 and 292 kb, *t*-test  $p < 0.001$ ). While the overall number of prophages did not differ between the two *P. gallaeciensis* clades, different taxa of prophages were present. Both Pacific strains contained two Myoviridae (labeled p1 and p2; **Figure 3**) whereas all members of the aquaculture clade harbored a representative each of the Siphoviridae (p4) and of the Podoviridae (p5). One aquaculture strain, P75, in addition contained a member of the Myoviridae (p3), which was totally dissimilar on nucleotide and protein level to the Myoviridae detected in the Pacific *P. gallaeciensis*. Interestingly, this latter prophage occurred in a large 100 kb-region which contains several transferred genetic elements but is completely lacking in the Pacific clade (**Figure 3**). At this particular position, other *Phaeobacter* species also contained different mobile elements (Freese et al., submitted) which may indicate a selective loss in the Pacific clade after its divergence.

All *P. gallaeciensis* strains contained the TDA gene cluster and genes involved in surface attachment and biofilm formation (like motility, chemotaxis, polysaccharide biosynthesis and transport, *rfb* genes). On the other hand, comparative analysis of the functional gene content identified 170 and 316 clade-specific orthologs in the Pacific and aquaculture strains of *P. gallaeciensis*, respectively (**Table 3**). Functional assignments were possible for a limited number of these orthologs (**Table 4**). In the aquaculture strains these innovative orthologs often represented functional orphans, i.e., they occur isolated in metabolic pathways like in the terpenoid metabolism, or as part of xenobiotics biodegradation or host infections. The aquaculture strains may also contain

**TABLE 3 |** Characteristics of *Phaeobacter gallaeciensis* genomes.

	Strains	Size (Mb)	No. of plasmids	mol% G+C	Genome sequence dissimilarity (%)	DDH (%)	No. of genes	% core genes	% clade core	% unique genes	Clade specific genes
Median	7	4.54	6	59.49	0.01	100.0	4339	89.6	na*	0.02	na
Minimum	7	4.35	5	59.44	0	81.2	4098	87.7	na	0.00	na
Maximum	7	4.61	7	59.62	2.19	100.0	4432	94.7	na	1.22	na
Median pacific	2	4.36	5.5	59.61	0	100.0	4106	94.5	99.8	0.06	170
Median other strains	5	4.54	7	59.46	0.01	100.0	4359	89.1	96.1	0.02	316

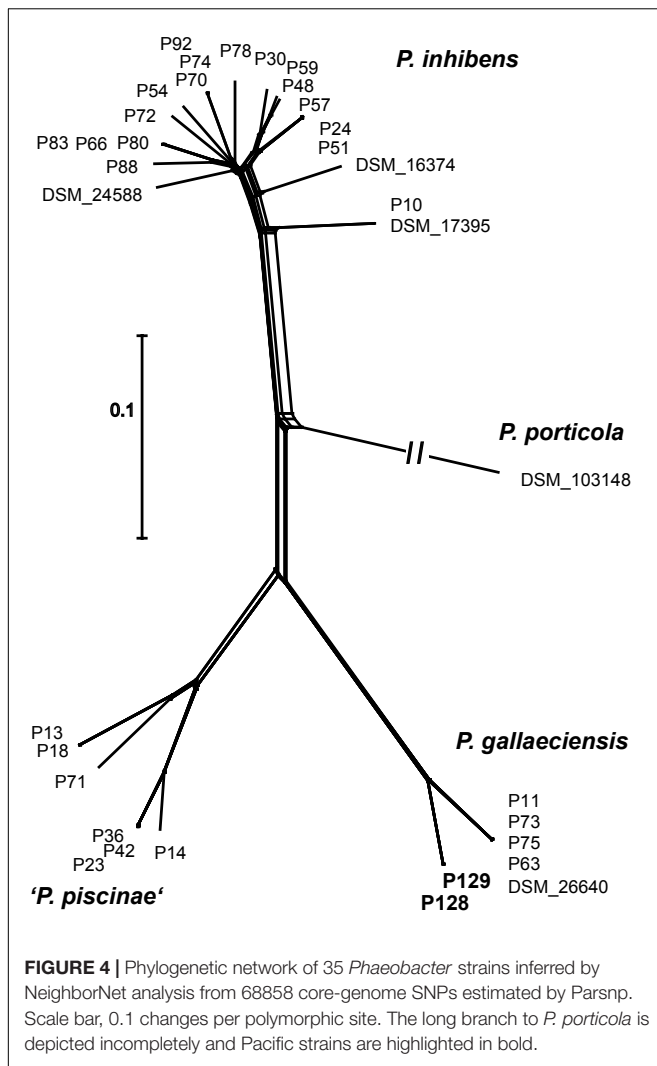
\*na, not applicable.  
DDH, whole genome nucleotide similarity estimated as digital DNA:DNA hybridisation.



few clade-specific genes coding for alternative reactions in carbohydrate or amino acid metabolism like a glutamate dehydrogenase (K00262), which may be involved in fixation or release of ammonia.

In contrast to the aquaculture strains, the major innovative function of the Pacific strains is a complete Type I restriction/modification system (K03427, K01154, K01153) which occurred twice (i.e., PhaeoP128\_00141 to 00144 and PhaeoP128\_00349 to 00351). A detailed analysis of methylation patterns revealed two previously unknown Type I restriction motifs which occur exclusively in the Pacific clade

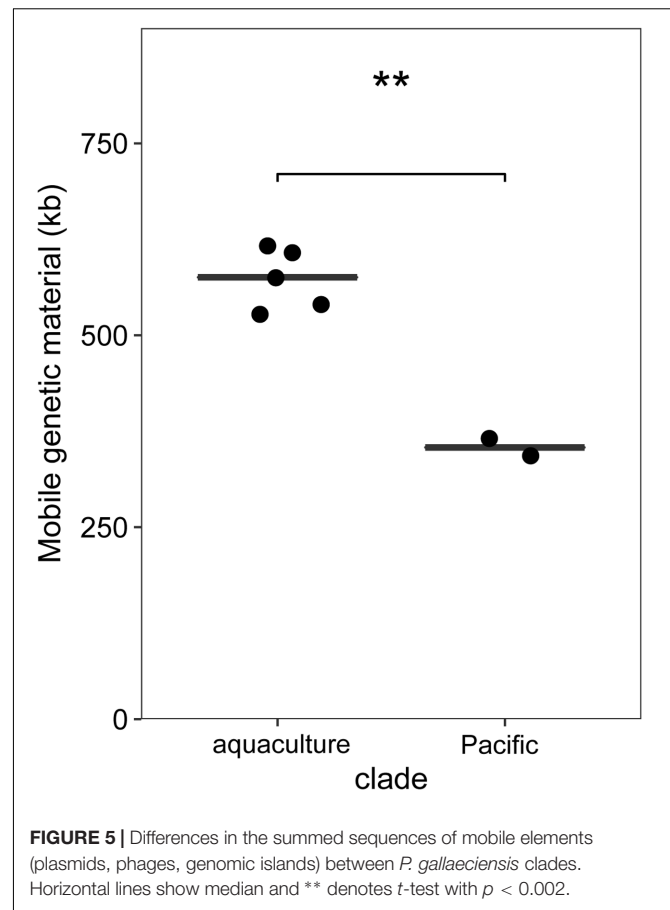
(AGCN<sub>6</sub>GTCY and AGCAN<sub>8</sub>TTYG; Table 5). On average, 99.1% (98.2–99.7%) of these two motifs were found to be methylated in the genomes of the Pacific strains, indicating that the Type I restriction modification systems are active. All *P. gallaeciensis* strains encode an identical, active Type II restriction system (modification methylase BabI, ribonuclease HII). Some further methylated motifs of restrictions systems were detected but the corresponding genes for methylases could not be identified. Remarkably, the Pacific clade of *P. gallaeciensis* also acquired a second, homolog phosphonate transport system (PhnCDE) in addition to the phosphonate transport system/C–P



lyase enzyme complex (PhnGHJKLM) present in all other *P. gallaeciensis*.

## DISCUSSION

*Phaeobacter* spp. were detected in the >300  $\mu\text{m}$  size plankton fraction which contained mostly mesozooplankton except for station SO248\_12 at 34°N where samples were dominated by radiolarian macrocolonies. In contrast, *Phaeobacter* spp. was neither detected in the free-living bacterioplankton nor in the plankton fraction >100  $\mu\text{m}$ . Due to the overall larger biovolume of the smaller plankton, the mesozooplankton (>300  $\mu\text{m}$ ) was highly diluted in the size fraction >100  $\mu\text{m}$  which likely is the reason for the lack of *Phaeobacter* in these samples. Our results reveal that *Phaeobacter* occurs associated with zooplankton in the open ocean but in contrast to our initial expectations we found no evidence for their association with phytoplankton. Although *Phaeobacter* was previously shown to colonize macroalgae and laboratory cultures of dinoflagellates (Rao et al., 2005; Frank et al., 2015) a symbiotic interaction



with algae was so far only demonstrated for the coccolithophore *E. huxleyi* (Seyedsayamdost et al., 2011a,b). In addition to zooplankton, *Phaeobacter* spp. may therefore also be specifically associated with some phytoplankton taxa which were not prevalent during the time of our cruise (e.g., Alvain et al., 2008).

Whereas, little information is available on the bacterial colonization of mesozooplankton in the open ocean, the total abundance of bacteria associated with estuarine or North Sea zooplankton was found to range between  $1.2 \cdot 10^8$  and  $3.6 \cdot 10^{12}$  (ml zooplankton biovolume)<sup>-1</sup> (Møller et al., 2007; Bickel and Tang, 2014). Based on these total bacterial cell numbers, the abundance of *Phaeobacter* spp. detected by our cultivation-based approach would amount to  $\leq 0.1\%$  of all zooplankton-associated bacteria. By comparison, marine vibrios constitute between 1 and 26% of the associated bacterial community (Heidelberg et al., 2002). The low abundance of *Phaeobacter* spp. on Pacific mesozooplankton resembles the low percentage of *P. inhibens* sequence reads determined in biofilms colonizing inert surface or marine animals in harbors (Gram et al., 2015) and thus may constitute a typical feature of the genus. As *Phaeobacter* can exert antibacterial activities against marine pathogens and prevent biofouling at low cell densities (Rao et al., 2007) they are likely of ecological relevance for their hosts even at the low abundances deduced in the present study.



**TABLE 4 |** Specific functional orthologs of *Phaeobacter gallaeciensis* clades.

Functional category	Pacific clade	Aquaculture clade
Amino acid metabolism		K01580, K00262, K01697
Carbohydrate metabolism	K00853, K01214	K01569, K00177, K05857, K03737
Catalase		K03781 (3x)
Cell growth and death		K03386, K03364
Energy metabolism		K02147
Glycan biosynthesis and metabolism		K05365
Host infection		K04646, K15352, K05692
Metabolism of cofactors and vitamins	K02549, K02551	
Metabolism of terpenoids and polyketides		K11731, K00587
Nervous system		K12460
Nucleotide metabolism		K02319
Replication and repair	K07466	K06223
Restriction and modification system (R-M system)	K03427 (2x), K01154 (2x), K01153 (2x), K07448	
Signal transduction	K04854, K12293	
Toxin–antitoxin system (TA system)		K07154
Transcription	K03142	
Translation	K14564	K14293, K03243
Transport and catabolism		K05940
Xenobiotics biodegradation and metabolism		K10221, K16046, K00517

So far, *Phaeobacter* spp. have neither been found associated with zooplankton in coastal and shelf areas (Møller et al., 2007; Tang et al., 2009) or in oligotrophic open ocean (Shoemaker and Moisaner, 2015). Elevated temperatures (above  $\sim 18^{\circ}\text{C}$ ) were considered as a main factor determining the occurrence of *Phaeobacter* in harbors (Gram et al., 2015) but we could not confirm this for the open ocean. It remains to be investigated if *Phaeobacter* is preferentially associated with specific zooplankton taxa and changes in abundance according to the seasonal abundance of their hosts as was described for other bacterial groups (Turner et al., 2009; Tang et al., 2010). However, considering that zooplankton-associated bacteria typically constitute only  $<0.1$ – $0.3\%$  of all water column bacteria in the marine environment (Møller et al., 2007; Bickel and Tang, 2014), *Phaeobacter* spp. must represent a very rare bacterial group in marine water samples with abundances lower ( $<10^{-3}\%$ ) than the detection limit of many studies employing next generation sequencing (0.64%; Pochon et al., 2013). Commensurate with these results the metagenomes of a large number of oceanic samples typically do not contain any 16S rRNA gene sequence specific for *Phaeobacter* spp. In contrast, our cultivation-based high-throughput approach in optimized media enabled us to recover representative genotypes despite their low *in situ* abundance and to study their specific features.

Even by cultivation-based methods, *Phaeobacter* spp. have so far only been detected in anthropogenic habitats like aquacultures and harbors (Hjelm et al., 2004; Prado et al., 2009; Gram et al., 2015; Xue et al., 2016) or were sporadically isolated from shoreline samples (Rao et al., 2005; Martens et al., 2006; Penesyan et al., 2009). In general solid marine broth agar media have been employed to isolate *Phaeobacter*. Although we used Pacific samples for which *in situ* temperatures were close to the temperature optimum of growth of *P. gallaeciensis* DSM 26640<sup>T</sup> and *P. inhibens* DSMZ 16374<sup>T</sup> (23–27 and 27–29°C, respectively; Ruiz-Ponte et al., 1998; Martens et al., 2006), our attempts to isolate *Phaeobacter* on solid media were not successful. This suggests that agar media might be less suitable for recovering *Phaeobacter* from the open ocean. The *Phaeobacter* strains obtained in the present study employing optimized liquid enrichment strategies represent the first isolates from the open ocean environment. These genotypes constitute a phylogenomically separate subclade and hence do not represent known variants that link the distant coastal populations across the globe. Therefore their genomes were investigated for potential mechanisms of adaption to the open ocean habitat and to the inferred association to mesozooplankton.

All known genetic elements characterizing the surface- and host-associated lifestyle of *Phaeobacter* (Thole et al., 2012; Frank et al., 2015) are generally present in both *P. gallaeciensis* clades. The most prominent genomic features of the Pacific strains are their two complete, exclusive, and active Type I restriction modification systems which act as phage defense mechanism (Stern and Sorek, 2011; Loenen et al., 2014). Unlike their aquaculture counterparts, the Pacific strains only contained one type of prophage despite the high abundance and diversity of bacteriophages present in the ocean (Breitbart, 2012), suggesting an effective protection of the host against a broader range of bacteriophages. These restriction systems also are likely to prevent the horizontal transfer of foreign DNA through gene transfer agents (Stern and Sorek, 2011; Lang et al., 2012) which represent a major mechanism of gene acquisition of the genus *Phaeobacter* (Freese et al. submitted). This may explain the lower number and functional diversity of mobile elements observed in the Pacific strains of *P. gallaeciensis* compared to their aquaculture relatives. In addition, the slight reduction of the genome size of the Pacific strains by 0.18 Mb maybe related to an incipient genome streamlining as a consequence of nutrient limitation (Giovannoni et al., 2014; Luo and Moran, 2014). In this context it is remarkable that the genotypes of the Pacific *P. gallaeciensis* clade not only encode the complete phosphonate transport and degradation complex (cf. Villarreal-Chiu et al., 2012) like the aquaculture strains but in addition acquired and maintain a second set of phosphonate transporter (PhnCDE) despite their genome reduction. Phosphonates are particular prominent among marine invertebrate (Quin, 2000). Phosphonates also constitute a third of the dissolved organic phosphorus in oceanic waters and marine bacteria from the *Roseobacter* group have been shown to be capable of utilizing these compounds

**TABLE 5 |** Presence of restriction system motifs in *Phaeobacter gallaeciensis* strains.

Motif*	R-M characteristics			Strain						
	Modification	Type	Sub type	P128	P129	DSM 26640	P11	P63	P73	P75
AGCN <sub>6</sub> GTCY	m6 <sub>A</sub>	I	nd	+	+	na	–	–	–	–
AGCAN <sub>8</sub> TTYG	m6 <sub>A</sub>	I	nd	+	+	na	–	–	–	–
GANTC	m6 <sub>A</sub>	II	beta	+	+	na	+	+	+	+
TTCGAG	m6 <sub>A</sub>	II	G,S, gamma	–	–	na	+	+	+	+
CACN <sub>3</sub> GTCG	m6 <sub>A</sub>	II	gamma	–	+	na	–	–	–	–
CGATCG	m6 <sub>A</sub>	II	nd	–	–	na	+	+	+	+
GGCGAG	m6 <sub>A</sub>		nd	+	+	na	–	–	–	–
CGCRGAC	m6 <sub>A</sub>		nd	–	–	na	–	+	+	–

\*Methylated position within the motif is highlighted in bold and underscored when occurring on the complementary strand.

na: genome of strain DSM 26640 was not done by SMRT sequencing therefore information on methylation pattern is not available.

(Martinez et al., 2010). Pacific genotypes of *P. gallaeciensis* may therefore be particularly adapted to this alternative source of phosphorus.

Zooplankton surfaces are colonized by distinct bacterial communities which vary between zooplankton species and body regions (Tang et al., 2010). It has yet to be determined whether *P. gallaeciensis* can preferentially colonize the body surface or the intestinal tract of mesozooplankton or is also present on other organisms and particles >300  $\mu\text{m}$ . Zooplankton functions as microbial hotspots since it provides a nutrient-rich environment than the surrounding water (Tang et al., 2010). A possible reciprocal advantage of *P. gallaeciensis* for the host could be related to the formation of TDA and other antibiotics that might affect the success of other bacteria to colonize zooplankton.

The genomes of the two Pacific strains were highly similar (97.8% nucleotide similarity; fraction of core genome, 89.6%) to those of all other *P. gallaeciensis* strains from aquacultures. A genome conservation such as that observed in *Phaeobacter* is rare and even exceeds that of species with specialized ecological niches like the oligotrophic pelagic SAR11 subclade 1a (77.7% fraction of the core genome; Grote et al., 2012) or symbiotic *Vibrio fischeri* (80.4% core genome; Bongrand et al., 2016). Only some small obligate intracellular bacteria like the *Chlamydia psittaci* group (89.5% core genome; Voigt et al., 2012) and *Rickettsia* (84–93% core genome; Fuxelius et al., 2008) reach a similarly high genome conservation, but have a much smaller genome size. This raises the question which mechanisms underlie the unexpected high similarity of the *P. gallaeciensis* genomes.

Based on the homogenous random distribution of SNPs over the genome, recombination events and gene-specific sweeps through the population of *P. gallaeciensis* are unlikely (Marttinen et al., 2012; Shapiro et al., 2012). Instead, the different lineages must have diversified by neutral mutations. The two Pacific *P. gallaeciensis* strains detected in the mesozooplankton fraction differed by only 17 SNPs and were isolated from water samples 1500 km apart. Water currents have been shown to transport Pacific copepods over large distances up to 5000 km (Tätebe et al., 2010). The coastal strains of *P. gallaeciensis* for which genomes are available were isolated from aquacultures in Spain

and France (Supplementary Table S1) which are located at a distance of 8,600–10,200 km to the habitats of the Pacific *P. gallaeciensis* isolates. These aquaculture strains differed by 83,953 SNPs from the Pacific strains. Our data allow to infer hypotheses regarding the mechanism of genomic differentiation within the species *P. gallaeciensis*. The rate of spontaneous mutation has recently been determined for *Ruegeria pomeroyi*, another member of the *Roseobacter* group, as  $1.39 \cdot 10^{-10}$  base<sup>-1</sup> generation<sup>-1</sup> (Sun et al., 2017). Accordingly, *P. gallaeciensis* with a genome size of 4.54 Mb would take 1585 generations to acquire a SNP. *R. pomeroyi* reaches 45 generations per year in its marine environment (Sun et al., 2017). Due to the higher substrate supply, bacteria associated with marine copepods grow at rates which are 3–18 times higher than free-living bacteria (Tang et al., 2010). They can attain growth rates between 0.7 and 1.2 d<sup>-1</sup>, corresponding to generation times of 0.58–0.99 days, depending on the feeding status of the zooplankton (Tang, 2005; Möller et al., 2007). Based on these faster growth rates, the 1585 generations would take between 919 and 1569 days or 2.5 to 4.3 years. Correspondingly, the 17 SNPs that distinguish the two genomes of the Pacific strains may have accumulated over 20 to 38.7 years. By a similar calculation, it would take between 164,940 and 180,501 years to acquire the larger number of SNPs that distinguish the Pacific from the phylogenetically related aquaculture strains. This time frame is substantially larger than the time required for the global overturn of ocean waters by surface currents and thermohaline circulation (1000–2000 years; Döös et al., 2012) which renders geographic isolation in different water bodies a highly unlikely mechanism of genomic divergence in *P. gallaeciensis*.

## CONCLUSION

Our data indicate that a previously unknown, distinct clade of *P. gallaeciensis* acquired a limited number of clade-specific genes which may be relevant for its association with mesozooplankton and the colonization of the marine pelagial. The divergence of the Pacific clade was most likely driven by the adaptation to this novel ecological niche rather than by geographic isolation. The distribution pattern observed in the present study also provides

first indications for possible biotic interactions between the pelagic lineage of *P. gallaeciensis* and marine mesozooplankton.

## AUTHOR CONTRIBUTIONS

HF and AM performed the sampling and enrichments. AM conducted the screening and isolation. HF performed the analysis of genome sequences and other data. HF and JO designed the study and wrote the manuscript.

## ACKNOWLEDGMENTS

The work was supported by the Deutsche Forschungsgemeinschaft (DFG) within the Transregional Collaborative Research Centre “*Roseobacter*” (TRR 51/2 TA07). We thank the crew of RV Sonne for their great support during the cruise SO248 which was funded by the German Federal Ministry of Education and Research (BMBF) within the BacGeoPac project (03G0248A),

and Maria Pinto Gomes Ribeiro Teixeira who pre-processed samples of the horizontal haul. We are grateful for the fruitful discussions with Boyke Bunk and Johannes Sikorski who also helped with **Supplementary Figure S2**. We thank Cathrin Spröer for genome sequencing, Franziska Klann, Nicole Heyer and Simone Severitt for excellent technical assistance, and Isabel Schober for bioinformatics support and submission of genomes to NCBI.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01659/full#supplementary-material>

**FIGURE S1** | Flow scheme of combined methods applied in the current study.

**FIGURE S2** | Distribution of nucleotide diversity ( $\Pi$ ) per gene between *P. gallaeciensis* clades ordered along the genome of P75 by means of the locus\_tag number.

## REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. *Glob. Biogeochem. Cycles* 22:GB3001. doi: 10.1029/2007GB003154
- Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21. doi: 10.1093/nar/gkw387
- Badewien, T. H., Winkler, H., Arndt, K. L., and Simon, M. (2016). *Physical Oceanography during SONNE Cruise SO248 (BacGeoPac)*. Oldenburg: University of Oldenburg. doi: 10.1594/PANGAEA.864673
- Balcazar, J. L., Lee, N. M., Pintado, J., and Planas, M. (2010). Phylogenetic characterization and in situ detection of bacterial communities associated with seahorses (*Hippocampus guttulatus*) in captivity. *Syst. Appl. Microbiol.* 33, 71–77. doi: 10.1016/j.syapm.2009.11.005
- Bickel, S. L., and Tang, K. W. (2014). Zooplankton-associated and free-living bacteria in the York River, Chesapeake Bay: comparison of seasonal variations and controlling factors. *Hydrobiologia* 722, 305–318. doi: 10.1007/s10750-013-1725-0
- Billerbeck, S., Wemheuer, B., Voget, S., Poehlein, A., Giebel, H. A., Brinkhoff, T., et al. (2016). Biogeography and environmental genomics of the *Roseobacter*-affiliated pelagic CHAB-I-5 lineage. *Nat. Microbiol.* 1:16063. doi: 10.1038/nmicrobiol.2016.63
- Bongrand, C., Koch, E. J., Moriano-Gutierrez, S., Cordero, O. X., McFall-Ngai, M., Polz, M. F., et al. (2016). A genomic comparison of 13 symbiotic *Vibrio fischeri* isolates from the perspective of their host source and colonization behavior. *ISME J.* 10, 2907–2917. doi: 10.1038/ismej.2016.69
- Breider, S., Freese, H. M., Spröer, C., Simon, M., Overmann, J., and Brinkhoff, T. (2017). *Phaeobacter porticola* sp. nov., an antibiotic-producing bacterium isolated from a sea harbour. *Int. J. Syst. Evol. Microbiol.* doi: 10.1099/ijsem.0.001879 [Epub ahead of print].
- Breitbart, M. (2012). Marine viruses: truth or dare. *Annu. Rev. Mar. Sci.* 4, 425–448. doi: 10.1146/annurev-marine-120709-142805
- Brinkhoff, T., Bach, G., Heidorn, T., Liang, L., Schlingloff, A., and Simon, M. (2004). Antibiotic production by a *Roseobacter* clade-affiliated species from the German Wadden Sea and its antagonistic effects on indigenous isolates. *Appl. Environ. Microbiol.* 70, 2560–2565. doi: 10.1128/AEM.70.4.2560-2565.2003
- Bruns, A., Phillip, H., Cypionka, H., and Brinkhoff, T. (2003). *Aeromicrobium marinum* sp. nov., an abundant pelagic bacterium isolated from the German Wadden Sea. *Int. J. Syst. Evol. Microbiol.* 53, 1917–1923. doi: 10.1099/ijms.0.02735-0
- Buchan, A., González, J. M., and Moran, M. A. (2005). Overview of the marine *Roseobacter* lineage. *Appl. Environ. Microbiol.* 71, 5665–5677. doi: 10.1128/AEM.71.10.5665-5677.2005
- D’Alvise, P. W., Lillebø, S., Prol-García, M. J., Wergeland, H. I., Nielsen, K. F., Bergh, O., et al. (2012). *Phaeobacter gallaeciensis* reduces *Vibrio anguillarum* in cultures of microalgae and rotifers, and prevents vibriosis in cod larvae. *PLoS ONE* 7:e43996. doi: 10.1371/journal.pone.0043996
- Darling, A. E., Mau, B., and Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* 5:e11147. doi: 10.1371/journal.pone.0011147
- Dhillon, B. K., Chiu, T. A., Laird, M. R., Langille, M. G., and Brinkman, F. S. (2013). IslandViewer update: improved genomic island discovery and visualization. *Nucleic Acids Res.* 41, W129–W132. doi: 10.1093/nar/gkt394
- Dickschat, J. S., Zell, C., and Brock, N. L. (2010). Pathways and substrate specificity of DMSP catabolism in marine bacteria of the *Roseobacter* clade. *ChemBiochem* 11, 417–425. doi: 10.1002/cbic.200900668
- Doghri, I., Rodrigues, S., Bazire, A., Dufour, A., Akbar, D., Sopena, V., et al. (2015). Marine bacteria from the French Atlantic coast displaying high forming-biofilm abilities and different biofilm 3D architectures. *BMC Microbiol.* 15:231. doi: 10.1186/s12866-015-0568-4
- Döös, K., Nilsson, J., Nycander, J., Brodeau, L., and Ballarotta, M. (2012). The World Ocean Thermohaline Circulation. *J. Phys. Oceanogr.* 42, 1445–1460. doi: 10.1175/JPO-D-11-0163.1
- Federhen, S. (2015). Type material in the NCBI taxonomy database. *Nucleic Acids Res.* 43, D1086–D1098. doi: 10.1093/nar/gku1127
- Fisher, M. M., and Triplett, E. W. (1999). Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Appl. Environ. Microbiol.* 65, 4630–4636.
- Frank, O., Michael, V., Pauker, O., Boedeker, C., Jögler, C., Rohde, M., et al. (2015). Plasmid curing and the loss of grip - The 65-kb replicon of *Phaeobacter inhibens* DSM 17395 is required for biofilm formation, motility and the colonization of marine algae. *Syst. Appl. Microbiol.* 38, 120–127. doi: 10.1016/j.syapm.2014.12.001
- Fuxelius, H. H., Darby, A. C., Cho, N. H., and Andersson, S. G. (2008). Visualization of pseudogenes in intracellular bacteria reveals the different tracks to gene destruction. *Genome Biol.* 9:R42. doi: 10.1186/gb-2008-9-2-r42



- Gascuel, O. (1997). BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol. Biol. Evol.* 14, 685–695. doi: 10.1093/oxfordjournals.molbev.a025808
- Giovannoni, S. J., Thrash, J., and Temperton, B. (2014). Implications of streamlining theory for microbial ecology. *ISME J.* 8, 1553–1565. doi: 10.1038/ismej.2014.60
- Gram, L., Rasmussen, B. B., Wemheuer, B., Bernbom, N., Ng, Y. Y., Porsby, C. H., et al. (2015). *Phaeobacter inhibens* from the *Roseobacter* clade has an environmental niche as a surface colonizer in harbors. *Syst. Appl. Microbiol.* 38, 483–493. doi: 10.1016/j.syapm.2015.07.006
- Grote, J., Thrash, J., Huggett, M. J., Landry, Z. C., Carini, P., Giovannoni, S. J., et al. (2012). Streamlining and core genome conservation among highly divergent members of the SAR11 clade. *mBio* 3:e00252–12. doi: 10.1128/mBio.00252-12
- Guy, L., Roat Kultima, J., and Andersson, S. G. E. (2010). genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* 26, 2334–2335. doi: 10.1093/bioinformatics/btq413
- Heidelberg, J. F., Heidelberg, K. B., and Colwell, R. R. (2002). Bacteria of the gamma-subclass *Proteobacteria* associated with zooplankton in Chesapeake Bay. *Appl. Environ. Microbiol.* 68, 5498–5507. doi: 10.1128/AEM.68.11.5498-5507.2002
- Hjelm, M., Riiza, A., Formoso, F., Melchiorson, J., and Gram, L. (2004). Seasonal incidence of autochthonous antagonistic *Roseobacter* spp. and Vibrionaceae strains in a turbot larva (*Scophthalmus maximus*) rearing system. *Appl. Environ. Microbiol.* 70, 7288–7294. doi: 10.1128/AEM.70.12.7288-7294.2004
- Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030
- Jarvis, B., Wilrich, C., and Wilrich, P. T. (2010). Reconsideration of the derivation of Most Probable Numbers, their standard deviations, confidence bounds and rarity values. *J. Appl. Microbiol.* 109, 1660–1667. doi: 10.1111/j.1365-2672.2010.04792.x
- Karim, M., Zhao, W., Rowley, D., Nelson, D., and Gomez-Chiarri, M. (2013). Probiotic strains for shellfish aquaculture: protection of eastern oyster, *Crassostrea virginica*, larvae and juveniles against bacterial challenge. *J. Shellfish Res.* 32, 401–408. doi: 10.2983/035.032.0220
- Karsten, G. R., and Drake, H. L. (1995). Comparative assessment of the aerobic and anaerobic microfloras of earthworm guts and forest soils. *Appl. Environ. Microbiol.* 61, 1039–1044.
- Kesarodi-Watson, A., Miner, P., Nicolas, J. L., and Robert, R. (2012). Protective effect of four potential probiotics against pathogen-challenge of the larvae of three bivalves: pacific oyster (*Crassostrea gigas*), flat oyster (*Ostrea edulis*) and scallop (*Pecten maximus*). *Aquaculture* 34, 29–34. doi: 10.1016/j.aquaculture.2012.02.029
- Lane, D. J. (1991). “16S/23S rRNA sequencing,” in *Nucleic Acid Techniques in Bacterial Systematics*, eds E. Stackebrandt and M. Goodfellow (Chichester: John Wiley & Sons), 115–175.
- Lang, A. S., Zhaxybayeva, O., and Beatty, J. T. (2012). Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* 10, 472–482. doi: 10.1038/nrmicro2802
- Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., and Prohaska, S. J. (2011). Proteinortho: detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12:124. doi: 10.1186/1471-2105-12-124
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Loenen, W. A. M., Dryden, D. T. F., Raleigh, E. A., and Wilson, G. G. (2014). Type I restriction enzymes and their relatives. *Nucleic Acids Res.* 42, 20–44. doi: 10.1093/nar/gkt847
- Lopes, A., Tavares, P., Petit, M. A., Guerois, R., and Zinn-Justin, S. (2014). Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics* 15:1027. doi: 10.1186/1471-2164-15-1027
- Luo, H., and Moran, M. A. (2014). Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol. Mol. Biol. Rev.* 78, 573–587. doi: 10.1128/MMBR.00020-14
- Martens, T., Gram, L., Grossart, H. P., Kessler, D., Muller, R., Simon, M., et al. (2007). Bacteria of the *Roseobacter* clade show potential for secondary metabolite production. *Microb. Ecol.* 54, 31–42. doi: 10.1007/s00248-006-9165-2
- Martens, T., Heidorn, T., Pukall, R., Simon, M., Tindall, B. J., and Brinkhoff, T. (2006). Reclassification of *Roseobacter gallaeciensis* Ruiz-Ponte et al. 1998 as *Phaeobacter gallaeciensis* gen. nov., comb. nov., description of *Phaeobacter inhibens* sp. nov., reclassification of *Ruegeria algicola* (Lafay et al. 1995) Uchino et al. 1999 as *Marinovum algicola* gen. nov., comb. nov., and emended descriptions of the genera *Roseobacter*, *Ruegeria* and *Leisingera*. *Int. J. Syst. Evol. Microbiol.* 56, 1293–1304. doi: 10.1099/ijes.0.63724-0
- Martinez, A., Tyson, G. W., and DeLong, E. F. (2010). Widespread known and novel phosphonate utilization pathways in marine bacteria revealed by functional screening and metagenomic analyses. *Environ. Microbiol.* 12, 222–238. doi: 10.1111/j.1462-2920.2009.02062.x
- Martinen, P., Hanage, W. P., Croucher, N. J., Connor, T. R., Harris, S. R., Bentley, S. D., et al. (2012). Detection of recombination events in bacterial genomes from large population samples. *Nucleic Acids Res.* 40:e6. doi: 10.1093/nar/gkr928
- Meier-Kolthoff, J. P., Auch, A. F., Klenk, H. P., and Göker, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14:60. doi: 10.1186/1471-2105-14-60
- Møller, E. F., Riemann, L., and Søndergaard, M. (2007). Bacteria associated with copepods: abundance, activity and community composition. *Aquat. Microb. Ecol.* 47, 99–106. doi: 10.3354/ame047099
- Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A. C., and Kanehisa, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35, W182–W185. doi: 10.1093/nar/gkm321
- Newton, R. J., Griffin, L. E., Bowles, K. M., Meile, C., Gifford, S., Givens, C. E., et al. (2010). Genome characteristics of a generalist marine bacterial lineage. *ISME J.* 4, 784–798. doi: 10.1038/ismej.2009.150
- Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290. doi: 10.1093/bioinformatics/btg412
- Penesyan, A., Marshall-Jones, Z., Holmstrom, C., Kjelleberg, S., and Egan, S. (2009). Antimicrobial activity observed among cultured marine epiphytic bacteria reflects their potential as a source of new drugs. *FEMS Microbiol. Ecol.* 69, 113–124. doi: 10.1111/j.1574-6941.2009.00688.x
- Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., and Lercher, M. J. (2014). PopGenome: an efficient swiss army knife for population genomic analyses in R. *Mol. Biol. Evol.* 31, 1929–1936. doi: 10.1093/molbev/msu136
- Pochon, X., Bott, N. J., Smith, K. F., and Wood, S. A. (2013). Evaluating detection limits of next-generation sequencing for the surveillance and monitoring of international marine pests. *PLoS ONE* 8:e73935. doi: 10.1371/journal.pone.0073935
- Porsby, C. H., Nielsen, K. F., and Gram, L. (2008). *Phaeobacter* and *Ruegeria* species of the *Roseobacter* clade colonize separate niches in a Danish turbot (*Scophthalmus maximus*)-rearing farm and antagonize *Vibrio anguillarum* under different growth conditions. *Appl. Environ. Microbiol.* 74, 7356–7364. doi: 10.1128/AEM.01738-08
- Prado, S., Montes, J., Romalde, J. L., and Barja, J. L. (2009). Inhibitory activity of *Phaeobacter* strains against aquaculture pathogenic bacteria. *Int. Microbiol.* 12, 107–114.
- Quin, L. D. (2000). *A Guide to Organophosphorus Chemistry*. New York, NY: John Wiley & Sons.
- Rao, D., Skovhus, T., Tujula, N., Holmstrom, C., Dahllof, I., Webb, J. S., et al. (2010). Ability of *Pseudoalteromonas tunicata* to colonize natural biofilms and its effect on microbial community structure. *FEMS Microbiol. Ecol.* 73, 450–457. doi: 10.1111/j.1574-6941.2010.00917.x
- Rao, D., Webb, J. S., Holmstrom, C., Case, R., Low, A., Steinberg, P., et al. (2007). Low densities of epiphytic bacteria from the marine alga *Ulva australis* inhibit settlement of fouling organisms. *Appl. Environ. Microbiol.* 73, 7844–7852. doi: 10.1128/AEM.01543-07
- Rao, D., Webb, J. S., and Kjelleberg, S. (2005). Competitive interactions in mixed-species biofilms containing the marine bacterium *Pseudoalteromonas tunicata*. *Appl. Environ. Microbiol.* 71, 1729–1736. doi: 10.1128/AEM.71.4.1729-1736.2005
- Roberts, R. J., Vincze, T., Posfai, J., and Macelis, D. (2015). REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* 43, D298–D299. doi: 10.1093/nar/gku1046
- Ruiz-Ponte, C., Cilia, V., Lambert, C., and Nicolas, J. L. (1998). *Roseobacter gallaeciensis* sp. nov., a new marine bacterium isolated from rearings and collectors of the scallop *Pecten maximus*. *Int. J. Syst. Microbiol.* 48, 537–542.



- Ruiz-Ponte, C., Samain, J. F., Sánchez, J. L., and Nicolas, J. L. (1999). The benefit of a *Roseobacter* species on the survival of scallop larvae. *Mar. Biotechnol.* 1, 52–59. doi: 10.1007/PL00011751
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Seyedsayamdost, M. R., Carr, G., Kolter, R., and Clardy, J. (2011a). Roseobacticides: small molecule modulators of an algal-bacterial symbiosis. *J. Am. Chem. Soc.* 133, 18343–18349. doi: 10.1021/ja207172s
- Seyedsayamdost, M. R., Case, R. J., Kolter, R., and Clardy, J. (2011b). The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis*. *Nat. Chem.* 3, 331–335. doi: 10.1038/nchem.1002
- Shapiro, B. J., Friedman, J., Cordero, O. X., Preheim, S. P., Timberlake, S. C., Szabó, G., et al. (2012). Population genomics of early events in the ecological differentiation of bacteria. *Science* 336, 48–51. doi: 10.1126/science.1218198
- Shoemaker, K. M., and Moisan, P. H. (2015). Microbial diversity associated with copepods in the North Atlantic subtropical gyre. *FEMS Microbiol. Ecol.* 91:fiv064. doi: 10.1093/femsec/fiv064
- Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Dobler, I., Ulbrich, M., et al. (2017). Phylogenomics of *Rhodobacteraceae* reveals evolutionary adaptation to marine and non-marine habitats. *ISME J.* 11, 1483–1499. doi: 10.1038/ismej.2016.198
- Sonnenschein, E. C., Nielsen, K. F., D'Alvise, P., Porsby, C. H., Melchiorson, J., Heilmann, J., et al. (2017). Global occurrence and heterogeneity of the *Roseobacter*-clade species *Ruegeria mobilis*. *ISME J.* 11, 569–583. doi: 10.1038/ismej.2016.111
- Stern, A., and Sorek, R. (2011). The phage-host arms-race: shaping the evolution of microbes. *Bioessays* 33, 43–51. doi: 10.1002/bies.201000071
- Sun, Y., Powell, K. E., Sung, W., Lynch, M., Moran, M. A., and Luo, H. (2017). Spontaneous mutations of a model heterotrophic marine bacterium. *ISME J.* 11, 1713–1718. doi: 10.1038/ismej.2017.20
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tang, K., Dziallas, C., Hutalle-Schmelzer, K., and Grossart, H. P. (2009). Effects of food on bacterial community composition associated with the copepod *Acartia tonsa* Dana. *Biol. Lett.* 5, 549–553. doi: 10.1098/rsbl.2009.0076
- Tang, K. W. (2005). Copepods as microbial hotspots in the ocean: effects of host feeding activities on attached bacteria. *Aquat. Microb. Ecol.* 38, 31–40. doi: 10.3354/ame038031
- Tang, K. W., Turk, V., and Grossart, H. P. (2010). Linkage between crustacean zooplankton and aquatic bacteria. *Aquat. Microb. Ecol.* 61, 261–277. doi: 10.3354/ame01424
- Tatebe, H., Yasuda, I., Saito, H., and Shimizu, Y. (2010). Horizontal transport of the calanoid copepod *Neocalanus* in the North Pacific: the influences of the current system and the life history. *Deep Sea Res. I* 57, 409–419. doi: 10.1016/j.dsr.2009.11.009
- Thole, S., Kalhoefer, D., Voget, S., Berger, M., Engelhardt, T., Liesegang, H., et al. (2012). *Phaeobacter gallaeciensis* genomes from globally opposite locations reveal high similarity of adaptation to surface life. *ISME J.* 6, 2229–2244. doi: 10.1038/ismej.2012.62
- Treangen, T. J., Ondov, B. D., Koren, S., and Phillippy, A. M. (2014). The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol.* 15, 524. doi: 10.1186/s13059-014-0524-x
- Tscheck, A., and Pfennig, N. (1984). Growth yield increase linked to caffeate reduction in *Acetobacterium woodii*. *Arch. Microbiol.* 137, 163–167. doi: 10.1007/BF00414460
- Turner, J. W., Good, B., Cole, D., and Lipp, E. K. (2009). Plankton composition and environmental factors contribute to *Vibrio* seasonality. *ISME J.* 3, 1082–1092. doi: 10.1038/ismej.2009.50
- Villarreal-Chiu, J. F., Quinn, J. P., and McGrath, J. W. (2012). The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. *Front. Microbiol.* 3:19. doi: 10.3389/fmicb.2012.00019
- Voigt, A., Schöfl, G., and Saluz, H. P. (2012). The *Chlamydia psittaci* genome: a comparative analysis of intracellular pathogens. *PLoS ONE* 7:e35097. doi: 10.1371/journal.pone.0035097
- Wagner-Döbler, I., and Biebl, H. (2006). Environmental biology of the marine *Roseobacter* lineage. *Annu. Rev. Microbiol.* 60, 255–280. doi: 10.1146/annurev.micro.60.080805.142115
- Widdel, F., Kohring, G. W., and Mayer, F. (1983). Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids. *Arch. Microbiol.* 134, 286–294. doi: 10.1007/BF00407804
- Xue, M., Liang, H. F., He, Y. Y., and Wen, C. Q. (2016). Characterization and in-vivo evaluation of potential probiotics of the bacterial flora within the water column of a healthy shrimp larviculture system. *Chin. J. Oceanol. Limnol.* 34, 484–491. doi: 10.1007/s00343-016-5024-2

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Freese, Methner and Overmann. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The Biogeographical Distribution of Benthic *Roseobacter* Group Members along a Pacific Transect Is Structured by Nutrient Availability within the Sediments and Primary Production in Different Oceanic Provinces

## OPEN ACCESS

### Edited by:

Karla B. Heidelberg,  
University of Southern California,  
United States

### Reviewed by:

Adam R. Rivers,  
Agricultural Research Service (USDA),  
United States  
Andreas Teske,  
University of North Carolina at Chapel  
Hill, United States

### \*Correspondence:

Bert Engelen  
engelen@icbm.de

### † Present Address:

Bernd Wemheuer,  
Centre for Marine Bio-Innovation,  
The University of New South Wales,  
Sydney, NSW, Australia

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 27 June 2017

**Accepted:** 08 December 2017

**Published:** 18 December 2017

### Citation:

Pohlner M, Degenhardt J,  
von Hoyningen-Huene AJE,  
Wemheuer B, Erlmann N,  
Schnetger B, Badewien TH and  
Engelen B (2017) The Biogeographical  
Distribution of Benthic *Roseobacter*  
Group Members along a Pacific  
Transect Is Structured by Nutrient  
Availability within the Sediments and  
Primary Production in Different  
Oceanic Provinces.  
Front. Microbiol. 8:2550.  
doi: 10.3389/fmicb.2017.02550

Marion Pohlner<sup>1</sup>, Julius Degenhardt<sup>1</sup>, Avril J. E. von Hoyningen-Huene<sup>2</sup>,  
Bernd Wemheuer<sup>2†</sup>, Nora Erlmann<sup>3</sup>, Bernhard Schnetger<sup>3</sup>, Thomas H. Badewien<sup>4</sup> and  
Bert Engelen<sup>1\*</sup>

<sup>1</sup> Paleomicrobiology Group, Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany, <sup>2</sup> Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, University of Göttingen, Göttingen, Germany, <sup>3</sup> Microbiogeochemistry Group, Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany, <sup>4</sup> Group "Marine Sensor Systems", Institute for Chemistry and Biology of the Marine Environment, University of Oldenburg, Oldenburg, Germany

By now, only limited information on the *Roseobacter* group thriving at the seafloor is available. Hence, the current study was conducted to determine their abundance and diversity within Pacific sediments along the 180° meridian. We hypothesize a distinct biogeographical distribution of benthic members of the *Roseobacter* group linked to nutrient availability within the sediments and productivity of the water column. Lowest cell numbers were counted at the edge of the south Pacific gyre and within the north Pacific gyre followed by an increase to the north with maximum values in the highly productive Bering Sea. Specific quantification of the *Roseobacter* group revealed on average a relative abundance of 1.7 and 6.3% as determined by catalyzed reported deposition-fluorescence *in situ* hybridization (CARD-FISH) and quantitative PCR (qPCR), respectively. Corresponding Illumina tag sequencing of 16S rRNA genes and 16S rRNA transcripts showed different compositions containing on average 0.7 and 0.9% *Roseobacter*-affiliated OTUs of the DNA- and RNA-based communities. These OTUs were mainly assigned to uncultured members of the *Roseobacter* group. Among those with cultured representatives, *Sedimentitalea* and *Sulfitobacter* made up the largest proportions. The different oceanic provinces with low nutrient content such as both ocean gyres were characterized by specific communities of the *Roseobacter* group, distinct from those of the more productive Pacific subarctic region and the Bering Sea. However, linking the community structure to specific metabolic processes at the seafloor is hampered by the dominance of so-far uncultured members of the *Roseobacter* group, indicating a diversity that has yet to be explored.

**Keywords:** diversity, next-generation sequencing, CARD-FISH, qPCR, RV Sonne

## INTRODUCTION

The *Roseobacter* group within the family *Rhodobacteraceae* consists of nearly 90 genera and approximately 300 species (Pujalte et al., 2014; <http://www.bacterio.net/>). They thrive in a broad variety of marine habitats (Luo and Moran, 2014) and were found free-living in seawater (Giovannoni and Stingl, 2005), associated with micro and macro algae, marine sponges and invertebrates (González et al., 2000; Ivanova et al., 2004) as well as in biofilms, sea ice and sediments (Brinkmeyer et al., 2003; Inagaki et al., 2003). This wide habitat range is attributed to the broad metabolic versatility of different members within the group (Wagner-Döbler and Biebl, 2006). Most of the *Roseobacter*-affiliated bacteria are characterized as ecological generalists (Newton et al., 2010). However, some *Roseobacter* species are specialized to e.g., aerobic anoxygenic phototrophy, sulfur transformations, aromatic compound degradation or secondary metabolite production (Shiba, 1991; González et al., 1996, 1999; Brinkhoff et al., 2004). Due to their metabolic flexibility, the *Roseobacter* group can contribute in high proportions to the bacterial community composition in various marine habitats. For instance, they can account for up to 16% of bacterioplankton communities in polar and temperate waters (Selje et al., 2004). In a North Atlantic algal bloom, on average 23% of all 16S rRNA genes were affiliated to members of the *Roseobacter* group (González et al., 2000). Even subclasses within this group, e.g., the *Roseobacter* clade-affiliated cluster, can account for 36% of bacterial communities in coastal Antarctic regions (Giebel et al., 2009).

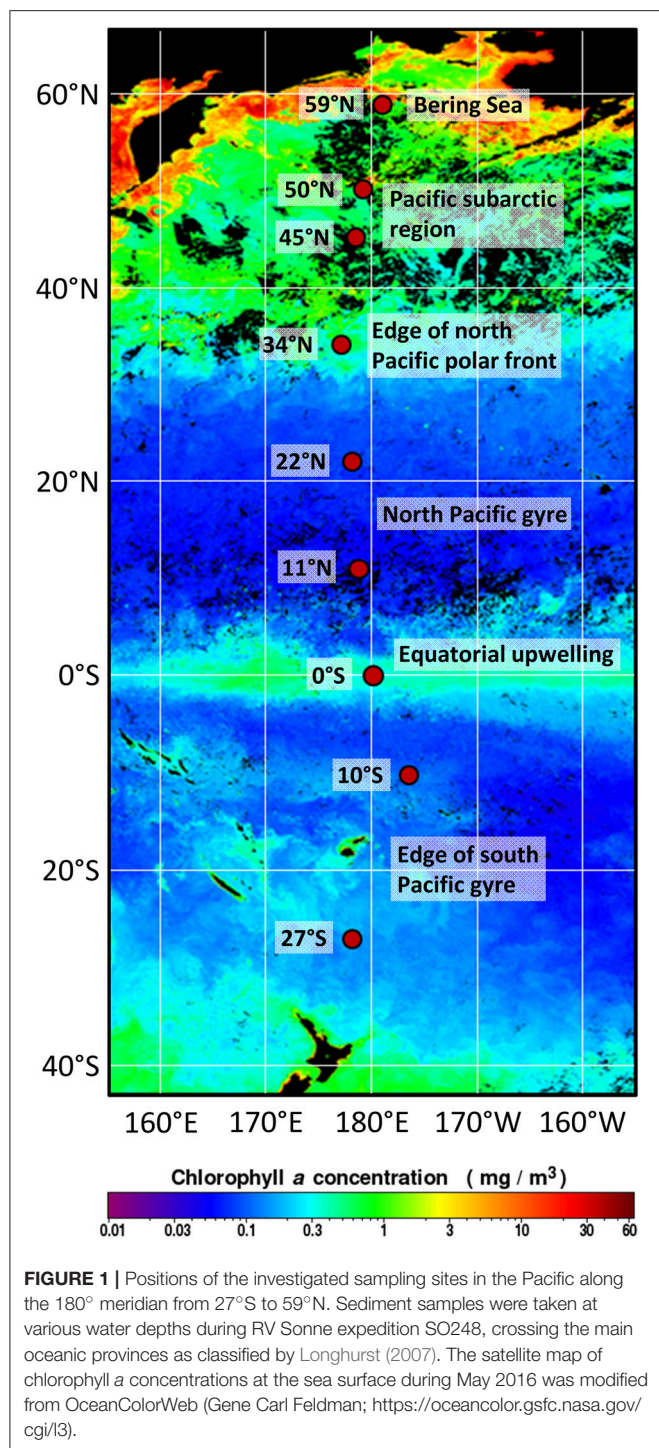
However, most studies on the abundance and diversity of the *Roseobacter* group were conducted on pelagic samples (Giebel et al., 2011; Wemheuer et al., 2014; Billerbeck et al., 2016; Zhang et al., 2016). In contrast, the distribution of this group in sediments is far less understood, even though 28% of all described *Roseobacter*-affiliated species which were known in 2014 are of benthic origin (Pujalte et al., 2014). This may be due to the fact that many studies in sediment microbiology focus on specific metabolic processes such as fermentation, nitrogen cycling, sulfate reduction, methanogenesis, or anaerobic methane oxidation (Orphan et al., 2001; Llobet-Brossa et al., 2002; Mills et al., 2008; Graue et al., 2012b). In diversity studies on marine sediments, the *Roseobacter* group is frequently neglected as their relative proportion on the benthic communities is often <10%. However, direct quantifications of the *Roseobacter* group at the sediment surface of North Sea tidal flats by CARD-FISH showed that their cell numbers exceed those in the pelagic environment by a factor of 1000 (Lenk et al., 2012). In reports on other coastal sediments, the *Roseobacter* group accounts for 1 to 4% of the entire bacterial community (Buchan et al., 2005; Kanukollu et al., 2016) or even 10% in brackish river sediments (González et al., 1999). Assuming that surface sediments contain approximately  $10^9$  cells  $\times$  cm<sup>-3</sup>, their absolute abundance is still orders of magnitudes higher than in corresponding water samples, which usually exhibit  $10^6$  cells  $\times$  ml<sup>-1</sup> (Kallmeyer et al., 2012). The *Roseobacter* group does not only differ numerically between benthic and

pelagic systems, but also in the community structure (Stevens et al., 2005). In a recent study on the distribution of the *Roseobacter* group in coastal North Sea sediments and water samples, we have shown that the diversity within the group increases from the sea surface to the seafloor revealing specific compositions of free-living and attached fractions (Kanukollu et al., 2016).

As all other benthic studies on the *Roseobacter* group were performed on relatively eutrophic coastal sites or high-energy systems such as hydrothermal vents (Buchan et al., 2005), we now describe their abundance, distribution and diversity in oligotrophic deep-sea sediments. In this study, we investigated sampling sites of a Pacific transect (**Figure 1**) spanning six distinct oceanic provinces as defined by Longhurst (2007). These provinces are principally formed by ocean circulation patterns leading to varying nutrient concentrations in the water column. The availability of nutrients directly effects the phytoplankton composition and pelagic bacterial communities (Longhurst, 2007). For instance, the north and south Pacific gyres are established by circular currents that cut off these oceanic regions from a continental nutrient inflow. These oligotrophic conditions lead to low primary production, extremely low sedimentation rates of organic matter and therefore to decreasing cell numbers and associated microbial activities at the seafloor (Rullkötter, 2006; Kallmeyer et al., 2012; Røy et al., 2012). Between the mid-ocean gyres, the upwelling along the equator transports nutrients to the sea surface and thereby stimulates primary production in the photic zone (Schulz and Zabel, 2006). Further north, within the Pacific subarctic region, water column productivity increases toward the Bering Sea. In the highly productive Bering Sea, proximity to land in combination with the coastal runoff leads to increased phytoplankton activities (Longhurst, 2007). The stimulation of primary production causes high organic matter sedimentation and therefore elevated nutrient concentrations at the seafloor (Schulz and Zabel, 2006; Wehrmann et al., 2011).

Our transpacific survey now offers the opportunity to investigate the *Roseobacter* group in deep-sea sediments exhibiting a wide range of environmental conditions from low to relatively high nutrient concentrations. We hypothesize a distinct biogeographical distribution of benthic *Roseobacter* group members corresponding to the composition and nutrient availability within the sediments and primary production in the water column. Thus, sediments were sampled during RV Sonne expedition SO248 along the 180° meridian covering the main oceanic provinces of the central Pacific. In total, seafloor samples (0–1 cm below seafloor, cmbsf) were collected at nine sites from 27°S to 59°N exhibiting water depths of 3,258–5,909 m. Total cell counts were compared to the specific abundance of *Bacteria* and the *Roseobacter* group as quantified by CARD-FISH and qPCR. Next-generation sequencing of 16S rRNA genes and 16S rRNA transcripts revealed deeper insights into the diversity and distribution of the *Roseobacter* group within Pacific sediments. Finally, sediments were characterized by their geochemical composition to correlate the distribution patterns of *Roseobacter* group members to the different environmental settings.





## MATERIALS AND METHODS

### Origin and Sampling of Sediments

Sediment samples were collected in May 2016 during RV Sonne expedition SO248 along a transect from Auckland, New Zealand to Dutch Harbor, Alaska, USA (Figure 1). The transect included very productive as well as highly oligotrophic regions in the

**TABLE 1 |** Origin of sediment samples.

Site	Oceanic province	Latitude	Longitude	Depth (mbsl)	Sum chlorophyll (mg/m <sup>3</sup> )
27°S	Edge of south Pacific gyre	26° 60' S	178° 14' E	4185	42
10°S	Edge of south Pacific gyre	10° 20' S	176° 29' W	4130	88
0°S	Equatorial upwelling	0° 2' S	179° 59' W	5285	102
11°N	North Pacific gyre	10° 58' N	179° 0' E	5402	67
22°N	North Pacific gyre	21° 58' N	178° 19' E	3258	49
34°N	Edge of north Pacific polar front	33° 58' N	177° 21' E	3486	52
45°N	Pacific subarctic region	45° 0' N	178° 45' E	5909	123
50°N	Pacific subarctic region	49° 60' N	179° 33' E	5621	121
59°N	Bering Sea	58° 55' N	178° 56' W	3300	119

The oceanic provinces and water depths are indicated. Chlorophyll data were summed up for the top 500 m of the water column.

Pacific Ocean and covered the main oceanic provinces along the 180° meridian from 27°S to 59°N (Table 1). Sediment cores were taken by a multicorer (Octopus, Germany) at nine sites, exhibiting water depths between 3,258 and 5,909 meter below sea level (mbsl). Subsampling of the sediment surface (0–1 cmbsf) was performed using sterile, cutoff syringes, and rhizons for porewater collection (Rhizosphere, Netherlands; Seeberg-Elverfeldt et al., 2005). Porewater was stored at –20°C, while samples for DNA and RNA extraction were frozen at –80°C for molecular analysis. Samples for cell counting (0.5 cm<sup>3</sup>) and CARD-FISH (1 cm<sup>3</sup>) were fixed for at least 1 h with 3% glutaraldehyde or 3% formaldehyde following the protocols of Lunau et al. (2005) and Ravenschlag et al. (2001), respectively. Afterwards, samples were centrifuged at 13,000 rpm for 1 min and washed two times with 1x TAE-buffer (40 mM NaCl, 1 mM Na<sub>2</sub>EDTA, pH 7.4 adjusted with acetic acid). Finally, samples were stored in TAE:EtOH (1:1) at –20°C.

### Measurement of Chlorophyll Concentrations

During the expedition, chlorophyll concentrations were measured using a CTD-rosette equipped with a fluorometer (FluoroWetlabECO\_AFL\_FL, SN: FLNTURTD-4111). Data were recorded and stored using the standard software Seasave V 7.23.2 and processed by means of ManageCTD. The sum of chlorophyll measured in the top 500 m of the water column was used to characterize the productivity in the water column at the different sampling sites along the transect. Raw data are available at PANGAEA: <https://doi.pangaea.de/10.1594/PANGAEA.864673> (Badewien et al., 2016).

### Geochemical Analyses of Bulk Sediments and Porewaters

Sediments were freeze-dried (Beta 1-8 LDplus, Christ, Germany) and ground as well as homogenized in a Mixer Mill MM 400 (3 Hz, 50 min; Retsch, Germany). Total carbon (TC) and sulfur (S) was analyzed with an elemental analyzer (Eltra CS-800, Germany) with a precision and accuracy of <3% (1σ). Inorganic carbon (IC) was analyzed using an acidification



module (Eltra CS-580) and the content of total organic carbon (TOC) was calculated by difference (TC–IC). Major and trace element analysis in sediments were performed by wavelength dispersive X-ray fluorescence (XRF, Panalytical AXIOS plus) on fused borate glass beads (detailed method in data repository of Eckert et al., 2013). For XRF measurements, accuracy and precision were tested with several SRM's being <4 rel-% (1 $\sigma$ ). Nutrient concentrations of ammonium (NH<sub>4</sub>), nitrate (NO<sub>3</sub>), phosphate (PO<sub>4</sub>) and silicic acid in porewaters were determined photometrically using the Multiscan GO Microplate Spectrophotometer (Thermo Fisher Scientific, USA). The method for measuring ammonium was modified after Benesch and Mangelsdorf (1972), NO<sub>3</sub> was quantified with the method described by Schnetger and Lehnert (2014). Concentrations of PO<sub>4</sub> and silicic acid were determined following the protocol of Grasshoff et al. (1999). Precision and accuracy were tested using solutions of known analyte concentrations (independently prepared) and were <10% (1 $\sigma$ ).

## Total Cell Counts

To compare total cell counts to previous studies on the abundance of benthic prokaryotes (Kallmeyer et al., 2012; Engelhardt et al., 2014), counting was conducted by epifluorescence microscopy using SYBR Green I as fluorescent dye. The protocol was performed according to Graue et al. (2012a) with slight modifications. Those included sonication (three times for 1 min followed by 1 min cooling on ice to prevent overheating), settling of the sample for 1 min to remove sand grains and an initial 10x to 50x dilution of the supernatant with TAE:EtOH (1:1). 10  $\mu$ l of the diluted sample were dispensed on a microscopic slide and dried. A SybrGreen I staining solution was prepared using 5  $\mu$ l of the concentrated SybrGreen I stock solution (Molecular Probes, Eugene, USA) diluted in 200  $\mu$ l moviol mounting medium (Lunau et al., 2005). A freshly prepared 1 M ascorbic acid solution, dissolved in TAE buffer was added at a final concentration of 1% as an antioxidant. Cells were stained with 8  $\mu$ l of the SybrGreen I staining solution. Counting was performed using an epifluorescence microscope (Zeiss, Germany) and twenty randomly selected fields were counted for each sediment sample.

## Catalyzed Reported Deposition-Fluorescence *in situ* Hybridization (CARD-FISH)

CARD-FISH quantification was performed as described by Pernthaler et al. (2002). First, 200  $\mu$ l sediment slurry was diluted 1:4 with 1x PBS (phosphate-buffered saline, 137 mM NaCl, 2.7 mM KCl, 19 mM NaH<sub>2</sub>PO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4) and sonicated for 15 min at 35°C. The sample was further diluted 1:200 in 1x PBS, filtered through a 0.2  $\mu$ m filter (Nucleopore, Whatman) and washed with 30 ml of sterile 1x PBS. Hybridization was carried out with 35% formamide at 46°C. The horseradish peroxidase-labeled *Roseobacter*-specific probe Roseo536 (5'-CAA CGC TAA CCC CCT CCG-3' plus competitor; Brinkmeyer et al., 2000) was diluted 1:100 in hybridization buffer. Samples were counterstained using 3  $\mu$ l

of DAPI (4',6-diamidino-2-phenylindole) and stored at –20°C until microscopic analysis. Microscopic images were taken semi-automatically at 55 randomly chosen spots using an epifluorescence microscope (AxioImager.Z2m, software package AxioVisionVs 40 V4.8.2.0; Carl Zeiss, Germany). Signals were counted using the automated image analysis software ACMEtool3 (M. Zeder; www.technobiology.ch).

## Extraction of Nucleic Acids

DNA extraction from sediment samples was performed using the DNeasy PowerSoil Kit (Qiagen, Germany) according to manufacturer's instructions. DNA was extracted from 0.5 cm<sup>3</sup> sediment and eluted from the columns using 30  $\mu$ l of PCR-grade water. RNA was extracted from 1 cm<sup>3</sup> sediment using the AllPrep DNA/RNA Mini Kit (Qiagen, Germany) following the manufacturer's protocol with some modifications prior to the disruption and homogenization of the cells: an equal volume (1 ml) of Bacterial RNA protect (Qiagen, Germany) was mixed with the sediment by vortexing for 5 s to prevent RNase-digestion. After incubation for 5 min and centrifugation (15 min, 10,000 rpm), the supernatant was discarded. Then, 1 g of zirconia-silica beads (diameter: 0.1 mm; Roth, Germany) and 700  $\mu$ l of RLT-buffer (amended with 0.7  $\mu$ l  $\beta$ -mercaptoethanol) were added to the pellet. The cells were mechanically disrupted by vigorously shaking using a Mini-Beadbeater (BiospecProducts, USA) for 90 s. The homogenized mixture was centrifuged for 3 min at 13,400 rpm and the supernatant was transferred to an AllPrep DNA spin column provided with the kit. After centrifugation for 30 s at 13,400 rpm, the flow through was mixed with 470  $\mu$ l of ice cold absolute ethanol and used for RNA purification on the RNeasy Mini spin column following the original protocol. Finally, the RNA was eluted in 40  $\mu$ l of PCR-grade water. The concentration and purity of the DNA and RNA extracts was determined spectrophotometrically (Nanodrop 2000c, Thermo Fisher Scientific, USA).

## Quantification of 16S rRNA Gene Targets

To determine the amount of 16S rRNA gene targets in the sediment samples, quantitative PCR (qPCR) was performed using the DyNAmo HS SYBR Green qPCR Kit (Thermo Fisher Scientific, USA) and the Light Cycler480II (Roche, Germany). The reaction was performed in 25  $\mu$ l setups: 12.5  $\mu$ l 2x master mix (supplied by the kit), 0.5  $\mu$ l BSA (10 mg/ $\mu$ l), 0.5  $\mu$ l of forward and 0.5  $\mu$ l of reverse primer (each 10 pmol/ $\mu$ l), 1  $\mu$ l PCR-grade water and 10  $\mu$ l of template DNA (1:100 and 1:200 dilutions for EUB, 1:100 dilution for Roseo). For quantification of *Bacteria* the specific primer pair 519f/907r (Lane, 1991; Muyzer et al., 1995) was used. The *Roseobacter* group was quantified by the *Roseobacter*-specific primers Roseo536f (reverse complementary to Roseo536r; Brinkmeyer et al., 2000) and GRb735R (Giuliano et al., 1999). Cycler settings for the *Bacteria*-specific qPCR included an activation step for 15 min at 95°C, followed by 50 cycles of denaturation for 10 s at 94°C, annealing for 20 s at 55°C and elongation for 30 s at 72°C. Afterwards a constant temperature was set for 2 min at 50°C with a subsequent melting curve analysis from 50 to 99°C (15 s/10 s). For *Roseobacter*-quantification, a two-step qPCR was performed as described

above with the same time intervals, but different annealing temperatures (annealing at 66°C for 5 cycles and 63°C for 50 cycles). All samples were analyzed in four independent runs with triplicates, each. Quantification standards were produced as described by Süß et al. (2004) using genomic DNA of *Dinoroseobacter shibae* as template. The resulting 16S rRNA gene amplicons were purified by the QIAquick PCR purification kit according to the manufacturer's instructions (Qiagen, Germany).

## Next-Generation Sequencing

The bacterial diversity was analyzed by next-generation sequencing. While the total bacterial community was assessed by 16S rRNA gene sequencing, the potentially active fraction was determined by an RNA-based approach analyzing 16S rRNA transcripts. For RNA analysis, first a DNase digestion and purification was performed as described by Schneider et al. (2017) and cDNA was generated from DNA-free RNA by SuperScript III (Thermo Fisher Scientific, USA) reverse transcription as described by Wemheuer et al. (2015) using the reverse primer S-D-Bact-0785-a-A-21 without MiSeq adapter (5'-GAC TAC HVG GGT ATC TAA TCC-3'; Klindworth et al., 2013). Extracted DNA was treated with RNase A as described by Schneider et al. (2017) and purified using the GeneRead Size Selection kit (Qiagen, Germany) according to manufacturer's instructions with one modification: samples were eluted in two steps, first using DEPC-treated water and subsequently elution buffer (supplied by the kit).

16S rRNA amplicon libraries were generated from DNA and cDNA by PCR using Phusion Polymerase as described by Wemheuer et al. (2015) with the forward primer S-D-Bact-0341-b-S-17 (5'-CCT ACG GGN GGC WGC AG-3') and the reverse primer S-D-Bact-0785-a-A-21 (5'-GAC TAC HVG GGT ATC TAA TCC-3'; both Klindworth et al., 2013) with Illumina Nextera adapters for sequencing. Each sample was subjected to three independent amplifications and pooled in equal amounts. Amplification products of cDNA were purified using NucleoMag NGS Clean-up and Size Select (Macherey-Nagel, Germany). Both, DNA and cDNA samples were quantified with a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, USA) and barcoded using the Nextera XT-Index kit (Illumina, USA) and the Kapa HIFI Hot Start polymerase (Kapa Biosystems, USA). Sequencing was performed at the Goettingen Genomics Laboratory on an Illumina MiSeq System (paired end 2 × 300 bp) using the MiSeq Reagent Kit v3 (Illumina, USA).

## Processing and Analysis of Illumina-Datasets

Generated datasets of 16S rRNA genes and transcripts were processed according to Granzow et al. (2017). The Trimmomatic version 0.32 (Bolger et al., 2014) was initially used to truncate low quality reads if quality dropped below 20 in a sliding window of 10 bp. Datasets were subsequently processed with Usearch version 8.0.1623 (Edgar, 2010) as described in Wemheuer and Wemheuer (2017). In brief, paired-end reads were merged and quality-filtered. Filtering included the

removal of low quality reads (maximum number of expected errors >1 and more than 1 ambiguous base, respectively) and those shorter than 200 bp. Processed sequences of all samples were joined and clustered in operational taxonomic units (OTUs) at 3% genetic divergence using the UPARSE algorithm implemented in Usearch. A *de novo* chimera removal was included in the clustering step. All OTUs consisting of one single sequence (singletons) were removed. Afterwards, remaining chimeric sequences were removed using the Uchime algorithm in reference mode with the most recent RDP training set (version 15) as reference dataset (Cole et al., 2009). Afterwards, OTU sequences were taxonomically classified using QIIME (Caporaso et al., 2010) by BLAST alignment against the SILVA database (SILVA SSURF 128 NR) and the QIIME release of the UNITE database (version 7.1; August 2016), respectively. All non-bacterial OTUs were removed based on their taxonomic classification in the respective database. Subsequently, processed sequences were mapped on OTU sequences to calculate the distribution and abundance of each OTU in every sample. Sequence data were deposited in the sequence read archive (SRA) of the National Center for Biotechnology Information (NCBI) under the accession numbers SRR5740175–SRR5740185 and SRR5740189–SRR5740196 (BioProject PRJNA391422). Details can be found in the Table S1.

## Phylogenetic Analysis

To display the next relatives of the OTUs identified as “uncultured,” the consensus sequences of these OTUs were aligned with the integrated aligner of ARB (version 6.0.2; Ludwig et al., 2004). Afterwards the sequences were added to the SILVA-backbone tree (SSURF NR99 128) by the maximum-parsimony method using the Quick-Add function in ARB. Genera outside the family *Rhodobacteraceae* and the genera which were not related to the defined OTUs were removed to simplify the tree. Fifty sequences of *Rhizobium* sp. served as root.

## Statistical Data Analysis

Statistical analyses were performed in R (version 3.4.0; R Core Team, 2015). Differences were considered statistically significant with  $p \leq 0.05$ . Environmental data (including concentrations of TOC, S, NO<sub>3</sub>, NH<sub>4</sub>, PO<sub>4</sub>, silicic acid, Fe<sub>2</sub>O<sub>3</sub>, and MnO<sub>2</sub> as well as chlorophyll concentrations) were normalized using the “scale” function in R. Non-metric multidimensional scaling (NMDS) was done with GUniFrac (version 1.0; Chen et al., 2012) and weight parameter  $\alpha = 0$ . A rooted phylogenetic tree for the analysis was generated with muscle (version 3.8.31). Environmental fit was calculated using the vegan package (version 2.4-3; Oksanen et al., 2017). The NMDS plots were generated with the standard plot function in R. Significant environmental factors ( $p \leq 0.05$ ) were plotted onto the graph as arrows. Members of the *Roseobacter* group as well as the oceanic provinces were clustered by similarity in standard R heat maps. All files and scripts used to generate the plots can be found in the supplements (data sheet 1). Image editing was done in inkscape (version 0.91).

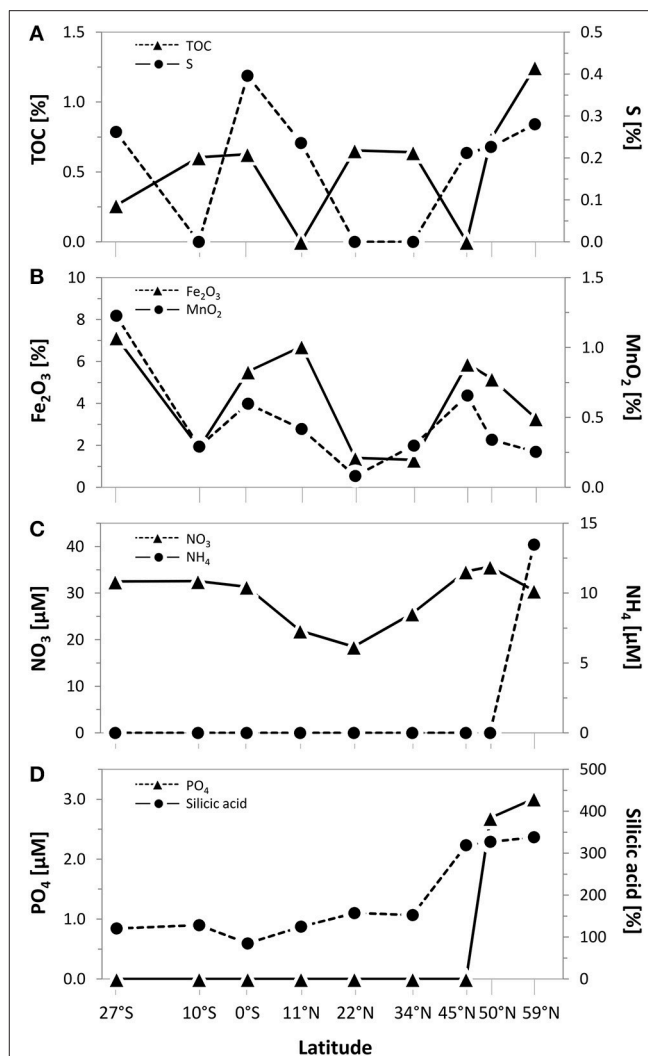
## RESULTS

### Nutrient Concentrations Increase along the Transect toward the Northern Pacific

Biogeographical oceanic provinces as defined by Longhurst (2007) are characteristic by varying nutrient regimes. The chlorophyll concentrations (summed over the upper 500 m; **Table 1**), reflecting the activity in the water column, were lowest at the sampling sites associated to both Pacific gyres (27°S and 22°N), while the amount of chlorophyll was twice as high at the equator (102 mg/m<sup>3</sup>). At the northernmost sites located in the Pacific subarctic region and the Bering Sea highest amounts were found (approximately 120 mg/m<sup>3</sup>). Within the sediments, the total organic carbon (TOC) content as a general indicator for nutrient availability fluctuated around 0.5% of the sediment dry weight, showing an opposing trend with sulfur at most sites of the transect. A maximum TOC content of 1.3% was found in the Bering Sea at 59°N (**Figure 2A**). Iron(III) oxide and manganese dioxide showed an equal trend and were increased at the southernmost site, the equator and within the Pacific subarctic region (**Figure 2B**). Nitrate concentrations in porewaters fluctuated around 30 μM along the transect, showing a maximum of approximately 35 μM at the sampling sites located in the Pacific subarctic region (45°N and 50°N; **Figure 2C**). Concentrations of ammonia and phosphate were 13.5 μM and 2.7–3.0 μM at the northernmost sites, while the concentrations were below quantification limit at all other sites of the transect. Silicic acid was relatively constant at a level of approximately 130 μM from 27°S to 34°N, followed by an increase to >310 μM in the Pacific subarctic region and remained at this high level further north (**Figure 2D**). In general, most parameters showed an increasing trend toward the north with highest concentrations in the Pacific subarctic region and the Bering Sea.

### Total Cell Numbers at the Seafloor Correlate to Increasing Nutrient Concentrations

In general, SYBR Green I counting revealed total cell numbers of  $1.3 \times 10^9$  to  $1.4 \times 10^{10}$  cells  $\times$  cm<sup>-3</sup> of sediment (**Figure 3A**). Lowest cell numbers were counted at the edge of the south Pacific gyre and within the north Pacific gyre (27°S and 11°N) with a slight increase at the equator. Cell numbers increased further north, showing maximum values in the highly productive Bering Sea. All DAPI counts (**Figure 3B**) that were obtained from counterstaining of the CARD-FISH filters were approximately one order of magnitude lower than the SYBR Green I counts. Even though the variations along the transect were less pronounced, both techniques showed a comparable trend. While lowest DAPI counts were also found at 27°S and 11°N ( $1.0$ – $1.3 \times 10^8$  cells  $\times$  cm<sup>-3</sup>), a maximum was detected in the Bering Sea ( $5.6 \times 10^8$  cells  $\times$  cm<sup>-3</sup>). Molecular quantification by qPCR revealed even lower numbers ranging between  $10^4$  and  $10^6$  bacterial 16S rRNA gene targets per cm<sup>3</sup> of sediment. But, independent of the quantification method, the counts correlated to the nutrient concentrations



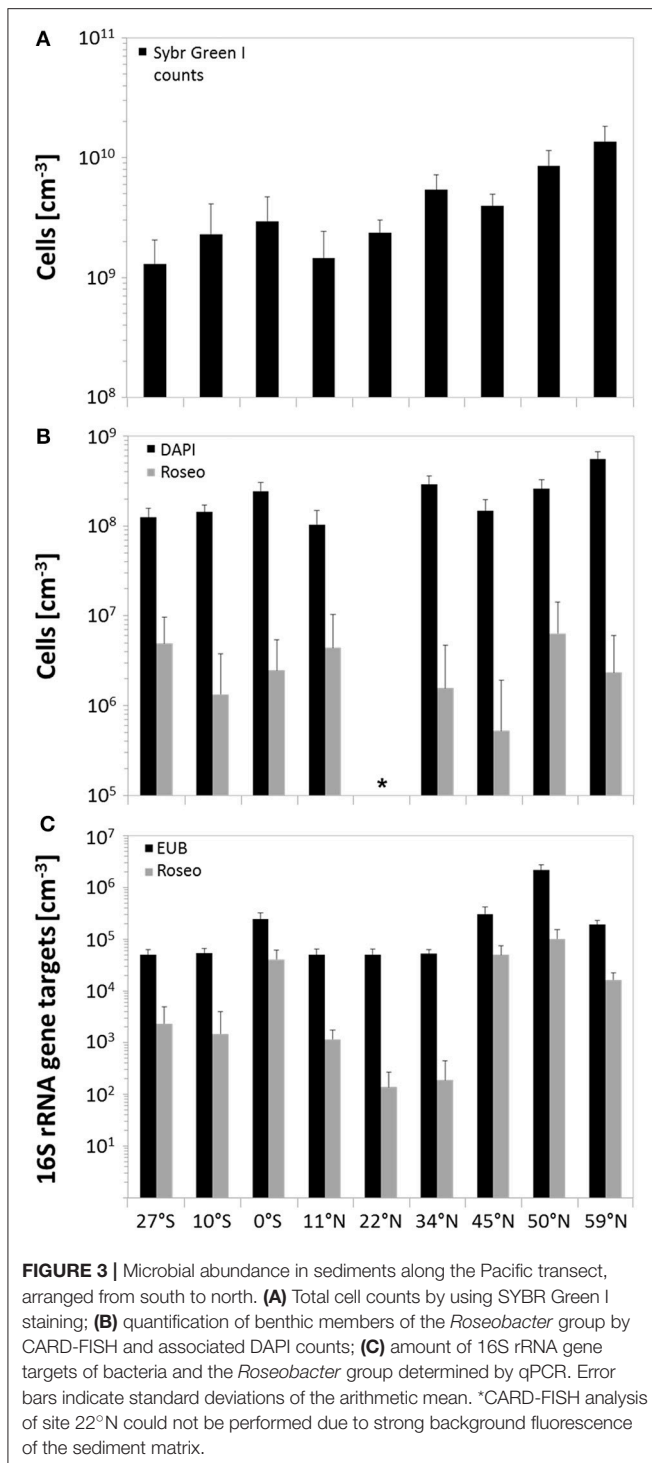
**FIGURE 2 |** Geochemical composition of surface sediments (0–1 cmbsf) along the Pacific transect from 27°S to 59°N. Concentrations of (A) total organic carbon (TOC) and sulfur as well as (B) iron(III) oxide and manganese dioxide in bulk sediments. Porewater concentrations of (C) nitrate and ammonium as well as (D) phosphate and silicic acid. Values plotted as 0 were below quantification limit.

within the sediments and the productivity of the overlying water column with increasing numbers toward the Bering Sea.

### Relative and Direct Quantification of the *Roseobacter* Group Indicates Their Contribution to the Total Bacterial Communities

The relative abundance of the *Roseobacter* group within the total bacterial community can be estimated from the results of next-generation sequencing. While Illumina sequencing of 16S rRNA genes was used to determine the DNA-based bacterial





diversity, the active community was identified by an RNA-based approach targeting 16S rRNA transcripts. Sequencing resulted in a total number of 382,651 sequences over all sampling sites affiliated to 4455 different OTUs (at 3% genetic divergence). Following decreasing phylogenetic levels, *Alphaproteobacteria* accounted for 5% (59°N) to 19% (22°N) of the DNA-based bacterial community and even up to 53%

of the active community at the edge of the south Pacific gyre (27°S). The family *Rhodobacteraceae*, mainly composed of the *Roseobacter* group, contributed up to 1.3% of the DNA-based and 3% of the RNA-based bacterial community (both 45°N). As only one *Paracoccus*-affiliated OTU was detected and the OTU identified as “uncultured V” was probably affiliated to the genus *Halovulum* (Figure S1), the *Roseobacter* group (including all other uncultured representatives) made up similar proportions. It constituted 0.3 to 1.3% of the DNA-based and 0.1 to 3% of the RNA-based community (Table 2). Interestingly, the active community contained a higher proportion of *Roseobacter* group members (average: 0.9%) in comparison to the DNA-based community (average: 0.7%).

As Illumina sequencing only results in estimates of relative numbers, direct quantification of the *Roseobacter* group was performed. Microscopic quantification using a *Roseobacter*-specific CARD-FISH probe revealed an average of  $3 \times 10^6$  cells  $\times$  cm<sup>-3</sup> of sediment, corresponding to a relative abundance of 1.7%. At 11°N, up to 4.3% of all DAPI counts were affiliated to the *Roseobacter* group, while at 45°N this group only made up a proportion of 0.4% (Figure 3B). Molecular analysis by *Roseobacter*-specific qPCR mostly exhibited lower values compared to CARD-FISH quantification. Quantitative PCR revealed 10<sup>2</sup> to 10<sup>5</sup> 16S rRNA gene targets per cm<sup>3</sup> of sediment along the transect, showing highest values at the equator and 45°N (Figure 3C). On average, *Roseobacter* gene targets accounted for 6.3% of all bacterial 16S rRNA genes.

### The Community Composition of the *Roseobacter* Group Is Dominated by Uncultured Representatives

Each sampling site was characterized by individual community patterns. Remarkably, out of the 15 different *Roseobacter*-affiliated OTUs, eight of them could not be assigned to known genera and were classified as “uncultured” during the processing of the Illumina dataset (Figure 4, blue bars “uncultured I–VIII”). Phylogenetical analysis (Figure S1) showed that the OTUs identified as “uncultured I–III” interestingly clustered with each other and did not show any relation to known genera within the family *Rhodobacteraceae*. This indicates the lack of isolates and the amount of unexplored diversity within the *Roseobacter* group. Notably, the OTU assigned to “uncultured I” made up the highest proportion within the DNA-based as well as in the RNA-based community (Figure 4). The next cultured relatives to the OTU “uncultured IV” were within the genus *Loktanella*, while the OTU assigned to “uncultured VI” clustered with the genus *Pacifibacter*. The OTUs affiliated to “uncultured VII and VIII” were related to the genera *Litorimicrobium* and *Ruegeria*, respectively. The OTU classified as “uncultured V” was distantly related to the genus *Halovulum*, a member of the *Amaricoccus* group (Figure S1).

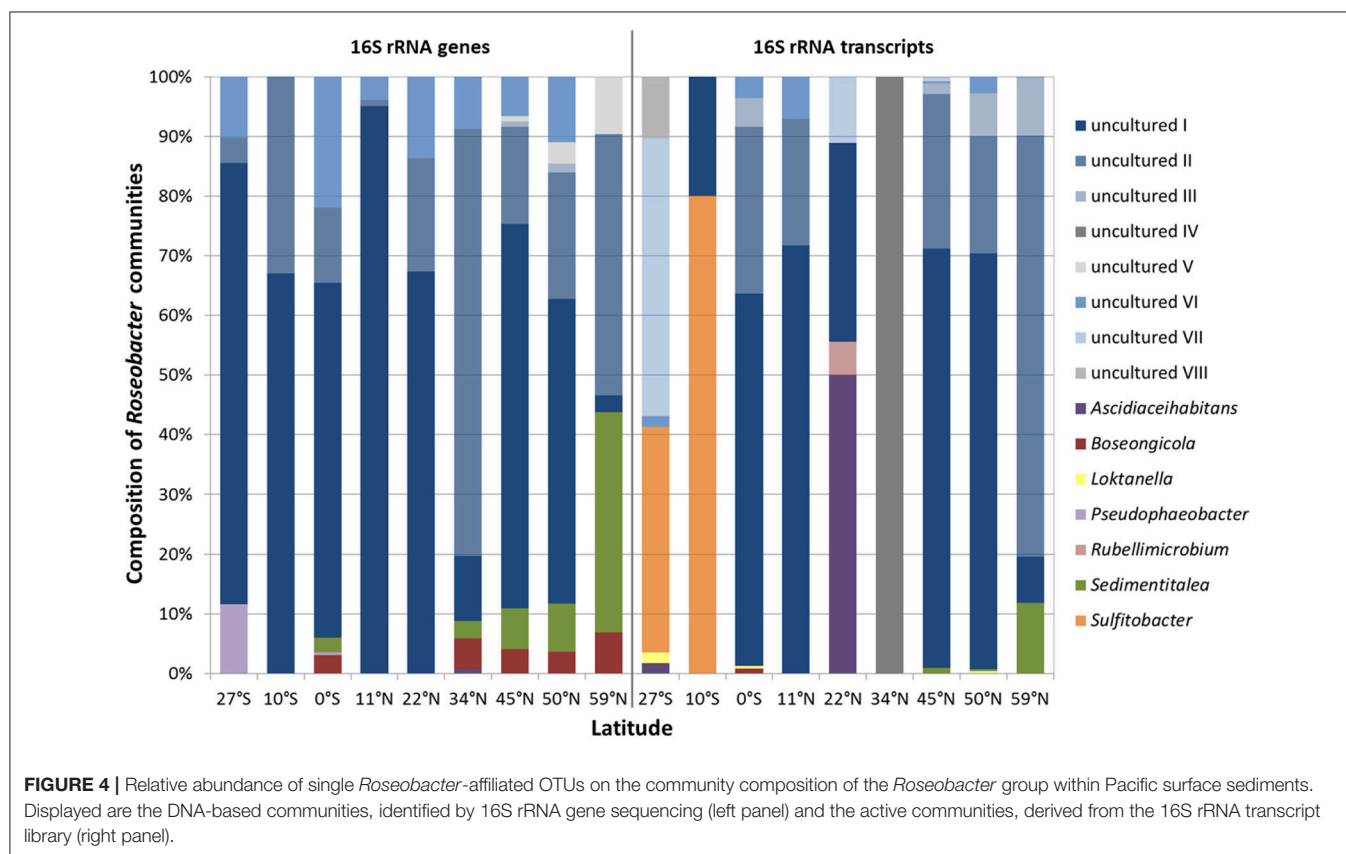
Focusing on the relative abundance of known genera that contribute to the diversity of the *Roseobacter* group within the DNA-based community, especially the genera



**TABLE 2** | Proportion of *Roseobacter*-affiliated OTUs on the total bacterial communities along the Pacific transect.

Site	DNA-based community (16S rRNA genes)			RNA-based community (16S rRNA transcripts)		
	Total OTUs	<i>Roseobacter</i> -affiliated OTUs		Total OTUs	<i>Roseobacter</i> -affiliated OTUs	
		No.	%		No.	%
27°S	25740	69	0.27	21345	58	0.27
10°S	26955	170	0.63	20729	5	0.02
0°S	18856	200	1.06	26743	393	1.47
11°N	22665	102	0.45	25985	99	0.38
22°N	22954	95	0.41	29070	18	0.06
34°N	26421	137	0.52	14637	20	0.14
45°N	24009	318	1.32	22071	662	3.00
50°N	15695	132	0.84	18584	291	1.57
59°N	11025	66	0.60	9167	102	1.11

The complete dataset contained 382,651 OTUs.



*Sedimentitalea* and *Boseongicola* were found (Figure 4 left panel). Most genera were broadly spread across the transect, however *Pseudophaeobacter* was only detected at the equator and further south. Within the active community, *Loktanella* and *Sedimentitalea* were distributed over one third of all sampling sites, while the latter together with *Sulfitobacter* made up the largest proportions of benthic *Roseobacter* group members with cultured representatives (Figure 4 right panel).

## Oceanic Provinces That Differ in Primary Production and Nutrient Availability within the Sediments Exhibit Distinct Communities of Benthic Members of the *Roseobacter* Group

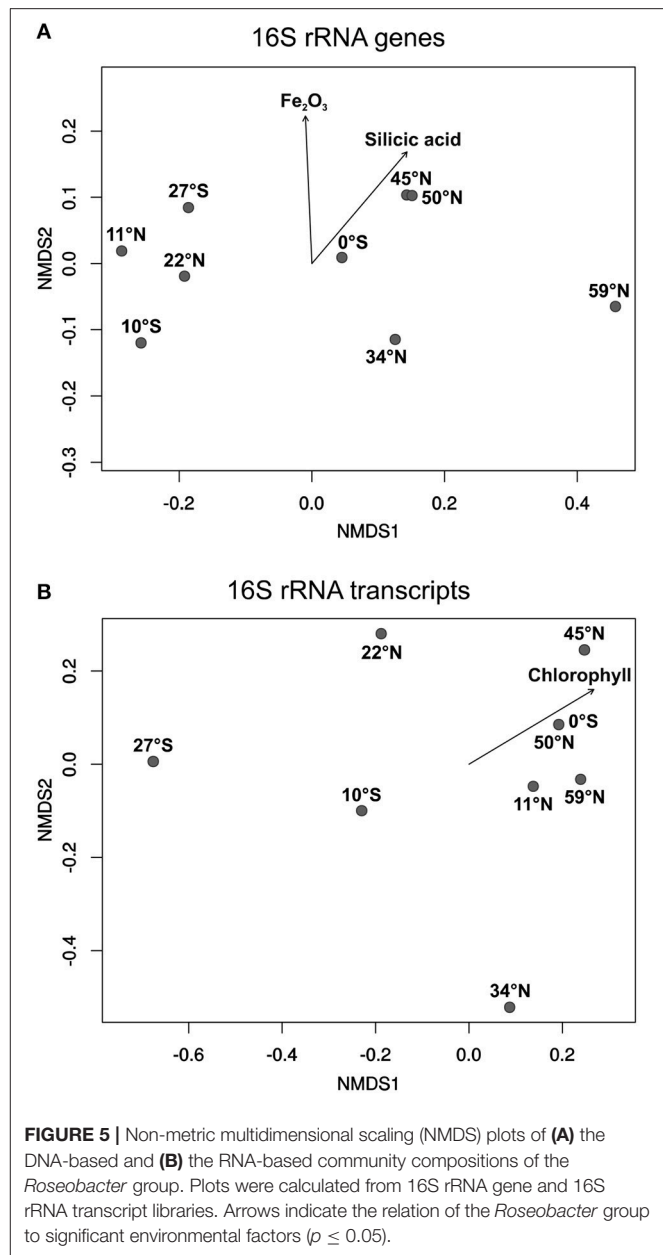
To express the specific distribution patterns, non-metric multidimensional scaling (NMDS) was performed to compare the community composition at the different sampling sites

to the respective environmental parameters. The DNA-based communities of benthic members of the *Roseobacter* group at the various sampling sites clearly clustered separately depending on local nutrient availability in the sediments, and on the primary production in the water column (Figure 5A). Sampling sites from both oligotrophic Pacific gyres (27°S, 10°S, 11°N, 22°N) formed one cluster, while the equator, characterized by higher nutrient content in the sediment and the water column, clustered with all other more eutrophic sampling sites. A most distinct community composition was found at the northernmost sampling site in the Bering Sea, exhibiting elevated nutrient concentrations. The iron(III) oxide content of the sediment as well as the concentrations of silicic acid in the porewater were found to significantly influence the community composition of benthic members of the *Roseobacter* group ( $p \leq 0.05$ ). The varying high  $\text{MnO}_2$  content of the deep-sea sediments indicate oxic environmental conditions at all sites and possibly a hydrothermal contribution at some sites. Surprisingly, this parameter had no influence on the community composition.

All active, RNA-based, communities of the *Roseobacter* group generally showed a less uniform composition compared to the DNA-based communities (Figure 5B). Apart from the sampling site at 11°N, the other three oligotrophic sites (27°S, 10°S, 22°N) were also separated from the more eutrophic regions. Here, site 34°N was more distinct as it was exclusively composed of the OTU identified as “uncultured IV” (phylogenetically assigned to the genus *Loktanella*). Interestingly, only the chlorophyll content of the overlying water column was identified as a significant environmental control on the diversity of active *Roseobacter* group members.

### The Oligotrophic Gyres and the More Eutrophic Regions Are Distinguished by Different Genera within the DNA- and RNA-Based Fraction of the *Roseobacter* Group Communities

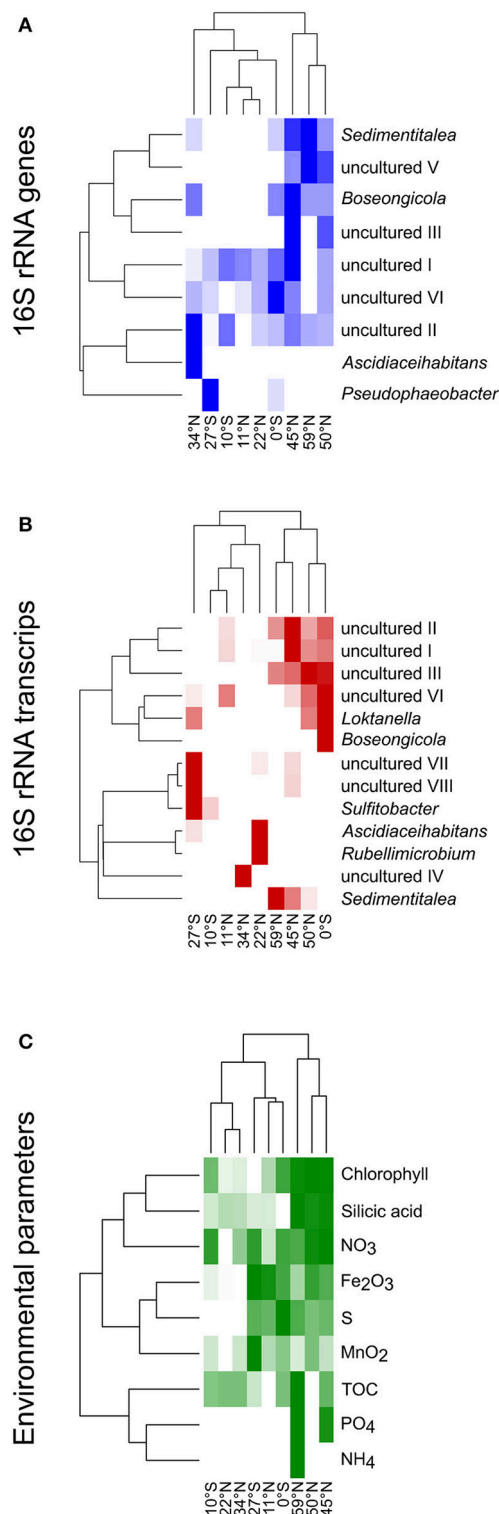
Heat maps were calculated for a more detailed view on the contribution of single genera to the community composition of benthic members of the *Roseobacter* group at the different sampling sites (Figure 6). Both, the DNA- and RNA-based communities of the sites in the North Pacific gyre (11°N and 22°N) clustered together with the site at 10°S, while the northernmost sites (45°N, 50°N, 59°N) formed another cluster that branched separately. In contrast, the equator and the sites at the edge of the south Pacific gyre (27°S) and at the north Pacific polar front varied in their assignment. In the DNA-based communities of the *Roseobacter* group, the branching of the northernmost sites was mainly due to the presence of the OTUs affiliated to *Sedimentitalea*, *Boseongicola* and “uncultured V” (Figure 6A). The DNA-based communities of the North Pacific gyre and the site at 10°S were in turn characterized by OTUs assigned to “uncultured I and II,” the OTU “uncultured VI” (phylogenetically affiliated to *Pacificbacter*) and the absence of other OTUs. *Ascidiaiehabitans*



**FIGURE 5 |** Non-metric multidimensional scaling (NMDS) plots of (A) the DNA-based and (B) the RNA-based community compositions of the *Roseobacter* group. Plots were calculated from 16S rRNA gene and 16S rRNA transcript libraries. Arrows indicate the relation of the *Roseobacter* group to significant environmental factors ( $p \leq 0.05$ ).

and *Pseudophaeobacter*-affiliated OTUs only contributed to the DNA-based community composition of the varying sites at 27°S, 0°S, and 34°N.

The active community compositions were generally more diverse than those of the DNA-based communities (Figure 6B). Here, the equatorial community branched with the communities of the northernmost sites. While all of these sites were characterized by OTUs assigned to some uncultured representatives (“uncultured I-III and VI”), *Sedimentitalea* and *Boseongicola*-affiliated OTUs were only found at the three northernmost sites and the equator, respectively. Branching of both Pacific gyres and site 34°N was mainly due to the presence and absence of single OTUs (e.g., *Rubellimicrobium*, *Sulfitobacter*,



**FIGURE 6 |** Heat maps for the different sampling sites according to the compositions of **(A)** the DNA-based, **(B)** the RNA-based community composition of the *Roseobacter* group and **(C)** nutrients in sediments and porewaters as well as chlorophyll in the water column. The clustering on the left side reflects the relative frequencies of the different OTUs/nutrient concentrations, while on top the clustering based on the similarity of the sampling sites is shown.

*Ascidiahabitans* or “uncultured IV”, phylogenetically affiliated to *Loktanella*).

The heat map for single environmental parameters in relation to the sampling sites showed a slightly different branching of the investigated sites (**Figure 6C**). Here, the sites at 10°S and 22°N clustered with the site located at the edge of the north Pacific polar front due to low concentrations of all nutrients analyzed. The northernmost sites branched together, but separate of all other sites, as the Pacific subarctic region and the Bering Sea are characterized by elevated nutrient concentrations.

## DISCUSSION

### Microbial Abundance within the Sediments Is Related to Nutrient Availability and Primary Production in the Different Oceanic Provinces

Organic matter input and quality, but also water depth and marine productivity in the water column, influence the cell abundance within sediments. Distinct oceanic provinces were crossed along the Pacific transect that are characterized by varying primary production, resulting in changing nutrient availability at the seafloor (Longhurst, 2007). The *in situ* chlorophyll measurements (summed over the upper 500 m) were higher compared to the concentrations determined by satellite imaging for the time period of the expedition (**Figure 1**) as the satellite only detects surface chlorophyll. The results of the satellite imaging are very useful for the classification of the oceanic provinces, but the *in situ* measurements include the deep chlorophyll maximum and thus reflect the total primary production.

Variations in benthic cell numbers depend on the primary production in the water column and are strongly correlated to sedimentation rates and distance from land (Kallmeyer et al., 2012). In general, total cell numbers determined by SYBR Green I counting of around  $10^9$  cells  $\times$  cm<sup>-3</sup> were in the expected range for seafloor sediments (Parkes et al., 2000; D’Hondt et al., 2004). DAPI counts performed as counterstaining for CARD-FISH quantification showed the same trend, but were approximately one magnitude lower than SYBR Green I counts as already observed by Weinbauer et al. (1998) and Morono et al. (2009). Furthermore, CARD-FISH quantification, targeting the RNA, might be influenced by low numbers of active cells or unspecific fluorescence of the sediment matrix, as e.g., observed for the site located at 22°N. This bias can partially be compensated by other quantification attempts such as qPCR or identifying the relative amount of specific OTUs in the Illumina sequence libraries. While qPCR significantly depends on the DNA extraction efficiency and might underestimate total numbers (e.g., sites at 0°N and 45°N), the results of next-generation sequencing should be interpreted as relative values and not as absolute numbers. Thus, each method should be evaluated separately, but trends in abundances may be confirmed (Lloyd et al., 2013).

All quantification methods used to estimate the total amount of bacteria showed a general trend following the primary production in the different oceanic provinces and the nutrient

availability within the sediments. Cell numbers were lowest in sediments of the north Pacific gyre and at the edge of the south Pacific gyre. Both provinces are characterized by low primary production and limiting nutrient concentrations, e.g. nitrate content (Rullkötter, 2006; Longhurst, 2007), resulting in extremely low sedimentation rates (Røy et al., 2012). By all general quantification methods, we confirmed that the upwelling along the equator leads to higher cell numbers at the seafloor due to elevated primary production in the water column. The more pronounced increase in bacterial abundance from the edge of the north Pacific polar front to the Bering Sea also follows the productivity in the water column. The highest cell numbers, which were found in the Bering Sea, are presumably due to the proximity to land in combination with the coastal runoff, leading to increased phytoplankton activities in the water column (Longhurst, 2007; Kallmeyer et al., 2012). The high primary production causes high sedimentation rates and therefore elevated nutrient concentrations at the seafloor (Schulz and Zabel, 2006; Wehrmann et al., 2011). The measured TOC content of the sediments, for instance, was around 1% as also reported by Seiter et al. (2004). In general, as all other nutrient concentrations increased in the northernmost sites, total cell counts followed this trend.

### Pacific Sediments Show Similar Maximum Proportions of the *Roseobacter* Group as Coastal Areas, but Exhibit Lower Average Abundances

As all investigated sampling sites of the Pacific transect exhibited water depths of several thousand meters, organic matter reaching the seafloor is supposed to be more recalcitrant due to microbial degradation while sinking (Martin et al., 1987; Karl et al., 1988). Although only 1% of the primary production reaches the seafloor, 97% of this material is decomposed by microbial activities and returned as dissolved matter to the water column (Zabel and Hensen, 2006). Interestingly, chlorophyll was identified as significant environmental control on the diversity of active members of the *Roseobacter* group (Figure 5B). This proves that not only the nutrient availability in the sediment alone is triggering the distribution of different benthic *Roseobacter* genera, but also the productivity in the water column has major influence.

The fact that the average abundance of *Roseobacter*-affiliated OTUs in our dataset was much lower than the maximum proportions is probably due to the oligotrophic nature of the investigated oceanic provinces. Furthermore, while the Pacific transect includes deep-sea sediments distinct from landmasses, exhibiting low nutrient contents, most other benthic studies were performed in coastal, nutrient-rich sediments (Buchan et al., 2005). However, the relative amount of *Roseobacter*-affiliated OTUs in our 16S rRNA transcript library of up to 3% is comparable to proportions found in coastal sediments from the North Sea and cold seeps in the Nankai Trough, both around 2% of all 16S rRNA genes (Li et al., 1999; Kanukollu et al., 2016), as well as in volcanic sediments of the Sea of Okhotsk (Inagaki et al.,

2003) and Antarctic Shelf sediments (Bowman and McCuaig, 2003). Differences in abundance between coastal and open ocean sites are not only visible at the seafloor. Even in the water column of the North Pacific, the *Roseobacter* group only made up 5% of the community associated to phytoplankton blooms (Tada et al., 2011), representing a lower abundance as described in previous studies on coastal regions (González et al., 2000; Selje et al., 2004; Giebel et al., 2009). The maximum amount of *Roseobacter* group members, as quantified by CARD-FISH (up to 4.3%), is within the range that was previously reported for tidal-flat sediments by Lenk et al. (2012). The proportion of *Roseobacter* group members quantified by qPCR with an average of 6.3% is again in the same range. However, the very high percentages of approximately 16% at sites 0°N and 45°N might be due to an underestimation of the total abundance by the bacteria-specific qPCR.

### The Dominance of Uncultured Members of the *Roseobacter* Group in Sediments Hampers to Directly Link Their Function to Environmental Settings

Almost all of the detected genera that have cultivated representatives were described as aerobic heterotrophs and were previously isolated from coastal marine sediments. For instance, *Sedimentitalea*, which is found in both, the DNA- and RNA-based communities, was originally isolated from sandy sediments of the South China Sea (Sun et al., 2010; Breider et al., 2014). The type strain of *Boseongicola* derived from tidal-flat sediments (Park et al., 2014) and *Loktanella* from a deep subseafloor sediment off the coast of Japan (Van Trappen et al., 2004; Tsubouchi et al., 2013). While *Loktanella* and *Sulfitobacter*-affiliated OTUs were found in coastal North Sea sediments (Kanukollu et al., 2016), none of the above mentioned genera were previously detected in deep-sea sediments of open ocean sites. The presence of *Loktanella* and *Sulfitobacter* in coastal sediments may suggest that they depend on high nutrient concentrations. However, our results show that they also constitute to the community of active *Roseobacter* group members in oligotrophic open-ocean environments.

The Illumina datasets of benthic members of the *Roseobacter* group were composed of OTUs affiliated to the above mentioned species, but mainly of OTUs assigned to “uncultured” (84% on average). The consensus sequences of these OTUs of around 440 bp length, allowed a phylogenetic classification on family level, but the exact affiliation to a specific genus still remains unclear. For a fixed assignment, the complete 16S rRNA sequence or even an isolate is needed. Nevertheless, the database entries of the sequence branching together with OTUs assigned to “uncultured I to III” reveal “Pacific Ocean sediments” as isolation source (e.g., accession numbers JX227630, AJ567557; JX226780, EU491756; KM071672; Figure S1). The large proportion of *Roseobacter*-affiliated OTUs in both sequence libraries that are classified as “uncultured” indicates the presence of a specific and active community of benthic members of the *Roseobacter* group that is well adapted to conditions at the seafloor of the deep sea. These bacteria were not yet cultured



as they might grow very slow, form microcolonies or depend on the co-occurrence of other microorganisms. Furthermore, this proves the assumption that the environmental conditions can hardly be mimicked in the laboratory (Amann et al., 1995; Kaeberlein et al., 2002). Due to a lack of isolates, their physiological characteristics are largely unknown and hamper further interpretation concerning their metabolism in sediments. However, as we have now correlated the environmental settings to the abundance and co-occurrence of single OTUs, specific media for the isolation of so-far uncultured benthic members of the *Roseobacter* group can be designed. Furthermore, the sequence information allows developing primers and probes for the detection and quantification of the different uncultured *Roseobacter* representatives in enrichment cultures and environmental samples.

## AUTHOR CONTRIBUTIONS

MP, JD, and BE performed the sediment sampling during RV Sonne expedition SO248. JD did the molecular quantification (cell counting, CARD-FISH, qPCR) and extraction of nucleic acids. NE did the nutrient measurements, supervised and interpreted by BS. AvH-H prepared samples for Illumina sequencing and did statistical analyses. BW processed the Illumina sequence data. TB collected and provided chlorophyll data. MP and BE wrote the first draft of the manuscript including

figure design. All authors contributed were involved in critical revision and approval of the final version.

## FUNDING

This study was supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the collaborative research center TRR51. The German Federal Ministry of Education and Research (BMBF) generously provided the main funding of the RV Sonne expedition SO248.

## ACKNOWLEDGMENTS

We thank the crew and the scientific party of RV Sonne expedition SO248 for their support during sampling as well as Jana Schmidt and Frank Meyerjürgens for technical assistance. Carola Lehnert is acknowledged for support during evaluation of geochemical data. Furthermore, Insa Bakenhus and Leon Dlugosch are acknowledged for help during CARD-FISH quantification and statistical analysis, respectively.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2017.02550/full#supplementary-material>

## REFERENCES

- Amann, R. L., Ludwig, W., and Schleifer, K. H. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* 59, 143–169.
- Badewien, T.H., Winkler, H., Arndt, K.L., and Simon, M., (2016). *Physical oceanography during SONNE cruise SO248 (BacGeoPac)*. Institute for Chemistry and Biology of the Marine Environment, Carl-von-Ossietzky University of Oldenburg, Germany, Dataset #864673.
- Benesch, R., and Mangelsdorf, P. (1972). Method for colorimetric determination of Ammonia in seawater. *Helgoländer Wissenschaftliche Meeresuntersuchungen* 23, 365–375. doi: 10.1007/BF01609682
- Billerbeck, S., Wemheuer, B., Voget, S., Poehlein, A., Giebel, H. A., Brinkhoff, T., et al. (2016). Biogeography and environmental genomics of the *Roseobacter*-affiliated pelagic CHAB-I-5 lineage. *Nat. Microbiol.* 1:16063. doi: 10.1038/nmicrobiol.2016.63
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Bowman, J. P., and McCuaig, R. D. (2003). Biodiversity, community structural shifts, and biogeography of prokaryotes within Antarctic continental shelf sediment. *Appl. Environ. Microbiol.* 69, 2463–2483. doi: 10.1128/AEM.69.5.2463-2483.2003
- Breider, S., Scheuner, C., Schumann, P., Fiebig, A., Petersen, J., Pradella, S., et al. (2014). Genome-scale data suggest reclassifications in the *Leisingera-Phaeobacter* cluster including proposals for *Sedimentitalea* gen. nov. and *Pseudophaeobacter* gen. nov. *Front. Microbiol.* 5:416. doi: 10.3389/fmicb.2014.00416
- Brinkhoff, T., Bach, G., Heidorn, T., Liang, L., Schlingloff, A., and Simon, M. (2004). Antibiotic production by a *Roseobacter* clade-affiliated species from the German Wadden Sea and its antagonistic effects on indigenous isolates. *Appl. Environ. Microbiol.* 70, 2560–2565. doi: 10.1128/AEM.70.4.2560-2565.2003
- Brinkmeyer, R., Rappé, M., Gallacher, S., and Medlin, L. (2000). Development of clade- (*Roseobacter* and *Alteromonas*) and taxon-specific oligonucleotide probes to study interactions between toxic dinoflagellates and their associated bacteria. *Eur. J. Phycol.* 35, 315–329. doi: 10.1080/09670260010001735921
- Brinkmeyer, R., Knittel, K., Jürgens, J., Weyland, H., Amann, R., and Helmke, E. (2003). Diversity and structure of bacterial communities in arctic versus antarctic pack ice. *Appl. Environ. Microbiol.* 69, 6610–6619. doi: 10.1128/AEM.69.11.6610-6619.2003
- Buchan, A., González, J. M., and Moran, M. A. (2005). Overview of the marine *Roseobacter* lineage. *Appl. Environ. Microbiol.* 71, 5665–5677. doi: 10.1128/AEM.71.10.5665-5677.2005
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303
- Chen, J., Bittinger, K., Charlson, E. S., Hoffmann, C., Lewis, J., Wu, G. D., et al. (2012). Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics* 28, 2106–2113. doi: 10.1093/bioinformatics/bts342
- Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., et al. (2009). The Ribosomal database project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res.* 37, D141–D145. doi: 10.1093/nar/gkn879
- D'Hondt, S., Jørgensen, B. B., Miller, D. J., Batzke, A., Blake, R., Cragg, B. A., et al. (2004). Distributions of microbial activities in deep subseafloor sediments. *Science* 306, 2216–2221. doi: 10.1126/science.1101155
- Eckert, S., Brumsack, H. J., Severmann, S., Schnetger, B., März, C., and Fröhlje, H. (2013). Establishment of euxinic conditions in the Holocene Black Sea. *Geology* 41, 431–434. doi: 10.1130/G33826.1
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461
- Engelhardt, T., Kallmeyer, J., Cypionka, H., and Engelen, B. (2014). High virus-to-cell ratios indicate ongoing production of viruses in deep subsurface sediments. *ISME J.* 8, 1503–1509. doi: 10.1038/ismej.2013.245

- Giebel, H. A., Brinkhoff, T., Zwisler, W., Selje, N., and Simon, M. (2009). Distribution of *Roseobacter*, R. C. A., and SAR11 lineages and distinct bacterial communities from the subtropics to the Southern Ocean. *Environ. Microbiol.* 11, 2164–2178. doi: 10.1111/j.1462-2920.2009.01942.x
- Giebel, H. A., Kallhoefer, D., Lemke, A., Thole, S., Gahl-Janssen, R., Simon, M., et al. (2011). Distribution of *Roseobacter*, R. C. A., and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J.* 5, 8–19. doi: 10.1038/ismej.2010.87
- Giovannoni, S. J., and Stingl, U. (2005). Molecular diversity and ecology of microbial plankton. *Nature* 437, 343–348. doi: 10.1038/nature04158
- Giuliano, L., De Domenico, M., De Domenico, E., Höfle, M. G., and Yakimov, M. M. (1999). Identification of culturable oligotrophic bacteria within naturally occurring bacterioplankton communities of the Ligurian sea by 16S rRNA sequencing and probing. *Microb. Ecol.* 37, 77–85. doi: 10.1007/s002489900132
- González, J. M., Whitman, W. B., Hodson, R. E., and Moran, M. A. (1996). Identifying numerically abundant culturable bacteria from complex communities: an example from a lignin enrichment culture. *Appl. Environ. Microbiol.* 62, 4433–4440.
- González, J. M., Kiene, R. P., and Moran, M. A. (1999). Transformation of sulfur compounds by an abundant lineage of marine bacteria in the alpha-subclass of the class *Proteobacteria*. *Appl. Environ. Microbiol.* 65, 3810–3819.
- González, J. M., Simó, R., Massana, R., Covert, J. S., Casamayor, E. O., Pedros-Alíó, C., et al. (2000). Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl. Environ. Microbiol.* 66, 4237–4246. doi: 10.1128/AEM.66.10.4237-4246.2000
- Granzow, S., Kaiser, K., Wemheuer, B., Pfeiffer, B., Daniel, R., Vidal, S., et al. (2017). The effects of cropping regimes on fungal and bacterial communities of wheat and faba bean in a greenhouse pot experiment differ between plant species and compartment. *Front. Microbiol.* 8:902. doi: 10.3389/fmicb.2017.00902
- Grasshoff, K., Kremling, K., and Ehrhardt, M. (1999). *Methods of Seawater Analysis, 3rd Edn*. New York, NY: Wiley-VC, 159–198.
- Graue, J., Engelen, B., and Cypionka, H. (2012a). Degradation of cyanobacterial biomass in anoxic tidal-flat sediments: a microcosm study of metabolic processes and community changes. *ISME J.* 1–10. doi: 10.1038/ismej.2011.120
- Graue, J., Kleindienst, S., Lueders, T., Cypionka, H., and Engelen, B. (2012b). Identifying fermenting bacteria in anoxic tidal-flat sediments by a combination of microcalorimetry and ribosome-based stable-isotope probing. *FEMS Microbiol. Ecol.* 81, 78–87. doi: 10.1111/j.1574-6941.2011.01282.x
- Inagaki, F., Suzuki, M., Takai, K., Oida, H., Sakamoto, T., Aoki, K., et al. (2003). Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. *Appl. Environ. Microbiol.* 69, 7224–7235. doi: 10.1128/AEM.69.12.7224-7235.2003
- Ivanova, E. P., Gorshkova, N. M., Sawabe, T., Zhukova, N. V., Hayashi, K., Kurilenko, V. V., et al. (2004). *Sulfitobacter delicatus* sp. nov. and *Sulfitobacter dubius* sp. nov., respectively from a starfish (*Stellaster equestris*) and sea grass (*Zostera marina*). *Int. J. Syst. Evol. Microbiol.* 54, 475–480. doi: 10.1099/ijs.0.02654-0
- Kaerberlein, T., Lewis, K., and Epstein, S. S. (2002). Isolating “uncultivable” microorganisms in pure culture in a simulated natural environment. *Science* 296, 1127–1129. doi: 10.1126/science.1070633
- Kallmeyer, J., Pockalny, R., Adhikari, R. R., Smith, D. C., and D'Hondt, S. (2012). Global distribution of microbial abundance and biomass in subseafloor sediment. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16213–16216. doi: 10.1073/pnas.1203849109
- Kanukollu, S., Wemheuer, B., Herber, J., Billerbeck, S., Lucas, J., Daniel, R., et al. (2016). Distinct compositions of free-living, particle-associated and benthic communities of the *Roseobacter* group in the North Sea. *FEMS Microbiol. Ecol.* 92:fiv145. doi: 10.1093/femsec/fiv145
- Karl, D. M., Knauer, G. A., and Martin, J. H. (1988). Downward flux of particulate organic matter in the ocean: a particle decomposition paradox. *Nature* 332, 438–441. doi: 10.1038/332438a0
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41, 11. doi: 10.1093/nar/gks808
- Lane, D. J. (1991). “16S/23S rRNA Sequencing,” in *Nucleic acid Techniques in Bacterial Systematics*, eds M. Goodfellow and E. Stackebrandt (Chichester: Wiley).
- Lenk, S., Moraru, C., Hahnke, S., Arnds, J., Richter, M., Kube, M., et al. (2012). *Roseobacter* clade bacteria are abundant in coastal sediments and encode a novel combination of sulfur oxidation genes. *ISME J.* 6, 2178–2187. doi: 10.1038/ismej.2012.66
- Li, L., Guenzennec, J., Nichols, P., Henry, P., Yanagibayashi, M., and Kato, C. (1999). Microbial diversity in Nankai Trough sediments at a depth of 3,843 m. *J. Oceanogr.* 55, 635–642. doi: 10.1023/A:1007897020042
- Llobet-Brossa, E., Rabus, R., Böttcher, M. E., Könneke, M., Finke, N., Schramm, A., et al. (2002). Community structure and activity of sulfate-reducing bacteria in an intertidal surface sediment: a multi-method approach. *Aquat. Microbial Ecol.* 29, 211–226. doi: 10.3354/ame029211
- Lloyd, K. G., May, M. K., Kevorkian, R. T., and Steen, A. D. (2013). Meta-analysis of quantification methods shows that archaea and bacteria have similar abundances in the subseafloor. *Appl. Environ. Microbiol.* 79, 7790–7799. doi: 10.1128/AEM.02090-13
- Longhurst, A. (2007). *Ecological Geography of the Sea, 2nd Edn*. San Diego, CA: Academic press
- Ludwig, W., Strunk, O., Westram, R., Richter, L., and Meier, H., Yadhukumar, et al. (2004). ARB: A software environment for sequence data. *Nucleic Acids Res.* 32, 1363–1371. doi: 10.1093/nar/gkh293
- Lunau, M., Lemke, A., Walther, K., Martens-Habbena, W., and Simon, M. (2005). An improved method for counting bacteria from sediments and turbid environments by epifluorescence microscopy. *Environ. Microbiol.* 7, 961–968. doi: 10.1111/j.1462-2920.2005.00767.x
- Luo, H. W., and Moran, M. A. (2014). Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol. Mol. Biol. Rev.* 78, 573–587. doi: 10.1128/MMBR.00020-14
- Martin, J. H., Knauer, G. A., Karl, D. M., and Broenkow, W. W. (1987). VERTEX: carbon cycling in the northeast Pacific. *Deep Sea Res.* 34, 267–285. doi: 10.1016/0198-0149(87)90086-0
- Mills, H. J., Hunter, E., Humphrys, M., Kerkhof, L., McGuinness, L., Huettel, M., et al. (2008). Characterization of nitrifying, denitrifying, and overall bacterial communities in permeable marine sediments of the northeastern Gulf of Mexico. *Appl. Environ. Microbiol.* 74, 4440–4453. doi: 10.1128/AEM.02692-07
- Morono, Y., Terada, T., Masui, N., and Inagaki, F. (2009). Discriminative detection and enumeration of microbial life in marine subsurface sediments. *ISME J.* 3, 503–511. doi: 10.1038/ismej.2009.1
- Muyzer, G., Teske, A., Wirsén, C. O., and Jannasch, H. W. (1995). Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. *Arch. Microbiol.* 164, 165–172. doi: 10.1007/BF02529967
- Newton, R. J., Griffin, L. E., Bowles, K. M., Meile, C., Gifford, S., Givens, C. E., et al. (2010). Genome characteristics of a generalist marine bacterial lineage. *ISME J.* 4, 784–798. doi: 10.1038/ismej.2009.150
- Oksanen, J., Guillaume Blanchet, F., Friendly, M., Kindt, R., Legendre, R., McGinn, D., et al. (2017). *vegan: Community Ecology Package*. version 2.4–3. Available Online at: <https://cran.r-project.org/web/packages/vegan/vegan.pdf>
- Orphan, V. J., Hinrichs, K. U., Ussler, W., Paull, C. K., Taylor, L. T., Sylva, S. P., et al. (2001). Comparative analysis of methane-oxidizing archaea and sulfate-reducing bacteria in anoxic marine sediments. *Appl. Environ. Microbiol.* 67, 1922–1934. doi: 10.1128/AEM.67.4.1922-1934.2001
- Park, S., Park, J. M., Lee, K. C., Bae, K. S., and Yoon, J. H. (2014). *Boseongicola aestuarii* gen. nov., sp. nov., isolated from a tidal flat sediment. *Int. J. Syst. Evol. Microbiol.* 64, 2618–2624. doi: 10.1099/ijs.0.061960-0
- Parkes, R. J., Cragg, B. A., and Wellsbury, P. (2000). Recent studies on bacterial populations and processes in subseafloor sediments: a review. *Hydrogeol. J.* 8, 11–28. doi: 10.1007/PL00010971
- Pernthaler, A., Pernthaler, J., and Amann, R. (2002). Fluorescence *in situ* hybridization and catalyzed reporter deposition for the identification of marine bacteria. *Appl. Environ. Microbiol.* 68, 3094–3101. doi: 10.1128/AEM.68.6.3094-3101.2002
- Pujalte, M. J., Lucena, T., Ruvira, M. A., Arahá, D. R., and Macián, M. C. (2014). “The family *Rhodobacteraceae*” in *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria, 4th Edn*, eds E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. L. Thompson (Berlin: Springer), 439–512.
- Ravenschlag, K., Sahm, K., and Amann, R. (2001). Quantitative molecular analysis of the microbial community in marine Arctic sediments (Svalbard). *Appl. Environ. Microbiol.* 67, 387–395. doi: 10.1128/AEM.67.1.387-395.2001

- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, VA: R Foundation for Statistical Computing. Available Online at <http://www.R-project.org/>
- Røy, H., Kallmeyer, J., Adhikari, R. R., Pockalny, R., Jørgensen, B. B., and D'Hondt, S. (2012). Aerobic microbial respiration in 86-million-year-old deep-sea red clay. *Science* 336, 922–925. doi: 10.1126/science.1219424
- Rullkötter, J. (2006). "Organic matter: the driving force for early diagenesis," in *Marine Geochemistry, 2nd Edn*, eds H. D. Schulz and M. Zabel. (Heidelberg: Springer), 125–168.
- Schneider, D., Wemheuer, F., Pfeiffer, B., and Wemheuer, B. (2017). "Extraction of total DNA and RNA from marine filter samples and generation of a cDNA as universal template for marker gene studies" in *Metagenomics - Methods and Protocols, 2nd Edn*, eds W. Streit and R. Daniel (New York, NY: Humana Press), 13–22.
- Schnetger, B., and Lehnert, C. (2014). Determination of nitrate plus nitrite in small volume marine water samples using vanadium(III)chloride as a reduction agent. *Mar. Chem.* 160, 91–98. doi: 10.1016/j.marchem.2014.01.010
- Schulz, H. D., and Zabel, M. (2006). *Marine Geochemistry, 2nd Edn*. Heidelberg: Springer.
- Seeborg-Elverfeldt, J., Schlüter, M., Feseker, T., and Kölling, M. (2005). Rhizon sampling of porewaters near the sediment-water interface of aquatic systems. *Limnol. Oceanogr. Methods* 3, 361–371. doi: 10.4319/lom.2005.3.361
- Seiter, K., Hensen, C., Schröter, E., and Zabel, M. (2004). Organic carbon content in surface sediments - defining regional provinces. *Deep Sea Res.* 51, 2001–2026. doi: 10.1016/j.dsr.2004.06.014
- Selje, N., Simon, M., and Brinkhoff, T. (2004). A newly discovered *Roseobacter* cluster in temperate and polar oceans. *Nature* 427, 445–448. doi: 10.1038/nature02272
- Shiba, T. (1991). *Roseobacter litoralis* gen. nov., sp. nov., and *Roseobacter denitrificans* sp. nov., aerobic pink-pigmented bacteria which contain bacteriochlorophyll *a*. *Syst. App. Microbiol.* 14, 140–145. doi: 10.1016/S0723-2020(11)80292-4
- Stevens, H., Brinkhoff, T., and Simon, M. (2005). Composition of free-living, aggregate-associated and sediment surface-associated bacterial communities in the German Wadden Sea. *Aquat. Microbial. Ecol.* 38, 15–30. doi: 10.3354/ame038015
- Süß, J., Engelen, B., Cypionka, H., and Sass, H. (2004). Quantitative analysis of bacterial communities from Mediterranean sapropels based on cultivation-dependent methods. *FEMS Microbiol. Ecol.* 51, 109–121. doi: 10.1016/j.femsec.2004.07.010
- Sun, F., Wang, B., Liu, X. P., Lai, Q., Du, Y., Li, G., et al. (2010). *Leisingera nanhaiensis* sp. nov., isolated from marine sediment. *Int. J. Syst. Evol. Microbiol.* 60, 275–280. doi: 10.1099/ijs.0.010439-0
- Tada, Y., Taniguchi, A., Nagao, I., Miki, T., Uematsu, M., Tsuda, A., et al. (2011). Differing growth responses of major phylogenetic groups of marine bacteria to natural phytoplankton blooms in the western north Pacific Ocean. *Appl. Environ. Microbiol.* 77, 4055–4065. doi: 10.1128/AEM.02952-10
- Tsubouchi, T., Shimane, Y., Mori, K., Miyazaki, M., Tame, A., Uematsu, K., et al. (2013). *Loktanella cinnabarina* sp. nov., isolated from a deep subseafloor sediment, and emended description of the genus *Loktanella*. *Int. J. Syst. Evol. Microbiol.* 63, 1390–1395. doi: 10.1099/ijs.0.043174-0
- Van Trappen, S., Mergaert, J., and Swings, J. (2004). *Loktanella salsilacus* gen. nov., sp. nov., *Loktanella fryxellensis* sp. nov. and *Loktanella vestfoldensis* sp. nov., new members of the Rhodobacter group, isolated from microbial mats in Antarctic lakes. *Int. J. Syst. Evol. Microbiol.* 54, 1263–1269. doi: 10.1099/ijs.0.03006-0
- Wagner-Döbler, I., and Biehl, H. (2006). Environmental biology of the marine *Roseobacter* lineage. *Ann. Rev. Microbiol.* 60, 255–280. doi: 10.1146/annurev.micro.60.080805.142115
- Wehrmann, L. M., Risgaard-Petersen, N., Schrum, H. N., Walsh, E. A., Huh, Y., Ikehara, M., et al. (2011). Coupled organic and inorganic carbon cycling in the deep subseafloor sediment of the northeastern Bering Sea Slope (IODP Exp. 323). *Chem. Geol.* 284, 251–261. doi: 10.1016/j.chemgeo.2011.03.002
- Weinbauer, M. G., Beckmann, C., and Höfle, M. G. (1998). Utility of green fluorescent nucleic acid dyes and aluminum oxide membrane filters for rapid epifluorescence enumeration of soil and sediment bacteria. *Appl. Environ. Microbiol.* 64, 5000–5003.
- Wemheuer, B., and Wemheuer, F. (2017). "Assessing bacterial and fungal diversity in the plant endosphere," in *Metagenomics - Methods and Protocols*, eds W. Streit and R. Daniel (New York, NY: Humana Press), 75–84.
- Wemheuer, B., Güllert, S., Billerbeck, S., Giebel, H. A., Voget, S., Simon, M., et al. (2014). Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol. Ecol.* 87, 378–389. doi: 10.1111/1574-6941.12230
- Wemheuer, B., Wemheuer, F., Hollensteiner, J., Meyer, F. D., Voget, S., and Daniel, R. (2015). The green impact: bacterioplankton response toward a phytoplankton spring bloom in the southern North Sea assessed by comparative metagenomic and metatranscriptomic approaches. *Front. Microbiol.* 6:05. doi: 10.3389/fmicb.2015.00805
- Zabel, M., and Hensen, C. (2006). "Quantification and regionalisation of benthic reflux," in *Marine Geochemistry, 2nd Edn*, eds H. D. Schulz and M. Zabel (Heidelberg: Springer), 429–456.
- Zhang, Y., Sun, Y., Jiao, N. Z., Stepanauskas, R., and Luo, H. W. (2016). Ecological genomics of the uncultivated marine *Roseobacter* lineage CHAB-I-5. *Appl. Environ. Microbiol.* 82, 2100–2111. doi: 10.1128/AEM.03678-15

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Pohlner, Degenhardt, von Hoyningen-Huene, Wemheuer, Erlmann, Schnetger, Badewien and Engelen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Complementary Metaproteomic Approaches to Assess the Bacterioplankton Response toward a Phytoplankton Spring Bloom in the Southern North Sea

Lars Wöhlbrand<sup>1\*</sup>, Bernd Wemheuer<sup>2</sup>, Christoph Feenders<sup>3</sup>, Hanna S. Ruppertsberg<sup>1</sup>, Christina Hinrichs<sup>1</sup>, Bernd Blasius<sup>3</sup>, Rolf Daniel<sup>2</sup> and Ralf Rabus<sup>1</sup>

<sup>1</sup> General and Molecular Microbiology, Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University of Oldenburg, Oldenburg, Germany, <sup>2</sup> Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August-University Göttingen, Göttingen, Germany, <sup>3</sup> Mathematical Modelling, Institute for Chemistry and Biology of the Marine Environment (ICBM), Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

## OPEN ACCESS

### Edited by:

Justin Robert Seymour,  
University of Technology, Sydney,  
Australia

### Reviewed by:

Martin Ostrowski,  
Macquarie University, Australia  
David H. Green,  
Scottish Association for Marine  
Science, UK

### \*Correspondence:

Lars Wöhlbrand  
lars.woehlbrand@uni-oldenburg.de

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 20 September 2016

**Accepted:** 03 March 2017

**Published:** 24 March 2017

### Citation:

Wöhlbrand L, Wemheuer B,  
Feenders C, Ruppertsberg HS,  
Hinrichs C, Blasius B, Daniel R and  
Rabus R (2017) Complementary  
Metaproteomic Approaches  
to Assess the Bacterioplankton  
Response toward a Phytoplankton  
Spring Bloom in the Southern North  
Sea. *Front. Microbiol.* 8:442.  
doi: 10.3389/fmicb.2017.00442

Annually recurring phytoplankton spring blooms are characteristic of temperate coastal shelf seas. During these blooms, environmental conditions, including nutrient availability, differ considerably from non-bloom conditions, affecting the entire ecosystem including the bacterioplankton. Accordingly, the emerging ecological niches during bloom transition are occupied by different bacterial populations, with *Roseobacter* RCA cluster and SAR92 clade members exhibiting high metabolic activity during bloom events. In this study, the functional response of the ambient bacterial community toward a *Phaeocystis globosa* bloom in the southern North Sea was studied using metaproteomic approaches. In contrast to other metaproteomic studies of marine bacterial communities, this is the first study comparing two different cell lysis and protein preparation methods [using trifluoroethanol (TFE) and in-solution digest as well as bead beating and SDS-based solubilization and in-gel digest (BB GeLC)]. In addition, two different mass spectrometric techniques (ESI-iontrap MS and MALDI-TOF MS) were used for peptide analysis. A total of 585 different proteins were identified, 296 of which were only detected using the TFE and 191 by the BB GeLC method, demonstrating the complementarity of these sample preparation methods. Furthermore, 158 proteins of the TFE cell lysis samples were exclusively detected by ESI-iontrap MS while 105 were only detected using MALDI-TOF MS, underpinning the value of using two different ionization and mass analysis methods. Notably, 12% of the detected proteins represent predicted integral membrane proteins, including the difficult to detect rhodopsin, indicating a considerable coverage of membrane proteins by this approach. This comprehensive approach verified previous metaproteomic studies of marine bacterioplankton, e.g., detection of many transport-related proteins (17% of the detected proteins). In addition, new insights into e.g., carbon and nitrogen metabolism were obtained. For instance, the C1 pathway was more prominent outside



the bloom and different strategies for glucose metabolism seem to be applied under the studied conditions. Furthermore, a higher number of nitrogen assimilating proteins were present under non-bloom conditions, reflecting the competition for this limited macro nutrient under oligotrophic conditions. Overall, application of different sample preparation techniques as well as MS methods facilitated a more holistic picture of the marine bacterioplankton response to changing environmental conditions.

**Keywords:** metaproteomics, bacterioplankton, algal bloom, sample preparation, mass spectrometry

## INTRODUCTION

Cultivation-independent analysis of marine bacterioplankton communities targeting 16S rRNA as well as environmental DNA and RNA applying next-generation sequencing techniques greatly advanced our understanding of their diversity and ecology (e.g., Venter et al., 2004; Frias-Lopez et al., 2008; Gifford et al., 2011; Vila-Costa et al., 2012). While metagenomic studies have provided insights into the metabolic potential of the communities, metatranscriptomic analysis generated functional data allowing first hints on microbial activity. In addition, *in situ* activity may be assessed by metaproteomics, analyzing the proteins, i.e., the catalytically active molecules, formed by the community in a given habitat (for overview see Hettich et al., 2012; Abraham et al., 2014). Metaproteomics has been successfully applied to diverse habitats ranging from low-complexity acid mine drainage biofilm (e.g., Ram et al., 2005), activated sludge (e.g., Wilmes and Bond, 2004), human microbiome (e.g., Chen et al., 2008) to the ocean (e.g., Giovannoni et al., 2005; Sowell et al., 2009; Morris et al., 2010; Teeling et al., 2012).

During phytoplankton blooms, large amounts of organic matter are generated by primary production (Arrigo, 2005; Bunse and Pinhassi, 2017). Marine bacteria play an important role in the decomposition of this organic matter, since they remineralize > 50% during and after bloom events (Cole et al., 1988; Kerner and Herndl, 1992; Ducklow et al., 1993). However, diverse environmental factors are influenced by the bloom, including limitation of nutrient availability for the marine bacterioplankton. Therefore, understanding the complex dynamics and interactions between bacterial communities and phytoplankton blooms is essential to assess the ecological impact of bloom events.

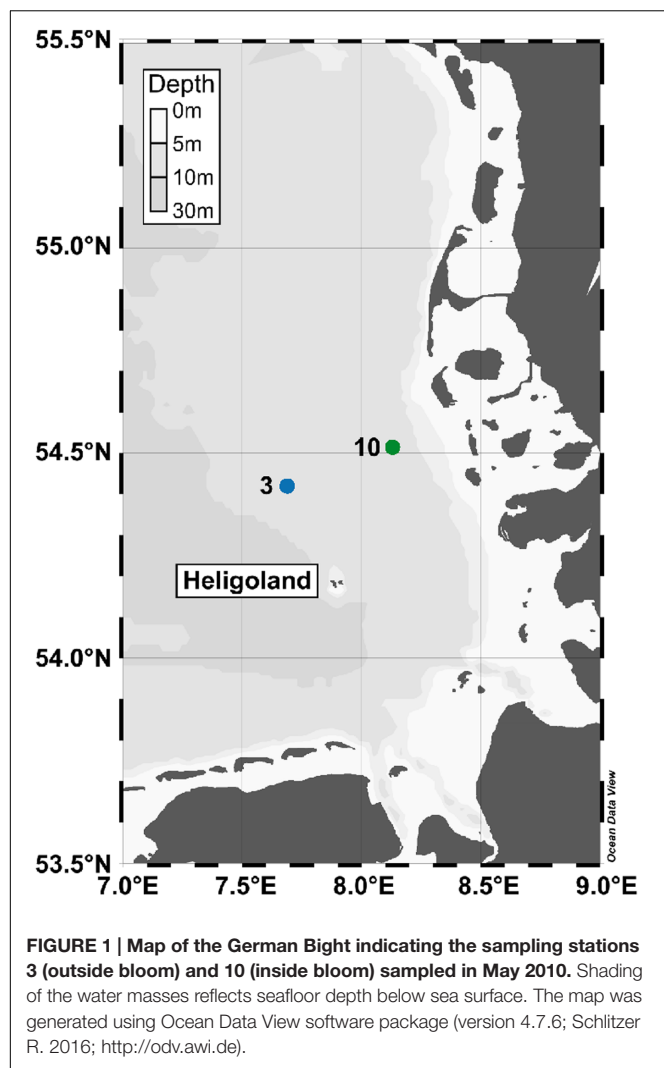
Annually recurring phytoplankton spring blooms can be observed in the North Sea, representing a typical coastal shelf sea of the temperate zone. Especially its southern region, the German Bight, is highly productive due to the continuous nutrient supply by rivers (McQuatters-Gollop et al., 2007; Wiltshire et al., 2008, 2010). A dynamic succession of distinct bacterial clades before, during, and after bloom events in the North Sea was observed in recent studies (Alderkamp et al., 2006; Alonso and Pernthaler, 2006a,b; Teeling et al., 2012). They

indicate that specialized bacterial populations occupy transitory ecological niches provided by phytoplankton-derived substrates. Metagenomic, -transcriptomic and -proteomic analysis of the diversity and activity of marine bacterioplankton during the same bloom event in the North Sea (Heligoland) showed that members of the *Rhodobacteraceae* and SAR92 clade exhibited high metabolic activity levels (Teeling et al., 2012; Klindworth et al., 2014). In two previous studies, structural and functional differences of the free-living bacterioplankton community in response to a *Phaeocystis globosa* bloom in the southern North Sea in spring 2010 were investigated using comparative metagenomic and metatranscriptomic approaches (Wemheuer et al., 2014, 2015). It was shown that the phytoplankton spring bloom significantly affected bacterioplankton community structures and the abundance of certain bacterial groups, e.g., significantly higher abundance of the *Roseobacter* RCA cluster and the SAR92 clade during a bloom. In addition, functional differences were investigated by comparative metagenomic and metatranscriptomic approaches revealing differences in bacterial gene expression inside and outside of the investigated bloom.

Metaproteomic analysis of environmental samples in particular is challenged by the sample's inherent high organismic diversity, coupled to the complexity and wide abundance range of the protein complement of each organism. While the latter demands highly accurate peptide separation and mass spectrometric analysis, the preceding cell lysis and subsequent sample processing determine the protein share that is accessible to analysis. Especially in the case of environmental prokaryotic samples, a high diversity of cell wall structures is encountered, ranging from "standard" Gram-positives/-negatives to highly specialized structures, e.g., in case of Planctomycetes (Fuerst, 2005). This diversity necessitates methods for cell lysis that cover all members of the prokaryotic community. Detergent-based methods for protein preparation allow for efficient cell lysis and protein solubilization. However, detergents have to be depleted (e.g., chaotropes as urea) or completely removed (e.g., strong ionic detergents as SDS) from the sample prior to proteolytic digest and subsequent MS analysis to allow for efficient proteolysis and mass acquisition. In previous studies, only a single method for cell lysis and protein preparation was applied to analyze the metaproteome of marine bacterioplankton communities (e.g., Sowell et al., 2009; Morris et al., 2010; Teeling et al., 2012).

In this study, two different types of cell lysis and subsequent protein/peptide separation were applied to analyze the protein complement of the bacterioplankton community within and outside a *P. globosa* dominated algal bloom in the southern North

**Abbreviations:** BB, bead beating; ESI, electrospray ionization; GeLC, SDS-PAGE decomplexation coupled to in-gel digest and nanoLC separation; MALDI, matrix assisted laser desorption ionization; nanoLC, nano liquid chromatography; SDS, sodium dodecyl sulfate; TFE, trifluoroethanol; TOF, time-of-flight.



Sea: (i) chemical lysis applying TFE coupled to in-solution digest and MS analysis as well as (ii) mechanical lysis using bead beating (BB) coupled to SDS-PAGE pre-fractionation, in-gel digest and MS analysis. In addition, two different mass spectrometric techniques were applied to analyze nanoLC-separation peptides of the TFE-lysed samples: (i) online by ESI-iontrap MS and (ii) offline by MALDI-TOF MS. Final protein identification was based on a corresponding metagenome/-transcriptome derived protein database. This methodological spectrum was applied to extend our understanding of microbial adaptation to nutrient limitation in the marine habitat.

## MATERIALS AND METHODS

### Sampling

Water samples were collected in the southern North Sea at 13 different stations in May 2010 on board RV Heinke (HE327) to investigate bacterial community composition (Wemheuer et al., 2014). At two of these stations additional

samples were collected inside (station 10) and outside (station 3) an algal bloom (2 m water depth) for this study (**Figure 1**). Water from eight CTD-mounted 5 L Niskin bottles was pooled in an ethanol-rinsed PE barrel and sequentially filtered through a 10  $\mu$ m prefilter prior to a sandwich of a precombusted (4 h at 450°C) glassfiber filter (47 mm diameter, Whatman GF/D; Whatman, Maidstone, UK) and a 3.0  $\mu$ m polycarbonate filter (47 mm diameter, Nuclepore; Whatman). One liter of the obtained filtrate was further filtered through a sandwich of a precombusted (4 h at 450°C) glassfiber filter (47 mm diameter, Whatman GF/F; Whatman) and a 0.2  $\mu$ m polycarbonate filter (47 mm diameter, Nuclepore, Whatman). Filters were frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until usage.

The presence of an algal bloom, as indicated by satellite data, was confirmed by determination of chlorophyll a and phaeopigments as described in detail by Wemheuer et al. (2014). In addition, suspended particulate matter (SPM), particulate organic carbon (POC), particulate organic nitrogen (PON), dissolved inorganic nutrients (nitrate, nitrite, and phosphate) and bacterioplankton cell numbers were determined as described Wemheuer et al. (2014).

### Processing and Analysis of Metagenomic and -Transcriptomic Datasets to Generate a Protein Database for Identification

Metagenomic and metatranscriptomic data sets used in this study were obtained from Wemheuer et al. (2015). Following data processing and assembly, all obtained contigs were joined and open reading frames (ORFs) were predicted for all contigs using Prodigal version 2.6 (Hyatt et al., 2010). Obtained protein sequences were clustered at 80% sequence similarity employing Usearch version 8.0.1623 (Edgar, 2010). During clustering, short peptide sequences (< 50 aa) were removed. Obtained protein sequences were functionally classified using UPProC version 1.2 (Meinicke, 2015). Corresponding nucleotide sequences were taxonomically classified using kraken (Wood and Salzberg, 2014).

To determine which ORFs were expressed, metatranscriptomic datasets were mapped on the assembled contigs using Bowtie 2 version 2.2.4 (Langmead and Salzberg, 2012) with one mismatch in the seed and multiple hits reporting enabled. Ribosomal RNA was removed from metatranscriptomic datasets prior to mapping employing SortMeRNA version 2.0 (Kopylova et al., 2012). The number of unique sequences per gene was calculated as described previously (Wemheuer et al., 2015). In addition, coverage was determined for the rRNA-depleted datasets using nonpareil version 2.4 (Rodríguez and Konstantinidis, 2014). Corresponding curves were subsequently generated in R (version 3.2.5; R Development Core Team 2014<sup>1</sup>).

### Protein Extraction

Protein extraction was performed either by direct cell lysis on the filter or by means of BB. In case of direct cell lysis (adapted from Wang et al., 2005), one eighth of a filter sandwich was cut, 200  $\mu$ L

<sup>1</sup><http://www.R-project.org/>

of 50% (v/v) TFE in 100 mM ammonium bicarbonate added and incubated for 5 min in an ultrasound bath. Disulfide bonds were reduced with 10 mM DTT by incubation for 45 min at 55°C followed by alkylation with 50 mM iodoacetamide for 30 min at room temperature in the dark. Subsequently, the suspension was centrifuged for 30 min at 20,000 g and 20°C followed by pooling the supernatant of two filter eighths. After addition of an equal amount of HPLC-grade water, a five-fold excess of ice-cold acetone was added and the solution was incubated over night at −20°C. Precipitated proteins were collected by centrifugation (5,700 g, 1 h at 4°C; Rotanta RP, Hettich, Beverly, MA, USA) and the obtained pellet was washed twice with 80% v/v ice-cold acetone. The pellet was air dried and resuspended in digestion buffer (50 mM Tris-HCl pH 8.0, 1 M urea). One microgram of trypsin (Trypsin GOLD, Promega, Mannheim, Germany) was added for proteolytic digest and incubated at 37°C over night. Aliquots were frozen in liquid nitrogen and stored at −80°C until use. Three half filters were prepared per station.

The second extraction method was modified from the one described by Teeling et al. (2012). Half a 47 mm filter was cut into small pieces and transferred into a BB tube containing 1.15 g 1.0 mm silica and 0.65 g 0.5 mm yttria-stabilized zirconium oxide spheres (MP Biomedical, Santa Ana, CA, USA). Following addition of 700 µL lysis buffer (50 mM Tris-HCl pH 8.0, 2% w/v SDS, 10% v/v glycerol, 0.1 M DTT) BB was performed for 15 s at a speed of 10 m/s (FastPrep-24 5G; MP Biomedical) with subsequent 5 min incubation on ice and repeated thrice. The suspension was incubated 10 min at 95°C. Following centrifugation (10 min, 20,000 g, 4°C) the supernatant was transferred into ultracentrifuge tubes and ultracentrifuged for 1 h at 100,000 g and 4°C (Ultima Max XP equipped with the MLA130 rotor; Beckman Coulter, Krefeld, Germany). Proteins of the supernatant were precipitated by addition of five volumes ice-cold acetone, incubation at −20°C over night and centrifugation for 2 h at 5,700 g and 4°C. The obtained pellet was washed with 80% v/v ice cold acetone twice and the pellet air dried prior to resuspension in sample buffer (50 mM Tris-HCl pH 8.0, 10% w/v SDS, 10% v/v glycerol, 0.1 M DTT). The entire protein preparation was subjected to separation by SDS-PAGE (12% acrylamide) using mini gels (7 cm long, 0.75 mm thick) and the Mini-PROTEAN Tetra System (Bio-Rad Laboratories, Munich, Germany). The PageRuler Unstained Protein Ladder (ThermoFisher Scientific, Rockford, IL, USA) was used as molecular size marker. Following electrophoresis, gels were stained with colloidal Coomassie Brilliant Blue according to the method described by Neuhoff et al. (1988). Three half filters were prepared per station.

## Peptide Analysis by nanoLC ESI-Iontrap MS/MS

Aliquots (15 µL) of the in-solution digested samples (TFE) were separated by nanoLC (Ultimate3000 nanoRSLC; ThermoFisher Scientific, Germering, Germany) operated in trap column mode (3 µm beads, 75 µm inner diameter, 2 cm length; ThermoFisher Scientific) equipped with a 25 cm separation column (2 µm beads, 75 µm inner diameter, ThermoFisher Scientific) and

applying a linear 240 min gradient from 2% v/v to 50% v/v acetonitrile with subsequent re-equilibration (eluent A: 0.1% v/v formic acid; eluent B: 80% v/v acetonitrile, 0.1% v/v formic acid). The nanoLC effluent was continuously analyzed by an online-coupled iontrap mass spectrometer (amaZon speed ETD; Bruker Daltonik GmbH, Bremen, Germany) using an electrospray ion source (Captivespray; Bruker Daltonik GmbH) operated in positive ion mode as described in detail before (Wöhlbrand et al., 2016).

In case of SDS-PAGE separated samples, the entire sample lane was cut into eight pieces, which were further cut into small pieces of ~1–2 mm<sup>2</sup> and subjected to in-gel digest as described by Kossmehl et al. (2013). Generated peptides were analyzed by nanoLC ESI-iontrap MS/MS (setup see above) using a linear 130 min gradient from 2% v/v to 50% v/v acetonitrile with subsequent re-equilibration.

## Peptide Analysis by nanoLC MALDI-TOF/TOF MS/MS

Peptide samples obtained from the direct extraction, i.e., in-solution digest (TFE), were additionally separated by nanoLC (setup see above) coupled to a fraction collection robot (ProteinPrep FC II; Bruker Daltonik GmbH). The latter continuously mixed the nanoLC eluent with a matrix solution (consisting of 22% v/v of a saturated  $\alpha$ -cyano-4-hydroxycinnamic acid solution in 90% v/v acetonitrile, 0.1% v/v trifluoroacetic acid, and 0.1% v/v ammonium-phosphate as well as 0.1% v/v trifluoroacetic acid in 95% v/v acetonitrile) and spotted onto a 384-spot anchochip target (Bruker Daltonik GmbH) at a flow rate of 0.3 µL/min. Peptides were separated applying a linear gradient of 120 min and fraction collection was performed from 20 to 84 min of this gradient. The peptide standard II mixture (Bruker Daltonik GmbH) was manually applied to calibration spots used for internal mass calibration. Prepared targets were analyzed by a MALDI-TOF/TOF mass spectrometer (ultrafleXtreme; Bruker Daltonik GmbH) using the WARP-LC software (version 1.3; Bruker Daltonik GmbH). Initially, MS analysis was performed for all prepared spots. Subsequently, masses of interest for MS/MS analysis were determined for the entire target that fulfilled the following criteria: signal-to-noise  $\geq 20$ , minimum mass distance to co-eluting compound 1.0 Da, compound mass tolerance 15 ppm. Masses present in more than 60% of all fractions were regarded as background and not considered. MS/MS of masses of interest were acquired for spots revealing the most intense MS signal. Following MS/MS measurement, all MS and MS/MS spectra were combined into a single file containing all compounds and transferred to ProteinScape for protein identification (see below).

## Protein Identification

Protein identification was performed via the ProteinScape platform (version 3.1; Bruker Daltonik GmbH) on an in house Mascot server (version 2.3; Matrix Science Ltd., London, UK) against the generated metagenome-/transcriptome-based database and applying a target-decoy strategy. Mascot search criteria were as follows: enzyme, trypsin; fixed modification,



carbamidomethylation (C); variable modification, oxidation (M); mass values, monoisotopic; max missed cleavages, 1; significance threshold,  $p < 0.05$ ; instrument type, ESI-TRAP or MALDI-TOF-TOF, respectively. Assessment of the Mascot search result was performed accepting only peptides with a Mascot score  $\geq 25.0$ , minimum peptide length of 5 and a false discovery rate  $< 1.0\%$ . In case of ESI-iontrap, the mass tolerance was set to 0.3 Da (MS) and 0.4 Da (MS/MS) and for MALDI-TOF to 100 ppm (MS) and 0.87 Da (MS/MS). Peptides were identified by at least two spectra meeting the identification criteria (on average 3.6).

In case of measured technical replicates (i.e., same sample preparation), respective peptides were compiled using the protein extractor implemented in ProteinScape. Similarly, peptides of all eight gel slices per sample were compiled yielding a single list of proteins per MS-technique and station sample preparation. A comparative table of all identified proteins with detailed identification information is provided as **Supplementary Table S1**. Protein functions were assessed by functional domains (Pfam) present in the metagenome-derived primary sequence (for details see section on metagenome generation). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PXD004944.

## Assessing Similarity of Samples (Multidimensional Scaling)

Sample dissimilarity was calculated according to the Bray–Curtis dissimilarity (Bray and Curtis, 1957) based on presence or absence (coded as 1 or 0, respectively) of each of the 585 assigned different proteins (similar to Kossmehl et al., 2013). The dissimilarity matrix of all pairwise distances was subsequently visualized using multidimensional scaling (MDS; Kruskal and Wish, 1978; Zech et al., 2011). Positions in the 2D plane were scanned for local minima with the Markov chain Monte Carlo sampling method parallel tempering (Swendsen and Wang, 1986; Geyer, 1991), and results were controlled both visually (Shepard diagrams) and through Kruskal's stress formula 1 (Kruskal, 1964). Differences between sample groups were tested for significance by permutational MANOVA (PERMANOVA) (Anderson, 2001), based on protein presence-absence.

## RESULTS AND DISCUSSION

### Characteristics of Analyzed Samples

In this study, the proteomic response of the pelagic bacterial community toward a phytoplankton bloom in the southern North Sea was analyzed. Samples were taken outside (station 3) and inside (station 10) a *Phaeocystis globosa* dominated algal bloom at 2 m water depth (**Figure 1** and **Table 1**). The microalgae *P. globosa* is globally distributed and its blooms have been observed in many marine environments such as the coast of the eastern English Channel, the southern North Sea and the south coast of China (Schoemann et al., 2005). Moreover, it is considered to be responsible for harmful algal blooms (Veldhuis and Wassmann, 2005). While the salinity of both stations was similar (31.4 and 31.1 psu at non-bloom and bloom, respectively),

temperature was slightly higher (9.4°C vs. 11.4°C) and the bacterial abundance decreased inside the bloom ( $2.57 \times 10^6$  vs.  $1.21 \times 10^6$  cells/mL). In addition, the chlorophyll a and phaeopigment contents were  $> 6$ -fold higher within the bloom (1.12 vs. 6.93  $\mu\text{g/L}$  and 0.25 vs. 2.10  $\mu\text{g/L}$ , respectively). Overall, most of the determined environmental parameters (**Table 1**) were significantly linked to the presence of an algal bloom, as outlined in detail by Wemheuer et al. (2014).

## The Metagenome-Based Protein Database

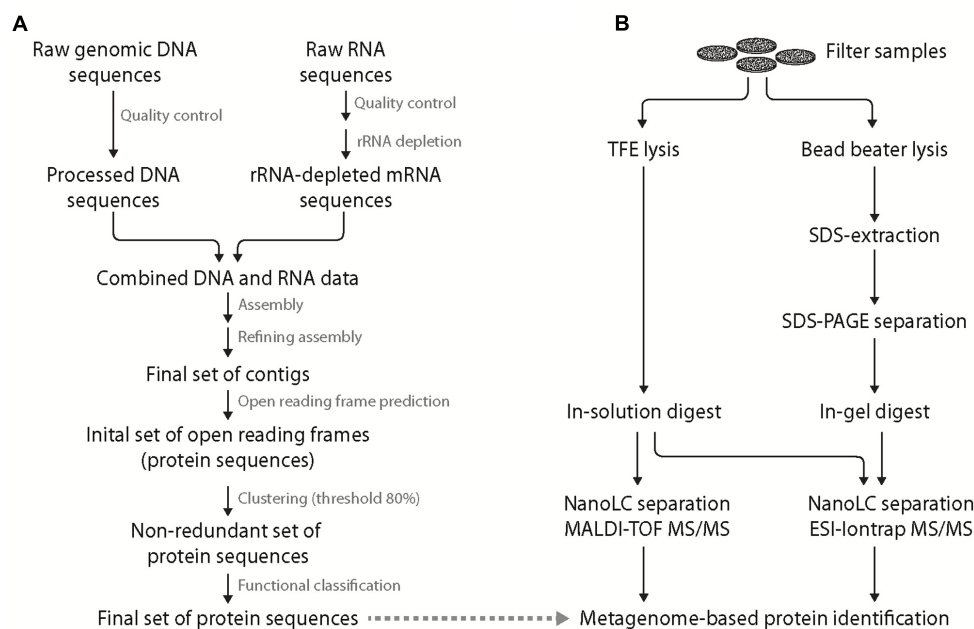
After quality filtering, nearly 564 million sequences derived from Illumina sequencing and pyrosequencing were assembled employing velvet and metavelvet at different kmer values (**Figure 2A**). After joining all obtained contigs, a total of 14.6 million contigs with an average insert length of 290 bp were retrieved with 13.8 million being unique across all assemblies. Nearly 6 million ORFs were subsequently predicted from these contigs using prodigal. After removal of redundant sequences,  $> 922,000$  proteins with an average length of 111 amino acids were obtained and used as backbone for protein calling. However, only 4% of all protein sequences within the database could be allocated to distinct phyla. The low number of taxonomically assigned reads might be related to the incompleteness of currently available databases or the shortness of some peptide reads as

**TABLE 1 | Position, time and environmental parameters at the sampling sites and of the retrieved samples.**

	Non-bloom (station 3)	Bloom (station 10)
Date of sampling	05/26/2010	05/28/2010
Latitude (°N)	54.4223	54.5135
Longitude (°E)	7.6833	8.128
Water depth (m)	2.0	2.0
Temperature (°C)	9.43	11.40
Salinity (psu)	31.42	31.11
Fluorescence (FU)	0.20	2.27
Transmission (%)	87.98	74.79
Density (g/L)	1024.27	1023.70
Chlorophyll a ( $\mu\text{g/L}$ )	1.12	6.93
Phaeopigments ( $\mu\text{g/L}$ )	0.25	2.10
Phaeopigments (% total Chl)	18.3	23.3
Bacterial abundance ( $10^6/\text{mL}$ )	2.57	1.21
Suspended particulate matter (mg/L)	4.60	7.50
Particulate organic carbon ( $\mu\text{g/L}$ )	291.2	737.5
Particulate organic nitrogen ( $\mu\text{g/L}$ )	43.0	95.3
Nitrate ( $\mu\text{M}$ )	7.4	3.7
Nitrite ( $\mu\text{M}$ )	0.25	0.29
Mononitrogen oxides ( $\mu\text{M}$ )	7.65	3.98
Phosphate ( $\mu\text{M}$ )	0.020	0.070

Significance of differences between the bloom and non-bloom conditions of all sites of the sampling campaign were tested by either the Student's two-sample *t*-test (homogeneous variances) or the Welch Two Samples *t*-test (heterogeneous variances) for normally distributed samples and with the Wilcoxon–Mann–Whitney test for non-normally distributed samples (for details see Wemheuer et al., 2014).





**FIGURE 2 | Integrated scheme of a “meta” protein sequence database generation and complementary metaproteomic approaches.** Generation of metagenome/-transcriptome-based, site-specific protein sequence database (**A**). Initial quality control of the obtained genomic DNA and enriched mRNA (including removal of adaptors and artifacts using Trimmatic) allowed the generation of an initial set of contigs (applying Velvet and 29-109 Kmers). Removal of chimeras and refinement of the assembly yielded the final set of contigs that was used for open reading frame (ORF) prediction. The resulting set of protein sequences was clustered to remove redundant sequences (threshold 80% identity) allowing for unambiguous protein identification. Finally, functional prediction for the non-redundant set of protein sequences was performed. Preparation of proteins from filtered bacterioplankton (**B**) was performed by either trifluoroethanol (TFE) lysis (left) coupled to in-solution digest and MALDI-TOF as well as ESI-iontrap mass spectrometry or bead beater lysis (right) with SDS protein extraction coupled to SDS-PAGE pre-fractionation, in-gel digest and nanoLC ESI-iontrap MS. In both cases, protein identification was based on the generated, site-specific metagenome/-transcriptome database.

these factors are still very problematic during taxonomic binning approaches (as reviewed by Mande et al., 2012).

To determine the number of proteins found in the meta-transcriptomic and meta-proteomic approaches, rRNA-depleted meta-transcriptomic data for the two sampling stations were mapped on all contigs obtained from the assemblies. Afterwards, the number of sequences per ORF was determined. A total of 239 genes or proteins (40.8%), respectively, were detected by both approaches (**Supplementary Table S1**) illustrating differences between, but also similarities of both approaches.

## Experimental Strategy

In this study, two different methods were applied for cellular lysis and protein extraction of bacterioplankton filter samples (**Figure 2B**). The first method was based on chemical cell lysis using TFE (Wang et al., 2005). The organic solvent TFE in hypotonic aqueous buffer allows for cell lysis on the filter and improves protein solubility and denaturation (Ferro et al., 2000; Chertov et al., 2004). Furthermore, its relatively high volatility allows for in-solution digest, direct nanoLC peptide separation and MS analysis thereby minimizing sample loss during sample preparation. The second method applies mechanical forces (BB) for cell lysis and the ionic detergent SDS for protein solubilization. Subsequent SDS-PAGE pre-fractionation was applied to reduce sample complexity as well

as to allow for protease treatment following detergent and salt removal (i.e., in-gel digest; method referred to as BB GeLC in the following). The larger peptide sample volume obtained in the TFE-lysis protocol allowed for peptide analysis by two different MS techniques following nanoLC separation, i.e., online-coupled ESI-iontrap MS and offline MALDI-TOF MS. Application of two MS techniques was performed to assess the influence of different ionization (ESI vs. MALDI) and mass analysis (iontrap vs. TOF) techniques on protein identification in case of the investigated environmental samples.

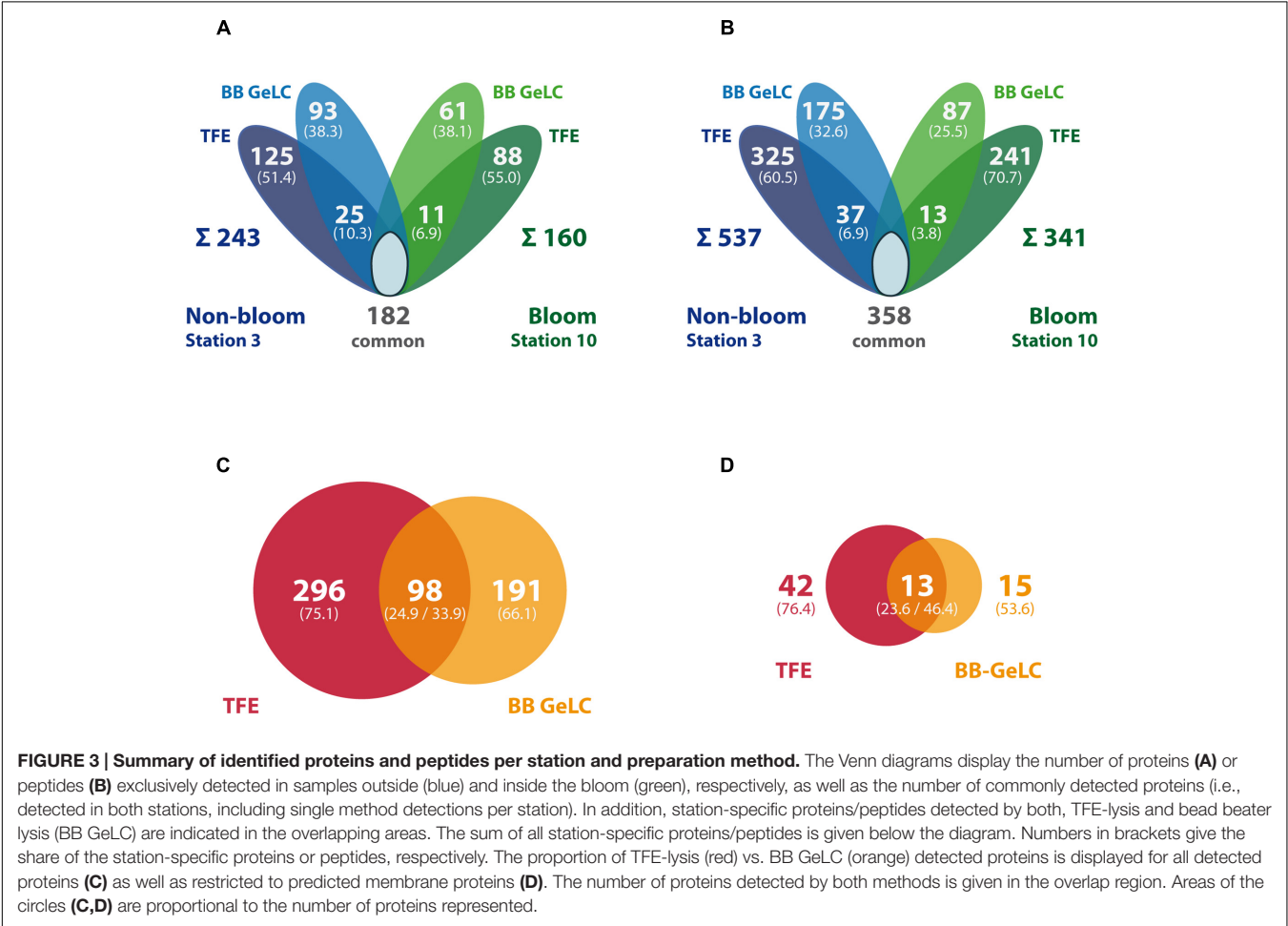
## Overview of the Proteomic Dataset

Combining all nine replicate analyses per station, a total of 585 different proteins was identified, based on the detection of 1,236 different peptides (**Table 2**). Identified proteins revealed an average Mascot score of 129.2, an average number of 1.9 protein identifying peptides (7.2% of all detected proteins identified by a single peptide) and an average sequence coverage of 23.2% (an overview of all identified proteins per station and method is given in **Supplementary Table S1**). Of all 1,936 detected proteins, about 60% were detected in two or more replicate analyses per station (**Supplementary Figure S1**). Single replicate detection of ~40% is larger than the previously reported 24% for replicate analysis of the yeast proteome applying 2D LC-MS/MS (Liu et al., 2004). The higher share may be attributed to

TABLE 2 | Summary of protein identification data.

	Total	ESI BB GeLC	Total TFE	ESI TFE	MALDI TFE
No. identified proteins	1,936	604	1,332	734	598
No. different proteins	585	289	394	289	236
No. identified peptides	3,681	959	2,722	1,429	1,293
Average protein score <sup>a</sup>	129.2	100.6	142.2	127.5	160.3
Average no. peptides <sup>b</sup>	1.9	1.8	2.0	1.9	2.2
Average seq. cov. (%) <sup>c</sup>	23.2	20.5	24.4	23.6	25.4
No. detected MemProt	225	53	173	102	70
No. different MemProt <sup>d</sup>	70	28	55	42	28

The protein identification parameters are given for the sum of all samples (total) as well as subdivided into the different cell lysis and MS techniques (in case of TFE lysis). BB GeLC, bead beater lysis with subsequent SDS-PAGE fractionation and in-gel digest; TFE, trifluoroethanol lysis with in-solution digest. <sup>a</sup>Mascot score. <sup>b</sup>Number of protein identifying peptides. <sup>c</sup>Seq. cov., sequence coverage. <sup>d</sup>MemProt, membrane proteins, only different proteins counted.



the large environmental database containing > 900,000 amino acid sequences (yeast 6,600), which makes it more likely that proteins miss the applied strict identification criteria (minimal ion score, 95% significance threshold of Mascot, and the target-decoy-based false discovery rate) and thus are not identified in each replicate. Additionally, detectability of peptides is influenced by technical constraints (e.g., ionization suppression of co-eluting peptides and limited number of MS/MS acquired per full scan MS), which are more likely to occur in case of highly complex

samples. The very high complexity of the analyzed environmental samples (covering a wide organismic diversity) thus promotes sporadic protein detection (Hettich et al., 2012). Out of the 585 different proteins, 243 were exclusively detected outside the bloom, while 160 were exclusively detected inside the bloom; 182 proteins were detected at both locations (Figure 3A). Correspondingly, more specific peptides were detected outside (537) as compared to inside the bloom (341; 358 peptides present at both stations) (Figure 3B). The differences in the number of

detected proteins and peptides, respectively, may be attributed to the (i) two-times higher bacterial abundance outside the bloom (**Table 1**) combined with (ii) a higher diversity in the total as well as active microbial community determined by metagenome and -transcriptome analysis (Wemheuer et al., 2015). Notably, 12% of all detected proteins represent predicted integral membrane proteins (**Table 2** and **Supplementary Table S1**) demonstrating coverage of the membrane protein fraction to some extent.

## Significance of Station-Specific Protein Complements

Identified proteins per station and preparation as well as MS detection method (in case of TFE samples) were tested for significance using a multivariate approach, which considers the protein data (i.e., protein presence or absence) of each station and method replicate (i.e., three replicates each of three distinct methods) as a multidimensional entity. To investigate how the different replicate samples relate to each other, a PERMANOVA (Anderson, 2001) comparing differences within and in-between groups was applied. Here, a group consists of the three replicates per station and measurement technique. In addition, similarities and dissimilarities between each prepared sample were visualized by MDS (**Figure 4A**). A detailed view on the relation (Bray–Curtis dissimilarity) of all samples to each other is provided in **Figure 4B**. Here, a value of 0 corresponds to completely identical samples (i.e., the same proteins were detected in both samples) and 1 to maximally different samples (i.e., samples share no proteins at all).

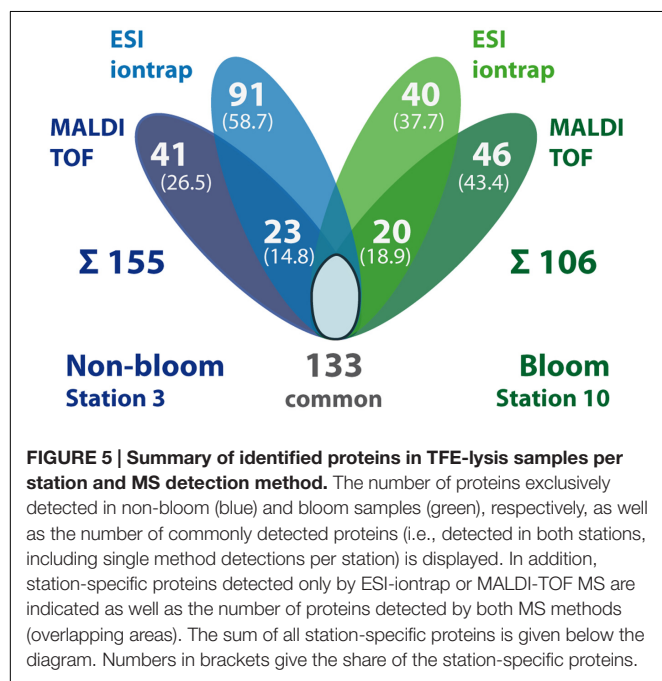
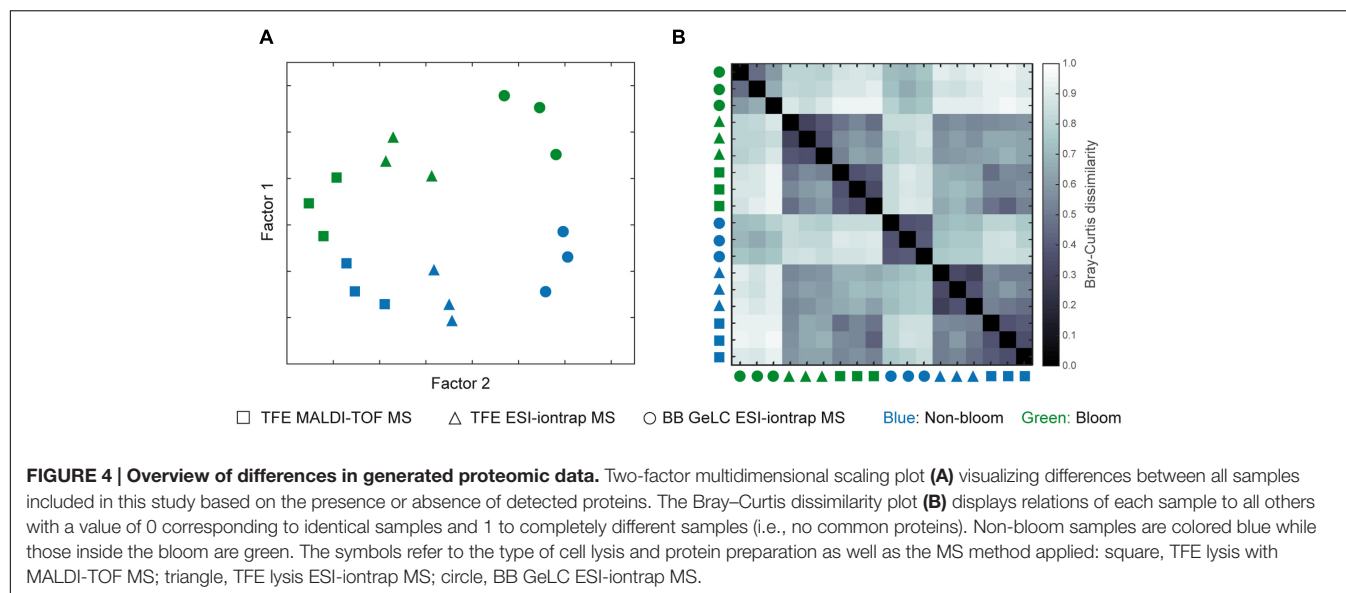
A clear separation of samples originating from out- and inside the bloom is visualized by the MDS-plot (**Figure 4A**) reflecting the calculated significant difference of the samples ( $p = 0.1$ ), irrespective of the cell lysis and protein preparation method used. Interestingly, replicate groups of the different preparation and measurement techniques of each station revealed clear differences in protein detection ( $p = 0.004$ ). Accordingly, proteins detected among replicates (any combination of station and measurement technique) are highly similar compared to those of other combinations (triplicate to all others  $< p = 0.0005$ ). This tight relationship is also visible in both the MDS- as well as the dissimilarity-plot, where replicate samples are in close proximity and display low Bray–Curtis dissimilarity values, respectively (**Figure 4**). Notably, despite their unique features, TFE lysis samples analyzed by ESI-iontrap or MALDI-TOF MS are more closely related to each other as compared to the BB GeLC treated samples.

## Impact of Protein Preparation and MS Methods on Protein Detection

Distinct sub-proteomes were detected applying the different lysis and protein preparation methods in case of both station samples (**Figure 3A**). Interestingly, the share of the method-specific protein detections of the respective station protein complement is very similar for both stations: 38% and 38% BB GeLC exclusive, 51% and 55% TFE exclusive and only 7 and 10% of the proteins were detected by both methods (values given for out- and inside the bloom, respectively). This reproducible

distinctiveness of the detected protein complements with respect to the extraction method but independent of sample origin demonstrates method specific effects on the protein extraction. This effect becomes more evident, considering all identified proteins irrespective of the station: 75% of the proteins detected by the TFE approach were exclusively identified by this method while 66% of the proteins detected by the BB GeLC approach were exclusively identified by this method (**Figure 3C**). Notably, nearly twice as many predicted membrane proteins have been identified applying the TFE method (55) as compared to the BB GeLC method (28) (**Figure 3D**). Hence, a pronounced difference of the TFE and BB GeLC prepared protein complement was shown, demonstrating a high degree of complementarity of these methods. The latter may be attributed to the different chemical attributes of the solvents used. TFE is a small fluoroorganic compound that acts phenol-like in many reactions, though the mechanism of protein solubilization by TFE is not completely understood at present (Kundu and Kishore, 2004). The anionic, micelle forming detergent SDS is a large molecule consisting of a long apolar alkyl chain ( $C_{12}$ ) esterified to sulfuric acid. Protein solubilization is achieved by hydrophobic interaction of its alkyl-moiety with (hydrophobic) amino acids of proteins causing unfolding due to electrostatic repulsion. Based on these chemical differences, the applied methods apparently have (i) a different effectiveness of cell lysis for different bacteria of the sampled community [i.e., in combination with ultrasound (TFE) or bead beating (BB GeLC), respectively] and (ii) a different protein accessibility ultimately affecting solubilization. Efficient protein extraction from environmental samples using TFE was previously reported for an acid mine drainage biofilm sample (Thompson et al., 2008), which, however, exhibits a rather low organismic diversity and comparison to SDS extraction has not been performed. It would be of interest to understand, if different bacterial (or archaeal) phyla are preferentially lysed by either of the methods. This could lead to misinterpretation of obtained metaproteomic data if only one method would be applied. Unfortunately, the very low share of phylogenetically allocable proteins within the present study does not allow for such a survey.

The rather large volume of sample obtained by the TFE-lysis protocol allowed for peptide analysis by two different mass spectrometric methods following initial nanoLC separation: (i) online by ESI-iontrap MS and (ii) offline by MALDI-TOF MS. In addition to the different ionization (ESI vs. MALDI) and mass analysis techniques (iontrap vs. TOF), the methods differ with respect to precursor selection for MS/MS fragmentation and fragment analysis (for overview see Wöhlbrand et al., 2013). About 50% of the detected proteins from the TFE-lysis samples were exclusively detected by either one of the two applied MS methods (**Figure 5**). The determined complementarity is even more pronounced than previously reported for the analysis of bovine mitochondrial ribosomes (63% overlap) applying both methods (Bodnar et al., 2003). This proportion is rather similar when only one condition is considered (**Figure 5**) as well as in the case of the predicted membrane proteins (not shown). While proteins detected by both ESI-iontrap MS and MALDI-TOF MS revealed on average two protein identifying peptides and average sequence coverages of  $\sim 24\%$  (23.6% ESI-MS vs.



25.4% MALDI-MS), the average Mascot score of ESI-iontrap MS identified proteins was 127, that of MALDI-TOF MS identified proteins was 160 (Table 2), demonstrating the superior mass accuracy of the TOF mass analyzer.

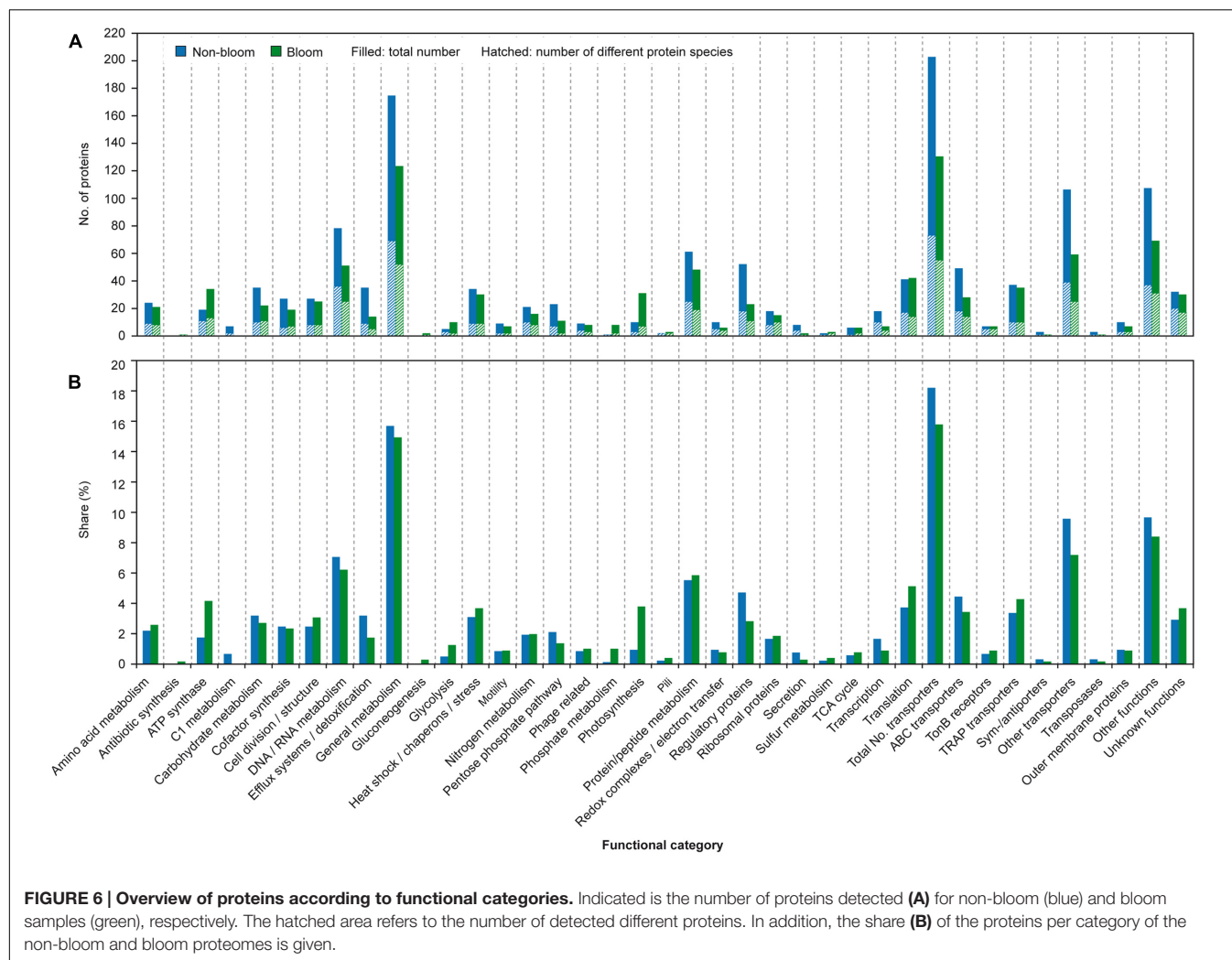
## Bacterioplankton Protein Complements are Affected by the Bloom

Identified proteins were categorized according to their predicted functions, revealing that nearly a third of all proteins belong to transporters (17.1%) and proteins of general metabolism (15.3%), respectively (Figure 6 and Supplementary Table S2). Other

abundant categories include DNA/RNA (6.7%), protein/peptide (5.6%) or carbohydrate metabolism (3.0%), translation (4.3%), regulatory proteins (3.9%) as well as proteins of unknown functions (3.2%). Although the relative proportion of all categories within the station protein complements is rather similar, the metaproteomic analysis revealed differences between bloom and non-bloom conditions with respect to the number of detected proteins of the different functional categories as well as their relative share of the respective station-proteome.

Outside of the bloom, the number (202 vs. 130) as well as the share of detected transport proteins among the respective station proteome (18.2% vs. 15.8%) is pronouncedly higher as compared to bloom conditions (Figure 6 and Supplementary Table S2). The high share of transport proteins is similar to the SAR11 metaproteome of the Sargasso Sea (17.4%; Sowell et al., 2009) as well as surface waters of the South Atlantic (Morris et al., 2010). The elevated number during non-bloom conditions determined in this study can be ascribed to a 30% higher share of ABC transporter-related proteins (as compared to the bloom conditions). Moreover, a nearly two-fold higher number of different unclassified transport proteins were detected outside the bloom. These unclassified transport proteins comprise a large number of unclassified periplasmic binding proteins that account for 9.5% of the station proteome (Figure 6 and Supplementary Table S2). Furthermore, the larger share of detected proteins involved in protein secretion outside the bloom (0.7% vs. 0.2%) agrees with the higher amount of extracellular binding proteins in this habitat. Overall, the high number of detected ABC-type transport proteins under the non-bloom condition is reminiscent of the situation in the oligotrophic Sargasso Sea (Sowell et al., 2009). Hence, formation of a large number of high affinity transport systems may be beneficial under nutrient limited conditions to scavenge the rare nutrients, a hypothesis previously also raised by Hirsch et al. (1979) and Button (1993). Furthermore, potential involvement of (some) periplasmic binding proteins in signal transduction





and chemotaxis (Mowbray and Sandgren, 1998) may facilitate migration toward preferred habitat conditions.

In addition to transport proteins, a higher number and share of regulatory proteins as well as proteins involved in transcription (mainly RNA polymerase subunits) and DNA/RNA metabolism were detected outside the algal bloom (**Figure 6** and **Supplementary Table S2**). The increased abundance of these proteins may represent a state of readiness during the oligotrophic conditions, to quickly respond to nutrients becoming available by promptly forming proteins for their utilization.

Within the bloom, the share of proteins related to photosynthesis was more than two times higher (2.5%) as compared to non-bloom conditions (0.9%) (**Figure 6** and **Supplementary Table S1**) agreeing with higher photosynthetic activity at this station. Notably, these proteins also include a bacteriorhodopsin detected at both stations (TFE lysis combined with MALDI-TOF detection), despite its challenging detectability due to its hydrophobicity and generally low abundance (Stapels et al., 2004). Rhodopsins were previously also detected by membrane targeting proteomics in nutrient-poor

surface waters of the southern Atlantic, the productive Benguela upwelling region (Morris et al., 2010) and coastal pacific waters (Giovannoni et al., 2005), suggesting a function not restricted to low-nutrient conditions. Besides photosynthetic proteins, a higher share of ATP synthase proteins (4.1% vs. 1.7%) and proteins involved in translation (5.1% vs. 3.7%) within the bloom may be indicative for a higher metabolic activity of the bacterial community, reflecting the higher nutrient availability. Correspondingly, a higher transcript level of tRNA synthetases within the bloom was previously determined for the same samples (Wemheuer et al., 2015).

Although a similar number of different proteins related to carbohydrate metabolism was detected at both stations (i.e., 11), both identified glycoside hydrolases were detected inside, but only one outside the bloom (**Figure 6** and **Supplementary Table S2**). A higher abundance of such carbohydrate-active enzymes (CAZymes) was previously reported for algal bloom conditions in the North Sea (Teeling et al., 2012) and respective transcripts were detected to be more abundant within the bloom for the same samples as studied here (Wemheuer et al., 2015). The rather low difference between

the bloom and non-bloom conditions may be due to the observed increase and subsequent decrease during the succession of the bloom (Teeling et al., 2012). The increased transcript level but rather low difference in protein amount may point to sampling during the increasing phase, so that transcripts may not yet be completely translated into proteins.

Interestingly, different catabolic strategies for glucose seem to be applied under the analyzed conditions. A higher share of proteins involved in the Enter-Doudoroff (ED) and pentose phosphate (PP) pathway was present outside the bloom (2.1% vs. 1.3%), while the share of glycolysis proteins was higher within the bloom (0.5% vs. 1.2%) (**Figure 6** and **Supplementary Table S2**). Such increased abundance of ED and PP pathway proteins was previously observed in batch culture for *Phaeobacter inhibens* DSM 17395, a member of the *Roseobacter* group, during growth in minimal medium as compared to complex medium (Zech et al., 2013) agreeing with the *in situ* observation.

Proteins of the C1 pathway were pronouncedly more abundant outside the bloom, comprising 1.0% of the station proteome (0.1% inside the bloom) (**Figure 6** and **Supplementary Table S2**) indicating activity of methylotrophic bacteria. Correspondingly, the active community outside the bloom revealed a pronouncedly higher share of methylotrophic OM43 clade members as compared to bloom conditions (Wemheuer et al., 2015) supporting the metaproteomic finding. Abundant detection of methanol dehydrogenase affiliated to OM43 members was previously also reported for the east Pacific during coastal upwelling (Sowell et al., 2011). Sowell et al. (2011) suggested that methanol, possibly produced by phytoplankton, may represent a significant source of carbon and energy in coastal ecosystems. Detection of C1 pathway related proteins under non-bloom conditions may indicate that also other one carbon compounds, including trimethylamine or dimethylsulfoniopropionate, potentially serve as substrates for methylotrophs in coastal areas.

In accordance with the competition of bacteria and algae for the macro-nutrient phosphate within the algal bloom, proteins involved in phosphate uptake and starvation were more numerous inside the bloom (1.0% vs. 0.1%), which is consistent with previous observations reported by Teeling et al. (2012). The limited amount of nitrogen outside the bloom is reflected by the detection of a higher number of glutamine synthetases (3 out of 4) as well as amidohydrolases (4 out of 5) at this station (only 2 and 1 inside the bloom, respectively) (**Figure 6** and **Supplementary Table S1**). Similarly, proteins involved in nitrogen metabolism, including glutamine synthetase and the nitrogen regulatory protein PII (effecting glutamine synthase activity) were detected in other oligotrophic marine environments (Sowell et al., 2009; Georges et al., 2014). Increased amounts of ammonia assimilating glutamine synthase were reported for diverse bacteria in response to nitrogen limitation (for overview see Leigh and Dodsworth, 2007), which is in accordance with the *in situ* observations of this study.

The similar number of phage-related proteins present in both investigated samples indicates a comparable phage

impact under both environmental conditions (**Figure 6** and **Supplementary Table S1**). Detection of station-exclusive phage proteins may be attributed to the differing composition of the community (Wemheuer et al., 2015) which might provoke activity of phages specific for respective phyla.

## CONCLUSION

This study demonstrates the complementarity and, hence, the value of using two different sample preparation and two different ionization and mass analysis techniques for a comprehensive characterization of environmental samples. Furthermore, a considerable number of membrane proteins were covered without specifically preparing the membrane fraction. This approach revealed that the non-bloom microbial community applies the ED and PP as well as C1 pathway for carbon metabolism and is prepared to adapt to changing conditions as evident from the high share of proteins involved in e.g., high affinity transport, regulation or transcription. In contrast, the bloom community focusses on metabolism of the nutrients present and simultaneously competes with the algae for limited macro-nutrients. Overall, however, the usually limited amount of environmental samples as well as the commonly low cell numbers on the filters restricts repeated analyses accomplishable per sample.

## AUTHOR CONTRIBUTIONS

LW and RR conceived and designed the experiments; BW performed sampling on BB Heinke and generated the metagenomic/-transcriptomic database; LW, HR, and CH performed the proteomic experiments; LW, CF, and BW analyzed the data; LW, BW, CF, and RR wrote the paper; all authors reviewed, edited and approved the manuscript.

## ACKNOWLEDGMENT

This study was supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the collaborative research center TRR51.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.00442/full#supplementary-material>

**FIGURE S1 | Share of replicate protein identifications per station.** Given is the number of repetitive protein identifications per station sample (non-bloom, blue; bloom, green) and their respective share of the station proteome.

**TABLE S1 | Summary of all identified proteins.**

**TABLE S2 | Functional categorization of identified proteins.** Given is the number and share of identified proteins in total as well as per station (in alphabetical order).

## REFERENCES

- Abraham, P. E., Giannone, R. J., Xiong, W., and Hettich, R. L. (2014). Metaproteomics: extracting and mining proteome information to characterize metabolic activities in microbial communities. *Curr. Protoc. Bioinformatics* 46, 13.26.1–14.13.26. doi: 10.1002/0471250953.bi1326s46
- Alderkamp, A.-C., Sintes, E., and Herndl, G. (2006). Abundance and activity of major groups of prokaryotic plankton in the coastal North Sea during spring and summer. *Aquat. Microb. Ecol.* 45, 237–246. doi: 10.3354/ame045237
- Alonso, C., and Pernthaler, J. (2006a). Concentration-dependent patterns of leucine incorporation by coastal picoplankton. *Appl. Environ. Microbiol.* 72, 2141–2147. doi: 10.1128/AEM.72.3.2141-2147.2006
- Alonso, C., and Pernthaler, J. (2006b). Roseobacter and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environ. Microbiol.* 8, 2022–2030. doi: 10.1111/j.1462-2920.2006.01082.x
- Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Aust. Ecol.* 26, 32–46.
- Arrigo, K. R. (2005). Marine microorganisms and global nutrient cycles. *Nature* 437, 349–355. doi: 10.1038/nature04159
- Bodnar, W. M., Blackburn, R. K., Krise, J. M., and Moseley, M. A. (2003). Exploiting the complementary nature of LC/MALDI/MS/MS and LC/ESI/MS/MS for increased proteome coverage. *J. Am. Soc. Mass Spectrom.* 14, 971–979. doi: 10.1016/S1044-0305(03)00209-5
- Bray, J., and Curtis, J. (1957). An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* 27, 325–349. doi: 10.2307/1942268
- Bunse, C., and Pinhassi, J. (2017). Marine bacterioplankton seasonal succession dynamics. *Trends Microbiol.* doi: 10.1016/j.tim.2016.12.013 [Epub ahead of print].
- Button, D. K. (1993). Nutrient-limited microbial growth kinetics: overview and recent advances. *Antonie Van Leeuwenhoek* 63, 225–235. doi: 10.1007/BF00871220
- Chen, J., Ryu, S., Gharib, S. A., Goodlett, D. R., and Schnapp, L. M. (2008). Exploration of the normal human bronchoalveolar lavage fluid proteome. *Proteom. Clin. Appl.* 2, 585–595. doi: 10.1002/prca.200780006
- Chertov, O., Biragyn, A., Kwak, L. W., Simpson, J. T., Boronina, T., Hoang, V. M., et al. (2004). Organic solvent extraction of proteins and peptides from serum as an effective sample preparation for detection and identification of biomarkers by mass spectrometry. *Proteomics* 4, 1195–1203. doi: 10.1002/pmic.200300677
- Cole, J. J., Findlay, S., and Pace, M. L. (1988). Bacterial production in fresh and saltwater ecosystems: a cross-system overview. *Mar. Ecol. Prog. Ser.* 43, 1–10. doi: 10.3354/meps043001
- Ducklow, H., Kirchmann, D. L., Quinby, H. L., Carlson, C. A., and Dam, H. G. (1993). Stocks and dynamics of bacterioplankton carbon during the spring bloom in the eastern North Atlantic Ocean. *Deep Sea Res. II* 40, 245–263. doi: 10.1016/0967-0645(93)90016-G
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461
- Ferro, M., Seigneurin-Berny, D., Rolland, N., Chapel, A., Salvi, D., Garin, J., et al. (2000). Organic solvent extraction as a versatile procedure to identify hydrophobic chloroplast membrane proteins. *Electrophoresis* 21, 3517–3526. doi: 10.1002/1522-2683(20001001)21:16<3517::AID-ELPS3517>3.0.CO;2-H
- Frias-Lopez, J., Shi, Y., Tyson, G. W., Coleman, M. L., Schuster, S. C., Chisholm, S. W., et al. (2008). Microbial community gene expression in ocean surface waters. *Proc. Natl. Acad. Sci. U.S.A.* 105, 3805–3810. doi: 10.1073/pnas.0708897105
- Fuerst, J. A. (2005). Intracellular compartmentation in planctomycetes. *Annu. Rev. Microbiol.* 59, 299–328. doi: 10.1146/annurev.micro.59.030804.121258
- Georges, A. A., El-Swaiss, H., Craig, S. E., Li, W. K., and Walsh, D. A. (2014). Metaproteomic analysis of a winter to spring succession in coastal northwest Atlantic Ocean microbial plankton. *ISME J.* 8, 1301–1313. doi: 10.1038/ismej.2013.234
- Geyer, C. (1991). “Markov chain Monte Carlo maximum likelihood,” in *Computing Science and Statistics Proceeding of the 23rd Symposium Interface*, eds E. M. Keramidas and S. M. Kaufman (Fairfax Station, VA: Interface Foundation of North America), 156–163.
- Gifford, S. M., Sharma, S., Rinta-Kanto, J. M., and Moran, M. A. (2011). Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. *ISME J.* 5, 461–472. doi: 10.1038/ismej.2010.141
- Giovannoni, S. J., Bibbs, L., Cho, J. C., Stapels, M. D., Desiderio, R., Vergin, K. L., et al. (2005). Proteorhodopsin in the ubiquitous marine bacterium SAR11. *Nature* 438, 82–85. doi: 10.1038/nature04032
- Hettich, R. L., Sharma, R., Chourey, K., and Giannone, R. J. (2012). Microbial metaproteomics: identifying the repertoire of proteins that microorganisms use to compete and cooperate in complex environmental communities. *Curr. Opin. Microbiol.* 15, 373–380. doi: 10.1016/j.mib.2012.04.008
- Hirsch, P., Bernhard, M., Cohen, S. S., Ensign, J. C., Jannasch, H. W., Koch, A. L., et al. (1979). “Life under conditions of low nutrient concentrations,” in *Strategies of Microbial Life in Extreme Environments*, ed. M. Shilo (Weinheim: Verlag Chemie), 357–372.
- Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119
- Kerner, M., and Herndl, G. (1992). Extracellular enzymatic activity and secondary production in free-living and marine-snow-associated bacteria. *Mar. Biol.* 113, 341–347.
- Klindworth, A., Mann, A. J., Huang, S., Wichels, A., Quast, C., Waldmann, J., et al. (2014). Diversity and activity of marine bacterioplankton during a diatom bloom in the North Sea assessed by total RNA and pyrotag sequencing. *Mar. Genomics* 18(Pt B), 185–192. doi: 10.1016/j.margen.2014.08.007
- Kopylova, E., Noé, L., and Touzet, H. (2012). SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28, 3211–3217. doi: 10.1093/bioinformatics/bts611
- Kossmehl, S., Wöhlbrand, L., Drüppel, K., Feenders, C., Blasius, B., and Rabus, R. (2013). Subcellular protein localization (cell envelope) in *Phaeobacter inhibens* DSM 17395. *Proteomics* 13, 2743–2760. doi: 10.1002/pmic.201300112
- Kruskal, J., and Wish, M. (1978). *Multidimensional Scaling*. Thousand Oaks, CA: SAGE Publication. doi: 10.4135/9781412985130
- Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29, 1–27. doi: 10.1007/BF02289565
- Kundu, A., and Kishore, N. (2004). Interaction of 2,2,2-trifluoroethanol with proteins: calorimetric, densimetric and surface tension approach. *Biophys. Chem.* 109, 427–442. doi: 10.1016/j.bpc.2003.12.009
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923
- Leigh, J. A., and Dodsworth, J. A. (2007). Nitrogen regulation in bacteria and archaea. *Annu. Rev. Microbiol.* 61, 349–377. doi: 10.1146/annurev.micro.61.080706.093409
- Liu, H., Sadygov, R. G., and Yates, J. R. III. (2004). A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* 76, 4193–4201. doi: 10.1021/ac0498563
- Mande, S. S., Mohammed, M. H., and Ghosh, T. S. (2012). Classification of metagenomic sequences: methods and challenges. *Brief. Bioinform.* 13, 669–681. doi: 10.1093/bib/bbs054
- McQuatters-Gollop, A., Raitos, D. E., Edwards, M., Pradhan, Y., Mee, L. D., and Lavender, S. J. (2007). A long-term chlorophyll data set reveals regime shift in North Sea phytoplankton biomass unconnected to nutrient trends. *Limnol. Oceanogr.* 52, 635–648. doi: 10.4319/lo.2007.52.2.0635
- Meinicke, P. (2015). UProC: tools for ultra-fast protein domain classification. *Bioinformatics* 31, 1382–1388. doi: 10.1093/bioinformatics/btu843
- Morris, R. M., Nunn, B. L., Frazer, C., Goodlett, D. R., Ting, Y. S., and Rocap, G. (2010). Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J.* 4, 673–685. doi: 10.1038/ismej.2010.4
- Mowbray, S. L., and Sandgren, M. O. (1998). Chemotaxis receptors: a progress report on structure and function. *J. Struct. Biol.* 124, 257–275. doi: 10.1006/jsbi.1998.4043
- Neuhoff, V., Arold, N., Taube, D., and Ehrhardt, W. (1988). Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. *Electrophoresis* 9, 255–262. doi: 10.1002/elps.1150090603
- Ram, R. J., Verberkmoes, N. C., Thelen, M. P., Tyson, G. W., Baker, B. J., Blake, R. C., et al. (2005). Community proteomics of a natural microbial biofilm. *Science* 308, 1915–1920. doi: 10.1126/science.1109070
- Rodriguez, R. L., and Konstantinidis, K. T. (2014). Nonpareil: a redundancy-based approach to assess the level of coverage in metagenomic datasets. *Bioinformatics* 30, 629–635. doi: 10.1093/bioinformatics/btt584

- Schoemann, V., Becquevort, S., Stefels, J., Rousseau, V., and Lancelot, C. (2005). *Phaeocystis* blooms in the global ocean and their controlling mechanisms: a review. *J. Sea Res.* 53, 43–66. doi: 10.1016/j.seares.2004.01.008
- Sowell, S. M., Abraham, P. E., Shah, M., Verberkmoes, N. C., Smith, D. P., Barofsky, D. F., et al. (2011). Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *ISME J.* 5, 856–865. doi: 10.1038/ismej.2010.168
- Sowell, S. M., Wilhelm, L. J., Norbeck, A. D., Lipton, M. S., Nicora, C. D., Barofsky, D. F., et al. (2009). Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J.* 3, 93–105. doi: 10.1038/ismej.2008.83
- Stapels, M. D., Cho, J. C., Giovannoni, S. J., and Barofsky, D. F. (2004). Proteomic analysis of novel marine bacteria using MALDI and ESI mass spectrometry. *J. Biomol. Tech.* 15, 191–198.
- Swendsen, R. H., and Wang, J. S. (1986). Replica Monte Carlo simulation of spin glasses. *Phys. Rev. Lett.* 57, 2607–2609. doi: 10.1103/PhysRevLett.57.2607
- Teeling, H., Fuchs, B. M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C. M., et al. (2012). Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* 336, 608–611. doi: 10.1126/science.1218344
- Thompson, M. R., Chourey, K., Froelich, J. M., Erickson, B. K., Verberkmoes, N. C., and Hettich, R. L. (2008). Experimental approach for deep proteome measurements from small-scale microbial biomass samples. *Anal. Chem.* 80, 9517–9525. doi: 10.1021/ac801707s
- Veldhuis, M. J. W., and Wassmann, P. (2005). Bloom dynamics and biological control of a high biomass HAB species in European coastal waters: a *Phaeocystis* case study. *Harmful Algae* 4, 805–809. doi: 10.1016/j.hal.2004.12.004
- Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304, 66–74. doi: 10.1126/science.1093857
- Vila-Costa, M., Gasol, J. M., Sharma, S., and Moran, M. A. (2012). Community analysis of high- and low-nucleic acid-containing bacteria in NW Mediterranean coastal waters using 16S rDNA pyrosequencing. *Environ. Microbiol.* 14, 1390–1402. doi: 10.1111/j.1462-2920.2012.02720.x
- Vizcaino, J. A., Csordas, A., Del-Toro, N., Dianas, J. A., Griss, J., Lavidas, I., et al. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res.* 44, D447–D456. doi: 10.1093/nar/gkv1145
- Wang, H., Qian, W.-J., Mottaz, H. M., Clauss, T. R. W., Anderson, D. J., Moore, R. J., et al. (2005). Development and evaluation of a micro- and nano-scale proteomic sample preparation method. *J. Proteome Res.* 4, 2397–2403. doi: 10.1021/pr050160f
- Wemheuer, B., Güllert, S., Billerbeck, S., Giebel, H. A., Voget, S., Simon, M., et al. (2014). Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol. Ecol.* 87, 378–389. doi: 10.1111/1574-6941.12230
- Wemheuer, B., Wemheuer, F., Hollensteiner, J., Meyer, F. D., Voget, S., and Daniel, R. (2015). The green impact: bacterioplankton response toward a phytoplankton spring bloom in the southern North Sea assessed by comparative metagenomic and metatranscriptomic approaches. *Front. Microbiol.* 6:805. doi: 10.3389/fmicb.2015.00805
- Wilmes, P., and Bond, P. L. (2004). The application of two-dimensional polyacrylamide gel electrophoresis and downstream analyses to a mixed community of prokaryotic microorganisms. *Environ. Microbiol.* 6, 911–920. doi: 10.1111/j.1462-2920.2004.00687.x
- Wiltshire, K. H., Kraberg, A., Bartsch, I., Boersma, M., Franke, H.-D., Freund, J., et al. (2010). Helgoland roads: 45 years of change. *Estuaries Coasts* 33, 295–310. doi: 10.1007/s12237-009-9228-y
- Wiltshire, K. H., Malzahn, A. M., Wirtz, K., Greve, W., Janisch, S., Mangelsdorf, P., et al. (2008). Resilience of North Sea phytoplankton spring bloom dynamics: an analysis of long-term data at Helgoland Roads. *Limnol. Oceanogr.* 53, 1294–1302. doi: 10.4319/lo.2008.53.4.1294
- Wöhlbrand, L., Ruppertsberg, H. S., Feenders, C., Blasius, B., Braun, H. P., and Rabus, R. (2016). Analysis of membrane-protein complexes of the marine sulfate reducer *Desulfobacula toluolica* Tol2 by 1D blue native-PAGE complexome profiling and 2D blue native-/SDS-PAGE. *Proteomics* 16, 973–988. doi: 10.1002/pmic.201500360
- Wöhlbrand, L., Trautwein, K., and Rabus, R. (2013). Proteomic tools for environmental microbiology—a roadmap from sample preparation to protein identification and quantification. *Proteomics* 13, 2700–2730. doi: 10.1002/pmic.201300175
- Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46
- Zech, H., Echtermeyer, C., Wöhlbrand, L., Blasius, B., and Rabus, R. (2011). Biological versus technical variability in 2-D DIGE experiments with environmental bacteria. *Proteomics* 11, 3380–3389. doi: 10.1002/pmic.201100071
- Zech, H., Hensler, M., Kossmehl, S., Drüppel, K., Wöhlbrand, L., Trautwein, K., et al. (2013). Adaptation of *Phaeobacter inhibens* DSM 17395 to growth with complex nutrients. *Proteomics* 13, 2851–2868. doi: 10.1002/pmic.201200513

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Wöhlbrand, Wemheuer, Feenders, Ruppertsberg, Hinrichs, Blasius, Daniel and Rabus. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Exoproteome Analysis of the Seaweed Pathogen *Nautella italica* R11 Reveals Temperature-Dependent Regulation of RTX-Like Proteins

Melissa Gardiner<sup>1\*</sup>, Adam M. Bournazos<sup>1</sup>, Claudia Maturana-Martinez<sup>1</sup>, Ling Zhong<sup>2</sup> and Suhelen Egan<sup>1\*</sup>

<sup>1</sup> School of Biological Earth and Environmental Sciences—Centre for Marine Bio-Innovation, The University of New South Wales, Sydney, NSW, Australia, <sup>2</sup> Bioanalytical Mass Spectrometry Facility, Mark Wainwright Analytical Centre, The University of New South Wales, Sydney, NSW, Australia

## OPEN ACCESS

### Edited by:

Rolf Daniel,  
University of Göttingen, Germany

### Reviewed by:

Anna Carratalà,  
École Polytechnique Fédérale  
de Lausanne, Switzerland  
Haiwei Luo,  
The Chinese University of Hong Kong,  
Hong Kong

### \*Correspondence:

Suhelen Egan  
s.egan@unsw.edu.au  
Melissa Gardiner  
melissa.gardiner@sydney.edu.au

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 10 April 2017

**Accepted:** 13 June 2017

**Published:** 29 June 2017

### Citation:

Gardiner M, Bournazos AM,  
Maturana-Martinez C, Zhong L and  
Egan S (2017) Exoproteome Analysis  
of the Seaweed Pathogen *Nautella*  
*italica* R11 Reveals  
Temperature-Dependent Regulation  
of RTX-Like Proteins.  
Front. Microbiol. 8:1203.  
doi: 10.3389/fmicb.2017.01203

Climate fluctuations have been linked to an increased prevalence of disease in seaweeds, including the red alga *Delisea pulchra*, which is susceptible to a bleaching disease caused by the bacterium *Nautella italica* R11 under elevated seawater temperatures. To further investigate the role of temperature in the induction of disease by *N. italica* R11, we assessed the effect of temperature on the expression of the extracellular proteome (exoproteome) in this bacterium. Label-free quantitative mass spectrometry was used to identify 207 proteins secreted into supernatant fraction, which is equivalent to 5% of the protein coding genes in the *N. italica* R11 genome. Comparative analysis demonstrated that expression of over 30% of the *N. italica* R11 exoproteome is affected by temperature. The temperature-dependent proteins include traits that could facilitate the ATP-dependent transport of amino acid and carbohydrate, as well as several uncharacterized proteins. Further, potential virulence determinants, including two RTX-like proteins, exhibited significantly higher expression in the exoproteome at the disease inducing temperature of 24°C relative to non-inducing temperature (16°C). This is the first study to demonstrate that temperature has an influence exoproteome expression in a macroalgal pathogen. The results have revealed several temperature regulated candidate virulence factors that may have a role in macroalgal colonization and invasion at elevated sea-surface temperatures, including novel RTX-like proteins.

**Keywords:** Roseobacter, virulence, secretome, marine bacteria, RTX-toxin, macroalgae, symbiosis, proteomics

## INTRODUCTION

Seaweeds (macroalgae) are key ecosystem engineers of temperate marine coastal habitats. They provide a protective habitat and nursery for many other species and as primary producers they act as a major food source (Wernberg et al., 2011). Unfortunately these important marine habitat formers are also susceptible to disease as a result of environmental stress and/or microbial pathogens (Gachon et al., 2010; Hollants et al., 2013; Egan et al., 2014). For example the red macroalga *Delisea pulchra* undergoes bleaching of its thallus corresponding to increased sea-surface temperatures (Campbell et al., 2011; Case et al., 2011). A combination of laboratory and

field experiments have demonstrated that under increased host stress (such as that imposed by higher temperatures) *D. pulchra* has reduced levels of its natural chemical defense (furanones) (Campbell et al., 2011; Case et al., 2011). This reduced algal defense results in the increased susceptibility to infection by bacterial pathogens, including the Roseobacter species *Nautella italica* R11 (formally *Ruegeria* sp. R11) under the elevated sea-surface temperatures of 24°C (Campbell et al., 2011; Case et al., 2011; Fernandes et al., 2011; Kumar et al., 2016).

Genomic analysis of *N. italica* R11 has provided insight into the potential virulome of the marine pathogen, which includes a range of transporter systems that could facilitate both the uptake of algal metabolites and the secretion of toxins and/or degradative proteins (Fernandes et al., 2011). The *N. italica* R11 genome encodes for several proteins with homology to repeats-in-toxin (RTX) proteins that function as cytotoxins in other pathogens (Linhartova et al., 2010), and these proteins may mediate tissue damage on a *D. pulchra* host (Fernandes et al., 2011). However, with the exception of recent studies demonstrating the role of oxidative stress resistance and quorum sensing-mediated colonization (Gardiner et al., 2015a,b), the relative importance of other virulence traits in *N. italica* R11 remains unknown.

High throughput mass spectrometric techniques permit the identification of the subset of bacterial proteins secreted in the extracellular milieu (the exoproteome) (Desvaux et al., 2009; Christie-Oleza and Armengaud, 2010). Recent data for Roseobacter clade members has highlighted the benefit of exoproteome analysis for the identification of proteins previously overlooked by genomic studies but important for bacterial physiology/host interactions (Zech et al., 2009; Christie-Oleza and Armengaud, 2010; Christie-Oleza et al., 2012b; Kossmehl et al., 2013; Christie-Oleza and Armengaud, 2015). Moreover, exoproteome analysis of human, animal, and plant pathogens has revealed new putative virulence determinants (Madec et al., 2014; Campbell et al., 2015; He et al., 2015; Mandelc and Javornik, 2015), and provided insight into the temperature-dependent regulation of protein secretion (Kimes et al., 2012). However, to date, nothing is known about the exoproteome composition or the influence of temperature on protein expression in seaweed pathogens.

Here, we use label free mass spectrometry to investigate for the first time the exoproteome of the seaweed pathogen *N. italica* R11. As elevated sea-surface temperature plays a key role in the induction of the seaweed bleaching disease by *N. italica* R11, we also compared the exoproteome of *N. italica* R11 exposed to disease (24°C) and non-disease (16°C) inducing temperatures to investigate potential virulence factors that could facilitate pathogenesis in this macroalgal pathogen.

## MATERIALS AND METHODS

### Bacterial Culture Conditions Used in the Study

*Nautella italica* R11 cultures were routinely grown in half strength (18.4 g/l) marine broth (Difco, Becton Dickinson, United States) with shaking (180 rpm) at room temperature (21°C) from

cultures stored at −80°C in 10% glycerol stocks. The cultures used for protein analysis were inoculated at 10% v/v using an overnight culture (grown directly from a frozen stock) harvested at an OD<sub>600</sub> = 1 and rinsed twice in fresh media before being incubated in 200 ml of media at either 24°C (disease-inducing temperature) or 16°C (non-disease-inducing temperature). Cells were harvested by centrifugation at early stationary phase, as determined using spectrophotometric analysis (Supplementary Figure S1). Early stationary phase was selected for analysis as previous studies investigating the virulence of *N. italica* R11 were conducted using bacteria at this growth phase (Case et al., 2011; Fernandes et al., 2011) and the expression of virulence and secondary metabolites has been linked to stationary growth phase in other marine bacteria (Egan et al., 2002; Vanden Bergh et al., 2013).

### Label-Free Quantitation of the *N. italica* R11 Exoproteome

To obtain the supernatant fraction (exoproteome) of *N. italica* R11 cells, 200 ml of culture was treated with protease inhibitor cocktail (Sigma) before filtration through a 0.2 µm Millex syringe filter unit (Merck Millipore). Peptides ≥ 3 kDa were then obtained by centrifugation through an Amicon Ultra-4 unit (Merck Millipore) according to the manufacturer's instructions. The concentration of the eluted proteins was determined using a 2-D Quant Kit (GE Healthcare) according to the manufacturer's instructions.

The supernatant proteins were separated and digested in-gel using SDS-PAGE due to the presence of low molecular weight peptides from the cell growth media (Additional file, Supplementary Figure S2). Fifteen micrograms of protein were denatured as described previously (Burgos-Portugal et al., 2012), separated on an any kD™ Mini-PROTEAN® TGX™ Precast Gel (Bio-Rad) by electrophoresis for 1.5 h at 100 V and stained using SimplyBlue™ SafeStain (Invitrogen). Once visualized the proteins on the SDS-PAGE gel were cut into three sections (Supplementary Figure S2), placed into three microfuge tubes, and destained (with 25% acetonitrile/25% NH<sub>4</sub>HCO<sub>3</sub>). Fifty microliters of the reducing agent (10 mM DTT, 50 mM NH<sub>4</sub>HCO<sub>3</sub>) was added before incubation at 37°C for 30 min. The reducing agent was removed and 50 µl of the cysteine-blocking reagent (200 mM Iodoacetamide, 100 mM NH<sub>4</sub>HCO<sub>3</sub>) was added before incubated at 37°C for 30 min. The gel pieces were dehydrated by adding 100% acetonitrile to each tube. The acetonitrile was then removed and 60 µl of 2 ng/µl trypsin in 20 mM NH<sub>4</sub>HCO<sub>3</sub> was added to digest the proteins at 37°C overnight. The digested peptides were then solubilized in 1% formic acid, 0.05 HFBA (heptafluorobutyric acid) and the peptides (corresponding to the three gel slices) were pooled. Digested peptides were separated by nano-LC using an Ultimate 3000 HPLC and autosampler system (Dionex, Amsterdam, Netherlands) as described previously (Kaakoush et al., 2015). High voltage (2000 V) was applied to low volume tee (Upchurch Scientific) and the column tip positioned ~0.5 cm from the heated capillary (*T* = 280°C) of an Orbitrap Velos (Thermo Electron, Bremen, Germany) mass spectrometer. Positive ions

were generated by electrospray and the Orbitrap operated in data dependent acquisition mode (DDA). A survey scan  $m/z$  350–1750 was acquired in the Orbitrap (Resolution = 30,000 at  $m/z$  400, with an accumulation target value of 1,000,000 ions) with lockmass enabled. Up to the 10 most abundant ions ( $>5,000$  counts) with charge states  $> +2$  were sequentially isolated and fragmented within the linear ion trap using collisionally induced dissociation with an activation  $q = 0.25$  and activation time of 30 ms at a target value of 30,000 ions.  $M/z$  ratios selected for tandem mass spectrometry (MS/MS) were dynamically excluded for 30 s. Mass spectrometry peak lists were generated using Mascot Daemon/extract\_msn (Matrix Science) using the default parameters and a decoy false discovery rate of  $p$ -value  $< 0.01$ , and were submitted to the database search program Mascot (version 2.1). Mascot determined the peptides with ion score cut-off at 20. Search parameters were: precursor tolerance 4 ppm and product ion tolerances  $\pm 0.4$  Da, alkylation of cysteine as fixed modification, methionine oxidation as variable modification, enzyme specificity was trypsin, 1 missed cleavage was possible, and the NCBI non-redundant (nr) database searched (accessed May 2009). Each MS/MS spectrum was compared to an in-house *N. italica* R11 genome (available via IMG, Genome ID 647533206) (Markowitz et al., 2006). Three technical replicates were performed for each biological replicate.

## Analysis of the Exoproteome Data

Progenesis<sup>®</sup> QI for Proteomics (version 2.0, Nonlinear Dynamics, United Kingdom) was used to analyze the mass spectrometry data. The acquired spectra were loaded into the Progenesis<sup>®</sup> software and the ion intensities of six runs (three replicates at 16°C, three at 24°C) were examined and label-free quantification was performed. Replicate 1 for temperature condition 16°C was chosen as the reference and the retention times of all six samples were aligned. Features with only one charge and over four changes were excluded from further analysis. Samples were grouped to their experimental condition (16°C versus 24°C). One-way analysis of variance (ANOVA) statistical analysis was performed using transformed normalized abundances. Only proteins represented by two or more peptides, including at least one unique peptide, were considered in the exoproteome analysis. Proteins with a fold change  $\pm 2$  and  $p$ -value  $< 0.05$  were considered significantly differentially expressed across the two temperature conditions. Identified *N. italica* R11 proteins are described as GenBank accession numbers (Wheeler et al., 2008) and assignment of clusters of orthologous groups (COG) was performed using the IMG/ER database (Markowitz et al., 2006). The predicted exoproteome was determined using the automated SignalP 3.0 analysis (Bendtsen et al., 2004) of the *N. italica* R11 genome in IMG/ER and by manually searching the genome for proteins containing the TAT (twin-arginine translocation) pathway signal sequence [TIGRfam family (TIGR01409)] in the translated nucleotide sequence (Berks et al., 2000). Supernatant proteins without a characterized signal peptide were scanned using the SecretomeP 2.0 server (Bendtsen et al., 2005), with a SecP score  $> 0.5$  considered as evidence of non-classical protein

secretion. The subcellular localization (SCL) of the supernatant proteins was assessed using the open source web-based predictor tool, PSORTb version 3.0 (Yu et al., 2010). Proteins of interest were scanned for protein families against the Integrated resource of protein families, domains and functional sites (InterPro) database (Mulder et al., 2007).

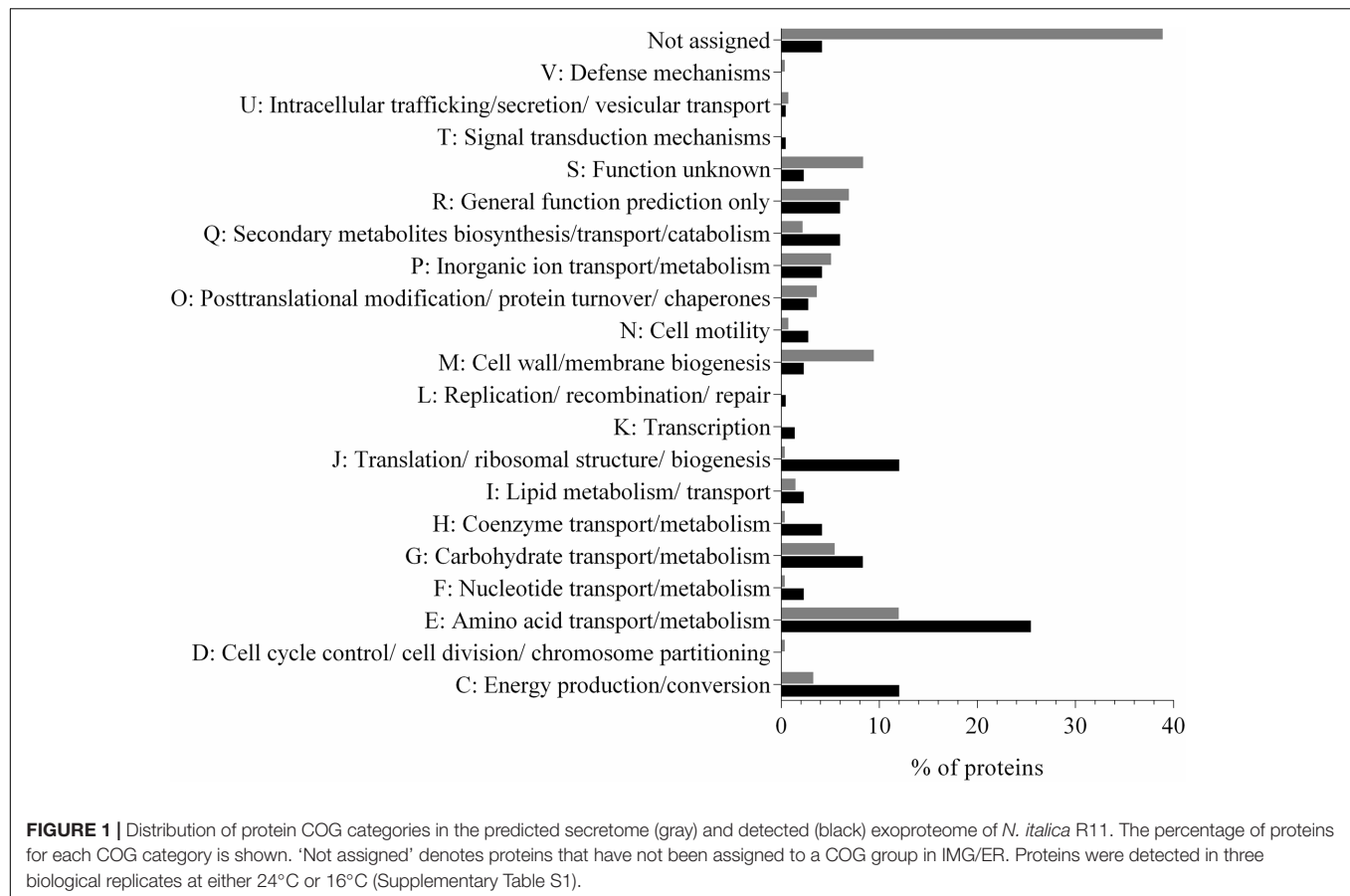
## RESULTS AND DISCUSSION

Evidence suggests that environmental perturbations, including increased seawater temperatures, contribute to the susceptibility of macroalgae to microbial disease (Gachon et al., 2010; Campbell et al., 2011; Koch et al., 2013). Therefore, an understanding of the impact of increased temperature on the virulence of bacterial seaweed pathogens is important, particularly within temperate ecosystems that are vulnerable to warming ocean currents (Wernberg et al., 2011; Egan et al., 2014). Here, we examined the influence of temperature on the exoproteome of the seaweed pathogen *N. italica* R11 using label-free LC-MS/MS.

### The Exoproteome of *N. italica* R11

In this study, we identified 207 proteins in the supernatant fraction of *N. italica* R11 (Supplementary Table S1), which corresponds to over 5% of the protein coding genes in the genome. The most abundant protein detected in the supernatant fraction (at  $>27\%$ ) was a flagellin domain protein (EEB70216); all other proteins had an individual abundance of  $<5\%$  (Supplementary Table S1). Twenty-two percent of the predicted exoproteome [annotated based on the presence of a TAT or Sec signal using IMG/ER (Markowitz et al., 2006)] was identified in the supernatant of *N. italica* R11. While many (16%) of the supernatant proteins with a signal peptide were predicted by PSORTb *in silico* analysis to be located in the periplasm (Yu et al., 2010), 16 proteins had a unknown cellular location and 12 proteins were identified as extracellular factors in agreement with the mass spectrometry data. Of the proteins that were predicted to be secreted but were not identified in the mass spectrometry analysis, many (116) proteins are uncharacterized factors [assigned to either COG S or no COG category (NA), **Figure 1**], and it is possible that these proteins may be expressed by *N. italica* R11 under growth conditions not tested here, for example, when directly associated with a *D. pulchra* host.

In contrast, a larger proportion of proteins associated with amino acid metabolism (COG E), protein translation (COG J) and energy production (COG C) were detected in the exoproteome of *N. italica* R11 than had been predicted from genome data (**Figure 1**). PSORTb analysis of the supernatant proteins suggested that at least 62% are localized in the cytoplasm, including, for example, ribosomal proteins (e.g., GenBank: EEB71264, EEB70057) and enzymes involved in the tricarboxylic acid (TCA) cycle (e.g., GenBank: EEB69552, EEB69852) (Supplementary Table S1). These typically cytosolic proteins are included in the list of detected supernatant proteins,



however, they are not considered to be part of the 'true exoproteome' of this bacterium. Instead these proteins were likely introduced into the supernatant by cell lysis during sample processing as has been previously reported for bacterial exoproteome data (Wang et al., 2013; Gotz et al., 2015). Future work may utilize subcellular fractionation to fully resolve the cellular location of putative virulence factors in *N. italica* R11 under the disease-inducing temperature.

Other studies have demonstrated that bacterial exoproteomes commonly contain a high proportion of proteins that do not possess a Sec or TAT signal sequence (Christie-Oleza and Armengaud, 2010; Kaakoush et al., 2010; Zijngje et al., 2012). For example in the Roseobacter clade bacterium *R. pomeroyi* DSS-3, 65% of exoproteome proteins did not have a signal peptide (Christie-Oleza and Armengaud, 2010). In well characterized pathogenic bacteria proteins such as antioxidant and metabolic enzymes have been observed to be secreted into the extracellular milieu via non-classical pathways (Bendtsen et al., 2005; Emanuelsson et al., 2007; Wang et al., 2016). Over ten percent of the proteins detected in the *N. italica* R11 exoproteome were predicted by SecretomeP analysis to be secreted via a non-classical pathway. These include two enzymes, pyruvate dehydrogenase (EEB70022) and glyceraldehyde-3-phosphate dehydrogenase (EEB70882), which are best known for their role as cytoplasmic glycolytic enzymes

and were localized to the cytoplasm by PSORT analysis (Yu et al., 2010). Interestingly both of these proteins have been reported in Gram-positive bacteria to be non-classically secreted (Wang et al., 2016) and shown to act as cell surface adhesions and for iron sequestration in a number of bacterial pathogens (Modun et al., 2000; Egea et al., 2007; Chauhan et al., 2015). Experimental investigation of the subcellular localization and membrane transport of proteins would provide insight into the role of non-classically secreted proteins in the physiology of *N. italica* R11, as has been modeled for the Roseobacter species *Phaeobacter inhibens* DSM 17395 (Kossmehl et al., 2013).

Numerous proteins associated with the transport of amino acids (e.g., GenBank: EEB71458), carbohydrates (e.g., GenBank: EEB69824), phosphates (e.g., GenBank: EEB70646), secondary metabolites (e.g., GenBank: EEB71180) and yet uncharacterized factors (e.g., GenBank: EEB69742) were detected in the *N. italica* R11 exoproteome (Figure 1 and Supplementary Table S1). This finding is in line with previous suggestions that this bacterium is proficient in assimilating metabolites from the environment (Fernandes et al., 2011). Fifteen percent of the proteins detected in the exoproteome data for *N. italica* R11 are homologous to bacterial Type I and Type II transporter proteins, indicating that ATP-dependent transport plays a key role in the physiology and metabolism of this pathogenic



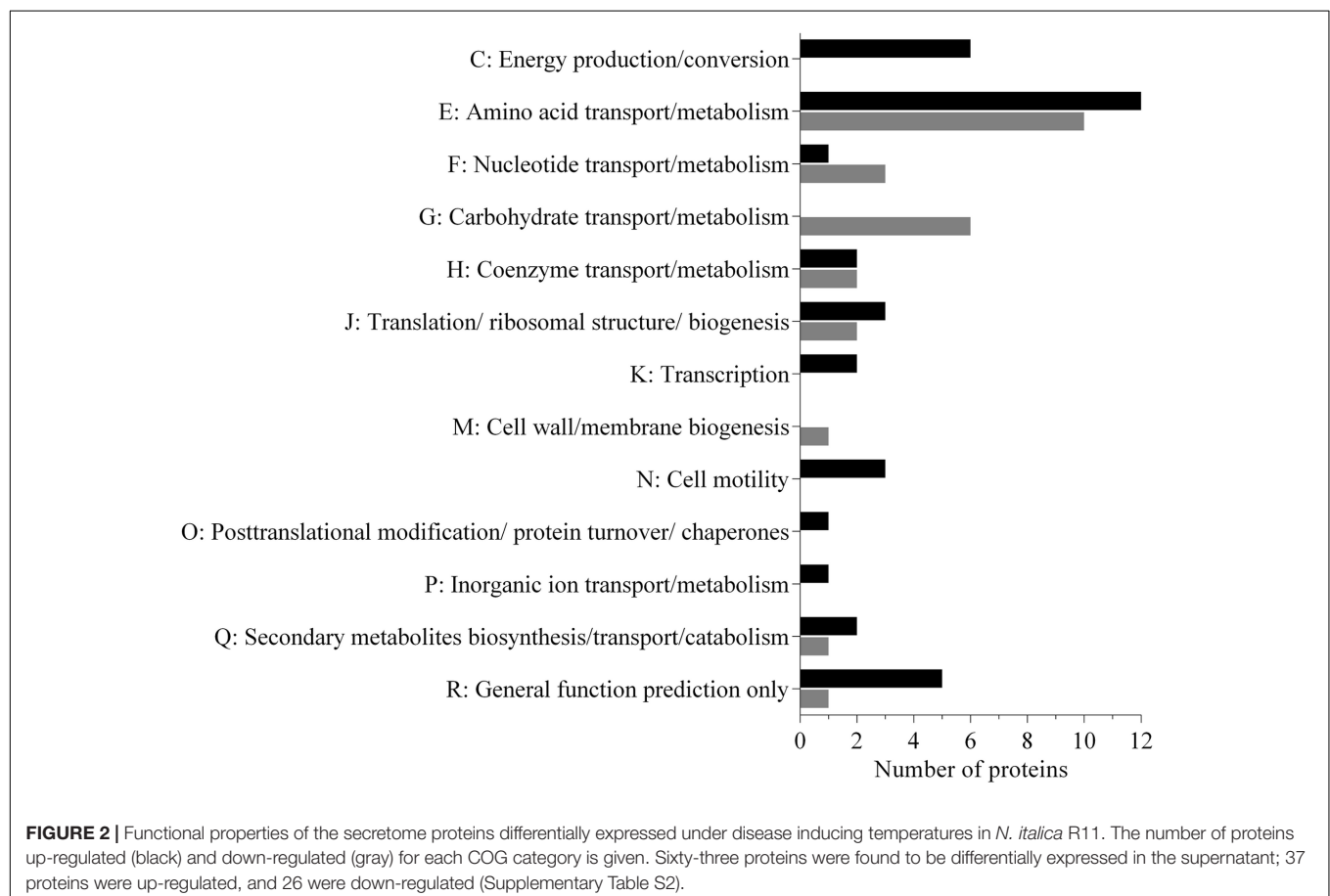
bacterium. Overrepresentation of Type I and II secretion systems has previously been reported for *Roseobacter* clade members, including *Ruegeria pomeroyi* DSS-3 and *Phaeobacter* DSM 17395 (Christie-Oleza and Armengaud, 2010; Christie-Oleza et al., 2012b; Durighello et al., 2014), where the abundance of specialized transport systems is similarly hypothesized to facilitate the assimilation of a diverse range of low concentration substrates, including metabolites, from host organisms (Christie-Oleza and Armengaud, 2010; Christie-Oleza et al., 2012a; Durighello et al., 2014).

## Putative Temperature-Dependent Virulence Factors Secreted by *N. italica* R11

Analysis of the data revealed that a subset (30%) of the *N. italica* R11 exoproteome was differentially expressed in cells at 24°C relative to those grown at 16°C (Figure 2 and Supplementary Table S2). The temperature-dependent secreted proteins are associated with transport, biogenesis and virulence related functions (Figure 2: COG E, G, J, Q), and numerous proteins (15%) are annotated as factors that mediate the binding and transport of a range of substrates including carbohydrates (e.g., GenBank: EEB71800) and amino acids (e.g., GenBank:

EEB71544) (Supplementary Table S2). Many of these transport factors were down-regulated at 24°C, including a leucine-binding (InterPro: IPR028081) extracellular receptor protein (GenBank: EEB71945) that was three-fold down-regulated. These data suggest that *N. italica* R11 modulates the uptake and/or transport of a range of metabolites and nutrients from an algal host or the surrounding environment in response to temperature conditions.

The exoproteome of *N. italica* R11 up-regulated at 24°C also includes factors potentially involved in motility, specifically, two flagella hook proteins (GenBank: EEB71275, EEB70112) (Figure 2, COG N), one protein containing the flagellar hook-length control domain (GenBank: EEB71644) and one putative flagella-associated protein (GenBank: EEB72697) (Supplementary Table S2). Motility has previously been suggested as a putative virulence determinant for *N. italica* R11 (Fernandes et al., 2011; Egan et al., 2014), and flagella hook proteins orthologous to those over-represented in the *N. italica* R11 exoproteome have a key role in the virulence of other bacterial pathogens (Mariappan et al., 2011; Haiko and Westerlund-Wikström, 2013). Increased abundance of flagella hook proteins in the exoproteome at 24°C could suggest that temperature positively effects motility in this bacterium, and future work should further investigate the link

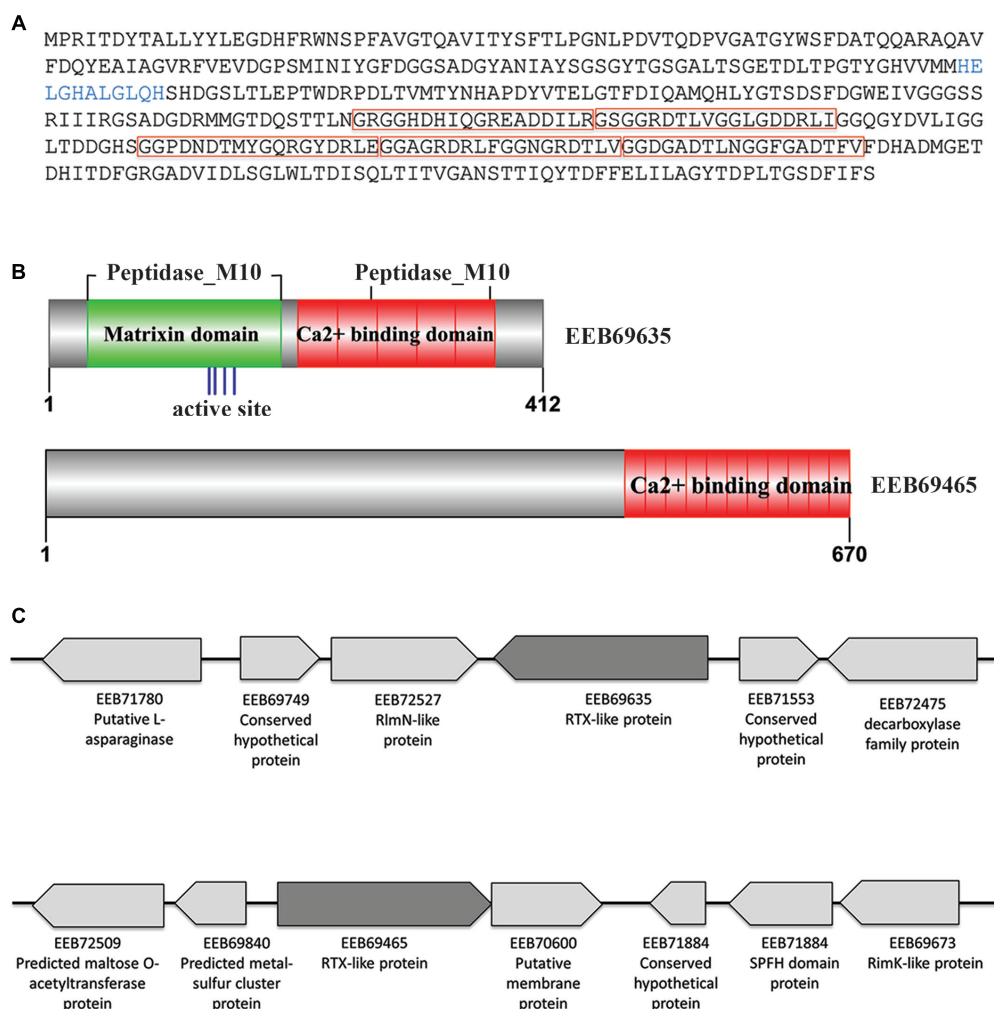


between motility, temperature, and virulence gene expression in *N. italica* R11.

Several of the proteins up-regulated in the supernatant of *N. italica* R11 under disease-inducing temperatures have a general predicted function only (GenBank: EEB72449, EEB72375, EEB72304, EEB72352) (Supplementary Table S2) (Figure 2: COG R) and thus require further investigation to elucidate their precise function in *N. italica* R11. For example, a putative zinc-dependent M16 peptidase domain (pfam00675) protein (GenBank: EEB72304) (Supplementary Table S2) that was up-regulated two-fold at 24°C has homology to an uncharacterized family of (insulinase-like) metalloproteases that are hypothesized to be involved in the degradation of small polypeptides (Fricke et al., 1995).

## *N. italica* R11 Secretes Two RTX-Like Toxins under Disease-Inducing Temperatures

The data generated in this study shows that *N. italica* secretes at least eight proteins that possess a hemolysin-type calcium-binding domain (COG Q: COG2931) characteristic of RTX proteins (GenBank: EEB69413, EEB69465, EEB69635, EEB69729, EEB71736, EEB70003, EEB70215, EEB71599) (Supplementary Table S2) (Linhartova et al., 2010). Together these eight proteins comprise over 5% of the detected exoproteome of *N. italica* R11 at 24°C (Supplementary Table S1). RTX proteins encompass a range of proteins, including leukotoxins, adhesins, and proteases (Linhartova et al., 2010) and have been well studied in other bacterial pathogens as virulence factors that



**FIGURE 3 |** Characteristics of the *N. italica* R11 RTX-like proteins secreted under disease inducing temperatures. **(A)** Translated amino acid sequence of EEB69635. The HExxH zinc-binding site/active site is shown in blue. The predicted protein sequence contains five calcium-binding repeat domains (pfam00353 HemolysinCabin domain) as indicated by red boxes. **(B)** Conserved domain structure of the RTX-like proteins EEB69635 and EEB69465, respectively. The zinc binding site/active site of EEB69635 is indicated by blue lines, with the HemolysinCabin domains in red and the matrixin domain (pfam00413) in green. **(C)** Genomic context of RTX-like toxins (EEB69635 and EEB69465 in dark gray) showing neighboring genes (not to scale); GenBank accession numbers and protein prediction are indicated.

mediate colonization, invasion and host damage (Lee et al., 2008; Li et al., 2008; Vigil et al., 2012; Wiles and Mulvey, 2013). While the *N. italica* R11 RTX-like proteins possess the RTX toxins related domain (COG2931), the proteins exhibit no significant sequence homology (<25%) to RTX proteins characterized in pathogenic bacteria (Linhartova et al., 2010; Wiles and Mulvey, 2013). Further, unlike well-characterized examples of RTX proteins that possess a Sec signal peptide (Linhartova et al., 2010), the *N. italica* R11 RTX-like proteins are predicted to be secreted via a non-classical system (SecP scores provided in Supplementary Table S1) (Bendtsen et al., 2005). Secretion of atypical RTX-like proteins has been demonstrated for Roseobacter clade members related to *N. italica* R11 (Christie-Oleza et al., 2012b; Durighello et al., 2014). This method of secretion has been proposed for the highly abundant RTX-like proteins in *Phaeobacter* strain DSM 17395 and *Ruegeria pomeroyi* DSS-3 with each protein comprising over half of the total exoproteome in these bacteria (Christie-Oleza and Armengaud, 2010; Durighello et al., 2014). Despite the significance of the atypical RTX-like proteins in the exoproteomes of Roseobacter species, the precise biological roles of these proteins have not yet been elucidated.

Two of the *N. italica* R11 RTX-like proteins were significantly up-regulated in the exoproteome under disease inducing temperatures (GenBank: EEB69465, EEB69635) (Figure 3 and Supplementary Table S2). To the best of the authors' knowledge, this constitutes the first evidence for temperature-regulation of RTX-like proteins in a Roseobacter clade member. The *N. italica* R11 RTX-like protein with the largest fold increase in the exoproteome [i.e., 5.5-fold up-regulated at the disease-inducing temperature (Supplementary Table S2)] was EEB69635. Phylogenetic analyses with homologous sequences found EEB69635 clustered with uncharacterized RTX-like proteins derived from other Roseobacters and was distantly related to characterized RTX-proteins, such as the serralsysin protein zapA from *Proteus mirabilis* (Q11137) (Supplementary Figure S3A). Analysis of the domain structure of EEB69635 revealed that in addition to containing the Ca<sup>2+</sup> binding protein RTX toxin-related domain (COG2931) typical for RTX-toxins, EEB69635 possesses an N-terminal peptidase\_M10 domain (Pfam domain: pfam00413) with a metal-binding HEXXH motif, characteristic of metalloproteases related to eukaryotic matrixins that degrade components of the extracellular matrix (Visse and Nagase, 2003). EEB69635 also has a peptidase\_M10\_C serralsysin-like C-terminal domain (Pfam domain: pfam08548). This C-terminal pfam08548 domain forms a "corkscrew" structure that is predicted to be important for protein secretion. While the M10 peptidase domain is yet to be fully characterized for bacterial proteins, it is interesting to hypothesize that the RTX-like protein EEB69635 may function in *N. italica* R11 to aid in the degradation of algal host tissue. Analysis of neighborhood genes (Figure 3) indicates EEB69635 is not located within a larger operon and annotation of surrounding genes does not provide any further functional insights of this protein.

The second differentially expressed RTX-like protein (EEB69465) was over three-fold upregulated (Supplementary

Table S2) and while this protein contains a Ca<sup>2+</sup> binding protein RTX toxin-related domain (COG2931) at the C-terminus, there is no evidence of sequence homology at the N-terminus to other characterized domains structures (Figure 3). Phylogenetic analyses with homologous sequences showed that EEB69465 clustered closely with similar uncharacterized RTX-like proteins from related Roseobacter species (Supplementary Figure S3B), thus did not provide further insight into its function. Directly downstream of the gene encoding for EEB69465 is a putative membrane protein and these two genes are likely to be encoded within a single operon. However, as with EEB69635, the gene neighborhood of EEB69465 provides little information related to the function of this protein. It is therefore difficult to speculate a specific role for EEB69465 in the pathogenesis of *N. italica* R11 beyond a predicted calcium binding function. Future work should elucidate the molecular function of these temperature-regulated RTX-like proteins in the pathogenesis and/or physiology of *N. italica* R11.

## CONCLUSION

This study is the first to demonstrate that *N. italica* R11 modulates the expression of a subset of its exoproteome in response to temperature, and it provides the foundation for future investigations into the function of the temperature-dependent secreted proteins in the pathogenicity and/or environmental persistence of *N. italica* R11. Further studies using RNA-seq techniques maybe undertaken in the future to assess virulence expression in this pathogen, for example, by characterizing transcription when the bacterium is exposed to host metabolites or exudates. In this study, the proteins that were most highly secreted from *N. italica* R11 under disease-inducing temperature, including the RTX-like proteins, constitute novel virulence factors that may play an important role in the colonization and bleaching of *D. pulchra* cells. While future studies are required to verify the expression levels and subcellular location of the proteins identified here, this foundational work highlights the potential importance of temperature for the expression of virulence factors in the macroalgal pathogen, *N. italica* R11. A relevant finding, given that increasing ocean temperatures and climate change are predicted to cause greater host stress and more extensive disease events in macroalgae.

## AUTHOR CONTRIBUTIONS

MG participated in the study design, contributed to the data analysis and wrote the manuscript. AB completed the protein experiments and data analysis, and drafted parts of the manuscript. CM-M participated in the data analysis and drafted parts of the manuscript. LZ participated in the study design and performed the protein analysis. SE conceived the study, and participated in the study design and drafting of the manuscript. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

This work was supported by funding from the Australian Research Council (ARC grant: DP1096464) and the Centre for Marine Bio-Innovation UNSW Australia.

## REFERENCES

- Bendtsen, J. D., Kiemer, L., Fausboll, A., and Brunak, S. (2005). Non-classical protein secretion in bacteria. *BMC Microbiol.* 5:58. doi: 10.1186/1471-2180-5-58
- Bendtsen, J. D., Nielsen, H., von Heijne, G., and Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* 340, 783–795. doi: 10.1016/j.jmb.2004.05.028
- Berks, B. C., Sargent, F., and Palmer, T. (2000). The Tat protein export pathway. *Mol. Microbiol.* 35, 260–274. doi: 10.1046/j.1365-2958.2000.01719.x
- Burgos-Portugal, J. A., Kaakoush, N. O., Raftery, M. J., and Mitchell, H. M. (2012). Pathogenic potential of *Campylobacter ureolyticus*. *Infect. Immun.* 80, 883–890. doi: 10.1128/iai.06031-11
- Campbell, A. H., Harder, T., Nielsen, S., Kjelleberg, S., and Steinberg, P. D. (2011). Climate change and disease: bleaching of a chemically defended seaweed. *Global Change Biol.* 17, 2958–2970. doi: 10.1111/j.1365-2486.2011.02456.x
- Campbell, L. T., Simonin, A. R., Chen, C., Ferdous, J., Padula, M. P., Harry, E., et al. (2015). *Cryptococcus* strains with different pathogenic potentials have diverse protein secretomes. *Eukaryot. Cell* 14, 554–563. doi: 10.1128/ec.00052-15
- Case, R. J., Longford, S. R., Campbell, A. H., Low, A., Tujula, N., Steinberg, P. D., et al. (2011). Temperature induced bacterial virulence and bleaching disease in a chemically defended marine macroalga. *Environ. Microbiol.* 13, 529–537. doi: 10.1111/j.1462-2920.2010.02356.x
- Chauhan, A. S., Rawat, P., Malhotra, H., Sheokand, N., Kumar, M., Patidar, A., et al. (2015). Secreted multifunctional Glyceraldehyde-3-phosphate dehydrogenase sequesters lactoferrin and iron into cells via a non-canonical pathway. *Sci. Rep.* 5:18465. doi: 10.1038/srep18465
- Christie-Oleza, J. A., and Armengaud, J. (2010). In-depth analysis of exoproteomes from marine bacteria by shotgun liquid chromatography-tandem mass spectrometry: the *Ruegeria pomeroyi* DSS-3 case-study. *Mar. Drugs* 8, 2223–2239. doi: 10.3390/md8082223
- Christie-Oleza, J. A., and Armengaud, J. (2015). Proteomics of the *Roseobacter clade*, a window to the marine microbiology landscape. *Proteomics* 15, 3928–3942. doi: 10.1002/pmic.201500222
- Christie-Oleza, J. A., Fernandez, B., Nogales, B., Bosch, R., and Armengaud, J. (2012a). Proteomic insights into the lifestyle of an environmentally relevant marine bacterium. *ISME J.* 6, 124–135. doi: 10.1038/ismej.2011.86
- Christie-Oleza, J. A., Pina-Villalonga, J. M., Bosch, R., Nogales, B., and Armengaud, J. (2012b). Comparative proteogenomics of twelve *Roseobacter* exoproteomes reveals different adaptive strategies among these marine bacteria. *Mol. Cell. Proteomics* 11, M111.013110. doi: 10.1074/mcp.M111.013110
- Desvaux, M., Hebraud, M., Talon, R., and Henderson, I. R. (2009). Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends Microbiol.* 17, 139–145. doi: 10.1016/j.tim.2009.01.004
- Durighello, E., Christie-Oleza, J. A., and Armengaud, J. (2014). Assessing the exoproteome of marine bacteria, lesson from a RTX-toxin abundantly secreted by *Phaeobacter* strain DSM 17395. *PLoS ONE* 9:e89691. doi: 10.1371/journal.pone.0089691
- Egan, S., Fernandes, N. D., Kumar, V., Gardiner, M., and Thomas, T. (2014). Bacterial pathogens, virulence mechanism and host defence in marine macroalgae. *Environ. Microbiol.* 16, 925–938. doi: 10.1111/1462-2920.12288
- Egan, S., James, S., and Kjelleberg, S. (2002). Identification and characterization of a putative transcriptional regulator controlling the expression of fouling inhibitors in *Pseudoalteromonas tunicata*. *Appl. Environ. Microbiol.* 68, 372–378. doi: 10.1128/AEM.68.1.372-378.2002
- Egea, L., Aguilera, L., Gimenez, R., Sorolla, M. A., Aguilar, J., Badia, J., et al. (2007). Role of secreted glyceraldehyde-3-phosphate dehydrogenase in the infection mechanism of enterohemorrhagic and enteropathogenic *Escherichia coli*: interaction of the extracellular enzyme with human plasminogen and fibrinogen. *Int. J. Biochem. Cell Biol.* 39, 1190–1203. doi: 10.1016/j.biocel.2007.03.008
- Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* 2, 953–971. doi: 10.1038/nprot.2007.131
- Fernandes, N., Case, R. J., Longford, S. R., Seyedsayamdost, M. R., Steinberg, P. D., Kjelleberg, S., et al. (2011). Genomes and virulence factors of novel bacterial pathogens causing bleaching disease in the marine red alga *Delisea pulchra*. *PLoS ONE* 6:e27387. doi: 10.1371/journal.pone.0027387
- Fricke, B., Betz, R., and Friebe, S. (1995). A periplasmic insulin-cleaving proteinase (ICP) from *Acinetobacter calcoaceticus* sharing properties with protease III from *Escherichia coli* and IDE from eucaryotes. *J. Basic Microbiol.* 35, 21–31. doi: 10.1002/jobm.3620350107
- Gachon, C. M. M., Sime-Ngando, T., Strittmatter, M., Chambouvet, A., and Kim, G. H. (2010). Algal diseases: spotlight on a black box. *Trends Plant Sci.* 15, 633–640. doi: 10.1016/j.tplants.2010.08.005
- Gardiner, M., Fernandes, N. D., Nowakowski, D., Raftery, M., Kjelleberg, S., Zhong, L., et al. (2015a). VarR controls colonization and virulence in the marine macroalgal pathogen *Nautella italica* R11. *Front. Microbiol.* 6:1130. doi: 10.3389/fmicb.2015.01130
- Gardiner, M., Thomas, T., and Egan, S. (2015b). A Glutathione peroxidase (GpoA) plays a role in the pathogenicity of *Nautella italica* strain R11 towards the red alga *Delisea pulchra*. *FEMS Microbiol. Ecol.* 91, fiv021. doi: 10.1093/femsec/fiv021
- Gotz, F., Yu, W., Dube, L., Prax, M., and Ebner, P. (2015). Excretion of cytosolic proteins (ECP) in bacteria. *Int. J. Med. Microbiol.* 305, 230–237. doi: 10.1016/j.ijmm.2014.12.021
- Haiko, J., and Westerlund-Wikström, B. (2013). The role of the bacterial flagellum in adhesion and virulence. *Biology* 2, 1242–1267. doi: 10.3390/biology2041242
- He, Y., Wang, H., and Chen, L. (2015). Comparative secretomics reveals novel virulence-associated factors of *Vibrio parahaemolyticus*. *Front. Microbiol.* 6:707. doi: 10.3389/fmicb.2015.00707
- Hollants, J., Leliaert, F., De Clerck, O., and Willems, A. (2013). What we can learn from sushi: a review on seaweed–bacterial associations. *FEMS Microbiol. Ecol.* 83, 1–16. doi: 10.1111/j.1574-6941.2012.01446.x
- Kaakoush, N. O., Deshpande, N. P., Man, S. M., Burgos-Portugal, J. A., Khattak, F. A., Raftery, M. J., et al. (2015). Transcriptomic and proteomic analyses reveal key innate immune signatures in the host response to the gastrointestinal pathogen *Campylobacter concisus*. *Infect. Immun.* 83, 832–845. doi: 10.1128/iai.03012-14
- Kaakoush, N. O., Man, S. M., Lamb, S., Raftery, M. J., Wilkins, M. R., Kovach, Z., et al. (2010). The secretome of *Campylobacter concisus*. *FEBS J.* 277, 1606–1617. doi: 10.1111/j.1742-4658.2010.07587.x
- Kimes, N. E., Grim, C. J., Johnson, W. R., Hasan, N. A., Tall, B. D., Kothary, M. H., et al. (2012). Temperature regulation of virulence factors in the pathogen *Vibrio coralliilyticus*. *ISME J.* 6, 835–846. doi: 10.1038/ismej.2011.154
- Koch, M., Bowes, G., Ross, C., and Zhang, X. H. (2013). Climate change and ocean acidification effects on seagrasses and marine macroalgae. *Global Change Biol.* 19, 103–132. doi: 10.1111/j.1365-2486.2012.02791.x
- Kossmehl, S., Wohlbrand, L., Druppel, K., Feenders, C., Blasius, B., and Rabus, R. (2013). Subcellular protein localization (cell envelope) in *Phaeobacter inhibens* DSM 17395. *Proteomics* 13, 2743–2760. doi: 10.1002/pmic.201300112
- Kumar, V., Zozaya-Valdes, E., Kjelleberg, S., Thomas, T., and Egan, S. (2016). Multiple opportunistic pathogens can cause a bleaching disease in the red seaweed *Delisea pulchra*. *Environ. Microbiol.* 18, 3962–3975. doi: 10.1111/1462-2920.13403
- Lee, B. C., Choi, S. H., and Kim, T. S. (2008). *Vibrio vulnificus* RTX toxin plays an important role in the apoptotic death of human intestinal epithelial cells exposed to *Vibrio vulnificus*. *Microbes Infect.* 10, 1504–1513. doi: 10.1016/j.micinf.2008.09.006

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01203/full#supplementary-material>



- Li, L., Rock, J. L., and Nelson, D. R. (2008). Identification and characterization of a Repeat-in-Toxin gene cluster in *Vibrio anguillarum*. *Infect. Immun.* 76, 2620–2632. doi: 10.1128/iai.01308-07
- Linhartova, I., Bumba, L., Masin, J., Basler, M., Osicka, R., Kamanova, J., et al. (2010). RTX proteins: a highly diverse family secreted by a common mechanism. *FEMS Microbiol. Rev.* 34, 1076–1112. doi: 10.1111/j.1574-6976.2010.00231.x
- Madec, S., Pichereau, V., Jacq, A., Paillard, M., Boisset, C., Guerard, F., et al. (2014). Characterization of the secretomes of two vibrios pathogenic to mollusks. *PLoS ONE* 9:e113097. doi: 10.1371/journal.pone.0113097
- Mandelc, S., and Javornik, B. (2015). The secretome of vascular wilt pathogen *Verticillium albo-atrum* in simulated xylem fluid. *Proteomics* 15, 787–797. doi: 10.1002/pmic.201400181
- Mariappan, V., Vellasamy, K. M., Hashim, O. H., and Vadivelu, J. (2011). Profiling of *Burkholderia cepacia* secretome at mid-logarithmic and early-stationary phases of growth. *PLoS ONE* 6:e26518. doi: 10.1371/journal.pone.0026518
- Markowitz, V. M., Korzeniewski, F., Palaniappan, K., Szeto, E., Werner, G., Padki, A., et al. (2006). The integrated microbial genomes (IMG) system. *Nucleic Acids Res.* 34, D344–D348. doi: 10.1093/nar/gkj024
- Modun, B., Morrissey, J., and Williams, P. (2000). The staphylococcal transferrin receptor: a glycolytic enzyme with novel functions. *Trends Microbiol.* 8, 231–237. doi: 10.1016/S0966-842X(00)01728-5
- Mulder, N. J., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., et al. (2007). New developments in the InterPro database. *Nucleic Acids Res.* 35, D224–D228. doi: 10.1093/nar/gkl841
- Vanden Bergh, P., Heller, M., Braga-Lagache, S., and Frey, J. (2013). The *Aeromonas salmonicida* subsp. *salmonicida* exoproteome: global analysis, moonlighting proteins and putative antigens for vaccination against furunculosis. *Proteome Sci.* 11:44. doi: 10.1186/1477-5956-11-44
- Vigil, P. D., Wiles, T. J., Engstrom, M. D., Prasov, L., Mulvey, M. A., and Mobley, H. L. (2012). The repeat-in-toxin family member TosA mediates adherence of uropathogenic *Escherichia coli* and survival during bacteremia. *Infect. Immun.* 80, 493–505. doi: 10.1128/iai.05713-11
- Visse, R., and Nagase, H. (2003). Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ. Res.* 92, 827–839. doi: 10.1161/01.res.0000070112.80711.3d
- Wang, G., Chen, H., Xia, Y., Cui, J., Gu, Z., Song, Y., et al. (2013). How are the non-classically secreted bacterial proteins released into the extracellular milieu? *Curr. Microbiol.* 67, 688–695. doi: 10.1007/s00284-013-0422-6
- Wang, G., Xia, Y., Song, X., and Ai, L. (2016). Common non-classically secreted bacterial proteins with experimental evidence. *Curr. Microbiol.* 72, 102–111. doi: 10.1007/s00284-015-0915-6
- Wernberg, T., Russell, B. D., Moore, P. J., Ling, S. D., Smale, D. A., Campbell, A., et al. (2011). Impacts of climate change in a global hotspot for temperate marine biodiversity and ocean warming. *J. Exp. Mar. Biol. Ecol.* 400, 7–16. doi: 10.1016/j.jembe.2011.02.021
- Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V., et al. (2008). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 36, D13–D21. doi: 10.1093/nar/gkm1000
- Wiles, T. J., and Mulvey, M. A. (2013). The RTX pore-forming toxin alpha-hemolysin of uropathogenic *Escherichia coli*: progress and perspectives. *Future Microbiol.* 8, 73–84. doi: 10.2217/fmb.12.131
- Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R., et al. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26, 1608–1615. doi: 10.1093/bioinformatics/btq249
- Zech, H., Thole, S., Schreiber, K., Kalhofer, D., Voget, S., Brinkhoff, T., et al. (2009). Growth phase-dependent global protein and metabolite profiles of *Phaeobacter gallaeciensis* strain DSM 17395, a member of the marine Roseobacter-clade. *Proteomics* 9, 3677–3697. doi: 10.1002/pmic.200901020
- Zijnge, V., Kieselbach, T., and Oscarsson, J. (2012). Proteomics of protein secretion by *Aggregatibacter actinomycetemcomitans*. *PLoS ONE* 7:e41662. doi: 10.1371/journal.pone.0041662

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Gardiner, Bournazos, Maturana-Martinez, Zhong and Egan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Evolution of Dimethylsulfoniopropionate Metabolism in Marine Phytoplankton and Bacteria

Hannah A. Bullock<sup>1\*</sup>, Haiwei Luo<sup>2</sup> and William B. Whitman<sup>1</sup>

<sup>1</sup> Department of Microbiology, University of Georgia, Athens, GA, USA, <sup>2</sup> School of Life Sciences, The Chinese University of Hong Kong, Hong Kong, Hong Kong

## OPEN ACCESS

### Edited by:

Bernd Wemheuer,  
University of New South Wales,  
Australia

### Reviewed by:

Andrew WB Johnston,  
University of East Anglia, UK  
Yin Chen,  
University of Warwick, UK

### \*Correspondence:

Hannah A. Bullock  
bullockh@uga.edu

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 14 January 2017

**Accepted:** 28 March 2017

**Published:** 19 April 2017

### Citation:

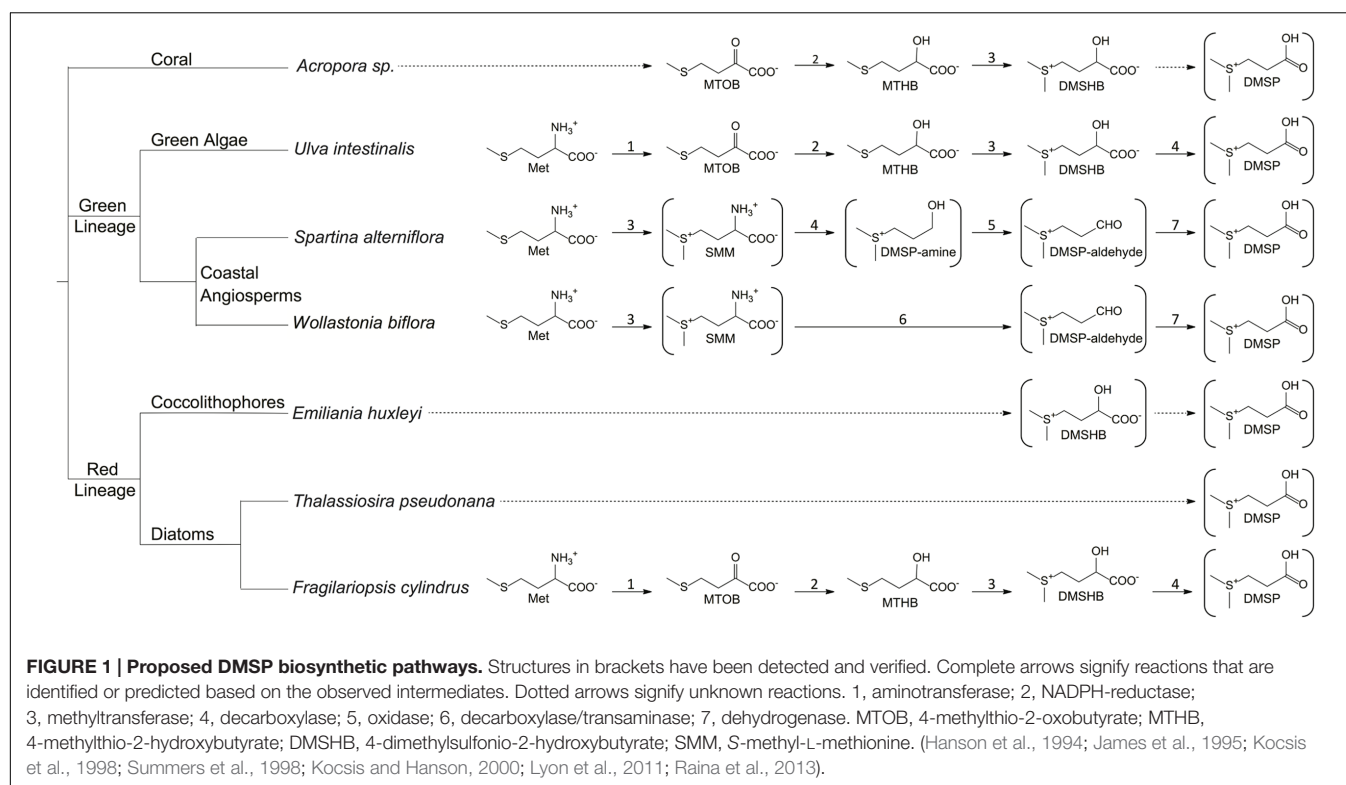
Bullock HA, Luo H and Whitman WB  
(2017) Evolution  
of Dimethylsulfoniopropionate  
Metabolism in Marine Phytoplankton  
and Bacteria. *Front. Microbiol.* 8:637.  
doi: 10.3389/fmicb.2017.00637

The elucidation of the pathways for dimethylsulfoniopropionate (DMSP) synthesis and metabolism and the ecological impact of DMSP have been studied for nearly 70 years. Much of this interest stems from the fact that DMSP metabolism produces the climatically active gas dimethyl sulfide (DMS), the primary natural source of sulfur to the atmosphere. DMSP plays many important roles for marine life, including use as an osmolyte, antioxidant, predator deterrent, and cryoprotectant for phytoplankton and as a reduced carbon and sulfur source for marine bacteria. DMSP is hypothesized to have become abundant in oceans approximately 250 million years ago with the diversification of the strong DMSP producers, the dinoflagellates. This event coincides with the first genome expansion of the Roseobacter clade, known DMSP degraders. Structural and mechanistic studies of the enzymes of the bacterial DMSP demethylation and cleavage pathways suggest that exposure to DMSP led to the recruitment of enzymes from preexisting metabolic pathways. In some cases, such as DmdA, DmdD, and DddP, these enzymes appear to have evolved to become more specific for DMSP metabolism. By contrast, many of the other enzymes, DmdB, DmdC, and the acrylate utilization hydratase AcuH, have maintained broad functionality and substrate specificities, allowing them to carry out a range of reactions within the cell. This review will cover the experimental evidence supporting the hypothesis that, as DMSP became more readily available in the marine environment, marine bacteria adapted enzymes already encoded in their genomes to utilize this new compound.

**Keywords:** DMSP, dimethylsulfoniopropionate, evolution, phytoplankton, Roseobacter

## INTRODUCTION

Dimethylsulfoniopropionate (DMSP) was first identified in 1948 and has since been found to be not only abundant in marine surface waters but also a valuable resource for many marine organisms and an integral part of the global sulfur cycle (Challenger and Simpson, 1948; van Duyl et al., 1998; Stefels et al., 2007). DMSP is the precursor of the climate-active gas dimethyl sulfide (DMS), which upon release into the atmosphere aids in the formation of cloud condensation nuclei (Lovelock et al., 1972; Hatakeyama et al., 1982). Additionally, DMS is the largest natural source of sulfur to the atmosphere, comparable in magnitude to the sulfur dioxide formed during the burning of coal.

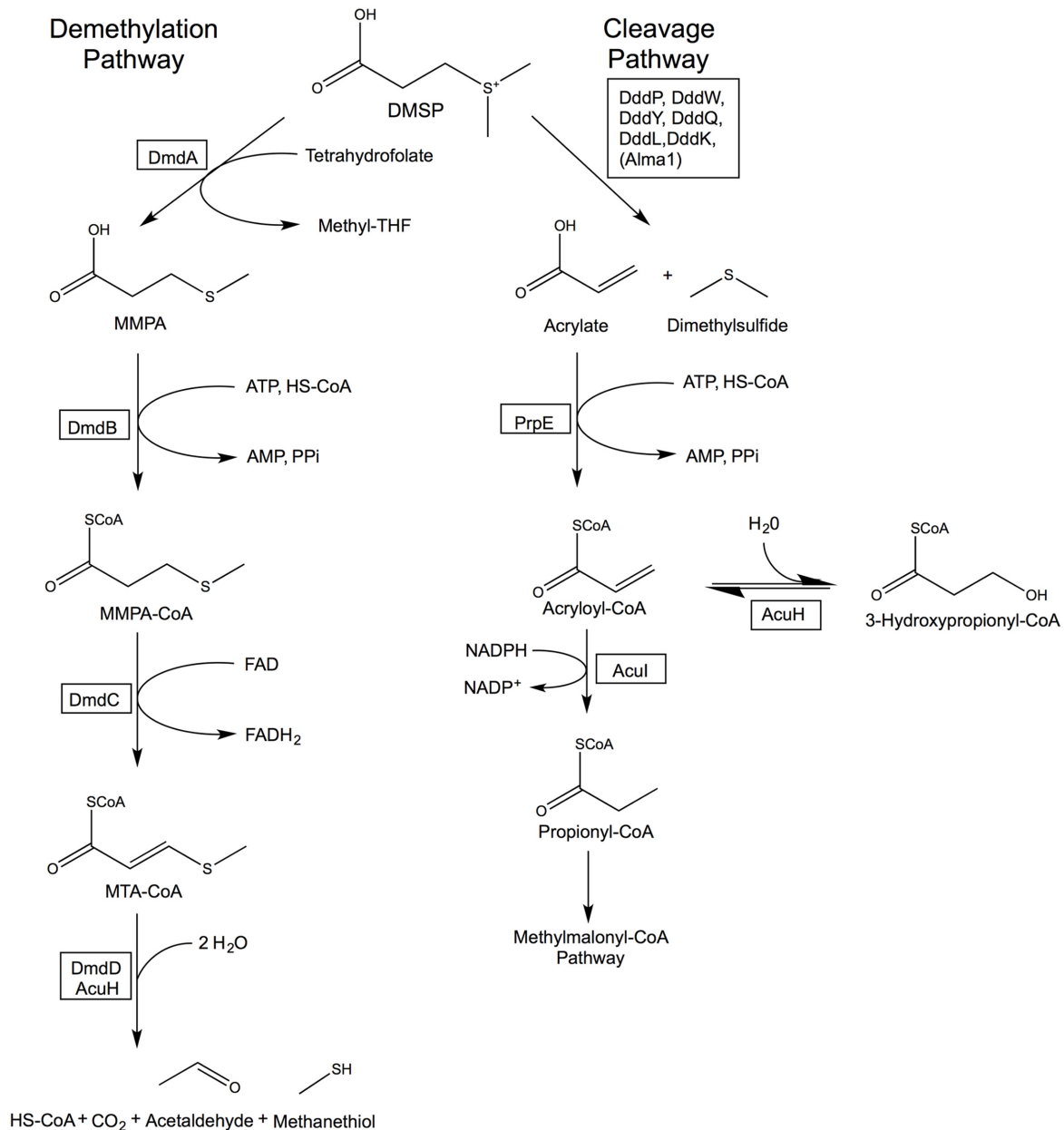


As DMS oxidation products display a longer residence time in the atmosphere than anthropogenic sulfur dioxide, their contribution to the global sulfur burden is also greater (Lovelock et al., 1972; Chin and Jacob, 1996).

From an organismal viewpoint, DMSP is equally important. The ability to produce and metabolize DMSP is concentrated into specific classes of life. The main producers of DMSP are phytoplankton, mostly the classes Dinophyceae (dinoflagellates) and Prymnesiophyceae (coccolithophores) (Keller, 1989). DMSP production has also been noted in diatoms (Lyon et al., 2011; Kettles et al., 2014), the green algae *Ulva intestinalis* (Gage et al., 1997), corals (Raina et al., 2013), and certain higher plants like sugarcane (Paquet et al., 1994), and the coastal angiosperms *Spartina alterniflora* (Kocsis et al., 1998) and *Wollastonia biflora* (Hanson et al., 1994). Recently, DMSP biosynthesis was detected in several marine Alphaproteobacteria (Curson et al., 2017). The basis of the need for DMSP is not entirely understood. Several physiological functions for DMSP in phytoplankton and green algae have been demonstrated, including roles as an osmolyte, antioxidant, predator deterrent, and cryoprotectant (Kirst et al., 1990; Karsten et al., 1996; Wolfe and Steinke, 1997; Sunda et al., 2002). At present, each of the proposed pathways for DMSP biosynthesis begins with methionine, although subsequent steps vary (Figure 1). The pathways proposed for phytoplankton, algae, corals, and perhaps the DMSP-producing Alphaproteobacteria share similar reactions and intermediates which differ distinctly from those predicted in the coastal angiosperms (Hanson et al., 1994; Gage et al., 1997; Kocsis et al., 1998; Lyon et al., 2011; Curson et al., 2017). These variations indicate that the

ability to synthesize DMSP has evolved at least twice (Stefels, 2000).

Bacteria may metabolize DMSP via two pathways, the cleavage or the demethylation pathway (Figure 2). The cleavage pathway results in the formation of DMS, while the demethylation pathway produces methanethiol (MeSH). The DMSP demethylation and cleavage pathway enzymes are hypothesized to be adapted versions of enzymes that were already contained within bacterial genomes and developed in response to the availability of this substrate (Reisch et al., 2011a,b). In this review, we investigate the likely evolutionary path that led to the development of DMSP biosynthesis and subsequently the specialized DMSP catabolic pathways. The members of the Alphaproteobacteria, specifically members of the Roseobacter clade, appear to be uniquely adapted to utilize this valuable source of reduced carbon and sulfur. Bacteria within the Roseobacter and SAR11 clades possess enzymes that specifically and efficiently catalyze reactions of the demethylation pathway (Reisch et al., 2008, 2011b; Curson et al., 2011b; Tan et al., 2013; Bullock et al., 2014; Johnston et al., 2016; Sun et al., 2016). Bacteria are also responsible for the majority of DMSP catabolism via the cleavage pathway (Figure 2). There is additional evidence suggesting the use of DMSP as an osmolyte and antioxidant in marine bacteria (Kiene et al., 2000; Simo et al., 2002; Lesser, 2006; Reisch et al., 2011b; Salgado et al., 2014). Many microorganisms encode enzymes that share a great deal of similarity to the demethylation pathway enzymes (Figure 3), demonstrating their adaptability and plasticity. The many roles of DMSP may have helped to drive the adaptation of existing enzymes for DMSP metabolism.



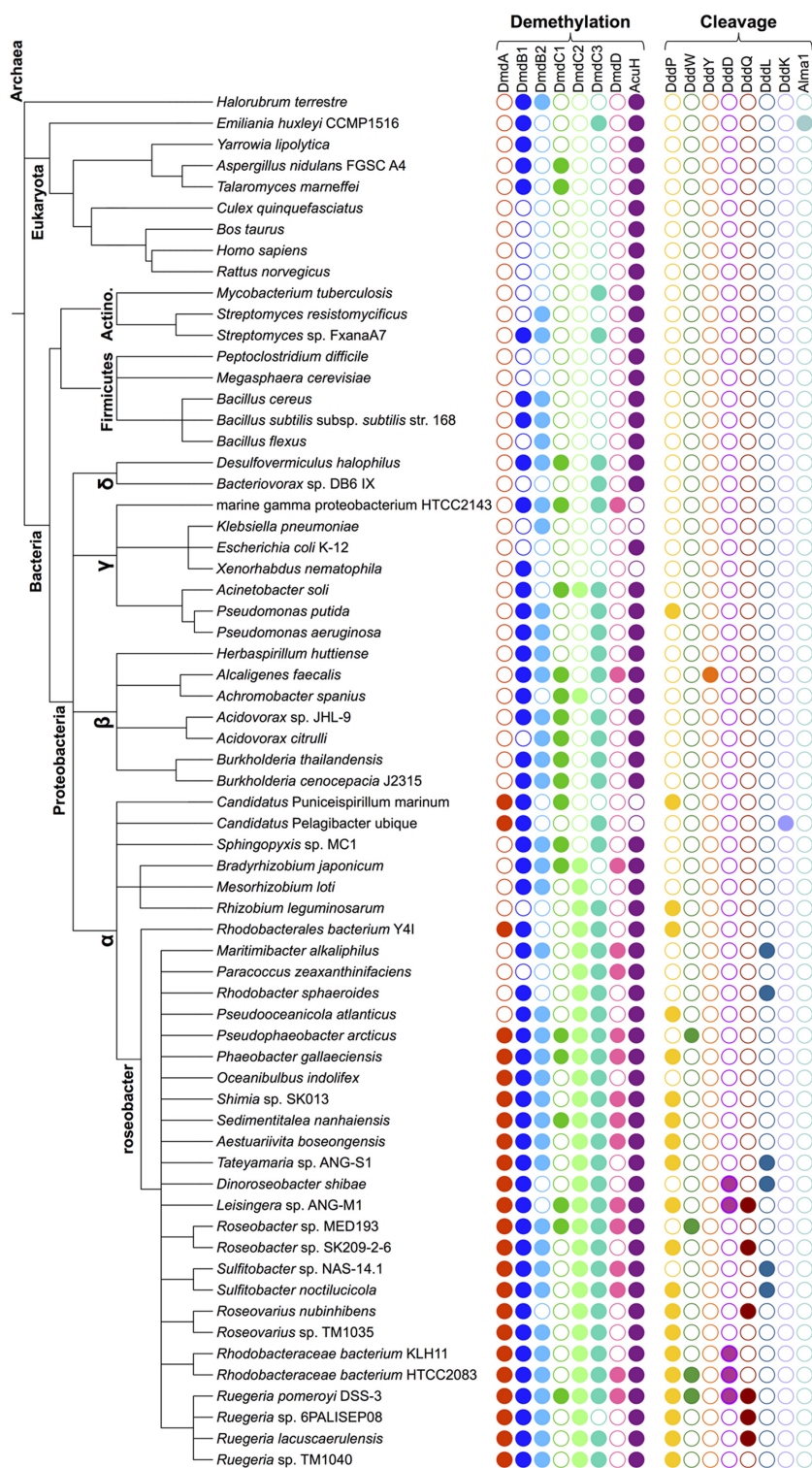
**FIGURE 2 | Bacterial DMSP demethylation and cleavage pathways.** The DMSP demethylation pathway is catalyzed by the DMSP demethylase (DmdA), MMPA-CoA ligase (DmdB), MMPA-CoA dehydrogenase (DmdC), and either the MTA-CoA hydratase (DmdD) or acrylate utilization hydratase (AcuH). The cleavage pathway is catalyzed by a DMSP lyase (DddP, DddW, DddY, DddQ, DddL, DddK, or the algal Alma1), an acrylate-CoA ligase (PrpE), and an acryloyl-CoA reductase (AcuL). AcuH catalyzes a side reaction forming 3-hydroxypropionyl-CoA in the cleavage pathway. Revised from Reisch et al. (2011b, 2013).

## EVOLUTION OF MODERN PHYTOPLANKTON

The first photosynthetic eukaryotes developed as the result of the acquisition of a cyanobacterium endosymbiont by a eukaryotic host, creating a membrane bound plastid (Bhattacharya and Medlin, 1998; Palmer, 2003; Yoon et al., 2004). Further diversification led to the formation of three clades from this original photosynthetic eukaryote, the green algae (green

plastid lineage), the red algae (red plastid lineage), and the microbial algae glaucophytes (Delwiche, 1999). These lineages are distinguished by the chlorophyll present in their plastids. All the plastids contain chlorophyll a, but the green plastids also contain chlorophyll b, and the red plastids contain phycobilin (Keeling, 2010). Members of the Charophyta branch of the green plastid lineage colonized the land approximately 430 million year ago (mya). The Chlorophyta branch evolved into the green algae species seen today, including the Euglenoids





**FIGURE 3 | Phylogenetic species tree representing the diversity of organisms that possess enzymes from the DMSP demethylation pathway and the DMSP lyases.** The relatedness of representative lineages is indicated schematically on the left. Colored filled circles represent the presence of the indicated protein-encoding gene. Protein designations and query sequences are as follows: From *R. pomeroyi*, DmdA (SPO1913); DmdB1 (SPO0677); DmdB2 (SPO2045); DmdC1 (SPO3805); DmdC2 (SPO0298); DmdC3 (SPO2915); DmdD (SPO3804); and AcuH (SL1157\_0807) from *R. lacuscaerulensis*. DMSP lyase protein designations and query sequences are *R. pomeroyi* DddP (SPO2299), DddW (SPO0453), DddQ (SPO1596), DddD (SPO1703); *A. faecalis* DddY (ADT64689.1); *R. sphaeroides* DddL (RSP1433), *P. ubique* DddK (SAR11\_0394); *E. huxleyi* Alma1 (XP\_005784450). The  $e$  value cut off used in all cases was  $< e 10^{-70}$  with the exception of DddW (cut off of  $< e 10^{-40}$ ) and DddQ (cut off of  $< e 10^{-30}$ ). See **Figure 2** for the names of the enzymes.

and Chlorarachniophytes (Sanderson, 2003; Lewis and McCourt, 2004; McCourt et al., 2004). Meanwhile, today's marine phytoplankton are largely descended from the red plastid lineage. The red plastid lineage phytoplankton, including coccolithophores, diatoms, and most dinoflagellates, first began to increase in abundance after the end-Permian extinction about 250 mya (Falkowski et al., 2004a,b).

The coccolithophores and dinoflagellates both began appearing in the fossil record about 250 mya in the Mesozoic period, while diatoms first appeared during the Early Cretaceous. All three groups saw extensive subsequent diversification in the Mesozoic period (250–65 mya) (Harwood and Nikolaev, 1995; Moldowan et al., 1996; Stover et al., 1996; Moldowan and Talyzina, 1998; Moldowan and Jacobson, 2000; Bown et al., 2004). The red lineage first began to proliferate in the benthic coastal regions, which were the first consistently oxic marine habitats. The breakup of Pangea increased sea levels and the total length of coastal area available for phytoplankton to colonize. This event also allowed nutrients that had been locked in the interior portions of continents to reach coastal waters (Vail et al., 1977; Haq et al., 1987). Changes in ocean redox chemistry from more reducing conditions that favored the green plastid lineage prior to the end-Permian extinction to the higher oxidation states of the Mesozoic ocean further contributed to the success of the red plastid lineage (Whitfield, 2001). Quigg et al. (2003) present evidence for the role of trace element availability in the proliferation of the red plastid lineage based on differences in the trace element composition between members of the red and green plastid lineages (Quigg et al., 2003). Members of the green plastid lineage have much higher requirements for iron, zinc, and copper while members of the red plastid lineage have high requirements for manganese, cobalt, and cadmium. It has been predicted that these differences in trace element requirements reflect differences in green vs. red plastid biochemistry (Quigg et al., 2003; Falkowski et al., 2004a).

The dominance of the red plastid lineage is such that all but one of the eight major taxa of eukaryotic phytoplankton in the present day oceans contains the red plastid (Falkowski et al., 2004a). The diversity of the red plastid lineage also greatly expanded as a result of secondary and tertiary endosymbiotic events, which are evident from the presence of multiple membranes surrounding some plastids of modern day phytoplankton. These events involved the engulfment of an algal cell by another eukaryote via endocytosis (Delwiche, 1999; Archibald and Keeling, 2002; Palmer, 2003; Keeling, 2010). The majority of the phytoplankton present today are the result of secondary and sometimes tertiary endosymbiotic events (Archibald, 2009). Today's phytoplankton play key roles in global nutrient cycles and particularly in the global sulfur cycle as producers of DMSP and DMS (Lovelock et al., 1972).

## PHYTOPLANKTON AND DMSP

Marine phytoplankton and algae live in an environment that is continually changing based on shifts in ocean currents. Living in this dynamic environment requires that these organisms

adapt continually to varying temperatures, light, and nutrient availability (Cortes et al., 2001; Wagner et al., 2005; Allen et al., 2006). These changes may have been even more extreme in the Paleozoic and Mesozoic oceans. Abiotic forces have been found to have a large impact on the population variability of *Emiliania huxleyi* and *Florisphaera profunda* (Cortes et al., 2001). Phytoplankton in general, however, adapt quickly and relatively readily to environmental changes due to their rapid cell division rates and large population sizes (Simo, 2001; Allen, 2005; Wagner et al., 2005; Allen et al., 2006).

One specific adaptation that may help phytoplankton deal with their ever changing environment is the ability to synthesize and utilize DMSP and DMS. DMSP makes up about 90% of the reduced sulfur found in algae, but much about the regulation of its biosynthesis and uptake is still not well understood (Gage et al., 1997). Nevertheless, many of the proposed roles for these compounds would be beneficial to phytoplankton trying to survive in an ever-changing environment. DMSP is proposed to have roles as an osmolyte (Kirst et al., 1990), an antioxidant (Sunda et al., 2002), and as a means of balancing excess cellular energy (Stefels, 2000; Allen, 2005). Additionally, polar diatoms and algae are thought to produce DMSP as a cryoprotectant (Karsten et al., 1996). This hypothesis is supported by the higher levels of DMSP in sea ice diatoms compared with those from more temperate climates (Lyon et al., 2011; Kettles et al., 2014). DMSP is also a predator/grazing deterrent owing to its cleavage to acrylate (Stefels, 2000). New studies of the coral genus *Acropora* have generated still more uses for DMSP. Reef building coral juveniles increase DMSP production when subject to thermal stress and may also use DMSP as a bacterial signaling molecule, attracting particular microbial communities that are necessary for coral health (Raina et al., 2013).

The role of DMSP and DMS as antioxidants could be particularly useful for phytoplankton as plastids are typically hyperoxic and produce reactive oxygen species (ROS) during oxygenic photosynthesis. Other stresses such as exposure to ultraviolet radiation (UVR) and thermal stress can further increase ROS production (Asada and Takahashi, 1987; Fridovich, 1998; Lesser, 2006). The production of ROS by plastids might explain why DMSP and DMS production are observed in both phytoplankton and land plants. There is also evidence to suggest that the final step of DMSP synthesis in the flowering plant *W. biflora* takes place in the plastid (chloroplast) (Trossat et al., 1996). DMSP, DMS, and acrylate are all able to quench HO• radicals, although acrylate and DMS are more efficient than DMSP. The resultant product of HO• quenching is dimethylsulfoxide (DMSO), which subsequently reacts with additional HO• radicals to form methane sulfinic acid and then methane sulfonic acid. In contrast to DMSP and acrylate, DMS is uncharged and can diffuse through biological membranes, acting as an antioxidant nearly anywhere in the cell (Sunda et al., 2002; Lesser, 2006; Husband et al., 2012).

Another impetus for the production of DMSP may be the need for an osmolyte that does not contain nitrogen. Nitrogen is often limiting in ocean surface waters, which may in turn limit the production of the nitrogen-containing osmolyte glycine betaine. Ito et al. (2011) observed that under conditions where

sulfate limited growth of the marine algae *Ulva pertusa*, the sulfur from methionine was used primarily for the synthesis of S-adenosyl methionine and methionyl-tRNA, rather than for DMSP synthesis. However, when the salinity and abundance of sulfate increased; the sulfur from methionine was increasingly used for DMSP biosynthesis, and the intracellular DMSP levels increased (Ito et al., 2011).

One additional hypothesis for the origin of DMSP biosynthesis proposes that it developed as a means of dispelling excess energy, carbon and reducing equivalents when growth becomes unbalanced due to nutrient limitation (Stefels, 2000). Rapid changes in the ocean environment can require phytoplankton to have an equally rapid response to imbalances between photosynthesis and growth (Allen, 2005; Wagner et al., 2005; Allen et al., 2006). Since photon capture cannot be quickly stopped, production of nitrogen or phosphorous poor molecules when growth is limited by these nutrients is a means of consuming extra carbon, energy, and reducing equivalents that cannot be used for protein biosynthesis or cell division (Stefels, 2000; Simo, 2001; Allen, 2005). Further, the continued production of DMSP may also serve to regenerate and redistribute nitrogen for the production of new amino acids and to stimulate continued sulfate assimilation by keeping the cellular concentration of methionine and cysteine low (Gage et al., 1997; Stefels, 2000). Thus, DMSP may have originally been produced as a means of dissipating excess energy and carbon and was then adapted for other functions.

## SYNTHESIS OF DMSP BY MARINE PHYTOPLANKTON AND ALGAE

The main producers of DMSP are phytoplankton, mostly in the classes Dinophyceae (dinoflagellates) and the Prymnesiophyceae (which includes the coccolithophores). Certain members of the Chrysophyceae and Bacillariophyceae (diatoms) can also produce DMSP (Keller, 1989). DMSP likely first became abundant in ocean environments about 250 mya in conjunction with the increasing abundance of dinoflagellates and coccolithophores. Based on a comparison of literature reports for 95 DMSP-producing species (Table 1), it was determined that dinoflagellates produced the highest amounts of DMSP, with intracellular levels ranging from 0.00011 to 14.7 pmol/cell (Keller, 1989; Wiesemeier and Pohnert, 2007; Caruana, 2010; Caruana and Malin, 2014). In particular, *Alexandrium minutum* and *Protoperdinium pellucidum* produced 14.2 and 14.7 pmol DMSP/cell, respectively. Diatoms have intracellular DMSP levels ranging from 0.0006 to 0.257 pmol/cell, while haptophytes (coccolithophores) contained from 0.00037 to 0.148 pmol DMSP/cell (Keller, 1989; Caruana and Malin, 2014). DMSP production is less common among the higher plants, although it has been observed in *Spartina* species (Kocsis et al., 1998), certain sugarcane (Paquet et al., 1994), and the flowering plant *W. biflora* (Hanson et al., 1994; James et al., 1995). DMSP production has also been observed in members of the coral genus *Acropora* in the absence of their algal endosymbiont *Symbiodinium*, also a known DMSP producer (Raina et al., 2013).

Thus, while DMSP is widely distributed in a large number of phototrophs, only a few groups produce very high amounts, and it is likely that DMSP only became widely available as a nutrient in marine environments following the evolution of these groups.

Little is understood about the biosynthetic pathways for DMSP in marine phytoplankton and corals. The first complete DMSP biosynthetic pathways were described in the green algae *U. intestinalis* (Gage et al., 1997), the marine cordgrass *S. alterniflora* (Kocsis et al., 1998), and coastal plant *W. biflora* (Hanson et al., 1994; James et al., 1995) (Figure 1). Each pathway identified thus far begins with methionine and includes a deamination reaction, supporting the hypothesis that DMSP biosynthesis is used by these organisms to regenerate nitrogen from methionine. The DMSP biosynthetic pathways of *S. alterniflora* and *W. biflora* are more similar to each other than they are to the pathway in *U. intestinalis*, suggesting that the plant pathways evolved independently from those in marine algae, corals, and phytoplankton. If true, this would indicate that there was selective pressure for the evolution of DMSP biosynthetic pathways even in very different organisms.

The DMSP biosynthetic pathways of the major producers in the marine environment are still largely unknown, but they are likely similar to the pathway described in *U. intestinalis*. The *U. intestinalis* pathway begins with methionine and utilizes an aminotransferase, a NADPH-linked reductase, a methyltransferase, and an oxidative decarboxylase to produce DMSP (Gage et al., 1997; Summers et al., 1998). The commitment step is hypothesized to be the third step, the conversion of 4-methylthio-2-hydroxybutyrate (MTHB) to 4-dimethylsulfonio-2-hydroxybutyrate (DMSHB) by a methyltransferase (Figure 1). The key intermediate DMSHB has been identified in *U. intestinalis*, *U. pertusa*, *E. huxleyi*, *Tetraselmis* sp., and *Melosira nummuloides*, indicating that this pathway is present in a range of phytoplankton (Gage et al., 1997; Stefels, 2000; Ito et al., 2011). Lyon et al. (2011) identified candidate proteins and genes for this four-step pathway in the

**TABLE 1 | Levels of DMSP production for different phytoplankton groups<sup>a</sup>.**

Taxonomic group	Number of species examined	Concentration of DMSP (median pmol/cell)	Range (pmol/cell)
Dinoflagellates	40	0.1725	0.00011–14.7
Diatoms	15	0.00745	0.0006–0.257
Chlorophytes	9	0.001	0.00015–0.012
Golden algae	4	$8.36 \times 10^{-04}$	0.000149–0.02
Haptophytes	15	0.0158	0.000373–0.148
Cryptophytes	1	0.0213	NA <sup>b</sup>
Rhodophyta	1	0.00231	NA
Cyanobacteria	1	$8.94 \times 10^{-06}$	$7.45\text{--}10.4 \times 10^{-6}$
Coral	6	0.0826	0.021–3.331
Zooxanthellae	4	0.14	0.048–0.285

<sup>a</sup>The data in this table was produced by synthesizing data from the following sources: (Keller, 1989; Corn et al., 1996; Yoch, 2002; Wiesemeier and Pohnert, 2007; Breckels et al., 2010; Caruana, 2010; Caruana and Malin, 2014). <sup>b</sup>Not applicable.



sea-ice diatom *Fragilariopsis cylindrus*. Proteins from the same enzyme classes proposed in the *U. intestinalis* pathway were more abundant when *F. cylindrus* was exposed to conditions that increased DMSP production. However, the activities of these proteins still need to be verified (Lyon et al., 2011). Orthologs for the genes encoding a NADPH-reductase and an AdoMet-dependent methyltransferase have also been found in the corals *Acropora millepora* and *Acropora digitifera* and in the coral dinoflagellate symbiont *Symbiodinium*, all known DMSP producers. Based on the collective data, Raina et al. (2013) hypothesized that the enzymes of the DMSP biosynthetic pathway are conserved between diatoms, alveolates, green algae, and corals. Interestingly, a study of the diatom *Thalassiosira pseudonana* did not identify any of the same proteins proposed for the *F. cylindrus* biosynthetic pathway under conditions that increased intracellular DMSP levels, suggesting that it may contain an alternative pathway (Kettles et al., 2014).

A recent study has reported the biosynthesis of DMSP by marine bacteria. DMSP production was observed from *Oceanicola batsensis* HTCC2597, *Pelagibaca bermudensis* HTCC2601, *Sediminimonas qiaohouensis* DSM21189, *Amorphus coralli* DSM18348, *Sagittula stellata* E-37, *Labrenzia aggregata* LZB033, *Labrenzia aggregata* IAM12614, and *Thalassobaculum salexigens* DSM19539 (Curson et al., 2017). DMSP biosynthesis in marine bacteria proceeds in a similar manner to that observed in algae and phytoplankton, via the methionine transamination based pathway. A methyltransferase gene, *dysB*, was identified in marine Alphaproteobacteria and appears to be the key enzyme for DMSP biosynthesis in these microorganisms. When *dysB* was cloned into the non-DMSP producer *Rhizobium leguminosarum*, the ability to synthesize DMSP was conferred. Thus, the addition of a single gene, in certain cases, is sufficient to enable the production of DMSP. Further, *dysB* expression from *L. aggregata* LZB033 is up-regulated during increased salinity, nitrogen limitation, and at low temperatures, conditions already predicted to stimulate DMSP production in marine phytoplankton and algae. Selective pressures, like changes in salinity or nitrogen limitation, could result in the acquisition of *dysB* by marine bacteria to enable DMSP biosynthesis and gain a competitive advantage in their environment (Curson et al., 2017).

## DMSP CLEAVAGE BY MARINE PHYTOPLANKTON

While the demethylation pathway appears to be unique to marine bacteria, several marine phytoplankton lyse DMSP into DMS. Multiple studies have reported significant DMSP lyase activity within phytoplankton blooms and among individual phytoplankton, including *Phaeocystis* sp., *Heterocapsa triquetra*, *Scrippsiella trochoidea*, and several *Symbiodinium* strains (Stefels et al., 1995; Niki et al., 1997, 2000; Yoch, 2002). To date, while several marine phytoplankton have been observed to produce DMS from DMSP, the genes responsible for this activity have not been identified in most cases. It has been known for many years that *E. huxleyi* cleaves DMSP into DMS and acrylate (Yoch, 2002), but only recently was the responsible gene, *Alma1*,

identified (Alcolombri et al., 2015). *Alma1* is a member of the aspartate racemase superfamily. Based on sequence similarity, *Alma1* and its paralogs from *E. huxleyi* are present in a wide range of phytoplankton as well as certain bacteria, highlighting the diversity of this protein (Yost and Mitchelmore, 2009; Alcolombri et al., 2015). There are seven *Alma1* paralogs within the *E. huxleyi* genome. *Alma1* paralogs from *E. huxleyi*, *Phaeocystis Antarctica*, *A. millepora* (coral), and *Symbiodinium* sp. were synthesized and tested for activity toward DMSP. Of those tested, however, only one *E. huxleyi* paralog, *Alma2*, and a *Symbiodinium* paralog had DMSP lyase activity, indicating that there is still much to learn about the phytoplankton DMSP lyases (Alcolombri et al., 2015).

## BACTERIAL PATHWAYS FOR DMSP METABOLISM

Marine bacteria have developed many uses for DMSP, from a source of reduced sulfur and carbon (Kiene et al., 1999, 2000), to use as an osmolyte (Sunda et al., 2002; Salgado et al., 2014), and potentially a cryoprotectant (Karsten et al., 1996). The details of the bacterial catabolism of DMSP have only recently come to light (Figure 2). The characterization of the enzymes involved in the DMSP demethylation pathway as well as the identification of several DMSP lyases from the DMSP cleavage pathway have provided new insights into the evolution of these enzymatic activities. Some of the enzymes of the demethylation pathway have likely roots in fatty acid  $\beta$ -oxidation (Reisch et al., 2011a,b; Bullock et al., 2014). The DMSP lyases are widely distributed and varied in sequence, structure, and activity (Curson et al., 2011b; Johnston et al., 2016). Many of the enzymes involved in the microbial DMSP catabolic pathways are widespread, particularly among the Proteobacteria (Figure 3). Even those Roseobacters with reduced genomes, such as the lineages SAG-O19, DC5-80-3, and NAC11-7, have been found to encode *dmdA* and at least one DMSP lyase (Zhang et al., 2016). Presumably, the relatively modern evolution of phytoplankton producing high levels of DMSP provided the impetus for developing and maintaining these functions. To learn more about how the degradation pathways evolved, the structural and functional characteristics of the DMSP catabolic enzymes were examined to posit how they may have been adapted from existing enzymes.

## ENZYMATIC CLEAVAGE OF DMSP

The enzymatic cleavage of DMSP produces DMS and acrylate. To date, eight DMSP lyases have been identified (Table 2). The lyases were recently reviewed (Johnston et al., 2016). Except *DddD* which produces 3-hydroxypropionate, these enzymes all carry out the same reaction to form DMS and acrylate despite differing drastically in sequence and size (Todd et al., 2010; Curson et al., 2011b). Based upon a survey of lyase encoding genes in representative genomes of marine bacteria, *dddP* is the most widely distributed (Figure 3). However, *dddD*, *dddW*, *dddQ*, and *dddL* are also relatively common. In contrast, *dddY*,



**TABLE 2 | Identified DMSP lyases and their  $K_m$  for DMSP.**

Protein	Organism	$K_m$ for DMSP	Reference
DddY	<i>A. faecalis</i> , <i>D. acrylicus</i>	1.4 mM, 0.4 mM	de Souza and Yoch, 1995; van der Maarel et al., 1996; Curson et al., 2011a
DddD	<i>Marinomonas</i>	>40 mM <sup>a</sup>	Todd et al., 2007; Alcolombri et al., 2014
DddL	<i>Sulfitobacter</i> EE-36	ND <sup>b</sup>	Curson et al., 2008
DddP	<i>R. lacuscaerulensis</i> , <i>R. nubinihibens</i>	17 mM, 14 mM	Todd et al., 2009; Kirkwood et al., 2010; Wang et al., 2015
DddW	<i>R. pomeroyi</i>	8.7 mM	Todd et al., 2012b; Brummett et al., 2015
DddQ	<i>R. lacuscaerulensis</i>	22 mM	Todd et al., 2011; Li et al., 2013
DddK	<i>P. ubique</i> HTCC1062	82 mM	Sun et al., 2016
Alma1	<i>E. huxleyi</i>	9.0 mM	Alcolombri et al., 2015

<sup>a</sup>Saturation was not observed at 40 mM. <sup>b</sup>No data.

*dddK* and *Alma1* are rare in marine bacteria. There are now several reports of DMSP lyase activity being induced by the presence of DMSP. In *Ruegeria pomeroyi* DSS-3 and *Roseovarius nubinihibens*, *dddP* and *dddQ* expression was induced when cells were pre-grown with DMSP as compared to cells not exposed to DMSP. Likewise, expression of *dddY* increased following growth of *Alcaligenes faecalis* with DMSP (Todd et al., 2007, 2009). Further, a field study in Monterey Bay, California, found that expression of *dddP* increased during mixed-community DMSP-producing phytoplankton blooms (Varaljay et al., 2015). Expression of *R. pomeroyi dddW* also increased after exposure to DMSP in growth medium (Todd et al., 2012b). These observations are consistent with a role of these enzymes in DMSP cleavage.

Evidence for the physiological relevance of the two best studied lyases, *DddP* and *DddQ*, has been mounting. The *dddP* and *dddQ* genes are the most abundant of the bacterial DMSP lyase genes in the marine metagenome as determined by the Global Ocean Sampling Expedition (GOS) (Rusch et al., 2007; Todd et al., 2011). The role of *DddP* and *DddQ* from *R. pomeroyi* DSS-3 and *R. nubinihibens* in DMSP cleavage has been clearly demonstrated. Studies using <sup>14</sup>C or <sup>13</sup>C labeled DMSP show that *Escherichia coli* extracts expressing *dddP* and *dddQ* are able to produce DMS and acrylate from DMSP (Kirkwood et al., 2010; Todd et al., 2011). Additionally, *dddP* and *dddQ* mutants in *R. pomeroyi* produce significantly less DMS when compared with wild-type cells, 50% less in the case of *dddP* and 97% less in the case of *dddQ*. A *dddQ* mutant from *R. nubinihibens* produced 20% less DMS from DMSP, while a *dddP* mutant produced only 10% of the wild-type levels (Todd et al., 2009, 2011; Kirkwood et al., 2010).

The structures of the DMSP lyases provide insights into their evolutionary roots. The crystal structures of *DddP* and *DddQ* from *Ruegeria lacuscaerulensis* and *DddP* from *Roseobacter denitrificans* have been solved (Li et al., 2013; Wang et al., 2015). Data gathered from the available structures suggests that subtle changes in the active sites of these lyases make sulfur containing substrates, like DMSP, the preferred substrates for these enzymes. The sequence and structure of *DddP* most closely resembles that of M24 peptidase. Typically, an M24 peptidase hydrolyzes C-N bonds. *DddP*, however, cleaves C-S bonds (Todd et al., 2009; Wang et al., 2015). Wang and coworkers expressed the recombinant *R. lacuscaerulensis dddP* in *E. coli* and found

it displayed no measurable activity toward the M24 peptidase substrate valine-proline, but it did exhibit DMSP lyase activity, producing acrylate and DMS (Wang et al., 2015). *DddP* is a homodimeric protein in which one monomer has a metal center containing Fe, while the other monomer generally contains Fe, but may also contain Ni, Zn, or Cu instead (Hehemann et al., 2014; Wang et al., 2015). The explanation for the change in substrate preference and activity appears to be due to the change of the active ion from Co or Mn coordinated by five residues in the M24 peptidases to Fe coordinated by six residues in *DddP*. The two metal ions in *DddP* are coordinated with three aspartates, two glutamates, and a histidine residue, which are conserved in the known functional *DddPs* (Hehemann et al., 2014; Wang et al., 2015). The substitution of any of the active site residues for alanine in *DddP* results in the elimination of DMSP lyase activity, indicating that all six are necessary for activity (Kirkwood et al., 2010; Wang et al., 2015). Additionally, two conserved histidine residues in M24 peptidases that help to bind and stabilize substrates are exchanged for aspartate and phenylalanine in *DddP* (Hehemann et al., 2014; Wang et al., 2015). Wang et al. (2015) suggest that this change abolishes the peptidase activity of *DddP* and allows the active site aspartate to act as a nucleophilic base for DMSP cleavage (Wang et al., 2015). It is further proposed that *DddP* is a case of divergent evolution from the M24 peptidases as *DddPs* cluster in a separate clade in phylogenetic analyses. In support of this hypothesis, the M24 peptidase conserved C-domain has up to 31% sequence identity with the C-domain of the *R. lacuscaerulensis dddP*. The N-domain of *DddP*, by contrast is structurally different than the N-domains of M24 peptidases and allows for the formation of a compact dimer and a smaller catalytic cavity for DMSP binding (Wang et al., 2015). In conclusion, *DddP* appears to have acquired specific adaptations for DMSP lyase activity, supporting the assertion that this is its major role.

A structure for *DddQ* from *R. lacuscaerulensis* has recently been solved (Li et al., 2013). *DddQ* is one of the cupin motif containing DMSP lyases, along with *DddW* and *DddL* (Curson et al., 2011b; Johnston et al., 2016). *DddQs* have been identified in a number of *Roseobacters*, but they display substantial amino acid sequence variation, even when multiple copies are present in the same organism (Todd et al., 2011). Despite this variation, certain amino acids in the cupin motifs, two histidines and a glutamate in cupin motif 1 and a histidine in cupin motif

2 are conserved in DddQ, DddW, and DddL. In addition to these conserved amino acids, two tyrosines in motif 1 are highly conserved in all the cupin protein DMSP lyases but not among other cupin proteins. These conserved active site residues are predicted to play a role in DMSP cleavage as substitution at any of these residues decreased activity toward DMSP (Li et al., 2013).

The formation of DMS and acrylate from DMSP is proposed to be the result of a  $\beta$ -elimination reaction (Li et al., 2013; Wang et al., 2015). The DMSP lyases appear to have developed different catalytic mechanisms for carrying out the same reaction, indicating separate evolutionary paths to this activity. DddP is proposed to implement an ion shift. When DMSP enters the active site, a moveable Fe binds to the carboxyl group of DMSP, stabilizing the molecule in the active site, while two other conserved residues, tryptophan and tyrosine, bind to the sulfur in DMSP. This orientation allows for the abstraction of a proton by aspartate from the alpha carbon of the DMSP carboxyl group, cleavage of the C-S bond, and the subsequent formation of a double bond between the alpha and beta carbons of DMSP to produce acrylate (Hehemann et al., 2014; Wang et al., 2015). In DddQ, it has been proposed that binding of DMSP to the metal cofactor causes a conserved tyrosine residue to shift closer to the DMSP molecule. This shift allows the oxygen atom of one of the conserved tyrosine residues to interact with the alpha carbon of DMSP. The resultant conformational change enables the abstraction of a proton from the DMSP carboxyl group by the oxygen atom of tyrosine (Li et al., 2013). The algal DMSP lyase, Alma1, is proposed to function in a similar manner, abstracting a proton from the carbon adjacent to the carboxylate to cause  $\beta$ -elimination and the subsequent release of DMS and acrylate (Alcolombri et al., 2015). Further investigations into the structures and mechanisms of the other DMSP lyases, like the algal Alma1 or DddY, may yield still more variability in reaction mechanisms.

DddY from *A. faecalis* M3A was the first identified DMSP lyase (de Souza and Yoch, 1995). It is the only DMSP lyase that is a periplasmic protein and has no similarity to any other enzyme of known function. DddYs have been identified in *A. faecalis* M3A and *Desulfovibrio acrylicus*, as well as in several *Shewanella* species and *Arcobacter nitrofigilis* DSM7299. DMS production from DMSP was observed in *Shewanella halifaxensis* HAW-EB4, *Shewanella putrefaciens* CN-32, and *A. nitrofigilis* DSM7299 (Curson et al., 2011a). *S. halifaxensis* and *S. putrefaciens* are found in marine sediments and shale sandstone, respectively, while *A. nitrofigilis* can be found in sediment around *Spartina* roots. It is likely that *dddY* was spread via horizontal gene transfer (HGT) among these distantly related bacteria. In addition to *dddY*, *A. faecalis* also has acrylate utilization (*acu*) genes that resemble those used for DMSP and acrylate metabolism in other DMSP-utilizing bacteria (Curson et al., 2011a). More in depth studies of DddY have not been undertaken.

Despite convincing evidence for the physiological role of the DMSP lyases, the affinities for DMSP of the currently known lyases are lower than expected for a natural substrate, displaying  $K_m$ s for DMSP in the millimolar range (Table 2) (Johnston et al., 2016). The  $K_m$ s for the most widely distributed lyases, DddP and DddQ, are among the highest (Rusch et al., 2007;

Todd et al., 2011). The lowest  $K_m$  for DMSP observed thus far is for DddY. The DddYs from *A. faecalis* and *D. acrylicus* have  $K_m$ s for DMSP of 1.4 and 0.4 mM, respectively (Table 2) (de Souza and Yoch, 1996; van der Maarel et al., 1996). Both of these organisms are found in coastal marine sediments and likely obtain DMSP from *Spartina* spp. (Curson et al., 2011a). High  $K_m$  values for DMSP are also shared with the DMSP demethylases (see DmdA below), which is the first committed step of the demethylation pathway. Thus, the low affinities of the lyases may simply reflect the requirement for high intracellular concentrations of DMSP to initiate its metabolism. If DMSP serves as an osmolyte in bacterioplankton, cells should maintain high concentrations in the cell. For instance, during growth on DMSP, a concentration of 70 mM has been observed in *R. pomeroyi*. Under these conditions, low  $K_m$ s for DMSP are not necessary for DMSP lyases to function effectively *in vivo* (Reisch et al., 2008). Concentrations of DMSP in ocean surface waters range from less than 1 nM in the open ocean to micromolar levels within phytoplankton blooms (van Duyl et al., 1998). Senescence and autolysis of DMSP producers like *Spartina* or phytoplankton can also produce microenvironments with high concentrations of DMSP (de Souza and Yoch, 1995). Provided a bacterium has the necessary transporters for the uptake of DMSP, intracellular concentrations of DMSP have the potential to reach to millimolar levels (Kiene, 1998; Kiene and Williams, 1998; Kiene et al., 1998; Todd et al., 2010).

In conclusion, the sequence and structural variability of the DMSP lyases that have been identified so far indicates that they likely evolved independently. For this to happen, the new activity must be readily acquired in evolution from multiple ancestral enzymes. Moreover, there must be strong selective pressures to maintain this function in very different groups of organisms. In addition, some bacteria contain multiple DMSP lyases, and it is possible that their physiological functions are somewhat different. This would allow cells to maintain lyases with the same catalytic activity but different regulatory or other functional properties.

## BACTERIAL DEMETHYLATION OF DMSP

The DMSP demethylation pathway consists of a series of reactions that convert DMSP into methanethiol (MeSH), HS-CoA, CO<sub>2</sub>, and acetaldehyde (Reisch et al., 2011a,b). While DMS production from phytoplankton has been observed, there is no indication that these organisms possess the demethylation pathway. Instead, the demethylation pathway is restricted to the Alphaproteobacteria (Figure 3). Based on the current evidence, it seems likely that the individual steps of the demethylation pathway may have evolved independently.

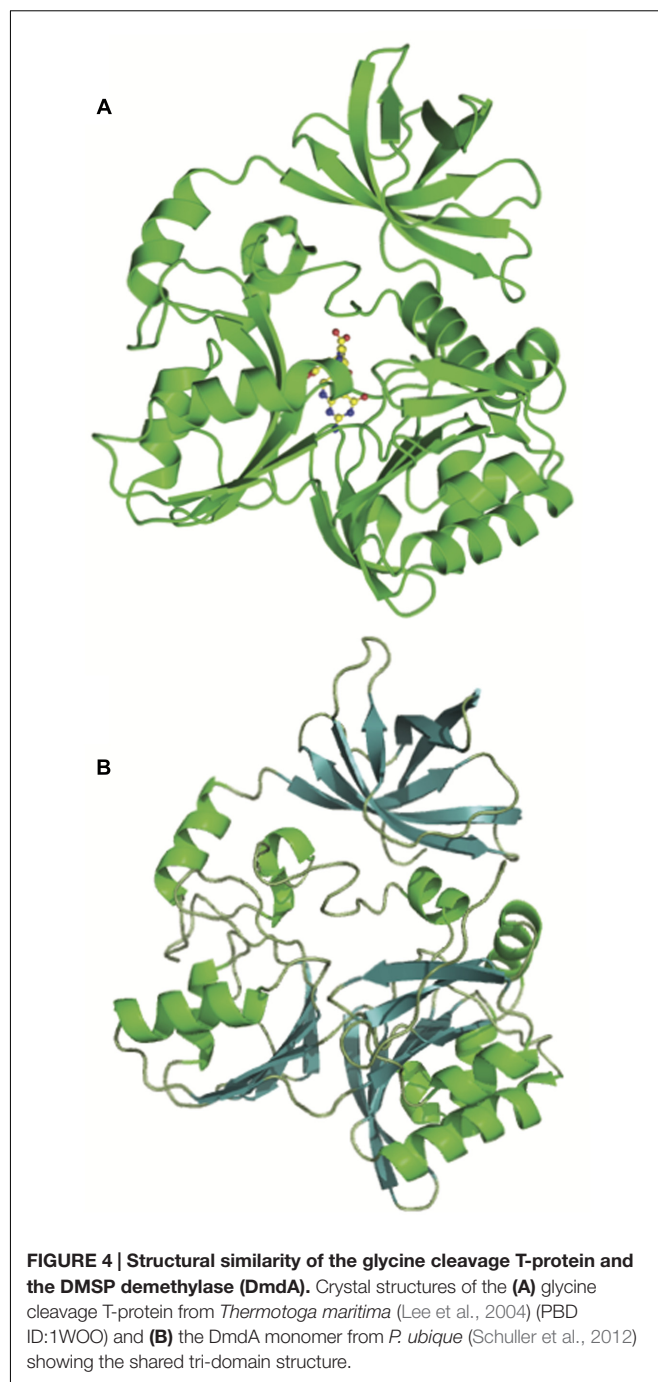
### DmdA: AN ADAPTED GLYCINE CLEAVAGE T-PROTEIN

The initial step of the demethylation pathway is mediated by the DMSP demethylase DmdA (Figure 2). This step also commits

DMSP to the demethylation pathway because demethylation precludes the formation of DMS (Howard et al., 2006; Reisch et al., 2008, 2011b). As with the DMSP lyases, the  $K_m$ s for DMSP of the two characterized DmdAs from *R. pomeroyi* and *Pelagibacter ubique* are relatively high, 5.4 and 13.2 mM, respectively. The deletion of *dmdA* from *R. pomeroyi*, however, results in a mutant incapable of producing MeSH, indicating that this gene encodes the only protein in *R. pomeroyi* able to perform this reaction (Howard et al., 2006; Reisch et al., 2008). Additionally, field measurements indicate that *dmdA* expression is upregulated during blooms of DMSP-producing phytoplankton (Varaljay et al., 2015). DmdA in *R. pomeroyi* was initially annotated as a glycine cleavage T-protein (GcvT) (Reisch et al., 2011a). However, when analyzed phylogenetically, DmdA-like proteins share sequence identity ranging from 22 to 26% with GcvT, dimethylglycine oxidase and sarcosine oxidase, but form a separate clade from known GcvTs.

The crystal structure of DmdA from *P. ubique* provides further evidence supporting a common ancestry for DmdA and GcvT (Figure 4). Schuller et al. (2012) described the structure of DmdA, noting that while DmdA is structurally similar to GcvT, the low sequence similarity between the two indicated that the enzymes are evolutionarily distant (Schuller et al., 2012). Both proteins possess a very similar tri-domain structure (Figure 4) with the conserved residues between the proteins being mainly involved with tetrahydrofolate (THF) binding. Specifically, the residues that interact with the folate moiety and those involved in the ring stacking of THF are highly conserved (Lee et al., 2004; Reisch et al., 2008; Schuller et al., 2012). In contrast, DmdA possesses high substrate specificity for DMSP and closely related compounds, so the binding site for this substrate must differ from that of GcvT. Despite structural similarity, DmdA and GcvT are mechanistically distinct. DmdA produces 5-methyl-THF from DMSP as the result of a redox-neutral methyl transfer while GcvT converts glycine to 5,10-methylene-THF (Howard et al., 2006; Reisch et al., 2008; Schuller et al., 2012). Small changes to the THF-binding fold in DmdA allow for hydrogen bond formation between amino acid residues in the fold and THF, enabling DmdA to carry out a redox-neutral methyl transfer to produce 5-methyl-THF. Overall, the mechanism of DmdA catalysis appears to be more similar to the S-adenosylmethionine SAM-dependent N-methyltransferases than the more closely related GcvTs (Schuller et al., 2012).

Phylogenetic analysis reveals that GcvTs and other similar proteins are nearly universally distributed among the prokaryotes, while DmdA proteins cluster separately. DmdA appears to be most prevalent among members of the Alphaproteobacteria (Figure 3) (Reisch et al., 2008; Moran et al., 2012). DmdA may have originally been a GcvT, but the development of a new activity and substrate preference has uniquely adapted this enzyme for DMSP metabolism (Reisch et al., 2008). Other organisms without DmdA may simply maintain DMSP as an osmolyte or utilize one of the many DMSP lyases identified so far to metabolize it.



## THE FLEXIBILITY OF DmdB AND DmdC

Once methylmercaptopropionate (MMPA) is produced by DmdA, it is converted first to MMPA-CoA by the MMPA-CoA ligase or DmdB and then to methylthioacryloyl (MTA)-CoA by the MMPA-CoA dehydrogenase DmdC (Reisch et al., 2011a). In contrast to the narrower distribution of DmdA, DmdB and DmdC are found in up to 60% of surface ocean bacteria, assuming one copy per cell, as well as in bacteria from terrestrial and other environments (Figure 3) (Reisch et al., 2011b).



The DmdB and DmdC enzymes characterized thus far show activity with a wide range of substrates, mostly with small to medium chain length fatty acids and their CoA derivatives (Reisch et al., 2011b; Bullock et al., 2014). These enzymes probably did not originate specifically for DMSP metabolism, potentially having been recruited from the pathways of methionine degradation and  $\beta$ -fatty acid oxidation (Reisch et al., 2011a,b). The ability of DmdB and DmdC to act upon MMPA and MMPA-CoA is a demonstration of the plasticity and flexibility of these enzymes.

*R. pomeroyi* possesses more than 20 CoA ligases, but not all are predicted to have activity with MMPA. *R. pomeroyi* has two DmdB isozymes, RPO\_DmdB1 and RPO\_DmdB2 (Table 3). RPO\_DmdB1 has a  $K_m$  of 0.08 mM for MMPA but even lower  $K_m$ s for butyrate and propionate, 0.02 and 0.04 mM, respectively. RPO\_DmdB2 has a  $K_m$  for MMPA similar to that of RPO\_DmdB1, 0.07 mM, but this was the lowest  $K_m$  it displayed with any of the substrates tested (Table 3) (Bullock et al., 2014). There are distinct differences between the DmdB enzymes from marine and non-marine microorganisms. Particularly, only the DmdBs from marine microorganisms are inhibited by concentrations of DMSP likely to be present in the cell. The *R. pomeroyi* DmdB isozymes exhibit different regulatory mechanisms to reverse this inhibition. RPO\_DmdB1 responds to changes in cellular energy charge, while RPO\_DmdB2 responds to increases in MMPA concentration (Bullock et al., 2014). These regulatory mechanisms may have developed during the specialization of the DmdB isozymes for DMSP rather than fatty acid metabolism. Because they are not found in the DmdBs from terrestrial bacteria, they appear to be specific adaptations to the importance of DMSP as a nutrient for marine bacteria.

Three DmdC isozymes were identified in *R. pomeroyi* and verified to have activity toward MMPA-CoA (Reisch et al., 2011b). The  $K_m$  of one of the DmdC isozymes (SPO3804; DmdC1) from *R. pomeroyi* for MMPA-CoA is low at 0.03 mM. However, lower  $K_m$ s were observed for this enzyme with caproyl-CoA, valeryl-CoA, and butyryl-CoA. Thus, MMPA-CoA is not necessarily the preferred substrate for this enzyme. Instead, the substrate specificity of DmdC appears, like DmdB, to be based primarily on the length of the carbon chain of a substrate.

DmdB and DmdC isozymes are more widely distributed than DmdA, suggesting that these enzymes may be important in organisms that either metabolize only MMPA but not DMSP

or possess pathways that form MMPA from substrates other than DMSP. Methionine degradation is one potential source of MMPA (Steele and Benevenga, 1979), a side reaction of the methionine salvage pathway can also produce MMPA (Sekowska et al., 2004; Albers, 2009). *Xanthomonas campestris* produces MMPA to induce bacterial blight in cassava (Perreux et al., 1982; Ewbank and Maraite, 1990), and many plants, particularly fruiting plants, produce sulfur volatiles which closely resemble MMPA in structure (e.g., 3-methylthio-propanol, 3-methylthio-propanal, and ethyl-3-methylthio-propionate) (Gonda et al., 2013). These compounds might also be substrates of DmdB. Alternatively, the primary function of DmdB and DmdC in many bacteria may be fatty acid oxidation, and MMPA may only be an occasional substrate.

## DMSP SPECIFIC ENOYL-CoA HYDRATASES: DmdD AND AcuH

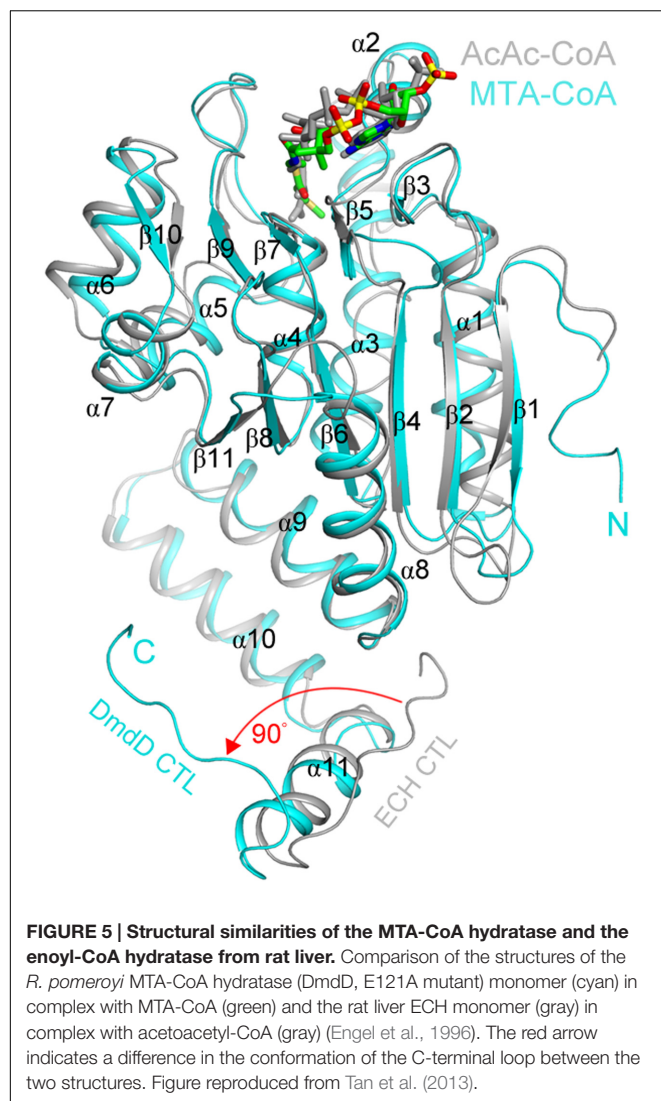
DmdD, a member of the crotonase superfamily, appears to be uniquely adapted for the metabolism of DMSP. DmdD has a crystal structure largely similar to that of other crotonases, a hexamer made up of a dimer of trimers. Figure 5 shows an overlay of one DmdD monomer with a monomer of rat liver enoyl-CoA hydratase (ECH). DmdD is similar to the rat liver ECH, sharing 32% amino acid identity. The main difference is that in DmdD the C-terminal loops of one of the trimers is oriented so that it can interact with the phosphate groups of CoA (Figure 5) (Tan et al., 2013). The same glutamate residues that are conserved and important for catalysis in the rat liver ECHs are also conserved in DmdD. However, DmdD is not nearly as efficient as an ECH at catalyzing the hydration of crotonyl-CoA, with a catalytic efficiency of  $2100 \text{ mM}^{-1}\text{s}^{-1}$  compared with the typical values of  $45000\text{--}119000 \text{ mM}^{-1}\text{s}^{-1}$  of other crotonases. DmdD instead displays a  $K_m$  of 0.008 mM for MTA-CoA and a high catalytic efficiency,  $5400 \text{ mM}^{-1}\text{s}^{-1}$  (Kiema et al., 1999; Feng et al., 2002; Tan et al., 2013). This greater catalytic efficiency only applies to reactions with MTA-CoA and appears to be due, at least in part, to the structure of MTA-CoA. The combination of the double bond and sulfur atom in MTA-CoA appear to be key for high rates DmdD hydrolysis activity as reactions with

**TABLE 3 | Apparent kinetic constants for *R. pomeroyi* DSS-3 MMPA-CoA ligases RPO\_DmdB1 and RPO\_DmdB2<sup>a</sup>.**

Substrate	Enzyme	$K_m$ (mM) <sup>a</sup>	$k_{cat}$ (s <sup>-1</sup> )	$k_{cat}/K_m$ (mM <sup>-1</sup> s <sup>-1</sup> )
Methylmercaptopropionate	RPO_DmdB1	$0.08 \pm 0.02$	18.7	233
	RPO_DmdB2	$0.07 \pm 0.02$	14.9	213
Butyrate	RPO_DmdB1	$0.02 \pm 0.01$	14.4	1031
	RPO_DmdB2	$0.12 \pm 0.03$	7.2	71
Propionate	RPO_DmdB1	$0.04 \pm 0.01$	10.8	271
	RPO_DmdB2	$3.11 \pm 1.13$	3.7	1.2
Acrylate	RPO_DmdB1	$0.9 \pm 0.2$	14.3	16
	RPO_DmdB2	$5.25 \pm 2.1$	1.0	0.2

<sup>a</sup> $K_m$  (mM) is shown ( $\pm$ SE) from three independent experiments.  $k_{cat}$  is expressed in units of s<sup>-1</sup> and  $k_{cat}/K_m$  in units of mM<sup>-1</sup> s<sup>-1</sup>. Reproduced from Bullock et al. (2014).





MMPA-CoA and crotonyl-CoA occur at lower rates (Tan et al., 2013).

While DmdD is highly efficient at catalyzing the hydration of MTA-CoA, it is not widely distributed (Reisch et al., 2011b). DmdD is absent from the majority of marine bacteria that utilize the demethylation pathway, i.e., possess DmdA (Figure 3). An ortholog of DmdD has been identified in the DmdD negative *R. lacuscaerulensis* as well as in *R. pomeroyi* (Reisch et al., 2011b). This enzyme, now designated AcuH for acrylate utilization hydratase, is an ECH with high activity toward acryloyl-CoA and crotonyl-CoA, but also displays activity toward MTA-CoA. The designation is similar to that of the acryloyl-CoA reductase AcuI. As a result of its activity toward acryloyl-CoA and MTA-CoA, AcuH is predicted to play an important role in the metabolism of acrylate formed from the cleavage pathway as well as MTA-CoA formed from the demethylation pathway (Figure 2) (Sullivan et al., 2011; Todd et al., 2012a; Reisch et al., 2013). By contrast, DmdD has no activity toward acryloyl-CoA. AcuH is less efficient than DmdD at hydrolyzing MTA-CoA,

however, it is far more common, being found in a wide range of microorganisms, including those in the Roseobacter clade (Figure 3). AcuH appears to be a more versatile enzyme than DmdD and has maintained more of its functional similarity to other ECHs. Since AcuH likely functions in both the cleavage and the demethylation pathways, this strategy gives the cells increased metabolic flexibility and may also protect against acryloyl toxicity. DmdD, by contrast, has adapted specifically to function in the demethylation pathway, possibly allowing organisms which possess DmdD to utilize DMSP more efficiently.

## LINKS BETWEEN THE BACTERIAL DMSP CLEAVAGE AND DEMETHYLATION PATHWAYS

Interactions between the cleavage pathway and demethylation pathways in organisms that contain both are an ongoing field of study. One proposal is that a 'bacterial switch' allows bacteria possessing both pathways to alternate between producing more or less DMS and MeSH (Kiene et al., 2000; Simo, 2001). While there is currently no consensus as to what signal controls the switch, the identification of the acrylate utilization enzymes AcuH and AcuI has begun to shed light on the topic. AcuH, as mentioned above, may function in both the cleavage and demethylation pathways (Reisch et al., 2011b, 2013). AcuI is an acryloyl-CoA reductase whose gene has been found immediately downstream of *dmdA* in many members of the Roseobacter clade. In *R. pomeroyi*, *dmdA* and *acuI* are co-regulated with acrylate acting as an inducer (Sullivan et al., 2011; Todd et al., 2012a). Since acrylate and acryloyl-CoA are inhibitory for bacterial growth, it has been proposed that AcuI maintains cellular acrylate concentrations below inhibitory levels. Thus, when acrylate concentrations increase as a result of DMSP lyase activity, AcuI and DmdA co-regulation results in increased activity of both enzymes. Elevated AcuI activity then alleviates inhibition caused by the build-up of acrylate, while increases in DmdA activity stimulate the demethylation pathway, allowing DMSP to be utilized in a manner that does not produce acrylate (Todd et al., 2012a). The activity of the demethylation pathway may also respond to carbon and energy limitation, with regulation resulting from changes in the energy charge of the cell (Bullock et al., 2014). Further research is still needed to investigate the physiological cues for balancing the demethylation and cleavage pathways.

## CONCLUSION: EVOLUTION OF DMSP METABOLISM

It is unclear what the original impetus for the development of DMSP biosynthetic pathway may have been. The proposed roles for DMSP in marine phytoplankton as an osmolyte, antioxidant, predator deterrent, cryoprotectant, and as an energy overflow mechanism each could provide great benefits, particularly in consistently changing marine environments. If DMSP was originally produced as part of an overflow mechanism for dealing

with unbalanced growth due to nutrient limitation, the other benefits provided by this compound may have selected for the maintenance of this pathway. The case has been made for the co-evolution of the marine Roseobacter and the DMSP producing-phytoplankton. Members of the Roseobacter clade are abundant in coastal waters and are one of the main bacterial groups enriched during DMSP-producing phytoplankton blooms (Gonzalez et al., 2000; Zubkov et al., 2001; Moran et al., 2007). Based on independent time estimates assisted by the cyanobacterial fossil calibration and estimates derived from the mutation rate clock method, the Roseobacter ancestor likely underwent a genome expansion, coincident with the increase in abundance and diversification of the dinoflagellates and coccolithophores around 250 mya (Luo et al., 2013; Luo and Moran, 2014; Sun et al., 2017). Thus, the radiation of dinoflagellates and coccolithophores may have provided new environments for members of the Roseobacter clade, in much the same way that the breakup of Pangea and changes in ocean redox chemistry created new environments for the proliferation of the red plastid lineage members (Whitfield, 2001; Quigg et al., 2003; Luo et al., 2013).

Research into the diversity of bacterial DMSP utilization enzymes and their regulation is still ongoing. New Roseobacter isolates showing adaptations to their particular environmental niches are continually being discovered. Recently, two new members of the Roseobacter clade were isolated from deep-sea water, *Thiobacimonas profunda* JLT2016 and *Pelagibaca abyssii* JLT2014. While these isolates did not possess DMSP metabolic genes, their genomes included genes for inorganic sulfur oxidation and CO<sub>2</sub> fixation, further demonstrating the metabolic flexibility of this clade (Tang et al., 2016) and illustrating that members of the Roseobacter clade have exploited different routes to metabolically thrive in their environments (Luo et al., 2014). Since the first Roseobacter genome expansion 250 mya, many factors could have played a role in the development of metabolic pathways for the utilization of a specific carbon source like DMSP. There was not one path for these organisms to follow but many, allowing for a diversity of solutions to a single goal. As DMSP became more readily available in the environment, marine organisms, such as *R. pomeroyi* and other Roseobacters, likely adapted to utilize this compound as a source of carbon as well as reduced sulfur. Members of the Roseobacter clade were well poised for this task, being metabolically versatile bacteria and thus able to thrive in dynamic environments (Moran et al., 2007, 2012). Exposure to DMSP would be the driving force behind this evolution of function, putting pressure on organisms to adapt proteins already encoded in their genomes to utilize this new compound. Possible examples of this can be seen in the enzymes DmdA and DmdD from the demethylation pathway and DddP from the cleavage pathway. Each of these enzymes

became more specialized to function in DMSP metabolism. Although structurally similar to their likely ancestral enzymes, they have undergone major changes in substrate specificities and, in some cases, enzymatic mechanism. AcuH seems to have adopted a different strategy and can function as both a MTA-CoA hydratase in the demethylation pathway and an acryloyl-CoA hydratase in acrylate metabolism. DmdB and DmdC also maintained activity with a wide range of substrates while adapting to function efficiently in DMSP metabolism as well. In these latter cases, there appears to have been minimal adaptations to DMSP metabolism, with changes in regulatory strategy as well as small changes in substrate specificity to accommodate a novel substrate.

Enzymes are known to diverge from a parental function to develop new substrate specificities, often via the duplication of genes encoding a multifunctional and multispecific enzymes that then undergo an alteration in substrate specificity (Noda-Garcia et al., 2013). Although a certain core set of amino acid residues are required for functionality and structure, there is also room for variation and change, allowing evolution of new functions and substrate specificities (Perona and Craik, 1997). Thus, the amino acid sequences and structures of the enzymes catalyzing the DMSP catabolic reactions do not vary greatly from their non-DMSP degrading counterparts. From this perspective, the enzymes of the DMSP demethylation and cleavage pathways are examples of the various processes of enzyme adaptation and evolution that occurred within the Roseobacter clade in the last 250 million years.

## AUTHOR CONTRIBUTIONS

WW and HB were responsible for the conceptualization and design of this manuscript. HB and HL collected data and analyzed the literature for this review. HB drafted the original manuscript. WW and HL reviewed and edited the manuscript. HB, WW, and HL provided final approval of the manuscript prior to submission.

## FUNDING

This work was supported in part by grants from the National Science Foundation MCB1158037 and OCE1342694.

## ACKNOWLEDGMENTS

We thank Norman Hassell for assistance with computational data analysis and Mak Yun Yi for assistance with literature collection.

## REFERENCES

Albers, E. (2009). Metabolic characteristics and importance of the universal methionine salvage pathway recycling methionine from 5'-methylthioadenosine. *IUBMB Life* 61, 1132–1142. doi: 10.1002/iub.278

Alcolombri, U., Ben-Dor, S., Feldmesser, E., Levin, Y., Tawfik, D. S., and Vardi, A. (2015). MARINE SULFUR CYCLE. Identification of the algal dimethyl sulfide-releasing enzyme: a missing link in the marine sulfur cycle. *Science* 348, 1466–1469. doi: 10.1126/science.aab1586

- Alcolombri, U., Laurino, P., Lara-Astiaso, P., Vardi, A., and Tawfik, D. S. (2014). DddD is a CoA-transferase/lyase producing dimethyl sulfide in the marine environment. *Biochemistry* 53, 5473–5475. doi: 10.1021/bi500853s
- Allen, A. E. (2005). Defining the molecular basis for energy balance in marine diatoms under fluctuating environmental conditions. *J. Phycol.* 6, 1073–1076. doi: 10.1111/j.1529-8817.2005.00156.x
- Allen, A. E., Vardi, A., and Bowler, C. (2006). An ecological and evolutionary context for integrated nitrogen metabolism and related signaling pathways in marine diatoms. *Curr. Opin. Plant Biol.* 9, 264–273. doi: 10.1016/j.pbi.2006.03.013
- Archibald, J. M. (2009). The puzzle of plastid evolution. *Curr. Biol.* 19, R81–R88. doi: 10.1016/j.cub.2008.11.067
- Archibald, J. M., and Keeling, P. J. (2002). Recycled plastids: a 'green movement' in eukaryotic evolution. *Trends Genet.* 18, 577–584. doi: 10.1016/S0168-9525(02)02777-4
- Asada, K., and Takahashi, M. (1987). "Production and scavenging of active oxygen in photosynthesis," in *Photoinhibition*, eds D. J. Kyle, C. B. Osmond, and C. J. Arntzen (Amsterdam: Elsevier), 227–287.
- Bhattacharya, D., and Medlin, L. (1998). Algal phylogeny and the origin of land plants. *Plant Physiol.* 116, 9–15. doi: 10.1104/pp.116.1.9
- Bown, P. R., Lees, J. A., and Young, J. R. (2004). "Calcareous nannoplankton diversity and evolution through time," in *Coccolithophores*, eds H. S. Thierstein and J. R. Young (Berlin: Springer-Verlag), 481–508.
- Breckels, M. N., Boakes, D. E., Codling, E. A., Malin, G., Archer, S. D., and Steinke, M. (2010). Modelling the concentration of exuded dimethylsulfoniopropionate (DMSP) in the boundary layer surrounding phytoplankton cells. *J. Plankton Res.* 32, 253–257. doi: 10.1093/plankt/fbp116
- Brummett, A. E., Schnicker, N. J., Crider, A., Todd, J. D., and Dey, M. (2015). Biochemical, kinetic, and spectroscopic characterization of *Ruegeria pomeroyi* DddW–A mononuclear iron-dependent DMSP lyase. *PLoS ONE* 10:e0127288. doi: 10.1371/journal.pone.0127288
- Bullock, H. A., Reisch, C. R., Burns, A. S., Moran, M. A., and Whitman, W. B. (2014). Regulatory and functional diversity of methylmercaptopyruvate coenzyme A ligases from the dimethylsulfoniopropionate demethylation pathway in *Ruegeria pomeroyi* DSS-3 and other proteobacteria. *J. Bacteriol.* 196, 1275–1285. doi: 10.1128/JB.00026-14
- Caruana, A. (2010). *DMS and DMSP Production by Marine Dinoflagellates*. Ph.D. dissertation, University of East Anglia, Norwich.
- Caruana, A., and Malin, G. (2014). The variability in DMSP content and DMSP lyase activity in marine dinoflagellates. *Prog. Oceanogr.* 120, 410–424. doi: 10.1016/j.pocean.2013.10.014
- Challenger, F., and Simpson, M. I. (1948). Studies on biological methylation; a precursor of the dimethyl sulphide evolved by *Polysiphonia fastigiata*; dimethyl-2-carboxyethylsulphonium hydroxide and its salts. *J. Chem. Soc.* 3, 1591–1597. doi: 10.1039/jr9480001591
- Chin, M., and Jacob, D. J. (1996). Anthropogenic and natural contributions to tropospheric sulfate: a global model analysis. *J. Geophys. Res.* 101, 18691–18699. doi: 10.1029/96jd01222
- Corn, M., Belviso, S., Partensky, F., Simon, N., and Christaki, U. (1996). "Origin and importance of picoplanktonic DMSP" in *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds*, ed. R. Kiene (New York, NY: Plenum Press), 191–201.
- Cortes, M. Y., Bollmann, J., and Thierstein, H. R. (2001). Coccolithophore ecology at the HOT station HOT. *Hawaii. Deep Sea Res. II* 48, 1957–1981. doi: 10.1016/S0967-0645(00)00165-X
- Curson, A. R., Liu, J., Bermejo Martinez, A., Green, R. T., Chan, Y., Carrion, O., et al. (2017). Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nat. Microbiol.* 2, 17009. doi: 10.1038/nmicrobiol.2017.9
- Curson, A. R., Rogers, R., Todd, J. D., Brearley, C. A., and Johnston, A. W. (2008). Molecular genetic analysis of a dimethylsulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine alpha-proteobacteria and *Rhodobacter sphaeroides*. *Environ. Microbiol.* 10, 757–767. doi: 10.1111/j.1462-2920.2007.01499.x
- Curson, A. R., Sullivan, M. J., Todd, J. D., and Johnston, A. W. (2011a). DddY, a periplasmic dimethylsulfoniopropionate lyase found in taxonomically diverse species of Proteobacteria. *ISME J.* 5, 1191–1200. doi: 10.1038/ismej.2010.203
- Curson, A. R., Todd, J. D., Sullivan, M. J., and Johnston, A. W. (2011b). Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nat. Rev. Microbiol.* 9, 849–859. doi: 10.1038/nrmicro2653
- de Souza, M. P., and Yoch, D. C. (1995). Comparative physiology of dimethyl sulfide production by dimethylsulfoniopropionate lyase in *Pseudomonas doadoroffii* and *Alcaligenes* sp. strain M3A. *Appl. Environ. Microbiol.* 61, 3986–3991.
- de Souza, M. P., and Yoch, D. C. (1996). Differential metabolism of dimethylsulfoniopropionate and acrylate in saline and brackish intertidal sediments. *Microb. Ecol.* 31, 319–330. doi: 10.1007/BF00171575
- Delwiche, C. F. (1999). Tracing the thread of plastid diversity through the tapestry of life. *Am. Nat.* 154, S164–S177. doi: 10.1086/303291
- Engel, C. K., Mathieu, M., Zeelen, J. P., Hiltunen, J. K., and Wierenga, R. K. (1996). Crystal structure of enoyl-coenzyme A (CoA) hydratase at 2.5 angstroms resolution: a spiral fold defines the CoA-binding pocket. *EMBO J.* 15, 5135–5145.
- Ewbank, E., and Maraite, H. (1990). Conversion of methionine to phytotoxic 3-methylthiopropionic acid by *Xanthomonas campestris* pv. *manihotis*. *J. Gen. Microbiol.* 136, 1185–1189. doi: 10.1099/00221287-136-7-1185
- Falkowski, P. G., Katz, M. E., Knoll, A. H., Quigg, A., Raven, J. A., Schofield, O., et al. (2004a). The evolution of modern eukaryotic phytoplankton. *Science* 305, 354–360. doi: 10.1126/science.1095964
- Falkowski, P. G., Schofield, O., Katz, M. E., Van De Schootbrugge, B., and Knoll, A. H. (2004b). "Why is the land green and the ocean red?," in *Coccolithophores*, eds H. R. Thierstein and J. R. Young (Amsterdam: Elsevier), 429–453.
- Feng, Y., Hofstein, H. A., Zwahlen, J., and Tonge, P. J. (2002). Effect of mutagenesis on the stereochemistry of enoyl-CoA hydratase. *Biochemistry* 41, 12883–12890. doi: 10.1021/bi020382g
- Fridovich, I. (1998). Oxygen toxicity: a radical explanation. *J. Exp. Biol.* 201(Pt 8), 1203–1209.
- Gage, D. A., Rhodes, D., Nolte, K. D., Hicks, W. A., Leustek, T., Cooper, A. J., et al. (1997). A new route for synthesis of dimethylsulphoniopropionate in marine algae. *Nature* 387, 891–894. doi: 10.1038/43160
- Gonda, I., Lev, S., Bar, E., Sikron, N., Portnoy, V., Davidovich-Rikanati, R., et al. (2013). Catabolism of L-methionine in the formation of sulfur and other volatiles in melon (*Cucumis melo* L.) fruit. *Plant J.* 74, 458–472. doi: 10.1111/tip.12149
- Gonzalez, J. M., Simo, R., Massana, R., Covert, J. S., Casamayor, E. O., Pedros-Alio, C., et al. (2000). Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl. Environ. Microbiol.* 66, 4237–4246. doi: 10.1128/AEM.66.10.4237-4246.2000
- Hanson, A. D., Rivoal, J., Paquet, L., and Gage, D. A. (1994). Biosynthesis of 3-dimethylsulfoniopropionate in *Wollastonia biflora* (L.) DC. Evidence that S-methylmethionine is an intermediate. *Plant Physiol.* 105, 103–110. doi: 10.1104/pp.105.1.103
- Haq, B. U., Hardenbol, J., and Vail, P. R. (1987). Chronology of fluctuating sea levels since the triassic. *Science* 235, 1156–1167. doi: 10.1126/science.235.4793.1156
- Harwood, D. M., and Nikolaev, V. A. (1995). "Cretaceous diatoms: morphology, taxonomy, biostratigraphy," in *Siliceous microfossils*, eds C. D. Blome, P. M. Whalen, and K. M. Reed (Lawrence, KS: Paleontological Society), 81–106.
- Hatakeyama, S., Okuda, M., and Akimoto, H. (1982). Formation of sulfur-dioxide and methanesulfonic-acid in the photo-oxidation of dimethyl sulfide in the air. *Geophys. Res. Lett.* 9, 583–586. doi: 10.1029/GL009i005p00583
- Hehemann, J. H., Law, A., Redecke, L., and Boraston, A. B. (2014). The structure of RdDddP from *Roseobacter denitrificans* reveals that DMSP lyases in the DddP-family are metalloenzymes. *PLoS ONE* 9:e103128. doi: 10.1371/journal.pone.0103128
- Howard, E. C., Henriksen, J. R., Buchan, A., Reisch, C. R., Burgmann, H., Welsh, R., et al. (2006). Bacterial taxa that limit sulfur flux from the ocean. *Science* 314, 649–652. doi: 10.1126/science.1130657
- Husband, J. D., Kiene, R. P., and Sherman, T. D. (2012). Oxidation of dimethylsulfoniopropionate (DMSP) in response to oxidative stress in *Spartina alterniflora* and protection of a non-DMSP producing grass by exogenous DMSP plus acrylate. *Environ. Exp. Bot.* 79, 44–48. doi: 10.1016/j.envexpbot.2012.01.006
- Ito, T., Asano, Y., Tanaka, Y., and Takabe, T. (2011). Regulation of biosynthesis of dimethylsulfoniopropionate and its uptake in sterile mutant of *Ulva pertusa* (Chlorophyta). *J. Phycol.* 47, 517–523. doi: 10.1111/j.1529-8817.2011.00977.x



- James, F., Paquet, L., Sparace, S. A., Gage, D. A., and Hanson, A. D. (1995). Evidence implicating dimethylsulfoniopropionaldehyde as an intermediate in dimethylsulfoniopropionate biosynthesis. *Plant Physiol.* 108, 1439–1448. doi: 10.1104/pp.108.4.1439
- Johnston, A. W., Green, R. T., and Todd, J. D. (2016). Enzymatic breakage of dimethylsulfoniopropionate—a signature molecule for life at sea. *Curr. Opin. Chem. Biol.* 31, 58–65. doi: 10.1016/j.cbpa.2016.01.011
- Karsten, U., Kuck, K., Vogt, C., and Kirst, G. O. (1996). “Dimethylsulphoniopropionate production in phototrophic organisms and its physiological function as a cryoprotectant,” in *Biological and Environmental Chemistry of DMSP and Related Sulphonium Compounds*, eds R. P. Kiene, P. T. Visscher, M. D. Keller, and G. O. Kirst (New York, NY: Plenum Press), 143–153.
- Keeling, P. J. (2010). The endosymbiotic origin, diversification and fate of plastids. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 729–748. doi: 10.1098/rstb.2009.0103
- Keller, M. (1989). Dimethyl sulfide production and marine phytoplankton. *Biol. Oceanogr.* 6, 375–382. doi: 10.1021/bk-1989-0393.ch011
- Kettles, N. L., Kopriva, S., and Malin, G. (2014). Insights into the regulation of DMSP synthesis in the diatom *Thalassiosira pseudonana* through APR activity, proteomics and gene expression analyses on cells acclimating to changes in salinity, light and nitrogen. *PLoS ONE* 9:e94795. doi: 10.1371/journal.pone.0094795
- Kiema, T. R., Engel, C. K., Schmitz, W., Filppula, S. A., Wierenga, R. K., and Hiltunen, J. K. (1999). Mutagenic and enzymological studies of the hydratase and isomerase activities of 2-enoyl-CoA hydratase-1. *Biochemistry* 38, 2991–2999. doi: 10.1021/bi981646v
- Kiene, R., Linn, L. J., and Bruton, J. A. (2000). New and important roles for DMSP in marine microbial communities. *J. Sea Res.* 43, 209–224. doi: 10.1016/S1385-1101(00)00023-X
- Kiene, R. P. (1998). Uptake of choline and its conversion to glycine betaine by bacteria in estuarine waters. *Appl. Environ. Microbiol.* 64, 1045–1051.
- Kiene, R. P., Linn, L. J., Gonzalez, J., Moran, M. A., and Bruton, J. A. (1999). Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton. *Appl. Environ. Microbiol.* 65, 4549–4558.
- Kiene, R. P., and Williams, L. P. H. (1998). Glycine betaine uptake, retention, and degradation by microorganisms in seawater. *Limnol. Oceanogr.* 43, 1592–1603. doi: 10.4319/lo.1998.43.7.1592
- Kiene, R. P., Williams, L. P. H., and Walker, J. E. (1998). Seawater microorganisms have a high affinity glycine betaine uptake system which also recognizes dimethylsulfoniopropionate. *Aquat. Microb. Ecol.* 15, 39–51. doi: 10.3354/ame015039
- Kirkwood, M., Le Brun, N. E., Todd, J. D., and Johnston, A. W. (2010). The *dddP* gene of *Roseovarius nubinhibens* encodes a novel lyase that cleaves dimethylsulfoniopropionate into acrylate plus dimethyl sulfide. *Microbiology* 156(Pt 6), 1900–1906. doi: 10.1099/mic.0.038927-0
- Kirst, G. O., Thiel, C., Wolff, H., Nothnagel, J., Wanzek, M., and Ulmke, R. (1990). Dimethylsulfoniopropionate (DMSP) in icealgae and its possible biological role. *Mar. Chem.* 35, 381–388. doi: 10.1016/S0304-4203(09)90030-5
- Kocsis, M. G., and Hanson, A. D. (2000). Biochemical evidence for two novel enzymes in the biosynthesis of 3-dimethylsulfoniopropionate in *Spartina alterniflora*. *Plant Physiol.* 123, 1153–1161. doi: 10.1104/pp.123.3.1153
- Kocsis, M. G., Nolte, K. D., Rhodes, D., Shen, T. L., Gage, D. A., and Hanson, A. D. (1998). Dimethylsulfoniopropionate biosynthesis in *Spartina alterniflora*. Evidence that S-methylmethionine and dimethylsulfoniopropylamine are intermediates. *Plant Physiol.* 117, 273–281. doi: 10.1104/pp.117.1.273
- Lee, H. H., Kim, D. J., Ahn, H. J., Ha, J. Y., and Suh, S. W. (2004). Crystal structure of T-protein of the glycine cleavage system. Cofactor binding, insights into H-protein recognition, and molecular basis for understanding nonketotic hyperglycinemia. *J. Biol. Chem.* 279, 50514–50523. doi: 10.1074/jbc.M409672200
- Lesser, M. P. (2006). Oxidative stress in marine environments: biochemistry and physiological ecology. *Annu. Rev. Physiol.* 68, 253–278. doi: 10.1146/annurev.physiol.68.040104.110001
- Lewis, L. A., and McCourt, R. M. (2004). Green algae and the origin of land plants. *Am. J. Bot.* 91, 1535–1556. doi: 10.3732/ajb.91.10.1535
- Li, C., Wei, T., Zhang, S., Chen, X., Gao, X., Wang, P., et al. (2013). Molecular insight into bacterial cleavage of oceanic dimethylsulfoniopropionate into demethyl sulfide. *Proc. Natl. Acad. Sci. U.S.A.* 111, 1026–1031. doi: 10.1073/pnas.1312354111
- Lovelock, J. E., Maggs, R. J., and Rasmussen, R. A. (1972). Atmospheric dimethyl sulfide and natural sulfur cycle. *Nature* 237, 452–453. doi: 10.1038/237452a0
- Luo, H., Csuros, M., Hughes, A. L., and Moran, M. A. (2013). Evolution of divergent life history strategies in marine alphaproteobacteria. *MBio* 4:e00373-13. doi: 10.1128/mBio.00373-13
- Luo, H., and Moran, M. A. (2014). Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol. Mol. Biol. Rev.* 78, 573–587. doi: 10.1128/MMBR.00020-14
- Luo, H., Swan, B. K., Stepanauskas, R., Hughes, A. L., and Moran, M. A. (2014). Evolutionary analysis of a streamlined lineage of surface ocean *Roseobacters*. *ISME J.* 8, 1428–1439. doi: 10.1038/ismej.2013.248
- Lyon, B. R., Lee, P. A., Bennett, J. M., Ditullio, G. R., and Janech, M. G. (2011). Proteomic analysis of a sea-ice diatom: salinity acclimation provides new insight into the dimethylsulfoniopropionate production pathway. *Plant Physiol.* 157, 1926–1941. doi: 10.1104/pp.111.185025
- McCourt, R. M., Delwiche, C. F., and Karol, K. G. (2004). Charophyte algae and land plant origins. *Trends Ecol. Evol.* 19, 661–666. doi: 10.1016/j.tree.2004.09.013
- Moldowan, J. M., Dahl, J., Jacobson, S. R., Huizinga, B. J., Fago, F. J., Shetty, R., et al. (1996). Chemostratigraphic reconstruction of biofacies: molecular evidence linking cyst-forming dinoflagellates with pre-Triassic ancestors. *Geology* 24, 159–162. doi: 10.1130/0091-7613(1996)024<0159:CROBME>2.3.CO;2
- Moldowan, J. M., and Jacobson, S. R. (2000). Chemical signals for early evolution of major taxa: biosignatures and taxon-specific biomarkers. *Int. Geol. Rev.* 42, 805–812. doi: 10.1080/00206810009465112
- Moldowan, J. M., and Talyzina, N. M. (1998). Biogeochemical evidence for dinoflagellate ancestors in the Early Cambrian. *Science* 281, 1168–1170. doi: 10.1126/science.281.5380.1168
- Moran, M. A., Belas, R., Schell, M. A., Gonzalez, J. M., Sun, F., Sun, S., et al. (2007). Ecological genomics of marine *Roseobacters*. *Appl. Environ. Microbiol.* 73, 4559–4569. doi: 10.1128/AEM.02580-06
- Moran, M. A., Reisch, C. R., Kiene, R. P., and Whitman, W. B. (2012). Genomic insights into bacterial DMSP transformations. *Annu. Rev. Mar. Sci.* 4, 523–542. doi: 10.1146/annurev-marine-120710-100827
- Niki, T., Kohata, K., and Otsuki, A. (2000). DMSP-lyase activity in five marine phytoplankton species: its potential importance in DMS production. *Mar. Biol.* 136, 759–764. doi: 10.1007/s002279900235
- Niki, T., Kunugi, M., Kohata, K., and Otsuki, A. (1997). Annual monitoring of DMS-producing bacteria in Tokyo Bay, Japan in relation to DMSP. *Mar. Ecol. Prog. Ser.* 156, 17–24. doi: 10.3354/meps156017
- Noda-Garcia, L., Camacho-Zarco, A., Medina-Ruiz, S., Gaytan, P., Carrillo-Tripp, M., Fulop, V., et al. (2013). Evolution of substrate specificity in a recipient's enzyme following horizontal gene transfer. *Mol. Biol. Evol.* 30, 2024–2034. doi: 10.1093/molbev/mst115
- Palmer, J. D. (2003). The symbiotic birth and spread of plastids: how many time and whodunit. *J. Phycol.* 39, 4–11. doi: 10.1046/j.1529-8817.2003.02185.x
- Paquet, L., Rathinasabapathi, B., Saini, H., Zamir, L., Gage, D. A., Huang, Z. H., et al. (1994). Accumulation of the compatible solute 3-dimethylsulfoniopropionate in sugarcane and its relatives, but not other gramineous crops. *Aust. J. Plant Physiol.* 21, 37–48. doi: 10.1071/PP9940037
- Perona, J. J., and Craik, C. S. (1997). Evolutionary divergence of substrate specificity within the chymotrypsin-like serine protease fold. *J. Biol. Chem.* 272, 29987–29990. doi: 10.1074/jbc.272.48.29987
- Perreux, D., Maraite, H., and Meyer, J. A. (1982). Identification of 3(methylthio) propionic acid as a blight inducing toxin produced by *Xanthomonas campestris* pv. *manihotis* in vitro. *Physiol. Plant Path.* 20, 313–319. doi: 10.1016/0048-4059(82)90056-X
- Quigg, A., Finkel, Z. V., Irwin, A. J., Rosenthal, Y., Ho, T. Y., Reinfelder, J. R., et al. (2003). The evolutionary inheritance of elemental stoichiometry in marine phytoplankton. *Nature* 425, 291–294. doi: 10.1038/nature01953
- Raina, J. B., Tapiolas, D. M., Forest, S., Lutz, A., Abrego, D., Ceh, J., et al. (2013). DMSP biosynthesis by an animal and its role in coral thermal stress response. *Nature* 502, 677–680. doi: 10.1038/nature12677
- Reisch, C. R., Crabb, W. M., Gifford, S. M., Teng, Q., Stoudemayer, M. J., Moran, M. A., et al. (2013). Metabolism of dimethylsulphoniopropionate by *Ruegeria pomeroyi* DSS-3. *Mol. Microbiol.* 89, 774–791. doi: 10.1111/mmi.12314



- Reisch, C. R., Moran, M. A., and Whitman, W. B. (2008). Dimethylsulfoniopropionate-dependent demethylase (DmdA) from *Pelagibacter ubique* and *Silicibacter pomeroyi*. *J. Bacteriol.* 190, 8018–8024. doi: 10.1128/JB.00770-08
- Reisch, C. R., Moran, M. A., and Whitman, W. B. (2011a). Bacterial catabolism of dimethylsulfoniopropionate (DMSP). *Front. Microbiol.* 2:172. doi: 10.3389/fmicb.2011.00172
- Reisch, C. R., Stoudemayer, M. J., Varaljay, V. A., Amster, I. J., Moran, M. A., and Whitman, W. B. (2011b). Novel pathway for assimilation of dimethylsulfoniopropionate widespread in marine bacteria. *Nature* 473, 208–211. doi: 10.1038/nature10078
- Rusch, D. B., Halpern, A. L., Sutton, G., Heidelberg, K. B., Williamson, S., Yooseph, S., et al. (2007). The sorcerer II global ocean sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* 5:e77. doi: 10.1371/journal.pbio.0050077
- Salgado, P., Kiene, R., Wiebe, W., and Magalhaes, C. (2014). Salinity as a regulator of DMSP degradation in *Ruegeria pomeroyi* DSS-3. *J. Microbiol.* 52, 948–954. doi: 10.1007/s12275-014-4409-1
- Sanderson, M. J. (2003). Molecular data from 27 proteins do not support a Precambrian origin of land plants. *Am. J. Bot.* 90, 954–956. doi: 10.3732/ajb.90.6.954
- Schuller, D. J., Reisch, C. R., Moran, M. A., Whitman, W. B., and Lanzilotta, W. N. (2012). Structures of dimethylsulfoniopropionate-dependent demethylase from the marine organism *Pelagibacter ubique*. *Protein Sci.* 21, 289–298. doi: 10.1002/pro.2015
- Sekowska, A., Denervaud, V., Ashida, H., Michoud, K., Haas, D., Yokota, A., et al. (2004). Bacterial variations on the methionine salvage pathway. *BMC Microbiol.* 4:9. doi: 10.1186/1471-2180-4-9
- Simo, R. (2001). Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and evolutionary links. *Trends Ecol. Evol.* 16, 287–294. doi: 10.1016/S0169-5347(01)02152-8
- Simo, R., Archer, S., Pedros-Ali, C., Gilpin, L., and Stelfox-Widdicombe, C. (2002). Coupled dynamics of dimethylsulfoniopropionate and dimethyl sulfide cycling and the microbial food web in surface waters of the North Atlantic. *Limnol. Oceanogr.* 47, 53–61. doi: 10.4319/lo.2002.47.1.0053
- Steele, R. D., and Benevenga, N. J. (1979). The metabolism of 3-methylthiopropionate in rat liver homogenates. *J. Biol. Chem.* 254, 8885–8890.
- Stefels, J. (2000). Physiological aspects of the production and conversion of DMSP in marine algae and higher plants. *J. Sea Res.* 43, 183–197. doi: 10.1016/S1385-1101(00)00030-7
- Stefels, J., Dijkhuizen, L., and Gieskes, W. C. (1995). DMSP-lyase activity in a spring phytoplankton bloom off the Dutch coast, related to *Phaeocystis* sp. abundance. *Mar. Ecol. Prog. Ser.* 123, 235–243. doi: 10.3354/meps123235
- Stefels, J., Steinke, M., Turner, S., Malin, G., and Belviso, S. (2007). Environmental constraints on the production and removal of the climatically active gas dimethylsulphide (DMS) and implications for ecosystem modeling. *Biogeochemistry* 83, 245–275. doi: 10.1007/s10533-007-9091-5
- Stover, L. E., Brinkhuis, H., Damassa, S. P., De Verteuil, L., and Helby, R. J. (1996). “Mesozoic-Tertiary dinoflagellates, acritarchs and prasinophyte,” in *Palynology: Principles and Applications*, eds D. C. Jansonius and J. McGregor (Washington DC: American Association of Stratigraphic Palynologists Foundation), 641–750.
- Sullivan, M. J., Curson, A. R., Shearer, N., Todd, J. D., Green, R. T., and Johnston, A. W. (2011). Unusual regulation of a leaderless operon involved in the catabolism of dimethylsulfoniopropionate in *Rhodobacter sphaeroides*. *PLoS ONE* 6:e15972. doi: 10.1371/journal.pone.0015972
- Summers, P. S., Nolte, K. D., Cooper, A. J. L., Borgeas, H., Leustek, T., Rhodes, D., et al. (1998). Identification and stereospecificity of the first three enzymes of 3-dimethylsulfoniopropionate biosynthesis in a chlorophyte alga. *Plant Physiol.* 116, 369–378. doi: 10.1104/pp.116.1.369
- Sun, J., Todd, J. D., Thrash, J. C., Qian, Y., Qian, M. C., Temperton, B., et al. (2016). The abundant marine bacterium *Pelagibacter* simultaneously catabolizes dimethylsulfoniopropionate to the gases dimethyl sulfide and methanethiol. *Nat. Microbiol.* 1, 16065. doi: 10.1038/nmicrobiol.2016.65
- Sun, Y., Powell, K. E., Sung, W., Lynch, M., Moran, M. A., and Luo, H. (2017). Spontaneous mutations of a model heterotrophic marine bacterium. *ISME J.* doi: 10.1038/ismej.2017.20 [Epub ahead of print].
- Sunda, W., Kieber, D. J., Kiene, R. P., and Huntsman, S. (2002). An antioxidant function for DMSP and DMS in marine algae. *Nature* 418, 317–320. doi: 10.1038/nature00851
- Tan, D., Crabb, W. M., Whitman, W. B., and Tong, L. (2013). Crystal structure of DmdD, a crotonase superfamily enzyme that catalyzes the hydration and hydrolysis of methylthioacryloyl-CoA. *PLoS ONE* 8:e63870. doi: 10.1371/journal.pone.0063870
- Tang, K., Yang, Y., Lin, D., Li, S., Zhou, W., Han, Y., et al. (2016). Genomic, physiologic, and proteomic insights into metabolic versatility in *Roseobacter* clade bacteria isolated from deep-sea water. *Sci. Rep.* 6:35528. doi: 10.1038/srep35528
- Todd, J. D., Curson, A. R., Dupont, C. L., Nicholson, P., and Johnston, A. W. (2009). The *dddP* gene, encoding a novel enzyme that converts dimethylsulfoniopropionate into dimethyl sulfide, is widespread in ocean metagenomes and marine bacteria and also occurs in some Ascomycete fungi. *Environ. Microbiol.* 11, 1376–1385. doi: 10.1111/j.1462-2920.2009.01864.x
- Todd, J. D., Curson, A. R., Kirkwood, M., Sullivan, M. J., Green, R. T., and Johnston, A. W. (2011). DddQ, a novel, cupin-containing, dimethylsulfoniopropionate lyase in marine roseobacters and in uncultured marine bacteria. *Environ. Microbiol.* 13, 427–438. doi: 10.1111/j.1462-2920.2010.02348.x
- Todd, J. D., Curson, A. R., Nikolaidou-Katsaraidou, N., Brearley, C. A., Watmough, N. J., Chan, Y., et al. (2010). Molecular dissection of bacterial acrylate catabolism—unexpected links with dimethylsulfoniopropionate catabolism and dimethyl sulfide production. *Environ. Microbiol.* 12, 327–343. doi: 10.1111/j.1462-2920.2009.02071.x
- Todd, J. D., Curson, A. R., Sullivan, M. J., Kirkwood, M., and Johnston, A. W. (2012a). The *Ruegeria pomeroyi* *acuI* gene has a role in DMSP catabolism and resembles *yhdH* of *Escherichia coli* and other bacteria in conferring resistance to acrylate. *PLoS ONE* 7:e35947. doi: 10.1371/journal.pone.0035947
- Todd, J. D., Kirkwood, M., Newton-Payne, S., and Johnston, A. W. (2012b). DddW, a third DMSP lyase in a model *Roseobacter* marine bacterium, *Ruegeria pomeroyi* DSS-3. *ISME J.* 6, 223–226. doi: 10.1038/ismej.2011.79
- Todd, J. D., Rogers, R., Li, Y. G., Wexler, M., Bond, P. L., Sun, L., et al. (2007). Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria. *Science* 315, 666–669. doi: 10.1126/science.1135370
- Trossat, C., Nolte, K. D., and Hanson, A. D. (1996). Evidence that the pathway of dimethylsulfoniopropionate biosynthesis begins in the cytosol and ends in the chloroplast. *Plant Physiol.* 111, 965–973. doi: 10.1104/pp.111.4.965
- Vail, P. R., Mitchum, R. M., and Thompson, I. I. S. (1977). “Seismic stratigraphy and global changes of sea level, part 4: global cycles of relative changes of sea level,” in *Seismic Stratigraphy-Applications to Hydrocarbon Exploration*, ed. C. E. Payton (Tulsa, OK: American Association of Petroleum Geologists Memoirs), 83–97.
- van der Maarel, M., Van Bergeijk, S., Van Werkhoven, A., Laverman, A., Meijer, W., Stam, W., et al. (1996). Cleavage of dimethylsulfoniopropionate and reduction of acrylate by *Desulfovibrio acrylicus* sp. nov. *Arch. Microbiol.* 166, 109–115. doi: 10.1007/s002030050363
- van Duyl, F. C., Gieskes, W. C., Kop, A. J., and Lewis, W. E. (1998). Biological control of short-term variations in the concentration of DMSP and DMS during *Phaeocystis* spring bloom. *J. Sea Res.* 40, 221–231. doi: 10.1016/S1385-1101(98)00024-0
- Varaljay, V. A., Robidart, J., Preston, C. M., Gifford, S. M., Durham, B. P., Burns, A. S., et al. (2015). Single-taxon field measurements of bacterial gene regulation controlling DMSP fate. *ISME J.* 9, 1692. doi: 10.1038/ismej.2015.94
- Wagner, H., Jakob, T., and Wilhelm, C. (2005). Balancing the energy flow from captured light into biomass under fluctuating light conditions. *New Phytol.* 169, 95–108. doi: 10.1111/j.1469-8137.2005.01550.x
- Wang, P., Chen, X. L., Li, C. Y., Gao, X., Zhu, D. Y., Xie, B. B., et al. (2015). Structural and molecular basis for the novel catalytic mechanism and evolution of DddP, an abundant peptidase-like bacterial dimethylsulfoniopropionate lyase: a new enzyme from an old fold. *Mol. Microbiol.* 98, 289–301. doi: 10.1111/mmi.13119
- Whitfield, M. (2001). Interactions between phytoplankton and trace metals in the ocean. *Adv. Mar. Biol.* 41, 3–128. doi: 10.1016/S0065-2881(01)41002-9
- Wiesemeier, T., and Pohnert, G. (2007). Direct quantification of dimethylsulfoniopropionate (DMSP) in marine micro- and macroalgae

- using HPLC or UPLC/MS. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 850, 493–498. doi: 10.1016/j.jchromb.2006.12.023
- Wolfe, G. V., and Steinke, M. (1997). Grazing-activated chemical defense in a unicellular marine alga. *Nature* 387, 894–897. doi: 10.1038/43168
- Yoch, D. C. (2002). Dimethylsulfoniopropionate: its sources, role in the marine food web, and biological degradation to dimethylsulfide. *Appl. Environ. Microbiol.* 68, 5804–5815. doi: 10.1128/AEM.68.12.5804-5815.2002
- Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G., and Bhattacharya, D. (2004). A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* 21, 809–818. doi: 10.1093/molbev/msh075
- Yost, D. M., and Mitchelmore, C. L. (2009). Dimethylsulfoniopropionate (DMSP) lyase activity in different strains of the symbiotic alga *Symbiodinium microadriaticum*. *Mar. Ecol. Prog. Ser.* 386, 61–70. doi: 10.3354/meps08031
- Zhang, Y., Sun, Y., Jiao, N., Stepanauskas, R., and Luo, H. (2016). Ecological genomics of the uncultivated marine *Roseobacter* lineage CHAB-I-5. *Appl. Environ. Microbiol.* 82, 2100–2111. doi: 10.1128/AEM.03678-15
- Zubkov, M. V., Fuchs, B. M., Archer, S. D., Kiene, R. P., Amann, R., and Burkill, P. H. (2001). Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. *Environ. Microbiol.* 3, 304–311. doi: 10.1046/j.1462-2920.2001.00196.x
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Bullock, Luo and Whitman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# FnrL and Three Dnr Regulators Are Used for the Metabolic Adaptation to Low Oxygen Tension in *Dinoroseobacter shibae*

Matthias Ebert<sup>1</sup>, Sebastian Laaß<sup>2</sup>, Andrea Thürmer<sup>3</sup>, Louisa Roselius<sup>4</sup>, Denitsa Eckweiler<sup>4</sup>, Rolf Daniel<sup>3</sup>, Elisabeth Härtig<sup>1\*</sup> and Dieter Jahn<sup>4</sup>

<sup>1</sup> Institute of Microbiology, Technische Universität Braunschweig, Braunschweig, Germany, <sup>2</sup> Institute for Molecular Biosciences, Goethe-University Frankfurt, Frankfurt, Germany, <sup>3</sup> Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August University Göttingen, Göttingen, Germany, <sup>4</sup> Braunschweig Integrated Centre of Systems Biology, Technische Universität Braunschweig, Braunschweig, Germany

## OPEN ACCESS

### Edited by:

Alison Buchan,  
University of Tennessee, USA

### Reviewed by:

Andrew S. Lang,  
Memorial University of Newfoundland,  
Canada

Yonghui Zeng,  
Aarhus University, Denmark

### \*Correspondence:

Elisabeth Härtig  
e.haertig@tu-bs.de

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 15 January 2017

**Accepted:** 29 March 2017

**Published:** 20 April 2017

### Citation:

Ebert M, Laaß S, Thürmer A,  
Roselius L, Eckweiler D, Daniel R,  
Härtig E and Jahn D (2017) FnrL and  
Three Dnr Regulators Are Used for the  
Metabolic Adaptation to Low Oxygen  
Tension in *Dinoroseobacter shibae*.  
Front. Microbiol. 8:642.  
doi: 10.3389/fmicb.2017.00642

The heterotrophic marine bacterium *Dinoroseobacter shibae* utilizes aerobic respiration and anaerobic denitrification supplemented with aerobic anoxygenic photosynthesis for energy generation. The aerobic to anaerobic transition is controlled by four Fnr/Crp family regulators in a unique cascade-type regulatory network. FnrL is utilizing an oxygen-sensitive Fe-S cluster for oxygen sensing. Active FnrL is inducing most operons encoding the denitrification machinery and the corresponding heme biosynthesis. Activation of gene expression of the high oxygen affinity *cbb<sub>3</sub>*-type and repression of the low affinity *aa<sub>3</sub>*-type cytochrome c oxidase is mediated by FnrL. Five regulator genes including *dnrE* and *dnrF* are directly controlled by FnrL. Multiple genes of the universal stress protein (USP) and cold shock response are further FnrL targets. DnrD, most likely sensing NO via a heme cofactor, co-induces genes of denitrification, heme biosynthesis, and the regulator genes *dnrE* and *dnrF*. DnrE is controlling genes for a putative Na<sup>+</sup>/H<sup>+</sup> antiporter, indicating a potential role of a Na<sup>+</sup> gradient under anaerobic conditions. The formation of the electron donating primary dehydrogenases is coordinated by FnrL and DnrE. Many plasmid encoded genes were DnrE regulated. DnrF is controlling directly two regulator genes including the Fe-S cluster biosynthesis regulator *iscR*, genes of the electron transport chain and the glutathione metabolism. The genes for nitrate reductase and CO dehydrogenase are repressed by DnrD and DnrF. Both regulators in concert with FnrL are inducing the photosynthesis genes. One of the major denitrification operon control regions, the intergenic region between *nirS* and *nosR2*, contains one Fnr/Dnr binding site. Using regulator gene mutant strains, *lacZ*-reporter gene fusions in combination with promoter mutagenesis, the function of the single Fnr/Dnr binding site for FnrL-, DnrD-, and partly DnrF-dependent *nirS* and *nosR2* transcriptional activation was shown. Overall, the unique regulatory network of the marine bacterium *D. shibae* for the transition from aerobic to anaerobic growth composed of four Crp/Fnr family regulators was elucidated.

**Keywords:** anaerobic energy metabolism, *Dinoroseobacter shibae*, regulation, oxygen-dependent gene expression, denitrification, FnrL, Dnr, Crp/Fnr regulator

## INTRODUCTION

The heterotrophic Alphaproteobacterium *Dinoroseobacter shibae* DFL12<sup>T</sup> is a member of the *Roseobacter* group, which are highly abundant in the marine ecosystem and possess a large metabolic diversity (Buchan et al., 2005; Wagner-Döbler and Biebl, 2006; Simon et al., 2017). *D. shibae* DFL12<sup>T</sup> utilizes carbon sources usually via the Entner-Doudoroff-pathway instead of standard glycolysis (Fürch et al., 2009). The marine bacterium is able to perform aerobic anoxygenic photophosphorylation to gain additional energy (Biebl et al., 2005). Furthermore, anaerobic growth of *D. shibae* DFL12<sup>T</sup> using nitrate as terminal electron acceptor was proposed (Wagner-Döbler et al., 2010). It was shown, that upon depletion of the electron acceptor oxygen *D. shibae* DFL12<sup>T</sup> establishes the whole process of denitrification with the reduction of nitrate via nitrite, nitric oxide, nitrous oxide to dinitrogen (Laass et al., 2014). The corresponding denitrification gene cluster comprises 39 genes organized in six operons. A fine-tuned regulatory network was expected for the gene expression control in response to low oxygen tension (Zumft, 1997). Within the denitrification gene cluster two genes, Dshi\_3189 and Dshi\_3191, encoding members of the Crp/Fnr family of transcription factors were found. Within the whole genome of *D. shibae* a total of seven genes encoding Crp/Fnr-like regulators were identified (Wagner-Döbler et al., 2010).

The superfamily of Crp/Fnr-like transcription factors are known to respond to a broad spectrum of intracellular and exogenous stimuli (Körner et al., 2003). Despite their low amino acid sequence identity of 25% this group of transcription factors shares common structural features. Known Crp/Fnr family proteins usually consist of two functionally distinct domains, a DNA binding helix-turn-helix motif and an N-terminal region of multiple antiparallel  $\beta$ -strands forming the sensory domain (Schultz et al., 1991). The sensing regions are individually adapted for the detection of highly different signals (Green et al., 2001). *Escherichia coli* Crp reversibly binds cAMP to monitor the glucose status of the cell (Crothers and Steitz, 1992). Fnr, the global oxygen-dependent transcriptional regulator for fumarate and nitrate reduction was firstly characterized for *E. coli* (Spiro and Guest, 1990; Khoroshilova et al., 1997). The sensory region of *E. coli* Fnr possesses four conserved cysteine residues (C20, C23, C29, and C122) at its N-terminus that mediate *in vivo* activity via ligation of an oxygen-sensitive iron-sulfur cluster (Trageser and Unden, 1989; Kiley and Reznikoff, 1991; Green and Guest, 1993; Kiley and Beinert, 1998). An intact iron-sulfur cluster is necessary for dimerization of the regulator and subsequent binding to the palindromic Fnr binding site TTTGAT-N<sub>4</sub>-ATCAA (Green et al., 1996; Lazazzera et al., 1996). *E. coli* Fnr in cooperation with the nitrate/nitrite responsive two component regulatory systems NarX/L and NarP/Q and the redox regulator ArcA/B is controlling the complex network of multiple anaerobic primary dehydrogenases, terminal oxidases and mixed acid fermentation processes (Jahn and Jahn, 2012; Tielen et al., 2012). In *Bacillus subtilis* Fnr in combination with the two component redox responsive system ResDE, the nitric oxide sensor NsrR, the Rex regulator responding to changes in the cellular NAD<sup>+</sup>/NADH ratio and the acetate sensor AlsR are regulating the fine-tuned

regulatory network of the anaerobic energy metabolism (Härtig and Jahn, 2012). For *Pseudomonas aeruginosa* regulation of denitrification genes was found dependent on Anr, a homolog of Fnr, a second Crp/Fnr-family regulator Dnr, which senses NO and the nitrate responsive NarX/L system (Schreiber et al., 2007; Rinaldo et al., 2012). All three regulators induce the expression of the nitrate reductase genes *narGHJI* (Schreiber et al., 2007). Dnr activates the expression of residual denitrification genes including *nirS*, *nirQ*, *norC*, and *nosR* (Arai, 2003). Active Anr binds to a conserved palindromic sequence within the target promoters known as Anr box, which is similar to the Fnr binding site. The Dnr binding site is indistinguishable from the Anr box (Rompf et al., 1998). A fine-tuned interplay of four Crp/Fnr-like regulators, FnrA, DnrD, DnrE, and DnrS with NarX/L is required for the regulating the denitrification genes of *Pseudomonas stutzeri* (Härtig and Zumft, 1999; Vollack et al., 1999; Vollack and Zumft, 2001). In *Rhodobacteraceae* regulators including RegA/B, FnrL, AppA/PpsR, and PrrBA are responsible for the coordination of the anaerobic metabolism (Wu and Bauer, 2008; Winkler et al., 2013; Kumka and Bauer, 2015).

For *D. shibae* time resolved transcriptome and proteome analyses of a continuous culture that was shifted from aerobic to nitrate respiratory conditions revealed the induced expression of four potential Crp/Fnr-like genes Dshi\_0660, Dshi\_3189, Dshi\_3191, and Dshi\_3270 (Laass et al., 2014). Due to their potential involvement in the aerobic to anaerobic transition process, they were annotated as FnrL, DnrD, DnrE, and DnrF, respectively. However, only FnrL contains cysteine residues potentially involved in iron-sulfur cluster formation. No genes for other typical redox sensing or nitrate/nitrite responsive regulators were found in the *D. shibae* genome. Here, we elucidated the regulatory network controlled by these four Crp/Fnr-family regulators. Their regulons were defined using regulatory mutants and transcriptome analyses. Promoter activities crucial for the onset of denitrification in *D. shibae* were further investigated *in vivo* using promoter reporter gene fusions. The iron-sulfur cluster of FnrL was demonstrated for the recombinant purified protein. A novel cascade type regulatory scenario for the aerobic-anaerobic transition coordinated by four Crp/Fnr regulators was observed.

## MATERIALS AND METHODS

### Bacterial Strains and Growth Conditions

The type strain *D. shibae* DFL12<sup>T</sup> and corresponding mutant strains were grown routinely aerobically in Marine-Bouillon (MB, Roth, Karlsruhe, Germany) at 30°C in bottle flasks shaking at 200 rpm in the dark or on the same medium solidified with 1.5% agar. *E. coli* strains were routinely grown at 37°C and shaking at 200 rpm in Lysogenic Broth (LB) supplemented with the appropriate antibiotics and amino acids (Table 1). The growth behavior of the *D. shibae* strains was analyzed under aerobic and anaerobic conditions in artificial seawater medium (SWM; Tomasch et al., 2011) supplemented with 16.9 mM succinate in bottle flasks shaking at 200 rpm for aerobic growth. For anaerobic cultivation NaNO<sub>3</sub> to the final concentration of



**TABLE 1 | Strains and plasmids used in this study.**

Strain	Description	Sources/References
<b><i>D. shibae</i> STRAINS</b>		
DFL12 <sup>T</sup>	wild type	Biebl et al., 2005
DS001	$\Delta fnrL::aacC1$ ( <i>Gm</i> <sup>r</sup> )	This study
DS005	$\Delta fnrL::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pRhokSfnrL, <i>Cm</i> <sup>r</sup>	This study
DS002	$\Delta dnrD::aacC1$ ( <i>Gm</i> <sup>r</sup> ); GFP	This study
DS006	$\Delta dnrD::aacC1$ ( <i>Gm</i> <sup>r</sup> ); GFP; pRhokSdnrD, <i>Cm</i> <sup>r</sup>	This study
DS003	$\Delta dnrE::aacC1$ ( <i>Gm</i> <sup>r</sup> )	This study
DS007	$\Delta dnrE::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pRhokSdnrE, <i>Cm</i> <sup>r</sup>	This study
DS004	$\Delta dnrF::aacC1$ ( <i>Gm</i> <sup>r</sup> )	This study
DS008	$\Delta dnrF::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pRhokSdnrF, <i>Cm</i> <sup>r</sup>	This study
DS100	DFL12 <sup>T</sup> ; pBBR1nirS-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS101	$\Delta fnrL::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pBBR1nirS-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS102	$\Delta dnrD::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pBBR1nirS-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS103	$\Delta dnrE::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pBBR1nirS-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS104	$\Delta dnrF::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pBBR1nirS-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS126	DFL12 <sup>T</sup> ; pBBR1nirS(mu)-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS127	$\Delta fnrL::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pBBR1nosR2-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS128	$\Delta dnrD::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pBBR1nosR2-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS129	$\Delta dnrE::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pBBR1nosR2-lacZ, <i>Cm</i> <sup>r</sup>	This study
DS130	$\Delta dnrF::aacC1$ ( <i>Gm</i> <sup>r</sup> ); pBBR1nosR2-lacZ, <i>Cm</i> <sup>r</sup>	This study
<b><i>E. coli</i> STRAINS</b>		
DH10b	<i>F</i> -endA1 recA1 galE15 galK16 nupG rpsL $\Delta lacX74$ $\Phi 80lacZ\Delta M15$ araD139 $\Delta(ara, leu)7697$ <i>mcrA</i> $\Delta(mrr-hsdRMS-mcrBC)$ $\lambda$ -	Invitrogen
ST18	<i>E. coli</i> S17-1 $\Delta hemA$ <i>thi pro hsdR</i> -M- chromosomal integrated [RP4-2 Tc::Mu:Kmr::Tn7, <i>Tra</i> + <i>Trir Strr</i> ]	Thoma and Schobert, 2009
BL21(DE3) pLysS	<i>F</i> -ompT gal dcm lon hsdSB(rB- mB-) $\lambda$ (DE3) pLysS( <i>cm</i> <sup>r</sup> )	Stratagene
<b>PLASMIDS</b>		
pEX18Tc	Tc <sup>r</sup> ; oriT <sup>+</sup> , sacB <sup>+</sup> , <i>lacZ</i> ,	Hoang et al., 1998
pEX18Ap	<i>Amp</i> <sup>r</sup> ; oriT <sup>+</sup> , sacB <sup>+</sup> , <i>lacZ</i> ,	Hoang et al., 1998
pEX18Gm	<i>Gm</i> <sup>r</sup> ; oriT <sup>+</sup> , sacB <sup>+</sup> , <i>lacZ</i> ,	Hoang et al., 1998
pBBR1MCS	<i>Cm</i> <sup>r</sup> ; <i>lacZ</i> P <sub>lac</sub> P <sub>T7</sub> rep	Kovach et al., 1994
pBBR1MCS-5	<i>Gm</i> <sup>r</sup> ; <i>lacZ</i> P <sub>lac</sub> P <sub>T7</sub> rep	Kovach et al., 1995
pPS858	<i>Amp</i> <sup>r</sup> ; <i>Gm</i> <sup>r</sup> ; GFP	Hoang et al., 1998
pEX18 $\Delta fnrL$	Tc <sup>r</sup> ; <i>fnrL::aacC1</i>	This study
pEX18 $\Delta dnrD$	Tc <sup>r</sup> ; <i>dnrD::aacC1</i>	This study
pEX18 $\Delta dnrE$	Tc <sup>r</sup> ; <i>dnrE::aacC1</i>	This study
pEX18 $\Delta dnrF$	<i>Amp</i> <sup>r</sup> ; <i>dnrF::aacC1</i>	This study
pRhokSfnrL	Chl <sup>r</sup> ; <i>fnrL</i>	This study
pRhokSdnrD	Chl <sup>r</sup> ; <i>dnrD</i>	This study
pRhokSdnrE	Chl <sup>r</sup> ; <i>dnrE</i>	This study
pRhokSdnrF	Chl <sup>r</sup> ; <i>dnrF</i>	This study
pMA-T	<i>Amp</i> <sup>r</sup> ; Col E1 origin	Thermo Fisher
13AADNLP_seq3_dnrD_pMA-T	<i>Amp</i> <sup>r</sup> ; Col E1 origin; <i>dnrD</i>	Thermo Fisher
pET52FnrL	<i>Amp</i> <sup>r</sup> ; <i>fnrL</i>	This study
pBBRLIC-lacZ	<i>Cm</i> <sup>r</sup> , LIC-lacZ	This study
mini-CTX-lacZ	Tc <sup>r</sup> , Col E1 mini CTX vector carrying the <i>lacZ</i> fusion	Becher and Schweizer, 2000
pBBRnirS-lacZ	<i>Cm</i> <sup>r</sup> , <i>nirS-lacZ</i>	This study
pBBRnirS(mu)-lacZ	<i>Cm</i> <sup>r</sup> , <i>nirS(mu)-lacZ</i>	This study
pBBRnosR2-lacZ	<i>Cm</i> <sup>r</sup> , <i>nosR2-lacZ</i>	This study

25 mM was added and incubation was performed in serum flasks sealed with rubber stoppers shaking at 100 rpm (Ebert et al., 2013).

## Construction of Vectors for Recombinant FnrL Production

The *fnrL* gene (Dshi\_0660) was PCR amplified using primers oPT229 and oPT230 (Table 2) containing *SmaI* and *Eco53kI* restriction sites from *D. shibae* genomic DNA. The amplification product and the pET52b vector (Novagen, Darmstadt, Germany) were both digested using *SmaI* and *Eco53kI*. Ligation of both DNA fragments resulted in the plasmid pET52FnrL.

## Production and Purification of Recombinant FnrL

For production of heterologous FnrL *E. coli* BL21(DE3) pLysS strain carrying the pET52FnrL vector was grown in 500 ml LB-medium containing 100 µg/ml of ampicillin at 37°C and 200 rpm in a 1,000 ml flask. The medium was inoculated to a starting OD<sub>578 nm</sub> of 0.05 with a corresponding overnight culture. After reaching an OD<sub>578 nm</sub> of 0.5–0.6, isopropyl-β-D-thiogalactopyranosid (IPTG) was added to a final concentration of 50 µM to induce protein production. The cultures were shifted to 17°C and 100 rpm for 16 h. Next the cultures were shifted to anaerobiosis and incubated for 2 h. All of the following procedures were performed under strict anaerobic conditions. A cell pellet was obtained by centrifugation of the culture for 15 min at 4,000 × g at 4°C. The cell pellet was resuspended in 10 ml binding buffer (100 mM Tris-HCl pH 7.5, 150 mM NaCl). For cell disruption, a French press (1,200 p.s.i.) was used and a soluble protein fraction was obtained by ultracentrifugation (40,000 × g, 65 min, 4°C). The supernatant was loaded onto a 1 ml *Strep-Tactin*® Superflow® high capacity column (IBA GmbH, Goettingen, Germany). The column was washed two times with 10 ml of washing buffer (100 mM Tris-HCl pH 7.5, 150 mM NaCl). The bound proteins were eluted with 10 ml of elution buffer (100 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM desthiobiotin). The purified recombinant proteins were stored at 17°C under oxygen exclusion. Protein fractions were analyzed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE; Laemmli, 1970; Righeiti, 1990). Protein concentrations were determined using the Bradford Reagent (Sigma-Aldrich, St. Louis, USA) according to the manufacturer's instructions.

## Reconstitution of Iron-Sulfur Clusters in FnrL

The chemical reconstitution of iron-sulfur clusters in FnrL was performed with a final concentration of 40 µM protein solution under strictly anaerobic conditions in a Coy anaerobic chamber (Coy, Grass Lake, USA). Initially, 10 mM DTT were added to the protein solution and incubated for 1 h at 17°C. Afterwards, 1 mM ammonium iron citrate was slowly added under careful mixing. After 5 min of incubation 1 mM lithium sulfide was added. The reaction was stopped after reaching a brownish color (~15 min) by 5 min centrifugation at 14,000 × g. The reconstituted protein was purified using a NAP column (GE Healthcare, Solingen,

Germany). Spectroscopic analysis was performed using a UV/Vis Spectroscope V-550 (Jasco, Gross-Umstadt, Germany).

## Construction of *D. shibae* Deletion Mutants

To obtain deletion strains of *fnrL* (Dshi\_0660), *dnrD* (Dshi\_3189), *dnrE* (Dshi\_3191), and *dnrF* (Dshi\_3270) three different cloning strategies were used. For deletion of *dnrD* a synthetic insert was designed and synthesized by GeneArt (Thermo Fisher Scientific, Waltham, USA). The insert harbored 500 bp upstream sequences of *dnrD* and 500 bp downstream sequences. The *dnrD* gene was replaced on the insert by the gentamicin resistance cassette of plasmid pEX18Gm (Hoang et al., 1998). This synthetic insert was cloned by flanking *KpnI* and *HindIII* restriction sites into the equally cutted pMA-T vector (GeneArt, Thermo Fisher Scientific, Waltham, USA) resulting in vector 13AADNLP\_seq3\_dnrD\_pMA-T (GeneArt, Thermo Fisher Scientific, Waltham, USA). Due to a lack in functionality of the ordered gentamicin resistance cassette, the gene was excised again by *SacI* cleavage and replaced by the *aacC1* gene (*Gm<sup>r</sup>*) and the *gfp* gene from vector pPS858 (Hoang et al., 1998). The newly created insert was cleaved by *KpnI* and *HindIII* and ligated into the suicide vector pEX18Tc resulting in the final pEX18Δ*dnrD* vector. For construction of the *fnrL* mutant strain the CloneEZ® PCR cloning Kit (Genscript, New Jersey, USA) was used. A 2715 bp fragment containing the *fnrL* gene together with 993 bp upstream and 973 bp downstream sequences was amplified from genomic DNA of *D. shibae* using primers oPT204 and oPT205 (Table 2). This PCR fragment was cloned into a pEX18Tc vector cleaved by *Eco53kI* by blunt end ligation. The primers oPT206 and oPT207 were used to amplify the entire vector, lacking the full length open reading frame of *fnrL*. Next, the gentamicin resistance cassette from plasmid pBBR1MCS-5 (Kovach et al., 1995) was PCR amplified using primers oPT208 and oPT209. The obtained insert and vector backbone was assembled by using the CloneEZ PCR Cloning Kit (GenScript, Piscataway, USA) according to the manufacturer's instructions resulting in pEX18Δ*fnrL*. The construction of the *dnrE* knockout vector was performed by using the same experimental procedure and resulted in vector pEX18Δ*dnrE*. Using primers oPT216 and oPT217 the full length *dnrE* gene together with 1,178 bp upstream and 1,136 bp downstream sequences of *dnrE* were amplified. For vector amplification primers oPT218 and oPT219 were used. The gentamicin resistance cassette was amplified from pBBR1MCS-5 (Kovach et al., 1995) using oPT220 and oPT221. To obtain a *dnrF* mutant strain of *D. shibae* a *SacI* digested *aacC1* (*gm<sup>r</sup>*) gene of pPS858 was ligated between two PCR fragments of the upstream and downstream region of *dnrF*. Using primers oPT146, which contained a *KpnI* restriction site at the 5' end and oPT147 with a *SacI* restriction site, a 686 bp upstream fragment was obtained. The 541 bp downstream fragment of *dnrF* was amplified using oPT148 with a *SacI* restriction site and oPT149 which contained a *HindIII* restriction site. The resulting insert was cloned into a pEX18Ap vector (Hoang et al., 1998) resulting in vector pEX18Δ*dnrF*. The obtained suicide vectors for *fnrL*, *dnrD*, *dnrE*, and *dnrF* were transferred into *D. shibae* DFL12<sup>T</sup> via biparental mating using the *E. coli* ST18 donor strain as described previously (Thoma and Schobert, 2009; Ebert et al.,

**TABLE 2 | Primers used in this study.**

Name	Sequence (5'–3') <sup>a</sup>
oPT204	TCCGTGTAGCCCTGGAAATT
oPT205	ATGTGCTTTTCCTTGCGGTT
oPT206	TGGCCGGCATTGGATTGATC
oPT207	TCTCAGTTTGTCCCGTAGCG
oPT208	ATCCAATGCCGGCCAGACGCACACCGTGAAA
oPT209	GGGAACAACTGAGAGCGGCGTTGTGACAATTT
oPT222	GAAGTGCCTTCACCTTCGAG
oPT223	CGCACAGACATGCGTACTTT
oPT224	CAATTCGTTCAAGCCGAGAT
oPT239	CAGAGAGCTCATTTAGCCAGCAGGGTATC
oPT240	TGCACTGCAGGTGACGCTACGGGAACAAAC
oPT241	AGACCGACGAGATGGAATTG
oPT242	GGGGTGGTGTTTCATCATCTG
oPT227	GTTTCATGCGACGATTAAGCA
oPT228	CTCTACATGGTTGGCTGACG
oPT169	GCGCTGCGGCCCTACTGGAGCTGAATTGGGGATCTTGAAGTTC
oPT170	CGAGCGTCAGCCAACCGAGCTCGAATTAGCTTC
EH645	CAGAGAGCTCGGCGACCCGGTGAAGTGC
EH646	TGCACTGCAGCTCGCGGGCGCTACATCA
oPT235	TGTATCGAATCACCCCAAT
oPT236	CTTGTCATAGGGCAGCATCA
oPT216	AGCGTGTCCGAGATGAACTT
oPT217	GATCACCATCACCGTCACC
oPT218	CTTGTCATCCGGGTGAA
oPT219	ATTTGCCGACAACCGTGT
oPT220	CCCGGATGGACAAGGACGCACACCGTGAAA
oPT221	CGGTTGTCGGCAAATGCGGCGTTGTGACAATTT
DnrEFW	CAGAGAGCTCCTAGGACATCGGGGTGCTG
DnrEBW	TGCACTGCAGTGCGACGAAAAGGCTCAG
oPT176	CCAAGGCTAAAACCGCACT
oPT177	GAACGAACACGGGCTCAC
oPT146	GACGGTACCCCTGATGACGGGCGAGAAC
oPT147	GACGAGCTCCATCCTTGATCTGTGCCA
oPT148	GACGAGCTCGGAGAGGCACATGACACCT
oPT149	GACAAGCTTTGTAAGGCCGGTGACGATG
DnrFFW	CAAGAGAGCTCGTGGTGTCCCATCCACAAA
DnrFBW	TGCACTGCAGCCGCGCCAAACTAGCATC
oPT237	AGGGTCTTGTGCGAATGATG
oPT238	TCGGCAAGCTCAGGACTATT
oPT229	TCCCGGGGATGCCGGCCACCAG
oPT230	TCGAGCTCTCAGACCAGCGGGCC
EH635	GCTCTAGATTTTACCGCGGGCTTTCCCGGAAGGAGGAAGTACTATGACCATGATTACGGATTTC
EH636	CCGCTCGAGTTTCTTACGCGAAATACG
EH675	CCGCGGGCTTTCCAGCCTCAGGATGTCTCCATTT
EH676	GTTCTCTCTCCACCTGCATGGTGCAGAGTAT

<sup>a</sup>Restriction enzyme sites are underlined.

2013). The gene knockout was obtained by double-homologous recombination and was selected on half-concentrated marine agar plates containing 80 µg/ml gentamicin. The genomic structure of the mutant strains DS001( $\Delta fmrL$ ), DS002( $\Delta dnrD$ ),

DS003( $\Delta dnrE$ ), and DS004( $\Delta dnrF$ ) constructed in this study were confirmed by PCR (Table 2).

For complementation, the *fmrL* gene together with 324 bp upstream sequences was PCR amplified using genomic DNA of

*D. shibae* and the primers oPT239 and oPT240 containing a *SacI* and a *PstI* restriction site, respectively. The insert was ligated into the multiple cloning site of a *SacI* and *PstI* digested pRhokS vector (Katzke et al., 2010), resulting in vector pRhokS*fnrL*. The same procedure was used for construction of complementation vectors of *dnrD*, *dnrE*, and *dnrF*. For construction of the *dnrD* expression plasmid primers EH645 and EH646 were used to amplify a 1,022 bp DNA fragment containing the *dnrD* gene and 301 bp upstream sequences resulting in vector pRhokS*dnrD*. For the *dnrE* vector primers DnrEFW and DnrEBW were used to amplify a 1,050 bp DNA fragment which contained the *dnrE* gene together with 334 bp upstream sequences resulting in vector pRhokS*dnrE*. For the *dnrF* plasmid primers DnrFFW and DnrFBW were used to amplify a 995 bp insert containing the *dnrF* gene together with 232 bp upstream sequences resulting in vector pRhokS*dnrF*. Complementation vectors were transferred into the respective mutant strains via biparental mating using the *E. coli* ST18 donor strain. Selection was performed on half-concentrated marine agar plates containing 15 µg/ml chloramphenicol. The resulting complemented *D. shibae* strains were termed DS005( $\Delta$ *fnrL*, pRhokS*fnrL*), DS006( $\Delta$ *dnrD*, pRhokS*dnrD*), DS007( $\Delta$ *dnrE*, pRhokS*dnrE*), and DS008( $\Delta$ *dnrF*, pRhokS*dnrF*; Table 1).

## Growth Curve and Shift Experiments

To record growth curves a pre-culture of the *D. shibae* strain of interest was inoculated in SWM and grown overnight at 30°C and 200 rpm in dark. Then 125 ml main culture was inoculated to a final OD<sub>578nm</sub> of 0.05 in a 1,000 ml baffled flask. For anaerobic growth the culture was inoculated to a final OD<sub>578 nm</sub> of 0.3 in a serum flask sealed with a rubber stopper. For shift experiments the main culture was inoculated to a final OD<sub>578 nm</sub> of 0.05 in SWM in a shaking flask. After reaching on OD<sub>578 nm</sub> of ~0.5 the cultures were shifted to anaerobic conditions and 25 mM (f.c.) NaNO<sub>3</sub> was supplied. For anaerobic cultivation a serum flask was used. Oxygen tension was measured every 5 min using a PreSense Fibox 3 LCD trace v7 and an oxygen sensor Type Pst3 with an accuracy of ±0.15%. Samples were taken for RNA preparation directly before the shift, and after 30 and 60 min of anaerobic conditions.

## RNA Isolation

For RNA isolation a volume of 2 ml cell culture was removed from the culture and transferred in 4 ml of RNeasy Protect (Qiagen, Hilden, Germany). The solution was roughly mixed and incubated for 5 min at room temperature. Cells were collected by centrifugation at 4,000 × g for 10 min at 4°C, washed with 1 ml water and centrifuged at 4,000 × g for 5 min. The resulting cell pellet was frozen and stored at -80°C until further processing. Cells were enzymatically lysed (15 mg/ml lysozyme) and mechanically disrupted by using 70–150 µm glass beads and a vortex genie 2 (Scientific industries, Inc., New York, USA) for 30 min. Samples were centrifuged at 4,000 × g for 3 min at 4°C and the supernatant was transferred to fresh tube followed by the addition of 1:3 (v/v) 100% EtOH. Samples were loaded onto a RNeasy Mini Spin column (RNeasy mini kit, Qiagen, Hilden, Germany). Genomic DNA was removed using an on column

DNase I treatments for 15 min. Residual DNase I was removed by following the RNeasy mini kit manufacture guidelines. The eluted RNA fraction was treated a second time with DNase I for 15 min and purified by an additional RNeasy Mini Spin column according to the manufacture protocol (RNeasy mini kit, Qiagen, Hilden, Germany).

## RNA Microarray

Two micrograms of total RNA was labeled with Cy3 using the ULS-system (Kreatech, Amsterdam, Netherlands) according to the manufacturer's manual. The labeled RNA (600 ng) was fragmented and hybridized to the custom designed gene specific DNA microarray (Agilent, Santa Clara, USA 8 × 15K format) according to Agilent's one-color microarray protocol. Microarray analyses were performed with three technical and three biological replicates. Only genes with a logarithmic change of ≥0.8 when comparing aerobic (0 min) and anaerobic (60 min) expression levels of wild type and mutant strains with a *P* < 0.05 were considered in subsequent analyses. Generated data have been deposited in NCBI's Gene Expression Omnibus (Edgar, 2002) and are accessible through Geo Series accession number GSE93652.

## RNA Sequencing

RNA sequencing was performed by the Goettingen genomics laboratory (G2L). Library construction started with total *D. shibae* RNA. The rRNA was depleted with the Ribo-Zero rRNA Removal Kit (Illumina, San Diego, USA). Strand-directed RNA-seq libraries were generated using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (New England Biolabs). Sequencing was done with MiSeq Reagent Kit v3 (150 cycle) chemistry in a 2 × 75 bp PE-run using the MiSeq Instrument (Illumina, San Diego, USA).

## Promoter-*lacZ* Reporter Gene Fusions

We developed the vector pBBRLIC-*lacZ* to generate *lacZ* transcriptional fusions via ligation-independent cloning (LIC). Primers EH635 (containing the LIC sequence) and EH636 were used to amplify a 3,168 bp DNA fragment from the plasmid mini-CTX-*lacZ* harboring the *lacZ* gene (Becher and Schweizer, 2000). By this approach the *lacZ* gene was fused with the LIC sequence 5'-TTTTACCGCGGGCTTTCCCGGGAAGGAGGAAGT-3' (Botella et al., 2010) containing a central *SmaI* restriction site (bold letters) and a ribosomal binding site (underlined). The amplification product and plasmid pBBRL1MCS (Kovach et al., 1994) were both digested with *XhoI* and *XbaI*. Ligation of both DNA fragments resulted in plasmid pBBRLIC-*lacZ*. Promoter fragments suitable for ligation-independent cloning (LIC) were generated by amplification of chromosomal DNA fragments using primers with a 5'-CCGCGGGCTTTCCAGC-3' LIC tail sequence added to the forward primer and a 5'-GTTCTCCTTCCCACC-3' LIC tail sequence added to the reverse primer. The pBBRLIC-*lacZ* plasmid was linearized using *SmaI*, gel-purified and treated with T4 DNA polymerase in the presence of 2.5 mM dATP for 20 min at 22°C, followed by 30 min at 75°C to inactivate the enzyme. This generated 12 nt single-stranded overhangs on either side of the *SmaI*



restriction site. For ligation independent cloning 0.2 pmol of each amplified promoter fragment was incubated with 2.5 mM dTTP and T4 DNA polymerase. Promoter fragments were then treated with T4 polymerase in the presence of dTTP for 20 min at 22°C followed by enzyme inactivation for 30 min at 75°C. This approach generated inserts with single-stranded ends that are complementary to those of the treated vector. Treated plasmids (5 ng) and inserts (15 ng) were mixed, annealed at room temperature for 10 min and transformed into *E. coli*. The promoter region of *nosR2* was amplified using primers pair EH675 and EH676 and cloned as described above into pBBRLIC-*lacZ* resulting in the pBBR*nosR2-lacZ* plasmid. A 118 bp *nirS* DNA fragment corresponding to promoter sequences from position -88 to +30 with respect to the transcriptional start site of *nirS* together with the LIC tail sequences was ordered as synthetic double stranded GeneArt string (Thermo Fisher Scientific Inc., Waltham, USA) and cloned into pBBRLIC-*lacZ* resulting in plasmid pBBR*nirS-lacZ*. For *nirS(mu)* the DNA fragment as described for *nirS* was used, however, carrying base exchanges at position -48/-47 from TT to GC and at position -36/-35 from AA to GC. It was ordered as synthetic double stranded GeneArt string (Thermo Fisher Scientific Inc., Waltham, USA). Cloning into pBBRLIC-*lacZ* resulted in plasmid pBBR*nirS(mu)-lacZ*.

## β-Galactosidase Assay

For β-galactosidase assays, *D. shibae* cells were grown in salt water medium and harvested after 16 h in the mid-exponential growth phase. β-Galactosidase assays were performed as described previously (Miller, 1992; Härtig et al., 2004).

## RESULTS

### Classification of Fnr- and Dnr-Type Regulators of *D. shibae*

Previous investigations identified two different major life styles for the marine bacterium *D. shibae*. Under aerobic growth conditions oxygen-dependent respiration is combined with anoxygenic photosynthesis, while anaerobic growth utilizes denitrification as major path of energy generation (Laass et al., 2014). In many other bacteria Fnr- and Dnr-type regulators were found to coordinate the transition between these two life styles at the transcriptional level (Härtig and Jahn, 2012; Tielen et al., 2012). Initial database searches using the amino acid sequences of *Rhodobacter capsulatus* FnrL and *Pseudomonas aeruginosa* Dnr identified seven open reading frames (Dshi\_0660, Dshi\_0447, Dshi\_2521, Dshi\_2528, Dshi\_3189, Dshi\_3191, Dshi\_3270) encoding potential Crp/Fnr-family regulators.

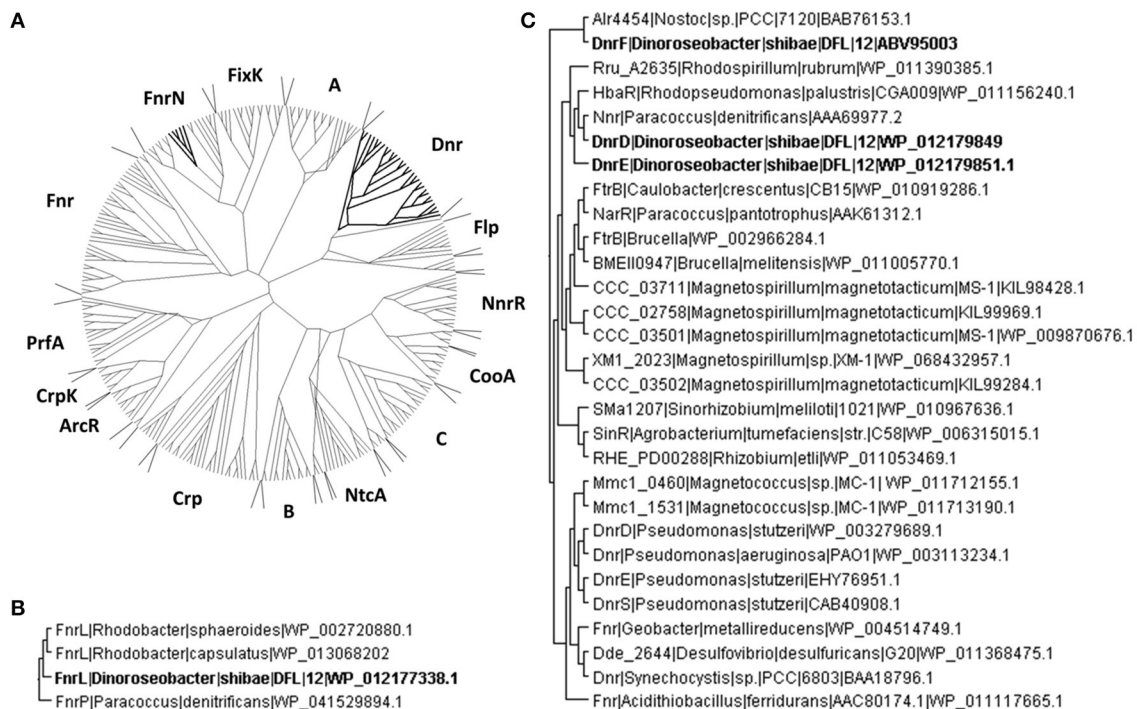
In order to group these Crp/Fnr regulators according to the classification developed by Körner et al. (2003) comprehensive amino acid sequence alignments were performed by using the multiple sequence alignment tool T-coffee (Notredame et al., 2000). A dataset of 287 pair-wise aligned sequences was generated (Figure 1A). The protein encoded by Dshi\_0660 affiliated with the FnrN subgroup (Figure 1B). The highest score of 70.16% identity was found for the transcriptional activator FnrL of *R. capsulatus* (WP\_013068202). Consequently,

the protein was named *D. shibae* FnrL. Moreover, a sequence identity of 64.66% was obtained for FnrP of *Paracoccus denitrificans* (WP\_041529894.1). The predicted transcription regulator encoded by Dshi\_0447 was named DnrA. It was found affiliated with the A subgroup of Crp/Fnr regulators with 25% overall amino acid sequence identity (Figure 1A). However, no experimental evidence for the regulation of the metabolic adaptation to anaerobiosis was obtained during this investigation. Similarly, the postulated Crp/Fnr family regulators encoded by Dshi\_2521 and Dshi\_2528, termed DnrB and DnrC respectively, also do not seem to be involved in the aerobic-anaerobic regulatory network. They belong to the subgroup B of Crp/Fnr regulators (Figure 1A). Moreover, functional predictions for the regulators of interest can sometimes be derived from the analysis of the genomic context (Vollack et al., 1999). However, corresponding analyses of the genomic context did not provide further information for the possible functions of DnrA, DnrB, and DnrC. All three of the corresponding genes are surrounded by genes coding for hypothetical proteins in the genome of *D. shibae*.

The DnrD protein encoded by Dshi\_3189 showed an amino acid sequence identity of 45.95% with Nnr of *P. denitrificans* (AAA69977.2), but only 27.48% identity with the amino acid sequence of Dnr from *P. aeruginosa*. DnrD and the protein encoded by Dshi\_3191 termed DnrE share 42.29% amino acid sequence identity and affiliated within the Dnr subgroup of the Crp/Fnr family (Figure 1C). The genes encoding DnrD and DnrE are localized within the denitrification gene cluster of *D. shibae*. This indicates a possible role in the regulation of denitrification as shown for Dnr from *P. aeruginosa* (Arai et al., 1995; Schreiber et al., 2007), *P. denitrificans* (Van Spanning et al., 1995), *R. sphaeroides* (Tosques et al., 1996), and *P. stutzeri* (Zumft, 1997; Vollack et al., 1999). A third Dnr-like regulator is encoded by Dshi\_3270 and was named DnrF. DnrF shared only 19.64% identity with Dnr of *P. aeruginosa* but 27.14% to transcriptional regulator *alr4454* of *Nostoc* sp. PCC7120 (BAB76153.1). With an overall sequence identity of 21.47% compared to the entire Dnr subfamily DnrF was placed at a basal branch point of the Dnr subfamily of the Crp/Fnr family tree (Figure 1C). It might represent a possible new subfamily of regulators within the Crp/Fnr family. In contrast to the relatively low overall amino acid sequence homology of DnrD, DnrE, and DnrF of *D. shibae* to other Dnr regulators, they show a significant amino acid sequence conservation within their DNA-binding domains and therefore may recognize similar DNA target sequences (Dufour et al., 2010). However, the function of a predicted upstream DNA target sequence has to be proven by transcriptome analyses.

### *D. shibae* FnrL Contains an Oxygen-Sensitive Fe-S Cluster

The two typical features of oxygen sensing Fnr proteins are a highly conserved HTH-motif for promoter recognition and an oxygen-sensitive [4Fe-4S]<sup>2+</sup> cluster for signal perception. A sequence alignment of *D. shibae* FnrL with homologous proteins of other *Rhodobacteraceae* species identified six highly



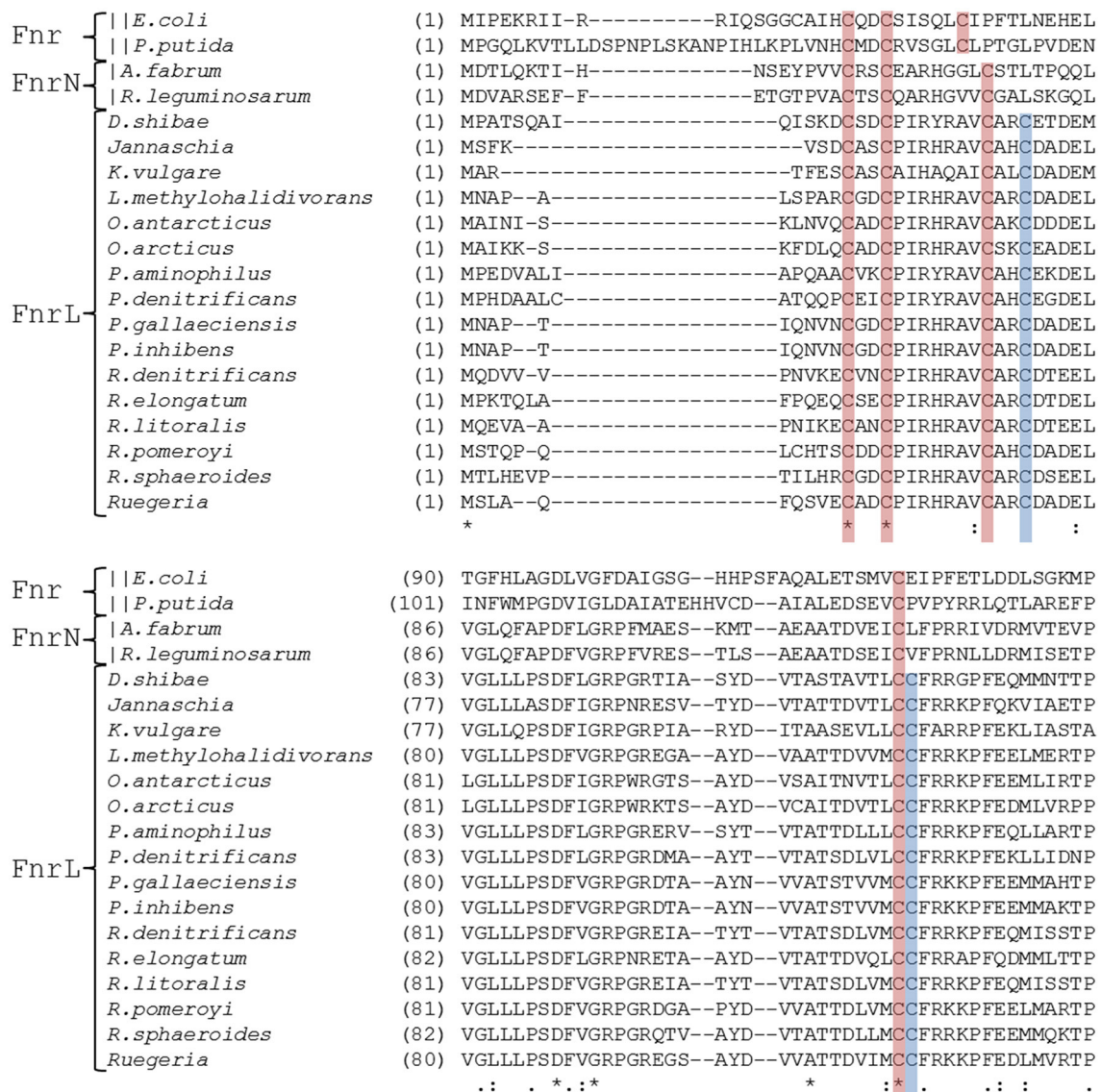
**FIGURE 1 | Phylogenetic affiliations of the *D. shibae* Crp/Fnr-like regulators. (A)** The unrooted neighbor joining trees containing *D. shibae* Crp/Fnr like regulators affiliated to the Crp/Fnr superfamily are shown. **(B)** FnrN subfamily of Fnr like regulators containing FnrL of *D. shibae*. **(C)** Dnr subgroup of Crp/Fnr like regulators containing three different Dnr regulators of *D. shibae*. The members of the tree are identified by the name of the host bacterium and the NCBI sequence identification number (accession number).

conserved cysteine residues (Figure 2). For *P. denitrificans* a mutational analysis identified four of these cysteines as functionally relevant for *in vivo* transcriptional activity, most likely as ligands for an Fe-S cluster (Hutchings et al., 2002). In order to investigate *D. shibae* FnrL for the presence of a Fe-S cluster, the regulator protein was produced as a Strep-tagged fusion protein and anaerobically purified to apparent homogeneity. UV/VIS spectroscopy performed with the purified protein revealed a broad shoulder at 420 nm, a typical spectrum for Fe-S cluster containing proteins (Figure 3A). After further reconstitution of Fe-S clusters, an increased absorption at 420 nm was detected. Fe-S cluster integrity upon exposure to air was examined in a time resolved manner using UV/VIS spectroscopy. The absorption at 420 nm at time points of 60, 120, and 240 min were shown in Figure 3B. The Fe-S cluster of *D. shibae* FnrL is obviously oxygen labile since absorption at 420 nm decreased over time. Thus, FnrL is an oxygen-dependent transcriptional regulator of *D. shibae*.

### *D. shibae* Knockout Mutants of *fnrL*, *dnrD*, *dnrE*, and *dnrF* Show Oxygen Tension-Dependent Growth Phenotypes

In order to determine the contribution of FnrL, DnrD, DnrE, and DnrF to aerobic and anaerobic growth, single gene knockout mutant strains were constructed by replacing the encoding

genes with gentamicin resistance cassettes. The respective mutant strains DS001( $\Delta fnrL$ ), DS002( $\Delta dnrD$ ), DS003( $\Delta dnrE$ ), DS004( $\Delta dnrF$ ), and the *D. shibae* DFL12<sup>T</sup> wild type strain were cultivated under aerobic and anaerobic growth conditions and their growth rates were determined (Table 3). The *fnrL* mutant strain DS001 ( $\Delta fnrL$ ) grew under aerobic and under anaerobic growth conditions comparable to the wild type strain (Table 3). In contrast, all *dnr* mutant strains DS002( $\Delta dnrD$ ), DS003( $\Delta dnrE$ ), and DS004( $\Delta dnrF$ ) showed reduced growth rates of 0.18, 0.3, and 0.25 per hour compared to the wild type strain with 0.42 per hour under aerobic conditions. Strongly retarded growth was observed for DS002 ( $\Delta dnrD$ ) and DS004 ( $\Delta dnrF$ ) under anaerobic growth conditions with growth rates of 0.02 and 0.03 per hour compared to the wild type with 0.1 per hour. Complementation of the mutant strains with their intact gene *in trans* (DS005, DS006, DS007, and DS008) resulted in restored aerobic and anaerobic growth to wild type levels for almost all mutant strains. Growth of the complemented  $\Delta dnrF$  strain DS008 under anaerobic conditions was fully restored compared to the wild type strain. In order to test for the contribution of the various regulators to the control of the aerobic to anaerobic transition, we performed shift experiments from aerobic culture in mid-exponential growth phase to anaerobic growth conditions in the presence of nitrate. Growth curves of DS001( $\Delta fnrL$ ), DS002( $\Delta dnrD$ ), DS003( $\Delta dnrE$ ), and DS004( $\Delta dnrF$ ) mutant strains compared to the *D. shibae* DFL12<sup>T</sup> wild type strain were recorded and

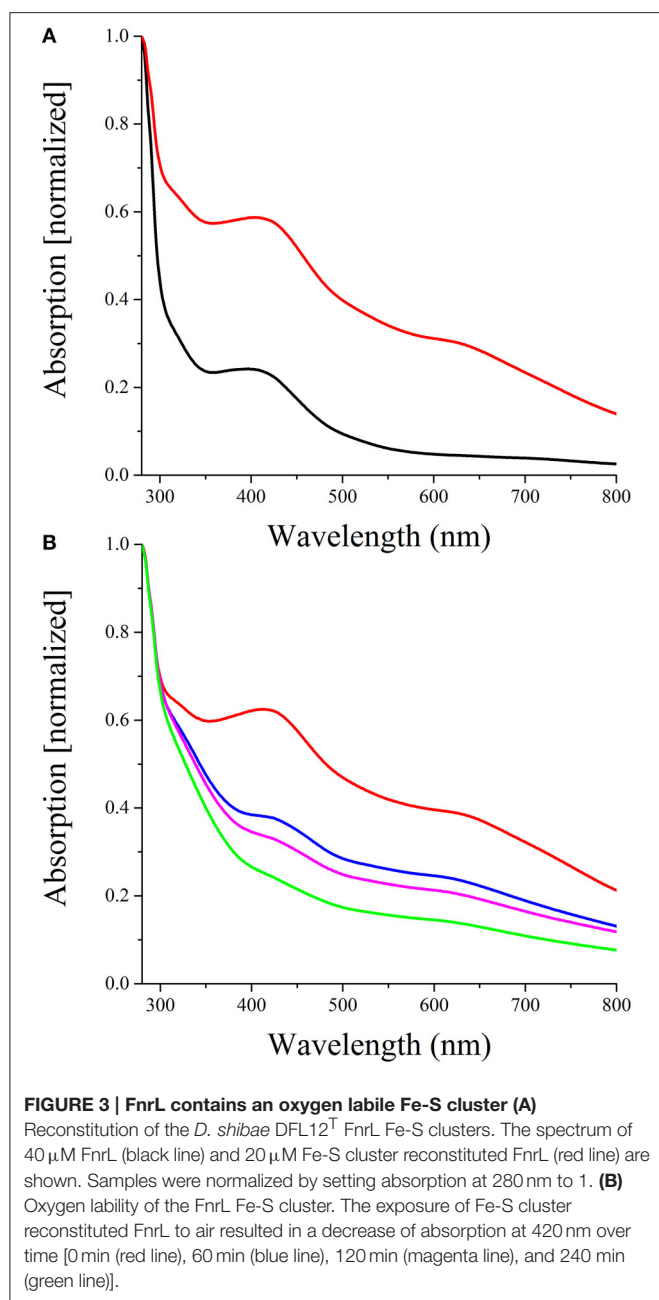


**FIGURE 2 | Alignment of the N-terminal regulatory domains of FnrL like regulators.** Cysteine residues which were known to be essential for activity in *P. denitrificans* are marked in red (Hutchings et al., 2002). Those cysteine residues which were additionally found conserved exclusively in the *Roseobacter* group are indicated in blue. The alignment comprises the amino acid sequences of FnrL proteins from *Dinoroseobacter shibae* DFL12<sup>T</sup> (WP\_012177338.1), *Paracoccus denitrificans* (WP\_041529894.1), *Rhodobacter sphaeroides* (WP\_002720880.1), *Roseobacter litoralis* (WP\_044025718.1), *Roseobacter denitrificans* (WP\_044032974.1), *Pheobacter gallaeciensis* (WP\_014881157.1), *Phaeobacter. inhibens* (WP\_040174583.1), *Roseobacterium elongatum* (WP\_025310970.1), *Ruegeria pomeroyi* (WP\_011049211.1), *Leisingera methylohalidivorans* (WP\_024088845.1), *Ruegeria* sp. (WP\_039983362.1), *Jannaschia* sp. (WP\_011456972.1), *Paracoccus aminophilus* (WP\_020949853.1), *Octadecabacter antarcticus* (WP\_015498258.1), *Octadecabacter arcticus* 238 (WP\_015493613.1), *Ketoglucicigernium vulgare* (WP\_013383280.1). In addition, Fnr from *Escherichia coli* (WP\_000611911.1) and Anr from *Pseudomonas putida* (WP\_010954166.1) were included. Symbols indicates an entirely conserved column (\*), a column comprising amino acids of same size and hydrophathy (:), and a column comprising amino acids of similar size or evolutionary preserved hydrophathy (.) (Notredame et al., 2000).

the oxygen consumption was recorded until the oxygen was no longer detectable. This was reached 15 min after the shift for each growth curve (Figure 4). Under anaerobic growth condition DS001( $\Delta$ fnrL) showed almost the same progression like the wild type strain but finally grew to higher cell density before entering the stationary growth phase (Figure 4). Thus, DS001( $\Delta$ fnrL) might have a growth advantage compared to the

wild type under anaerobic conditions. DS003( $\Delta$ dnrE) showed a slightly reduced growth compared to wild type resulting in a reduced cell number (Figure 4). Strain DS002( $\Delta$ dnrD) first showed a prolonged lag phase under aerobic conditions but in the exponential growth phase the growth rate was comparable to the wild type. After the shift to anaerobiosis DS002( $\Delta$ dnrD) stopped growing indicating a clear anaerobic phenotype (Figure 4). Strain





DS004( $\Delta dnrF$ ) also showed a prolonged lag phase but after the shift to anaerobic growth conditions, growth increased and the strain finally reached cell numbers higher than the wild type comparable to DS001( $\Delta fnrL$ ). In contrast to the *dnrD* mutant strain DS002( $\Delta dnrD$ ), DS001( $\Delta fnrL$ ), DS003( $\Delta dnrE$ ), DS004( $\Delta dnrF$ ) mutant strains were able to adapt to anaerobiosis and continued growth (Figure 4).

## Definition of the FnrL, DnrD, DnrE, and DnrF Regulons

In order to define the different regulons of the four regulators for the elucidation of their contribution to the regulatory network

**TABLE 3 | Growth rate of *D. shibae* DFL12<sup>T</sup> wild type strain and corresponding mutant strains under aerobic and anaerobic growth conditions.**

Strain	Growth rate aerob $\pm p$ ( $h^{-1}$ )	P-value aerob	Growth rate anaerob $\pm$ s.d. ( $h^{-1}$ )	P-value anaerob
DFL12 <sup>T</sup>	0.42 $\pm$ 0.02		0.1 $\pm$ 0.01	
DS001( $\Delta fnrL$ )	0.45 $\pm$ 0.02	0.5	0.11 $\pm$ 0	1
DS005( $\Delta fnrL + fnrL$ )	0.45 $\pm$ 0.03	0.2	0.11 $\pm$ 0	1
DS002( $\Delta dnrD$ )	0.18 $\pm$ 0	0	0.02 $\pm$ 0	0
DS006( $\Delta dnrD + dnrD$ )	0.39 $\pm$ 0	0.8	0.09 $\pm$ 0	1
DS003( $\Delta dnrE$ )	0.3 $\pm$ 0.04	0	0.1 $\pm$ 0	1
DS007( $\Delta dnrE + dnrE$ )	0.39 $\pm$ 0.02	0.8	0.1 $\pm$ 0	1
DS004( $\Delta dnrF$ )	0.25 $\pm$ 0.01	0	0.03 $\pm$ 0.01	0
DS008( $\Delta dnrF + dnrF$ )	0.24 $\pm$ 0.01	0	0.08 $\pm$ 0.01	0.9

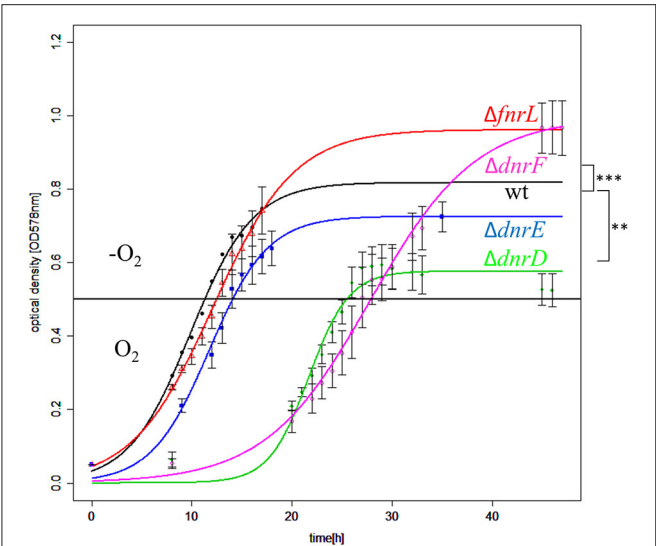
The growth rates values were determined in the logarithmic growth phase and represent the means of three technical and three biological replicates with appropriate standard deviation. P-values were determined by using regression and ANOVA followed by a Post-hoc-test with the wild type data as reference. A significant change in growth was assumed for  $P \leq 0.05$ .

for the anaerobic adaptation, transcriptome analyses were performed. First, the complete aerobic to anaerobic modulon for wild type *D. shibae* was investigated using RNA sequencing. Moreover, all involved open reading frames and corresponding transcriptional start sites became visible. The various mutant strains were analyzed using the DNA array technology. For this purpose shift experiments as outlined above were performed. RNA samples were collected directly before the shift representing the aerobic expression and 60 min after the shift. A differential gene expression was assumed for genes with an absolute log<sub>2</sub>-fold change higher or equal to 0.8 comparing wild type and mutant strains DS001( $\Delta fnrL$ ), DS002( $\Delta dnrD$ ), DS003( $\Delta dnrE$ ), DS004( $\Delta dnrF$ ). Under anaerobic conditions 329 genes were found induced in the DS001( $\Delta fnrL$ ) mutant strain compared to wild type indicating a major repressor function of FnrL. Concurrently, 160 genes were found repressed, indicating an activating function of FnrL (Table 4). For DS002( $\Delta dnrD$ ) in total 348 genes were found induced indicating a repression through DnrD and 138 were found repressed in a *dnrD* mutational background. Similar counts were found for DS004( $\Delta dnrF$ ) with 316 induced and 92 repressed genes under *dnrF* depletion. Only for DS003( $\Delta dnrE$ ) the regulon seemed to contain less genes. In total 144 genes were found induced and 69 were found repressed within the mutant strain.

## Identification of Potential Fnr and Dnr Binding Sites in the Genome of *D. shibae*

By using the Virtual Footprint tool of the PRODORIC database (Münch et al., 2005), for a global genome search of possible Anr, Fnr, and Dnr binding sites, 314 potential Fnr/Dnr binding sites were identified within the genome of *D. shibae*. We correlated these results with differential gene expression data derived from the DNA array analyses using wild type and the regulatory mutant strains. RNA sequencing analyses determined the corresponding transcriptional start sites. The localization





**FIGURE 4 | Growth behavior of *D. shibae* Dfl12<sup>T</sup> and oxygen regulatory mutants during an aerobic to anaerobic shift.** *D. shibae* DFL12<sup>T</sup> wildtype strain (black line), DS001( $\Delta$ *fnrL*) (red line), DS002( $\Delta$ *dnrD*) (green line), DS003( $\Delta$ *dnrE*) (blue line) and DS004( $\Delta$ *dnrF*) (magenta line) mutant strains were grown under aerobic conditions in artificial see water medium supplemented with 16.9 mM succinate. After reaching of OD<sub>578 nm</sub> of 0.5 cells were shifted to anaerobic growth conditions (black horizontal line) and 25 mM sodium nitrate was added. The optical density was measured in three independent replicates and error bars represent the standard deviation. *P*-values were determined by using ANOVA without a *Post-hoc*-test using the wild type values as reference. The symbols indicate *P*-values smaller or equal to 0 (\*\*\*) or 0.001 (\*\*).

**TABLE 4 | Comparative analysis of transcriptome results from *fnrL* and various *dnr* genes mutant strains under anaerobic conditions.**

	DS001( $\Delta$ <i>fnrL</i> )	DS002 ( $\Delta$ <i>dnrD</i> )	DS004 ( $\Delta$ <i>dnrF</i> )	DS003 ( $\Delta$ <i>dnrE</i> )
Activation	160	138	92	69
Repression	329	348	316	144

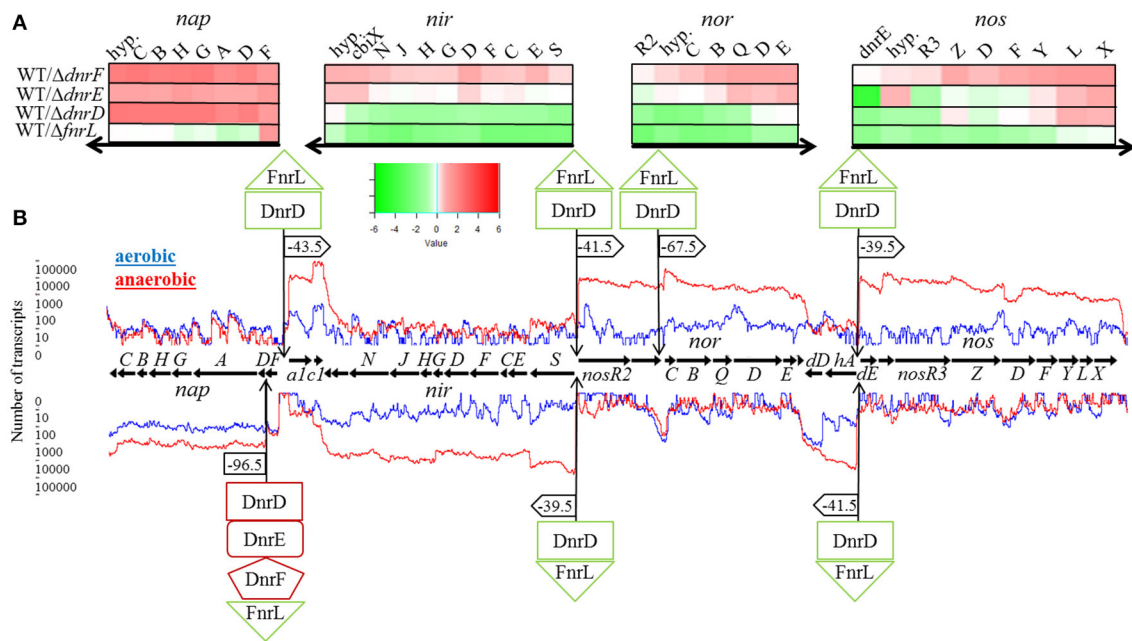
Number of differentially expressed genes (alteration in log 2-fold change  $\geq 0.8$ ) within the indicated knockout mutant strains compared to the wild type of *D. shibae* DFL12<sup>T</sup>.

of the identified binding sites were given with respect to the transcriptional start site of the regulated genes. Since, we got only a low coverage of transcript by RNA sequencing for the plasmid encoded genes, distances of the identified binding sites were given with respect to the translational start points. These analyses correlated 69 FnrL binding sites with induction or repression of the cognate genes by FnrL (Table S1). For DnrD 10 functional bindings sites, for DnrE 4 binding sites and for DnrF 8 binding sites were identified (Table S2).

**Interplay of FnrL, DnrD, DnrE, and DnrF for the Regulation of the Denitrification Gene Cluster**

We first focused on the role of the various regulators for the transcription of the denitrification gene cluster consisting of

the *nap* operon, encoding the periplasmic nitrate reductase (EC 1.7.99.4), the *nir* operon, encoding the nitrite reductase (EC 1.7.2.1), the *nor* operon encoding the nitric oxide reductase (EC 1.7.2.5), and the *nos* operon coding for the nitrous oxide reductase protein complex (EC 1.7.2.4; **Figure 5**). The different transcriptional units were deduced from the RNA sequencing data (**Figure 5**) and visualized using TRAV (Dietrich et al., 2014). The RNA sequencing experiments revealed a 10- to 100-fold increase in transcription for the denitrification genes under anaerobic conditions. Involved promoter regions were deduced. These results were combined with a bioinformatical transcription factor binding site prediction using the Virtual Footprint tool of the PRODORIC database (Münch et al., 2005). The expression of the *napDAGHBC* operon was found induced by FnrL and significantly repressed by DnrD, DnrE, and DnrF under anaerobic conditions (**Figure 5A**). This indicated an antagonistic function of FnrL and Dnr at the *napD* promoter. A possible regulatory binding site was found 96.5 bp upstream of the translational start site of *napD* within the *napF* transcript (**Figure 5B**). The palindromic sequence 5'-TTGAT-N<sub>4</sub>-ATCAA-3' exhibit a high degree of similarity to the Anr/Dnr binding site of *P. aeruginosa* (bold letters; Winteler and Haas, 1996; Rompf et al., 1998; Trunk et al., 2010). In contrast, FnrL/Dnr-independent transcription of the *napF* gene was observed. The gene *napF* encodes a cytoplasmic Fe-S protein that was also found in *E. coli* and *R. sphaeroides* 2.4.1. A strong cooperative anaerobic induction by both FnrL and DnrD was observed for *apbE1*, *cycA1*, and *hemA3* genes as well for the *nirSECFDGHJN*, *norCBQDE*, and *nosRZDFYLX* operons (**Figure 5**). *In silico* promoter analyses revealed the palindromic sequence 5'-TTAAT-N<sub>4</sub>-CAGAA-3' 43.5 bp upstream of the transcriptional start site of the *apbE1* gene as potential binding site for FnrL and/or DnrD (**Figure 5**). A coupled transcription with *cycA1* encoding a cytochrome C2 was not observed. Despite the anaerobic transcriptional activation of *cycA1* by FnrL and DnrD a potential regulator binding site was not found (**Figure 5**). Next, the *nirSECFDGHJN* operon and the divergently transcribed *nosR2* gene, encoding flavin-containing nitrous oxide reductase maturation factor (Zhang et al., 2017), were found anaerobically induced by FnrL and DnrD. Within their shared upstream regions one potential binding site 5'-TTAAC-N<sub>4</sub>-GTCAA-3' was identified. The sequence motive is located 39.5 bp upstream of the transcriptional start site of *nirS* and 41.5 bp upstream of the transcriptional start site of *nosR2*, both almost optimal positions for Fnr-dependent promoters (Tielen et al., 2012; **Figure 5**). A coupled transcription of *nosR2* and the following *norCBQDE* was indicated by the RNA sequencing data. Nevertheless, an additional Fnr/Dnr binding motif 5'-TTGAC-N<sub>4</sub>-GTAA-3' was found 67.5 bp upstream of the *norC* translational start site, allowing for *nosR2*-independent NorCB formation. Another Fnr binding site 5'-TTAAC-N<sub>4</sub>-GTCAA-3' was found located 41.5 bp upstream of the transcriptional start site of the *hemA3* gene, encoding the heme biosynthetic enzyme 5-aminolevulinic acid synthase and 39.5 bp upstream of the transcriptional start site of the *dnrE* gene, which is divergently transcribed to *hemA3* (**Figure 5**). RNA sequencing data indicate a coupled



**FIGURE 5 | Regulation of the denitrification operons by FnrL, DnrD, DnrE, and DnrF. (A)** Heat map representation of denitrification gene expression patterns of mutant strains DS001( $\Delta$ *fnrL*), DS002( $\Delta$ *dnrD*), DS003( $\Delta$ *dnrE*), and DS004( $\Delta$ *dnrF*) compared to *D. shibae* DFL12<sup>T</sup> wild type strain grown under anaerobic conditions. The colored bars represent the expression level in log<sub>2</sub> scale. Green indicates a relatively low expression level in the mutant strain which indicates activation by the regulator; red indicates relatively high expression levels in the mutant strain compared to the wild type indicating repression. **(B)** Denitrification operons with corresponding RNA sequencing data based transcriptional start sites and transcript quantification under aerobic (blue line) and anaerobic conditions (red line) of the wild type *D. shibae* DFL12<sup>T</sup>. Black horizontal arrows indicate open reading frames and the direction of transcription. Additionally, binding sites of FnrL, DnrD, DnrE, and DnrF are indicated by flags. If the corresponding binding site was located within a promoter sequence, the corresponding distance to the transcriptional start was given. Green boxes indicate an activation by the given regulator, red boxes indicate a repression. *a1*, *apbE1*; *c1*, *cycA1*; *hyp*, hypothetical gene; *dD*, *dnrD*; *hA*, *hemA3*; *dE*, *dnrE*.

transcription of *dnrE*, the open reading frame Dshi\_3192 and the *nosRZDFYLX* operon. The transcription of *hemA3*, *dnrE*, Dshi\_3192, and the *nosRZDFYLX* operon was found induced under anaerobic conditions by FnrL and partly by DnrD. Interestingly, *dnrD* gene expression was constitutive and not dependent on oxygen tension. Thus, we identified the first part of aerobic-anaerobic regulatory cascade, in which constitutively produced FnrL and DnrD induce *dnrE* transcription upon anaerobiosis and denitrification. Similarly, *dnrF* transcription is controlled by FnrL and Dnr. As outlined above DnrE and DnrF are repressing *nap* operon expression. Moreover, a significant solely DnrF-dependent transcriptional repression of the *nosRZDFYLX* operon was observed (summarized in **Figures 8, 10**). Since only DnrF mediates repression of the *nos* operon via this palindromic sequence, one could predict this regulation as DnrF specific. Noteworthy is the fact that those two bidirectional Fnr/Dnr binding sequences between *nirS* and *nosR2* as well as between *hemA3* and *dnrE* are controlling the entire denitrification process. These palindromic sequences exhibit an overall identity of 85% (TTGAC(G/T)TT(T/G)GTAA) with a single inversion of the central four nucleotides. Due to the promoter specificity of FnrL and DnrD an overlapping bidirectional active binding motive was suggested.

## Expression of the *nirS* and *nosR2* Promoter Is under the Control of One Central Bidirectional Functional FnrL/Dnr Binding Site

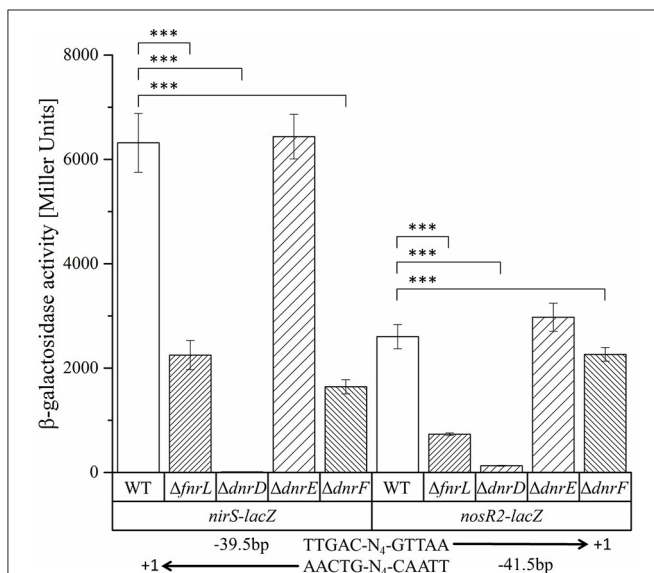
Obviously, the bidirectional employed promoter regions and corresponding Fnr/Dnr binding sites between *nirS* and *nosR2* and between *hemA3* and *dnrE* were predicted to control the onset of denitrification in *D. shibae*. To study the role of the centrally located potential bidirectional binding site between the divergently *nirS* and *nosR2* genes, we created appropriate promoter-*lacZ* reporter gene fusions. We choose a 118 bp DNA fragment spanning the promoter sequences between the transcriptional start site of *nirS* and *nosR2*. The DNA sequence was cloned in both directions upstream of the *lacZ* gene resulting in *nirS-lacZ* and *nosR2-lacZ* reporter gene fusions (**Figure 6**). The reporter gene fusions were transformed into the *D. shibae* wild type strain and all four regulator mutant strains: DS001( $\Delta$ *fnrL*), DS002( $\Delta$ *dnrD*), DS003( $\Delta$ *dnrE*), DS004( $\Delta$ *dnrF*). The resulting strains were grown under anaerobic conditions in the presence of 25 mM nitrate and  $\beta$ -galactosidase activities were determined in the mid-exponential growth phase. In the wild type strain *nirS-lacZ* expression resulted in  $6,318 \pm 564$  Miller Units. In the DS001( $\Delta$ *fnrL*) mutant strain still 2,250

$\pm 282$  Miller Units were determined while in DS002( $\Delta dnrD$ ) *nirS-lacZ* expression was totally abolished with  $11 \pm 0.45$  Miller Units. In the DS003( $\Delta dnrE$ ) mutant strain *nirS-lacZ* expression was comparable to wild type levels. In the DS004( $\Delta dnrF$ ) mutant strain  $1,643 \pm 428$  Miller Units were measured. Overall expression of *nosR2-lacZ* was lower compared to *nirS-lacZ*. In the wild type strain  $2,603 \pm 233$  Miller Units were measured. But again, expression was found reduced in DS001( $\Delta fnrL$ ) with  $732 \pm 26$  Miller units and almost absent in DS002( $\Delta dnrD$ ). In contrast, wild type levels were found in DS003( $\Delta dnrE$ ) and DS004( $\Delta dnrF$ ). These results identified DnrD as the main regulator of *nirS-lacZ* and *nosR2-lacZ* expression under denitrifying growth condition and FnrL as an additional co-regulator. To confirm the functional role of the potential Fnr/Dnr binding site within the *nirS* promoter, we mutated the terminal nucleotides of the palindrome TT to GC and AA to GC resulting in the mutant sequence 5'-GCAAC-N<sub>4</sub>-GTCGC-3'. Expression of the mutated *nirS(mu)-lacZ* reporter gene fusion was studied in the wild type strain under aerobic and anaerobic growth conditions. Under aerobic conditions neither *nirS-lacZ* nor *nirS(mu)-lacZ* expression were found, confirming strict anaerobic expression (Figure 7). Under anaerobic conditions *nirS-lacZ* expression was induced as expected for the wild type promoter, but no expression was found for *nirS(mu)-lacZ*. This indicated the importance of the Fnr/Dnr binding site and

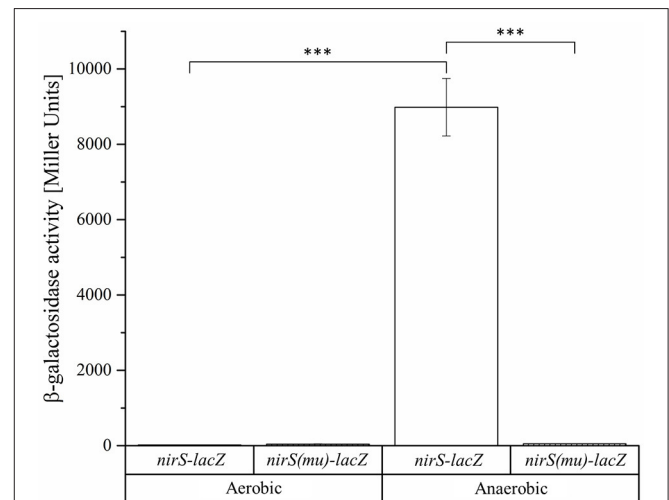
especially the relevance of the highly conserved TT and AA sequences for regulator binding.

## Genes Encoding the Electron Transport Chain Are Affected by FnrL and Dnr Regulators in *D. shibae*

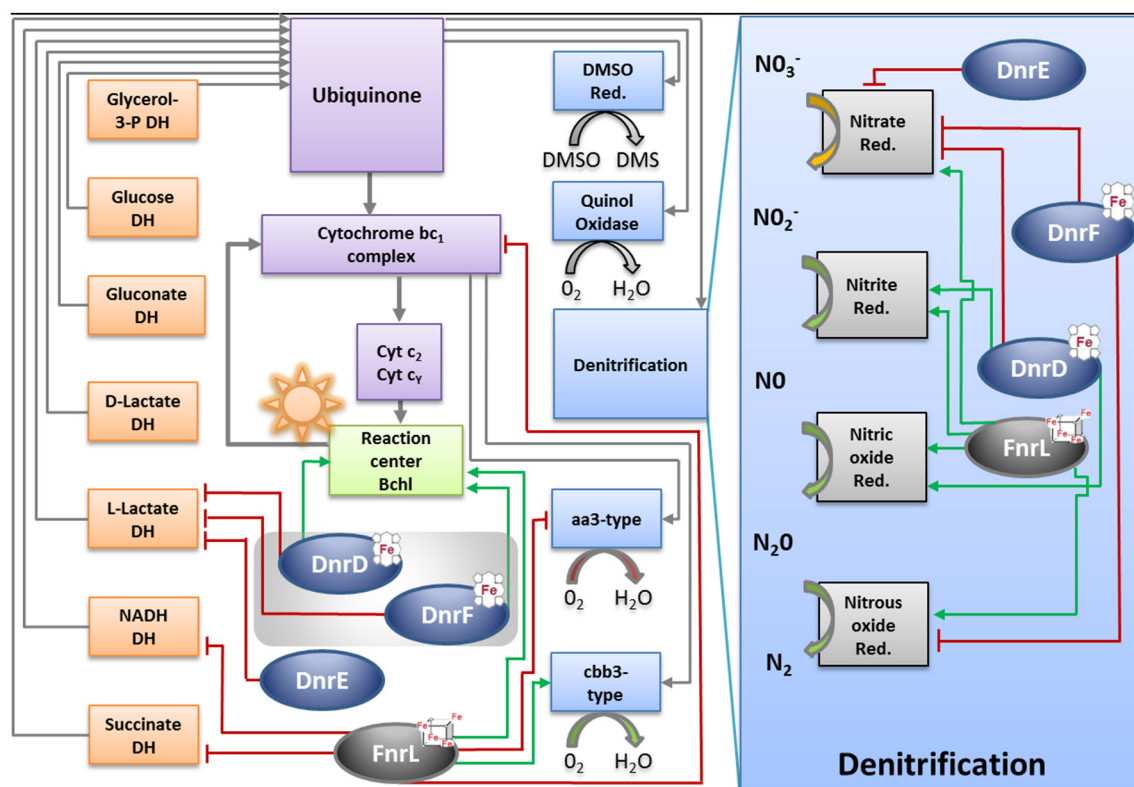
In addition to the denitrification genes, genes encoding components of the *D. shibae* electron transport chain were found differentially expressed under anaerobic conditions dependent on FnrL and DnrD (Table S3; Figure 8). The overall composition of the electron transport chain was deduced from the genome of *D. shibae* (Figure 8; Wagner-Döbler et al., 2010). As described previously, the acquisition of electrons occurred by different primary dehydrogenases (Laass et al., 2014). In other bacteria aerobic and anaerobic electron transport chains often employ different electron donor systems. An upregulation of the *nuo*-operon (Dshi\_1307-1326; Dshi\_1327-1330) encoding an NADH dehydrogenase I (EC 1.6.99.5), Dshi\_1390 encoding an alternative NADH dehydrogenase (EC 1.6.5.3 and EC 1.6.99.3) and the *sdh*-operon (Dshi\_2861-2867) encoding the succinate dehydrogenase (EC 1.3.99.1) was observed under conditions of anaerobiosis (Figure 8). Interestingly, an FnrL-mediated anaerobic repression of the NADH dehydrogenase (EC 1.6.99.5, EC 1.6.5.3, and EC 1.6.99.3) and the succinate dehydrogenases (EC 1.3.99.1) was found. Moreover, the repression of the L-lactate dehydrogenase (*lldD2*; EC 1.1.2.3) was found to be mediated by all three Dnr regulators (Figure 8). Genes for other known



**FIGURE 6 | Functional investigation of the *nirS-nosR2* intergenic promoter region using *nirS-lacZ* and *nosR2-lacZ* reporter gene fusions.** *D. shibae* DFL12<sup>T</sup> wild type (WT), DS001( $\Delta fnrL$ ) ( $\Delta fnrL$ ), DS002( $\Delta dnrD$ ) ( $\Delta dnrD$ ), DS003( $\Delta dnrE$ ) ( $\Delta dnrE$ ), and DS004( $\Delta dnrF$ ) ( $\Delta dnrF$ ) mutant strains carrying the *nirS-lacZ* and *nosR2-lacZ* reporter gene fusions were grown under oxygen-limited conditions, and  $\beta$ -galactosidase activity was measured. Error bars represent the observed standard deviation. *P*-values were determined using ANOVA and the Tukey-test (Tukey, 1949). The symbol \*\*\* indicate *P*-values smaller or equal to 0. The palindromic sequence of the Fnr/Dnr binding site and its position with respect to the transcriptional start site of *nirS* and *nosR2* are given.



**FIGURE 7 | Functional importance of the Fnr/Dnr binding site in the intergenic region of *nirS* and *nosR2*.** *D. shibae* wild type strains DFL12<sup>T</sup> carrying the *nirS-lacZ* and *nirS(mu)-lacZ* reporter gene fusions were grown under aerobic and anaerobic growth conditions and  $\beta$ -galactosidase activities were measured in three independent replicates. The *nirS-lacZ* reporter gene fusion is carrying the palindromic sequence 5'-TTAAC-N<sub>4</sub>-GTCAA-3' of the Fnr/Dnr binding site. The binding sequence was mutated to 5'-GCAAC-N<sub>4</sub>-GTCGC-3' in the *nirS(mu)-lacZ* reporter gene fusion. Error bars represent the observed standard deviation. *P*-values were determined using ANOVA and the Tukey-test (Tukey, 1949). The symbol \*\*\* indicate *P*-values smaller or equal to 0.



**FIGURE 8 | Model for the regulation of the energy metabolism of *D. shibae* by FnrL, DnrD, DnrE and DnrF during the aerobic to anaerobic transition.** DH, dehydrogenase; Cyt, cytochrome; bchl, bacteriochlorophyll; Q<sub>d</sub> and Q<sub>b</sub> oxidized and reduced form of reaction center quinone, respectively; Cox, cytochrome oxidase; DMSO, dimethyl sulfoxide; Red, reductase. Figure was adapted from Wagner-Döbler et al. (2010).

primary dehydrogenases remain unaffected by the Fnr and Dnr regulators. An FnrL-dependent repression of the genes encoding an electron transferring flavoprotein (*eftA/B*) was detected. An anaerobic activation of terpenoid backbone synthesis genes (*ispA*, *crtE*, *bchP*) was observed which might result in an increase of ubiquinone biosynthesis.

Significant FnrL-dependent repression of transcription of *petC/B* genes encoding the cytochrome *bc*<sub>1</sub> complex (EC 1.10.2.2) was observed. Cytochrome *bc*<sub>1</sub> channels electrons toward the photosynthesis machinery and two types of cytochrome oxidases using a cytochrome *c* pool (Figure 8; Tables S1, S3). A considerable FnrL-dependent anaerobic activation of bacteriochlorophyll and carotenoid biosynthesis was noticed. Genes encoding the chlorophyllide a reductase (*bchCXYZ*), the light-independent protochlorophyllide reductase (*bchNB*) and the photosynthetic reaction center (*pufQBALMC*) were found activated by FnrL. Due to a compact genomic organization of photoactive pigments synthesis machinery similar regulatory patterns were observed for phytoene desaturase (*crtBI*), spirilloxanthin synthesis (*crtCDE*), and the spheroidene monooxygenase (*crtA*). Furthermore, an FnrL-mediated induced transcription of the *fixNOQP*GHIS operon encoding a high-affinity *cbb*<sub>3</sub> type cytochrome oxidase (EC 1.9.3.1) was observed (Figure 8; Table S1, S3). On the other hand a significant anaerobic transcriptional repression of *ctaCBGE* and *ctaD* encoding a low affinity *aa*<sub>3</sub>-type cytochrome oxidase (EC 1.9.3.1)

was found. The final ATP generation by an *F*<sub>0</sub>*F*<sub>1</sub>-ATP synthases gene clusters (*atpHAGDC/atpIBEXF*; EC 3.6.3.14) was found transcriptionally repressed through FnrL and DnrD (Figure 8; Table S3).

## The Essential Role of FnrL for the Adaptation of *D. shibae* to Low Oxygen Tension

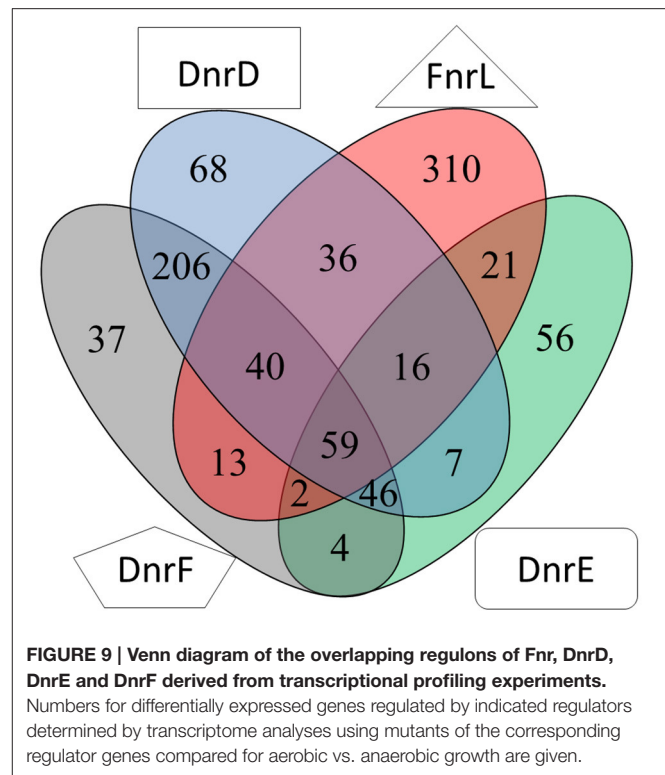
Under anaerobic conditions 477 genes organized within 268 transcriptional units were regulated by FnrL. Taken into account, that only 69 promoters carried a potential Fnr binding site, we assume that the majority of the other genes is regulated in an indirect manner. Thus, the FnrL-specific regulon comprises 69 transcriptional units (Table S2). We performed cluster enrichment analysis of orthologous groups (Blanka et al., 2014; Table S4). In addition to the outlined genes involved in denitrification and corresponding electron transport chains, genes encoding various transcriptional regulators were part of the FnrL regulon. Beside the autoregulation of the *fnrL* gene, an anaerobic activation of the *dnrE* and *dnrF* gene was observed. Moreover, the gene encoding the benzyl coenzyme A (CoA) reduction regulator (*bedM/rnf2*) and for quorum sensing regulators (*traR/dksA*) were found activated by FnrL. Repression of genes for an ABC transport systems regulator (*rbsB*) and two *luxR* family regulators was observed. In addition,



various biosynthetic pathways were FnrL controlled. Genes encoding enzymes of the tetrapyrrole biosynthesis were found activated by FnrL. These include 5-aminolaevulinic acid synthase HemA (Dshi\_1182/Dshi\_2190) and the coproporphyrinogen III dehydrogenase HemN (Dshi\_0541/Dshi\_0659), the light-independent protochlorophyllide reductase BchFNBH (Dshi\_3533-3536) and chlorophyllide reductase BchXYZF (Dshi\_3517-3519/Dshi\_3533). Obviously, FnrL is also controlling the restructuring of the cell envelope in response to oxygen limitation. Beside the activation of a gene encoding a putative membrane protein (Dshi\_1454) the highest log2-fold change was found for transcription of a gene encoding the outer membrane protein (*ompW*; Bouchal et al., 2010). On the other hand the genes for a part of the protein translocation channel (*secE*), an ABC transporter of unknown function (Dshi\_1404) and two metallic cation/iron-siderophore transporters (*znuA/sitA*) were found repressed. In addition, the gene of an iron-sulfur cluster assembly accessory protein (Dshi\_1730) was found repressed by FnrL. Finally, genes of the cellular stress response are part of the Fnr regulon. For example, genes for universal stress proteins Usp (Dshi\_1338/Dshi\_2213/Dshi\_2686) and genes encoding a cytochrome-c peroxidase (*ccpA*; EC 1.11.1.5) were found highly activated by direct FnrL interaction.

### The Role of the Three Different Dnrs of *D. shibae* for the Aerobic-Anaerobic Transition

Based on the transcriptomic experiments cluster enrichment analyses of orthologous groups were performed (Blanka et al., 2014; Table S4). The determined regulons were merged in a Venn diagram (Figure 9). Especially the DnrD regulator was assumed to be the major player for the regulation of the denitrification processes and the corresponding electron transport chain. We found a set of 68 genes, which were regulated exclusively by DnrD under anaerobic conditions. However, cluster enrichment analyses revealed no orthologous group of genes that was regulated exclusively by DnrD. Nevertheless, a significant activation was observed for detoxification systems like a sulfite export system (Dshi\_0205) and a di-haem cytochrome c peroxidase (Dshi\_2749). Moreover, we were able to identify 10 potential regulator binding sites for DnrD in front of DnrD regulated genes. A palindromic sequence was found centered 42.5 bp upstream of the transcriptional start of a gene encoding a mineral and organic ion transporter (Dshi\_1384). On the other hand pilus assembly genes (Dshi\_1130) were found repressed by DnrD (Table S2). A huge overlap of 53% was found for the DnrD and DnrF regulon. Pathways like the carotenoid biosynthesis, the metabolism of terpenoids and polyketides, the degradation of aromatic compounds, cell motility, sulfur metabolism, oxidative phosphorylation, and methane metabolism were found regulated by DnrD and DnrF (Table S2). Furthermore, we found genes for porine biosynthesis (*ompR*/Dshi\_0212), environmental information processing (*kspTE*) and homologs recombination (*recA*) differentially expressed in DnrD and DnrF depleted strains. Due to the observed cascade regulation, in which DnrD induces the *dnrF* gene, it was not possible to distinguish between



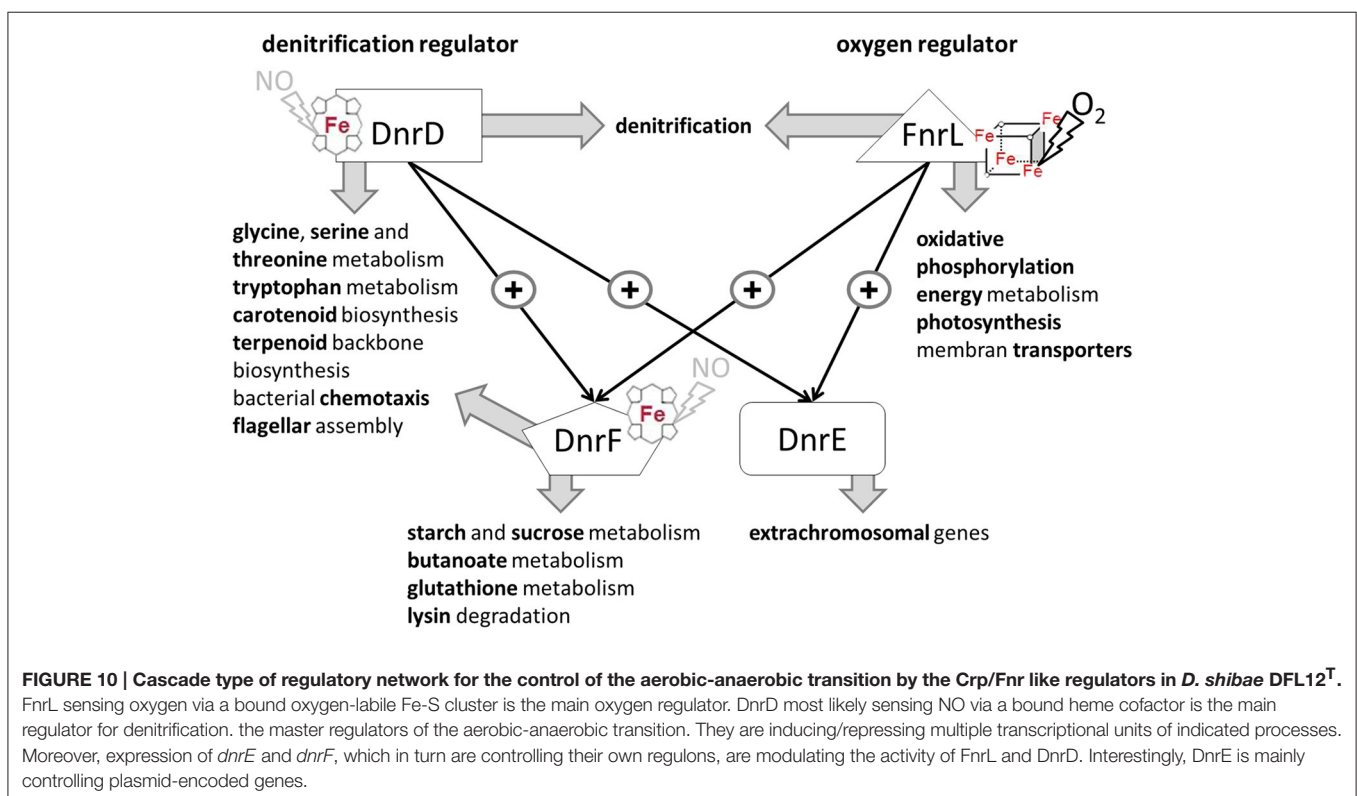
direct DnrD activation and indirect DnrD activity via *dnrF* activation and subsequent DnrF activity. A surprisingly small DnrF specific regulon was deduced with 37 genes. Processes like the biosynthesis of secondary metabolites, signal transduction, iron homeostasis (*hmuS*, *iscR*, Dshi\_0882) and the formation of the quinoprotein glucose dehydrogenase (*gcd*; EC: 1.1.5.2) were found controlled by DnrF. Potential Fnr/Dnr regulator binding sites and significant transcriptional activation were found for eight promoters. Repression of fermentation processes via the downregulation of the genes of an alcohol dehydrogenase class III (*adhC*/Dshi\_0473) and changes in the osmotic adaptation via the repression of L-carnitine dehydratase formation (Dshi\_3269) were observed (Table S2). The genes of DnrD and DnrE were found in close neighborhood on the genome and exhibit a high amino acid sequence homology. Nevertheless, only a limited overlap of regulons of both regulators was observed (Figure 9). However, the enrichment analysis of DnrE revealed a major impact on translation, lipid metabolism and biosynthesis of amino acids. Interestingly, 73.58% of the specific DnrE regulated genes were located on the various plasmids of *D. shibae*. Genes for cytochromes *b<sub>561</sub>* (Dshi\_4169), heme uptake (Dshi\_4225) and proton/sodium antiporter synthesis (*phaACEFG*) were found exclusively altered in a *dnrE* mutant strain. Moreover, a significant correlation of the observed gene expression and found regulator binding site in the upstream region of the corresponding genes was observed. These data suggest that *D. shibae* possesses a specific regulator for extrachromosomal genes which are essential for the adaptation process to various oxygen tensions (Ebert et al., 2013).

## DISCUSSION

Multiple alternative respiratory and fermentative systems enable bacteria to grown in the absence of oxygen. A broad variety of electron donating primary dehydrogenases, various electron transferring quinones and many different terminal oxidases using alternative electron acceptor including nitrate, sulfate, fumarate, and various metals have been observed. Similarly, bacterial aerobic respiration is using multiple electron-donating and excepting enzyme systems. Depending on the ecological niche inhabited by the bacterium different combinations of these electron transport chains are employed. Often bacteria are capable of aerobic and anaerobic growth. Consequently, a shift from aerobic to anaerobic growth requires a finely-tuned restructuring of the involved electron transport chains with their multiple protein and cofactor components. Parameters like oxygen tension, the presence of nitrate, NO or the electron flux in the membrane provide the bacterium with necessary information to induce the required regulatory process. Species-specific combinations of multiple regulatory proteins are used by the different bacteria to induce a tailor-made regulatory response (see Section Introduction for details).

Here we describe for the first time a regulatory network solely composed of four Crp/Fnr-family regulators. It allows the marine bacterium *D. shibae* the transition from aerobic to anaerobic growth. Like in many other bacteria oxygen tension is detected via an oxygen-labile Fe-S cluster attached to an Fnr-type regulator, here FnrL (Härtig and Jahn, 2012; Tielen et al., 2012). Similar to the regulatory network of *P. aeruginosa* and

*P. stutzeri* a second Crp/Fnr family regulator termed Dnr, here DnrD, detecting NO via a bound heme cofactors, is required for the full induction of the alternative respiratory system of denitrification (Zumft, 1997; Schreiber et al., 2007; Trunk et al., 2010). However, the nitrate responsive NarX/L system, present in *Pseudomonas*, is missing in *D. shibae*. No genes for other known systems of nitrate detection or membrane associated electron flux measurements, like ArcAB, were detected. The periplasmic nitrate reductase of the Nap type convert available nitrate to nitrite. Chemical conversion of nitrite generates NO, the signal for DnrD. In the absence of oxygen this regulator in combination with FnrL induces the whole denitrification pathway (Figure 8). FnrL controls the formation or repression of other terminal oxidases. No dedicated regulator for the repression of genes of the aerobic energy metabolism was observed. This task was achieved by the whole network. For this purpose FnrL and DnrD are inducing the production of two additional Dnrs, namely DnrE and DnrF (Figure 10). This is also the pre-requisite for the fine-tuned adaption of the set of primary dehydrogenases, the coordination with the anoxygenic photosynthesis (Figure 8) and the usual stress response (Usps, cold shock, glutathione) accompanying major physiological changes. Major overlapping regulons were detected for the four Crp/Fnr family regulators of *D. shibae* (Figure 9) allowing for a tightly and cooperatively controlled gene expression. DnrF induces a Na<sup>+</sup>/H<sup>+</sup> antiporter system. In agreement, a transposon mutagenesis approach with *D. shibae* identified various genes of Na<sup>+</sup> gradient formation and utilization as essential for anaerobic growth (Ebert et al., 2013). This provides further evidence for a role of membrane



localized Na<sup>+</sup> gradient in the adaptation to anaerobic growth. The transposon mutagenesis approach also identified multiple essential genes of *D. shibae* for anaerobic growth encoded by the five plasmids of the bacterium (Ebert et al., 2013). In agreement, the activity of one of the Dnr regulators, DnrE, is dedicated to plasmid gene control. Obviously, due to the slower anaerobic growth protein biosynthesis, amino acid biosynthesis, and ATP generation are slowed down (Laass et al., 2014). Corresponding gene regulatory adaptations, with decreased ribosomal protein and ATPase formation have been observed. However, only parts of the observed gene regulatory scenario can be attributed to the direct activity of the four Crp/Fnr regulators, most likely other processes including stringent response might account for the observed adaptation. Overall, a novel type of gene regulatory network composed of one Fnr and three Dnr proteins is tightly controlling the aerobic to anaerobic transition of the marine bacterium *D. shibae*.

## AUTHOR CONTRIBUTIONS

ME, EH, and DJ: Substantial contributions to the conception and design of the work, data collection, data analysis, and

interpretation, drafting the article critical revision to the article, final approval to the version to be published. SL, DE, LR, and AT: Contributions to conception and design of the work, data collection, data analysis, and interpretation. RD: Contribution to design of the work, data analysis, and interpretation.

## FUNDING

This work was supported by the DFG in Transregio-SFB TRR51.

## ACKNOWLEDGMENTS

This work was supported by the DFG in Transregio-SFB TR51. We are grateful to the technical assistance of Anja Hartmann (Institute of Microbiology, Technische Universität Braunschweig).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.00642/full#supplementary-material>

## REFERENCES

- Arai, H. (2003). Transcriptional regulation of the *nos* genes for nitrous oxide reductase in *Pseudomonas aeruginosa*. *Microbiology* 149, 29–36. doi: 10.1099/mic.0.25936-0
- Arai, H., Igarashi, Y., and Kodama, T. (1995). Expression of the *nir* and *nor* genes for denitrification of *Pseudomonas aeruginosa* requires a novel CRP/ FNR-related transcriptional regulator, DNR, in addition to ANR. *FEBS Lett.* 371, 73–76. doi: 10.1016/0014-5793(95)00885-D
- Becher, A., and Schweizer, H. P. (2000). Integration-proficient *Pseudomonas aeruginosa* vectors for isolation of single-copy chromosomal *lacZ* and *lux* gene fusions. *Biotechniques* 29, 948–952.
- Biebl, H., Allgaier, M., Tindall, B. J., Koblizek, M., Lünsdorf, H., Pukall, R. (2005). *Dinoroseobacter shibae* gen. nov., sp. nov., a new aerobic phototrophic bacterium isolated from dinoflagellates. *Int. J. Syst. Evol. Microbiol.* 55, 1089–1096. doi: 10.1099/ijls.0.63511-0
- Blanka, A., Schulz, S., Eckweiler, D., Franke, R., Bielecka, A., Nicolai, T., et al. (2014). Identification of the alternative sigma factor SigX regulon and its implications for *Pseudomonas aeruginosa* pathogenicity. *J. Bacteriol.* 196, 345–356. doi: 10.1128/JB.01034-13
- Botella, E., Fogg, M., Jules, M., Piersma, S., Doherty, G., Hansen, A., et al. (2010). pBaSysBioII: an integrative plasmid generating *gfp* transcriptional fusions for high-throughput analysis of gene expression in *Bacillus subtilis*. *Microbiology* 156, 1600–1608. doi: 10.1099/mic.0.035758-0
- Bouchal, P., Struhárová, I., Budinská, E., Šedo, O., Vyhliďalová, T., Zdráhal, Z., et al. (2010). Unraveling an FNR based regulatory circuit in *Paracoccus denitrificans* using a proteomics-based approach. *Biochim. Biophys. Acta* 1804, 1350–1358. doi: 10.1016/j.bbapap.2010.01.016
- Buchan, A., Gonzalez, J. M., and Moran, M. A. (2005). Overview of the marine roseobacter lineage. *Appl. Environ. Microbiol.* 71, 5665–5677. doi: 10.1128/AEM.71.10.5665-5677.2005
- Crothers, D. M., and Steitz, T. A. (1992). “Transcriptional activation by *Escherichia coli* CAP protein,” in *Transcription Regulation*, eds S. L. McKnight and K. R. Yamamoto (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), 501–534.
- Dietrich, S., Wiegand, S., and Liesegang, H. (2014). TraV: a genome context sensitive transcriptome browser. *PLoS ONE* 9:e93677. doi: 10.1371/journal.pone.0093677
- Dufour, Y. S., Kiley, P. J., and Donohue, T. J. (2010). Reconstruction of the core and extended regulons of global transcription factors. *PLoS Genet.* 6:e1001027. doi: 10.1371/journal.pgen.1001027
- Ebert, M., Laaß, S., Burghartz, M., Petersen, J., Kofsmehl, S., Wohlbrand, L., et al. (2013). Transposon mutagenesis identified chromosomal and plasmid genes essential for adaptation of the marine bacterium *Dinoroseobacter shibae* to anaerobic conditions. *J. Bacteriol.* 195, 4769–4777. doi: 10.1128/JB.00860-13
- Edgar, R. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30, 207–210.
- Fürch, T., Preusse, M., Tomasch, J., Zech, H., Wagner-Döbler, I., Rabus, R., et al. (2009). Metabolic fluxes in the central carbon metabolism of *Dinoroseobacter shibae* and *Phaeobacter gallaeciensis*, two members of the marine Roseobacter clade. *BMC Microbiol.* 9:209. doi: 10.1186/1471-2180-9-209
- Green, J., Bennett, B., Jordan, B., Ralph, E. T., Thomson, A. J., and Guest, J. R. (1996). Reconstitution of the [4Fe-4S] cluster in FNR and demonstration of the aerobic-anaerobic transcription switch *in vitro*. *Biochem. J.* 316, 887–892. doi: 10.1042/bj3160887
- Green, J., and Guest, J. R. (1993). Activation of FNR-dependent transcription by iron: An *in vitro* switch for FNR. *FEMS Microbiol. Lett.* 113, 219. doi: 10.1111/j.1574-6968.1993.tb06517.x
- Green, J., Scott, C., and Guest, J. R. (2001). Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP. *Adv. Microb. Physiol.* 44, 1–34. doi: 10.1016/S0065-2911(01)44010-0
- Härtig, E., Geng, H., Hartmann, A., Hubacek, A., Munch, R., Ye, R. W., et al. (2004). *Bacillus subtilis* ResD induces expression of the potential regulatory genes *yclJ/K* upon oxygen limitation. *J. Bacteriol.* 186, 6477–6484. doi: 10.1128/JB.186.19.6477-6484.2004
- Härtig, E., and Jahn, D. (2012). Regulation of the anaerobic metabolism in *Bacillus subtilis*. *Adv. Microb. Physiol.* 61, 195–216. doi: 10.1016/B978-0-12-394423-8.00005-6
- Härtig, E., and Zumft, W. G. (1999). Kinetics of *nirS* Expression (Cytochrome cd<sub>1</sub> nitrite reductase) in *Pseudomonas stutzeri* during the transition from aerobic respiration to denitrification: evidence for a denitrification-specific nitrate- and nitrite-responsive regulatory system. *J. Bacteriol.* 181, 161–166.
- Hoang, T. T., Karkhoff-Schweizer, R. R., Kutchma, A. J., and Schweizer, H. P. (1998). A broad-host-range FLP-FRT recombination system for site-specific excision of chromosomally-located DNA sequences: application for

- isolation of unmarked *Pseudomonas aeruginosa* mutants. *Gene* 212, 77–86. doi: 10.1016/S0378-1119(98)00130-9
- Hutchings, M. I., Crack, J. C., Shearer, N., Thompson, B. J., Thomson, A. J., and Spiro, S. (2002). Transcription factor FnrP from *Paracoccus denitrificans* contains an iron-sulfur cluster and is activated by anoxia: identification of essential cysteine residues. *J. Bacteriol.* 184, 503–508. doi: 10.1128/JB.184.2.503-508.2002
- Jahn, M., and Jahn, D. (2012). “Electron transfer reactions and oxidative phosphorylation,” in *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, eds G. Michal and D. Schomburg (New York, NY: John Wiley & Sons), 183–187.
- Katzke, N., Arvani, S., Bergmann, R., Circolone, F., Markert, A., Svensson, V., et al. (2010). A novel T7 RNA polymerase dependent expression system for high-level protein production in the phototrophic bacterium *Rhodobacter capsulatus*. *Protein Expr. Purif.* 69, 137–146. doi: 10.1016/j.pep.2009.08.008
- Khoroshilova, N., Popescu, C., Münck, E., Beinert, H., and Kiley, P. J. (1997). Iron-sulfur cluster disassembly in the FNR protein of *Escherichia coli* by O<sub>2</sub>: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity. *Proc. Natl. Acad. Sci. U.S.A.* 94, 6087–6092. doi: 10.1073/pnas.94.12.6087
- Kiley, P. J., and Beinert, H. (1998). Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster. *FEMS Microbiol. Rev.* 22, 341. doi: 10.1111/j.1574-6976.1998.tb00375.x
- Kiley, P. J., and Reznikoff, W. S. (1991). Fnr mutants that activate gene expression in the presence of oxygen. *J. Bacteriol.* 173, 16–22. doi: 10.1128/jb.173.1.16-22.1991
- Körner, H., Sofia, H. J., and Zumft, W. G. (2003). Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs. *FEMS Microbiol. Rev.* 27, 559–592. doi: 10.1016/S0168-6445(03)00066-4
- Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, R. M. II, et al. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175–176. doi: 10.1016/0378-1119(95)00584-1
- Kovach, M. E., Phillips, R. W., Elzer, P. H., Roop, R. M., and Peterson, K. M. (1994). pBBR1MCS: a broad-host-range cloning vector. *Biotechniques* 16, 800–802.
- Kumka, J. E., and Bauer, C. E. (2015). Analysis of the FnrL regulon in *Rhodobacter capsulatus* reveals limited regulon overlap with orthologues from *Rhodobacter sphaeroides* and *Escherichia coli*. *BMC Genomics* 16:895. doi: 10.1186/s12864-015-2162-4
- Laass, S., Kleist, S., Bill, N., Druppel, K., Kossmehl, S., Wohlbrand, L., et al. (2014). Gene regulatory and metabolic adaptation processes of *Dinoroseobacter shibae* DFL12T during oxygen depletion. *J. Biol. Chem.* 289, 13219–13231. doi: 10.1074/jbc.M113.545004
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
- Lazazzera, B. A., Beinert, H., Khoroshilova, N., Kennedy, M. C., and Kiley, P. J. (1996). DNA binding and dimerization of the Fe-S-containing FNR protein from *Escherichia coli* are regulated by oxygen. *J. Biol. Chem.* 271, 2762–2768. doi: 10.1074/jbc.271.5.2762
- Miller, J. H. (1992). *A Short Course in Bacterial Genetics: A Laboratory Manual and Handbook for Escherichia coli and Related Bacteria*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., et al. (2005). Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. *Bioinformatics* 21, 4187–4189. doi: 10.1093/bioinformatics/bti635
- Notredame, C., Higgins, D. G., and Heringa, J. (2000). T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* 302, 205–217. doi: 10.1006/jmbi.2000.4042
- Righetti, P. G. (1990). *Immobilized PH Gradients: Theory and Methodology*. Amsterdam: Elsevier.
- Rinaldo, S., Castiglione, N., Giardina, G., Caruso, M., Arcovito, A., Longa, S., et al. (2012). Unusual heme binding properties of the dissimilative nitrate respiration regulator, a bacterial nitric oxide sensor. *Antioxid. Redox Signal.* 17, 1178–1189. doi: 10.1089/ars.2011.4226
- Rompf, A., Hungerer, C., Hoffmann, T., Lindenmeyer, M., Römmling, U., Groß, U., et al. (1998). Regulation of *Pseudomonas aeruginosa* hemF and hemN by the dual action of the redox response regulators Anr and Dnr. *Mol. Microbiol.* 29, 985–997. doi: 10.1046/j.1365-2958.1998.00980.x
- Schreiber, K., Krieger, R., Benkert, B., Eschbach, M., Arai, H., Schobert, M., et al. (2007). The anaerobic regulatory network required for *Pseudomonas aeruginosa* nitrate respiration. *J. Bacteriol.* 189, 4310–4314. doi: 10.1128/JB.00240-07
- Schultz, S., Shields, G., and Steitz, T. (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. *Science* 253, 1001–1007. doi: 10.1126/science.1653449
- Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., et al. (2017). Phylogenomics of *Rhodobacteraceae* reveals evolutionary adaptation to marine and non-marine habitats. *ISME J.* doi: 10.1038/ismej.2016.198. [Epub ahead of print].
- Spiro, S., and Guest, J. R. (1990). FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. *FEMS Microbiol. Rev.* 6, 399.
- Thoma, S., and Schobert, M. (2009). An improved *Escherichia coli* donor strain for diparental mating. *FEMS Microbiol. Lett.* 294, 127–132. doi: 10.1111/j.1574-6968.2009.01556.x
- Tielen, P., Schobert, M., Härtig, E., and Jahn, D. (2012). “Anaerobic regulatory networks in bacteria,” in *Bacterial Regulatory Networks*, ed A. Filloux (Norfolk, VA: Caister Academic Press), 273–305.
- Tomasch, J., Gohl, R., Bunk, B., Diez, M. S., and Wagner-Döbler, I. (2011). Transcriptional response of the photoheterotrophic marine bacterium *Dinoroseobacter shibae* to changing light regimes. *ISME J.* 5, 1957–1968. doi: 10.1038/ismej.2011.68
- Tosques, I. E., Shi, J., and Shapleigh, J. P. (1996). Cloning and characterization of *nmrR*, whose product is required for the expression of proteins involved in nitric oxide metabolism in *Rhodobacter sphaeroides* 2.4.3. *J. Bacteriol.* 178, 4958–4964. doi: 10.1128/jb.178.16.4958-4964.1996
- Trageser, M., and Uden, G. (1989). Role of cysteine residues and of metal ions in the regulatory functioning of FNR, the transcriptional regulator of anaerobic respiration in *Escherichia coli*. *Mol. Microbiol.* 3, 593–599. doi: 10.1111/j.1365-2958.1989.tb00206.x
- Trunk, K., Benkert, B., Quäck, N., Münch, R., Scheer, M., Garbe, J., et al. (2010). Anaerobic adaptation in *Pseudomonas aeruginosa*: definition of the Anr and Dnr regulons: *P. aeruginosa* Anr and Dnr regulon. *Environ. Microbiol.* 12, 1719–1733. doi: 10.1111/j.1462-2920.2010.02252.x
- Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics* 5, 99. doi: 10.2307/3001913
- Van Spanning, R. J. M., De Boer, A. P. N., Reijnders, W. N. M., Spiro, S., Westerhoff, H. V., Stouthamer, A. H., et al. (1995). Nitrite and nitric oxide reduction in *Paracoccus denitrificans* is under the control of NNR, a regulatory protein that belongs to the FNR family of transcriptional activators. *FEBS Lett.* 360, 151–154. doi: 10.1016/0014-5793(95)00091-M
- Vollack, K. U., Härtig, E., Körner, H., and Zumft, W. G. (1999). Multiple transcription factors of the FNR family in denitrifying *Pseudomonas stutzeri*: characterization of four Fnr-like Genes, regulatory responses and cognate metabolic processes. *Mol. Microbiol.* 31, 1681–1694. doi: 10.1046/j.1365-2958.1999.01302.x
- Vollack, K. U., and Zumft, W. G. (2001). Nitric oxide signaling and transcriptional control of denitrification genes in *Pseudomonas stutzeri*. *J. Bacteriol.* 183, 2516–2526. doi: 10.1128/JB.183.8.2516-2526.2001
- Wagner-Döbler, I., Ballhausen, B., Berger, M., Brinkhoff, T., Buchholz, I., Bunk, B., et al. (2010). The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker's guide to life in the sea. *ISME J.* 4, 61–77. doi: 10.1038/ismej.2009.94
- Wagner-Döbler, I., and Biebl, H. (2006). Environmental biology of the marine *Roseobacter* lineage. *Annu. Rev. Microbiol.* 60, 255–280. doi: 10.1146/annurev.micro.60.080805.142115
- Winkler, A., Heintz, U., Lindner, R., Reinstein, J., Shoeman, R. L., and Schlichting, I. (2013). A ternary AppA-PpsR-DNA complex mediates light regulation of photosynthesis-related gene expression. *Nat. Struct. Mol. Biol.* 20, 859–867. doi: 10.1038/nsmb.2597
- Winteler, H. V., and Haas, D. (1996). The homologous regulators ANR of *Pseudomonas aeruginosa* and FNR of *Escherichia coli* have overlapping but



- distinct specificities for anaerobically inducible promoters. *Microbiology* 142, 685–693. doi: 10.1099/13500872-142-3-685
- Wu, J., and Bauer, C. E. (2008). “RegB/RegA, a global redox-responding two-component system,” in *Bacterial Signal Transduction: Networks and Drug Targets*, Vol. 631, ed R. Utsumi (New York, NY: Springer), 131–148.
- Zhang, L., Trncik, C., Andrade, S. L. A., and Einsle, O. (2017). The flavinyl transferase ApbE of *Pseudomonas stutzeri* matures the NosR protein required for nitrous oxide reduction. *Biochim. Biophys. Acta* 1858, 95–102. doi: 10.1016/j.bbabi.2016.11.008
- Zumft, W. G. (1997). Cell biology and molecular basis of denitrification. *Microbiol. Mol. Biol. Rev.* 61, 533–616.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Ebert, Laaß, Thürmer, Roselius, Eckweiler, Daniel, Härtig and Jahn. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Large Diversity and Original Structures of Acyl-Homoserine Lactones in Strain MOLA 401, a Marine *Rhodobacteraceae* Bacterium

Margot Doberva<sup>1</sup>, Didier Stien<sup>1</sup>, Jonathan Sorres<sup>1</sup>, Nathalie Hue<sup>2</sup>, Sophie Sanchez-Ferandin<sup>3</sup>, Véronique Eparvier<sup>2</sup>, Yoan Ferandin<sup>1</sup>, Philippe Lebaron<sup>1</sup> and Raphaël Lami<sup>1\*</sup>

<sup>1</sup> Sorbonne Universités, UPMC Univ Paris 6, CNRS, Laboratoire de Biodiversité et Biotechnologies Microbiennes (LBBM), Observatoire Océanologique, Banyuls/Mer, France, <sup>2</sup> CNRS, Institut de Chimie des Substances Naturelles (ICSN), Université Paris-Sud, Gif-sur-Yvette, France, <sup>3</sup> Sorbonne Universités, UPMC Univ Paris 6, CNRS, Biologie Intégrative des Organismes Marins (BIOM), Observatoire Océanologique, Banyuls/Mer, France

## OPEN ACCESS

### Edited by:

Meinhard Simon,  
University of Oldenburg, Germany

### Reviewed by:

Eva Sonnenschein,  
Technical University of Denmark,  
Denmark  
Stefan Schulz,  
Technische Universität Braunschweig,  
Germany

### \*Correspondence:

Raphaël Lami  
raphael.lami@obs-banyuls.fr

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

Received: 12 January 2017

Accepted: 07 June 2017

Published: 22 June 2017

### Citation:

Doberva M, Stien D, Sorres J,  
Hue N, Sanchez-Ferandin S,  
Eparvier V, Ferandin Y, Lebaron P  
and Lami R (2017) Large Diversity  
and Original Structures  
of Acyl-Homoserine Lactones  
in Strain MOLA 401, a Marine  
*Rhodobacteraceae* Bacterium.  
Front. Microbiol. 8:1152.  
doi: 10.3389/fmicb.2017.01152

Quorum sensing (QS) is a density-dependent mechanism allowing bacteria to synchronize their physiological activities, mediated by a wide range of signaling molecules including N-acyl-homoserine lactones (AHLs). Production of AHL has been identified in various marine strains of Proteobacteria. However, the chemical diversity of these molecules still needs to be further explored. In this study, we examined the diversity of AHLs produced by strain MOLA 401, a marine *Alphaproteobacterium* that belongs to the ubiquitous *Rhodobacteraceae* family. We combined an original biosensors-based guided screening of extract microfractions with liquid chromatography coupled to mass spectrometry (MS), High Resolution MS/MS and Nuclear Magnetic Resonance. This approach revealed the unsuspected capacity of a single *Rhodobacteraceae* strain to synthesize 20 different compounds, which are most likely AHLs. Also, some of these AHLs possessed original features that have never been previously observed, including long (up to 19 carbons) and poly-hydroxylated acyl side chains, revealing new molecular adaptations of QS to planktonic life and a larger molecular diversity than expected of molecules involved in cell-cell signaling within a single strain.

**Keywords:** quorum sensing, acyl-homoserine lactone, marine bacteria, *Rhodobacteraceae*

## INTRODUCTION

Quorum sensing (QS) allows bacteria to sense their population density (Nealson, 1977) and coordinate their gene expression (Bassler, 1999; Fuqua and Greenberg, 2002) and physiology (Miller and Bassler, 2001). QS communication is based on the secretion and detection of small molecules by bacteria in their nearby environment (Atkinson and Williams, 2009). A large number of studies have demonstrated that QS regulates many different bacterial features including biofilm production (Parsek and Greenberg, 2005; Dickschat, 2010), nodulation (Cha et al., 1998;

Loh et al., 2002), bioluminescence (Waters and Bassler, 2005), virulence factor production (Smith and Iglewski, 2003) among others (Diggle et al., 2007). The coordination of bacterial community activities is supposed to confer an ecological advantage to the population (Case et al., 2008).

Among the various molecular signals used in QS systems, AHLs (acyl-homoserine lactone or autoinducer type-1, AI-1) constitute the major class of semiochemicals (Williams et al., 2007; Papenfort and Bassler, 2016) which has been widely studied (Fuqua et al., 1994). AHLs are homoserine lactone (HSL) linked to fatty acyl chains through an amide bond. The acyl chain length can vary from 4 to 18 carbons and sometimes includes a 3-oxo or a 3-hydroxy functional group (Fuqua and Greenberg, 2002). AHLs are usually saturated, but some unsaturated bonds in the fatty acyl chains are known. AHLs are synthesized by AHL-synthases, which catalyze the amide bond formation between the acyl chain carried by the ACP (acyl carrier protein) and the amine moiety precursor SAM (S-adenosyl-methionine) (Fuqua and Greenberg, 2002; Parsek and Greenberg, 2005). Three AHL synthase genes are currently known: *ainS*-like (Gilson et al., 1995), *luxI*-like (Engebrecht and Silverman, 1984), and *hdtS*-like (Loh et al., 2002). Of these three, *ainS*-like genes are found only in *Vibrio*, *luxI*-like genes are the most well-studied and are present in many *Proteobacteria* genomes (Gelencser et al., 2012) and little is known about *hdtS*-like genes (Doberva et al., 2015).

Many marine bacteria regulate some of their physiological traits using QS systems, among them the *Rhodobacteraceae*, a key bacterial family in marine environments that drives important biogeochemical reactions (Gram et al., 2002; Schaefer et al., 2002; Wagner-Döbler et al., 2005; Wagner-Döbler and Biebl, 2006). *Rhodobacteraceae* are abundant in the ocean and it has been demonstrated that 87% of completely sequenced genomes in this group encode *luxI*-like genes (Cude and Buchan, 2013; Zan et al., 2014). Known AHL-producing *Rhodobacteraceae* are diverse (Wagner-Döbler et al., 2005). Among them, *Ruegeria* spp. is found associated with sponges (Mohamed et al., 2008; Zan et al., 2012, 2015), *Dinoroseobacter* spp. lives in association with dinoflagellates (Patzelt et al., 2013), *Sulfitobacter* sp. is a diatom-associated bacteria. (Amin et al., 2015; Limardo and Worden, 2015), and *Phaeobacter gallaeciensis* proliferates in coastal waters (Wagner-Döbler et al., 2005; Berger et al., 2011). Interestingly, QS has been mainly identified in strains isolated from niches where bacteria can reach high concentrations (phycosphere, sponges tissues) (Rolland et al., 2016). More generally, it is commonly thought that QS is uncommon in marine oligotrophic strains as bacterial concentrations in oligotrophic environments (approximately  $10^5$  cells per mL) were presumably too low to trigger QS behaviors. However, a few recent publications report the occurrence of QS in marine bacterial strains isolated in oligotrophic environments. In their pioneering work, Moran et al. (2004) sequenced the full genome of *Silicibacter pomeroyi*, an oligotrophic *Roseobacter* and detected two QS systems. Similarly, in a previous study, we reported many translated gene sequences affiliated to *Roseobacters luxI* and *hdtS* from oligotrophic environments in the predicted proteome of

Global Ocean Sampling metagenomic dataset (Doberva et al., 2015).

Collectively, these preliminary observations suggest that QS could constitute an important physiological trait of *Rhodobacteraceae* in all types of aquatic environments. New AHLs are regularly described in this group (Schaefer et al., 2008; Thiel et al., 2009; Ziesche et al., 2015). This suggests that the real extent of AHL chemical diversity in marine bacteria is still unknown. However, to our knowledge, no studies had yet used bioguided microfractionation combined with thorough mass spectrometry-based and Nuclear Magnetic resonance spectroscopy (NMR) for an in-depth description of AHLs diversity emitted marine *Rhodobacteraceae*. In this study, we investigated the potential of strain MOLA 401, isolated in an oligotrophic lagoon, to produce different types of AHLs. The strain MOLA 401 was isolated in a tropical oligotrophic lagoon located in New Caledonia. A closely related strain (*Maribius pelagius* B5-6<sup>T</sup>; 96% 16S rRNA sequence identity) has been isolated in the oligotrophic Sargasso Sea (Atlantic Ocean) (Choi et al., 2007). We had sequenced the full genome of the strain MOLA 401, and previously reported the presence of *luxI*, *luxR*, and *hdtS* genes, revealing the potential of this strain to communicate by QS (Doberva et al., 2014). We report here on the chemical diversity of strain MOLA 401 AHLs.

## MATERIALS AND METHODS

### Culture of Strain MOLA 401

The strain MOLA 401 is from the MOLA culture collection (WDCM911<sup>1</sup>) and available (strain code BBCC401) upon request<sup>2</sup>. This strain was isolated on December 3, 2004 at 4 m depth from marine oligotrophic waters in the southwest lagoon of New Caledonia (France; 22°21.23'S/166°23.43'E) (Conan et al., 2008). Sampled waters harbored a Chl *a* concentration of 1.07  $\mu\text{g ml}^{-1}$  (F. Joux, pers. communication). All culturing steps were performed using Marine Broth (MB) 2216 (BD Difco, Sparks, MD, United States of America). The draft genome sequence has been published under accession number JQEY000000000 and revealed a quorum-sensing dependent physiology (Doberva et al., 2014).

### Phylogenetic Analyses of the Strain MOLA 401 16S rRNA and Quorum Sensing Genes

Phylogenetic analyses were conducted using MEGA 6 software (Tamura et al., 2013) to assess the position of strain MOLA 401 QS genes with respect to already published phylogenies and to infer their putative functional role. Four different genes were considered: 16S rRNA, *luxR* (an AHL receptor), *luxI* (an AHL synthase), and *hdtS* (another AHL synthase). All alignments were performed using ClustalW (Larkin et al., 2007) and trimmed manually. Phylogenetic trees were constructed using the

<sup>1</sup><http://collection.obs-banyuls.fr/index.php>

<sup>2</sup><http://collection.obs-banyuls.fr/ordering.php>

Neighbor-Joining (NJ) method with p-distance correction and 1000 bootstrap replicates.

## AHL Standards Used in This Study

N-acyl-homoserine lactones were obtained from Cayman Chemical (Ann Arbor, MI, United States). Stock solutions ( $10 \text{ mmol L}^{-1}$ ) of the analytes were prepared in dimethylsulfoxide (DMSO). Oxo-C14:1-HSL was dissolved in acetonitrile ( $\text{CH}_3\text{CN}$ ). The list of standards AHLs is provided in Supplementary Table S1.

## Extraction of AHL from Strain MOLA 401 Supernatants

The strain MOLA401 was pre-cultured in 30 mL (96 h,  $25^\circ\text{C}$ , 100 rpm) and then cultivated under aerobic conditions in 3 L of MB in 6 Erlenmeyer flasks under continuous shaking (200 rpm,  $25^\circ\text{C}$ , 72 h, 5 mL from the preculture). At late exponential cell growth phase (72 h), tert-butyl methyl ether (600 mL) was added in each flask. This mixture was shaken overnight at room temperature (150 rpm). The two phases were then separated and the organic phase was dried with  $\text{MgSO}_4$ . Solvent was filtered and removed with a rotary evaporator. The crude extract (170.6 mg) was dissolved in DMSO ( $57 \text{ mg mL}^{-1}$ ). A Phenomenex Strata C18, 55  $\mu\text{m}$ , 5 g column was equilibrated with  $\text{CH}_3\text{CN}$  (100 mL),  $\text{H}_2\text{O}:\text{CH}_3\text{CN}$  75:25 (35 mL), then  $\text{H}_2\text{O}$  (100 mL). The crude extract dissolved in DMSO was deposited on top of the column. Elution was carried out with  $\text{H}_2\text{O}$  (50 mL),  $\text{CH}_3\text{CN}$  (50 mL, Fraction M). Fraction M was then evaporated in a rotary evaporator yielding 40.6 mg of material.

## HPLC Fractionation

Fraction M was then dissolved in DMSO ( $40 \text{ mg mL}^{-1}$ ) and fractionated on a preparative HPLC system with 2 Varian Prep Star pumps, a manual injector, a Dionex Ultimate 3000 RS variable wavelength detector and a Dionex Ultimate 3000 fraction collector. The column was a Phenomenex Luna C18, 5  $\mu\text{m}$ ,  $21.2 \times 250 \text{ mm}$ , and the flow rate was set to  $20 \text{ mL min}^{-1}$ . The solvent was gradient grade  $\text{H}_2\text{O}$  and  $\text{CH}_3\text{CN}$  (70:30 for 3 min, followed by a 12 min linear gradient from 70:30 to 0:100, followed by 100%  $\text{CH}_3\text{CN}$  for 10 min). The eluents were monitored at 214, 254, 274, and 280 nm, and were collected between 3 and 25 min (1 fraction  $\text{min}^{-1}$ , 22 fractions total referenced as M1–M22). The solvent was removed from each fraction with a genevac HT-4X system. Each fraction was dissolved in 100  $\mu\text{L}$  DMSO to perform biosensor tests.

## Culture of Biosensor Strains

The 22 fractions (M1–M22) previously prepared were tested in the AHL biosensor assay following previously described protocols using *Pseudomonas putida* and *Escherichia coli* based biosensors (Andersen et al., 2001; Riedel et al., 2001; Steindler and Venturi, 2007). Briefly, *P. putida* F117 (pRK-C12; Km<sup>r</sup>; *ppuI::npt*) was used for the detection of long-chain AHLs (Andersen et al., 2001) and *E. coli* MT102 (pJBA132) for the detection of short chain AHLs (Riedel et al., 2001). *E. coli* MT102 and *P. putida* F117 were cultivated in Luria–Bertani

(LB) Broth (Sigma L3022) overnight with continuous shaking (200 rpm), at  $37^\circ\text{C}$  supplemented with tetracycline ( $25 \mu\text{g mL}^{-1}$ ) and at  $30^\circ\text{C}$  supplemented with gentamicin ( $20 \mu\text{g mL}^{-1}$ ), respectively. An overnight culture of each biosensor strain (200  $\mu\text{L}$ ) was inoculated in 9.8 mL of fresh LB medium with the adapted antibiotics. This fresh biosensor culture was dispensed into 96-well microplates (180  $\mu\text{L}$  per well). Then, the microfractions in DMSO (20  $\mu\text{L}$ ) were added in each well in triplicate. Microplates were incubated at  $30^\circ\text{C}$  and  $37^\circ\text{C}$  depending of growth optimum of the selected biosensor strain, without shaking. After 0, 5 and 24 h of incubation, fluorescence was determined with a Victor1420 Multilabel Counter (Perkin–Elmer) at an excitation wavelength of 485 nm and a detection wavelength of 535 nm. OD620 was also measured to control for biosensor cell growth. Negative controls were biosensor cultures without extract, and sterile LB medium. Biosensor cultures with addition of commercial AHLs (C6-HSL for *E. coli* MT102 and oxo-C10-HSL for *P. putida* F117) were used as a positive control.

## LC-MS Analyses

UHPLC-MS analyses were performed with a Waters (Milford, CT, United States) Acquity UPLC-TQD (Triple Quadrupole Detector) system controlled by the MassLynx 4.1 software. Column was an Acquity HSS C18 ( $2.1 \times 50 \text{ mm}$ ) with 1.8  $\mu\text{m}$  particle size (Waters). The column oven was set to  $40^\circ\text{C}$ . The flow rate was maintained at  $0.6 \text{ mL min}^{-1}$  and the injection volume was 2  $\mu\text{L}$ . The mobile phase was composed of 0.1% formic acid in water (eluent A) and 0.1% formic acid in acetonitrile (B). A gradient profile was used, starting with 95% of A, keeping this composition constant for 0.5 min. Proportion of B was linearly increased to 100% in 6.5 min, and was left at 100% for 3 min.

The T.Q. Detector operated in ElectroSpray Ionization (ESI) in the positive and negative modes. First, the third quadrupole (Q3) has been used in scanning mode on the  $m/z$  50–800 mass range in order to confirm the molecular weight and the purity of our 26 standard AHLs ( $2 \text{ mg mL}^{-1}$  in DMSO, 2  $\mu\text{L}$  injected), but also to determine their retention time (RT) under our chromatographic conditions.

Two cone voltages (30 and 60 volts) were applied both in ESI<sup>+</sup> and ESI<sup>−</sup> modes. The other ion source parameters were as follows: capillary voltage 3.2 kV for positive mode (3 kV in negative mode), the source temperature was set at  $150^\circ\text{C}$  and the desolvation temperature was  $450^\circ\text{C}$ . Nitrogen was used as desolvation gas at a flow rate of  $800 \text{ L h}^{-1}$  and as cone gas at a flow rate of  $50 \text{ L h}^{-1}$ . The analytical approach first involved the study of mass spectra obtained for our standard molecules. These compounds ionized significantly better in the positive mode and the signal of the protonated molecule ( $[\text{M}+\text{H}]^+$ ) appeared more abundant when the cone voltage involved was lower. A peak corresponding to the cationized AHL with ubiquitous sodium ( $[\text{M}+\text{Na}]^+$ ) often had a significant intensity too. Applying a higher cone voltage led to fragmentation in the ion source. In particular, a fragment ion at  $m/z$  102 was specific for the HSL moiety. This signal was chosen as the specific ion indicating the presence of HSL-type compounds. In a second step, the first quadrupole (Q1) of the TQD instrument was used in scanning mode from  $m/z$  50–500 as mass range and several cone voltages



(10, 15, 20, and 25 volts) were applied in order to determine the best value to observe the more intense  $[M+H]^+$  signal for each standard AHL. The  $[M+H]^+$  ions were later used as the precursor ion for MS/MS experiments. Each ion of interest was selected by the first quadrupole (Q1) and then focused in the collision cell (Q2) where fragmentation reactions occurred. The resulting fragment ions were finally analyzed by the third quadrupole (Q3). The collision gas (argon) was introduced into the collision cell to maintain a pressure near to  $4.5 \times 10^{-3}$  mbar. The collision energy was optimized to lead to an attenuation of the precursor ion beam of almost 85%. The fragmentation pattern of each  $[M+H]^+$  standard ion (MS/MS spectrum) has been recorded with the most suitable parameters for a later comparison with those obtained for the signals of interest observed in samples.

## Molecular Formula Determination and High Resolution MS/MS

High-resolution MS/MS analyses were conducted with a Thermo UHPLC-HRMS system. Analyses of microfractions and standards (1.0  $\mu$ L injected) were performed in electrospray positive ionization mode in the 133.4–2000 Da range in centroid mode. The mass detector was an Orbitrap MS/MS FT Q-Exactive focus mass spectrometer. The analysis was conducted in FullMS data dependent MS2 mode. In FullMS, resolution was set to 70,000 and AGC target was  $3 \cdot 10^6$ . In MS2, resolution was 17,500, AGC target  $10^5$ , isolation window 0.4 Da, normalized collision energy 30, with 15 s dynamic exclusion. UHPLC column was a Phenomenex Luna Omega polar C-18  $150 \times 2.1$  mm, 1.6  $\mu$ m. The column temperature was set to 42 °C, and the flow rate was 0.5 mL min<sup>-1</sup>. The solvent system was a mixture of water (A) with increasing proportions of acetonitrile (B), both solvents modified with 0.1% formic acid. The gradient was as follows: 2% B 3 min before injection, then from 1 to 13 min, a shark fin gradient increase of B up to 100% (curve 2), followed by 100% B for 5 min. The flow was discarded (not injected into the mass spectrometer) before injection and up to 1 min after injection. The exact masses and corresponding molecular formulas are reported in **Table 2**. A full list of standards along with RTs and exact masses is provided in Supplementary Table S1.

## Molecular Networking

A molecular network was constructed based on UHPLC-HRMS/MS analyses using the GNPS platform<sup>3</sup>. Nodes from MOLA 401 microfractions are in yellow, those from standards appear in blue, and those detected in both are in green. The number of compared ions was set to 8, and the minimum cosine for linking two parent ions was set to 0.7. With these parameters, only short side chains AHLs were not clustered. Since the strain microfractions contained so many AHLs, the detection limit was set to 1000 in order to simplify the cluster.

## NMR Analyses

Nuclear magnetic resonance spectra were recorded in DMSO-*d*<sub>6</sub> on a Bruker 600 MHz NMR spectrometer equipped with a 1 mm

inverse detection probe. Chemical shifts ( $\delta$ ) are reported as ppm based on the tetramethylsilane signal.

# RESULTS AND DISCUSSION

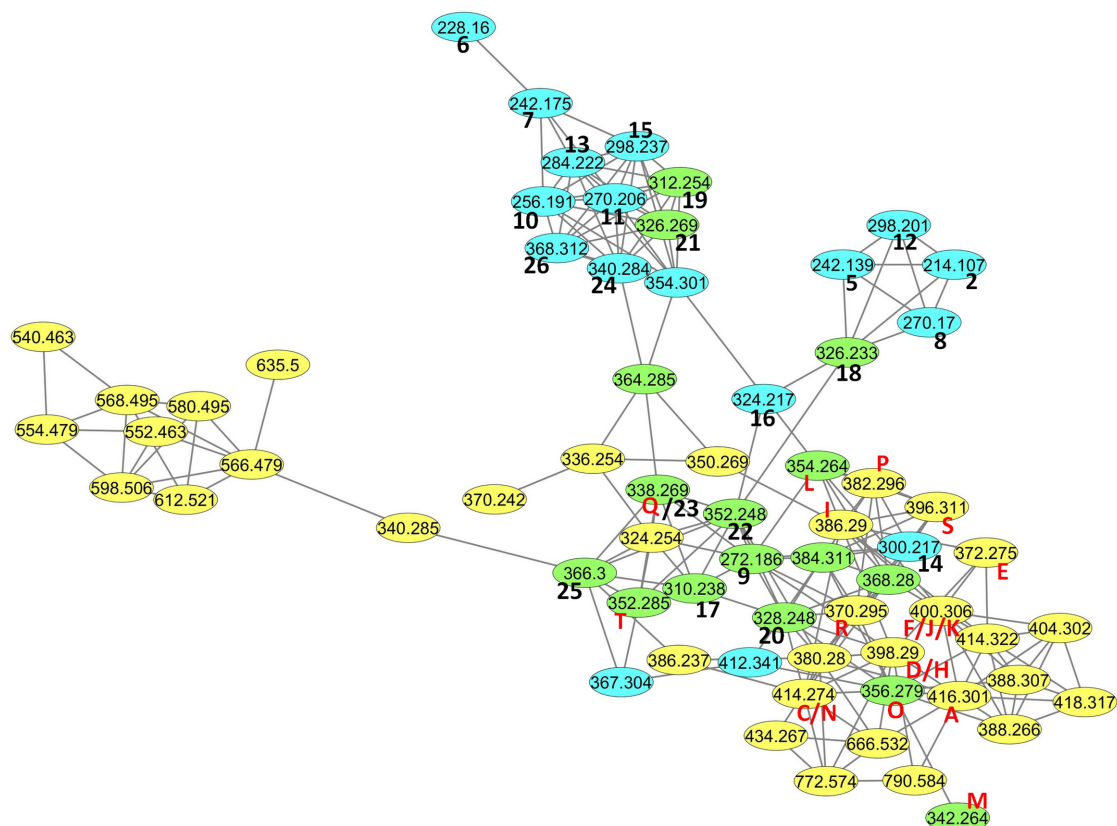
## Presence and Chemical Features of AHLs in Strain MOLA 401

MOLA 401 culture supernatant was extracted with *tert*-butyl methyl ether and fractionated in 22 fractions, M1–M22. These fractions were tested for AHL production using the biosensor strains *E. coli* MT102 and *P. putida* F117, which are GFP-based biosensors emitting light in presence of AHLs. Interestingly, 10 fractions were positive in these assays with both biosensors (M9, M10, M11, M12, M13, M14, M15, M16, M17, M18) (Supplementary Figures S1, S2). For detection of AHLs, we initially focused on LC-MS profiling with single ion recording at *m/z* 102, which corresponds to the mass of the protonated homoserine moiety. MS ionization conditions were optimized in order to favor the formation of this fragment. Then the full MS scan at the RTs pointed out in SIR102 allowed us to propose a list of pseudomolecular ion masses of putative AHLs. UHPLC coupled to high resolution MS/MS analyses were then conducted in the discovery mode. This allowed us to calculate the molecular formulas of the putative AHLs based high resolution masses, to obtain high resolution fragmentation analyses, and to obtain a molecular network including all the microfractions along with the 26 AHLs standards (**Figure 1**; Wang et al., 2016). All MS spectrum, SIR102 chromatograms, TIC chromatograms, MS spectrum, High Resolution MS spectrum are provided in Supplementary Figures S4–S61.

Overall, it was demonstrated that MOLA 401 produced at least 20 different AHLs out of the 21 putative ones detected by SIR102 (**Table 1**). The confirmation of the presence of the HSL subunit was obtained by the method described by Patel et al (Patel et al., 2016). In our case, all the AHLs had all 4 diagnostic fragments at *m/z* 102.055, 84.045, 74.061, and 56.050 in MS/MS. Also, compound **Q** was identical to standard **23** [*N*-hexadec-11(*Z*)-enoyl-L-homoserine lactone]. The molecular network shown in **Figure 1** further demonstrated that all MOLA 401 compounds identified in this study as potential AHLs clustered with the network defined by the standards. Much to our surprise, the molecular networking analysis uncovered AHLs in the strain although the detection limit was set to 1000, making the cluster much simpler (note that only **B** was not detected with this value). For the present article, restrained to the ones found in the SIR102 analysis, but many AHLs detected in the cluster did in fact present HSL diagnostic ions.

Eventually, it turned out that microfraction M17 essentially contained AHLs (**P**, **Q**, **R**). This fraction was analyzed by 1D and 2D NMR (**Table 2**). The <sup>1</sup>H NMR spectrum showed the presence of a methylene at  $\delta_H$  2.36 (m, 1H, 4a) and  $\delta_H$  2.11 (m, 1H, 4b), an oxomethylene at  $\delta_H$  4.33 (td, *J* = 8.9, 1.8, 1H, 5b) et  $\delta_H$  4.20 (m, 1H, 5a), and a methyne at  $\delta_H$  4.56 (m, 1H, 3). Long-range <sup>1</sup>H-<sup>13</sup>C correlations between H-3/H-5a and carbonyl C-2 at  $\delta_C$  175.3, as well as the sequence of

<sup>3</sup><https://gnps.ucsd.edu/>



**FIGURE 1 |** Molecular network containing the AHLs detected in strain MOLA401 microfractions (yellow), and in analytical standards (blue). Identical matches are in green. In the graph, the parent mass is reported on the node, and for each node the corresponding table entry codes are reported if applicable. Entry numbers for Supplementary Table S1 are in black, and **Table 1** entry letters are in red.

COZY correlations between H-3, H-4, and H-5 confirmed the presence of a lactone ring (**Figure 2**). Then, the  $^1\text{H}$ - $^{13}\text{C}$  HMBC correlations of amide proton at  $\delta_{\text{H}}$  8.29 (td,  $J = 8.9, 1.8, 1\text{H}$ ) with carbons C-3 ( $\delta_{\text{C}}$  47.4) and C-2' ( $\delta_{\text{C}}$  170.8) allowed us to position an acylamino group in C-3, therefore confirming that the 3 major compounds of fraction M17 were AHLs. For compounds with a hydroxyl group on the side chain (**P**, **R**), it was possible to ascertain the  $\text{CH}_2\text{CH}(\text{OH})\text{CH}_2$  partial sequence based on  $^1\text{H}$ - $^1\text{H}$  correlations of the oxomethine at  $\delta_{\text{H}}$  3.78 (m, 1H, 4') with the methylenes at  $\delta_{\text{H}}$  2.19 (m, 2H, 3') and at  $\delta_{\text{H}}$  1.37 (m, 1H, 5'a) /  $\delta_{\text{H}}$  1.30 (m, 1H, 5'b), and based on the long-range  $^1\text{H}$ - $^{13}\text{C}$  correlation of H-3' with C-2'. The rest of the side chain cannot be attributed due to extensive overlapping of the signals. Nevertheless, vinyl protons give key information on the double bonds in **P** and **Q**. The attribution of protons and carbons a-d in fragment B based on HSQC and COZY experiments was straightforward (**Figure 2**). Fragment B was constituted of two methylenes at  $\delta_{\text{H}}$  1.98 (m, 4H, a and d) and two vinyl protons at  $\delta_{\text{H}}$  5.32 (m, 2H, b and d). The shape of the vinyl protons signal confirmed the Z configuration of the double bond in **P** and **Q** (Frost and Gunstone, 1975). All NMR spectrum are provided in Supplementary Figures S62–S66.

To our knowledge, 20 or more AHLs is the highest diversity of AHLs reported to be produced by a single strain. Others studies on soil bacterial strains have detected up to five AHLs produced by only one strain. For example, *Sinorhizobium meliloti* produces C16-HSL, 3-oxo-C14-HSL, C16:1-HSL, 3-oxo-C16-HSL and 3-oxo-C16:1-HSL (Gao et al., 2005), and *Azospirillum lipoferum* TVV3, synthesizes C8-HSL, 3-oxo-C8-HSL, 3-oxo-C10-HSL, 3-OH-C10-HSL, and 3-oxo-C10-HSL (Boyer et al., 2008). A recent study revealed 7 AHLs produced by the marine strain *P. gallaeciensis* isolated on the surface of the algae *Sargassum muticum* (C14:1-HSL, C14:2-HSL, C16:1-HSL, C16:2-HSL, C18:1-HSL, 2,11-C18:2-HSL, C18:2-HSL) (Ziesche et al., 2015).

The length of acyl chains in the detected AHLs ranged between 15 and 19 carbons. To our knowledge, this also the first report of acyl chains longer than 18 carbons. Also, 5 AHLs presented an odd number of carbons in their acyl side chain (**Table 1**). This observation also constitutes an interesting feature, as very few AHLs with acyl side chain presenting an odd number of carbons have been previously identified. More frequently, such AHLs were present as trace elements (C13:0-HSL, C15:0-HSL, C15:1-HSL, C15:2-HSL) (Wagner-Döbler et al., 2005), except in *Sulfitobacter* sp. D13 where the 9-C17:1-HSL is an AHL

**TABLE 1 |** List of AHLs detected in the microfractions of the strain MOLA401.

Entry	Fraction number	Acyl chain length	Acyl chain unsaturations	LC/MS retention time (min)	LC-HRMS retention time (min)	Experi-mental m/z [M+H] <sup>+</sup>	Calculated molecular formula	Calculated m/z [M+H] <sup>+</sup>
A	M9	C18	0	3.59	6.07	416.3009	C <sub>22</sub> H <sub>41</sub> NO <sub>6</sub>	416.3007
B	M10	C18	2	3.66	6.36	412.2697	C <sub>22</sub> H <sub>37</sub> NO <sub>6</sub>	412.2694
C	M11	C18	1	4.08	6.60	414.2854	C <sub>22</sub> H <sub>39</sub> NO <sub>6</sub>	414.2850
D	M11	C18	1	4.10	6.51	398.2905	C <sub>22</sub> H <sub>39</sub> NO <sub>5</sub>	398.2901
E	M11	C16	0	4.16	6.57	372.2745	C <sub>20</sub> H <sub>37</sub> NO <sub>5</sub>	372.2744
F	M12	C18	0	4.22	6.84	400.3057	C <sub>22</sub> H <sub>41</sub> NO <sub>5</sub>	400.3057
G	M12	C19	1	4.36	6.93	428.3013	C <sub>23</sub> H <sub>41</sub> NO <sub>6</sub>	428.3007
H	M12	C18	1	4.42	6.87	398.2902	C <sub>22</sub> H <sub>39</sub> NO <sub>5</sub>	398.2901
I	M13	C17	0	4.49	6.91	386.2904	C <sub>21</sub> H <sub>39</sub> NO <sub>5</sub>	386.2901
J	M13	C18	0	4.61	7.08	400.0366	C <sub>22</sub> H <sub>41</sub> NO <sub>5</sub>	400.3057
K	M13	C18	0	4.70	7.15	400.3061	C <sub>22</sub> H <sub>41</sub> NO <sub>5</sub>	400.3057
L	M15	C16	1	5.02	7.61	354.2642	C <sub>20</sub> H <sub>35</sub> NO <sub>4</sub>	354.2639
M	M15	C15	0	5.14	7.72	342.2645	C <sub>19</sub> H <sub>35</sub> NO <sub>4</sub>	342.2639
N	M16	C16	1	5.38	8.13	414.2676	C <sub>20</sub> H <sub>39</sub> NO <sub>4</sub> S	414.2673
O	M16	C16	0	5.49	8.18	356.2798	C <sub>20</sub> H <sub>37</sub> NO <sub>4</sub>	356.2795
P	M16/M17	C18	1	5.62	8.46	382.2954	C <sub>22</sub> H <sub>39</sub> NO <sub>4</sub>	382.2952
Q	M17	C16	1	5.71	8.68	338.2690	C <sub>20</sub> H <sub>35</sub> NO <sub>3</sub>	338.2690
R	M17/M18	C17	0	5.79	8.67	370.2952	C <sub>21</sub> H <sub>39</sub> NO <sub>4</sub>	370.2952
S	M18	C19	1	5.93	8.96	396.3113	C <sub>23</sub> H <sub>41</sub> NO <sub>4</sub>	396.3108
T	M18	C17	1	6.02	9.19	352.2846	C <sub>21</sub> H <sub>38</sub> NO <sub>3</sub>	352.2846

which appears synthesized in large quantities (Ziesche et al., 2015).

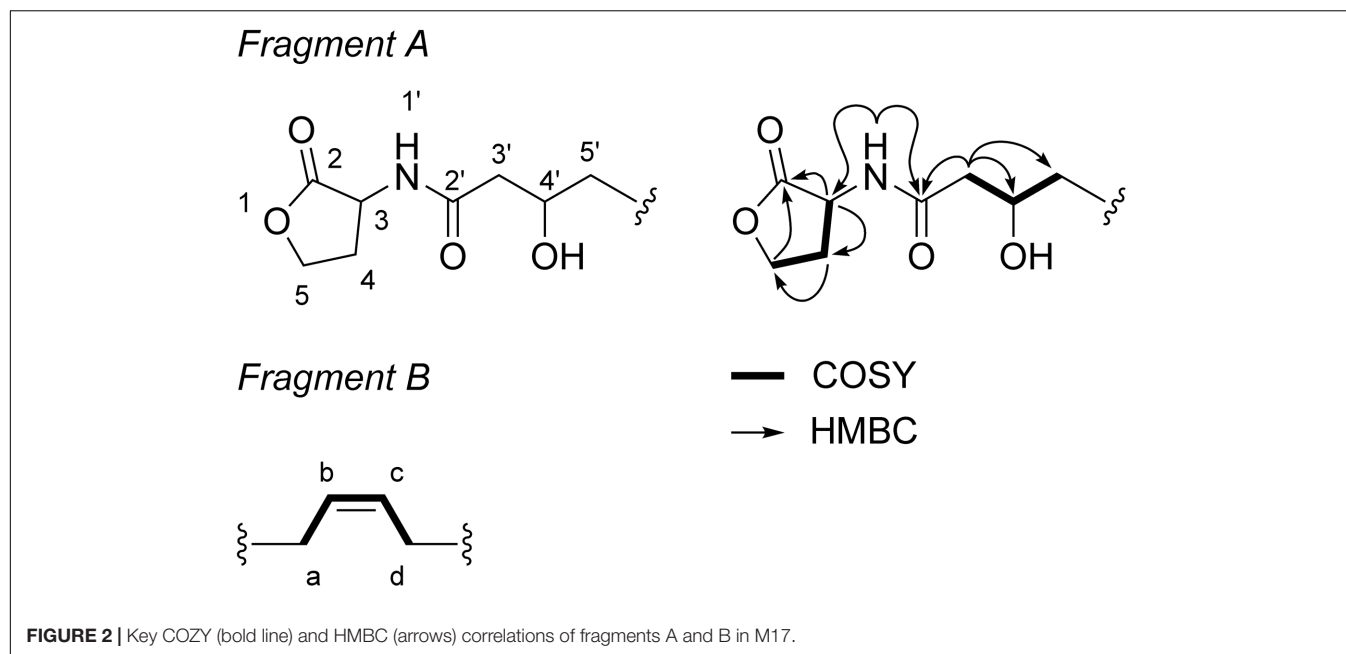
We also detected at least 6 AHLs with two or three hydroxyl groups along the acyl side chain. An examination of previously characterized AHLs revealed only single hydroxylation per acyl chain (Churchill and Chen, 2011) located at C-3.

**TABLE 2 |** <sup>1</sup>H and <sup>13</sup>C data for fragment A (recorded at 600 MHz and 150 MHz in DMSO-*d*<sub>6</sub>, respectively).

Atom	δ <sub>C</sub>	δ <sub>H</sub> (J in Hz)	COZY	HMBC
<b>Fragment A</b>				
2	175.3			
3	47.4	4.56, m	4, 1'	2, 4
4a	27.9	2.36, m	5a, 5b	
4b		2.11, m	5a, 5b	5
5a	64.9	4.33, dd (8.9, 1.8)	4a, 4b	2, 3, 5
5b		4.20, m	4a, 4b	2, 3, 5
1'		8.29, d (8.0)	3	3, 2'
2'	170.8			
3'	43.5	2.19, m	4'	2', 4', 5'
4'	67.0	3.78, bs	3'	
5'a	36.5	1.37, m	3'	
5'b		1.30, m	3'	
<b>Fragment B</b>				
a	26.4	1.98, m	b	a
b	129.3	5.32, m	a, c	b
c	129.3	5.32, m	b, d	c
d	26.4	1.98, m	c	d

This is the case for the AHL detected in the marine *Roseobacter* strains *Phaeobacter* sp. BS107 or *Loktanella* sp. F14 who produces 3-OH-C12:1-HSL (Ziesche et al., 2015). Thus, we report here another interesting new feature of marine AHLs, which is the existence of poly-hydroxylation of the acyl chain (Table 1). The position of the hydroxyl groups along the acyl chain could not be determined as these groups did not induce fragmentation of the side chain in MS/MS. NMR of the microfractions were very difficult to interpret due to the relatively low proportion of each AHL in these fractions. However, despite these limitations, our data unambiguously indicate that the strain MOLA 401 is able to synthesize a wide diversity of AHLs. Also, we detected at least 2 AHLs presenting one double bond in their acyl side chain (Table 1). The position and configuration of the double bound chain was confirmed in compound Q by the analytical standard 23. When there was oxygen and double bonds detected in the side chain, it was not possible to distinguish a carbonyl group or a hydroxyl and a carbon-carbon double bond, as the two would lead to the same molecular formula.

Short acyl chain molecules are more polar and soluble in seawater than those presenting long aliphatic chains, which are thus less hydrophilic. However, it appears that marine bacteria produce AHLs with long chains (Wagner-Döbler et al., 2005; Zan et al., 2012). Thus, our data confirm these previous observations. Also, our technical approach revealed that many AHL acyl chains were oxidized. Such observation indicates that these AHLs are adapted for signal release and diffusion in marine environments as acyl side chain modifications would increase water solubility and compatibility with active efflux pumps (Pearson et al., 1999).



Most of *Rhodobacteraceae* bacteria produce long chain AHLs with additional modifications (Cude and Buchan, 2013). For example, the marine free-living strain *Rhodobacter sphaeroides* produces C14:1-HSL (Puskas et al., 1997), the marine dinoflagellate associated bacterium *Dinoroseobacter shibae* synthesizes mainly C18:2-HSL and C18:1-HSL, but also traces of C16-HSL, C15-HSL and C14-HSL (Wagner-Döbler et al., 2005; Neumann et al., 2013; Patzelt et al., 2013), the sponge symbiont *Ruegeria* sp. emits OH-C14-HSL, OH-C14:1-HSL and OH-C12-HSL (Zan et al., 2012). *S. pomeroyi* produces the *p*-coumaroyl-HSL, a non-conventional AHL in which the acyl side chain is replaced by a coumaroyl moiety (Schaefer et al., 2008). Nevertheless, the poly-hydroxylation of acyl chain observed in strain MOLA 401 combined with the presence of unsaturation appears to be an original feature. We hypothesize that the AHL synthase produces a molecule with acyl chain containing 15 to 19 carbons, and that additional modifications of the acyl chain are mediated by cytochrome P450 (Chowdhary et al., 2007) (WP\_036181863.1), which oxidizes aliphatic chains, and by desaturases which produce double bonds (Aguilar and de Mendoza, 2006). Interestingly, we detected a cytochrome P450 homolog in the genome of the strain MOLA 401 (Doberva et al., 2014).

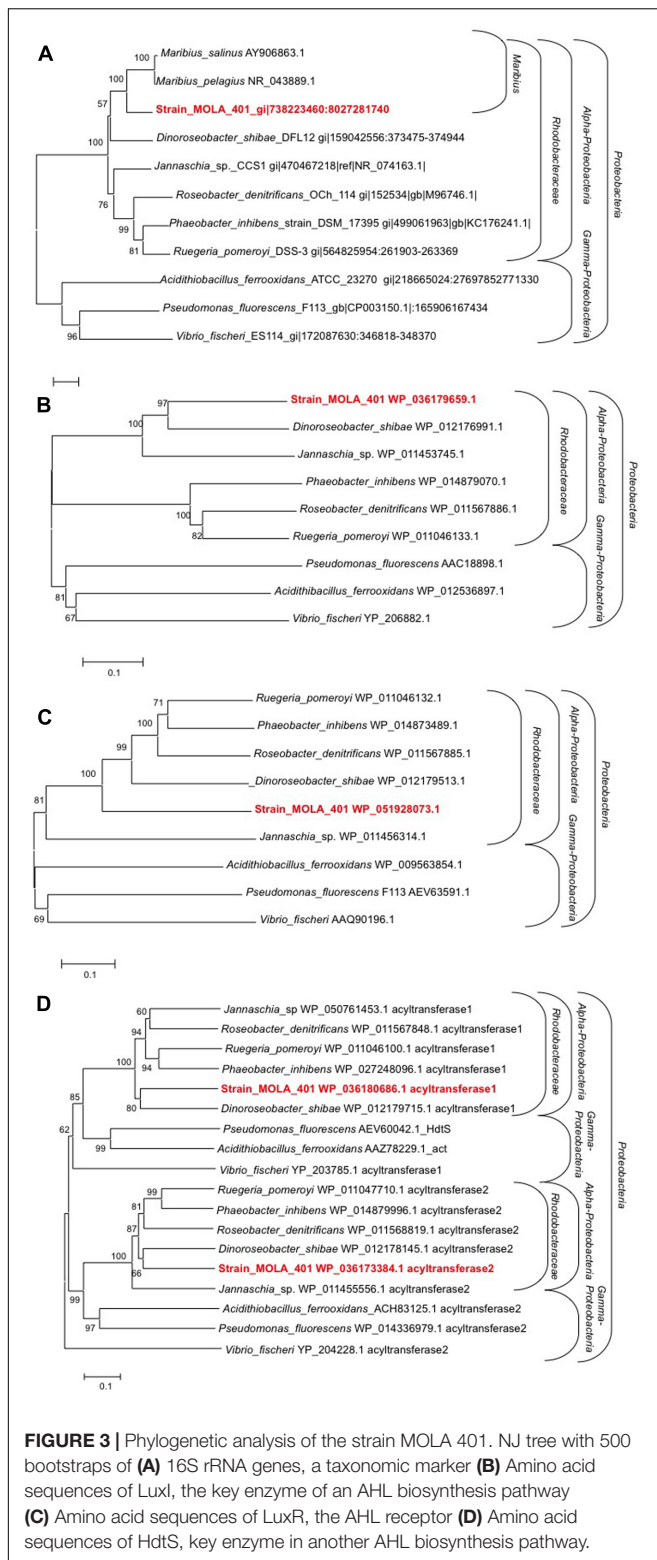
## Linking Genetic and Chemical Features

Phylogenetic analyses based on 16S rRNA and putative LuxI sequences confirm the position of strain MOLA 401 in the *Rhodobacteraceae* family and the *Proteobacteria* phylum (Figure 3A). The position of this strain, close to two *Maribius* isolates was well supported (BP<sub>NJ</sub> = 100) (Figure 3A) and confirmed affiliation to the *Rhodobacteraceae* family. The strain MOLA 401 putative LuxI protein sequence is closely related to other LuxI sequences of *Rhodobacteraceae* strains within *Alphaproteobacteria* (Figure 3B). Clustering of the

*Rhodobacteraceae* LuxI sequences (group 1 includes *Ruegeria pomeroyi*, *Roseobacter denitrificans*, *P. inhibens*; group 2 includes *D. shibae*, *Maribius* sp., *Jannaschia* sp.) were well supported (Figure 3B). Similarly, the phylogenetic tree based on the AHL receptor LuxR placed the strain MOLA 401 putative LuxR within the *Rhodobacteraceae* (Figure 3C). These data clearly indicate that the strain MOLA 401 strain belongs to the *Roseobacter* group with respect to its 16S rRNA or the genes encoding for AHL production and reception. This makes strain MOLA 401 an ideal model strain for future studies of QS in marine environments. Also these data confirmed previous observation based only on 16S rRNA genes (Choi et al., 2007).

Another protein potentially involved in AHL production is HdtS, of which two homologs have been detected in the full genome sequence of the strain MOLA 401 (Doberva et al., 2014). HdtS is a member of the lysophosphatidic acid acyltransferase family (Laue et al., 2000) and has a dual functionality, acylation of lysophosphatidic acid (Cullinane et al., 2005) and AHL synthesis (Laue et al., 2000). The HdtS-mediated production of AHL has been demonstrated experimentally in *P. fluorescens* (Laue et al., 2000) and *Acidithiobacillus ferrooxidans* (Rivas et al., 2007). *P. fluorescens* produces 3-OH-C14:1-HSL, C10-HSL and C6-HSL, while *A. ferrooxidans* produces a C14-HSL. The strain MOLA 401 putative HdtS sequences clustered into two groups, both of which were related to putative HdtS found in other *Rhodobacteraceae*, with strong bootstrap supports (Figure 3D). Interestingly, one homolog was clustered with HdtS from the Gammaproteobacteria *A. ferrooxidans* and *P. fluorescens*, the only HdtS enzymes with confirmed AHL synthesis activity (acyltransferase2 sequences, Figure 3D). This suggests that the MOLA 401 putative HdtS is similarly contributing to the AHL pool produced by MOLA 401, in cooperation with putative LuxI. However, an experimental confirmation of such HdtS





based AHL production in strain MOLA 401 is required in future studies, also because the MOLA 401 strain does not produce similar AHL as those found in *A. ferrooxidans* and *P. fluorescens*.

The specificity of LuxI synthases varies, especially in regards to the type of acyl side chain recognized as substrate (Gould et al., 2004). For example, the LasI (a LuxI homolog) in *P. aeruginosa* produces different AHLs depending on the growth conditions and the host. By contrast, YspI and EsaI, respectively, found in *Yersinia pestis* and *Erwinia stewartii*, are specific to one type of acyl-ACP (Gould et al., 2006) producing defined AHLs. The strain MOLA401 putative LuxI synthase has the conserved the arginine and phenylalanine in positions 25 and 29 (two key aminoacids residues in this protein), respectively, similar to the *P. aeruginosa* LuxI (Supplementary Figure S3). Thus MOLA 401 putative LuxI sequence is consistent with a capacity to produce a large number of AHLs. One possible hypothesis is that the same LuxI-synthase may produce several AHLs with low side chain length specificity, as demonstrated by Neumann et al. (2013).

## Culture of Strain MOLA 401 and QS Abilities

The strain MOLA 401 is a bacterium from the *Rhodobacteraceae* family isolated in an oligotrophic lagoon. Phylogenetically close *Maribius* strains have also been isolated in such oligotrophic waters, like in the Sargasso Sea (Choi et al., 2007). The ability of bacteria isolated from oligotrophic waters to communicate could appear paradoxical (Moran et al., 2004) as cell densities in such environments are below the expected threshold that enables QS. However, our study demonstrates the ability of the strain MOLA 401 to synthesize diverse types of AHLs. We experimented on a MOLA 401 strain cultured under rich nutrient conditions (Marine Broth). Thus, we could suggest that the large spectrum of AHL produced by MOLA 401 might give this strain the ability to exploit organic matter by a complex coordination of the bacterial population (Rolland et al., 2016). This observation is in line with previous hypothesis suggesting that such coordination allows particle-attached bacteria to exploit marine organic matter (Moran et al., 2016). Future studies need to be conducted to evaluate the capacity of *Rhodobacteraceae* to produce AHLs when cultured in oligotrophic media.

Collectively, our technical approach based on a bioguided search of AHL in bacterial extracts and the obtained data reveal that the *Rhodobacteraceae* strain MOLA 401 isolated in an oligotrophic lagoon is able to produce a very large number of different AHLs. The AHLs characterized in this study possessed interesting and original features including variable acyl chain length and multiple-hydroxylation sites. The strain MOLA 401 strain provides new insights into the breadth of possible AHL diversity, suggesting the existence of original adaptations of bacterial dialogs to marine environments.

## AUTHOR CONTRIBUTIONS

MD, DS, JS, NH, SS-F, VE, YF, and RL conducted the experimental work. MD, DS, PL, SS-F, and RL designed the experiments. All authors wrote the manuscript.

## FUNDING

This work was supported by Emergence UPMC, CNRS-EC2CO, and SECIL ANR-15-CE21-0016 grants.

## ACKNOWLEDGMENTS

We thank Sarah Bennai and Laurent Intertaglia for technical help. We thank Fabien Joux for providing strain MOLA 401.

## REFERENCES

- Aguilar, P. S., and de Mendoza, D. (2006). Control of fatty acid desaturation: a mechanism conserved from bacteria to humans. *Mol. Microbiol.* 62, 1507–1514. doi: 10.1111/j.1365-2958.2006.05484.x
- Amin, S., Hmelo, L., Van Tol, H., Durham, B., Carlson, L., Heal, K., et al. (2015). Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature* 522, 98–101. doi: 10.1038/nature14488
- Andersen, J. B., Heydorn, A., Hentzer, M., Eberl, L., Geisenberger, O., Christensen, B. B., et al. (2001). gfp-based N-acyl homoserine-lactone sensor systems for detection of bacterial communication. *Appl. Environ. Microbiol.* 67, 575–585. doi: 10.1128/AEM.67.2.575-585.2001
- Atkinson, S., and Williams, P. (2009). Quorum sensing and social networking in the microbial world. *J. R. Soc. Interface* 6, 959–978. doi: 10.1098/rsif.2009.0203
- Bassler, B. L. (1999). How bacteria talk to each other: regulation of gene expression by quorum sensing. *Curr. Opin. Microbiol.* 2, 582–587. doi: 10.1016/S1369-5274(99)00025-9
- Berger, M., Neumann, A., Schulz, S., Simon, M., and Brinkhoff, T. (2011). Tropodithietic acid production in *Phaeobacter gallaeciensis* is regulated by N-acyl homoserine lactone-mediated quorum sensing. *J. Bacteriol.* 193, 6576–6585. doi: 10.1128/JB.05818-11
- Boyer, M., Bally, R., Perrotto, S., Chaintreuil, C., and Wisniewski-Dye, F. (2008). A quorum-quenching approach to identify quorum-sensing-regulated functions in *Azospirillum lipoferum*. *Res. Microbiol.* 159, 699–708. doi: 10.1016/j.resmic.2008.08.003
- Case, R. J., Labbate, M., and Kjelleberg, S. (2008). AHL-driven quorum-sensing circuits: their frequency and function among the Proteobacteria. *ISME J.* 2, 345. doi: 10.1038/ismej.2008.13
- Cha, C., Gao, P., Chen, Y. C., Shaw, P. D., and Farrand, S. K. (1998). Production of acyl-homoserine lactone quorum-sensing signals by gram-negative plant-associated bacteria. *MPMI* 11, 1119–1129. doi: 10.1094/MPMI.1998.11.11.1119
- Choi, D. H., Cho, J. C., Lanoil, B. D., Giovannoni, S. J., and Cho, B. C. (2007). *Maribius salinus* gen. nov., sp. nov., isolated from a solar saltern and *Maribius pelagius* sp. nov., cultured from the Sargasso Sea, belonging to the *Roseobacter* clade. *Int. J. Syst. Evol. Microbiol.* 57, 270–275. doi: 10.1099/ijs.0.64552-0
- Chowdhary, P. K., Keshavan, N., Nguyen, H. Q., Peterson, J. A., Gonzalez, J. E., and Haines, D. C. (2007). *Bacillus megaterium* CYP102A1 oxidation of acyl homoserine lactones and acyl homoserines. *Biochemistry* 46, 14429–14437. doi: 10.1021/bi701945j
- Churchill, M. E., and Chen, L. (2011). Structural basis of acyl-homoserine lactone-dependent signaling. *Chem. Rev.* 111, 68–85. doi: 10.1021/cr1000817
- Conan, P., Joux, F., Torretton, J. P., Pujo-Pay, M., Douki, T., Rochelle-Newall, E., et al. (2008). Effect of solar ultraviolet radiation on bacterio- and phytoplankton activity in a large coral reef lagoon (southwest New Caledonia). *Aquat. Microb. Ecol.* 52, 83–98. doi: 10.3354/ame01204
- Cude, W. N., and Buchan, A. (2013). Acyl-homoserine lactone-based quorum sensing in the *Roseobacter* clade: complex cell-to-cell communication controls multiple physiologies. *Front. Microbiol.* 4:336. doi: 10.3389/fmicb.2013.00336
- Cullinane, M., Baysse, C., Morrissey, J. P., and O'gara, F. (2005). Identification of two lysophosphatidic acid acyltransferase genes with overlapping function in *Pseudomonas fluorescens*. *Microbiology* 151, 3071–3080. doi: 10.1099/mic.0.27958-0
- Dickschat, J. S. (2010). Quorum sensing and bacterial biofilms. *Nat. Prod. Rep.* 27, 343–369. doi: 10.1039/b804469b
- Diggle, S. P., Cruz, S. A., and Cámara, M. (2007). Quorum sensing. *Curr. Biol.* 17, R907–R910. doi: 10.1016/j.cub.2007.08.045
- Doberva, M., Sanchez-Ferandin, S., Ferandin, Y., Intertaglia, L., Joux, F., Lebaron, P., et al. (2014). Genome sequence of *Maribius* sp. strain MOLA 401, a marine *Roseobacter* with a quorum-sensing cell-dependent physiology. *Genome Announc.* 2, e00997–14. doi: 10.1128/genomeA.00997-14
- Doberva, M., Sanchez-Ferandin, S., Toulza, E., Lebaron, P., and Lami, R. (2015). Diversity of quorum sensing autoinducer synthases in the Global Ocean Sampling metagenomic database. *Aquat. Microb. Ecol.* 74, 107–119. doi: 10.3354/ame01734
- Engelbrecht, J., and Silverman, M. (1984). Identification of genes and gene products necessary for bacterial bioluminescence. *Proc. Natl. Acad. Sci. U.S.A.* 81, 4154–4158. doi: 10.1073/pnas.81.13.4154
- Frost, D. J., and Gunstone, F. D. (1975). The PMR analysis of non-conjugated alkenoic and alkyenoic acids and esters. *Chem. Phys. Lipids* 15, 53–85. doi: 10.1016/0009-3084(75)90032-8
- Fuqua, C., and Greenberg, E. P. (2002). Listening in on bacteria: acyl-homoserine lactone signalling. *Nat. Rev. Mol. Cell Biol.* 3, 685–695. doi: 10.1038/nrm907
- Fuqua, W. C., Winans, S. C., and Greenberg, E. P. (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J. Bacteriol.* 176, 269–275. doi: 10.1128/jb.176.2.269-275.1994
- Gao, M., Chen, H., Eberhard, A., Gronquist, M. R., Robinson, J. B., Rolfe, B. G., et al. (2005). *sinI* and *expR*-dependent quorum sensing in *Sinorhizobium meliloti*. *J. Bacteriol.* 187, 7931–7944. doi: 10.1128/JB.187.23.7931-7944.2005
- Gelencser, Z., Choudhary, K. S., Coutinho, B. G., Hudaiberdiev, S., Galbats, B., Venturi, V., et al. (2012). Classifying the topology of AHL-driven quorum sensing circuits in proteobacterial genomes. *Sensors* 12, 5432–5444. doi: 10.3390/s120505432
- Gilson, L., Kuo, A., and Dunlap, P. V. (1995). *AinS* and a new family of autoinducer synthesis proteins. *J. Bacteriol.* 177, 6946–6951. doi: 10.1128/jb.177.23.6946-6951.1995
- Gould, T. A., Herman, J., Krank, J., Murphy, R. C., and Churchill, M. E. (2006). Specificity of acyl-homoserine lactone synthases examined by mass spectrometry. *J. Bacteriol.* 188, 773–783. doi: 10.1128/JB.188.2.773-783.2006
- Gould, T. A., Schweizer, H. P., and Churchill, M. E. (2004). Structure of the *Pseudomonas aeruginosa* acyl-homoserine lactone synthase *LasI*. *Mol. Microbiol.* 53, 1135–1146. doi: 10.1111/j.1365-2958.2004.04211.x
- Gram, L., Grossart, H. P., Schlingloff, A., and Kiorboe, T. (2002). Possible quorum sensing in marine snow bacteria: production of acylated homoserine lactones by *Roseobacter* strains isolated from marine snow. *Appl. Environ. Microbiol.* 68, 4111–4116. doi: 10.1128/AEM.68.8.4111-4116.2002
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404
- Laue, B. E., Jiang, Y., Chhabra, S. R., Jacob, S., Stewart, G. S., Hardman, A., et al. (2000). The biocontrol strain *Pseudomonas fluorescens* F113 produces the Rhizobium small bacteriocin, N-(3-hydroxy-7-cis-tetradecenoyl)-homoserine lactone, via HdtS, a putative novel N-acyl-homoserine lactone synthase. *Microbiology* 146, 2469–2480. doi: 10.1099/00221287-146-10-2469
- Limardo, A. J., and Worden, A. Z. (2015). Microbiology: exclusive networks in the sea. *Nature* 522, 36–37. doi: 10.1038/nature14530
- Loh, J., Pierson, E. A., Pierson, L. S. III, Stacey, G., and Chatterjee, A. (2002). Quorum sensing in plant-associated bacteria. *Curr. Opin. Plant Biol.* 5, 285–290. doi: 10.1016/S1369-5266(02)00274-1

- Miller, M. B., and Bassler, B. L. (2001). Quorum sensing in bacteria. *Annu. Rev. Microbiol.* 55, 165–199. doi: 10.1146/annurev.micro.55.1.165
- Mohamed, N. M., Cicirelli, E. M., Kan, J., Chen, F., Fuqua, C., and Hill, R. T. (2008). Diversity and quorum-sensing signal production of *Proteobacteria* associated with marine sponges. *Environ. Microbiol.* 10, 75–86. doi: 10.1111/j.1462-2920.2007.01431.x
- Moran, M. A., Buchan, A., Gonzalez, J. M., Heidelberg, J. F., Whitman, W. B., Kiene, R. P., et al. (2004). Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* 432, 910–913. doi: 10.1038/nature03170
- Moran, M. A., Kujawinski, E. B., Stubbins, A., Fatland, R., Aluwihare, L. I., Buchan, A., et al. (2016). Deciphering ocean carbon in a changing world. *Proc. Natl. Acad. Sci. U.S.A.* 113, 3143–3151. doi: 10.1073/pnas.1514645113
- Nealson, K. H. (1977). Autoinduction of bacterial luciferase. *Arch. Microbiol.* 112, 73–79. doi: 10.1007/BF00446657
- Neumann, A., Patzelt, D., Wagner-Döbler, I., and Schulz, S. (2013). Identification of new N-acylhomoserine lactone signalling compounds of *Dinoroseobacter shibae* DFL-12T by overexpression of luxI Genes. *Chembiochem* 14, 2355–2361. doi: 10.1002/cbic.201300424
- Papenfert, K., and Bassler, B. L. (2016). Quorum sensing signal-response systems in Gram-negative bacteria. *Nat. Rev. Microbiol.* 14, 576–588. doi: 10.1038/nrmicro.2016.89
- Parsek, M. R., and Greenberg, E. (2005). Sociomicrobiology: the connections between quorum sensing and biofilms. *Trends Microbiol.* 13, 27–33. doi: 10.1016/j.tim.2004.11.007
- Patel, N. M., Moore, J. D., Blackwell, H. E., and Amador-Noguez, D. (2016). Identification of Unanticipated and Novel N-Acyl L-Homoserine Lactones (AHLs) using a sensitive non-targeted LC-MS/MS method. *PLoS ONE* 11:e0163469. doi: 10.1371/journal.pone.0163469
- Patzelt, D., Wang, H., Buchholz, I., Rohde, M., Gröbe, L., Pradella, S., et al. (2013). You are what you talk: quorum sensing induces individual morphologies and cell division modes in *Dinoroseobacter shibae*. *ISME J.* 7, 2274–2286. doi: 10.1038/ismej.2013.107
- Pearson, J. P., Van Delden, C., and Iglewski, B. H. (1999). Active efflux and diffusion are involved in transport of *Pseudomonas aeruginosa* cell-to-cell signals. *J. Bacteriol.* 181, 1203–1210.
- Puskas, A., Greenberg, E. P., Kaplan, S., and Schaefer, A. L. (1997). A quorum-sensing system in the free-living photosynthetic bacterium *Rhodobacter sphaeroides*. *J. Bacteriol.* 179, 7530–7537. doi: 10.1128/jb.179.23.7530-7537.1997
- Riedel, K., Hentzer, M., Geisenberger, O., Huber, B., Steidle, A., Wu, H., et al. (2001). N-acylhomoserine-lactone-mediated communication between *Pseudomonas aeruginosa* and *Burkholderia cepacia* in mixed biofilms. *Microbiology* 147, 3249–3262. doi: 10.1099/00221287-147-12-3249
- Rivas, M., Seeger, M., Jedlicki, E., and Holmes, D. S. (2007). Second acyl homoserine lactone production system in the extreme acidophile *Acidithiobacillus ferrooxidans*. *Appl. Environ. Microbiol.* 73, 3225–3231. doi: 10.1128/AEM.02948-06
- Rolland, J. L., Stien, D., Sanchez-Ferandin, S., and Lami, R. (2016). Quorum sensing and quorum quenching in the phycosphere of phytoplankton: a case of chemical interactions in Ecology. *J. Chem. Ecol.* 42, 1201–1211. doi: 10.1007/s10886-016-0791-y
- Schaefer, A. L., Greenberg, E. P., Oliver, C. M., Oda, Y., Huang, J. J., Bittan-Banin, G., et al. (2008). A new class of homoserine lactone quorum-sensing signals. *Nature* 454, 595–599. doi: 10.1038/nature07088
- Schaefer, A. L., Taylor, T. A., Beatty, J. T., and Greenberg, E. P. (2002). Long-chain acyl-homoserine lactone quorum-sensing regulation of *Rhodobacter capsulatus* gene transfer agent production. *J. Bacteriol.* 184, 6515–6521. doi: 10.1128/JB.184.23.6515-6521.2002
- Smith, R. S., and Iglewski, B. H. (2003). *P. aeruginosa* quorum-sensing systems and virulence. *Curr. Opin. Microbiol.* 6, 56–60. doi: 10.1016/S1369-5274(03)00008-0
- Steindler, L., and Venturi, V. (2007). Detection of quorum-sensing N-acyl homoserine lactone signal molecules by bacterial biosensors. *FEMS Microbiol. Lett.* 266, 1–9. doi: 10.1111/j.1574-6968.2006.00501.x
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Thiel, V., Kunze, B., Verma, P., Wagner-Döbler, I., and Schulz, S. (2009). New structural variants of homoserine lactones in bacteria. *Chembiochem* 10, 1861–1868. doi: 10.1002/cbic.200900126
- Wagner-Döbler, I., and Biehl, H. (2006). Environmental biology of the marine Roseobacter lineage. *Annu. Rev. Microbiol.* 60, 255–280. doi: 10.1146/annurev.micro.60.080805.142115
- Wagner-Döbler, I., Thiel, V., Eberl, L., Allgaier, M., Bodor, A., Meyer, S., et al. (2005). Discovery of complex mixtures of novel long-chain quorum sensing signals in free-living and host-associated marine alphaproteobacteria. *Chembiochem* 6, 2195–2206. doi: 10.1002/cbic.200500189
- Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* 34, 828–837. doi: 10.1038/nbt.3597
- Waters, C. M., and Bassler, B. L. (2005). Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.* 21, 319–346. doi: 10.1146/annurev.cellbio.21.012704.131001
- Williams, P., Winzer, K., Chan, W. C., and Camara, M. (2007). Look who's talking: communication and quorum sensing in the bacterial world. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362, 1119–1134. doi: 10.1098/rstb.2007.2039
- Zan, J., Choi, O., Meharena, H., Uhls, C. L., Churchill, M. E., Hill, R. T., et al. (2015). A solo luxI-type gene directs acylhomoserine lactone synthesis and contributes to motility control in the marine sponge symbiont *Ruegeria* sp. KLH11. *Microbiology* 161, 50–56. doi: 10.1099/mic.0.083956-0
- Zan, J., Cicirelli, E. M., Mohamed, N. M., Sibhatu, H., Kroll, S., Choi, O., et al. (2012). A complex LuxR–LuxI type quorum sensing network in a roseobacterial marine sponge symbiont activates flagellar motility and inhibits biofilm formation. *Mol. Microbiol.* 85, 916–933. doi: 10.1111/j.1365-2958.2012.08149.x
- Zan, J., Liu, Y., Fuqua, C., and Hill, R. T. (2014). Acyl-homoserine lactone quorum sensing in the *Roseobacter* clade. *Int. J. Mol. Sci.* 15, 654–669. doi: 10.3390/ijms15010654
- Ziesche, L., Bruns, H., Dogs, M., Wolter, L., Mann, F., Wagner-Döbler, I., et al. (2015). Homoserine lactones, methyl oligohydroxybutyrates, and other extracellular metabolites of macroalgae-associated bacteria of the *Roseobacter* clade: identification and functions. *Chembiochem* 16, 2094–2107. doi: 10.1002/cbic.201500189

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Doberva, Stien, Sorres, Hue, Sanchez-Ferandin, Eparvier, Ferandin, Lebaron and Lami. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# The Composite 259-kb Plasmid of *Martellella mediterranea* DSM 17316<sup>T</sup>—A Natural Replicon with Functional RepABC Modules from *Rhodobacteraceae* and *Rhizobiaceae*

Pascal Bartling, Henner Brinkmann, Boyke Bunk, Jörg Overmann, Markus Göker and Jörn Petersen\*

Leibniz-Institute DSMZ—German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany

## OPEN ACCESS

### Edited by:

Bernd Wemheuer,  
University of New South Wales,  
Australia

### Reviewed by:

William Martin,  
University of Dusseldorf Medical  
School, Germany

Andrew W. B. Johnston,  
University of East Anglia,  
United Kingdom

### \*Correspondence:

Jörn Petersen  
joern.petersen@dsMZ.de

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 22 March 2017

**Accepted:** 05 September 2017

**Published:** 21 September 2017

### Citation:

Bartling P, Brinkmann H, Bunk B,  
Overmann J, Göker M and Petersen J  
(2017) The Composite 259-kb  
Plasmid of *Martellella mediterranea*  
DSM 17316<sup>T</sup>—A Natural Replicon with  
Functional RepABC Modules from  
*Rhodobacteraceae* and *Rhizobiaceae*.  
*Front. Microbiol.* 8:1787.  
doi: 10.3389/fmicb.2017.01787

A multipartite genome organization with a chromosome and many extrachromosomal replicons (ECRs) is characteristic for *Alphaproteobacteria*. The best investigated ECRs of terrestrial rhizobia are the symbiotic plasmids for legume root nodulation and the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens*. RepABC plasmids represent the most abundant alphaproteobacterial replicon type. The currently known homologous replication modules of rhizobia and *Rhodobacteraceae* are phylogenetically distinct. In this study, we surveyed type-strain genomes from the One Thousand Microbial Genomes (KMG-I) project and identified a roseobacter-specific RepABC-type operon in the draft genome of the marine rhizobium *Martellella mediterranea* DSM 17316<sup>T</sup>. PacBio genome sequencing demonstrated the presence of three circular ECRs with sizes of 593, 259, and 170-kb. The rhodobacterial RepABC module is located together with a rhizobial equivalent on the intermediate sized plasmid pMM259, which likely originated in the fusion of a pre-existing rhizobial ECR with a conjugated roseobacter plasmid. Further evidence for horizontal gene transfer (HGT) is given by the presence of a roseobacter-specific type IV secretion system on the 259-kb plasmid and the rhodobacteracean origin of 62% of the genes on this plasmid. Functionality tests documented that the genuine rhizobial RepABC module from the *Martellella* 259-kb plasmid is only maintained in *A. tumefaciens* C58 (*Rhizobiaceae*) but not in *Phaeobacter inhibens* DSM 17395 (*Rhodobacteraceae*). Unexpectedly, the roseobacter-like replication system is functional and stably maintained in both host strains, thus providing evidence for a broader host range than previously proposed. In conclusion, pMM259 is the first example of a natural plasmid that likely mediates genetic exchange between roseobacters and rhizobia.

**Keywords:** RepABC-type plasmids, compatibility, type IV secretion systems, plasmid fusion, comparative genomics, horizontal gene transfer



## INTRODUCTION

RepABC-type plasmids play a crucial role for the multipartite genome organization and the lifestyle of rhizobia (Pappas and Cevallos, 2011). Long-known examples are the pathogenic tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* and the symbiotic nodulation (pSym) plasmids of the genus *Rhizobium*. RepABC-type plasmids comprise up to 50% of the rhizobial genome and represent the by far most abundant replicon type of these soil bacteria (Pappas and Cevallos, 2011). *Rhizobium etli* CFN42 harbors eight RepABC operons that are located on six extrachromosomal replicons (ECRs; González et al., 2006). The relevance of ECRs for the marine roseobacter group (*Rhodobacteraceae*) is exemplified by photosynthesis, flagellar and biofilm plasmids (Petersen et al., 2012; Michael et al., 2016). Roseobacters contains at least four different plasmid types (RepA, RepB, DnaA-like, RepABC) with more than 20 compatibility groups (Petersen, 2011). Nine different compatibility groups of RepABC-type plasmids, which can stably coexist in the same cell, have been identified in this lineage (Petersen et al., 2009).

RepABC modules are specific for *Alphaproteobacteria* and contain three genes, the *repA* and *repB* partitioning genes as well as the replication gene *repC*, which are arranged in a characteristic operon (Pinto et al., 2012). The structure of RepABC type plasmids coincides with the localization of the origin of replication (*ori*) within the protein coding part of the replicase (*repC*) and the presence of a regulatory antisense RNA between *repB* and *repC* (Weaver, 2007). Conserved palindromes of RepABC-type plasmids that serve as cis-acting anchors for RepB proteins are indispensable for the successful partitioning of these low copy number replicons. The RepA/RepB system of RepABC-type plasmids is homologous to the universal ParA/ParB partitioning system of the bacterial chromosome and other tripartite plasmids (Petersen et al., 2011).

Horizontal gene transfer (HGT) correlates with major evolutionary transitions (Nelson-Sathi et al., 2015) and the process is *inter alia* mediated by phages and conjugative plasmids. Plasmid mobilization is ensured by type IV secretion systems (T4SS), representing a conserved export apparatus that is also used by several pathogenic bacteria for DNA and protein secretion (Cascales and Christie, 2003). The role of plasmid conjugation for the rapid adaptability of bacteria is exemplified by the emergence of multi-resistant hospital strains as a consequence of massive antibiotic (mis) use in medicine and livestock husbandry (Palmer et al., 2010). Many extrachromosomal elements are adapted to their host and do not stably maintain in distantly related bacteria (narrow-host-range plasmid; Kües and Stahl, 1989). A prominent example is the Ti plasmid from *A. tumefaciens* C58 that can be conjugated into *Escherichia coli* but requires an additional replication system for stable maintenance in *Enterobacteriaceae* (*Gammaproteobacteria*; Holsters et al., 1978). Naturally occurring broad-host-range vectors from *E. coli* were

genetically engineered and serve as crucial tools for molecular biology (see e.g., Kovach et al., 1995).

Conjugative plasmid transfer is an important factor in the evolution of rhizobia (Ding and Hynes, 2009; López-Guerrero et al., 2012) and has been proposed as the major driving force for the rapid adaptation of roseobacters to novel ecological niches (Petersen et al., 2013). In nitrogen-fixing rhizobia, symbiotic plasmids mediate legume root nodulation and define the particular plant host (Perret et al., 2000; Gibson et al., 2008). The pSym plasmid has been horizontally exchanged between different sympatric *Rhizobium* species, thereby transferring the capacity to nodulate the same host cell, i.e., the common bean *Phaseolus vulgaris* (Pérez Carrascal et al., 2016). Comparative genome analyses of *Rhodobacteraceae* showed that T4S systems for plasmid mobilization are typically located on RepABC plasmids, and experimental conjugation of natural plasmids has recently become feasible (Patzelt et al., 2016). The wealth of more than 400 draft genomes allowed tracing natural plasmid transfer among genus barriers in the roseobacter group (Petersen and Wagner-Döbler, 2017). However, the recombination rate between bacteria exponentially drops with increasing sequence divergence (Fraser et al., 2007), which depends, e.g., on the limited host range of mobilizable plasmids.

One purpose of the current study was the experimental validation of our *in silico* prediction that RepABC-type plasmids can not be transferred stably between *Rhizobiaceae* and *Rhodobacteraceae*. They are equivalent in structure and function but can clearly be distinguished by phylogenetic analyses (Petersen et al., 2009). Eight RepABC compatibility groups from roseobacters (*Rhodobacteraceae*) have a common origin and were once recruited from a rhizobial donor, but the respective plasmids have not been identified in rhizobia. The strict phylogenetic separation of these plasmid modules allows for a reliable genomic differentiation between both alphaproteobacterial orders, thus providing indirect evidence for functional constraints resulting in a narrow-host-range. However, the supposed ecological separation between the soil and the ocean, which would limit the physical contact for conjugation, is less pronounced than *a priori* assumed. Several rhizobial lineages, such as the genus *Martellella* are adapted to saline habitats (Rivas et al., 2005) and roseobacters represent a paraphyletic group associated with non-marine *Rhodobacteraceae* including the genus *Paracoccus* (Simon et al., 2017).

In the current study we experimentally document that rhizobial RepABC plasmids do not replicate in *Phaeobacter inhibens* DSM 17395 (*Rhodobacteraceae*), but we provide the first example of a natural plasmid that can be stably maintained in both rhizobia and roseobacters. This composite plasmid from the marine rhizobium *Martellella mediterranea* DSM 17316<sup>T</sup> originated from a plasmid fusion and still harbors rhizobial and rhodobacterial RepABC cassettes thus overcoming the limits of their host range.

## RESULTS AND DISCUSSION

### Identification of a *Rhodobacteraceae*-Specific RepABC Plasmid Replication Module in *Martellella mediterranea* DSM 17316<sup>T</sup>

Extensive data mining in roseobacters (*Rhodobacteraceae*) was the basis for the detection of nine compatibility groups of RepABC plasmids (Petersen et al., 2009), and recent studies indicated that this replicon type is crucial for HGT via conjugation (Petersen et al., 2013; Frank et al., 2015; Patzelt et al., 2016). The discoveries of the current study benefit from the Genomic Encyclopedia of Bacteria and Archaea (GEBA) genome sequencing project, which was aimed to fill phylogenetic gaps in the tree of life (Wu et al., 2009; Mukherjee et al., 2017), and the follow-up study of one thousand microbial genomes (KMG-I) that was focused on type strains (Kyrpides et al., 2014). More than a third of the selected strains from the latter project were *Proteobacteria* with 107 alphaproteobacterial representatives and 17 *Rhodobacteraceae* that were used to improve our reference data set of plasmid modules. However, we also investigated the distribution of RepABC operons in 35 novel rhizobial genomes, because this order previously served as a distantly related natural outgroup for *Rhodobacteraceae*-specific ECRs (Petersen et al., 2009). BLASTP searches with the RepC-2 replicase of the 126-kb RepABC plasmid from *Dinoroseobacter shibae* DFL12<sup>T</sup> (pDSHI03; WP\_012187065.1) allowed us to identify typical rhizobial homologs with a moderate protein identity of up to 40%, which is exemplified by the Ti-plasmid of *A. tumefaciens* C58 (36% identity). The sole and striking exception is a highly conserved RepABC module from the marine rhizobium *M. mediterranea* DSM 17316<sup>T</sup> (Rivas et al., 2005), whose replicase exhibits a conspicuous conservation of 58% identity. This finding was independently validated by analogous BLASTP searches with the adjacent RepA and RepB partitioning proteins, thereby documenting that *Martellella*'s RepABC module is indistinguishable from genuine operons of *Rhodobacteraceae*.

### Genome Sequencing of *Martellella mediterranea* DSM 17316<sup>T</sup>

#### Establishment of a Finished Genome with the PacBio Technique

We identified the roseobacter-specific RepABC plasmid replication operon of *Martellella* on a linear 105-kb DNA fragment (scaffold\_16.17, NZ\_AQWH01000017), which has been established in the first phase of the type strain sequencing project One Thousand Microbial Genomes (KMG-I; Kyrpides et al., 2014). The scaffold also contains a complete type IV secretion system (T4SS; Cascales and Christie, 2003) and a characteristic post-segregational killing system (PSK) encoding a stable toxin and an unstable antitoxin (Zielenkiewicz and Ceglowski, 2001) thus indicating that it might still represent a functional plasmid. However, meaningful analyses of extrachromosomal elements and the systematic investigation of HGT via conjugation essentially depend on a complete genome

sequence without any gaps and uncertain contig-affiliations. Accordingly, the genome of *M. mediterranea* DSM 17316<sup>T</sup> was sequenced with our PacBio platform. Based on a 130-fold sequence coverage and subsequent Illumina correction, we obtained a finished genome of highest quality with a size of 5.7 Mb, harboring four circular replicons representing the 4.7 Mb chromosome and three large RepABC-type ECRs with sizes of 593-, 259-, and 170-kb (**Figure S1**; CP020330 to CP020333). Four scaffolds including the 105-kb fragment perfectly match with the 259-kb plasmid pMM259 thus providing independent evidence for the technical reliability of both sequencing approaches, i.e., the initial Illumina assembly and our newly established PacBio genome. Hereby, tandem repeats of transposase and recombinase genes flanking the Illumina scaffolds are limiting *de novo* genome assembly approaches based on short reads. In return the newly established complete *Martellella* genome exemplifies the repeat-resolving power of long reads established by PacBio sequencing (Bleidorn, 2016).

#### Gene Content of the Extrachromosomal Elements

The 170-kb replicon pMM170 contains many sugar transport systems and the 593-kb ECR pMM593 holds a striking amount of TRAP- and ABC-transporters (Davidson et al., 2008; Fischer et al., 2010). Accordingly, both extrachromosomal elements might have an important function for *Martellella*'s metabolite exchange with the environment. The composite plasmid pMM259, which harbors the rhodobacterial RepABC module of interest and a rhizobial equivalent (**Figure S1**), contains two operons for copper export (Cu<sup>+</sup>; Mame\_05011-05014, Mame\_05047, Mame\_05054), two operons for cadmium or zinc export (Mame\_04956-04962, Mame\_04977-04982) and an arsenate-resistance cassette (Mame\_05000-05003), indicating that it represents a resistance plasmid for the detoxification of heavy metals. A specific exposure to toxic metal-ions has not been reported for the natural habitat of *M. mediterranea*, the subterranean Lake Martel on Mallorca (Rivas et al., 2005), but microorganisms are very sensitive to long time exposures of even moderate concentrations of heavy metals (Gadd and Griffiths, 1978).

### Phylogenetic Analysis of *Martellella*'s *Rhodobacteraceae*-Specific Plasmid RepABC Operon

#### Global Phylogeny of RepC-Type Replicases

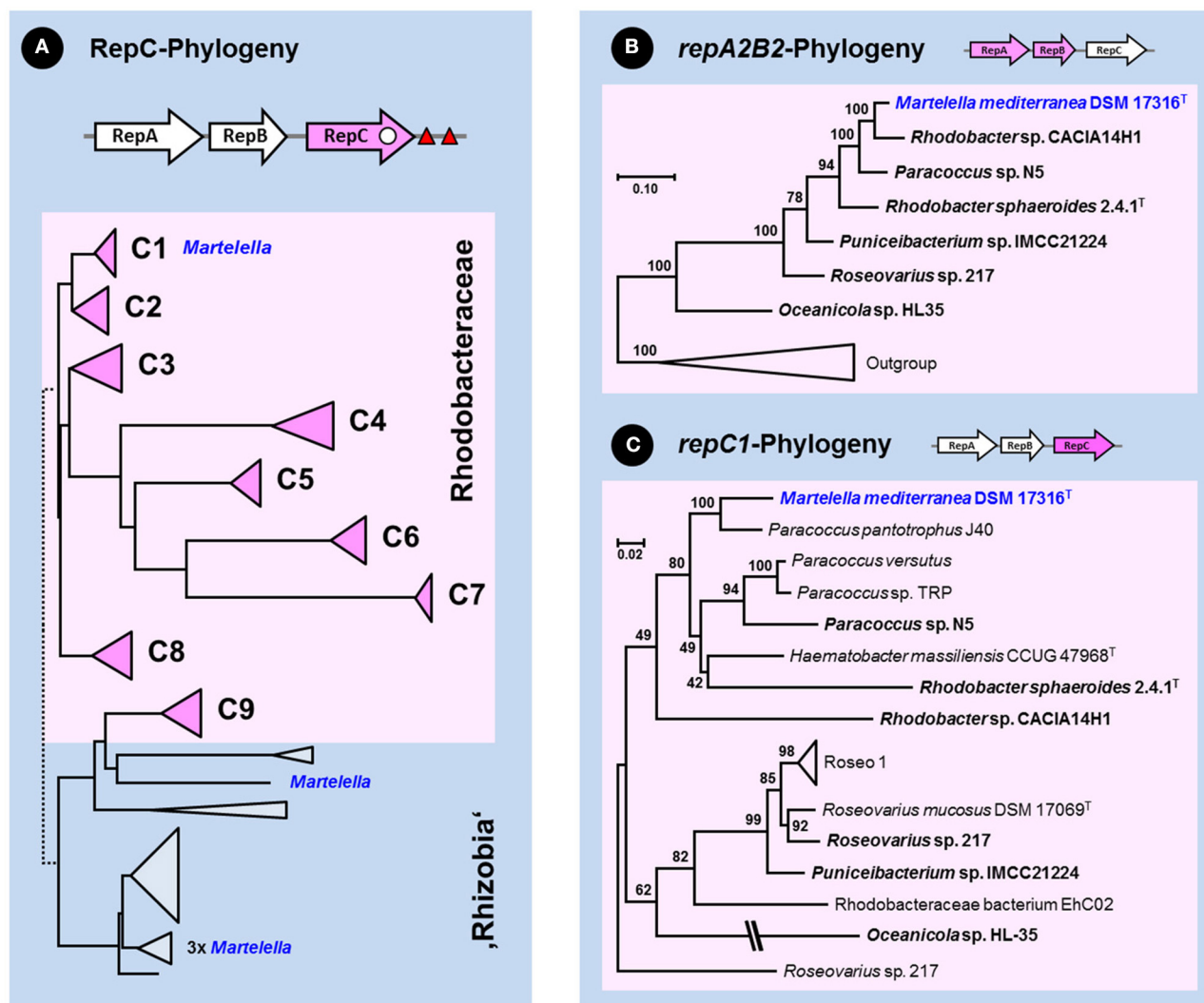
The genome of *M. mediterranea* harbors four RepABC replication modules [pMM593 (Mame\_04351-04353, Mame\_04880-04882), pMM259 (Mame\_04944-04946), pMM169 (Mame\_05119-05121)] and a solitary replicase gene [pMM259 (Mame\_04960)]. The phylogenetic position of the respective RepC proteins was determined based on a set of roseobacterial and mostly rhizobial reference sequences largely corresponding to those of our former study (Petersen et al., 2009), which allowed us to differentiate between the nine different compatibility groups from *Rhodobacteraceae* originating from two ancient acquisitions (I: C1 to C8, II: C9). The phylogenetic tree, which was calculated based on 121 RepC replicases, showed that

four *Martellella* sequences are located in the rhizobial part of the tree (blue color; **Figure 1A**, **Figure S2**). In contrast, the newly identified RepC protein of interest is placed amidst other rhodobacterial sequences in the distinct subtree C1 statistically supported by a 100% bootstrap proportion (BP; rose color; **Figure S2**). The internal branching pattern of the distinct subtree C1 is only poorly resolved due to the phylogenetically broad sequence sampling including the extremely divergent subtrees C4–C7, which resulted in only 145 comparable amino acid (aa) alignment positions (**Figure S2**). Accordingly, further analyses were focused on the *Rhodobacteraceae*-specific subtrees C1 and C2, whose sister-group relationship is solidly supported

(85% BP; **Figure 1**, **Figure S2**). We systematically searched the public sequence databases and investigated 81 different complete RepABC operons of the compatibility groups 1 and 2 (Petersen et al., 2009). Comprehensive phylogenetic subanalyses of all three genes were performed in order to detect the closest relative(s) of the roseobacterial RepABC module from *M. mediterranea*.

### Phylogenetic Analyses of Partitioning Proteins (RepA, RepB) and the Replicase RepC

Our subanalyses of proteins from RepABC modules belonging to the compatibility groups -1 and -2 resulted in a major improvement of the phylogenetic resolution with increased



**FIGURE 1 |** Phylogenetic analyses of the RepABC-type plasmid replication modules from *Rhodobacteraceae*. The pink color of the schematic RepABC operon above the respective phylogenetic tree indicates the analyzed gene. The origin of replication (*ori*) and conserved palindromes for plasmid partitioning are indicated by a white dot and red triangles, respectively. The localization of rhizobial RepC sequences from *Martellella mediterranea* DSM 17316<sup>T</sup> is highlighted in blue. **(A)** Schematic Neighbor Joining tree of 121 RepC replicase protein sequences from *Rhodobacteraceae* and rhizobia (see **Figure S2**). Subtrees of the nine compatibility groups C1–C9 from roseobacters are shown by pink triangles and rhodobacteracean subtrees are highlighted by a rose box. **(B,C)** Subanalysis of nucleotide sequences from concatenated *repA2B2* genes and the *repC1* gene. *Rhodobacteraceae* with a reshuffled RepA2B2A1 replication modules are highlighted in bold (**Figures S6, S7**). The final taxon sampling for the localization of *M. mediterranea* was determined by comprehensive analyses of RepA, RepB and RepC proteins belonging to compatibility groups 1 and 2 (**Figures S3–S5**).



bootstrap support (**Figures S3–S5**). The RepA and RepB trees have comparable branching patterns, which mirrors their concerted evolution in a functional partitioning operon, and the best resolution was obtained in the RepB analysis (**Figure S4**). Both partitioning proteins of *Martellella* (Mame\_04880–04881) are located in subtrees A2 and B2 (green sequences, 100% BP), which is *a priori* surprising because they show a deviating localization compared to the replicase RepC that is located in subtree C1 (**Figures S2, S5**). Our previous study showed a synchronous evolution of all three genes of the RepABC operon (Petersen et al., 2009), which is here exemplified by *Dinoroseobacter shibae* DFL12<sup>T</sup>, *Pseudooceanicola batsensis* HTCC 2597<sup>T</sup>, *Roseovarius indicus* B108<sup>T</sup> and *R. atlanticus* R12B<sup>T</sup>, four roseobacter type-strains that harbor RepABC operons of both compatibility groups (A1B1C1, A2B2C2; **Figures S3–S5**). Recombination events between a partitioning module of one compatibility group with a replicase of another compatibility group are rare, but they have previously been reported for the A2B2C1 module from *Roseovarius* sp. 217 (Petersen et al., 2009). The respective partitioning proteins of *M. mediterranea* (RepA, RepB) group together with *Roseovarius* sp. 217 and five other *Rhodobacteraceae* in a distinct branch of subtree -2 (highlighted in green, 100% BP; **Figures S3, S4**), in contrast to their replicases (RepC) that are all located in subtree -1, which reflects the common origin of the reshuffled A2B2C1 module.

### Detection of RepABC-1 and -2 Specific Palindromes

An independent criterion for the classification of RepABC-type plasmid replication modules is the presence of specific palindromes that also allow for differentiating between the nine compatibility groups in *Rhodobacteraceae* (Petersen et al., 2009). The highly conserved inverted repeats with a length of 14 nucleotides are in RepABC-1 and -2 modules typically located in close proximity of the RepABC operon downstream of RepC (**Figure 1**). Accordingly, we investigated the sequences of 36 RepABC modules starting 500 base pairs (bp) upstream of the RepA start codon and ending 500 bp downstream of the RepC stop codon. The sampling was focused on the A2B2C1 module of *M. mediterranea* DSM 17316<sup>T</sup> and our model organism *D. shibae* DFL12<sup>T</sup>, which served as a reference due to the presence of two plasmids with characteristic A1B1C1 and A2B2C2 modules (see above; Petersen et al., 2013). All but three of the investigated plasmid replication systems contain two adjacent copies of the specific palindrome separated by only 11 to 42 bp (**Table 1**). Necessities of partitioning appeared to result in a nearly universal conservation of the palindrome motifs TTAACAG/CTGT<sup>T</sup>A for compatibility group -1 and TTCACAG/CTGT<sup>T</sup>GAA for compatibility group -2. A single reciprocal nucleotide exchange in the third and third to last palindrome position (A:C, T:G) is responsible for plasmid compatibility and thus the stable co-existence e.g., of the 86-kb RepABC-1 and the 126-kb RepABC-2 replicons in *D. shibae* (Wagner-Döbler et al., 2010). *Martellella* and five additional *Rhodobacteraceae* with A2B2C1 modules harbor the characteristic doublet of compatibility group -2 palindromes, which is indistinguishable from those of genuine A2B2C2

plasmids. The phylogenetic localization of their partitioning genes in subtrees A2 and B2 reflects, in combination with the presence of compatibility group -2 palindromes, a case example of co-evolution based on functional constraints (**Figures S3, S4; Table 1**). The palindrome represents the highly specific cis-acting DNA recognition site for the partitioning protein ParB, whose smooth interaction is the prerequisite for successful plasmid distribution during bacterial cell division (Pinto et al., 2012). The lack of conserved palindromes in *Oceanicola* sp. HL-35, the sixth strain with an A2B2C1 operon (**Figures S3–S5**), might reflect the inactivation of its RepABC module. This prediction is supported by the early-branching position of strain HL-35 in the RepC tree (**Figure S5**) that could represent a phylogenetic long-branch attraction artifact (LBA, Philippe et al., 2005). In contrast, our phylogenetic analyses showed neither a conspicuous position nor prolonged branches for *M. mediterranea*'s RepA, RepB and RepC sequences (**Figures S3–S5**), thus indicating that its A2B2C1 plasmid module should be still functional at least in roseobacters (*Rhodobacteraceae*). Further phylogenies clearly documented a common origin of the *repA2B2* partitioning operon together with *Rhodobacter* sp. CACIA14H1 and of the *repC1* gene together with *Paracoccus pantotrophus* J40 (**Figures 1B,C; Supplemental Material S1**). This distribution reflects the frequent reshuffling of the replicase in A2B2C1 modules and moreover showed that the genuine rhodobacteracean donor of *Martellella*'s RepABC cassette has not been detected yet.

### The Composite Plasmid pMM259–A Chimera of Rhizobial and Rhodobacterial ECRs

Our plasmid of interest pMM259 possesses, apart from the conspicuous rhodobacterial RepABC module (A2B2C1; see above), a complete RepABC module as well as a solitary replicase (RepC), both of rhizobial origin, thus documenting that it represents a composite plasmid with replication systems from two alphaproteobacterial orders (**Figure 2A**). Replicons that originate from plasmid fusion events have previously been detected in completely sequenced genomes of other rhizobia, such as *Rhizobium etli* CFN 42, *Rhizobium leguminosarum* biovar *viciae* 3841 and *Rhizobium* sp. NT-26 (González et al., 2006; Young et al., 2006; Andres et al., 2013). Both replication modules of a composite plasmid might still be functional as experimentally documented, e.g., for the 107-kb replicon of *Paracoccus versutus* UW1 (Bartosik et al., 1998). However, a second replication system is generally not required for the stable maintenance of these low copy number plasmids and can be lost again. Experimental testing of the respective modules is hence the prerequisite for determining their functionality (see below), and it allows for drawing conclusions about the intrinsic potential of plasmid fission resulting in two operative replicons.

### Holistic Classification of *Martellella mediterranea*'s Extrachromosomal Replicons

The “chromid” concept of Harrison et al. (2010) introduced an evolutionary dimension into the classification of ECRs based on codon usage (CU) analyses. In brief, so-called chromids are



essential ECRs with a CU comparable to that of the chromosome, which mirrors their long-lasting co-evolution, whereas true plasmids are frequently exchanged via conjugation and thus exhibit a largely deviating CU (Petersen et al., 2013). We investigated the affiliation of all replicons from *M. mediterranea* DSM 17316<sup>T</sup>, *M. endophytica* YC6887<sup>T</sup>, and *Martellella* sp. AD-3, which represent the three completely sequenced genomes of this genus, in a principal component analysis (PCA) of the relative synonymous codon usage (RSCU; **Figures 2B,C**). The two-dimensional PCA, which explains 90.0% of the CU variance, shows a clear affiliation of pMM593 and pMM170 with *Martellella*'s chromosome, thus justifying their classification as chromids, whereas the composite replicon pMM259 represents a genuine plasmid. The capacity of horizontal exchange of this 259-kb plasmid is indicated by the presence of two T4S systems (see below) and furthermore supported by its RepABC replication system of rhodobacterial origin (A2B2C1 type; **Figure 1**).

### Identification of Horizontally Transferred Genes in pMM259

The presence of a composite plasmid with rhizobial and rhodobacterial replication systems in *M. mediterranea* indicated that this ECR might contain additional horizontally acquired

genes from roseobacters. However, a reliable detection of authentic HGTs would need to be based on time-consuming phylogenetic analyses, as documented for the RepABC modules (**Figure 1, Figures S2–S7**). We used a customized version of HGTector (Zhu et al., 2014) as a rapid discovery tool for the detection of potential HGT-derived genes on the plasmids of *Marteella*. The program allowed us to identify many putative HGTs (**Tables S1–S3**), but the two chromids of *Marteella* indicate that the number of authentic vertically evolving rhizobial genes may be underestimated. HGTector proposed a comparably low number of genes that are vertically transmitted (no HGT; 41% pMM593, 49% pMM170), whereas the best BLAST hits revealed a rhizobial affiliation for a larger part of these genes (55% pMM593, 64% pMM170; **Tables S1–S3**). Accordingly, we used the more conservative best BLAST hits for the differentiation between vertical inherited and horizontally acquired genes (**Figure 2, Figure S1**).

## Identification of Rhizobial and Rhodobacterial Genes on the *Martelella* Plasmid pMM259

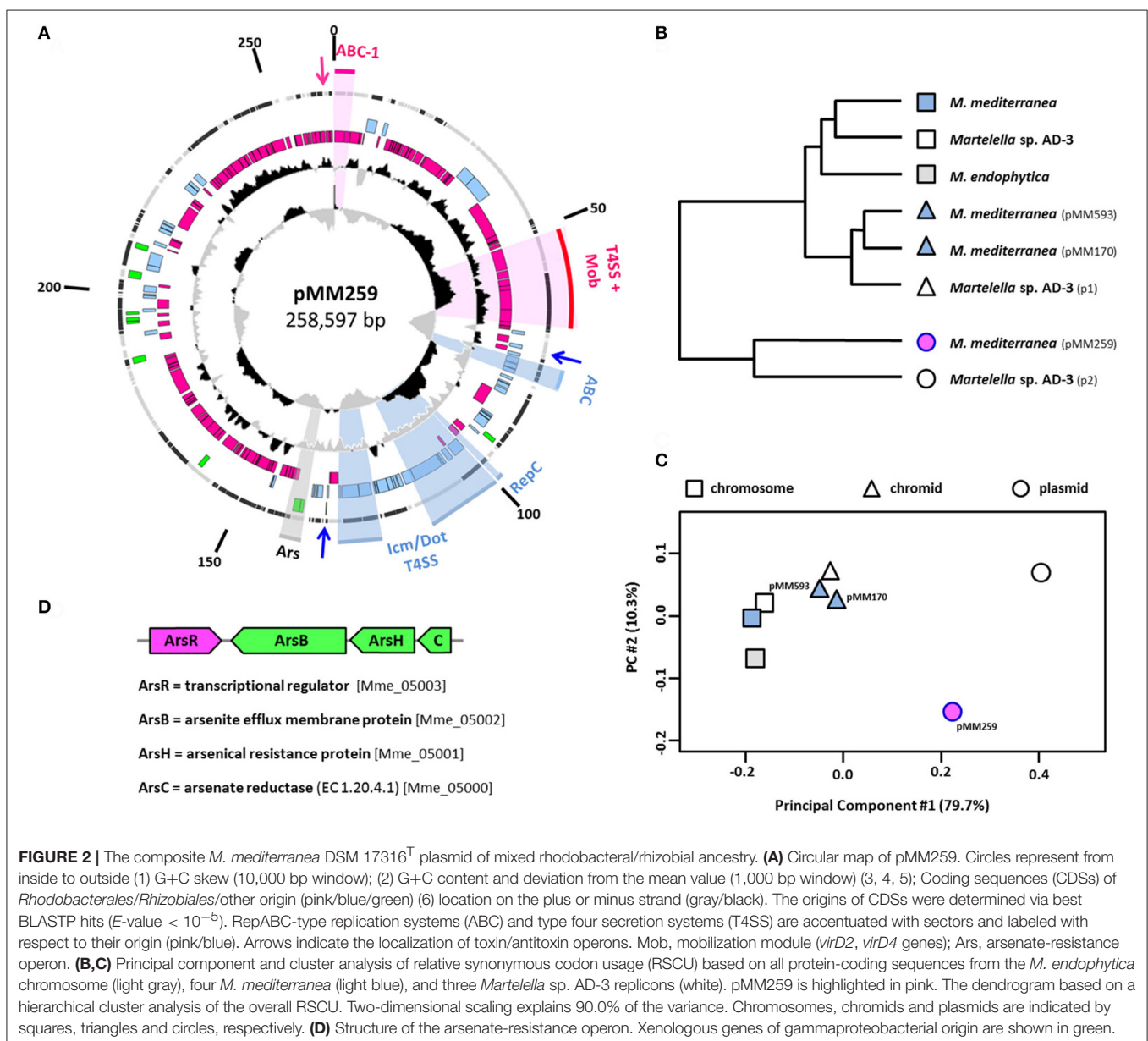
Our comparison of ECRs from *M. mediterranea* recovered a genuine rhizobial affiliation for the majority of chromid-located genes [291/529 (55%) pMM593, 100/155 (65%) pMM170],

**TABLE 1** | Conserved palindromes associated with RepABC-1 and -2 plasmids.

RepABC-Type	Palindromic 1		Palindromic 2	
	-----><-----		-----><-----	
A1-B1-C1	CTTAACAGCTGTTAAG.....42nt.....		CTTAACAGCTGTTAAG	Dinoroseobacter shibae DFL12 <sup>T</sup> (86-kb; pDSHI04)
A1-B1-C1	TTAACAGCTGTATA.....32nt.....		TTAACAGCTGTATA	Pseudoceanicola batsensis HTCC 2597 <sup>T</sup>
A1-B1-C1	CTTAACAGCTGTTAAG.....28nt.....		GACTTAACAGCTGTTAAGTC	Roseovarius indicus B108 <sup>T</sup>
A1-B1-C1	CTTAACAGCTGTTAAG.....42nt.....		CTTAACAGCTGTTAAG	Roseovarius atlanticus R12B <sup>T</sup>
A1-B1-C1	CTTAACAGCTGTTAAG.....29nt.....		CTTAACAGCTGTTAAG	Roseovarius sp. 217 (I)
A1-B1-C1	CTTAACAGCTGTTAAG.....42nt.....		CTTAACAGCTGTTAAG	Oceanicola sp. S124
A1-B1-C1	CTTAACAGCTGTTAAG.....42nt.....		CTTAACAGCTGTTAAG	Donghicola xiamenensis DSM 18339 <sup>T</sup>
A1-B1-C1	CTTAACAGCTGTTAAG.....42nt.....		CTTAACAGCTGTTAAG	Thioclava sp. 13D2W2
A1-B1-C1	CTTAACAGCTGTTAAG.....42nt.....		CTTAACAGCTGTTAAG	Pseudooceanicola atlanticus 22II-S11g <sup>T</sup>
A1-B1-C1	CTTAACAGCTGTTAAG.....42nt.....		CTTAACAGCTGTTAAG	Phaeobacter sp. S60
A1-B1-C1	TTAACAGCTGTATA.....42nt.....		TTAACAGCTGTATA	Sulfitobacter sp. NAS-14.1
A1-B1-C1	CaTTAACAGCTGTATAAG.....29nt....		GACTTtACAGCTGTAAAGTC	Haematobacter massiliensis CCUG 47968 <sup>T</sup>
A1-B1-C1	TTAACAGCTGTATA.....30nt.....		ACTTAACAGCTGTATAAGT	Shimia marina CECT7688 <sup>T</sup>
A1-B1-C1	TTAACAGCTGTATA.....30nt.....		ACTTAACAGCTGTATAAGT	Marinovum algicola DG898
A1-B1-C1	TTAACAGCTGTATA.....30nt.....		ACTTAACAGCTGTATAAGT	Marinovum algicola DSM 10251 <sup>T</sup>
A1-B1-C1	TTAACAGCTGTATA.....32nt.....		TTAACAGCTGTATA	Maritimibacter alkaliphilus HTCC 2654 <sup>T</sup>
A1-B1-C1	TTAACAGCTGTATA.....30nt.....		ACTTAACAGCTGTATAAGT	Roseivivax isoporaе LMG 25204 <sup>T</sup>
A1-B1-C1	CaTTAACAGCTGTATAAG no 2nd palindrome			Paracoccus pantotrophus J40
A1-B1-C1	TTAACAGCTGTATA.....32nt.....		TTAACAGCTGTATA	Ruegeria mobilis S1611
A1-B1-C1	TTAACAGCTGTATA.....32nt.....		TTAACAGCTGTATA	Thalassobius mediterraneus CECT5118
A1-B1-C1	aTTAACAGCTGTATA.....27nt.....		TTtgCAGCTGttAA	Paracoccus versutus
	:: :::::::::: ::		:: :::::::::: ::	
A2-B2-C2	AGTTACAGCTGTCAACT.....26nt.....		AGTTACAGCTGTCAACT	Dinoroseobacter shibae DFL12 <sup>T</sup> (126-kb; pDSHI03)
A2-B2-C2	GTTACAGtTGtCAAC.....27nt.....		GTTACAGCTGTCAAC	Pseudooceanicola batsensis HTCC 2597 <sup>T</sup>
A2-B2-C2	GAGTTACAGCTGTCAACTC no 2nd palindrome			Roseovarius indicus B108 <sup>T</sup>
A2-B2-C2	GAGTTACAGCTGTCAACTC.....26nt.....		AGTTACAGCTGTCAACT	Roseovarius atlanticus R12B <sup>T</sup>
A2-B2-C1	GACAGTTACAGCTGTCAACTGTC.....15nt.....		GTTACAGCTGTCAAC	Martellella mediterranea DSM 17316 <sup>T</sup>
A2-B2-C1	AGTTACAGCTGTCAACT.....30nt.....		GTTACAGCTGTCAAC	Rhodobacter sp. CACIA14H1
A2-B2-C1	AGTTACAGCTGTCAACT.....18nt.....		GTTACAGCTGTCAAC	Paracoccus sp. N5
A2-B2-C1	GTTACAGCTGTCAAC.....29nt.....		GTTACAGCTGTCAAC	Rhodobacter sphaeroides 2.4.1 <sup>T</sup>
A2-B2-C1	AGTTACAGCTGTCAACT.....29nt.....		GTTACAGCTGTCAAC	Puniceibacterium sp. IMCC21224
A2-B2-C1	GTTACAGCTGTCAAC.....28nt.....		GTTACAGCTGTCAAC	Roseovarius sp. 217 (II)
A2-B2-C2	AGTTACAGtTGtCAACT.....29nt.....		AGTTACAGCTGTCAACT	Sagittula stellata E 37 <sup>T</sup>
A2-B2-C2	GTTACAGCTGTCAAC.....29nt.....		GTTACAGCTGTCAAC	Roseovarius nubinihibens ISM <sup>T</sup>
A2-B2-C2	TTACAGCTGTCAA.....11nt.....		TTACAGCTGTCAA	Pelagibaca bermudensis HTCC 2601 <sup>T</sup>
A2-B2-C2	GAGTTACAGCTGTCAACTC.....29nt.....		AGTTACAGCTGTCAACT	Rhodobacter sp. CCB-MM2
A2-B2-C2	GTTACAGCTGTCAAC no 2nd palindrome			Paenirhodobacter ensiensis DW2-9 <sup>T</sup>

but only for 29% of the genes from the composite plasmid pMM259 (69/239), whose largest portion of genes [147/239 (62%)] is of rhodobacterial origin (**Figure S1**). A 30-kb stretch between 183- and 213-kb on pMM259 exhibits a rather scattered distribution of rhizobial and non-rhizobial genes (**Figure 2A**), but many of the rhizobial genes represent transposases that might have recently been acquired by intragenomic transposition events. Accordingly, the general composition of pMM259 clearly documents that the 259-kb plasmid harbors a backbone of roseobacter-associated genes and a rhizobial insertion of about 60-kb starting upstream of the blue RepABC module and ending downstream of the Icm/Dot type T4S system [**Figure 2A**, **Table S2** (Mame\_04940 to Mame\_04999)]. This spatial separation of genes with a vertical and non-vertical history indicates that the present day plasmid still reflects the

fusion event of a conjugated roseobacter plasmid with a size of about 200-kb with a 60-kb equivalent from the rhizobial host. Both partners have a different nucleotide composition as illustrated by deviations of the G+C content (**Figure 2A**). The 60-kb remnant of the rhizobial plasmid has a remarkably low G+C content of just 58% compared to 62% in the rhodobacterial part, a proportion that is, coincidentally, comparable to those of the two chromids (62%) and the chromosome (63%). The observed difference nearly reflects the natural G+C range of rhizobial genomes, thus documenting that *Martellella* is not the natural host of the rhizobial part of pMM259. This conclusion is supported by a codon-usage subanalysis (RSCU) of the rhizobial and rhodobacterial parts of pMM259 (**Figure S8**). The clustering clearly documents that (i) the CUs of rhizobial and roseobacter-specific genes largely differ and that (ii) both fusion partners can



be classified as plasmids with a CU largely deviating from those of the chromosome. The G+C skew plot in **Figure 2A** allows one to pinpoint the origin of replication of the rhizobial 60-kb fragment, which is located within the *repC* gene of the blue RepABC module (Pinto et al., 2012), and it further shows the leading and lagging strand for DNA replication (Lobry, 1996; Grigoriev, 1998). *In silico* ligation of the 60-kb fragment even allows for predicting the former terminus of replication within the *icm/dot* operon for pilus formation of the rhizobial T4SS. Our analyses documented that it is still possible to detect specific molecular imprints in the genuine rhizobial plasmid, thus we conclude that the plasmid fusion was, from an evolutionary point of view, a rather recent event.

### Further Xenologous Genes of pMM259

The HGT analyses showed that between 4 and 11% of the genes encoded on the three ECRs from *M. mediterranea* have a distinct affiliation that is neither related to rhizobia nor to *Rhodobacteraceae* (Tables S1–S3; Figure S1). These genes are highlighted in green within the outermost colored circle of **Figure 2A** and **Figure S1**. Gene clusters with comparable best BLASTP hits are especially interesting because they indicate that whole DNA modules and not only single genes have been horizontally transferred. One example for the 259-kb plasmid is an operon with two adjacent genes from *Rhodospirillales* encoding a thiol-disulfide interchange protein precursor and a lipoprotein signal peptidase II involved in protein export (Mame\_05059, Mame\_05058). However, the most conspicuous finding is the xenologous arsenate-resistance operon *arsCHB* with the adjacent transcriptional regulator *arsR* (Figures 2A,D; Mame\_05000–05003), which is crucial for the oxidative detoxification of the highly poisonous methylarsenite (III) to methylarsenate (V) by ArsH (Mukhopadhyay et al., 2002; Yang and Rosen, 2016). This operon is, from an evolutionary perspective, of remarkable interest, because it exemplifies that the recombination of distantly related genes from *Alpha*- and *Gammaproteobacteria* resulted in the formation of a functional unit. The transcriptional regulator ArsR originate—as indicated by its pink color—from *Rhodobacteraceae*, whereas the genes marked in green of the resistance operon have a gammaproteobacterial origin with *Halomonas zhanjiangensis* DSM 21076<sup>T</sup> (*Oceanospirillales*, *Halomonadaceae*) as closest relative. The chronology of HGT and plasmid fusion is difficult to estimate because the module is located within the transition zone between the rhizobial and the rhodobacterial part of plasmid pMM259 (**Figure 2A**). The horizontal transfer of the gammaproteobacterial *arsCHB* operon might thus either reflect a rather recent event in the genus *Martellella* or it already occurred within roseobacters prior to plasmid conjugation.

## Origin and Distribution of the Composite Plasmid

### Type IV Secretion Systems of the Composite Plasmid pMM259

We investigated the origin of the two type IV secretion systems (T4SS) located on the composite 259-kb plasmid

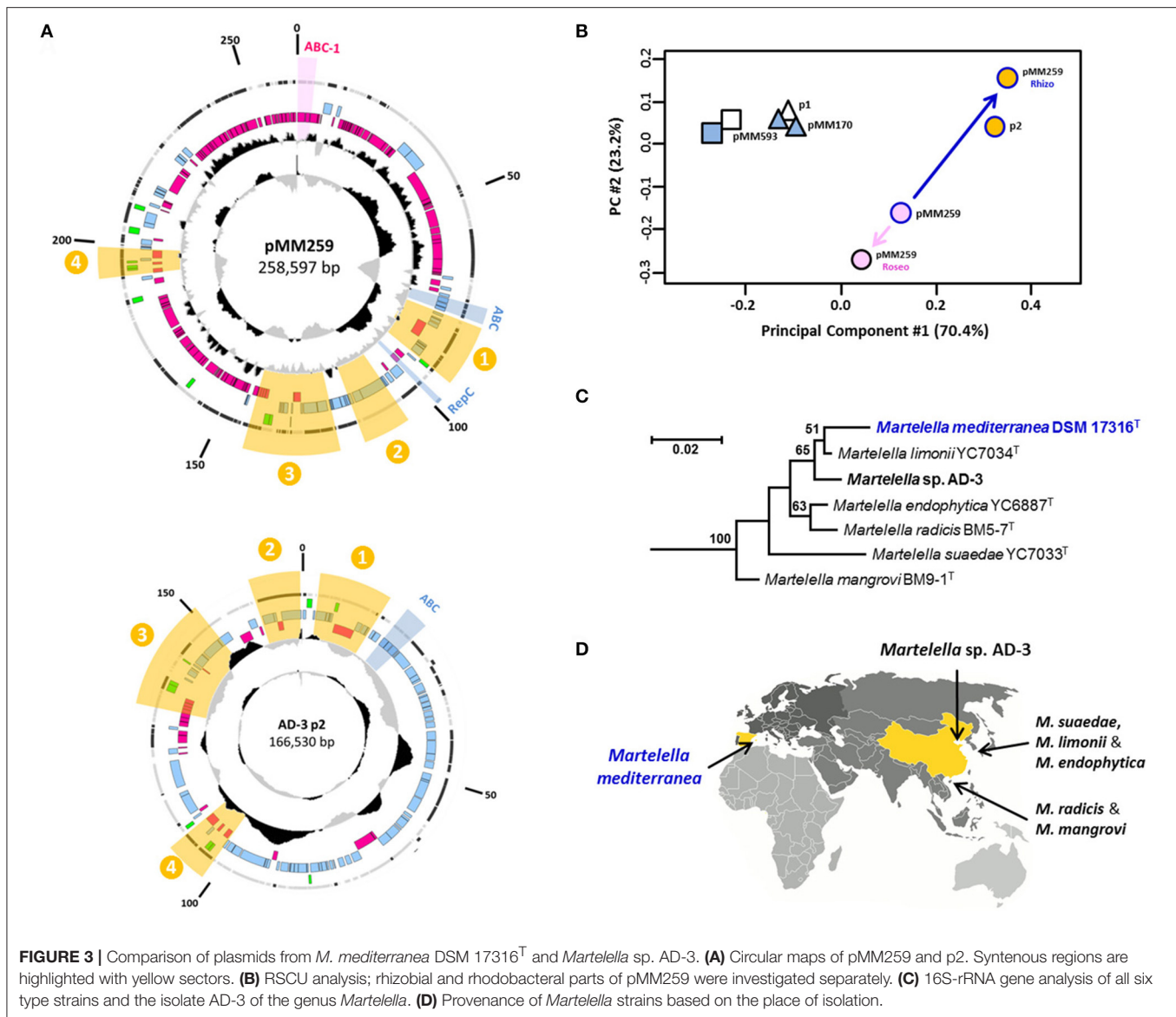
(**Figure 2A**) based on the assumption that one of them might have mediated the conjugational transfer of the roseobacter-specific A2B2C1-type RepABC plasmid into *M. mediterranea* (**Figure 1**, **Table 1**). The superoperon marked in blue of rhizobial origin represents an Icm/Dot T4SS with characteristic *icm* and *dot* genes (Mame\_04961–04992; Juhas et al., 2008). This extremely divergent secretion system harbors a conjugative transfer relaxase gene *traA* (Mame\_04988; **Table S2**), thus indicating that it is responsible for plasmid conjugation, whereas comparable conserved systems of *Legionella pneumophila* and *Coxiella burnetii* are utilized for bacterial pathogenesis (Segal et al., 2005). Syntenous superoperons including the mobilization genes have been identified in other rhizobia, such as *Sinorhizobium* sp. CCBAU 05631 or *Ochrobactrum anthropi* OAB. We were surprised that the “pink” T4S system of rhodobacterial origin also contains all essential genes for conjugational plasmid transfer. Its module structure, which comprises the *virB* secretion apparatus (Mame\_04915–04935) and a cluster of mobilization genes including the crucial relaxase and the coupling protein [*virD2* (Mame\_04917), *virD4* (Mame\_04918)], is absolutely conserved regarding homologs from other *Rhodobacteraceae* (Petersen et al., 2013). One example is the duplicated T4SS from the 191-kb and 126-kb sister plasmids of *D. shibae* DFL12<sup>T</sup>, whose conjugation across genus barriers has recently been demonstrated (Patzelt et al., 2016).

Our analyses suggest that the “pink” T4S system mediated plasmid conjugation from a still unknown roseobacter (donor) into the rhizobial recipient *Martellella*, thus explaining the large portion the rhodobacteracean genes on pMM259 (**Figure 2A**). Moreover, the structural integrity of the investigated T4SS systems indicates that the 259-kb plasmid from *Martellella* is still conjugative. Accordingly, horizontally transferred syntenous plasmids are waiting to be discovered in other marine bacteria. This aim seems to be like searching for a needle in the haystack, but horizontal plasmid transfer in the ocean has—concomitant with the exponential increase of whole genome sequences—very recently been reported for two roseobacters i.e., *D. shibae* DFL12<sup>T</sup> and *Confluentimicrobium naphthalenivorans* NS6<sup>T</sup> (Petersen and Wagner-Döbler, 2017).

### The Closest Relative of the Composite Plasmid pMM259

The conspicuous separation of rhodobacterial and rhizobial genes on the composite plasmid is indicative of a rather recent fusion event (see above; **Figure 2**). Accordingly, we tried to identify close syntenous relatives of pMM259 with BLASTN searches in the non-redundant (nr) and whole-genome shotgun (wgs) nucleotide databases of the NCBI. This approach allows for the identification of conserved genetic modules and is based on the detection of silent mutations more sensitive than a standard BLASTP search. However, our analyses revealed no highly specific *Rhodobacteraceae* hits with more than 95% sequence identity. This outcome documents that the donor for plasmid fusion is still undetected, which is in agreement with the phylogenetic analyses of the A2B2C1 plasmid-replication module (see above). In contrast, the composite *M. mediterranea*





plasmid pMM259 specifically matches with the 167-kb plasmid “p2” from *Martellella* sp. AD-3 (CP014277.1), and the four syntenic regions with a total size of 47-kb exhibit an average sequence identity between 95 and 99% (highlighted in yellow, **Figure 3A**). Their close affiliation is independently shown by the RSCU comparison of plasmid p2 with the rhizobial part from pMM259 (**Figures 3B,C**). Three of the conserved areas including region “three,” which contains the arsenate-resistance operon (**Figure 2D**), are matching the rhizobial part of *M. mediterranea*’s 259-kb plasmid, but region “four” shares 97% sequence identity with the rhodobacterial part of pMM259. This distribution indicates that the 167-kb plasmid from *Martellella* sp. AD-3 might also originate from the composite rhizobial/rhodobacterial fusion plasmid and secondarily lost the majority of roseobacter-specific genes including the A2B2C1-type RepABC replication module and the T4SS.

## pMM259—A Natural Plasmid for Horizontal Gene Transfer between *Rhodobacteraceae* and *Rhizobiaceae*

### Rationale for the Experiments

The presence of two complete RepABC-type replication systems of rhizobial and rhodobacterial origin on the composite *M. mediterranea* plasmid is of particular interest because it suggests that pMM259 might represent a natural replicon mediating HGT between two alphaproteobacterial orders. Accordingly, we established a transformation assay that allowed us to monitor the replication of RepABC-type plasmid modules based on antibiotic selection and chose the model organisms *P. inhibens* DSM 17395 and *A. tumefaciens* C58 DSM 5172 (synonyms *Agrobacterium radiobacter*, *Rhizobium radiobacter*; Tindall, 2014) as test strains.



## Cloning of RepABC Modules

Two RepABC operons from the 259-kb *M. mediterranea* plasmid, i.e., the roseobacter-specific RepABC module (4415 bp; A2B2C1) and the genuine rhizobial RepABC module (4986 bp), were cloned into the commercial vector pCR2.1 (see Experimental Procedures). Furthermore, we searched for a rhodobacterial positive control for our stability tests and analogously cloned the RepABC-8 operon from *P. inhibens* T5<sup>T</sup> (DSM 16374<sup>T</sup>; 3909 bp, A8B8C8). The respective module is specific for the type strain and located on an 88-kb plasmid, which is missing in other isolates, such as *P. inhibens* DSM 17395 (Thole et al., 2012; Dogs et al., 2014). The resulting plasmids pPI88-Roseo, pMM259-Rhizo and pMM259-Roseo that are shown in **Figure 4A** represent artificial shuttle vectors with a host-specific copy number. In *Escherichia coli* (*Enterobacteriaceae*, *Gammaproteobacteria*) they replicate based on the modified pUC origin derived from a ColE1/pMB1 vector as high copy number plasmids [500–700 copies per chromosome (HCNP); Gelfand et al., 1978; Lee et al., 2006], in contrast to the alphaproteobacterial host(s) where the respective RepABC system ensures a stable maintenance as a low copy number plasmid [1 copy per chromosome (LCNP); Pappas, 2008].

## Functionality Tests of RepABC Modules in *P. inhibens* DSM 17395 and *A. tumefaciens* C58

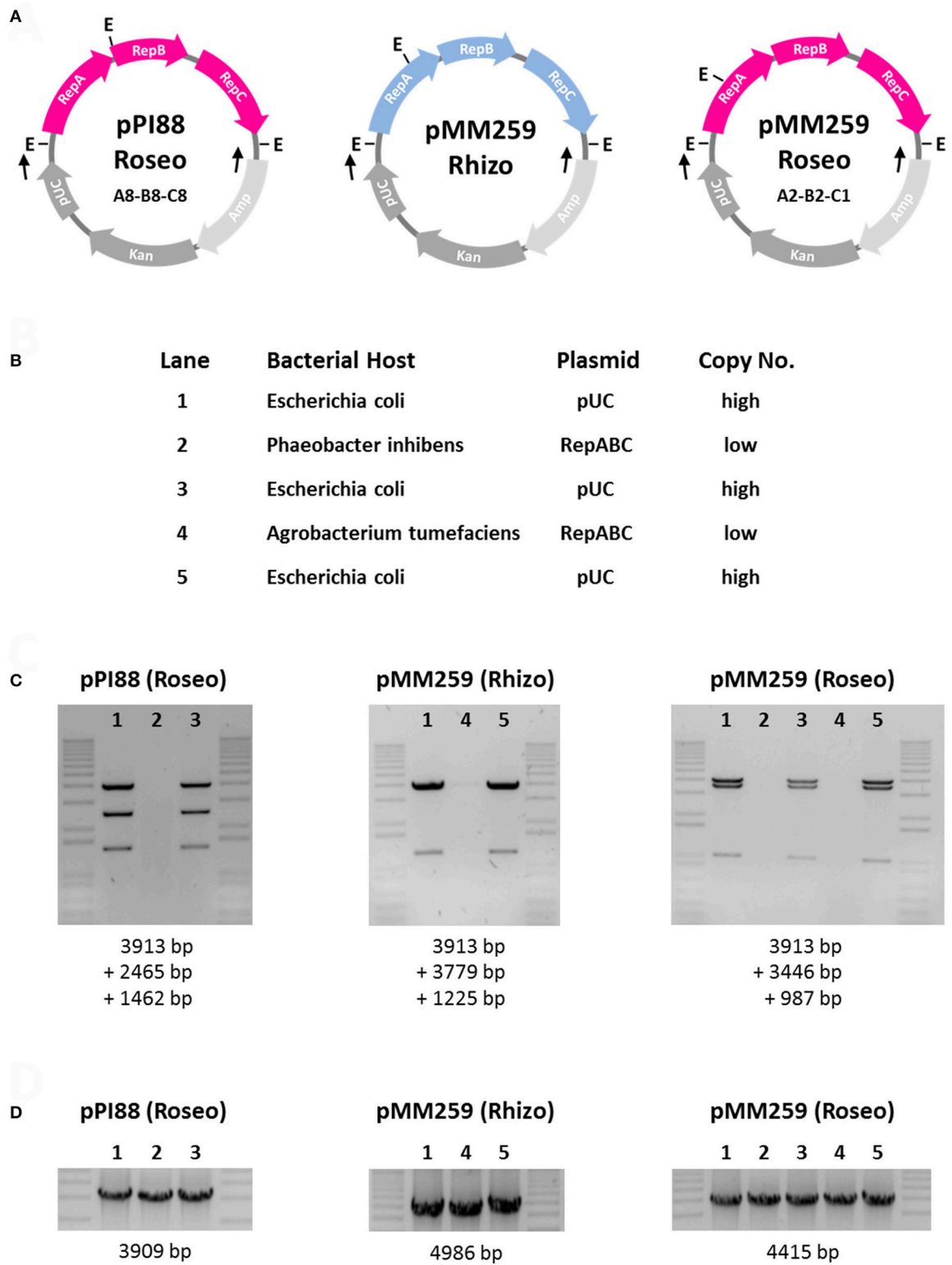
*Phaeobacter inhibens* DSM 17395 and *A. tumefaciens* C58 are both sensitive to the antibiotic kanamycin, and we accordingly used the respective resistance gene of pCR2.1 as a selection marker for our experiments (**Figure 4A**). Transformation of a circular pCR2.1 plasmid without an insert was used as negative control and confirmed that the *E. coli* cloning vector does not replicate in *Alphaproteobacteria*. A potential pitfall of the functionality test is a stable integration of the pCR2.1 construct into the chromosome, which would also result in kanamycin-resistant transformants. Accordingly, and as an ultimate proof of functional plasmid replication, we isolated the LCNPs from the alphaproteobacterial host, retransformed them into *E. coli* and showed that the *Eco*RI restriction patterns of isolated plasmid DNA are identical to those of the original digests (lane 1&3, lane 1&5; **Figure 4C**). The absence of DNA fragments in lanes 2 and 4 mirrors the low copy number of RepABC-type plasmids in *Alphaproteobacteria* but control PCRs showed that the respective LCNPs are present in all three (five) samples (**Figure 4D**).

Based on this experimental setup, we were able to document the functionality of our assay including the selected test strains. The positive control pPI88-Roseo mediated—based on its RepABC-8 module—stable plasmid replication in *P. inhibens* DSM 17395, but it does not replicate in *A. tumefaciens* C58 (**Figure 4**). The rhizobial RepABC module from *Martelella* (pMM259-Rhizo) showed a reciprocal pattern and is only replicated in *Agrobacterium*. This finding is in agreement with the strict phylogenetic separation of rhizobial and rhodobacterial RepABC replication systems (**Figure 1**), which led to the *in silico* prediction of functional incompatibility (Petersen et al., 2009). Furthermore, our assay did not only validate the functionality of *Martelella*'s xenologous A2B2C1 plasmid replication system

in *Phaeobacter*, it surprisingly also showed that pMM259-Roseo is—at least under kanamycin selection—replicated and stably maintained in *A. tumefaciens* (**Figures 4C,D**). The outcome is contradictory to the experiments with the RepABC-8 operon of *P. inhibens* T5<sup>T</sup> and indicates that some RepABC-type plasmids of *Rhodobacteraceae* might have a broader host range than previously assumed. This prediction is supported by the presence of a rhodobacterial operon on the composite 322-kb plasmid from *Rhizobium* sp. NT-26 (Andres et al., 2013) but especially by former host-range tests with the RepABC-1 operon of the composite 107-kb plasmid pTAV1 from *P. versutus* UW1, which documented stable replication in *Rhizobium etli* CE3 and *Rhizobium leguminosarum* 1062 (Bartosik et al., 1998). In contrast, the rhizobial RepABC module from *Martelella* (pMM259-Rhizo; **Figure 4**) showed the expected host range limited to rhizobia. This outcome was independently validated by the respective operon from the Ti-plasmid (*A. tumefaciens* C58) that does also not replicate in *P. inhibens* DSM 17395 (data not shown), thus indicating that functional constraints prevent the replication of rhizobial RepABC plasmids in *Rhodobacteraceae*.

## Stability Tests of RepABC Modules Replicating in *Phaeobacter* and *Agrobacterium*

The presence of two replication systems on *Martelella*'s 259-kb plasmid is surprising, because the rhizobial module should be sufficient for replication. Accordingly, we proposed that the stability of pMM259-Roseo is reduced in rhizobia and tested this hypothesis experimentally based on the four previously established transformants (**Figure 4**). The tests of pPI88-Roseo and pMM259-Roseo in *Phaeobacter*, which served as a reference, showed that about 5% of the cells, i.e., two of 40 tested colonies, lost their RepABC-type plasmid over night during exponential growth under non-selective growth conditions (**Figure S9**). The comparable stability of both constructs documented that the roseobacter-specific module from *Martelella* is not only functional in *Rhodobacteraceae*, it is even unaffected in its viability. The outcome of analogous tests with the two *Martelella* constructs pMM259-Roseo and pMM259-Rhizo in *Agrobacterium* was completely unexpected, because it showed that 90% of the host cells lost the genuine rhizobial RepABC construct spontaneously (36/40 colonies), whereas the xenologous rhodobacterial module was lost in just one of the 40 tested colonies (2.5%; **Figure S9**). We validated the presence of the respective construct for two resistant colonies to exclude any sample mix up and repeated the experiment, which resulted in comparable rates of spontaneous plasmid loss (pMM259-Roseo: 0/40; pMM259-Rhizo: 32/40). Yet, pMM259-Rhizo is still functional and maintained in *Agrobacterium* under selective pressure, but the high frequency of loss under non-selective growth conditions might reflect an ongoing degeneration of the RepABC-system into a “pseudogene module.” Accordingly, the most probable evolutionary scenario predicts that the selective pressure exclusively remains on the functional rhodobacterial RepABC cassette. The inactivated



**FIGURE 4 |** Functionality tests of rhizobial and rhodobacterial RepABC-type plasmids in *Phaeobacter inhibens* DSM 17395 and *Agrobacterium tumefaciens* C58. **(A)** Schematic plasmid maps of the tested constructs. Roseobacter and rhizobial RepABC-type replication systems are shown in pink and blue, respectively. The genes of the cloning vector pCR2.1 are shown in gray. pUC, origin of replication; Kan and Amp, kanamycin and ampicillin resistance genes; E, *EcoRI* restriction site; Black arrows indicate primer binding sites for PCR amplification. **(B)** Specificity and copy number of composite plasmids based on the host range of the plasmid replication system. **(C)** Plasmid restriction assay with *EcoRI*. The 1 kb Plus DNA ladder from Invitrogen was used as a marker. **(D)** PCR assay with pCR2.1-specific primers.

rhizobial module will get lost soon thus erasing the plasmid-specific molecular footprint of one fusion partner in the composite plasmid pMM259.

The rate of plasmid loss observed in the current study correlates with an exponential growth of the host cell in extremely nutrient-rich medium and is thus not representative for the natural habitat. Stable maintenance of natural plasmids is promoted by beneficial and sometimes even essential genes and furthermore ensured by toxin/antitoxin systems (Zielenkiewicz and Ceglowski, 2001), which is exemplified by three respective modules on pMM259 (Figure 2). Taken together, the replication module pMM259-Roseo has a broader host range than its equivalent pMM259-Rhizo (Figure 4) and it moreover showed an unexpected stability in both tested host strains [*Phaeobacter* (*Rhodobacteraceae*), *Agrobacterium* (*Rhizobiaceae*); Figure S9]. Accordingly, this A2B2C1-cassette represents the “functional heart” of a natural plasmid that should mediate stable genetic exchange between alphaproteobacterial orders essentially based on the presence of two different T4SSs (Figure 2).

### Significance of the Composite Plasmid pMM259 and Conclusion

In the current study we established the complete genome sequence of the rhizobium *M. mediterranea* DSM 17316<sup>T</sup>. Its composite 259-kb replicon, which originated from a plasmid fusion, still harbors two functional RepABC modules that ensure plasmid replication in *Rhizobiaceae* and *Rhodobacteraceae* (Figures 1, 2, 4). *M. mediterranea* has been isolated from the subterranean Lake Martel in the Dragon Cave on the Spanish island Mallorca (Rivas et al., 2005). This saline karst lake is located in very close proximity of the Mediterranean Sea and represents an ideal location for the intimate contact of a halotolerant rhizobium with roseobacters. Accordingly, there are no ecological boundaries preventing trans-order conjugation. The presence of two conserved T4S systems strongly indicates that pMM259 is still mobilizable (Figure 2), and it is thus likely that it mediates HGT from the globally occurring marine genus *Martelella* (Figure 3) into new rhizobial as well as roseobacter recipients. This plasmid is to our knowledge the first example of a natural replicon bridging the phylogenetic gap between these alphaproteobacterial orders. The presence of two replication systems on pMM259 overcomes the problem of the narrow-host-range of the rhizobial RepABC-type plasmids, and we thus propose that analogous plasmid fusions facilitate the genetic exchange even between bacterial classes. Previously, an outsourcing of the complete photosynthesis gene cluster for aerobic anoxygenic photosynthesis (AAnP) from the chromosome to a plasmid has been documented within the genus *Roseobacter* (Petersen et al., 2012). According to the “Think Pink” scenario (Petersen et al., 2013), plasmid conjugation could explain the presence of a homologous superoperon for AAnP in the marine gammaproteobacterium *Congregibacter litoralis* KT71<sup>T</sup> (Fuchs et al., 2007). Natural shuttle vectors would hence connect distantly related bacterial lineages from the same habitat thereby providing access to the metabolic potential of the marine pan-genome.

## EXPERIMENTAL PROCEDURES

### Bacterial Strains, Plasmids, and Growth Conditions

Bacterial strains and plasmids used in this study are listed in Table S4. For preparation of competent cells and isolation of genomic DNA all *Rhodobacteraceae* and *Rhizobiaceae* strains were cultured in 40 g/l Marine Broth medium (MB, Carl Roth) at 28°C and 120 rpm. ½ MB with 120 µg/ml kanamycin (Carl Roth) was used for antibiotic selection.

### Host Range Tests of RepABC Replication Systems

The RepABC replication systems of *P. inhibens* T5<sup>T</sup> (= DSM 16374<sup>T</sup>) and *M. mediterranea* DSM 17316<sup>T</sup> were amplified from genomic DNA by PCR using the specific primers P1093 (5'-ACCGGCGACACAACACTCACC-3') and P1094 (3'-ACGCGTGATCTTTCTGCTCTT-5') for pPI88-Roseo, P1245 (5'-CGTCGAGCAGGTAAAGAACG-3') and P1246 (3'-GTTTCGACCCCTTCAGCATC-5') for pMM259-Roseo and P1289 (5'-GCTCATCGTACCGTTTGTCC-3') and P1290 (3'-GCGAAATCCACGGTAATGCT-5') for pMM259-Rhizo with the Phusion proof-reading polymerase (Thermo-Fischer Scientific). The obtained PCR products were subsequently cloned into the *E. coli* vector pCR2.1 with a kanamycin resistance and a pUC origin of replication, which is not functional in *Alphaproteobacteria*. Control sequencing documented the integrity of the modules and the absence of PCR errors. We chose *P. inhibens* DSM 17395 (*Rhodobacteraceae*) and *A. tumefaciens* C58 DSM 5172 (*Rhizobiaceae*) as representative hosts for plasmid stability experiments. Electrocompetent cells were generated as previously described (Dower et al., 1988). Electroporation was conducted using 50 ng plasmid DNA in a 2 mm cuvette and 2.5 kV. Colonies grown were passaged three times on fresh agar plates under constant antibiotic pressure to eradicate residual untransformed plasmids from the culture. Plasmid DNA was isolated with the NucleoSpin Plasmid kit from Macherey-Nagel. PCR with the generic pCR2.1 vector primers P022 (5'-GGAAACAGCTATGACCATGATTAC-3') and P023 (5'-CGTAATACGACTCATATAGGGC-3') was performed to detect low copy number plasmids. Retransformation of the isolated plasmid DNA into *E. coli* allowed for excluding false positives resulting from genomic integration of the kanamycin resistance gene and thus to verify the functionality of the tested RepABC replication systems. The integrity of retransformed constructs was documented by *EcoRI* digestion and gel electrophoresis.

### Stability Tests of RepABC Replication Systems

Bacterial transformants (*P. inhibens*, *A. tumefaciens*) harboring RepABC modules cloned in pCR2.1 were grown in a test tube with 3 ml MB medium and kanamycin (120 µg/ml) overnight. 10 µl of the culture was transferred in a 50 ml Erlenmeyer flask with 10 ml MB medium without antibiotics and grown for 16 h. The cultures were streaked out on MB plates and incubated for 2 days. Single colonies have been resuspended in 20 µl MB medium and 3 µl of these cells were in parallel spotted on MB

plates with and without kanamycin. We investigated the presence of 40 independent colonies of each transformant and could thus monitor the stability of the RepABC-type plasmid in the respective host bacterium.

## PacBio Library Preparation and Sequencing

A SMRTbell™ template library was prepared according to the instructions from PacificBiosciences, Menlo Park, CA, USA, following the Procedure & Checklist- >10 kb Template Preparation Using Ampure® PB Beads. Briefly, for preparation of 10 kb libraries 8 µg genomic DNA was sheared using g-tubes™ from Covaris, Woburn, MA, USA according to the manufacturer's instructions. DNA was end-repaired and ligated overnight to hairpin adapters applying components from the DNA/Polymerase Binding Kit P6 from Pacific Biosciences, Menlo Park, CA, USA. Reactions were carried out according to the manufacturer's instructions. BluePippin™ Size-Selection to 7 kb was performed according to the manufacturer's instructions (Sage Science, Beverly, MA, USA). Conditions for annealing of sequencing primers and binding of polymerase to purified SMRTbell™ template were assessed with the Calculator in RS Remote, PacificBiosciences, Menlo Park, CA, USA. SMRT sequencing of two SMRT cells was carried out on the PacBio RSII (PacificBiosciences, Menlo Park, CA, USA) taking 240-min movies.

## Genome Assembly, Error Correction, and Annotation

*De novo* genome assembly of *M. mediterranea* DSM 17316<sup>T</sup> was carried out based on 67,093 post-filtered PacBio reads with an average read length of 13,478 bp using the “RS\_HGAP\_Assembly.3” protocol included in SMRT Portal version 2.3.0 applying default parameters. The assembly process revealed one circular chromosome and three ECRs. End trimming and circularization was performed, where the chromosome was adjusted to *dnaA* and all ECRs to their replication genes. Finally, each genome was error-corrected by a mapping of Illumina reads onto finished genomes using BWA (Li and Durbin, 2009) with subsequent variant and consensus calling using VarScan (Koboldt et al., 2012). Correct replicon structures and a consensus concordance of QV60 were confirmed by using the “RS\_Bridgemappper.1” protocol. Finally, an annotation was generated using Prokka 1.8 with subsequent manual reannotation of all replication genes (Seemann, 2014). Complete genomes were deposited at NCBI GenBank under the accession numbers CP020330 to CP020333.

## Analysis of Horizontally Transferred Genes

HGT analysis of *M. mediterranea* DSM 17316<sup>T</sup> plasmids was conducted using HGTector.py (Zhu et al., 2014) with BLASTP against the NCBI non-redundant (nr) sequence database (download: October, 12th 2016), the taxonDMP (October, 12th 2016) and release 78 of MultispeciesAutonomousProtein2taxname from RefSeq. To exclude self-hits the corresponding TaxID (293088) was defined as self-group of *M. mediterranea*. The close group was defined as

TaxID 356 (*Rhizobiales*) respectively, all other organisms in the nr database made up the distal group. Best hits from up to 500 blast results with a  $10^{-5}$  *e*-value cutoff were used to determine the origin of genes on order level. Further analysis and creation of circle plots was accomplished by custom R scripts utilizing ggbio, GRanges, ggplot2, rentrez, and taxize packages.

## Phylogenetic Analyses

The amino acid and nucleotide alignments of RepABC genes obtained with ClustalW (Thompson et al., 1997) were manually refined using the ED option of the MUST program package (Philippe, 1993). Gblocks was used to eliminate both highly variable and/or ambiguous portions of the alignments (Talavera and Castresana, 2007). Maximum likelihood (ML) analyses were performed with RAXML version 8.2.4 (Stamatakis, 2014) applying Pthreads to use multiple shared memory nodes and SSE3 vector instructions, which together allow for substantially speeding up the computations depending on the number of nodes used. In RAXML a rapid bootstrap analysis with 100 replicates followed by a thorough search of the ML tree was conducted under the LG+F+4I<sup>†</sup> model. For protein analyses of the RepABC modules the neighbor-joining algorithm with gamma-corrected distances under the JTT model including 100 bootstrap replicates was used as described in Petersen et al. (2011). The calculations were performed in the program MEGA version 5 (Tamura et al., 2011) in an interactive way via the graphical user interface (GUI).

## AUTHOR CONTRIBUTIONS

JP and PB designed research. PB, HB, and BB contributed new data. PB, HB, BB, and JP performed analyses. BB and MG contributed software tools. JP, HB, and PB drafted manuscript and all authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

We would like to thank Claire Ellebrandt, Simone Severitt and Nicole Heyer for excellent technical assistance, Cathrin Spröer for PacBio sequencing support and three reviewers for their constructive criticism. This work including the PhD stipend for PB was supported by the Transregional Collaborative Research Center “*Roseobacter*” (Transregio TRR 51) of the Deutsche Forschungsgemeinschaft.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01787/full#supplementary-material>

**Figure S1** | Extrachromosomal replicons of *M. mediterranea* DSM 17316<sup>T</sup>. Circles represent from inside to outside (1) G+C Skew (10,000 bp window); (2) G+C content and deviation from the mean value (1,000 bp window); (3, 4, 5) Coding sequences (CDSs) of *Rhodobacterales*/*Rhizobiales*/other origin (pink/blue/green); (6) location on plus or minus strand (gray/black). The origins of CDSs were determined via best BLASTP hits (*E*-value <  $10^{-5}$ ). The actual scale



between plasmids is not taken into account. Toxin/antitoxin operons for plasmid stability are indicated by stars.

**Figure S2** | Composite Neighbor Joining tree of 121 **RepC** replicases from RepABC-type plasmids representing all nine *Rhodobacteraceae*-specific **compatibility groups** (C1 to C9). The upper subtree based on 50 sequences from *Rhodobacteraceae* and 145 amino acid position ( $\alpha = 0.92$ ; JTT) and the lower subtree based on 71 mostly rhizobial sequences including rhodobacteracean RepC-9 proteins and 226 amino acid positions ( $\alpha = 1.03$ ; JTT). The statistical support for the internal nodes was determined by 100 bootstrap replicates (BR) and values >50% are shown. Internal rooting was performed according to the RepC-tree of Petersen et al. (2009). Rhodobacteracean subtrees and the "rhizobial" tree are highlighted by pink and blue boxes, respectively.

**Figure S3** | Neighbor Joining tree (p-distances; 100 BR) of **RepA** partitioning proteins from the RepABC plasmid replication operon of the rhodobacteracean compatibility groups 1 and 2 based on 81 sequences and 386 amino acid positions. *Martellella mediterranea* DSM 17316<sup>T</sup> is highlighted in blue. Strains for phylogenetic subanalyses are highlighted in green (**Figure 1B**, **Figure S5**).

**Figure S4** | Neighbor Joining tree (p-distances; 100 BR) of **RepB** partitioning proteins from the RepABC plasmid replication operon of the rhodobacteracean compatibility groups 1 and 2 based on 81 sequences and 225 amino acid positions.

**Figure S5** | Neighbor Joining tree (p-distances; 100 BR) of **RepC** replicases from the RepABC plasmid replication operon of the rhodobacteracean compatibility groups 1 and 2 based on 81 sequences and 352 amino acid positions.

**Figure S6** | Phylogenetic positioning of the *repAB* partitioning module from *Martellella mediterranea* DSM 17316<sup>T</sup>. Strains with A2B2C1-type plasmid replication systems (**Figures S2–S4**) are highlighted in bold and green. **(A)** Maximum Likelihood [ML] tree (RAxML, LG+F+4 $\Gamma$ ; 100 BR) of **RepA2** proteins based on nine sequences and 394 amino acid positions. **(B)** ML tree (RAxML,

LG+F+4 $\Gamma$ ; 100 BR) of **RepB2** proteins based on 312 amino acid positions. **(C)** ML tree (RAxML, LG+F+4 $\Gamma$ ; 100 BR) of concatenated **RepA2** and **RepB2** proteins based on 706 amino acid positions. **(D)** ML tree (RAxML, GTR+4 $\Gamma$ ; 100 BR) of concatenated *repA2* and *repB2* genes based on 2164 nucleotide positions.

**Figure S7** | Phylogenetic positioning of the *repC* replication gene from *Martellella mediterranea* DSM 17316<sup>T</sup>. Strains with A2B2C1-type plasmid replication systems (**Figures S2–S4**) are highlighted in bold and green. **(A)** Maximum Likelihood tree (ML; RAxML, LG+F+4 $\Gamma$ ; 100) of **RepC1** proteins based on 39 sequences and 403 amino acid positions. **(B)** ML tree (RAxML, LG+F+4 $\Gamma$ ; 100 BR) of RepC1 proteins based on 17 sequences and 402 amino acid positions. **(C)** ML tree (RAxML, GTR+4 $\Gamma$ ; 100 BR) of *repC1* genes based on 17 sequences and 1212 nucleotide positions.

**Figure S8** | Principal component and cluster analysis of relative synonymous codon usage (RSCU) based on all protein-coding sequences from the four *M. mediterranea* replicons. The rhodobacterial (Roseo) and rhizobial (Rhizo) specific genes of pMM259 were also analyzed separately. Their distribution is highlighted in the plasmid map in pink or in blue, respectively. Two-dimensional scaling explains 98.0% of the variance. Chromosomes, chromids and plasmids are indicated by squares, triangles and circles, respectively.

**Figure S9** | Plasmid stability tests of the rhizobial and rhodobacteracean RepABC-type plasmids pMM259-Roseo, pMM259-Rhizo and pP188-Roseo in *Phaeobacter inhibens* DSM 17395 and *Agrobacterium tumefaciens* C58 (see **Figure 4**). 3  $\mu$ l of resuspended bacterial colonies were in parallel spotted on two agar plates with and without the antibiotic kanamycin.

**Table S1** | HGT Analysis of *Martellella mediterranea* plasmid pMM593.

**Table S2** | HGT Analysis of *Martellella mediterranea* plasmid pMM259.

**Table S3** | HGT Analysis of *Martellella mediterranea* plasmid pMM170.

**Table S4** | Strains and vectors used in this study.

## REFERENCES

- Andres, J., Arsène-Ploet, F., Barbe, V., Brochier-Armanet, C., Cleiss-Arnold, J., Coppée, J. Y., et al. (2013). Life in an arsenic-containing gold mine: genome and physiology of the autotrophic arsenite-oxidizing bacterium *Rhizobium* sp. NT-26. *Genome Biol. Evol.* 5, 934–953. doi: 10.1093/gbe/evt061
- Bartosik, D., Baj, J., and Włodarczyk, M. (1998). Molecular and functional analysis of pTAV320, a repABC-type replicon of the *Paracoccus versutus* composite plasmid pTAV1. *Microbiology* 144, 3149–3157. doi: 10.1099/00221287-144-11-3149
- Bleidorn, C. (2016). Third generation sequencing: technology and its potential impact on evolutionary biodiversity research. *Syst. Biodivers.* 14, 1–8. doi: 10.1080/14772000.2015.1099575
- Cascales, E., and Christie, P. J. (2003). The versatile bacterial type IV secretion systems. *Nat. Rev. Microbiol.* 1, 137–149. doi: 10.1038/nrmicro753
- Davidson, A. L., Dassa, E., Orelle, C., and Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiol. Mol. Biol. Rev.* 72, 317–364. doi: 10.1128/MMBR.00031-07
- Ding, H., and Hynes, M. F. (2009). Plasmid transfer systems in the rhizobia. *Can. J. Microbiol.* 55, 917–927. doi: 10.1139/W09-056
- Dogs, M., Voget, S., Teshima, H., Petersen, J., Davenport, K., Dalingault, H., et al. (2014). Genome sequence of *Phaeobacter inhibens* type strain (T5T), a secondary metabolite producing representative of the marine *Roseobacter* clade, and emendation of the species description of *Phaeobacter inhibens*. *Stand. Genomic Sci.* 9, 334–350. doi: 10.4056/sigs.4448212
- Dower, J., Miller, J. F., and Ragsdale, C. W. (1988). High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* 16, 6127–6145. doi: 10.1093/nar/16.13.6127
- Fischer, M., Zhang, Q. Y., Hubbard, R. E., and Thomas, G. H. (2010). Caught in a TRAP: substrate-binding proteins in secondary transport. *Trends Microbiol.* 18, 471–478. doi: 10.1016/j.tim.2010.06.009
- Frank, O., Göker, M., Pradella, S., and Petersen, J. (2015). Ocean's Twelve: flagellar and biofilm chromids in the multipartite genome of *Marinovum algicola* DG898 exemplify functional compartmentalization. *Environ. Microbiol.* 17, 4019–4034. doi: 10.1111/1462-2920.12947
- Fraser, C., Hanage, W. P., and Spratt, B. G. (2007). Recombination and the nature of bacterial speciation. *Science* 315, 476–480. doi: 10.1126/science.1127573
- Fuchs, B. M., Spring, S., Teeling, H., Quast, C., Wulf, J., Schattenhofer, M., et al. (2007). Characterization of a marine gammaproteobacterium capable of aerobic anoxygenic photosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2891–2896. doi: 10.1073/pnas.0608046104
- Gadd, G. M., and Griffiths, A. J. (1978). Microorganisms and heavy metal toxicity. *Microb. Ecol.* 4, 303–317. doi: 10.1007/BF02013274
- Gelfand, D. H., Shepard, H. M., O'Farrell, P. H., and Polisky, B. (1978). Isolation and characterization of a ColE1-derived plasmid copy-number mutant. *Proc. Natl. Acad. Sci. U.S.A.* 75, 5869–5873. doi: 10.1073/pnas.75.12.5869
- Gibson, K. E., Kobayashi, H., and Walker, G. C. (2008). Molecular determinants of a symbiotic chronic infection. *Annu. Rev. Genet.* 42, 413–441. doi: 10.1146/annurev.genet.42.110807.091427
- González, V., Santamaría, R. I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., et al. (2006). The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3834–3839. doi: 10.1073/pnas.0508502103
- Grigoriev, A. (1998). Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Res.* 26, 2286–2290. doi: 10.1093/nar/26.10.2286
- Harrison, P. W., Lower, R. P. J., Kim, N. K. D., and Young, J. P. W. (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol.* 18, 141–148. doi: 10.1016/j.tim.2009.12.010
- Holsters, M., Silva, B., Van Fliet, F., Hernalsteens, J. P., Genetello, C., Van Montagu, M., et al. (1978). *In vivo* transfer of the Ti-plasmid of *Agrobacterium tumefaciens* to *Escherichia coli*. *Mol. Gen. Genet.* 163, 335–338. doi: 10.1007/BF00271963
- Juhas, M., Crook, D. W., and Hood, D. W. (2008). Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. *Cell. Microbiol.* 10, 2377–2386. doi: 10.1111/j.1462-5822.2008.01187.x
- Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., et al. (2012). VarScan 2: somatic mutation and copy number alteration

- discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576. doi: 10.1101/gr.129684.111
- Kovach, M. E., Elzer, P. H., Hill, D. S., Robertson, G. T., Farris, M. A., Roop, I. L., R. M., et al. (1995). Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* 166, 175–176. doi: 10.1016/0378-1119(95)00584-1
- Kües, U., and Stahl, U. (1989). Replication of plasmids in gram-negative bacteria. *Microbiol. Rev.* 53, 491–516.
- Kyrpides, N. C., Hugenholtz, P., Eisen, J. A., Woyke, T., Göker, M., Parker, C. T., et al. (2014). Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol.* 12:e1001920. doi: 10.1371/journal.pbio.1001920
- Lee, C. L., Ow, D. S. W., and Oh, S. K. W. (2006). Quantitative real-time polymerase chain reaction for determination of plasmid copy number in bacteria. *J. Microbiol. Methods* 65, 258–267. doi: 10.1016/j.mimet.2005.07.019
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324
- Lobry, J. R. (1996). Asymmetric substitution patterns in the two DNA strands of bacteria. *Mol. Biol. Evol.* 13, 660–665. doi: 10.1093/oxfordjournals.molbev.a025626
- López-Guerrero, M. G., Ormeño-Orrillo, E., Acosta, J. L., Mendoza-Vargas, A., Rogel, M. A., Ramírez, M. A., et al. (2012). Rhizobial extrachromosomal replicon variability, stability and expression in natural niches. *Plasmid* 68, 149–158. doi: 10.1016/j.plasmid.2012.07.002
- Michael, V., Frank, O., Bartling, P., Scheuner, C., Göker, M., Brinkmann, H., et al. (2016). Biofilm plasmids with a rhamnose operon are widely distributed determinants of the “swim-or-stick” lifestyle in roseobacters. *ISME J.* 10, 2498–2513. doi: 10.1038/ismej.2016.30
- Mukherjee, S., Seshadri, R., Varghese, N. J., Eloë-Fadrosch, E. A., Meier-Kolthoff, J. P., Göker, M., et al. (2017). 1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life. *Nat. Biotechnol.* 35, 676–683. doi: 10.1038/nbt.3886
- Mukhopadhyay, R., Rosen, B. P., Phung, L. T., and Silver, S. (2002). Microbial arsenic: from geocycles to genes and enzymes. *FEMS Microbiol. Rev.* 26, 311–325. doi: 10.1111/j.1574-6976.2002.tb00617.x
- Nelson-Sathi, S., Sousa, F. L., Roettger, M., Lozada-Chávez, N., Thiergart, T., Janssen, A., et al. (2015). Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 517, 77–80. doi: 10.1038/nature13805
- Palmer, K. L., Kos, V. N., and Gilmore, M. S. (2010). Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Curr. Opin. Microbiol.* 13, 632–639. doi: 10.1016/j.mib.2010.08.004
- Pappas, K. M. (2008). Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. *Plasmid* 60, 89–107. doi: 10.1016/j.plasmid.2008.05.003
- Pappas, K. M., and Cevallos, M. Á. (2011). “Plasmids of the *Rhizobiaceae* and their role in interbacterial and transkingdom interactions,” in *Biocommunication in Soil Microorganisms*, ed G. Witzany (Berlin, Heidelberg: Springer-Verlag), 403–412.
- Patzelt, D., Michael, V., Püker, O., Ebert, M., Tielen, P., Jahn, D., et al. (2016). Gene flow across genus barriers—Conjugation of *Dinoroseobacter shibae*’s 191-kb killer plasmid into *Phaeobacter inhibens* and AHL-mediated expression of type IV secretion systems. *Front. Microbiol.* 7:742. doi: 10.3389/fmicb.2016.00742
- Pérez Carrascal, O. M., VanInsberghe, D., Juárez, S., Polz, M. F., Vinuesa, P., and González, V. (2016). Population genomics of the symbiotic plasmids of sympatric nitrogen-fixing *Rhizobium* species associated with *Phaseolus vulgaris*. *Environ. Microbiol.* 18, 2660–2676. doi: 10.1111/1462-2920.13415
- Perret, X., Staehelin, C., and Broughton, W. J. (2000). Molecular basis of symbiotic promiscuity. *Microbiol. Mol. Biol. Rev.* 64, 180–201. doi: 10.1128/MMBR.64.1.180-201.2000
- Petersen, J. (2011). Phylogeny and compatibility: plasmid classification in the genomics era. *Arch. Microbiol.* 193, 313–321. doi: 10.1007/s00203-011-0686-9
- Petersen, J., Brinkmann, H., Berger, M., Brinkhoff, T., Püker, O., and Pradella, S. (2011). Origin and evolution of a novel DnaA-like plasmid replication type in *Rhodobacterales*. *Mol. Biol. Evol.* 28, 1229–1240. doi: 10.1093/molbev/msq310
- Petersen, J., Brinkmann, H., Bunk, B., Michael, V., Püker, O., and Pradella, S. (2012). Think pink: photosynthesis, plasmids and the *Roseobacter* clade. *Environ. Microbiol.* 14, 2661–2672. doi: 10.1111/j.1462-2920.2012.02806.x
- Petersen, J., Brinkmann, H., and Pradella, S. (2009). Diversity and evolution of repABC type plasmids in *Rhodobacterales*. *Environ. Microbiol.* 11, 2627–2638. doi: 10.1111/j.1462-2920.2009.01987.x
- Petersen, J., Frank, O., Göker, M., and Pradella, S. (2013). Extrachromosomal, extraordinary and essential—the plasmids of the *Roseobacter* clade. *Appl. Microbiol. Biotechnol.* 97, 2805–2815. doi: 10.1007/s00253-013-4746-8
- Petersen, J., and Wagner-Döbler, I. (2017). Plasmid transfer in the ocean—a case study from the *Roseobacter* group. *Front. Microbiol.* 8:1350. doi: 10.3389/fmicb.2017.01350
- Philippe, H. (1993). MUST, a computer package of Management Utilities for Sequences and Trees. *Nucleic Acids Res.* 21, 5264–5272. doi: 10.1093/nar/21.22.5264
- Philippe, H., Lartillot, N., and Brinkmann, H. (2005). Multigene analyses of bilaterian animals corroborate the monophyly of *Ecdysozoa*, *Lophotrochozoa*, and *Protostomia*. *Mol. Biol. Evol.* 22, 1246–1253. doi: 10.1093/molbev/msi111
- Pinto, U. M., Pappas, K. M., and Winans, S. C. (2012). The ABCs of plasmid replication and segregation. *Nat. Rev. Microbiol.* 10, 755–765. doi: 10.1038/nrmicro2882
- Rivas, R., Sánchez-Márquez, S., Mateos, P. F., Martínez-Molina, E., and Velázquez, E. (2005). *Marteella mediterranea* gen. nov., sp. nov., a novel  $\alpha$ -proteobacterium isolated from a subterranean saline lake. *Int. J. Syst. Evol. Microbiol.* 55, 955–959. doi: 10.1099/ijs.0.63438-0
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Segal, G., Feldman, M., and Zusman, T. (2005). The Icm/Dot type-IV secretion systems of *Legionella pneumophila* and *Coxiella burnetii*. *FEMS Microbiol. Rev.* 29, 65–81. doi: 10.1016/j.femsre.2004.07.001
- Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., et al. (2017). Phylogenomics of *Rhodobacteraceae* reveals evolutionary adaptation to marine and non-marine habitats. *ISME J.* 11, 1–17. doi: 10.1038/ismej.2016.198
- Stamatakis, A. (2014). RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56, 564–577. doi: 10.1080/10635150701472164
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28, 2731–2739. doi: 10.1093/molbev/msr121
- Thole, S., Kalhoefer, D., Voget, S., Berger, M., Engelhardt, T., Liesegang, H., et al. (2012). *Phaeobacter gallaeciensis* genomes from globally opposite locations reveal high similarity of adaptation to surface life. *ISME J.* 6, 1–16. doi: 10.1038/ismej.2012.62
- Thompson, J. D., Gilson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997). The Clustal\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882. doi: 10.1093/nar/25.24.4876
- Tindall, B. J. (2014). *Agrobacterium radiobacter* (Beijerinck and van Delden 1902) Conn 1942 has priority over *Agrobacterium tumefaciens* (Smith and Townsend 1907) Conn 1942 when the two are treated as members of the same species based on the principle of priority and Rule 23a, Note 1 as applied to the corresponding specific epithets. *Opinion* 94. *Int. J. Syst. Evol. Microbiol.* 64, 3590–3592. doi: 10.1099/ijs.0.069203-0
- Wagner-Döbler, I., Ballhausen, B., Berger, M., Brinkhoff, T., Buchholz, I., Bunk, B., et al. (2010). The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker’s guide to life in the sea. *ISME J.* 4, 61–77. doi: 10.1038/ismej.2009.94
- Weaver, K. E. (2007). Emerging plasmid-encoded antisense RNA regulated systems. *Curr. Opin. Microbiol.* 10, 110–116. doi: 10.1016/j.mib.2007.03.002
- Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., et al. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462, 1056–1060. doi: 10.1038/nature08656

- Yang, H. C., and Rosen, B. P. (2016). New mechanisms of bacterial arsenic resistance. *Biomed. J.* 39, 5–13. doi: 10.1016/j.bj.2015.08.003
- Young, J. P. W., Crossman, L. C., Johnston, A. W., Thomson, N. R., Ghazoui, Z. F., Hull, K. H., et al. (2006). The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol.* 7:R34. doi: 10.1186/gb-2006-7-4-r34
- Zhu, Q., Kosoy, M., and Dittmar, K. (2014). HGTector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics* 15:717. doi: 10.1186/1471-2164-15-717
- Zielenkiewicz, U., and Ceglowski, P. (2001). Mechanisms of plasmid stable maintenance with special focus on plasmid addiction systems. *Acta Biochim. Pol.* 48, 1003–1023.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Bartling, Brinkmann, Bunk, Overmann, Göker and Petersen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Plasmid Transfer in the Ocean – A Case Study from the Roseobacter Group

Jörn Petersen<sup>1\*</sup> and Irene Wagner-Döbler<sup>2\*</sup>

<sup>1</sup> Research Group Plasmids and Protists, Leibniz-Institute DSMZ – German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany, <sup>2</sup> Research Group Microbial Communication, Helmholtz – Center for Infection Research, Braunschweig, Germany

## OPEN ACCESS

### Edited by:

Bernd Wemheuer,  
University of New South Wales,  
Australia

### Reviewed by:

Haiwei Luo,  
The Chinese University of Hong Kong,  
Hong Kong  
Yonghui Zeng,  
Aarhus University, Denmark

### \*Correspondence:

Jörn Petersen  
joern.petersen@dsMZ.de  
Irene Wagner-Döbler  
iwd@helmholtz-hzi.de

### Specialty section:

This article was submitted to  
Aquatic Microbiology,  
a section of the journal  
Frontiers in Microbiology

**Received:** 20 March 2017

**Accepted:** 03 July 2017

**Published:** 18 July 2017

### Citation:

Petersen J and Wagner-Döbler I  
(2017) Plasmid Transfer  
in the Ocean – A Case Study from  
the Roseobacter Group.  
Front. Microbiol. 8:1350.  
doi: 10.3389/fmicb.2017.01350

Plasmid mediated horizontal gene transfer (HGT) has been speculated to be one of the prime mechanisms for the adaptation of roseobacters (*Rhodobacteraceae*) to their ecological niches in the marine habitat. Their plasmids contain ecologically crucial functional modules of up to ~40-kb in size, e.g., for aerobic anoxygenic photosynthesis, flagellar formation and the biosynthesis of the antibiotic tropodithietic acid. Furthermore, the widely present type four secretion system (T4SS) of roseobacters has been shown to mediate conjugation across genus barriers, albeit in the laboratory. Here we discovered that *Confluentimicrobium naphthalenivorans* NS6<sup>T</sup>, a tidal flat bacterium isolated in Korea, carries a 185-kb plasmid, which exhibits a long-range synteny with the conjugative 126-kb plasmid of *Dinoroseobacter shibae* DFL12<sup>T</sup>. Both replicons are stably maintained by RepABC operons of the same compatibility group (–2) and they harbor a homologous T4SS. Principal component analysis of the codon usage shows a large similarity between the two plasmids, while the chromosomes are very distinct, showing that neither of the two bacterial species represents the original host of those RepABC-2 type plasmids. The two species do not share a common habitat today and they are phylogenetically only distantly related. Our finding demonstrates the first clear-cut evidence for conjugational plasmid transfer across biogeographical and phylogenetic barriers in *Rhodobacteraceae* and documents the importance of conjugative HGT in the ocean.

**Keywords:** plasmid synteny, type IV secretion systems, conjugation, horizontal gene transfer, evolution

## INTRODUCTION

Horizontal gene transfer (HGT) dominates prokaryotic evolution (Koonin, 2016). In addition to plasmid transfer by conjugation, two other mechanisms of HGT are known in the ocean, namely, transformation via direct uptake of DNA and transduction by phages or gene transfer agents (GTAs). Plasmids have an important role because of the large amount of genetic material that can be transferred in a single conjugation event. Thus, uptake of a plasmid may allow the host to colonize an entirely new niche. For example, the virulence plasmid of *Shigella* converts a harmless *Escherichia coli* commensal into a deadly pathogen (The et al., 2016). The spread of antibiotic resistance plasmids across the globe powered by the selective pressure exerted by antibiotic misuse and overuse, represents an unsolved challenge to our health-care system (Palmer et al., 2010; Perry and Wright, 2013). The rapidity of this process is exemplified by the spread of resistance against colistin, the antibiotic of last resort, encoded by the plasmid-located *mrc-1* gene. It was



first discovered in November 2015 in commensal *E. coli* isolates from food animals in China (Liu et al., 2016) and six month later the first colistin-resistant *E. coli* strain was isolated from a patient in Pennsylvania, United States (McGann et al., 2016).

Little is known about HGT by conjugation in the ocean. Since it requires close physical contact between donor and recipient, it is thought to be restricted to hot spots like aggregates or surfaces, where bacteria form biofilms (Thibault and Top, 2016), have high cell densities and are metabolically active (Sobecky and Hazen, 2009). Studies until now have been restricted to microcosms (Aminov, 2011).

Roseobacters, a marine subgroup of the *Rhodobacteraceae*, carry up to 12 plasmids per cell (Pradella et al., 2010). The maintenance of these low copy number plasmids is ensured by a characteristic tripartite module comprising a replicase initiating their duplication at the origin of replication (*ori*) and two partitioning genes that mediate the coordinated anchorage of the replicons at the cell poles during cell division (Petersen, 2011). The most abundant plasmid type of roseobacters is represented by RepABC-type replicons comprising at least nine different compatibility groups that ensure their stable coexistence within the same cell (Petersen et al., 2009). The genetic functions encoded on Roseobacter plasmids are potential drivers of adaptation to ecological niches in the ocean, e.g., aerobic anoxygenic photosynthesis, synthesis of the antibiotic tropodithetic acid, biofilm formation, or killing of dinoflagellate cells (Petersen et al., 2013; Wang et al., 2015). Many of those plasmids carry type IV secretion systems (T4SS) (Chandran Darbari and Waksman, 2015), suggesting that they might be conjugative. Indeed it could be demonstrated that two syntenic sister plasmids of *Dinoroseobacter shibae* DFL12<sup>T</sup>, which originate from a plasmid duplication and the recruitment of a novel RepABC-type replication module (Wagner-Döbler et al., 2010), can be transferred to *Phaeobacter inhibens* DSM 17395 by conjugation *in vitro* (Patzelt et al., 2016). *Dinoroseobacter* is among the deepest branching genera within the Roseobacter group, while *Phaeobacter* is located in the most distant subclade (Simon et al., 2017), showing that these plasmids have a very broad host range in *Rhodobacteraceae*.

Here we provide insights into the evolutionary history of the conjugative 126-kb plasmid of *D. shibae*. We show that a syntenic plasmid is naturally present in two phylogenetically distant species of the Roseobacter group, i.e., *Confluentimicrobium naphthalenivorans* and *Roseovarius indicus*. Codon usage analysis shows that the plasmids have been obtained from different hosts, and thus have probably been transferred multiple times – they have “tramped” through the phylogenetic tree of *Rhodobacteraceae*. To the best of our knowledge, these observations provide the first example for HGT by conjugation in the ocean.

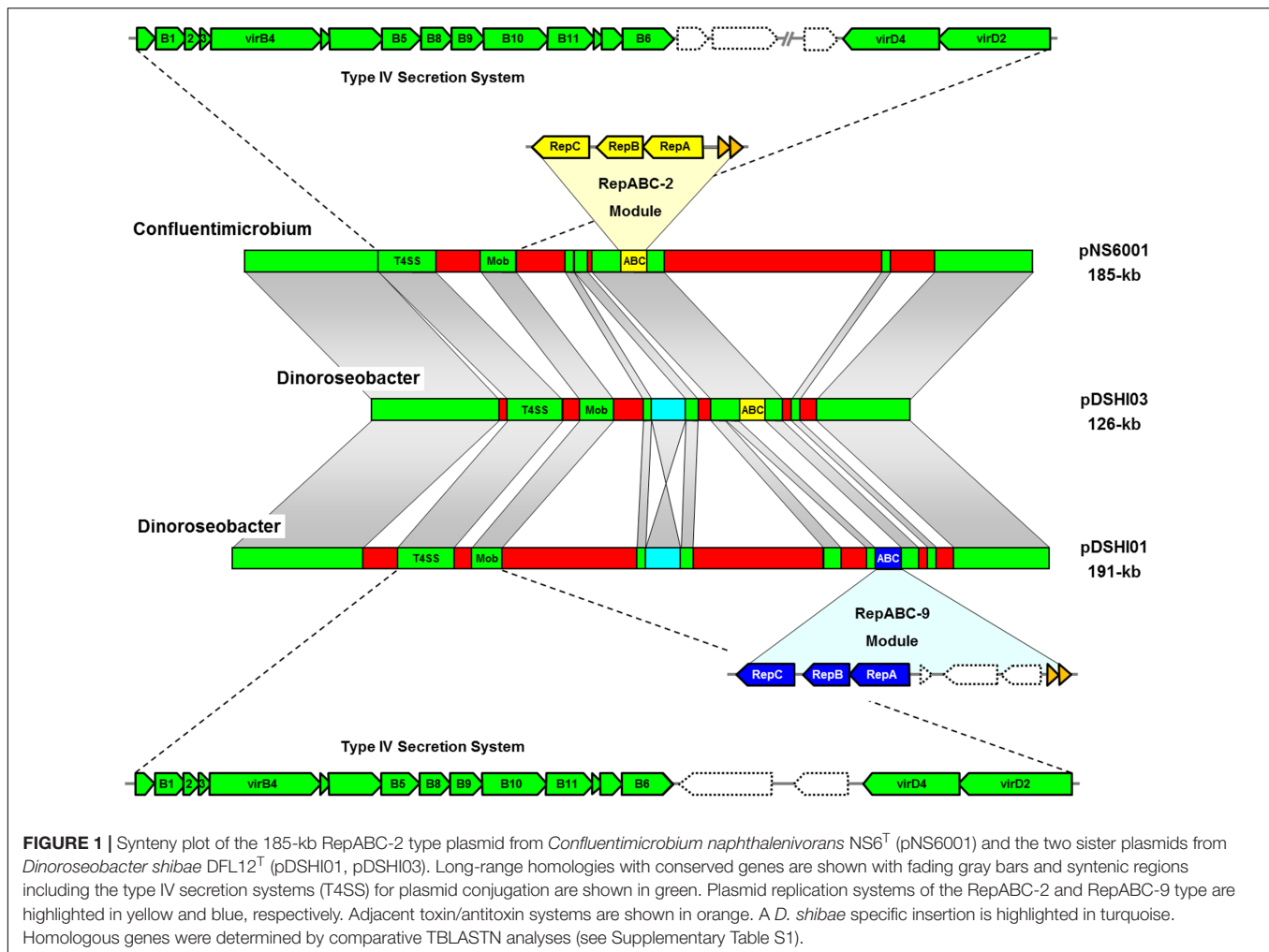
## RESULTS AND DISCUSSION

### Detection of Syntenic Plasmids

The improvement of sequencing technologies in the last decade resulted in an exponential increase of completely deciphered

bacterial genomes with more than 400 *Rhodobacteraceae* currently deposited in public data bases. We used the crucial replicases of the two sister plasmids from *D. shibae* for BLASTP searches and could identify a 185-kb plasmid from *Confluentimicrobium naphthalenivorans* NS6<sup>T</sup> (pNS6001; Jeong et al., 2015) that exhibits a long range synteny with the 126-kb plasmid pDSHI03 (**Figure 1**). Both replicons share the type IV secretion system (T4SS), toxin/antitoxin modules and a RepABC-2 type replication operon for plasmid maintenance (yellow), which has been replaced in the 191-kb sister plasmid pDSHI01 by a compatible RepABC-9 type equivalent (blue). Cytochrome *c* biosynthesis and heavy metal detoxification genes are the most conspicuous shared life style determinants of these extrachromosomal replicons (ECRs; Supplementary Table S1). Comprehensive TBLASTN comparisons revealed a protein sequence identity between 92 and 100% for most conserved genes, and the differences reflect their individual evolutionary history in the respective host cell. For example, the selective pressure on the functional genes is indirectly documented by an about 10% lower conservation of several hypothetical proteins, thus providing evidence that the plasmids pNS6001 and pDSHI03 were not recently transferred. However, 74% of the genes from pDSHI03 are shared with the homologous RepABC-2 plasmid from *Confluentimicrobium* (Supplementary Table S1). The structural backbone of both plasmids is absolutely conserved apart from some unique regions, such as the large 66-kb insertion in pNS6001 (**Figure 1**). The degree of sequence conservation even exceeds that of the stably coexisting 126-kb and 191-kb sister plasmids of *D. shibae* that harbor an inverted InDel (see below), but exhibit different compatible replication modules of the RepABC-2 and -9 type representing the ‘heart of a plasmid’ (Wagner-Döbler et al., 2010).

Our BLASTP searches also revealed the presence of a 135-kb contig with RepABC-2 replication module in the draft genome of *Roseovarius indicus* EhC03 (Rosana et al., 2016) that likely represents a second syntenic plasmid. The contig is less conserved than pNS6001, but still contains 45% of the genes from pDSHI03 including the T4SS (Supplementary Table S2). Systematic TBLASTN based comparisons of the respective replicons showed a slightly lower degree of protein conservation. The benefit of reciprocal BLAST analyses of conserved DNA modules is exemplified by the detection of four pseudogenes in *D. shibae*. The SAM-dependent methyltransferase from *R. indicus* (OAO03639.1) and *C. naphthalenivorans* (WP\_054540530.1) matches with two annotated genes of the 126-kb plasmid of *D. shibae* (Dshi\_4003 [WP\_050757929.1], Dshi\_4008 [WP\_012187351.1]) that are separated by a 6-kb fragment containing a type III restriction endonuclease and integrases (Dshi\_4004 to Dshi\_4007; highlighted in turquoise, **Figure 1**). This arrangement suggests that the functional gene in *Dinoroseobacter* was inactivated by a transposition event (Supplementary Table S2). A homologous but inverted insertion is present on the 191-kb plasmid (Dshi\_3695 to Dshi\_3698) resulting in two annotated SAM-dependent methyltransferase pseudogenes (Dshi\_3694 [WP\_012187109.1], Dshi\_3699 [WP\_012187104.1]), which is illustrated by the hourglass shaped syntenic area between the sister plasmids in **Figure 1**.

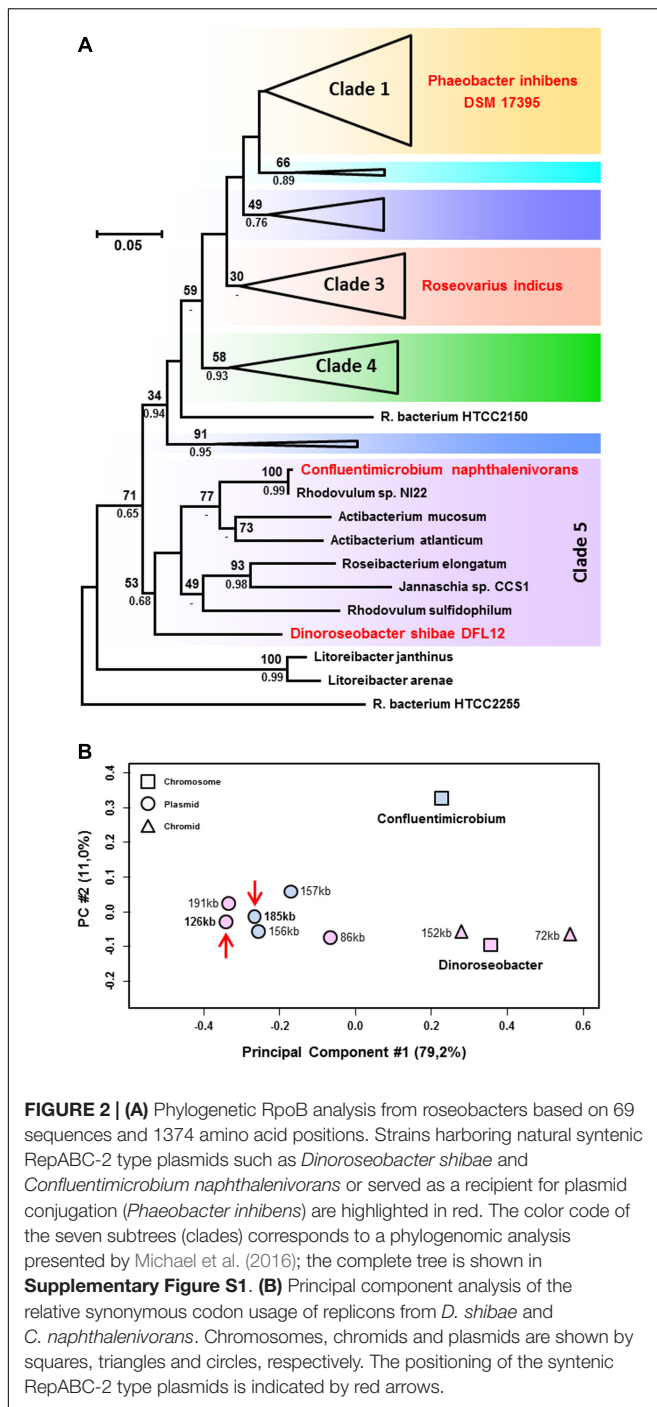


This case example documents that a thorough comparison of closely related sequences helps to improve the manual genome annotation of model organisms such as *D. shibae* DFL12<sup>T</sup> (Wagner-Döbler et al., 2010).

## RpoB Phylogeny as a First Proxy for the Taxonomic Positioning

A phylogenetic RaxML tree of the RNA polymerase beta subunit (RpoB) from 69 Roseobacter strains was calculated in order to reveal the relationships of the host cells of the syntenic RepABC-2 type plasmids (**Supplementary Figure S1**). The taxon sampling included the natural hosts *D. shibae*, *C. naphthalenivorans* as well as *R. indicus* and largely corresponded to that of a recent phylogenomic analysis of 65 Roseobacter genomes (Michael et al., 2016). Our single gene phylogeny recovered all seven clades as monophyletic groups and many subtrees were even supported with a solid bootstrap proportion (BP). Furthermore, it also mirrors the branching pattern of a phylogenomic analysis based on 44 roseobacters with a different taxon sampling including three single-cell genomes from an uncultured streamlined lineage from the ocean surface

(Luo et al., 2014). This outcome documents that the phylogenetic position of novel *Rhodobacteraceae* genomes can rapidly be estimated by a single gene RpoB analysis, which is based on 1374 amino acid (aa) alignment positions only, thus serving as a first proxy for complex phylogenomic studies with more than 200,000 aa positions (Case et al., 2007; Michael et al., 2016). The broader taxon sampling of eight genome sequenced strains from clade 5 (former analyses were restricted to two or three strains in this particular subgroup; Luo et al., 2014; Michael et al., 2016; Simon et al., 2017) recovered a very close branching of *Confluentimicrobium* with *Rhodovulum* sp. NI22 and a sistergroup relationship with the genus *Actibacterium* (77% BP; **Figure 2A**). The basal position of *Dinoroseobacter* shows a considerable evolutionary distance to *Confluentimicrobium*. The absence of syntenic RepABC-2 replicons and even plasmids of the same compatibility group from the remaining six Roseobacter strains in clade 5 is in agreement with a horizontal recruitment by *D. shibae* and *C. naphthalenivorans* from taxa outside of clade 5, a hypothesis that is based on the distinct codon usage (CU; see below) and thus more plausible than the alternative explanation of a common ancestry followed by differential plasmid losses. The phylogenetic position of *Roseovarius indicus* in clade 3 indicates



that the syntenic RepABC-2 plasmids (**Figure 1**; **Supplementary Tables S1, S2**) are migrating within the Roseobacter group. This prediction is supported by the experimental conjugation of pDSHI03 into *P. inhibens* DSM 17395 (Patzelt et al., 2016) thereby connecting clades 1 and 5 (**Figure 2A**). Syntenic RepABC-2 plasmid have thus been experimentally transferred between the most distant clades, and can naturally be detected in members of clade 5 and 3, thus documenting independent events of lateral gene transfer in the ocean.

## Codon Usage Analysis of *Confluentimicrobium* and *Dinoroseobacter* Replicons

The relative synonymous codon usage (RSCU) of the chromosome and ECRs from *D. shibae* and *C. naphthalenivorans* was analyzed using methods described previously (Petersen et al., 2013) and a plot of the averaged RSCU of protein-coding genes of all ten replicons is shown in **Figure 2B**. The principal component (PC) analysis confirmed the classification of *D. shibae*'s 152-kb and 72-kb ECRs as chromids, which exhibit a RSCU comparable to that of the chromosome (Harrison et al., 2010), whereas the three more distinct replicons of 86-kb, 126-kb and 191-kb represent authentic plasmids. The RSCU analysis of all *Confluentimicrobium* replicons showed a distinct localization of the three ECRs compared with the chromosome providing evidence that they also represent typical plasmids. This conclusion is independently supported by the presence of structurally conserved T4SS on the chromosome and all three ECRs (**Supplementary Table S3**) that contain RepABC operons of different compatibility groups (pNS6001 [185-kb, RepABC-2], pNS6002 [157-kb, RepABC-9], pNS6003 [156-kb, RepABC-5]). RepABC modules represent the most frequently mobilized plasmid type among the four characteristic replication systems of Roseobacter ECRs (RepA, RepB, DnaA-like, RepABC; Petersen et al., 2013). Based on the detection of a RepABC-9 operon on another replicon of *Confluentimicrobium*, the 157-kb plasmid pNS6002, which is homologous and functionally equivalent to that of the 191-kb sister plasmid pDSHI01 from *D. shibae* (blue module, **Figure 1**), we compared both plasmids via BLASTN, but observed only a limited structural conservation for the replication module and the generally conserved T4SS. However, the presence of a co-existing RepABC-2 and -9 plasmid pair in two Roseobacter species reflects a stable co-evolution of these ECRs. The close grouping of the syntenic 126-kb pDSHI03 and 185-kb pNS6001 plasmids in the RSCU, which is highlighted with red arrows in **Figure 2B**, suggests a common origin of both replicons. The large distance of both replicons to their respective chromosomes unequivocally documents that neither *D. shibae* nor *Confluentimicrobium* is the genuine host cell of this conjugative RepABC-2 type plasmid.

## Origin and Habitat of Roseobacter Isolates

*Dinoroseobacter shibae* DFL12<sup>T</sup> was isolated from a culture of the dinoflagellate *Prorocentrum lima* maintained at the Biological Research Station of the Alfred-Wegener-Institute for Polar Research (AWI) at Helgoland in the North Sea (Biebl et al., 2005). The algal culture was a gift from the Toralla Marine Science Station (ECIMAT) of the University of Vigo, Spain, where the dinoflagellate had been isolated from the North Atlantic Ocean. *C. naphthalenivorans* NS6<sup>T</sup> was isolated from heavily polluted tidal flat sediments in the South Sea, a part of the Pacific Ocean adjacent to South Korea (Jeong et al., 2015), at a distance of roughly 8000 km to Vigo, Spain. Thus, today those two species do not share a similar habitat or location. *Roseovarius indicus*

EhC03 was isolated from a culture of the coccolithophore alga *Emiliania huxleyi* M217 maintained at the Plymouth Algal Collection, United Kingdom. The algae had originally been isolated from surface water of the South Pacific (NCBI BioSample SAMN04965936; Rosana et al., 2016). The synteny of the *R. indicus* RepABC-2 plasmid with pDSHI03 of *D. shibae* is interesting and it could be speculated if those plasmids share conserved traits that are beneficial for the association of roseobacters with eukaryotic microalgae from phylogenetically distant phyla, e.g., haptophytes (*Emiliania huxleyi*) and dinoflagellates (*Prorocentrum lima*).

## PERSPECTIVE

The discovery of long-range synteny between conjugative plasmids in species from the Roseobacter group which are phylogenetically distant and geographically separated by 1000s of kilometers today provides the first proof for independent events of conjugation in the ocean that were stably maintained in the respective species. Plasmid mediated HGT likely plays an important role for the evolution of roseobacters.

## AUTHOR CONTRIBUTIONS

JP conceived the study and performed the analyses; JP and IWD wrote the manuscript.

## REFERENCES

- Aminov, R. I. (2011). Horizontal gene exchange in environmental microbiota. *Front. Microbiol.* 2:158. doi: 10.3389/fmicb.2011.00158
- Biebl, H., Allgaier, M., Tindall, B. J., Koblizek, M., Lünsdorf, H., Pukall, R., et al. (2005). *Dinoroseobacter shibae* gen. nov., sp. nov., a new aerobic phototrophic bacterium isolated from dinoflagellates. *Int. J. Syst. Evol. Microbiol.* 55, 1089–1096.
- Case, R. J., Boucher, Y., Dahllöf, I., Holmström, C., Doolittle, W. F., and Kjelleberg, S. (2007). Use of 16S rRNA and *rpoB* genes as molecular markers for microbial ecology studies. *Appl. Environ. Microbiol.* 73, 278–288. doi: 10.1128/AEM.01177-06
- Chandran Darbari, V., and Waksman, G. (2015). Structural biology of bacterial type IV secretion systems. *Annu. Rev. Biochem.* 84, 603–629. doi: 10.1146/annurev-biochem-062911-102821
- Harrison, P. W., Lower, R. P. J., Kim, N. K. D., and Young, J. P. W. (2010). Introducing the bacterial “chromid”: not a chromosome, not a plasmid. *Trends Microbiol.* 18, 141–148. doi: 10.1016/j.tim.2009.12.010
- Jeong, H. I., Jin, H. M., and Jeon, C. O. (2015). *Confluentimicrobium naphthalenivorans* sp. nov., a naphthalene-degrading bacterium isolated from sea-tidal-flat sediment, and emended description of the genus *Confluentimicrobium* Park et al. 2015. *Int. J. Syst. Evol. Microbiol.* 65, 4191–4195. doi: 10.1099/ijsem.0.000561
- Koonin, E. V. (2016). Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. *F1000Research* 5, 1–9. doi: 10.12688/f1000research.8737.1
- Liu, Y. Y., Wang, Y., Walsh, T. R., Yi, L. X., Zhang, R., Spencer, J., et al. (2016). Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect. Dis.* 16, 161–168. doi: 10.1016/S1473-3099(15)00424-7
- Luo, H., Swan, B. K., Stepanauskas, R., Hughes, A. L., and Moran, M. A. (2014). Evolutionary analysis of a streamlined lineage of surface ocean Roseobacters. *ISME J.* 8, 1428–1439. doi: 10.1038/ismej.2013.248

## FUNDING

This work was supported by the Transregional Collaborative Research Center “Roseobacter” (Transregio TRR 51) of the Deutsche Forschungsgemeinschaft.

## ACKNOWLEDGMENTS

We would like to thank Pascal Bartling and Henner Brinkmann for providing codon usage and phylogenetic analyses and two reviewers for their constructive comments.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2017.01350/full#supplementary-material>

**FIGURE S1 |** Phylogenetic Maximum Likelihood tree of the RNA-polymerase beta subunit from roseobacters (RaxML, LG + F + 4 $\Gamma$ ). The RpoB analysis was based on 69 sequences and 1374 amino acid positions. The statistical support for the internal nodes of the RaxML tree was determined by 100 bootstrap replicates (BR) and values  $\geq 30\%$  are shown (upper value). Posterior probabilities were calculated with PhyloBayes v3 (CATGTR + 4 $\Gamma$ ; lower value). Strains that harbor syntenic RepABC-2 type plasmids (*D. shibae*, *C. naphthalenivorans*, *R. indicus*) or served as conjugational recipient (*P. inhibens*) are highlighted in red. The color code of the seven clades corresponds to a phylogenomic analysis presented by Michael et al. (2016).

- McGann, P., Snesrud, E., Maybank, R., Corey, B., Ong, A. C., Hinkle, M., et al. (2016). *Escherichia coli* harboring *mcr-1* and *bla*<sub>CTX-M</sub> on a novel IncF plasmid: first report of *mcr-1* in the United States. *Antimicrob. Agents Chemother.* 60, 4420–4421. doi: 10.1128/AAC.01103-16
- Michael, V., Frank, O., Bartling, P., Scheuner, C., Göker, M., Brinkmann, H., et al. (2016). Biofilm plasmids with a rhamnase operon are widely distributed determinants of the “swim-or-stick” lifestyle in roseobacters. *ISME J.* 10, 2498–2513. doi: 10.1038/ismej.2016.30
- Palmer, K. L., Kos, V. N., and Gilmore, M. S. (2010). Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. *Curr. Opin. Microbiol.* 13, 632–639. doi: 10.1016/j.mib.2010.08.004
- Patzelt, D., Michael, V., Päufer, O., Ebert, M., Tielen, P., Jahn, D., et al. (2016). Gene flow across genus barriers - Conjugation of *Dinoroseobacter shibae*'s 191-kb killer plasmid into *Phaeobacter inhibens* and AHL-mediated expression of type IV secretion systems. *Front. Microbiol.* 7:742. doi: 10.3389/fmicb.2016.00742
- Perry, J. A., and Wright, G. D. (2013). The antibiotic resistance “mobilome”: searching for the link between environment and clinic. *Front. Microbiol.* 4:138. doi: 10.3389/fmicb.2013.00138
- Petersen, J. (2011). Phylogeny and compatibility: plasmid classification in the genomics era. *Arch. Microbiol.* 193, 313–321. doi: 10.1007/s00203-011-06869
- Petersen, J., Brinkmann, H., and Pradella, S. (2009). Diversity and evolution of *repABC* type plasmids in *Rhodobacterales*. *Environ. Microbiol.* 11, 2627–2638. doi: 10.1111/j.1462-2920.2009.01987.x
- Petersen, J., Frank, O., Göker, M., and Pradella, S. (2013). Extrachromosomal, extraordinary and essential - the plasmids of the Roseobacter clade. *Appl. Microbiol. Biotechnol.* 97, 2805–2815. doi: 10.1007/s00253-013-47468
- Pradella, S., Päufer, O., and Petersen, J. (2010). Genome organisation of the marine Roseobacter clade member *Marinovum algicola*. *Arch. Microbiol.* 192, 115–126. doi: 10.1007/s00203-009-05352
- Rosana, A. R. R., Orata, F. D., Xu, Y., Simkus, D. N., Bramucci, A. R., Boucher, Y., et al. (2016). Draft genome sequences of seven bacterial strains isolated from a



- polymicrobial culture of coccolith-bearing (C-Type) *Emiliana huxleyi* M217. *Genome Announc.* 4:e673-16. doi: 10.1128/genomeA.00673-16
- Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., et al. (2017). Phylogenomics of *Rhodobacteraceae* reveals evolutionary adaptation to marine and non-marine habitats. *ISME J.* 11, 1483–1499. doi: 10.1038/ismej.2016.198
- Sobecky, P. A., and Hazen, T. H. (2009). Horizontal gene transfer and mobile genetic elements in marine systems. *Methods Mol. Biol.* 532, 435–453. doi: 10.1007/978-1-60327-853-9\_25
- The, H. C., Thanh, D. P., Holt, K. E., Thomson, N. R., and Baker, S. (2016). The genomic signatures of *Shigella* evolution, adaptation and geographical spread. *Nat. Rev. Microbiol.* 14, 235–250. doi: 10.1038/nrmicro.2016.10
- Thibault, S., and Top, E. (2016). Plasmid transfer in biofilms: a perspective on limitations and opportunities. *NPJ Biofilms Microbiomes* 2:16022. doi: 10.1038/npjbiofilms.2016.22
- Wagner-Döbler, I., Ballhausen, B., Berger, M., Brinkhoff, T., Buchholz, I., Bunk, B., et al. (2010). The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker's guide to life in the sea. *ISME J.* 4, 61–77. doi: 10.1038/ismej.2009.94
- Wang, H., Tomasch, J., Michael, V., Bhujju, S., Jarek, M., Petersen, J., et al. (2015). Identification of genetic modules mediating the Jekyll and Hyde interaction of *Dinoroseobacter shibae* with the dinoflagellate *Prorocentrum minimum*. *Front. Microbiol.* 6:1262. doi: 10.3389/fmicb.2015.01262
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Copyright © 2017 Petersen and Wagner-Döbler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Advantages of publishing in Frontiers



## OPEN ACCESS

Articles are free to read  
for greatest visibility  
and readership



## FAST PUBLICATION

Around 90 days  
from submission  
to decision



## HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,  
and constructive  
peer-review



## TRANSPARENT PEER-REVIEW

Editors and reviewers  
acknowledged by name  
on published articles

## Frontiers

Avenue du Tribunal-Fédéral 34  
1005 Lausanne | Switzerland

**Visit us:** [www.frontiersin.org](http://www.frontiersin.org)

**Contact us:** [info@frontiersin.org](mailto:info@frontiersin.org) | +41 21 510 17 00



## REPRODUCIBILITY OF RESEARCH

Support open data  
and methods to enhance  
research reproducibility



## DIGITAL PUBLISHING

Articles designed  
for optimal readership  
across devices



## FOLLOW US

[@frontiersin](https://twitter.com/frontiersin)



## IMPACT METRICS

Advanced article metrics  
track visibility across  
digital media



## EXTENSIVE PROMOTION

Marketing  
and promotion  
of impactful research



## LOOP RESEARCH NETWORK

Our network  
increases your  
article's readership