Available online at www.sciencedirect.com
ScienceDirect
Systematic and Applied Microbiology 32 (2009) 124–139
SYSTEMATIC AND APPLIED MICROBIOLOGY
www.elsevier.de/syapm

Biogeography and phylogeny of the NOR5/OM60 clade of *Gammaproteobacteria*

Shi Yan[a], Bernhard M. Fuchs[a,*], Sabine Lenk[a], Jens Harder[b], Jörg Wulf[a], Nian-Zhi Jiao[c], Rudolf Amann[a]

[a]*Department of Molecular Ecology, Max Planck Institute for Marine Microbiology, Bremen D-28359, Germany*
[b]*Department of Microbiology, Max Planck Institute for Marine Microbiology, Bremen D-28359, Germany*
[c]*National Key Laboratory for Marine Environmental Sciences, University of Xiamen, Xiamen, Fujian 361005, China*

Abstract

The phylogeny, abundance, and biogeography of the NOR5/OM60 clade was investigated. This clade includes "*Congregibacter litoralis*" strain KT71, the first cultured representative of marine aerobic anoxygenic phototrophic *Gammaproteobacteria*. More than 500 16S rRNA sequences affiliated with this clade were retrieved from public databases. By comparative sequence analysis, 13 subclades could be identified, some of which are currently restricted to discrete habitat types. Almost all sequences in the largest subclade NOR5-1 and related subclade NOR5-4 originated from marine surface water samples. Overall, most of the NOR5/OM60 sequences were retrieved from marine coastal settings, whereas there were fewer from open-ocean surface waters, deep-sea sediment, freshwater, saline lakes and soil.

The abundance of members of the NOR5/OM60 clade in various marine sites was determined by fluorescence *in situ* hybridization using a newly designed and optimized probe set. Relative abundances in coastal marine waters off Namibia and the Yangtze estuary were up to 3% of the total 4′,6-diamidino-2-phenylindole (DAPI) counts, and in the German Bight off Helgoland the abundance was even up to 7%. In an open-ocean North Atlantic transect, between Iceland and the Azores, the NOR5/OM60 group was much less abundant (0.1–0.5%). Interestingly, the surface layer of North Sea intertidal sediments was very rich in NOR5/OM60, with absolute numbers $>10^8$ cells cm^{-3} (or 4% of the total DAPI). An analysis of the frequencies of NOR5/OM60 16S rRNA genes in the Global Ocean Survey datasets provided further support for a marine cosmopolitan occurrence of NOR5/OM60, and a clear preference for coastal marine waters.

Keywords: Microbial ecology; Bacteria; Fluorescence *in situ* hybridization; Phylogeny; Biogeography

Introduction

Aerobic anoxygenic phototrophic bacteria (AAnP) use light as an energy source, and appear to have an important role in marine carbon cycling [33,34]. They also seem to be highly abundant in the oceans. Recent studies based on infrared microscopy showed abundances of $4.5 \pm 2.4\%$ with a maximum of 13.5% in coastal waters, while in oceanic water the frequency was lower at $1.5 \pm 1.3\%$ [30,56,62]. For a long time, all cultured representatives of marine AAnP belonged to

*Corresponding author.
E-mail address: bfuchs@mpi-bremen.de (B.M. Fuchs).

0723-2020/$ - see front matter © 2008 Elsevier GmbH. All rights reserved.
doi:10.1016/j.syapm.2008.12.001

the class *Alphaproteobacteria*. This recently changed when genome analyses of marine gammaproteobacterial isolates revealed the presence of complete photosynthesis superoperons [22]. The North Sea isolate KT71 was identified as AAnP based on the expression of photosynthetic pigments [22]. In parallel, HTCC2080, one of three gammaproteobacterial strains obtained by high-throughput cultivation from coastal Pacific surface water off Oregon, was shown to grow faster and to have higher cell yield with light rather than in the dark [12]. All four isolates are from one monophyletic gammaproteobacterial clade that had been predicted based on metagenomics [8,58].

The history of this clade dates back to 1995, when a 284 bp gammaproteobacterial 16S rRNA sequence retrieved from Sargasso Sea surface water was referred to as SAR125 (L35466) [47]. In 1997, two almost full-length sequences were found to be closely related to SAR125; clones OM60 (U70696) and OM241 (U70702) were from a marine coastal site off North Carolina, USA [51]. In 1999, strain KT71 (AY007676) was isolated from marine surface water at the "Kabeltonne" station off the island of Helgoland, North Sea [19]. Strain KT71 was placed in a group named NOR5, and recently the binominal name "*Congregibacter litoralis*" has been suggested for this isolate [22]. Several related strains isolated by a novel high-throughput culturing method, including HTCC2080, were placed in the OM60/OM241 clade [14], which was later referred to as the OM60 clade [11]. Since all these clade names are now largely redundant, we refer to it here as the NOR5/OM60 clade [22] and, currently, many more related sequences can be found in various databases.

Based on comparative 16S rRNA sequence analysis, the NOR5/OM60 clade is most closely related to the genera *Endobugula*, *Microbulbifer*, *Teredinibacter* (all *Alteromonadales*), *Cellvibrio* (*Pseudomonadales*) and several other groups of oligotrophic marine *Gammaproteobacteria*, including the clades BD1-7, KI89A, OM182 and SAR92 [11]. A sequence retrieved from deep-sea sediment, BD2-7 (AB015537) [37], was considered to represent a sister clade to the NOR5/OM60 clade.

Data on the biogeography of the NOR5/OM60 clade are so far sparse. Besides the mostly qualitative evidence from sequence retrieval there is also some quantification by fluorescence *in situ* hybridization (FISH). The probe NOR5-730 has yielded counts in North Sea surface water of up to 8% [19], and even 11% [48] of all DAPI-stained cells. With the same probe, an abundance of $3.4 \pm 1.1\%$ was detected in a Pacific coastal transect along the Newport Hydroline [12], where maximum numbers of up to 1×10^5 cells ml^{-1} were linked to the chlorophyll maximum.

The goals of this study were to provide a more detailed description of the phylogeny of the NOR5/OM60 clade and to analyze its geographic distribution, as well as abundance, in the marine environment. Therefore, all available NOR5/OM60 16S rRNA sequences from public databases were mapped. This information originated mostly from PCR-based clone libraries, but also encompassed isolates and the metagenomic Global Ocean Survey (GOS) [52]. Based on this comprehensive dataset the specificity of published 16S rRNA-targeted oligonucleotide probes was checked, and new ones were designed. A probe mixture was optimized for a catalyzed reporter deposition FISH (CARD-FISH) assay. With this, the abundance of members of the NOR5/OM60 clade was quantified in coastal marine sites off China, Namibia, Spain, and Germany, as well as in an open-ocean transect in the North Atlantic.

Materials and methods

Sequence retrieval

NOR5/OM60-related 16S rRNA sequences were initially retrieved with the ARB program (http://www.arb-home.de/) [43] from the SILVA database (Version 91) [50] (http://www.arb-silva.de/) by targeting group-specific signatures (e.g. complement to probe NOR5-730 with 0–2 mismatches). Highly related sequences were also identified using Blastn (http://www.ncbi.nlm.nih.gov/BLAST/), and imported into the database. Additionally, the 16S rRNA subset of the Camera GOS Cruise, a metagenomic project that sampled various environments, including pelagic and coastal seawater, fresh water or hypersalinic environments [52], was downloaded (http://web.camera.calit2.net/) and analyzed in a similar way.

In total, more than 500 16S rRNA sequences including 147 almost full-length sequences with a length of >1400 nucleotides were collected. Sequences were manually checked for sequence quality and chimera, using the pintail value [5] provided with the SILVA 16S rRNA gene database and using the program "chimera check" at the website of the ribosomal database project (http://rdp8.cme.msu.edu/cgis/chimera.cgi?su=SSU) [13]. Twenty-eight probably chimeric NOR5/OM60 sequences were used only for biogeographic studies, but not for phylogenetic analysis or probe design. The sequences were aligned using the ARB aligner, and added to the universal parsimony tree using "ARB parsimony" with a "positional variability by parsimony" filter for *Bacteria* [50].

The 16S rRNA sequences from 22 newly isolated strains from sediment off the North Sea island Sylt (55.01°N, 8.26°E) were determined by standard molecular techniques (PCR, sequencing) and submitted

to GenBank under the accession numbers EU672847–EU672869. A detailed description of the strains will be published elsewhere (Harder, unpublished).

Phylogeny

All the qualified, almost full-length sequences (>1400 nt) that belonged to the NOR5/OM60 clade, as well as several closely related outgroup sequences (in total around 150 sequences), were selected for phylogenetic reconstructions. Three column filters were made in ARB and were used for selecting the columns with certain conservative levels for calculation. Each filter kept 1493, 1450 and 1393 bases, respectively. The sequences were filtered into different datasets, and each was then used as input for four different algorithms: "ARB neighbor joining" (Felsenstein correction), "ARB parsimony interactive", Maximum likelihood using AxML, and the posterior possibility algorithm using MrBayes (Version 3.1, http://mrbayes.csit.fsu.edu/) [27]. The MrBayes trees were built according to the manual (http://mrbayes.csit.fsu.edu/manual.php) with settings of a likelihood model in two parallel runs, each containing four chains. The program ran until the "average standard deviation of split frequencies" became less than 0.1. Then, the first thousand trees of the unstable generations were "burnt in" using "half-compat", and MrBayes consensus trees were constructed. Subsequently, the trees obtained using the three filtered sets and four methods were compared manually. Groups that were stable in all or most of the trees were named as subclades. Whenever the branching patterns varied in many of the trees, a multifurcation was introduced at that position [42].

Probe design, check and optimization

The "probe check" module in the ARB program was used to check probes NOR5-730 and NOR5-130 [19] for coverage and specificity, but also to design new probes for NOR5/OM60 subclades. Only sequences with >1400 nt were used for probe design. Candidate probes were also checked against the 16S rRNA sequence databases, including the partial sequences. Helper and competitor nucleotides were designed and tested as described before [21,46] (Table 1). The probes were optimized by performing hybridization with all the relevant helpers and competitors at varying formamide concentrations [23]. Optimizations were carried out at 46 °C on pure cultures with a fully complementary 16S rRNA.

Biogeographic analysis

For each of the NOR5/OM60 sequences retrieved from the databases, the geographic and environmental data were manually collected, either directly from

Table 1. Probes, helpers and competitors that were used in this study.

Name	Targeted group	Sequence (5′–3′)	Target site (16S rRNA *Escherichia coli* numbering)	Formamide	Reference
NOR5-730	NOR5/OM60 clade	TCG AGC CAG GAG GCC GCC	730–747	50%	[19]
NOR5-709h	n.a.	TTC GCC ACY GGT ATT CCT CCA	709–729	n.a.	This study
NOR5-659h	n.a.	GAA TTC TAC CTC CCT CTC YCG	659–679	n.a.	This study
NOR5-1238	NOR5/OM60 clade, excluding NOR5-1 and -4	CCC TCT GTG CGT TCC ATT	1238–1255	50–55%	This study
NOR5-1217h	n.a.	GTA GCA CGT GTG TAG CCC AGG	1217–1237	n.a.	This study
NOR5-1287h	n.a.	ATC CGG ACT ACG AAA CGT TTT	1287–1307	n.a.	This study
EUB I–III	*Bacteria*	GCT GCC TCC CGT AGG AGT, GCA GCC ACC CGT AGG TGT, GCT GCC ACC CGT AGG TGT	338–355	35%	[16]
NON	Negative control	ACT CCT ACG GGA GGC AGC	Reverse complement of EUB I	35%	[60]

Suffix "h" stands for helper oligonucleotide.
n.a.: not applicable.

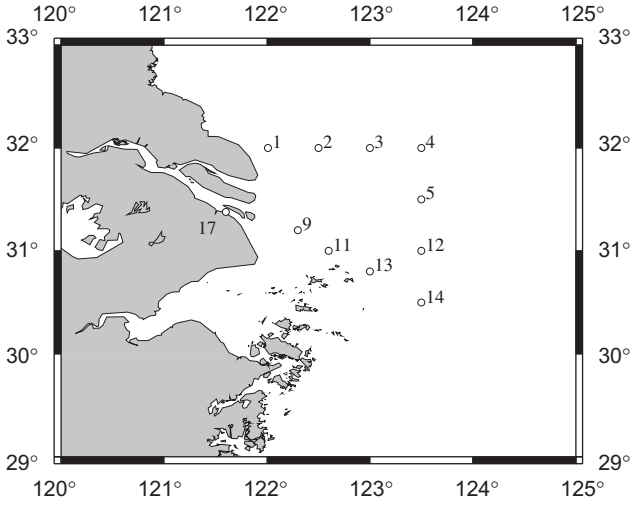

Fig. 1. Sampling stations off the Yangtze River estuary.

GenBank entries or from corresponding publications. Data included longitude, latitude, depth or altitude, and habitat information in categories such as marine water, hypersaline water or fresh water. The resulting table (SI Table 3) was the basis for creating the biogeographic map shown in Fig. 3. Results of the CARD-FISH with probe NOR5-730 and NOR5-1238 were also included in the map.

Sampling sites and procedures

Yangtze River estuary

On September 6–8, 2006, a small cruise was undertaken in the Yangtze River estuary (Fig. 1). Samples were taken from surface water, and immediately fixed with 1% paraformaldehyde (PFA) for 1 h, filtered onto polycarbonate filters (Millipore, 47 mm in diameter, 0.2 μm pore size) and stored frozen at −20 °C.

Namibian upwelling region

The cruise took place on March 22–23, 2003, along 23°S near Walvis Bay, from the coast into the Atlantic Ocean (14.4–12.0°E), through the Benguela Current. Surface water samples (10 m) from 13 stations, as well as three depth profiles, were collected and immediately fixed in 1% PFA for 1 h at room temperature or for 24 h at 4 °C. Subsequently, samples were kept frozen at −80 °C. For FISH, samples were carefully thawed, filtered onto polycarbonate filters (Millipore, 47 mm in diameter, 0.2 μm pore size) and further processed for FISH (see below).

Vision cruise

The Vision cruise was conducted in the period September 20–October 3, 2006. Sampling was carried out along the transect 30°W, from Iceland to the south of the Azores Islands, from surface waters (mostly at a depth of 10 m, SI Table 1). All water samples were fixed, filtered and stored as described for the Yangtze estuary.

German Bight

Samples were taken from a depth of 1 m at station "Kabeltonne", Helgoland (54.18°N, 7.90°E), German Bight, on seven separate days from May to July 2007 (see Table 2 for details) and from Cuxhaven in July 2007. The water samples from Helgoland were first pre-filtered at 10 μm to remove large particles and then fixed, filtered and stored as described for the Yangtze estuary. The water samples from Cuxhaven were treated identically but they were not pre-filtered.

North Sea sediment

The sediment samples were sampled from a sandy intertidal flat at Janssand (53.72°N, 7.68°E), German North Sea coast, in March and August 2007. Each time two adjacent cores were sampled for duplicates, and subsampled for each 1 cm range. The subsamples were fixed and sonicated, then the supernatant was filtered onto polycarbonate filters, as described previously [41,49].

Small sampling campaigns

A summary of all additional sampling stations is given in Table 2. Surface water samples from Xiamen, China were taken in September 2006 at the Xiamen ferry port, and in July 2007 from a sandy coast near Xiamen University. Other marine water samples were obtained from Southampton dock water, UK, and coastal water near Barcelona, Spain. Fresh water samples from the river Weser and freshwater ponds in Bremen were also checked for comparison. All the water samples were treated as described above for the Yangtze River sample. Other sediment samples were taken from intertidal sandy surface sediment from Sylt on the German North Sea coast. The samples were fixed, sonicated and processed as described above for North Sea sediment samples.

CARD-FISH

CARD-FISH was undertaken according to Pernthaler et al. [49] with the following modifications: agarose-embedded filters were permeabilized with 10 mg ml^{-1}

Table 2. Small sampling campaigns for determining the NOR5/OM60 distribution.

Location	Coordinates	NOR5/OM60 count	Method	Note[a]	Reference
Germany, North Sea, Helgoland, surface water	54.18°N 7.90°E	Up to 6–8% in early June and late July, 1998	FISH	NOR5-730, 30% FA	[19]
		6–7% in May, 11–13% in August, 2002	CARD-FISH	NOR5-730, 55% FA, 35 °C	[48]
		3–5%, unaffected in incubation after 0.8 μm pre-filtration, 2000	FISH	NOR5-730	[7]
		0.2% in February, 1.5–1.9% in summer not pre-filtered, 2007	CARD-FISH	NOR5-730 + NOR5-1238 with 4 helpers, 50% FA	This study
Germany, North Sea, Cuxhaven, surface water	53.887°N 8.641°E	0.9 ± 0.3% in July, 2007	CARD-FISH	NOR5-730 + NOR5-1238 with 4 helpers, 50% FA	This study
Spain, Blanes Bay, coastal surface water	41.67°N 2.80°E	0.6% in January, 1.3% in July, 2.6% in October, 2005	CARD-FISH	NOR5-730 + NOR5-1238 with 4 helpers, 50% FA	This study
		Detectable year round, low in winter, up to 5% in July, 2003–2004	CARD-FISH	NOR5-730, 50% FA, 35 °C	[3]
UK, Southampton dock water	50.9°N, 1.4°W	0.7%	CARD-FISH	NOR5-730, 30% FA	This study
China, Xiamen coastal surface water	24.450°N 118.074°E, 24.435°N 118.095°E	1.0–2.0% in summer, 2006 and 2007	CARD-FISH	NOR5-730 + NOR5-1238 with 4 helpers, 50% FA	This study
Pacific Ocean, Newport Hydroline, marine water, euphotic zone	44.65°N (124.42°W, 124.88°W, 125.60°W, 127.00°W)	3.4 ± 1.1%, only in euphotic zone	FISH	NOR5-730, 35% FA	[12]
Germany, North Sea, Sylt, intertidal sediment	55.04°N, 8.42°E	3% in 0–1 cm depth, 0.2% in 7–8 cm depth, 2007	CARD-FISH	NOR5-730 + NOR5-1238 with 4 helpers, 50% FA	This study
Germany, River Weser, in Bremen, fresh water	53.066°N 8.836°E	0.06%, in January 2008	CARD-FISH	NOR5-730 + NOR5-1238 with 4 helpers, 50% FA	This study
Germany, Bremen, MPI-Pond, fresh surface water	53.110°N 8.847°E	0.03%, in January 2008	CARD-FISH	NOR5-730 + NOR5-1238 with 4 helpers, 50% FA	This study
Germany, Bremen, Kuhgrabensee, salinity 2 psu, surface water	53.118°N, 8.852°E	<0.03%, in February 2008	CARD-FISH	NOR5-730 + NOR5-1238 with 4 helpers, 50% FA	This study

[a]All CARD-FISH experiments were performed at 46 °C unless stated otherwise.

lysozyme for 20 min at 35 °C. Hybridization was performed at 46 °C for 3 h, and washing was carried out at 48 °C for 15 min. For the quantification of most of the members of the NOR5/OM60 groups, a combination of NOR5-730 and NOR5-1238 with all four helpers was used (Table 1). Signal amplification was carried out for 15 min with a fluorescein-labeled tyramide. All CARD-FISH preparations were counterstained with DAPI. The relative abundance of hybridized cells was estimated as a ratio of hybridized cell counts to counts of DAPI-stained cells using epifluorescence microscopy. At least 500 DAPI-stained cells were counted. To check for autofluorescence or unspecific binding of the probe or tyramide, all samples were checked with the non-binding probe NON. The specificity of the NOR5-specific probes was checked with CARD-FISH on PFA-fixed cultures of *Congregibacter litoralis* KT71.

Results

Phylogeny

Based on an extensive comparison of trees obtained with various programs for phylogenetic reconstruction

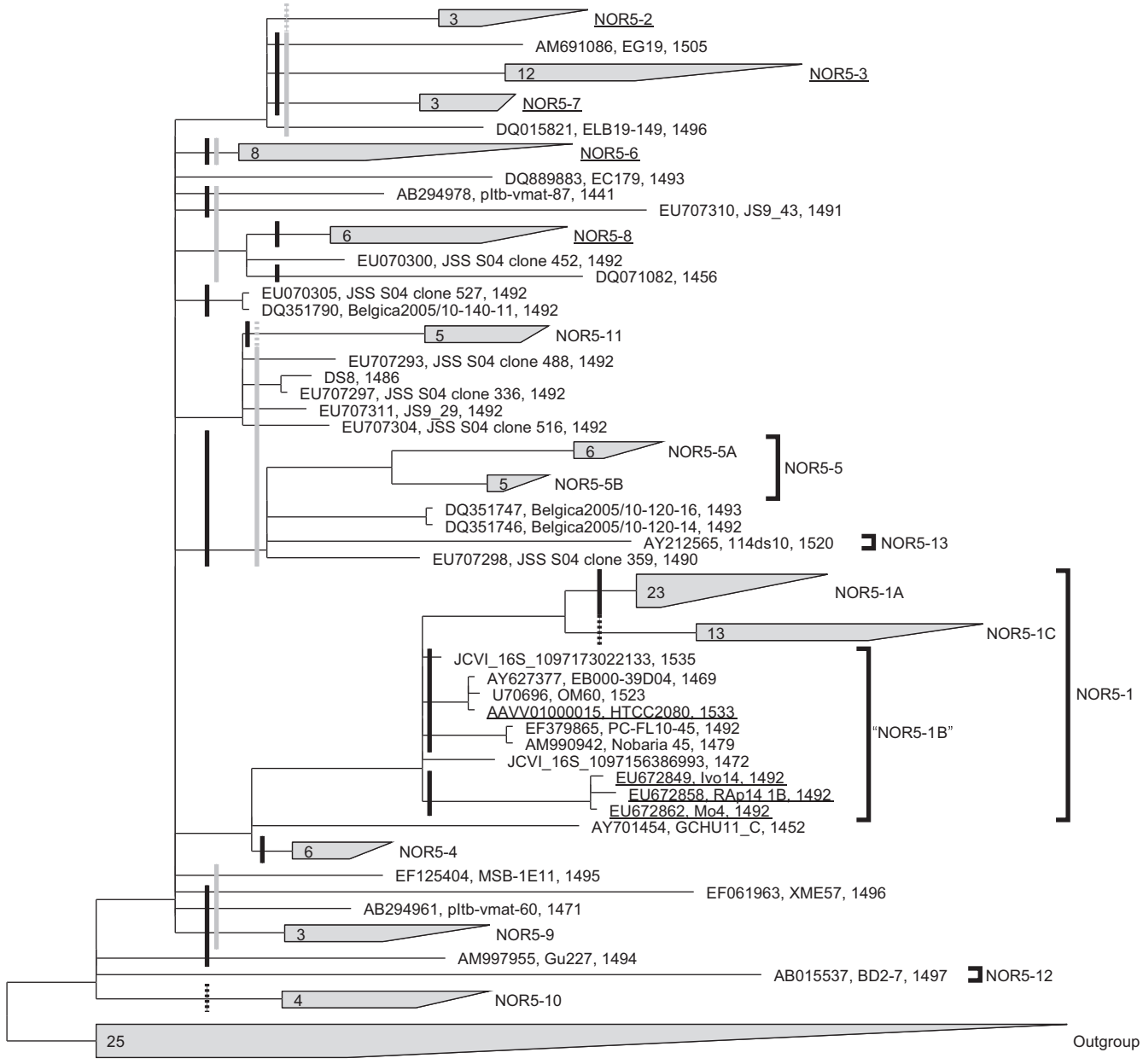

Fig. 2. Consensus tree reconstructed based on almost full-length (>1400 nt) 16S rRNA sequences of members of the NOR5/OM60 clade. Underlined names are cultured isolates and subclades that include cultured isolates. The black and grey bars on the left of the branches show the clades that can be targeted by probes NOR5-730 and NOR5-1238, respectively, and the dashed lines for partly targeted subclades.

on more than 150 almost full-length NOR5/OM60 and closely related 16S rRNA gene sequences, a new consensus tree was calculated (Fig. 2). With all treeing methods, the NOR5/OM60 clade formed a monophyletic group within *Gammaproteobacteria*. In contrast to earlier trees based on less sequences [11,22] the current reconstruction of the NOR5/OM60 clade now also includes, besides the sequences from strain KT71 and clones OM60 and OM241, a cluster of freshwater clones and BD2-7, a clone retrieved from the deep sea. Another deep-sea sequence, BD1-7, was still excluded from the NOR5/OM60 clade. Sequence identities within NOR5/OM60 were typically >92%, whereas identities to outgroup sequences were usually below 92%. However, exceptions did occur (e.g. due to imperfect sequence quality) and therefore sequence identity alone was insufficient to include or exclude a new sequence from the NOR5/OM60 clade.

The exact branching within the NOR5/OM60 clade depended on the algorithms and filters used for reconstruction. There were stable subclades in which the same sequences always clustered together. However,

the relationship between the subclades was unstable. The largest subclade was labeled as NOR5-1, and it comprised more than one-third of all the available full-length sequences, as well as many partial sequences. NOR5-1 showed two stable subgroups. The largest was NOR5-1A with more than 90 full and partial sequences (50% of NOR5-1). NOR5-1C also seemed to be monophyletic, which did not apply to the other sequences of NOR5-1 (“NOR5-1B”), which included strain HTCC2080 and several North Sea strains, such as Ivo14. Another stable subclade was NOR5-4, which was the sister group of NOR5-1 in most of the trees.

Subclade NOR5-3 included the 16S rRNA sequence of “*Congregibacter litoralis*” KT71, as well as that of 17 other NOR5/OM60 strains which have all been recently isolated from the oxic layer of marine surface sediment of the German island Sylt. However, only a few environmental clone sequences fell into this subclade. Subclades NOR5-2 and NOR5-7 were close to NOR5-3 in most phylogenetic reconstructions. They currently comprise only a few sequences, including those of the NEP isolates obtained from Japanese marine coastal sediments [45], and an isolate from coastal marine water sampled off Banyuls-sur-Mer [1,2].

Subclades NOR5-5, NOR5-6, NOR5-8, NOR5-9 and NOR5-11 together contained one-fifth of all the NOR5/OM60 sequences. Subgroups NOR5-10 and NOR5-12 were deeply branched in most of the trees and they were dominated by sequences obtained from the deep sea. BD2-7 was the only full-length sequence of NOR5-12, and showed low identity (usually $<92\%$) with other NOR5/OM60 sequences.

Clone 114ds10 (AY212565) [57] was the only full sequence in the terrestrial subclade NOR5-13, which also included 13 partial sequences recovered from fresh water, fresh water sediment or soil. Five more partial sequences retrieved from fresh water studies (EF192914, EF192886, EF192904, AY214643 and AY214720) [18,40] did not group in the NOR5-13 subclade.

About 30% of NOR5/OM60 sequences could not be grouped into any of the above-mentioned subclades, and most of these were partial sequences. The representatives of each subclade and closely related outgroups are listed in SI Table 2 as hallmarks for categorizing further sequences. The full list of all NOR5/OM60 sequences retrieved in this study, either full-length or partial, is shown in SI Table 3.

Biogeography

The geographic information was compiled for all identifications of members of the NOR5/OM60 clade (Fig. 3). This included identification by isolation, 16S rRNA gene libraries, metagenomic studies, and by fluorescence *in situ* hybridization (FISH). So far, only 14 isolates have been reported from marine water or coastal marine sediment (SI Table 4). Here, we report 22 more NOR5/OM60 strains isolated from surface sediments of an intertidal sandflat from the North Sea island

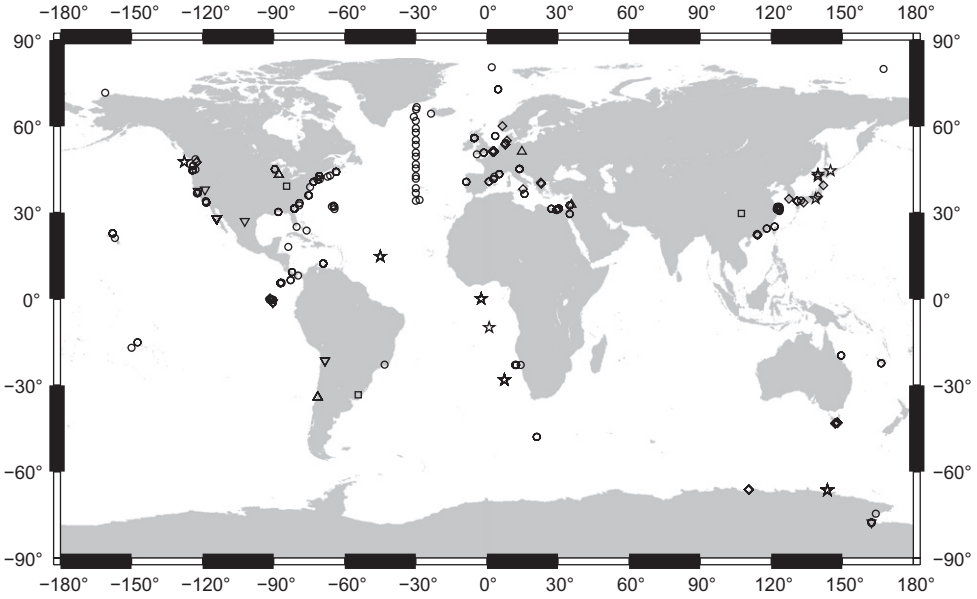

Fig. 3. Biogeography of the NOR5/OM60 clade. Sequence-, isolation-, and FISH-based identifications of NOR5/OM60 were marked on the world map. Signs refer to the habitat from where the sample was retrieved: marine water or other marine habitats – circle; marine coastal sediment – diamond; hypersaline – inverted triangle; soil – hexagon; fresh water – square; fresh sediment – triangle; deep sea – star. The map was created using the GMT (generic mapping tools) software package.

Table 3. Distribution of NOR5/OM60 subclades in different environments.

	Marine water and other marine habitats	Marine sediment	Hypersaline	Soil	Fresh water	Fresh sediment	Deep sea	Total
NOR5-1A	105	0	0	0	0	0	1	107
NOR5-1C	53	1	1	0	0	0	0	55
NOR5-1B	35	0	3	0	0	0	0	44
NOR5-4	21	0	0	0	0	0	0	23
NOR5-3	5	7	10	0	0	2	0	39
NOR5-2	2	2	0	0	0	0	0	4
NOR5-7	0	3	0	0	0	0	0	4
NOR5-5	13	7	2	2	0	0	1	30
NOR5-6	7	24	0	0	0	1	2	35
NOR5-8	12	8	0	0	0	0	0	20
NOR5-9	6	12	0	0	0	0	0	19
NOR5-11	1	6	0	0	0	0	0	11
NOR5-10	1	0	0	0	0	0	10	11
NOR5-12	0	1	0	0	0	0	2	3
NOR5-13	0	0	0	5	5	3	0	13
Other NOR5/OM60	91	56	7	1	0	1	14	179
Total	352	127	23	8	5	7	30	588

The numbers in the table give the number of 16S rRNA sequences retrieved from the public databases for a certain environment. The environmental conditions of several sequences could not be categorized; therefore the total number of a subclade can be higher than the sum of the listed numbers from different environments.

of Sylt, Germany, which belonged either to subclades NOR5-3 (17 strains) or to NOR5-1B (5 strains).

Table 3 lists the habitat preferences for each subclade according to source materials. The large subclades NOR5-1 and NOR5-4 appeared almost exclusively in the marine water column. Subclades NOR5-10 and NOR5-12 contained mainly identifications reported from deep-sea samples, and NOR5-13 was a freshwater clade. Sequences of the other NOR5/OM60 subclades were retrieved from marine sediment and the water column.

The NOR5/OM60 clade is cosmopolitan in the marine realm. Identifications have been reported from almost all oceans and at many coastal sites. In this respect, the American, European and East Asian coasts are particularly well covered with 16S rRNA gene libraries. There seems to be no latitudinal preference since NOR5/OM60 clones have been reported from mangrove [38,39] and coral reef [6,20,35], as well as sea-ice habitats [10].

NOR5/OM60 sequences were also reported in deep-sea sediments sampled near Antarctica [9] and Japan [4,28,37], as well as in the northeast Pacific [26] and the Atlantic (Schauer, unpublished). Additional reports on NOR5/OM60 sequences come from environments with different salinity: freshwater rivers [54,57], a rice paddy (DQ830363), freshwater sediments [44,61], activated sludge [32], soil [25,40], and also from hypersaline environments [24,36,52]. In addition, there are some sequences that cannot be placed on a world map since they are, for instance, from human plasma (clone NF37-A2; AY886614) [59]. Unlike the alphaproteobacterial RCA-1 cluster [55], it was still not possible in this study to detect biogeographic patterns for the various NOR5/OM60 subclades, neither latitudinal nor with respect to certain oceanic provinces.

NOR5/OM60 affiliated 16S rRNA genes in the GOS metagenomics dataset

The GOS dataset [52] contains 3728 16S rRNA gene sequences with lengths >300 nt. By comparative sequence analysis, 30 of these sequences (0.8%) could be unambiguously grouped within the NOR5/OM60 clade. Therein, 28 belonged to subclade NOR5-1, which is typical of marine surface water, and two belonged to its sister subclade NOR5-4. The sequences were found in 21 out of a total of 44 sampling stations. Except for station GS-04 (Atlantic, Canadian coast, salinity 28.3 psu), in which four out of a total of 31 16S rRNA gene sequences affiliated with the NOR5/OM60 clade, all the other libraries contained at most two NOR5/OM60 sequences. No sequences were found at low salinity stations (GS-12, salinity 3.5 psu, and GS-20,

freshwater). NOR5/OM60 16S rRNA sequences were clearly more frequent in coastal (20/1471; 1.4%) than in open-ocean samples (5/1516; 0.3%).

Design and optimization of probes for the quantification of the NOR5/OM60 clade

Our comprehensive collection of 16S rRNA sequences of the NOR5/OM60 clade facilitated a re-evaluation of oligonucleotide probes targeting this group. Probe NOR5-730 (Table 1) has been used in several studies for FISH-based quantification of members of the NOR5/OM60 clade [19]. This probe covered 131 out of 155 (84%) high-quality, almost full-length 16S rRNA sequences of the NOR5/OM60 clade included in the SILVA Ref dataset (Version 91) [50]. The probe design function of ARB [43] was used in an attempt to design new probes with an improved coverage of the NOR5/OM50 clade. However, it was not possible to design a single probe that perfectly matched all the NOR5/OM60 sequences without targeting outgroup sequences. The new probe NOR5-1238 (Table 1) targeted 46% of all high-quality NOR5/OM60 sequences, excluding the two major subclades NOR5-1 and NOR5-4. A combination of the probes NOR5-730 and NOR5-1238 increased the current coverage of the NOR5/OM60 clade to 92%, without any outgroup hits (Fig. 2). The combination failed to detect part of NOR5-1C, NOR5-2, NOR5-10, and all sequences in the NOR5-12 subclade.

Helper oligonucleotides were designed for all the above probes in an attempt to improve their hybridization efficiency (Table 1) [21]. Helpers are unlabeled oligonucleotides that bind in the vicinity of the probe, thereby presumably opening the secondary structure of the rRNA. The application of two helpers per probe significantly increased the intensity of monolabeled and CARD-FISH signals. Fixed cells of *Congregibacter litoralis* KT71 were used to determine the optimal formamide concentration for hybridization of probes NOR5-730 and NOR-1238 as 50%. These two probes were subsequently used at this formamide concentration in combination with helpers NOR5-659h, NOR5-709h, NOR5-1217h and NOR5-1287h for a specific and sensitive identification of members of the NOR5/OM60 clade in environmental samples.

Quantification of members of the NOR5/OM60 clade by CARD-FISH

The cells detected by CARD-FISH with the probe mixture NOR5-730/NOR5-1238 in marine plankton and benthos samples were pleomorphic (Fig. 4), often from coccoid to rod-shaped, sometimes were also bent in a vibrio shape. The length of the cells was between 0.5 and 3 μm, with a diameter between 0.5 and 1 μm. In plankton samples, single cells were mostly detected, suggesting that they were free-living. However, as described before [22] cells were also detected that were attached to microaggregates. In sediment samples, cells detected as NOR5/OM60 were also arranged in rosettes (Fig. 4f).

The optimized NOR5-730/NOR5-1238 probe/helper mixture was subsequently used for CARD-FISH-based quantifications in various marine samples. In the brackish to marine Yangtze River estuary (salinities of 22–32 psu), between 0% and 2.3% of all DAPI-stained cells were detected. Absolute numbers went up to 1.2×10^5 cells ml^{-1} (Table 4). No NOR5/OM60 cells were detected further up the Yangtze River at a freshwater reference site with 0.2 psu salinity.

Counts in surface waters obtained during an open-ocean North Atlantic transect in September 2006 (Vision cruise) were between 0.1% and 0.5% (3×10^3–1×10^4 cells ml^{-1}). NOR5/OM60 cells were present in all the samples, with no obvious trend from high to low latitude. The counts were only higher, at 0.9%, in a coastal sample taken during the same cruise off Iceland (SI Table 1).

In a transect in the Namibian coastal upwelling region along 23.00°S, the NOR5/OM60 counts at depths 10–15 m decreased with fluctuation from 3.0% (2.0×10^5 cells ml^{-1}) near the coast to 0.5% (1.3×10^4 cells ml^{-1}) in the open ocean (Fig. 5). Three depth profiles made at coastal (14.36°E), mid-shelf (13.15°E) and open-ocean stations (12.00°E) all clearly showed a steep decrease of the NOR5/OM60 abundances with depth. Below 70 m, the counts became marginal (<0.05%), and NOR5/OM60 cells were not detected below 300 m. With a number of 3.3×10^5 cells ml^{-1} (3.3% of DAPI) the counts were highest at the station closest to the coast (14.36°E) at a water depth of 5 m.

The highest relative abundance of NOR5/OM60 cells encountered in this study was recorded near the North Sea island Helgoland at station "Kabeltonne" (54.18°N 7.90°E). In surface water samples retrieved during May–July 2007, which were prefiltered through a 10 μm filter, members of the NOR5/OM60 clade ranged from 1.7% to 6.6% (8.2×10^3–1.2×10^5 cells ml^{-1}).

Counts were also high in sandy intertidal sediments taken at Janssand (53.72°N, 7.68°E), in the back-barrier region of the island of Spiekeroog close to the German North Sea coast. At this sample point, a range of 2.5–4.0% was detected in the top 3 cm of the sediment, and 1.4–3.1% at a depth of 3–12 cm (Fig. 6). Counts in March 2007 were generally lower than in August. The absolute number of NOR5/OM60 was in the order of

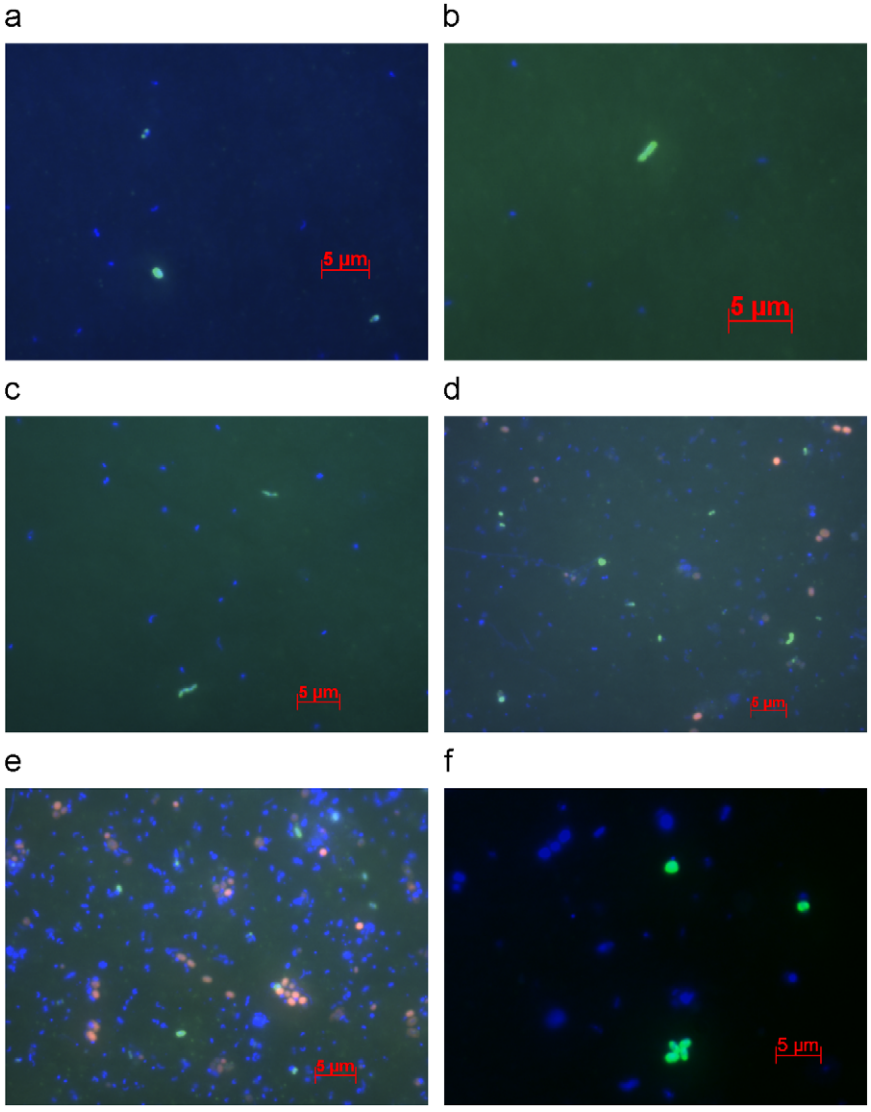

Fig. 4. Microscopic pictures of NOR5/OM60 cells in the environment: (a)–(c) North Atlantic open ocean, (d) and (e) from the Yangtze River estuary, (f) from Janssand sediment, North Sea. For all the pictures, blue: DAPI-stained DNA; green: fluorescein conferred probe signal for the NOR5/OM60 group; red: autofluorescence of cyanobacteria.

Table 4. NOR5/OM60 counts from the Yangtze River estuary cruise on September 6–8, 2006.

Station	Latitude (°N)	Longitude (°E)	Salinity (psu)	Temperature (°C)	NOR5/ OM60 count	DAPI cell count (10^6 cells ml^{-1})	Chlorophyll ($\mu g\,l^{-1}$)
CJ-1	32.00	122.01	25.1	27.8	0	n.d.	1.3
CJ-2	32.00	122.50	29.6	26.5	0.6±0.2%	4.1	0.4
CJ-3	32.00	123.00	31.2	25.2	1.6±0.3%	n.d.	2.2
CJ-4	32.00	123.50	31.2	25.6	0.9±0.2%	7.8	1.3
CJ-5	31.50	123.50	31.2	26.3	0.4±0.1%	10.8	0.9
CJ-12	31.00	123.50	31.8	28.5	0.10±0.05%	n.d.	0.5
CJ-14	30.50	123.50	32.3	28.0	0.5±0.1%	3.5	1.1
CJ-13	30.80	123.00	30.1	26.7	2.3±0.6%	4.1	3.9
CJ-11	31.00	122.60	28.1	27.1	1.4±0.3%	8.7	2.3
CJ-9	31.20	122.30	22.5	27.2	1.2±0.3%	3.7	1.0
CJ-17	31.38	121.60	0.2	28.9	0	n.d.	n.d.

n.d.: not determined.

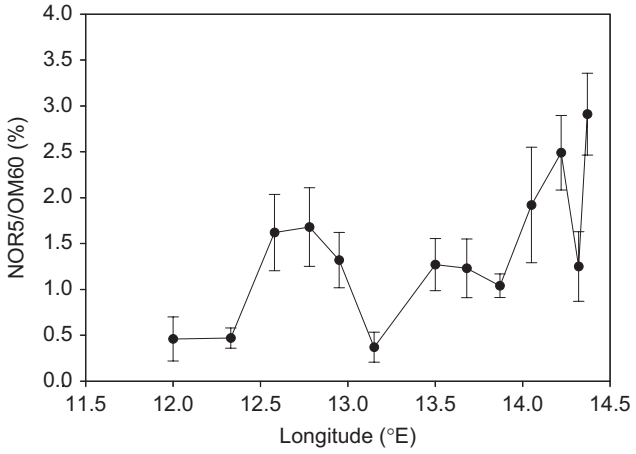

Fig. 5. NOR5/OM60 counts by CARD-FISH in the Namibian upwelling region on March 22–23, 2003, along 23°S near Walvis Bay. The counts were mostly at a depth of 10 m, with the exception of the two easternmost points that were at a depth of 15 m.

10^7 cells cm^{-3}, and in the surface sediments it was as high as 1.5×10^8 cells cm^{-3}.

Further quantifications of the NOR5/OM60 clade, either from this study or from former studies, are summarized in Table 3. A few preliminary quantifications of NOR5/OM60 in freshwater samples were also taken from Bremen, Germany. Abundances were less than 0.1% in the River Weser and from two ponds, one freshwater and the other with a salinity of 2 psu.

Discussion

The interest in the NOR5/OM60 clade has significantly increased with the discovery that this clade encompasses the elusive marine gammaproteobacterial branch of AAnPs [12,22]. The data obtained in this study clearly support the hypothesis of a cosmopolitan

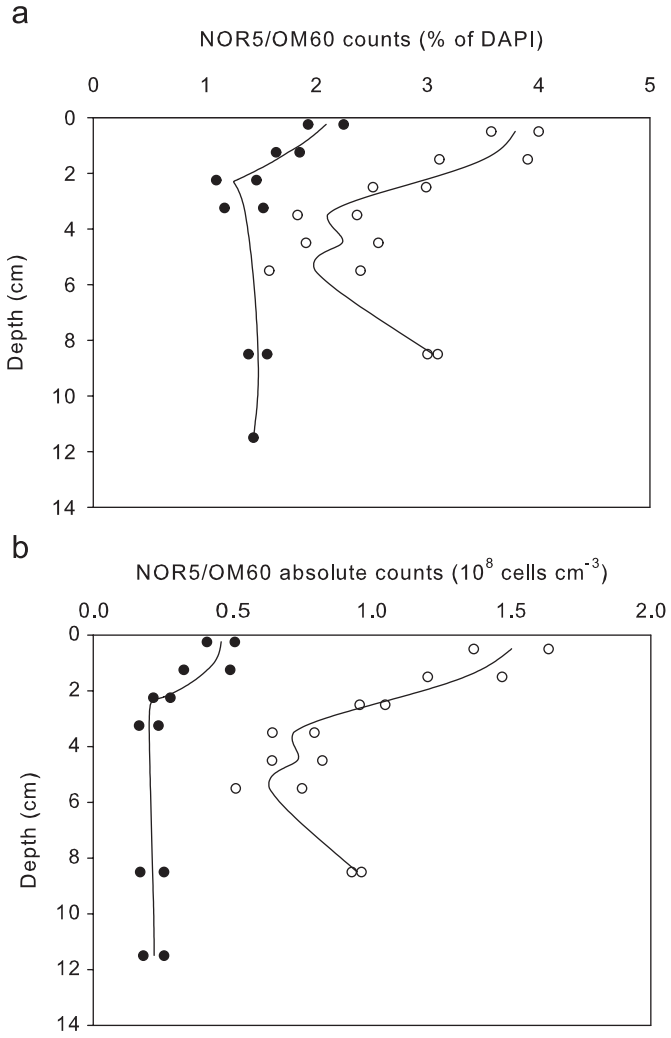

Fig. 6. NOR5/OM60 counts from Janssand sediment in March (black circles) and August (white circles) 2007. At each time point, two adjacent cores were sampled for duplication.

distribution of the NOR5/OM60 clade in marine surface waters. The presence of NOR5/OM60 is now confirmed for all oceans except the Indian Ocean, which has still not yet been examined in this regard. Several lines of evidence suggest that members of the NOR5/OM60 clade are generally more abundant in coastal areas than in open-ocean settings. CARD-FISH counts in coastal surface waters ($N = 30$) showed an average of $2.1 \pm 1.5\%$, whereas open-ocean surface water samples ($N = 36$) had an average of $0.5 \pm 0.4\%$. Fully independent support for this finding comes from the GOS dataset [52] in which the frequency of NOR5/OM60 16S rRNA gene sequences was significantly higher in coastal stations (1.4%) than in open-ocean stations (0.3%). Nevertheless, these data cannot be directly translated into cell frequencies since the number of rRNA operons per genome is quite different in marine bacteria [31]. The two fully sequenced strains of the NOR5/OM60 clade, "*Congregibacter litoralis*" KT71 and HTCC2080 contain two and one ribosomal operon, respectively. When the copy number of the NOR5/OM60 group members is less than the average, the frequency of 16S rRNA genes in the metagenomic library might underestimate the relative abundance of cells, and *vice versa*.

Our large CARD-FISH dataset also provides convincing support for a preference of members of the NOR5/OM60 clade for the euphotic zone, which had been reported for the coastal Pacific Newport Hydroline station [12]. Once again, support for this conspicuous depth distribution comes from metagenomics. In fosmid libraries constructed from bacterioplankton samples at Aloha Station, Hawaii, 16S rRNA genes of the NOR5/OM60 clade were detected at depths of 10 m and 70 m, but not at 130 m or deeper [17]. However, since the absence of NOR5/OM60 sequences is not valid proof of their absence *per se*, other methods, such as quantitative PCR, need to be applied to quantify NOR5/OM60 in these deep water layers.

There are indications for strong seasonal fluctuation of the NOR5/OM60 abundance in coastal waters. Three studies on North Sea surface water had reported NOR5/OM60 blooms up to 13% co-occurring with, for example, a dinoflagellate bloom [7,19,48]. Similar observations were recently reported for northwest Mediterranean coastal waters [3]. In our study, the samples from Xiamen, Barcelona and Helgoland showed the same trend, with high counts of NOR5/OM60 co-occurring with algal blooms. The photoheterotrophic members of the NOR5/OM60 clade benefit from algal photosynthesis, yet it is too early to speculate on a specific link to particular algal species. We hope that the more sensitive and specific probes and protocols developed in this study will facilitate more detailed investigations on the seasonality of NOR5/OM60 abundance in the future.

Besides season, water depth, and distance to the coast, we also searched for other parameters that might influence the distribution of NOR5/OM60 in the water column. Consequently, linear regression analysis was used to check for correlation between NOR5/OM60 abundance and other parameters. In the Namibian transect, NOR5/OM60 abundance was highly correlated to turbidity ($r^2 = 0.79$) (Fig. 7a). The NOR5/OM60 isolate *Congregibacter litoralis* KT71 is known to form aggregates in pure culture. Also, *in situ* attachment of NOR5/OM60 cells to aggregates has been demonstrated before [22]. NOR5/OM60 was positively correlated to chlorophyll fluorescence ($r^2 = 0.73$) (Fig. 7b). Similarly, in the surface waters of the Yangtze River estuary, the NOR5/OM60 abundance showed a strong positive correlation with chlorophyll concentration ($r^2 = 0.74$) (Fig. 7c). Algae are a source of fresh organic material, which in turn could serve as a substrate for NOR5/OM60. Indeed, the cultured strain KT71 prefers short oligomers and amino acids as substrates [22], which are generally rapidly consumed in the water column [63]. So, it might be advantageous to stay close to the site of production of these substrates. No other significant correlation of NOR5/OM60 abundance was detected with temperature, salinity or total bacterio-

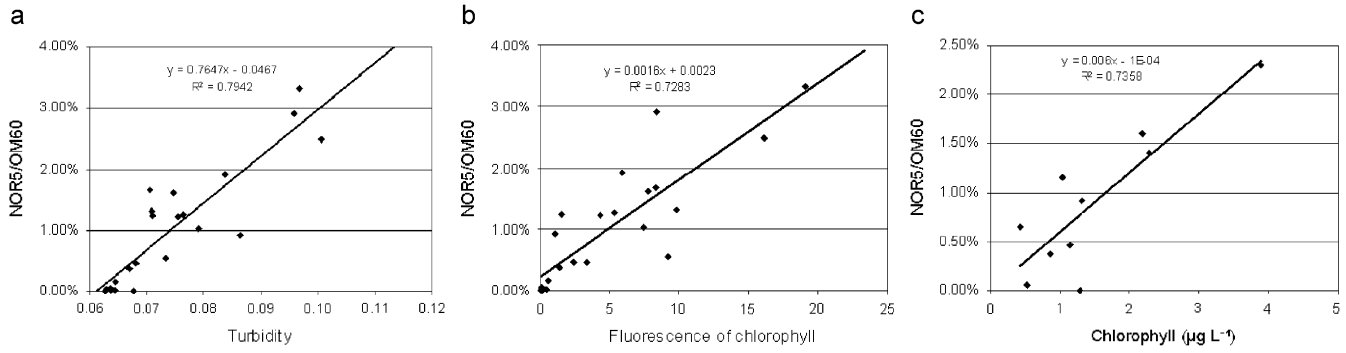

Fig. 7. Correlation of NOR5/OM60 to other environmental parameters: (a) turbidity in the Namibian transect, (b) chlorophyll fluorescence in the Namibian transect, (c) chlorophyll in the Yangtze River estuary.

plankton cell counts at either site. Also, during the North Atlantic Vision cruise, the NOR5/OM60 proportion in the surface water (10 m) showed no obvious correlation with any detected parameters (latitude, temperature, total DAPI cell count or chlorophyll fluorescence).

The NOR5/OM60 group and the AAnPs showed some common features for distribution, at least in some regions: they both occurred at a higher percentage in coastal water compared to the open ocean, and they were more abundant in summer or autumn than in winter or spring. Most of them appear in the euphotic zone in the marine water column and they are positively related to high chlorophyll concentrations [15,30,53,56,62]. It seems that many members of NOR5/OM60 may indeed be AAnPs. However, considering the rather high 16S rRNA sequence diversity within the NOR5/OM60 clade and its broad habitat range, it cannot be taken for granted that all members of NOR5/OM60 are AAnPs. The four strains isolated from marine surface waters KT71 (NOR5-3), HTCC2080 (NOR5-1B), HTCC2246 (not grouped) and HTCC2148 (NOR5-8) were shown to contain genes *puf*L and *puf*M coding for light-harvesting complex I (LHC I). However, *puf*L and *puf*M trees lacked congruence with the 16S rRNA phylogeny [12]. The *puf*M gene of HTCC2080 belongs to *puf*M Group K, the sister group of KT71, while the *puf*L and *puf*M genes of the HTCC2246 and HTCC2148 groups belong to different clusters of *Alphaproteobacteria* [62]. This suggests that the loss and gain of photosynthesis operons might be quite frequent in the NOR5/OM60 clade. For bacteria with a dominantly heterotrophic metabolism, photosynthesis might just be an accessory energy source that may not be required by all members of the NOR5/OM60 clade. Further studies that might combine the *in situ* identification of NOR5/OM60 cells by CARD-FISH with the direct identification of AAnP using infrared fluorescence microscopy [29,53] are therefore needed. Only then could we directly determine how many NOR5/OM60 are AAnPs, and, *vice versa*, how many of the AAnP are NOR5/OM60. In this respect, Yutin and colleagues [62] have recently reported on compositional changes within the AAnP between coastal and open-ocean sites. They have assessed the diversity of marine AAnPs based on comparative analysis of *puf*M sequences retrieved from the metagenomic libraries of the GOS [52]. Only six out of 85 scaffolds identified to include *puf*M in the oxic samples of the GOS dataset were highly similar to the *puf*M of *Congregibacter litoralis* KT71 and therefore likely to be from the NOR5/OM60 clade. These were from coastal samples, whereas the *puf*M of the alphaproteobacterial *Roseobacter*-type was found throughout the samples, and, additionally, as yet unidentified groups of AAnP seemed to dominate pelagic marine waters [62].

Acknowledgements

We thank Julia Arnds, Regina Schauer, Marc Mußmann, and Alaa Bakr for providing sequences. Pep Gasol, Paola Gomez and Ilaria Pizetti provided marine surface water samples from Barcelona, the North Atlantic (Vision cruise), and sampling station Kabeltonne off Helgoland, respectively. We acknowledge the staff of the State Key Laboratory at Xiamen University, especially Yao Zhang for organizing the cruise to the Yangtze River estuary, and Fan Zhang for providing chlorophyll data. The work was funded by the Max Planck Society, the German Academic Exchange Service (DAAD) and the China Scholarship Council (CSC).

Appendix A. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.syapm.2008.12.001.

References

[1] H. Agogué, E.O. Casamayor, M. Bourrain, I. Obernosterer, F. Joux, G.J. Herndl, P. Lebaron, A survey on bacteria inhabiting the sea surface microlayer of coastal ecosystems, FEMS Microbiol. Ecol. 54 (2005) 269–280.

[2] H. Agogué, F. Joux, I. Obernosterer, P. Lebaron, Resistance of marine bacterioneuston to solar radiation, Appl. Environ. Microbiol. 71 (2005) 5282–5289.

[3] L. Alonso-Sáez, V. Balagué, E.L. Sà, O. Sánchez, J.M. González, J. Pinhassi, R. Massana, J. Pernthaler, C. Pedrós-Alió, J.M. Gasol, Seasonality in bacterial diversity in north-west Mediterranean coastal waters: assessment through clone libraries, fingerprinting and FISH, FEMS Microbiol. Ecol. 60 (2007) 98–112.

[4] S. Arakawa, T. Sato, R. Sato, J. Zhang, T. Gamo, U. Tsunogai, A. Hirota, Y. Yoshida, R. Usami, F. Inagaki, C. Kato, Molecular phylogenetic and chemical analyses of the microbial mats in deep-sea cold seep sediments at the northeastern Japan Sea, Extremophiles 10 (2006) 311–319.

[5] K.E. Ashelford, N.A. Chuzhanova, J.C. Fry, A.J. Jones, A.J. Weightman, At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies, Appl. Environ. Microbiol. 71 (2005) 7724–7736.

[6] O. Barneah, E. Ben-Dov, E. Kramarsky-Winter, A. Kushmaro, Characterization of black band disease in Red Sea stony corals, Environ. Microbiol. 9 (2007) 1995–2006.

[7] C. Beardsley, J. Pernthaler, W. Wosniok, R. Amann, Are readily culturable bacteria in coastal North Sea waters suppressed by selective grazing mortality?, Appl. Environ. Microbiol. 69 (2003) 2624–2630.

[8] O. Béjà, M.T. Suzuki, J.F. Heidelberg, W.C. Nelson, C.M. Preston, T. Hamada, J.A. Eisen, C.M. Fraser, E.F. DeLong, Unsuspected diversity among marine aerobic anoxygenic phototrophs, Nature 415 (2002) 630–633.

[9] J.P. Bowman, R.D. McCuaig, Biodiversity, community structural shifts, and biogeography of prokaryotes within Antarctic continental shelf sediment, Appl. Environ. Microbiol. 69 (2003) 2463–2483.

[10] R. Brinkmeyer, K. Knittel, J. Jurgens, H. Weyland, R. Amann, E. Helmke, Diversity and structure of bacterial communities in Arctic versus Antarctic pack ice, Appl. Environ. Microbiol. 69 (2003) 6610–6619.

[11] J.C. Cho, S.J. Giovannoni, Cultivation and growth characteristics of a diverse group of oligotrophic marine *Gammaproteobacteria*, Appl. Environ. Microbiol. 70 (2004) 432–440.

[12] J.C. Cho, M.D. Stapels, R.M. Morris, K.L. Vergin, M.S. Schwalbach, S.A. Givan, D.F. Barofsky, S.J. Giovannoni, Polyphyletic photosynthetic reaction centre genes in oligotrophic marine *Gammaproteobacteria*, Environ. Microbiol. 9 (2007) 1456–1463.

[13] J.R. Cole, B. Chai, T.L. Marsh, R.J. Farris, Q. Wang, S.A. Kulam, S. Chandra, D.M. McGarrell, T.M. Schmidt, G.M. Garrity, J.M. Tiedje, The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy, Nucleic Acids Res. 31 (2003) 442–443.

[14] S.A. Connon, S.J. Giovannoni, High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates, Appl. Environ. Microbiol. 68 (2002) 3878–3885.

[15] M.T. Cottrell, A. Mannino, D.L. Kirchman, Aerobic anoxygenic phototrophic bacteria in the Mid-Atlantic Bight and the North Pacific Gyre, Appl. Environ. Microbiol. 72 (2006) 557–564.

[16] H. Daims, A. Bruhl, R. Amann, K.H. Schleifer, M. Wagner, The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set, Syst. Appl. Microbiol. 22 (1999) 434–444.

[17] E.F. DeLong, C.M. Preston, T. Mincer, V. Rich, S.J. Hallam, N.U. Frigaard, A. Martinez, M.B. Sullivan, R. Edwards, B.R. Brito, S.W. Chisholm, D.M. Karl, Community genomics among stratified microbial assemblages in the ocean's interior, Science 311 (2006) 496–503.

[18] C. Dorador, G. Castillo, K.P. Witzel, I. Vila, Bacterial diversity in the sediments of a temperate artificial lake, Rapel reservoir, Rev. Chil. Hist. Nat. 80 (2007) 213–224.

[19] H. Eilers, J. Pernthaler, J. Peplies, F.O. Glöckner, G. Gerdts, R. Amann, Isolation of novel pelagic bacteria from the German bight and their seasonal contributions to surface picoplankton, Appl. Environ. Microbiol. 67 (2001) 5134–5142.

[20] J. Frias-Lopez, A.L. Zerkle, G.T. Bonheyo, B.W. Fouke, Partitioning of bacterial communities between seawater and healthy black band diseased, and dead coral surfaces, Appl. Environ. Microbiol. 68 (2002) 2214–2228.

[21] B.M. Fuchs, F.O. Glöckner, J. Wulf, R. Amann, Unlabeled helper oligonucleotides increase the in situ accessibility to 16S rRNA of fluorescently labeled oligonucleotide probes, Appl. Environ. Microbiol. 66 (2000) 3603–3607.

[22] B.M. Fuchs, S. Spring, H. Teeling, C. Quast, J. Wulf, M. Schattenhofer, S. Yan, S. Ferriera, J. Johnson, F.O. Glöckner, R. Amann, Characterization of a marine gammaproteobacterium capable of aerobic anoxygenic photosynthesis, Proc. Natl. Acad. Sci. USA 104 (2007) 2891–2896.

[23] B.M. Fuchs, G. Wallner, W. Beisker, I. Schwippl, W. Ludwig, R. Amann, Flow cytometric analysis of the in situ accessibility of *Escherichia coli* 16S rRNA for fluorescently labeled oligonucleotide probes, Appl. Environ. Microbiol. 64 (1998) 4973–4982.

[24] R.E. Glatz, P.W. Lepp, B.B. Ward, C.A. Francis, Planktonic microbial community composition across steep physical/chemical gradients in permanently ice-covered Lake Bonney, Antarctica, Geobiology 4 (2006) 53–67.

[25] M. Hartmann, F. Widmer, Community structure analyses are more sensitive to differences in soil bacterial communities than anonymous diversity indices, Appl. Environ. Microbiol. 72 (2006) 7804–7812.

[26] J.A. Huber, H.P. Johnson, D.A. Butterfield, J.A. Baross, Microbial life in ridge flank crustal fluids, Environ. Microbiol. 8 (2006) 88–99.

[27] J.P. Huelsenbeck, F. Ronquist, MRBAYES: Bayesian inference of phylogenetic trees, Bioinformatics 17 (2001) 754–755.

[28] F. Inagaki, M. Suzuki, K. Takai, H. Oida, T. Sakamoto, K. Aoki, K.H. Nealson, K. Horikoshi, Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk, Appl. Environ. Microbiol. 69 (2003) 7224–7235.

[29] N.Z. Jiao, Y. Zhang, Y. Chen, Time series observation based InfraRed Epifluorescence Microscopic (TIREM) approach for accurate enumeration of bacteriochlorophyll-containing microbes in marine environments, J. Microbiol. Methods 65 (2006) 442–452.

[30] N.Z. Jiao, Y. Zhang, Y.H. Zeng, N. Hong, R.L. Liu, F. Chen, P.X. Wang, Distinct distribution pattern of abundance and diversity of aerobic anoxygenic phototrophic bacteria in the global ocean, Environ. Microbiol. 9 (2007) 3091–3099.

[31] J.A. Klappenbach, J.M. Dunbar, T.M. Schmidt, RRNA operon copy number reflects ecological strategies of bacteria, Appl. Environ. Microbiol. 66 (2000) 1328–1333.

[32] A.N. Klein, D. Frigon, L. Raskin, Populations related to Alkanindiges, a novel genus containing obligate alkane degraders, are implicated in biological foaming in activated sludge systems, Environ. Microbiol. 9 (2007) 1898–1912.

[33] Z.S. Kolber, F.G. Plumley, A.S. Lang, J.T. Beatty, R.E. Blankenship, C.L. van Dover, C. Vetriani, M. Koblizek, C. Rathgeber, P.G. Falkowski, Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean, Science 292 (2001) 2492–2495.

[34] Z.S. Kolber, C.L. van Dover, R.A. Niederman, P.G. Falkowski, Bacterial photosynthesis in surface waters of the open ocean, Nature 407 (2000) 177–179.

[35] O. Koren, E. Rosenberg, Bacteria associated with mucus and tissues of the coral *Oculina patagonica* in summer and winter, Appl. Environ. Microbiol. 72 (2006) 5254–5259.
[36] R.E. Ley, J.K. Harris, J. Wilcox, J.R. Spear, S.R. Miller, B.M. Bebout, J.A. Maresca, D.A. Bryant, M.L. Sogin, N.R. Pace, Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat, Appl. Environ. Microbiol. 72 (2006) 3685–3695.
[37] L.N. Li, C. Kato, K. Horikoshi, Bacterial diversity in deep-sea sediments from different depths, Biodivers. Conserv. 8 (1999) 659–677.
[38] J.B. Liang, Y.Q. Chen, C.Y. Lan, N.F.Y. Tam, Q.J. Zan, L.N. Huang, Recovery of novel bacterial diversity from mangrove sediment, Mar. Biol. 150 (2007) 739–747.
[39] P.C. Liao, B.H. Huang, S. Huang, Microbial community composition of the Danshui river estuary of northern Taiwan and the practicality of the phylogenetic method in microbial barcoding, Microb. Ecol. 54 (2007) 497–507.
[40] M.R. Liles, B.F. Manske, S.B. Bintrim, J. Handelsman, R.M. Goodman, A census of rRNA genes and linked genomic sequences within a soil metagenomic library, Appl. Environ. Microbiol. 69 (2003) 2684–2691.
[41] E. Llobet-Brossa, R. Rosselló-Mora, R. Amann, Microbial community composition of Wadden Sea sediments as revealed by fluorescence in situ hybridization, Appl. Environ. Microbiol. 64 (1998) 2691–2696.
[42] W. Ludwig, O. Strunk, S. Klugbauer, N. Klugbauer, M. Weizenegger, J. Neumaier, M. Bachleitner, K.H. Schleifer, Bacterial phylogeny based on comparative sequence analysis, Electrophoresis 19 (1998) 554–568.
[43] W. Ludwig, O. Strunk, R. Westram, L. Richter, H. Meier, Yadhukumar, A. Buchner, T. Lai, S. Steppi, G. Jobb, W. Forster, I. Brettske, S. Gerber, A.W. Ginhart, O. Gross, S. Grumann, S. Hermann, R. Jost, A. Konig, T. Liss, R. Lussmann, M. May, B. Nonhoff, B. Reichel, R. Strehlow, A. Stamatakis, N. Stuckmann, A. Vilbig, M. Lenke, T. Ludwig, A. Bode, K.H. Schleifer, ARB: a software environment for sequence data, Nucleic Acids Res. 32 (2004) 1363–1371.
[44] B.J. MacGregor, S. Toze, E.W. Alm, R. Sharp, C.J. Ziemer, D.A. Stahl, Distribution and abundance of Gram-positive bacteria in the environment: development of a group-specific probe, J. Microbiol. Methods 44 (2001) 193–203.
[45] T. Maeda, K. Hayakawa, M. You, M. Sasaki, Y. Yamaji, M. Furushita, T. Shiba, Characteristics of nonylphenol polyethoxylate-degrading bacteria isolated from coastal sediments, Microbes Environ. 20 (2005) 253–257.
[46] W. Manz, R. Amann, W. Ludwig, M. Wagner, K.H. Schleifer, Phylogenetic oligodeoxynucleotide probes for the major subclasses of *Proteobacteria*-problems and solutions, Syst. Appl. Microbiol. 15 (1992) 593–600.
[47] T.D. Mullins, T.B. Britschgi, R.L. Krest, S.J. Giovannoni, Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities, Limnol. Oceanogr. 40 (1995) 148–158.
[48] A. Pernthaler, J. Pernthaler, Diurnal variation of cell proliferation in three bacterial taxa from coastal North Sea waters, Appl. Environ. Microbiol. 71 (2005) 4638–4644.
[49] A. Pernthaler, J. Pernthaler, R. Amann, Fluorescence in situ hybridization and catalyzed reporter deposition for the identification of marine bacteria, Appl. Environ. Microbiol. 68 (2002) 3094–3101.
[50] E. Pruesse, C. Quast, K. Knittel, B.M. Fuchs, W. Ludwig, J. Peplies, F.O. Glöckner, SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB, Nucleic Acids Res. (2007).
[51] M.S. Rappé, P.F. Kemp, S.J. Giovannoni, Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina, Limnol. Oceanogr. 42 (1997) 811–826.
[52] D.B. Rusch, A.L. Halpern, G. Sutton, K.B. Heidelberg, S. Williamson, S. Yooseph, D.Y. Wu, J.A. Eisen, J.M. Hoffman, K. Remington, K. Beeson, B. Tran, H. Smith, H. Baden-Tillson, C. Stewart, J. Thorpe, J. Freeman, C. Andrews-Pfannkoch, J.E. Venter, K. Li, S. Kravitz, J.F. Heidelberg, T. Utterback, Y.H. Rogers, L.I. Falcon, V. Souza, G. Bonilla-Rosso, L.E. Eguiarte, D.M. Karl, S. Sathyendranath, T. Platt, E. Bermingham, V. Gallardo, G. Tamayo-Castillo, M.R. Ferrari, R.L. Strausberg, K. Nealson, R. Friedman, M. Frazier, J.C. Venter, The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through Eastern Tropical Pacific, PLoS Biol. 5 (2007) 398–431.
[53] M.S. Schwalbach, J.A. Fuhrman, Wide-ranging abundances of aerobic anoxygenic phototrophic bacteria in the world ocean revealed by epifluorescence microscopy and quantitative PCR, Limnol. Oceanogr. 50 (2005) 620–628.
[54] H. Sekiguchi, M. Watanabe, T. Nakahara, B.H. Xu, H. Uchiyama, Succession of bacterial community structure along the Changjiang River determined by denaturing gradient gel electrophoresis and clone library analysis, Appl. Environ. Microbiol. 68 (2002) 5142–5150.
[55] N. Selje, M. Simon, T. Brinkhoff, A newly discovered *Roseobacter* cluster in temperate and polar oceans, Nature 427 (2004) 445–448.
[56] M.E. Sieracki, I.C. Gilg, E.C. Thier, N.J. Poulton, R. Goericke, Distribution of planktonic aerobic anoxygenic photoheterotrophic bacteria in the northwest Atlantic, Limnol. Oceanogr. 51 (2006) 38–46.
[57] J.M. Simpson, J.W.S. Domingo, D.J. Reasoner, Assessment of equine fecal contamination: the search for alternative bacterial source-tracking targets, FEMS Microbiol. Ecol. 47 (2004) 65–75.
[58] M.T. Suzuki, C.M. Preston, O. Béjà, J.R. de la Torre, G.F. Steward, E.F. DeLong, Phylogenetic screening of ribosomal RNA gene-containing clones in bacterial artificial chromosome (BAC) libraries from different depths in Monterey Bay, Microb. Ecol. 48 (2004) 473–488.
[59] S.D. Vernon, S.K. Shukla, J. Conradt, E.R. Unger, W.C. Reeves, Analysis of 16S rRNA gene sequences and circulating cell-free DNA from plasma of chronic fatigue syndrome and non-fatigued subjects, BMC Microbiol. 2 (2002).
[60] G. Wallner, R. Amann, W. Beisker, Optimizing fluorescent in situ hybridization with ribosomal-RNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms, Cytometry 14 (1993) 136–143.

[61] A. Wobus, C. Bleul, S. Maassen, C. Scheerer, M. Schuppler, E. Jacobs, I. Roske, Microbial diversity and functional characterization of sediments from reservoirs of different trophic state, FEMS Microbiol. Ecol. 46 (2003) 331–347.
[62] N. Yutin, M.T. Suzuki, H. Teeling, M. Weber, J.C. Venter, D.B. Rusch, O. Béjà, Assessing diversity and biogeography of aerobic anoxygenic phototrophic bacteria in surface waters of the Atlantic and Pacific Oceans using the Global Ocean Sampling expedition metagenomes, Environ. Microbiol. 9 (2007) 1464–1475.
[63] M.V. Zubkov, B.M. Fuchs, P.H. Burkill, R. Amann, Comparison of cellular and biomass specific activities of dominant bacterioplankton groups in stratified waters of the Celtic Sea, Appl. Environ. Microbiol. 67 (2001) 5210–5218.