ORIGINAL ARTICLE

Global occurrence and heterogeneity of the *Roseobacter*-clade species *Ruegeria mobilis*

Eva C Sonnenschein¹, Kristian F Nielsen¹, Paul D'Alvise^{1,8}, Cisse H Porsby^{1,2}, Jette Melchiorsen¹, Jens Heilmann³, Panos G Kalatzis^{4,5}, Mario López-Pérez⁶, Boyke Bunk⁷, Cathrin Spröer⁷, Mathias Middelboe⁵ and Lone Gram¹

¹Department of Systems Biology, Technical University of Denmark, Kongens Lyngby, Denmark; ²Biogen Manufacturing, Biogen Idec Allé 1, Hillerød, Denmark; ³Technical University of Denmark, National Institute for Aquatic Resources, Charlottenlund, Denmark; ⁴Hellenic Centre for Marine Research, Institute of Aquaculture, Heraklion, Greece; ⁵Section for Marine Biology, University of Copenhagen, Helsingør, Denmark; ⁶División de Microbiología, Evolutionary Genomics Group, Universidad Miguel Hernández, San Juan, Alicante, Spain and ⁷Helmholtz Centre for Infection Research, Braunschweig, Germany

Tropodithietic acid (TDA)-producing Ruegeria mobilis strains of the Roseobacter clade have primarily been isolated from marine aquaculture and have probiotic potential due to inhibition of fish pathogens. We hypothesized that TDA producers with additional novel features are present in the oceanic environment. We isolated 42 TDA-producing R. mobilis strains during a global marine research cruise. While highly similar on the 16S ribosomal RNA gene level (99-100% identity), the strains separated into four sub-clusters in a multilocus sequence analysis. They were further differentiated to the strain level by average nucleotide identity using pairwise genome comparison. The four sub-clusters could not be associated with a specific environmental niche, however, correlated with the pattern of sub-typing using co-isolated phages, the number of prophages in the genomes and the distribution in ocean provinces. Major genomic differences within the sub-clusters include prophages and toxin-antitoxin systems. In general, the genome of R. mobilis revealed adaptation to a particle-associated life style and querying TARA ocean data confirmed that R. mobilis is more abundant in the particle-associated fraction than in the free-living fraction occurring in 40% and 6% of the samples, respectively. Our data and the TARA data, although lacking sufficient data from the polar regions, demonstrate that R. mobilis is a globally distributed marine bacterial species found primarily in the upper open oceans. It has preserved key phenotypic behaviors such as the production of TDA, but contains diverse sub-clusters, which could provide new capabilities for utilization in aquaculture.

The ISME Journal (2017) 11, 569-583; doi:10.1038/ismej.2016.111; published online 23 August 2016

Introduction

Aquaculture provides nearly 50% of the global supply of fish and shellfish for human consumption, however, production is constrained by bacterial diseases. To reduce the use of antibiotics, we and others have focused on use of probiotic strategies and used bacteria of the *Roseobacter* clade that can antagonize fish-pathogenic bacteria without harming the fish or their live feed (D'Alvise *et al.*, 2012, 2013). Specifically, strains belonging to the

species Ruegeria mobilis and Phaeobacter inhibens are potent antagonists of the fish pathogens (Planas et al., 2006; Porsby et al., 2008; D'Alvise et al., 2010) primarily caused by production of the antibacterial compound tropodithietic acid (TDA) (D'Alvise et al., 2010, 2012). TDA is also bactericidal against human pathogens including Salmonella enterica and Staphylococcus aureus (Porsby et al., 2011) and it has been recently been demonstrated that TDA acts by disrupting the proton motive force (Wilson *et al.*, 2016). TDA is produced by R. mobilis, P. inhibens and Phaeobacter gallaeciensis, as well as by other alphaproteobacteria belonging to Pseudovibrio (Porsby et al., 2008; Geng and Belas, 2010; Penesyan et al., 2011; Harrington et al., 2014). Strains of Ruegeria and Phaeobacter produce TDA when grown under stagnant conditions in marine broth (MB) and TDA production coincides with bacterial rosette formation, a biofilm at the air-liquid interface, and production of a brown pigment that is

Correspondence: EC Sonnenschein, Department of Systems Biology, Technical University of Denmark, Matematiktorvet 301, 2800 Kongens Lyngby, Denmark.

E-mail: evaso@bio.dtu.dk

⁸Current address: University of Hohenheim, Institute for Animal Science, Population genomics group, Garbenstr. 17, 70599 Stuttgart, Germany

Received 27 November 2015; revised 12 June 2016; accepted 3 July 2016; published online 23 August 2016

involved in quorum sensing (QS) (Wagner-Döbler and Biebl, 2006; Martens et al., 2007; Zan et al., 2014). In P. gallaeciensis, TDA production is influenced by acylated homoserine lactones (Berger et al., 2011) and in Ruegeria (Silicibacter) sp. TM1040, TDA upregulates its own production (Geng and Belas, 2010). The antibacterial activity of these Roseobacter-clade strains could have a role in their natural niche, for example, during growth on particles, such as algae and marine snow, and it allows them to invade already formed bacterial biofilms (Rao et al., 2005); however, it is unknown how wide these specific bioactive strains are distributed on a global scale and in which environmental niches they appear.

an iron-TDA complex (Bruhn et al., 2007; Geng

et al., 2008; Porsby et al., 2008; D'Alvise et al., 2015).

Several roseobacters produce compounds potentially

Alphaproteobacteria and species belonging to the *Roseobacter* clade are widely distributed in marine environments and have a major role in oceanic sulfur cvcling (Wagner-Döbler and Biebl, 2006). Roseobacters have been isolated from seawater, marine sediments, surfaces of marine organisms, and hypersaline ponds (Buchan and Moran, 2005) and can constitute up to 2-8% of surface water bacterioplankton (Wietz et al., 2010; Sunagawa et al., 2015), but the clade can represent as much as 20–40% of sequence data from 16S ribosomal RNA (rRNA) gene libraries especially during algal blooms (Moran et al., 2003; Buchan and Moran, 2005; Prabagaran et al., 2007). Many of the Roseobacter-clade strains are associated with algae and can metabolize dimethylsulfoniopropionate produced by algae (González et al., 2000; Alavi et al., 2001; Miller and Belas, 2004; Buchan and Moran, 2005; Todd et al., 2009). Cosmopolitan distribution has been demonstrated for the SAR11 clade also belonging to the alphaproteobacteria (Wietz et al., 2010; Giebel et al., 2011; Ghiglione et al., 2012). The clade constitutes the dominant prokaryotic population in surface waters (Morris et al., 2002). While the Roseobacter clade as such is also widespread, very little is known about its distribution at the species level. Selje *et al.* (2004) found that a particular phylotype of the *Roseobacter* clade affiliated cluster was widely distributed in temperate and polar waters, but was not detected in the tropical and sub-tropical areas. The abundance of the *Roseobacter* clade affiliated cluster correlates positively with phaeopigments and chlorophyll (Giebel *et al.*, 2011) indicating an algal association. We recently reported that *Roseobacter*-clade strains identified as *Ruegeria* spp. could be isolated from upper ocean waters (Gram et al., 2010), however, these strains were not further characterized.

The purpose of the present study was to analyze the global distribution of *R. mobilis* in terms of bioactivity and phylogeny and to determine if strain characteristics could be related to biogeographic patterns. In comparison to aquaculture-derived R. mobilis strains, new isolates from the marine environment could represent a novel diversity within the species.

Materials and methods

Isolation of Roseobacter-clade strains from marine samples

A total of 533 samples were analyzed: 224 seawater samples from 120 locations, 217 surface swab samples (54 locations) and 92 samples (18 locations) taken from material of a ship sea chest filter were collected during a global marine research cruise, Galathea 3 (Gram et al., 2010). Samples were 10-fold serially diluted in sterile seawater, plated on 50% marine agar (Difco 2216, Franklin Lakes, NJ, USA) and incubated at 20 °C until visible colony growth (3-7 days). Colonies were replica-plated onto a seawater agar cast with Vibrio anguillarum strain 90-11-287 (Skov et al., 1995) and incubated at 25 °C for 1-2 days. Colonies causing inhibition of Vibrio growth were isolated from the original plates. Details of sampling procedures, isolation and identification by 16S rRNA gene sequence analyses have been described (Gram et al., 2010). Strains isolated with this procedure were denoted S in Supplementary Table S1. Additionally, two techniques for the enrichment of biofilm-forming bacteria and TDAresistant strains-both criteria of TDA-producing Roseobacter strains (Bruhn et al., 2005; Porsby et al., 2008)—were utilized and the obtained isolates (denoted as F and G, respectively, in Supplementary Table S1) were included in the analyses. We attempted specific isolation of bacteria forming biofilm at broth surfaces as TDA-producing Roseo*bacter* strains form biofilms of star-shaped aggregates at the air-liquid interface (Bruhn et al., 2005). To enrich biofilm-forming bacteria, at some sampling stations. 0.1 ml seawater was mixed with 5 ml MB (Difco 2216) diluted 50% with filter-sterilized local seawater. Tubes were left stagnant at 20 °C for up to 3 weeks and enrichment cultures were tested for inhibitory activity by spotting 10 µl of a the enriched culture on a seawater agar cast with *V. anguillarum*. In several tubes, a characteristic brownish color developed (Supplementary Figure S1), and this is typical of iron-TDA complexes (D'Alvise et al., 2015). Enrichments were streaked on marine agar, and pure cultures were isolated and re-tested for anti-Vibrio activity. Strains isolated with this procedure are denoted F in Supplementary Table S1. To enrich TDA-resistant strains, a few samples were also mixed with sterile-filtered supernatant from a TDA-producing Roseobacter-clade strain (Phaeobacter sp. 27-4 grown 3-4 days in MB under stagnant conditions) to select for TDA-tolerant organisms. Samples from these enrichments were also streaked and replica-plated and pure cultures isolated from colonies with anti-Vibrio activity. Strains isolated with this procedure are denoted G in Supplementary Table S1.

Assessment of metadata obtained from the Galathea expedition

A number of physicochemical environmental parameters were measured by standard methods at each station where water samples were taken. These included temperature, salinity, *in vivo* fluorescence, oxygen concentration and concentrations of inorganic nutrients (total nitrogen [N] and total phosphorus [P], nitrite $[NO_2^-]$, nitrate $[NO_3^-]$, ammonium $[NH_4^+]$, phosphate $[PO_4^-]$ and silica $[SiO_2]$). The distance from each *R. mobilis* sampling site to the coast was calculated on the ERDDAP server (https:// bluehub.jrc.ec.europa.eu/erddap/griddap/hawaii_92 a1_b23f_19c0.html). Ocean provinces and biomes (Longhurst, 2007) of the sampling sites were identified on the ArcGIS platform (http://www.arcgis.com/ home/webmap/viewer.html?useExisting=1).

Identification and antibacterial activity of marine Roseobacter-clade strains and strains from fish farms and culture collections

All strains were identified as *R. mobilis* by initial 16S rRNA gene sequence analysis (Gram *et al.*, 2010) and antibacterial activity confirmed on V. anguillarumembedded plates, however, we cannot exclude the presence of non-bioactive strains of the species at these locations. The isolated R. mobilis strains were re-evaluated for production of antibacterial compound(s) in MB. Strains were incubated either with aeration (200 r.p.m.) or under stagnant conditions for 3 days at 25 °C. Sterile-filtered culture supernatants were tested for inhibitory activity against V. anguillarum strain 90-11-287 cast in an agar consisting of 3% Instant Ocean Sea Salt (Aquarium Systems Inc., Sarrebourg, France), 0.3% casamino acids and 0.4% glucose as previously described (Porsby et al., 2008). We included two R. mobilis from culture collections, and three isolates from turbot farms (Hjelm et al., 2004; Porsby et al., 2008) (M strains in Supplementary Table S1) in the comparison. Culture collection strains were *R. mobilis* NBRC101030^T and *R. mobilis* (formerly *R. pelagia*) NBRC102038^T. We earlier reported that R. mobilis and R. pelagia could not be distinguished based on 16S rRNA gene homology (Porsby et al., 2008) and R. pelagia is a later synonym of *R. mobilis* (Lai *et al.*, 2010). For comparative purposes, we included a broader range of Roseobacter-clade bacteria, including several TDA producers (P. inhibens strains DSM17395 and 2.10, Ruegeria (Silicibacter) sp. TM1040 and P. inhibens DSM16374^T (T5)) and several assumed non-TDA producers: *Roseobacter denitrificans* DSM7001^T, Roseobacter litoralis DSM6996^T, Roseovarius tolerans DSM11467^T, Leisingera methylohalidivorans DSM14336^T, Sulfitobacter pontiacus DSM10014^T, Ruegeria pomerovi DSM15171^T (DSS-3), Marinovum algicola DSM10251^T, Ruegeria atlantica DSM5823^T, Ruegeria lacuscaerulensis DSM11341^T and Ruegeria scottomollicae LGM24367^T.

Phylogenetic analyses

16S rRNA genes were extracted from genome sequences (Supplementary Table S2) using CLC Genomics Workbench, version 7 (CLC Bio, Aarhus, Denmark) or retrieved from Genbank. The 16S rRNA gene sequence of *V. anguillarum* strain 90-11-287 (Genbank acc. no. KP792697) was used as outgroup. MUSCLE alignment and neighbor joining tree were generated in MEGA 6.06 (Tamura *et al.*, 2013). Bootstrap values were based on 1000 replicates.

The multilocus sequence analysis (MLSA) was performed using nine housekeeping genes (16S rRNA gene, gapA, gyrB, ftsZ, mreB, pyrH, recA, rpoA and topA) (Sawabe et al., 2013). Sequences were extracted from genome sequences, aligned and concatenated using CLC Genomics Workbench. A neighbor joining tree was generated in MEGA 6.06 (Tamura et al., 2013) with bootstrap values based on 1000 replicates.

Pairwise genome comparison of the Ruegeria strains, all currently available Ruegeria genome sequences (R. pomeroyi DSS-3, R. lacuscaerulensis ITI-1157, *R. halocynthiae* MOLA R1/13b, R. conchae TW15, Ruegeria (Silicibacter) sp. TrichCH4B, TM1040, R11, ANG-R, ANG-S4) and inhibens DSM17395 was conducted with Р BLAST-based average nucleotide identity (ANI) using the 'calculate_ani' python script (Pritchard, 2014) based on work by Richter and Rosselló-Móra and Goris et al. (Goris et al., 2007; Richter and Rosselló-Móra, 2009). Heatmaps were generated in R using complete agglomeration clustering based on Euclidean distances (Boston, MA, USA).

Metagenomic data analysis

Metagenomic recruitments were carried out between bacterial genomes (five representative R. mobilis genomes (F1926, S1424, NBRC101030, NBRC102038, 270-3) and nine genomes of reference strains (R. pomeroyi DSS-3, 'Candidatus Pelagibacter ubique' HTCC1062, V. anguillarum 90-11-287 and the Galathea isolates *Pseudoalteromonas piscicida* S2040, P. rubra S2471, P. ruthenica S3137, Photobacterium galathea S2753, V. coralliilyticus S2052, V. galathea S2757) (Gram et al., 2010)) and the metagenomic data of the TARA expedition (two fractions: free-living (0.22–5 µm, 426 metagenomes) attached (5–20 µm, 370 metagenomes)) and (Sunagawa et al., 2015) using BLASTN (Altschul et al., 1997). A hit was considered only when it was at least 50 nucleotides long, with an identity of >95% and with an *E* value $\leq 1e-5$. These hits were used to compute the RPKG (reads recruited per kb of genome per Gb of metagenome) values, that provide normalization across various metagenomics а results. An RPKG of >1 was considered acceptable to demonstrate the presence of the strain.

Chemical analysis of TDA and acylated homoserine lactones

Production of TDA and OH-C₁₀ homoserine lactone (HSL) was tested from sterile-filtered culture supernatants by reversed phase liquid chromatographytandem mass spectrometry on a Agilent 1100 high performance liquid chromatography system (Agilent Technologies, Waldbronn, Germany) coupled to a Micromass Quattro Ultima triple quadrupole MS (Waters, Manchester, UK) for TDA as described by Porsby et al. (Porsby et al., 2008) and for acylhomoserine lactones (AHLs) as described in Rau et al. (Rau et al., 2010). OH-C₁₀-HSL lactone was identified by $272 \rightarrow 102@\ 20 \text{ eV}$ (quantifier) and 272 \rightarrow 74@ 30 eV (qualifier); and OH-C₁₀-HSL-acid (open lactone) $290 \rightarrow 102$ @ $20 \,\mathrm{eV}$ (quantifier) and $290 \rightarrow 171@30 \text{ eV}$ (qualifier). Specifically, TDA was confirmed against an authentic standard (BioViotica, Dransfeld, Germany), and so was OH-C₁₀-HSL (Nottingham University, UK, www.nottingham.ac.uk/ quorum/compounds.htm) (Rasmussen et al., 2014; D'Alvise *et al.*, 2015).

Bacteriophage isolation and sub-typing

Samples for isolation of *R. mobilis*-infecting bacteriophages were obtained from surface water (10 m water depth) at 120 stations along the Galathea route. From each station, 60 ml sample was 0.2 µm filtered and phages in the filtrate were concentrated by mild centrifugation (500 g, 15-30 min) using 15 ml Amicon Ultra spin devices (30 000 MW cutoff, Millipore, Billerica, MA, USA). The concentrate (1-1.5 ml) was mixed with an equal amount of suspension medium buffer (50 mM Tris-Cl pH 7.5, 99 mM NaCl, 8 mM MgSO₄, 0.01% gelatin) in a cryovial and kept at 5 °C until further analysis. Before phage isolation, the concentrates from individual cruise legs were pooled resulting in 10 groups of 10–15 phage concentrates each representing a different geographical area visited on the cruise (Supplementary Table S3). Phage proliferation and isolation was performed using enrichment cultures and subsequent spot assay using R. mobilis F1926 as target strain, basically following the protocol of Stenholm et al. (Stenholm et al., 2008). Firstly, 7 ml mid-log-phase bacterial culture were added to 250 µl phage concentrate and incubated for proliferation of potential R. mobilis-infecting phages present in the concentrate. After 24 h, the culture was centrifuged $(10\,000\,g,\,5\,\text{min})$ and the supernatant was $0.2\,\mu\text{m}$ filtered and stored at 5 °C. For isolation of specific phages, 5 µl of the 0.2 µm filtered enrichment cultures were spotted onto lawns of R. mobilis F1926 in soft agar, and incubated for 24 h. In the case of a clearing zone (plaque) in the spotted area, indicating phage lysis of the host, the plaque was transferred to 1 ml suspension medium buffer and eluted for at least 2 h. To purify individual lytic *R*. mobilis F1926 phages obtained from the spot assay, plaque assays were performed with serial dilutions of the isolated phages using the standard double layer method with *R. mobilis* F1926 as host strain. From these plaque assays, single plaques were picked with a sterile glass pipette and the phages were eluted in suspension medium buffer. Eight bacteriophages were isolated by three repeated rounds of plaque purification and reinfection.

To determine the susceptibility of the collection of *R. mobilis* isolates to the isolated phages, the phage host range was determined by spotting 2μ l of purified bacteriophage on top of a marine agar plate freshly prepared with 4 ml of soft agar inoculated with 0.3 ml of a mid-log-phase culture (OD₆₀₀ = 0.3) of the strain to be tested. The host-range was determined on three separate plates for each phagehost combination, and all 47 bacterial strains were tested against all the phage isolates.

Statistical analysis

One-way analysis of variance was used to evaluate if the *R. mobilis* sub-clusters identified by ANI were correlated to the oceanographic parameters of the Galathea cruise and those associated with the ocean provinces (productivity, chlorophyll, photic depth) (Longhurst, 2007), phage susceptibility, genome characteristics, number of prophages and CRISPRs and the presence of AHL synthase type 2. Whether the source of isolation was different between R. mobilis and non-R.mobilis isolates from the Galathea expedition and whether the isolation strategy, the distance to the coast, ocean provinces and biomes as defined by Longhurst (Longhurst, 2007) were different between the ANI sub-clusters of *R. mobilis* was assessed using the Fisher's exact test with the 'fisher.test' R script. Two-tailed *t*-test was applied for comparison of phage susceptibility between co-isolated strains and strains isolated outside leg 6. Furthermore, two-tailed t-test was used to evaluate the difference of oceanographic data (depth, nitrate, oxygen, salinity, chlorophyll, temperature; data obtained from NCBI) of the TARA expedition depending on the abundance of a strain (RPKG > 1: present or RPKG<1: low abundance or absent) and was also performed for the TARA recruitment data of surface samples from the free-living versus particle-associated fraction for each strain.

Results

Isolation and occurrence of R. mobilis

Of the 533 samples obtained from 294 sampling sites, 519 bioactive strains were isolated by surface and subsequent replica plating (Gram *et al.*, 2010). Out of these bioactive strains, eighteen isolates were identified as *R. mobilis* by 16S rRNA gene analysis and included in the present study (S strains, Supplementary Table S1). In addition, enrichment cultures with visible formation of a brown pigment (Supplementary Figure S1) were used for isolation. Sterile-filtered supernatants from 19 brownish enrichment cultures caused inhibition of *V. anguillarum*, whereas

anti-Vibrio activity was never found in nonpigmented enrichments. All 19 strains isolated from these enrichments were brown-pigmented and subsequently identified as *R. mobilis* (F strains, Supplementary Table S1). Five additional *R. mobilis* strains were isolated by enriching seawater samples in sterile-filtered supernatant from the TDA-producing *Phaeobacter* sp. 27-4 (G strains, Supplementary Table S1). Only *R. mobilis* and no other bacteria were isolated from these enrichments.

The total of 42 *R. mobilis* strains (8% of all Galathea isolates) were obtained from 29 of the 294 tested sampling sites (Figure 1). *R. mobilis* was primarily isolated from the open ocean (12 nautical miles from shore, 71%) and with a lesser extend from the high seas (200 nautical miles from shore, 19%) and coastal waters (10%) (Figure 1). The strains were obtained from surface waters from 5 to 120 m, besides one strain being isolated from 400 m depth (Supplementary Table S4). From five samples obtained from four sampling sites (one site represented by two depths), ten *R. mobilis* strains were isolated from the same sample by both direct plating (followed by replica plating) and by enrichment in

MB (marked in bold in Supplementary Table S3). However, in the majority of samples (the remaining 32), strains were isolated by only one of these procedures with the enrichment being more successful (19 of 32 isolates) compared with the plating/ replica plating (13 of 32 isolates).

The Galathea R. mobilis strains originated from samples with different oceanographic and chemical profiles (Supplementary Table S4). The sampling sites were located in all of the four biomes (coastal, polar, trades, westerlies) and in 14 of the 54 ocean provinces as defined by Longhurst (Longhurst, 2007). Temperature during sample isolation varied from 4 to 30 °C and oxygen concentrations from 0.1 to 7 mll⁻¹. Total nitrogen concentrations ranged from 6 to 50 µM and covered a wide range of NO_3^{-} concentrations from 0 to 40 µM at the South American West Coast associated with the coastal upwelling sites. *R. mobilis* strains were isolated from sites that varied in *in vivo* chlorophyll a fluorescence from 0 to 5 V (relative units). Thirty-seven of the R. mobilis strains were obtained from seawater samples and 2 and 3 from the surface of copepods and sediment, respectively. This is in contrast



Figure 1 Global appearance of *Ruegeria mobilis*: stations of the Galathea 3 cruise, where the herein described 42 strains of *R. mobilis* were isolated (blue dots) and stations of the TARA expedition (Sunagawa *et al.*, 2015), where *R. mobilis* was identified in the metagenomic data in the attached (green circles) and pelagic fraction (red circles). Mapbox, Data ODbL OpenStreetMap contributors.

(P < 0.0001) to the non-*Ruegeria* Galathea isolates that were obtained equally from each of the three sources of isolation (seawater, surface, sea chest) (Gram *et al.*, 2010). Interestingly, the surfaces investigated during the expedition included many different organisms (copepods, sponges, fish and so on) and materials (lava stones, sediment, and so on), however, the herein described *R. mobilis* strains were only isolated from copepods or sediment.

Querying metagenomic data of the TARA expedition (Sunagawa et al., 2015) with five selected *R. mobilis* strains revealed additional locations of *R. mobilis* appearance (Figure 1). Overall, the analyzed R. mobilis strains have a similar distribution pattern to R. pomerovi DSS-3 and are unique in comparison to the other analyzed species (Supplementary Figure S2). Occurrence of R. mobilis decreased with increasing depth in the particleassociated fraction. R. mobilis NBRC101030 was detected (RPKG>1) in the free-living fraction in 25 of 426 samples (6%) at 19 of 70 stations (Supplementary Table S5a) and in the particleassociated fraction in 146 of 370 samples (40%) at 42 of 47 stations (Supplementary Table S5b). In comparison, the stream lined species 'Candidatus Pelagibacter ubique' HTCC1062 is present in 57% of the samples of the free-living fraction and 71% of those of the particle-associated fraction. In contrast, the herein used target bacterium, the fish pathogen V. anguillarum 90-11-287 has a low abundance (4% in each fraction).

We further evaluated if the bacterial abundance found in the TARA metagenomic data was linked to oceanographic parameters (depth, nitrate, oxygen, salinity, chlorophyll, temperature) measured during the TARA expedition (Supplementary Figure S3, Supplementary Table S6). Comparing the data associated with a RPKG < 1 (no to low abundance) with those associated with a RPKG > 1 (strain is present) demonstrated no significant correlation of bacterial occurrence and oceanographic parameters for R. mobilis in the particle-associated fraction. In contrast, occurrence of *R. mobilis* NBRC101030 in the free-living fraction is linked to depth, nitrate and temperature, demonstrating that the species could have a preference for environments in greater depth, with lower temperature and higher nitrate concentration (median: 90 m, 10.5 °C, 28.6 µmol l⁻¹, respectively). Two of the five tested R. mobilis strains also demonstrate a significant preference for lower levels of chlorophyll (median: 0.03 (S1424) and 0.006 (270-3) versus $0.16 (both) \text{ mg m}^{-3}$.

Phylogeny and genome analysis of R. mobilis strains from the Galathea cruise

All 42 strains were Gram-negative motile rods, produced a brownish pigment when grown under static conditions in MB, and formed rosettes (phase contrast microscopy, ×1000). Initial phylogenetic analysis by alignment of 16S rRNA gene sequences demonstrated that the strains were very homogenous and grouped in two very closely related clusters (99% identity) (Supplementary Figure S4). The two R. mobilis clusters included the type strain of R. mobilis NBRC101030 (and the synonymous R. pelagia NBRC102038 (Lai et al., 2010)), the *R. mobilis* strains isolated from Danish and Spanish turbot farms, and was distinct from other Ruegeria and Roseobacter clade species. Ruegeria (formerly Silicibacter) TM1040 was closely related to the main cluster (98%), but did cluster separately from our *R. mobilis* strains. Some strains (marked in bold in Supplementary Table S1) were isolated by two methods from the same sample, for example, G1302 and S1424, and such strains may be clones based on 16S rRNA gene similarity analysis.

MLSA allowed further differentiation of the *R. mobilis* strains into four sub-clusters (Supplementary Figure S5). While sub-cluster I and II only contained isolates from the Galathea expedition, Galathea strains grouped with the *R. mobilis* type strain NBRC101030 in sub-cluster III and sub-cluster IV consists of the Danish turbot farm strains and the strain NBRC102038. The Spanish turbot farm strain 270-3 and strain TrichCH4B build unique branches within the overall *R. mobilis* cluster.

Distinct separation of individual strains was achieved by whole-genome analysis, in our case, ANI (Figure 2). According to the cutoff for species differentiation at 95% (Richter and Rosselló-Móra, 2009), it was confirmed that all our isolates belong to the species *R. mobilis*. The strains group into the same four phylogenetic sub-clusters as in the MLSA, though 270-3 and TrichCH4B now group within the branch of sub-cluster IV. Four of the five pairs of strains isolated from the same samples using different methods (G1302/S1424, G1303/S1425, S1689/ F1926, F3871/S3940) fall within the same subcluster, however, do not group together within the sub-cluster, while one pair (S2684/F2587) is even separated into two sub-clusters. These pairs thus represent individual strains and are not clones of the same strain. There was no significant correlation between the ANI clustering and the parameters measured during sampling (depth, temperature, salinity, oxygen, nutrient concentration) or the distance to the shore (as categorized by coastal, open ocean and high seas) or the biomes associated with the sampling sites (Supplementary Table S4, P > 0.05); however, there was a significant difference for the distribution of the isolation techniques (S, F and G, P < 0.0001) and the ocean provinces, from which the strains were obtained (P < 0.0001), within the sub-clusters. That is, most of the strains from sub-cluster I were isolated by biofilm enrichment, while most of the strains from sub-cluster II and III were obtained by replica plating. Most strains of sub-cluster I, II and III were obtained from ocean province 46 (Chile-Peru Current Coastal Province), 21 (Benguela Current Coastal Province) and 42 (Archipelagic Deep Basins Province), respectively,



Figure 2 Heatmap of average nucleotide identity (ANI) analysis of *Ruegeria mobilis* strains, *Ruegeria* type strains and selected *Roseobacter* type strains (Richter and Rosselló-Móra, 2009; Pritchard, 2014).

(sub-cluster IV was excluded, as it contained only reference strains) (Supplementary Figure S6). As defined by Longhurst (Longhurst, 2007), for each ocean province, a class of certain parameters (productivity, chlorophyll and photic depth) corresponding to five equal intervals of average values was assigned. Analysing these classes for each ANI sub-cluster revealed no significant difference. This confirms that while the geographic distribution of each ANI sub-cluster is unique, this cannot be linked to certain environmental parameters.

The R. mobilis draft genomes ranged in size between 4.7 and 4.9 Mb (with the exception of

F3183 attributed to weak sequencing quality) and had a GC content of 58.7 to 59.3% (Supplementary Table S2). The sequence data (with the exception of F3183) was assembled in 45 to 94 contigs. Total number of genes varied from 4541 to 4828 (with the exception of F3183). The pangenome of 46 *R. mobilis* (excluding F3183) comprises 6,797 genes and the core genome consists of 3,599 genes (53% of the pangenome) (Supplementary Figure S7). The closed genome of F1926 consists of a chromosome (3.3 Mb, GC content: 58.8%) and four plasmids (1,256, 147, 107, 58 kb). In comparison, the closed genomes of *R. pomeroyi* DSS-3 and *Ruegeria* sp. TM1040 consist

of a 3.2 Mb chromosome plus three plasmids (822, 135, 131 kb) (Sule and Belas, 2013) and a 4.1 Mb chromosome plus one megaplasmid (492 kb) with a GC content of 60% and 64.2%, respectively (Moran et al., 2004, 2007). Two further plasmid sequences are available from Ruegeria sp. PR1b (149 kb and 76 kb). Comparison of the nucleotide sequence of DSS-3, TM1040 and the plasmids of PR1b (16S rRNA sequence identities of the strains to F1926 are 96%, 99% and 96%, respectively) to F1926 revealed that DSS-3 and TM1040 share similarities with the chromosome and the plasmid A of F1926 (Supplementary Figure S8). The hits on plasmid A of F1926 are encoded on the megaplasmid of TM1040, however, in case of DSS-3 on the chromosome (data not shown). Sequence data of TM1040 is further comparable to sequences on plasmid D of F1926 (Supplementary Figure S8). As previously reported (Sule and Belas, 2013), a third plasmid of TM1040 containing the tda genes remains unsequenced potentially leading to the missing sequence similarity to plasmid B in our representation. Sequence similarities of the two plasmids of PR1b, pSD25 and pSD20, can be found on plasmid C and D of F1926, respectively (data not shown). From the comparable sizes of total sequence data (Supplementary Table S2) and the conservation of replicases on the three larger plasmids of F1926 (repA on plasmid D cannot be found in S1424, NBRC101030 and NBRC102038) (Supplementary Figure S9), we conclude that the organization of the genome is conserved within the species (Supplementary Figure S9) and that the herein

presented *R. mobilis* strains harbor at least one plasmid, possibly a megaplasmid (Supplementary Table S2). This is supported by preliminary analysis using DNA gel electrophoresis (data not shown). The alignment gaps to F1926 on the chromosome and plasmid A and D can partly be attributed to prophages (Supplementary Table S7, Figure 3). In F1926 and S1424, we found unique pathways such as toxin-antitoxin systems that are important regulatory elements for a cell and we also detected phthalate degradation in NBRC101030, which sparks interest for industrial application and bioremediation (Supplementary Table S7). In F1926, the toxinantitoxin system is located on the 108 kb-plasmid and thus could be responsible for its perpetuation during cell division.

The four representative *R. mobilis* strains of the four ANI sub-clusters are very similar on the metabolic level (92%, Figure 3a) and the profile agrees with the description of *Roseobacter* Clade 1 strains (Newton *et al.*, 2010): they are heterotrophic, carry the ability to biosynthesize biotin, while not capable to produce polyphosphate (Supplementary Table S8). Their genomes demonstrate an adaptation to a particle- and eukaryote-associated lifestyle, including biotin biosynthesis, motility, chemotaxis, antibiotic biosynthesis, secretion systems (Supplementary Table S8). Although the core of the four \hat{R} mobilis genomes representing the four sub-clusters carry most of the KO orthologs present in the model strains TM1040 and DSS-3 (862 KEGG orthologs, 48%, Figure 3b), they also incorporate a high level of novelty (683 KEGG orthologs, 38%, Figure 3b). In closer comparison, they



Figure 3 Genome comparison using (a) the KEGG orthologs of the four *Ruegeria mobilis* representatives F1926, S1424, NBRC101030 and NBRC102038 (Kanehisa, Sato and Morishima, 2015) and (b) comparing their core KOs (* = 1636 KOs of Figure 3a) to *R. pomeroyi* DSS-3 and *Ruegeria* sp. TM1040 (Kanehisa *et al.*, 2015) using Venny (Oliveros) (details can be found in Supplementary Tables S7 and S8).

lack certain features such as sulfur oxidization (soxABCXYZ), reduction of nitrous-oxide (norBC, nosZ) or the acyl homoserine lactone synthase lasI (all unique to DSS-3) and have acquired others, that is, type IV and VI secretion systems or chemotaxis (Supplementary Table S8). Unique features of the four *R. mobilis* representatives include: transporters for autoinducer 2 (lsrA-D), oligopeptides (oppA-D+F) and D-methionine (metINQ), chemotaxis to air (aer) and oxidative phosphorylation using the cytochrome bd complex (cydAB) (Supplementary Table S8, Figure 3b).

Bioactivity of R. mobilis

Sterile-filtered culture supernatants from all 42 *R. mobilis* strains inhibited *V. anguillarum*. This was also true for culture supernatants from R. mobilis isolated from turbot farms and for *P. gallaeciensis* or P. inhibens strains (Supplementary Table S1). When incubated under shaken conditions (200 r.p.m.), none of the R. mobilis strains produced pigment and none of the sterile-filtered supernatants were inhibitory to V. anguillarum. TDA, which is believed to be the prime agent of Roseobacter antibacterial activity (Brinkhoff et al., 2004; Bruhn et al., 2006), was detected in all *R. mobilis* by liquid chromatography-tandem mass spectrometry and correlates with the presence of the central genes required for the biosynthesis of TDA, tdaBCDE, in all strains (Supplementary Table S8). In F1926, tdaBCDE are located on a plasmid, the 147 kb plasmid, as has been demonstrated for other TDA producers (Thole et al., 2012). No TDA, or only traces, was detected in MB cultures from non-inhibitory strains or under conditions that abolished activity, such as aeration of the R. mobilis strains. As QS can be involved in regulating TDA production (Bruhn *et al.*, 2005; Geng and Belas, 2010; Zan et al., 2014), we specifically searched for OH-C₁₀-HSL, the AHL inducing TDA biosynthesis in P. inhibens (Berger et al., 2011; Wilson et al., 2016), and its known open lactone form in sterile-filtered supernatants, however, no AHL compounds were produced by *R. mobilis* under the tested conditions, but were easily detected in the three *Phaeobacter* reference strains using the same method (Supplementary Table S1). Furthermore, when concentrated acidified ethyl acetate extracts of selected *R. mobilis* strains of the four sub-clusters were analyzed by high performance liquid chromatography-time of flight MS, OH-C₁₀-HSL or other AHLs were not detected (data not shown).

To complement our experimental findings, we searched the genomic data of the R. mobilis strains for the presence of genes related to QS (Supplementary Table S8, Supplementary Figure S10). Although present in the reference strain R. pomeroyi DSS-3, none were found in the R. mobilis genomes using an analysis based on the KEGG orthologs. However, during the sequence-based analysis using BLAST, two different types of potential AHL synthases were identified in the R. mobilis genomes presenting the conserved residues of well described AHL synthases such as LasI, LuxI or TofI (Supplementary Figure S10). The two AHL synthase types differ from the AHL synthases of other Ruegeria species, including the previously described synthases SsaI and SsbI (Zan et al., 2012). Type I (equal to HSL1 identified by antiSMASH, Supplementary Table S8) was found in all herein presented R. mobilis strains including TrichCH4B and the type strain NBRC101030. Type II, however, was only identified in a sub-group that corresponds with the phylogenetic sub-clusters II, III and IV (P < 0.0001), excluding only strains NBRC101030, NBRC102038 and 270-3.

Phage isolation and subtyping of R. mobilis

R. mobilis-infecting phages were obtained from several of the viral concentrates, however, all from leg 6 (Accra to Cape Town) of the cruise (Supplementary Table S3). Thus, there was no systematic spatial pattern with respect to their occurrence in the water samples, and the nonquantitative approach (meaning the pooling of all concentrates from each leg) does not allow for speculations about the spatial distribution or quantitative importance of the isolated phages. Testing the host range of the eight phage isolates against 47 *R. mobilis* strains showed three different host range patterns: Phages L1, L2, L3 had an identical host range, phages L4, L5, L6, L7 also had an identical host range and phage S4 had a unique host range. From this it was concluded that among the eight phage isolates, only three different phages (denoted L1, L4 and S4) had been obtained. Of the 47 bacterial strains, 6 were susceptible to one or two of the phages and 24 were susceptible to all 3 phages (Figure 4). The number of the phages, to which each strain was susceptible to, correlated with the ANI clustering (P < 0.0001). Most of the strains that were susceptible to all three phages were found in the sub-cluster I, whereas non-susceptible strains mostly clustered together in sub-clusters II, III, and IV (Figure 4). Only sub-cluster II contained strains that were susceptible to one or two of the three phages. None of the five reference strains from culture collections (NBRC101030 and NBRC102038) or from aquaculture farms (270-3, M41-2.2, M43-2.3) were susceptible to the phages. Bacterial strains co-isolated on leg 6 were not more susceptible to the phages than strains obtained from a different region (P=0.687). Instead, the 24 isolates that were susceptible to all three phages were distributed among most of the stations of isolation along the cruise covering highly different environments such as oxygen rich surface water (for example, F3688, 10 m depth), oxygen poor deeper water at the oxygen minimum zone off Peru (S3899, 400 m depth), shallow sediment (F2158, 60 m depth) and copepod surfaces (S1942) (Supplementary Table S4, Figure 4). Altogether, these 24 susceptible R. mobilis strains were



Figure 4 Sub-typing of *Ruegeria mobilis* strains (F, G and S strains were isolated as described in this study and Gram *et al.* (Gram *et al.*, 2010), M strains were isolated from Danish turbot-rearing farms (Porsby *et al.*, 2008)) using ANI-based phylogeny and phage typing by three phages (L1, L4 and S4); nd = not defined.

found in environments covering a temperature range from 4.0 to 29.3 $^{\circ}\mathrm{C}.$

Several intact, incomplete and questionable prophages were detected in almost all strains with an average of four prophages per strain and a maximum of eight sequences in F3183 (Supplementary Table S9). No prophages were identified in the strains G930, F2587, S1847, S1611 (sub-cluster III), and the reference strain ANG-S4. Most commonly found were sequences similar to the *Roseobacter* phage RDJL Φ 1 (although as questionable) (Zhao *et al.*, 2009; Huang *et al.*, 2011), followed by the intact *Rhodobacter* phage RcapNL, the incomplete *Rhodobacter* phage RC1 and the *Brochothrix* phage BL3 (Kilcher *et al.*, 2010) and Enterobacteria phage Φ P27 (Recktenwald and Schmidt, 2002) (both questionable hits). Once more, ANI clustering correlated with total number of prophages found per genome (P < 0.0001). Sub-cluster I strains carry at least sequences similar to the prophages RDJL Φ 1, RcapNL and BL3. Sub-cluster II strains have at least sequences similar to RDJL Φ 1 and Dp-1. No pattern is observed in subcluster III, where only three of the seven strains demonstrated incomplete or questionable hits. All, but NBRC102038 in sub-cluster IV carry at least sequences similar to the prophages RDJL Φ 1 and RcapNL.

All hits obtained in CRISPRFinder were assigned as questionable, a term, which refers to a small CRISPR with only two to three direct repeats or structures where the repeated motifs are not 100% identical (Grissa et al., 2007). All the detected CRISPRs had only two direct repeats and correspondingly one spacer (Supplementary Table S10), but two CRISPRs with the same direct repeat consensus and one identical and one distinct spacer each were detected in F3183 (CRISPR no. 6 and 7, Supplementary Table S10). Only one CRISPR was detected in a reference strains (TM1040), which, however, was not detected in any of the *R. mobilis* strains. All, but four *R. mobilis* strains (G930, F2459, NBRC102038, TrichCH4B) contained one or two CRISPR sequences. In strain F3183, eight CRISPRs were identified, which we, however, attribute to the poor sequence quality of the genomic data. Although there is a tendency that total number of CRISPR sequences found per genome correlated with the ANI clustering, there was no statistical evidence (P = 0.4218). Sub-cluster I strains carry either CRISPR 2 and 4 or 4 and 8 or only 2 or 8. Most of the subcluster II strains are positive for CRISPR 14, three strains for CRISPR 1. Sub-cluster III strains have CRISPR 3 and/or 11, however, no pattern is visible in sub-cluster IV.

Discussion

The *Roseobacter* clade is widespread in oceanic surface waters (Buchan and Moran, 2005; Wagner-Döbler and Biebl, 2006; Brinkhoff et al., 2008) and has been detected using molecular techniques worldwide (González et al., 2000; Wietz et al., 2010; Ghiglione *et al.*, 2012). Some sub-groups appear to be associated with particular geographic areas such as temperate and polar waters (Selje *et al.*, 2004; Swan et al., 2013). Here, we demonstrated that culturable Roseobacter-clade organisms can be detected worldwide and that at least one particular culturable species occurs globally in marine waters. There is an on-going discussion about distribution of marine bacteria, that is, whether particular sub-groups occur only in certain niches or whether some have a cosmopolitan occurrence. Some studies have found only very few cosmopolitan organisms (operational

taxonomic units) (Pommier *et al.*, 2006), whereas others have hypothesized that many organisms are ubiquitous on a global scale (Finlay, 2002; Fenchel and Finlay, 2004).

Strains within our R. mobilis collection were remarkably similar based on genotypic, phenotypic and chemical analyses, and although we did not isolate *R. mobilis* in Arctic and Antarctic waters, the strains appear to be representatives of an almost cosmopolitan species. The global appearance is further supported by the results of the metagenomic analysis. The overall low number of polar stations during the Galathea and the TARA research cruise could have caused the lack of isolates/data from these regions. One could hypothesise that even within a species, phenotypic variations occur depending on geographic area. However, in agreement with studies on Salinispora (Jensen et al., 2006), we found that the secondary metabolite production was independent of the geographical site of isolation. With the techniques applied, we excluded the isolation of potential non-bioactive *R. mobilis* strains, it is, however, intriguing that the consistent production of TDA indicates a strong evolutionary pressure in their natural habitat to maintain the megaplasmid carrying the highly conserved TDA biosynthetic genes (Sule and Belas, 2013; Harrington et al., 2014). Similar observations have been made in very different niches, such as silage where Penicillium paneum and P. roqueforti produce marcfortines and mycophenolic acid, respectively, and share andrastins (O'Brien et al., 2006).

The four phylogenetic sub-clusters of the 47 R. mobilis strains that were identified by MLSA and ANI and confirmed by the susceptibility to bacteriophages and the genomic pattern of prophage sequences demonstrate the differentiation within the species *R. mobilis*. Phages have previously been shown to be important drivers of bacterial diversification at the strain level by selecting for phageresistant strains, thus leading to a local arms race and co-evolution between phage and host populations (Middelboe et al., 2009). Hence, the versatility within the species is prompted by the exchange of genetic material as it has been proposed for the Roseobacter clade (Newton et al., 2010) and is not associated with any oceanographic parameters, which has been confirmed by the analysis of the TARA data (Supplementary Table S6). This is furthermore supported by previous observations that genetically related phages infecting marine Vibrio parahaemolyticus were distributed across large geographic distances (Kellogg et al., 1995). Concluding, the R. mobilis strains of each sub-cluster were distributed across large geographical scales and environmental conditions, suggesting that the pheno- and genotypic properties of each sub-cluster are maintained across variable habitats (planktonic, benthic and zooplankton-associated) and large spatial scales. Thus, the distinction in the geographic distribution of each cluster could only be related to the prophages within the genomes; comparably to a link that has been drawn recently between the viral community composition and the ocean provinces using TARA data (Brum *et al.*, 2015).

The element herein referred to as Roseobacter phage RDJL Φ 1 (Supplementary Table S9) (Huang et al., 2011) due to PHAST annotation (Zhou et al., 2011) has also been described as gene transfer agent (Zhao et al., 2009; Thole et al., 2012). Gene transfer agents have first been described in Rhodobacteraceae (Marrs, 1974) and are widespread within the family (Biers et al., 2008; Huang et al., 2011; Thole et al., 2012). This specific gene transfer agent has been found in all to-date sequenced Roseobacter genome sequences, but strain HTCC1062 (Zhao et al., 2009). HTCC1062 is an exception within the Roseobacter clade as it represents a putative oligotroph with a small genome size of only 2.3 MB (most Roseobacter strains have a genome of about 4 MB (Newton et al., 2010)) and its genome does not carry any gene transfer agent sub-clusters. Surprisingly, this element was also not found in the seven strains of our *R*. *mobilis* sub-cluster III. Overall, sub-cluster III strains appear poor in prophages in the PHAST analysis, which might imply that they harbor unknown prophages and this remains to be further investigated. As *R. mobilis* is phylogenetically not closely related to HTCC1062 and they possess no recent common predicted ancestors (Newton et al., 2010; Luo *et al.*, 2013), we suggest independent loss of the gene cluster and no direct evolutionary relation.

The *R. mobilis* strains described herein can potentially be classified as copiotroph representatives of the *Roseobacter* clade due to their genome size and genomic features (Luo and Moran, 2014; Voget *et al.*, 2014). As they were isolated from a wide variety of ecological niches, they might contribute substantially to the total *Roseobacter* gene repertoire (Moran *et al.*, 2007). Thereby, *R. mobilis* stands in contrast to streamlined bacteria such as 'Candidatus Pelagibacter ubique' that has been described as example of a beneficiary in the Black Queen Hypothesis; a theory referring to the selective advantage of gene loss (Jeffrey *et al.*, 2012). 'Ca. Pelagibacter ubique' has a reduced genomic content and depends on the leaky production of reduced sulfur by the community. This could further be correlated with the lifestyle of these two different bacterial species. Although *R. mobilis* is much more abundant in the particle-associated fraction than the free-living, the distribution between both lifestyles is more similar in 'Ca. Pelagibacter ubique' (Supplementary Figure S2). Concluding that *R. mobilis* is well adapted to the surface-attached lifestyle with constant nutrient supply from larger particles (also confirmed by the isolation technique using biofilm formation for enrichment), however, can shift to a pelagic lifestyle for dispersal. Respectively, as discussed previously, 'Ca. Pelagibacter ubique' is independent of nutrient quantities

The ISME Journal

available from non-bacterial, organic origin (Steindler *et al.*, 2011; Giovannoni *et al.*, 2014). Two of our *R. mobilis* strains were isolated from copepods, however, chitinases were not detected in the genomes. Instead, we identified the complete uptake system for N-acetylglucosamine dimers including the sensor chiS (Li and Roseman, 2004; Machado *et al.*, 2015). This could present an example of *R. mobilis* not only being a supplier of nutrients, but also benefiting from the leaky production by others.

It has been hypothesized that QS regulates the production of TDA (Geng and Belas, 2010; Berger et al., 2011). One of the classical groups of QS molecules, the AHLs, was not detected in any of the tested R. mobilis strains. However, we identified one or two potential AHL synthases in each genome of the Galathea strains. They might not be activated under the tested conditions or their products could not be detected in the chemical analysis. In contrast, P. inhibens, which also produces TDA, does produce AHL signals (Bruhn et al., 2005). Hence, a common cross-species strategy for using AHLs to regulate TDA production is not likely. It has been demonstrated that TDA induces its own production in *Ruegeria* (Silicibacter) sp. TM1040 (Geng and Belas, 2010), and we speculate that this would also be true for the *R. mobilis* strains of the present study.

Both species, R. mobilis and P. inhibens, produce TDA and were predicted to have the same direct ancestor (Luo et al., 2013). Although R. mobilis and *P. inhibens* strains have been obtained from the same sources and similar environments (Buchan and Moran, 2005; Porsby et al., 2008; Alsmark et al., 2013), we only isolated R. mobilis strains on the Galathea expedition using a technique that would also allow growth of *P. inhibens* (Gram *et al.*, 2010). In contrast, we repeatedly have isolated *P. inhibens* from surface biofilms in Danish harbor areas (Gram et al., 2015). It is, therefore, tempting to speculate that *R. mobilis* is the open-water counterpart to the coastal, macrosurface-attached P. inhibens (Supplementary Table S1 and Supplementary material \$8 in (Thole et al., 2012)). In the TARA data set (Global Ocean Microbial Reference Gene Catalog), Ruegeria sp. represents 5% of all detected Rhodobacteraceae (which comprise 3% of all bacteria), while *Phaeobacter* sp. corresponds to only 0.3% (Sunagawa et al., 2015). Though this would support the hypothesis of niche-development of these two species, the amount of data is too low for a statistical proof and requires further analysis.

In conclusion, we have found that a culturable bioactive species of the *Roseobacter* clade occurs ubiquitously as pelagic marine bacteria, but potentially being rare in polar regions. The intra-species variation originates primarily from prophages, which is linked to the spatial distribution patterns. In comparison to isolates obtained from aquaculture and other *Ruegiera* model strains, the new marine isolates represent a novel diversity with potential for application as probionts. Based on the tested genotypic, phenotypic or chemical profiles, its preserved traits indicate a high degree of adaptation to a generalist ubiquitous lifestyle.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

The Carlsberg Foundation and the Danish Center for Food Research (LMC) are thanked for donations covering international travel. Funding from the Programme Committee for Food, Health and Welfare under the Danish Strategic Research Council (grant no. DSF 09-063047), from the Danish Council for Independent Research (grant no. FNU-09-072829) and the European Union (Seventh Framework Programmes MaCuMBA (KBBE-2012-6-311975) and PharmaSea (KBBE-2012-6-312184)) is acknowledged. We express our gratitude to Dr Jørn Smedsgaard, Dr Jesper B. Bruhn, technician Ellen Kirstine Lyhne, Dr Lasse Riemann and Karin Holmfeldt for participation in and sampling done during the cruise. We thank Katherine Richardson and Karen Marie Hilligsøe for providing physicochemical data, Nete Bernbom for initial work on sub-typing of the strains, Simone Severitt and Nicole Mrotzek for excellent technical assistance, Nikolaus Sonnenschein for help with computational analyses and three anonymous reviewers for valuable comments. The present work was carried out as part of the Galathea 3 expedition under the auspices of the Danish Expedition Foundation. This is Galathea 3 contribution no. P117.

References

- Alavi M, Miller T, Erlandson K, Schneider R, Belas R. (2001). Bacterial community associated with *Pfiesteria*like dinoflagellate cultures. *Environ Microbiol* 3: 380–396.
- Alsmark C, Strese Å, Wedén C, Backlund A. (2013). Microbial diversity of *Alcyonium digitatum*. *Phytochem Rev* 12: 531–542.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–3402.
- Berger M, Neumann A, Schulz S, Simon M, Brinkhoff T. (2011). Tropodithietic acid production in *Phaeobacter* gallaeciensis is regulated by N-acyl homoserine lactonemediated quorum sensing. J Bacteriol **193**: 6576–6585.
- Biers EJ, Wang K, Pennington C, Belas R, Chen F, Moran MA. (2008). Occurrence and expression of gene transfer agent genes in marine bacterioplankton. *Appl Environ Microbiol* **74**: 2933–2939.
- Brinkhoff T, Bach G, Heidorn T, Liang L, Schlingloff A, 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.

- Brinkhoff T, Giebel H-A, Simon M. (2008). Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch Microbiol* **189**: 531–539.
- Bruhn J, Nielsen K, Hjelm M, Hansen M, Bresciani J, Schulz S et al. (2005). Ecology, inhibitory activity, and morphogenesis of a marine antagonistic bacterium belonging to the *Roseobacter* clade. Appl Environ Microbiol 71: 7263–7270.
- Bruhn JB, Gram L, Belas R. (2007). Production of antibacterial compounds and biofilm formation by *Roseobacter* species are influenced by culture conditions. *Appl Environ Microbiol* **73**: 442–450.
- Bruhn JB, Haagensen JAJ, Bagge-Ravn D, Gram L. (2006). Culture conditions of *Roseobacter* strain 27-4 affect its attachment and biofilm formation as quantified by real-time PCR. *Appl Environ Microbiol* 72: 3011–3015.
- Brum JR, Ignacio-espinoza JC, Roux S, Doulcier G, Acinas SG, Alberti A *et al.* (2015). Patterns and ecological drivers of ocean viral communities. *Science* **348**: 1261498.
- Buchan A, Moran MA. (2005). Overview of the marine *Roseobacter* lineage. *Appl Environ Microbiol* 71: 5665–5677.
- D'Alvise PW, Lillebø S, Prol-Garcia MJ, Wergeland HI, Nielsen KF, Bergh Ø et al. (2012). Phaeobacter gallaeciensis reduces Vibrio anguillarum in cultures of microalgae and rotifers, and prevents vibriosis in cod larvae. PLoS One 7: e43996.
- D'Alvise PW, Lillebø S, Wergeland HI, Gram L, Bergh Ø. (2013). Protection of cod larvae from vibriosis by *Phaeobacter* spp.: a comparison of strains and introduction times. *Aquaculture* **384–387**: 82–86.
- D'Alvise PW, Melchiorsen J, Porsby CH, Nielsen KF, Gram L. (2010). Inactivation of Vibrio anguillarum by attached and planktonic Roseobacter cells. Appl Environ Microbiol **76**: 2366–2370.
- D'Alvise PW, Phippen CBW, Nielsen KF, Gram L. (2015). Influence of iron on the production of the antibacterial compound tropodithietic acid and its non-inhibitory analog in *Phaeobacter inhibens*. Appl Environ Microbiol 82: 502–509.
- Fenchel T, Finlay BJ. (2004). The ubiquity of small species: patterns of local and global diversity. *Bioscience* **54**: 777.
- Finlay BJ. (2002). Global dispersal of free-living microbial eukaryote species. *Science* **296**: 1061–1063.
- Geng H, Belas R. (2010). Expression of tropodithietic acid biosynthesis is controlled by a novel autoinducer. J Bacteriol 192: 4377–4387.
- Geng H, Bruhn JB, Nielsen KF, Gram L, Belas R. (2008). Genetic dissection of tropodithietic acid biosynthesis by marine roseobacters. *Appl Environ Microbiol* 74: 1535–1545.
- Ghiglione J-F, Galand PE, Pommier T, Pedros-Alio C, Maas EW, Bakker K *et al.* (2012). Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc Natl Acad Sci* **109**: 17633–17638.
- 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.
- Giovannoni SJ, Cameron Thrash J, Temperton B. (2014). Implications of streamlining theory for microbial ecology. *ISME J* 8: 1–13.
- González JM, Simó R, Massana R, Covert JS, Casamayor EO, Pedrós-Alió C *et al.* (2000). Bacterial community

structure associated with a dimethylsulfoniopropionateproducing North Atlantic algal bloom. *Appl Environ Microbiol* **66**: 4237–4246.

- Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* **57**: 81–91.
- Gram L, Melchiorsen J, Bruhn J. (2010). Antibacterial activity of marine culturable bacteria collected from a global sampling of ocean surface waters and surface swabs of marine organisms. *Mar Biotechnol (NY)* **12**: 439–451.
- Gram L, Rasmussen BB, Wemheuer B, Bernbom N, Ng YY, Porsby CH *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.
- Grissa I, Vergnaud G, Pourcel C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **35**: W52–W57.
- Harrington C, Reen FJ, Mooij MJ, Stewart FA, Chabot J-B, Guerra AF *et al.* (2014). Characterisation of nonautoinducing tropodithietic acid (TDA) production from marine sponge *Pseudovibrio* species. *Mar Drugs* **12**: 5960–5978.
- Hjelm M, Bergh O, Riaza A, Nielsen J, Melchiorsen J, Jensen S *et al.* (2004). Selection and identification of autochthonous potential probiotic bacteria from turbot larvae (*Scophthalmus maximus*) rearing units. *Syst Appl Microbiol* **27**: 360–371.
- Huang S, Zhang Y, Chen F, Jiao N. (2011). Complete genome sequence of a marine roseophage provides evidence into the evolution of gene transfer agents in alphaproteobacteria. *Virol J* 8: 124.
- Jeffrey MJ, Lenski RE, Zinser ER. (2012). The Black Queen Hypothesis : evolution of dependencies through adaptive gene loss. *mBio* **3**: 1–7.
- Jensen PR, Williams PG, Oh D-C, Zeigler L, Fenical W. (2006). Species-specific secondary metabolite production in marine actinomycetes of the genus *Salinispora*. *Appl Environ Microbiol* **73**: 1146–1152.
- Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. (2015). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44: D457–D462.
- Kanehisa M, Sato Y, Morishima K. (2015). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* **428**: 726–731.
- Kellogg C, Rose JB, Jiang SC, Thurmond JM, Paul JH. (1995). Genetic diversity of related vibriophages isolated from marine environments around Florida and Hawaii, USA. *Mar Ecol Prog Ser* **120**: 89–98.
- Kilcher S, Loessner MJ, Klumpp J. (2010). *Brochothrix thermosphacta* bacteriophages feature heterogeneous and highly mosaic genomes and utilize unique prophage insertion sites. *J Bacteriol* **192**: 5441–5453.
- Lai Q, Yuan J, Li F, Zheng T, Shao Z. (2010). Ruegeria pelagia is a later heterotypic synonym of Ruegeria mobilis. Int J Syst Evol Microbiol **60**: 1918–1920.
- Li X, Roseman S. (2004). The chitinolytic cascade in Vibrios is regulated by chitin oligosaccharides and a two-component chitin catabolic sensor/kinase. *Proc Natl Acad Sci USA* **101**: 627–631.
- Longhurst AR. (2007). *Ecological Geography of the Sea*. Academic Press.

- Luo H, Csuros M, Hughes AL, Moran MA. (2013). Evolution of divergent life history strategies in marine Alphaproteobacteria. *mBio* **4**: 1–8.
- Luo H, Moran MA. (2014). Evolutionary ecology of the marine *Roseobacter* clade. *Microbiol Mol Biol Rev* **78**: 573–587.
- Machado H, Sonnenschein EC, Melchiorsen J, Gram L. (2015). Genome mining reveals unlocked bioactive potential of marine Gram-negative bacteria. *BMC Genomics* **16**: 1365.
- Marrs B. (1974). Genetic recombination in *Rhodopseudo*monas capsulata. Proc Natl Acad Sci USA **71**: 971–973.
- Martens T, Gram L, Grossart H-P, Kessler D, Müller R, Simon M *et al.* (2007). Bacteria of the *Roseobacter* clade show potential for secondary metabolite production. *Microb Ecol* **54**: 31–42.
- Middelboe M, Holmfeldt K, Riemann L, Nybroe O, Haaber J. (2009). Bacteriophages drive strain diversification in a marine *Flavobacterium*: implications for phage resistance and physiological properties. *Environ Microbiol* **11**: 1971–1982.
- Miller TR, Belas R. (2004). Dimethylsulfoniopropionate metabolism by *Pfiesteria*-associated *Roseobacter* spp. *Appl Environ Microbiol* **70**: 3383–3391.
- Moran MA, Belas R, Schell MA, González JM, Sun F, Sun S et al. (2007). Ecological genomics of marine Roseobacters. Appl Environ Microbiol **73**: 4559–4569.
- Moran MA, Buchan A, González JM, Heidelberg JF, Whitman WB, Kiene RP *et al.* (2004). Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* **432**: 910–913
- Moran MA, González JM, Kiene RP. (2003). Linking a bacterial taxon to sulfur cycling in the sea: studies of the marine *Roseobacter* group. *Geomicrobiol J* **20**: 375–388.
- Morris RM, Rappe MS, Vergin KL, Siebold WA, Carlson CA, Giovannoni SJ. (2002). SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**: 806–810.
- Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, Givens CE *et al.* (2010). Genome characteristics of a generalist marine bacterial lineage. *ISME J* **4**: 784–798.
- Oliveros JC Venny. An interactive tool for comparing lists with Venn's diagrams. Available at: http://bioinfogp. cnb.csic.es/tools/venny/index.html.
- O'Brien M, Nielsen KF, O'Kiely P, Forristal PD, Fuller HT, Frisvad JC. (2006). Mycotoxins and other secondary metabolites produced in vitro by *Penicillium paneum* Frisvad and *Penicillium roqueforti* Thom isolated from baled grass silage in Ireland. J Agric Food Chem 54: 9268–9276.
- Penesyan A, Tebben J, Lee M, Thomas T, Kjelleberg S, Harder T *et al.* (2011). Identification of the antibacterial compound produced by the marine epiphytic bacterium *Pseudovibrio* sp. D323 and related spongeassociated bacteria. *Mar Drugs* **9**: 1391–1402.
- Planas M, Pérez-Lorenzo M, Hjelm M, Gram L, Uglenes Fiksdal I, Bergh Ø et al. (2006). Probiotic effect in vivo of Roseobacter strain 27-4 against Vibrio (Listonella) anguillarum infections in turbot (Scophthalmus maximus L.) larvae. Aquaculture 255: 323–333.
- Pommier T, Canback B, Riemann L, Bostrom KH, Simu K, Lundberg P *et al.* (2006). Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol* **16**: 867–880.

- Porsby CH, Nielsen KF, 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.
- Porsby CH, Webber MA, Nielsen KF, Piddock LJ, Gram L. (2011). Resistance and tolerance to tropodithietic acid, an antimicrobial in aquaculture, is hard to select. *Antimicrob Agents Chemother* **55**: 1332–1337.
- Prabagaran SR, Manorama R, Delille D, Shivaji S. (2007). Predominance of Roseobacter, Sulfitobacter, Glaciecola and Psychrobacter in seawater collected off Ushuaia, Argentina, Sub-Antarctica. FEMS Microbiol Ecol 59: 342–355.
- Pritchard L. (2014). calculate_ani.py. GitHub Repos.
- Rao D, Webb JS, Kjelleberg S. (2005). Competitive interactions in mixed-species biofilms containing the marine bacterium *Pseudoalteromonas tunicata*. *Appl Environ Microbiol* **71**: 1729–1736.
- Rasmussen BB, Nielsen KF, Machado H, Melchiorsen J, Gram L, Sonnenschein EC. (2014). Global and phylogenetic distribution of quorum sensing signals, acyl homoserine lactones, in the family of Vibrionaceae. *Mar Drugs* **12**: 5527–5546.
- Rau MH, Hansen SK, Johansen HK, Thomsen LE, Workman CT, Nielsen KF *et al.* (2010). Early adaptive developments of Pseudomonas aeruginosa after the transition from life in the environment to persistent colonization in the airways of human cystic fibrosis hosts. *Environ Microbiol* **12**: 1643–1658.
- Recktenwald J, Schmidt H. (2002). The nucleotide sequence of Shiga toxin (Stx) 2e-encoding phage ϕ P27 is not related to other Stx phage genomes, but the modular genetic structure is conserved. Infect Immun **70**: 1896–1908.
- Richter M, Rosselló-Móra R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* **106**: 19126–19131.
- Sawabe T, Ogura Y, Matsumura Y, Feng G, Amin AR, Mino S *et al.* (2013). Updating the *Vibrio* clades defined by multilocus sequence phylogeny: proposal of eight new clades, and the description of *Vibrio tritonius* sp. nov. *Front Microbiol* **4**: 1–414.
- Selje N, Simon M, Brinkhoff T. (2004). A newly discovered *Roseobacter* cluster in temperate and polar oceans. *Nature* **427**: 445–448.
- Skov MN, Pedersen K, Larsen JL. (1995). Comparison of pulsed-field gel electrophoresis, ribotyping, and plasmid profiling for typing of *Vibrio anguillarum* serovar O1. *Appl Environ Microbiol* **61**: 1540–1545.
- Steindler L, Schwalbach MS, Smith DP, Chan F, Giovannoni SJ. (2011). Energy starved *Candidatus* Pelagibacter ubique substitutes light-mediated ATP production for endogenous carbon respiration. *PLoS One* **6**: 1–10.
- Stenholm AR, Dalsgaard I, Middelboe M. (2008). Isolation and characterization of bacteriophages infecting the fish pathogen *Flavobacterium psychrophilum*. *Appl Environ Microbiol* **74**: 4070–4078.
- Sule P, Belas R. (2013). A novel inducer of *Roseobacter* motility is also a disruptor of algal symbiosis. *J Bacteriol* **195**: 637–646.
- Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G *et al.* (2015). Structure and function of the global ocean microbiome. *Science* **348**: 1–10.

- Swan BK, Tupper B, Sczyrba A, Lauro FM, Martinez-Garcia M, González JM et al. (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. Proc Natl Acad Sci USA 110: 11463–11468.
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. (2013). MEGA6: Molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* **30**: 2725–2729.
- 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.
- Todd JD, Curson ARJ, Dupont CL, Nicholson P, Johnston AWB. (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.
- Voget S, Wemheuer B, Brinkhoff T, Vollmers J, Dietrich S, Giebel H-A *et al.* (2014). Adaptation of an abundant Roseobacter RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. *ISME J* **9**: 371–384.
- Wagner-Döbler I, Biebl H. (2006). Environmental biology of the marine Roseobacter lineage. Annu Rev Microbiol 60: 255–280.

- Wietz M, Gram L, Jørgensen B, Schramm A. (2010). Latitudinal patterns in the abundance of major marine bacterioplankton groups. Aquat Microb Ecol 61: 179–189.
- Wilson MZ, Wang R, Gitai Z, Seyedsayamdost MR. (2016). Mode of action and resistance studies unveil new roles for tropodithietic acid as an anticancer agent and the γ -glutamyl cycle as a proton sink. *Proc Natl Acad Sci* **113**: 1630–1635.
- Zan J, Cicirelli EM, Mohamed NM, 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.
- Zan J, Liu Y, Fuqua C, Hill RT. (2014). Acyl-homoserine lactone quorum sensing in the *Roseobacter* clade. *Int J Mol Sci* **15**: 654–669.
- Zhao Y, Wang K, Budinoff C, Buchan A, Lang A, Jiao N et al. (2009). Gene transfer agent (GTA) genes reveal diverse and dynamic *Roseobacter* and *Rhodobacter* populations in the Chesapeake Bay. *ISME J* **3**: 364–373.
- Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. (2011). PHAST: a fast phage search tool. *Nucleic Acids Res* **39**: W347–W352.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)