

Novel estuarine bacterioplankton in rRNA operon libraries from the Chesapeake Bay

Jinjun Kan^{1,3}, Sarah E. Evans², Feng Chen¹, Marcelino T. Suzuki^{2,*}

¹Center of Marine Biotechnology, University of Maryland Biotechnology Institute, Baltimore, Maryland 21202, USA

²Chesapeake Biological Laboratory, University of Maryland Center for Environmental Science, Solomons, Maryland 20688, USA

³Present address: Department of Earth Sciences, University of Southern California, Los Angeles, California 90089, USA

ABSTRACT: Compared to open oceans and freshwater systems, less is known about the phylogenetic diversity in temperate estuaries. In the present study, 6 rRNA operon libraries constructed along the salinity gradient in the Chesapeake Bay in 2 different seasons were analyzed and sequenced. Phylogenetic analyses showed that Chesapeake Bay bacterioplankton represented a mixture of typical marine and freshwater clades, but several groups that had not been previously retrieved from either system were also found in the bay. Unique or novel SAR11 (*Pelagibacter ubique*), *Roseobacter*, SAR86 and *Actinobacteria* subclades present in the bay suggested that these microorganisms might be adapted to large temperate estuaries with long residence times, such as the Chesapeake Bay. These results represent the first comprehensive study of phylogenetic diversity in estuaries with long residence times.

KEY WORDS: Phylogenetic analysis · rRNA operon libraries · Estuarine bacterioplankton

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Bacterioplankton is an important component of aquatic ecosystems due to its abundance, biomass and contributions to the cycling of major elements. Thus, numerous studies have been conducted to measure bulk properties and activity of bacterioplankton in aquatic ecosystems. Recently, the use of molecular biology techniques has allowed the study of diversity, distribution and activity of aquatic bacterioplankton at finer phylogenetic resolution, particularly in oceanic (e.g. Giovannoni & Rappé 2000, Venter et al. 2004, DeLong et al. 2006), coastal (e.g. Suzuki et al. 2001, Buchan et al. 2005) and, more recently, freshwater environments (e.g. Crump et al. 1999, Zwart et al. 2002, Warnecke et al. 2004). Interestingly, significantly less is known regarding the diversity in estuarine environments, and in particular those with longer residence times, such as the Chesapeake Bay.

Estuarine bacterioplankton are subject to the influence of freshwater and seawater, and, depending on residence time, an estuarine community can also form (Crump et al. 1999). Few studies have shown that estuarine bacteria contain mixed populations of both freshwater and marine origins, as well as populations adapted to estuarine ecosystems (Crump et al. 1999, Henriques et al. 2004, Hewson & Fuhrman 2004), although most of these studies were conducted in relatively small systems or in estuaries with short residence times. A limited number of studies have targeted the composition of bacterioplankton in the Chesapeake Bay (e.g. Bidle & Fletcher 1995, Bouvier & del Giorgio 2002), one of the largest estuaries in the world, with an average residence time of about 7 mo (Nixon et al. 1996). However, these previous studies only characterized the Chesapeake Bay bacterioplankton community either on a broad level (phyla or classes) or at a narrow resolution, focusing on individ-

ual specific genera/species (e.g. Bidle & Fletcher 1995, Bouvier & del Giorgio 2002). With few systematic studies of sequencing 16S rRNA genes, currently little is known about Chesapeake Bay bacterioplankton phylogenetic diversity.

In order to better understand the population composition and seasonal variation of bacterioplankton in a large estuary with a long residence time, we constructed 6 rRNA operon clone libraries from northern, middle and southern Chesapeake Bay, in 2 different seasons. A total of 576 clones was screened by ITS (internal transcribed spacer)-LH (length heterogeneity)-PCR analysis, and the combined fragment sizes were compared to a database of previously sized fragments (Suzuki et al. 2004). Numerically representative clones and those containing unique and novel size combinations were sequenced and identified by phylogenetic analysis of full-length 16S rRNA genes. A detailed comparison of clonal composition and frequency in each of the libraries and comparison to other fingerprinting methods are described elsewhere (Kan et al. 2007); thus, the focus here is on a fine-resolution phylogenetic analysis of Chesapeake Bay bacterioplankton and other aquatic ecosystems.

MATERIALS AND METHODS

DNA sampling, library construction and clone screening by ITS-LH-PCR. Details of DNA sampling, construction of rRNA operon clone libraries and ITS-LH-PCR analysis are described elsewhere (Kan et al. 2007). Briefly, 6 rRNA operon libraries were constructed by cloning ca. 3.5 to 4 kb fragments amplified by PCR using primers 16S rRNA 27F and 23S rRNA 1933R (Kan et al. 2007). These fragments contained the 16S rRNA gene, ITS and a large part of 23S rRNA gene from bacterioplankton collected at 3 sites in September 2002 and March 2003 (Fig. 1, Table 2). The libraries were named CB0 (northern bay, 09/02), CB1 (mid-bay, 09/02), CB2 (southern bay, 09/02), CB3 (northern bay,

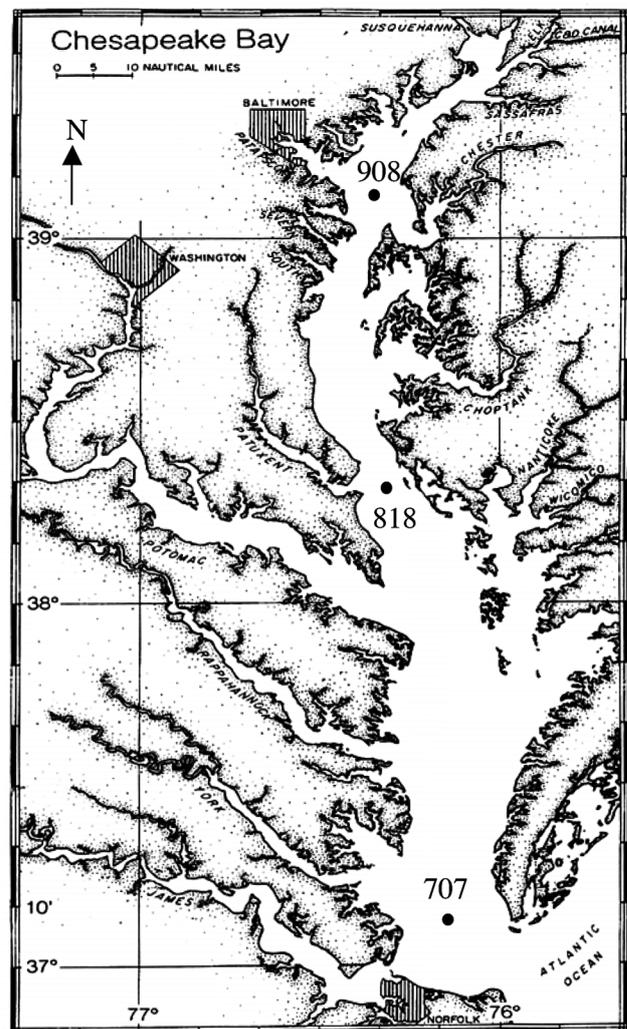


Fig. 1. Sampling stations in the Chesapeake Bay (modified from Smith et al. 1992)

03/03), CB4 (mid-bay, 03/03) and CB5 (southern bay, 03/03). Ninety-six clones per library were screened by ITS-LH-PCR (Suzuki et al. 2004), a method where the combined length of the ITS and the presence and posi-

Table 1. Primers used for sequencing in the present study

Primer	Sequence (5'–3')	Target site	Source
16S-27F	AGAGTTTGATCCTGGCTCAG	16S 8–27	DeLong et al. (1993)
16S-1074F	ATGGCTGTCGTCAGCTCGTG	16S 1055–1074	Suzuki et al. (2004)
16S-1100R	AGGGTTGCGCTCGTTG	16S 1100–1115	Suzuki et al. (2004)
16S-1541R	AAGGAGGTGATCCRGCCGCA	16S 1522–1541	Suzuki et al. (2000)
16S-1406F	TGYACACACCGCCCGT	16S 1391–1406	Lane (1991)
ITS-tRNAalaR	CTGCTTGCAAAGCAGGCGCTC	ITS-tRNA alanine	Suzuki et al. (2004)
ITS-tRNAalaF	GAGCGCCTGCTTTGCAAGCAG	ITS-tRNA alanine	Suzuki et al. (2004)
23S-139R	GCTGGGTTKTCTCATTCTCG	23S 121–139	Present study

tion of a tRNA-alanine in the spacer—measured by PCR with universal primers—is used to identify operational taxonomic units (OTUs). Clones representing all the OTUs were then sequenced.

Sequencing and phylogenetic analyses. For the vast majority of novel and unique ITS-LH-PCR fragment pairs, 16S rRNA genes of the clones were bidirectionally fully sequenced by the dideoxynucleotide termination reaction with 16S-targeting primers 27F, 1074F, 1100R and 1541R, ITS tRNA_{ala} primers and 23S primer 139R (Lane 1991, Suzuki et al. 2000, 2004; Table 1). Plasmids used as templates for sequencing were purified using either the Montage (Millipore), SprintPrep (Agencourt), or Fastplasmid (Eppendorf) kits following manufacturers' specifications. Sequencing reactions using Big Dye Chemistry v.3.1 (Applied Biosystems) were run in an AB3100 genetic analyzer with an 80 cm capillary and POP4 polymer, resulting in ca. 800 to 1000 bp per read. Chimeric sequences were screened by the 'CHIMERA DETECTION' program of the Ribosomal Database Project (Maidak et al. 1997) and removed from further analysis. Sequence alignments were constructed using the ARB_EDIT software (Ludwig et al. 2004), manually inspected and corrected based on the conserved secondary structure of 16S rRNA genes.

Preliminary placement of aligned sequences in a tree based on tree ssujun02.arb and containing ca. 29 000 sequences was performed by ARB_PARSIMONY using a universal mask, excluding positions where gaps outnumbered characters. Alignments were refined for specific clades, and distance matrices were calculated for near-complete and previously published sequences using clade-specific masks, excluding ambiguous positions and positions where gaps outnumbered characters. Phylogenetic distances were calculated using the Jukes & Cantor (1969) model in the PHYLIP package (Felsenstein 1989). This resulted in an analysis of 1421 characters for the SAR86 clade (Giovannoni & Rappé 2000) and related sequences, 1239 for the SAR11 clade (Giovannoni & Rappé 2000), 1143 for the SAR116 clade (Giovannoni & Rappé 2000), 1405 for the freshwater *Actinobacteria* Clade I (Warnecke et al. 2004), 1330 for the freshwater *Actinobacteria* Clades II plus III (Warnecke et al. 2004) and 1337 for the freshwater *Actinobacteria* Clade IV (Warnecke et al. 2004). For the *Roseobacter* clade, we included all but two of the reference sequences used in a review of this clade by Buchan et al. (2005), resulting in a slightly lower number (1196) of analyzed positions. Tree reconstruction was performed by neighbor-joining using the PHYLIP package. Initially trees were constructed for 5'- to 3'-end regions of the 16S rRNA for each of the clades, and putative chimeric genes from our, or previously published, libraries were removed from further

analysis based on the following criteria: (1) the sequences originated from a single clone library and had obviously distinct placement between 5'- and 3'-end-based trees and (2) the sequences were assigned as chimeric by Bellerophon analysis (Huber et al. 2004). Final full-length trees were constructed by neighbor-joining, and bootstrap analyses were performed based on 100 randomly re-sampled datasets. Short sequences obtained from our clone libraries were added to the preliminary trees by ARB_PARSIMONY.

Since the phylogenetic analysis defined several putative novel and unique clades, we performed BLAST searches against the GenBank September 2006 nucleotide database in order to verify the existence of short or unpublished sequences belonging to these so-called novel groups. Sequence positions between 356 and 906 of the *Escherichia coli* numbering system (commonly retrieved in surveys using denaturing gradient gel electrophoresis) of all full-length sequences belonging to each of the novel groups were used in the searches. Sequences retrieved by the BLAST searches and with 99 to 100% identity were added to phylogenetic trees by ARB_PARSIMONY, but, for the sake of clarity, these sequences were not included in the final trees.

Nucleotide sequence accession numbers. Sequences of Chesapeake Bay clones obtained in this study were deposited in the GenBank database under Accession Numbers EF471449 to EF471733.

RESULTS AND DISCUSSION

The phylogenetic diversity of bacterioplankton in estuarine systems has been the subject of several previous studies (Crump et al. 1999, Henriques et al. 2004, Hewson & Fuhrman 2004), but, as we expected, the diversity of bacterioplankton of a large estuary with long residence times, such as the Chesapeake Bay, was considerably different from that previously reported for smaller estuaries.

Table 2. Information on sampling and location of stations in the Chesapeake Bay. All stations were sampled at ~2 m depth

Library	Stn	Sampling date	Sampling time (h)	Coordinates
CB0	908	27 Sep 2002	12:05	39° 08' N, 76° 20' W
CB1	818	27 Sep 2002	17:50	38° 18' N, 76° 17' W
CB2	707	28 Sep 2002	16:10	37° 07' N, 76° 07' W
CB3	908	4 Mar 2003	11:15	39° 08' N, 76° 20' W
CB4	818	3 Mar 2003	15:20	38° 18' N, 76° 17' W
CB5	707	2 Mar 2003	10:30	37° 07' N, 76° 07' W

Alphaproteobacteria

Alphaproteobacteria was a major component in Chesapeake Bay bacterioplankton, and members of the SAR11, *Roseobacter* and *Rhodobacter* clades were the prevalent groups in the clone libraries. Most clones of *Alphaproteobacteria* were related to marine counterparts, but several novel clades were also retrieved that appear to be limited to coastal and estuarine systems.

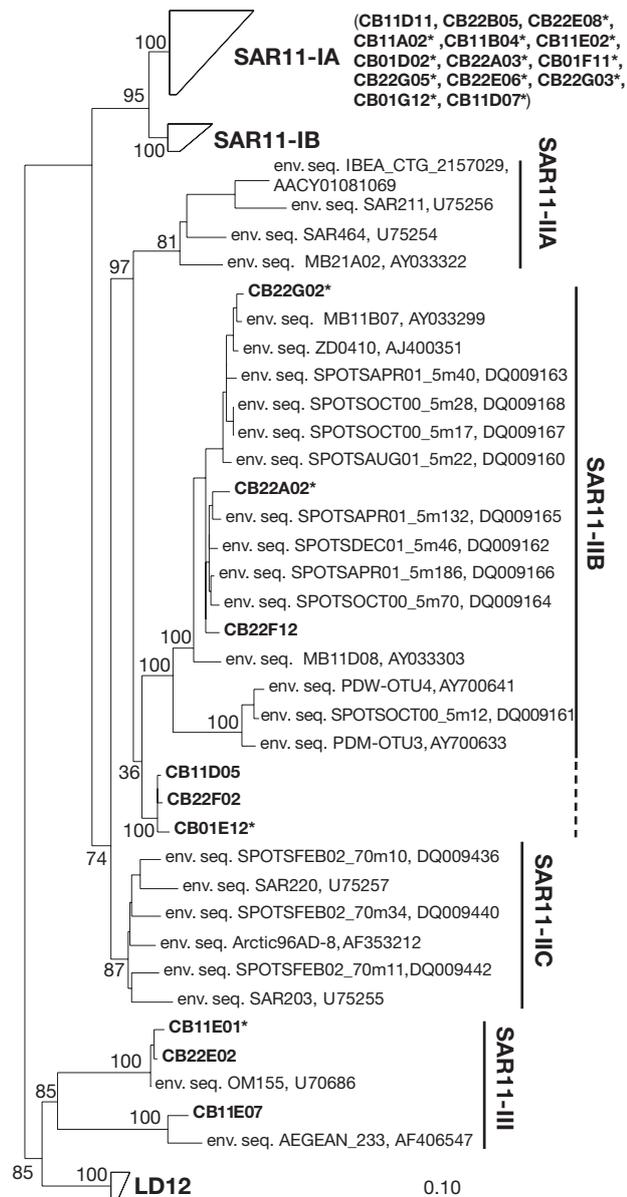


Fig. 2. Phylogenetic reconstruction (DNA distance and neighbor joining) of the SAR11 clade. Bootstrap values were based on 100 replicated trees. Clones in boldface were sequenced in the present study. *: short sequences added to the original tree by ARB_PARSIMONY. Scale bar indicates 10% estimated sequence divergence

Clones belonging to the SAR11 clade only appeared in the clone libraries sampled in September 2002, reflecting seasonal variation in microbial community structure (Kan et al. 2007). Four main subclades of the SAR11 clade, SAR11-I, -II, -III and LD12, were retrieved in the reconstructed phylogeny (Fig. 2). The phylogeny of SAR11-I and -II was in very good agreement with a previous classification scheme, supported by phylogeny of 16S rRNA gene, ITS and 23S rRNA gene phylogeny (Suzuki et al. 2001). Fourteen clones, mostly retrieved from the middle and southern bay, clustered with *Pelagibacter ubique* in SAR11-IA (Suzuki et al. 2001). SAR11-II showed a somewhat different picture than that previously described by Suzuki et al. (2004) (Fig. 2). SAR11-IIA contained mostly sequences from deep waters (80 to 500 m) in the Arctic (Bano & Hollibaugh 2002), along the southern California coast (Brown & Fuhrman 2005) and in Monterey Bay (Suzuki et al. 2004), but also 1 surface sample from the Sargasso Sea (Venter et al. 2004). SAR11-IIB contained primarily sequences retrieved from surface coastal samples, including Monterey Bay (Suzuki et al. 2004), the southern California coast (Brown & Fuhrman 2005), the North Sea (Zubkov et al. 2002) and the Chesapeake Bay. A subset of sequences, mostly from deeper water samples from the Sargasso Sea (Field et al. 1997), Arctic (Bano & Hollibaugh 2002) and the southern California coast (Brown & Fuhrman 2005) formed a third major subclade (SAR11-IIC), and the high support for this previously undefined clade was likely due to the removal of putatively chimeric sequences. Finally, 3 additional Chesapeake Bay clones were loosely affiliated with Clade SAR11-IIB and might represent a novel subclade of SAR11-II (Fig. 2). LD12 formed a monophyletic group supported by high bootstrap values in agreement with the original description of the group (Zwart et al. 2002). Our reconstruction showed that clones in this group are almost exclusively of freshwater (mostly lacustrine) origin, and BLAST searches indicated that only a single sequence belonging to the group has been retrieved from a mesohaline environment (AY145598, the Weser River turbidity maximum; Selje et al. 2005). A newly proposed group, SAR11-III, was separated from SAR11-I, SAR11-II and LD12 with high bootstrap support (Fig. 2). Subclade SAR11-III appeared to represent organisms mostly inhabiting coastal ecosystems and included 1 clone from the continental shelf off Cape Hatteras, North Carolina (Rappé et al. 1997), 1 clone from the Aegean Sea (Moeseneder et al. 2005) and 3 clones from Chesapeake Bay (Fig. 2). BLAST searches corroborated this coastal and estuarine distribution of SAR11-III as top hits included short and unpublished sequences from coastal environments such as the Baltic Sea, i.e. AY317115 (Simu & Hagstrom 2004), AF388881 (Kisand & Wikner 2003),

the Plum Island estuary, i.e. AY580584-5 (Acinas et al. 2004) and a salt marsh creek, i.e. DQ421658 (Mou et al. 2007). Interestingly, 3 clones belonging to this clade were retrieved from the high altitude (3203 m), polyhaline (high salinity) Lake Qinghai (AM182276, AM182277, AM182278; Wu et al. 2006). Thus, SAR11-III appears to be a novel clade with high bootstrap values separating it from other members of the marine SAR11 clade, and the sequences in this group appeared predominantly in coastal and estuarine environments, or high altitude saline lakes, but not in freshwater or the open ocean.

The *Roseobacter* clade represents one of the 9 major clades of marine bacteria (Gonzalez & Moran 1997, Giovannoni & Rappé 2000). Typically members of this group comprise up to 15–20% of the ocean and estuarine bacterial communities (Gonzalez & Moran 1997, Giovannoni & Rappé 2000) and represent a high within-group diversity (Buchan et al. 2005). In the March 2003 libraries, the *Roseobacter* clade clones represented 35.3 to 39.1% of the total clones (Kan et al. 2007) and showed a high diversity of phylotypes. The phylogenetic reconstruction of the *Roseobacter* clade was consistent with that performed by Buchan et al. (2005). High bootstrap values supported 5 novel *Roseobacter* lineages that were also resolved by phylogenetic reconstructions using parsimony and maximum likelihood (Kan 2006). These novel lineages were named clades ChesI to ChesV (Fig. 3). Sequences in the ChesI and ChesV clades were only retrieved in September 2002. The ChesI clade was further divided based on 16S rRNA gene reconstruction and ITS-LH-PCR fragment sizes, and also corroborated by ITS phylogeny (M. T. Suzuki & A. Buchan unpubl. results). ChesI-A contained clones from Chesapeake Bay, the Sargasso Sea metagenomic database (Venter et al. 2004), the North Pacific Subtropical Gyre metagenomic database (DeLong et al. 2006), and the southern California coast (Brown et al. 2005), indicating widespread distribution, much like groups ChesI-B and ChesI-D (Fig. 3). However, remarkably, very few other sequences (including short or unpublished ones) belonging to these 3 groups had been previously retrieved in earlier surveys of marine bacterioplankton 16S rRNA genes. ChesI-C contained exclusively clones recovered from the Chesapeake Bay, but BLAST searches retrieved short sequences from other coastal and estuarine systems, including a salt marsh creek (i.e. DQ421658; Mou et al. 2007) and waters adjacent to the Mobile River plume (i.e. AY904490; Pinhassi et al. 2005). Two Chesapeake Bay clones retrieved in September 2002 formed the novel ChesV clade with no other related sequences.

Chesapeake *Roseobacter* Clades ChesII, ChesIII and ChesIV were uniquely retrieved in the Chesapeake

Bay and, furthermore, only in March 2003 samples, initially suggesting that they could represent groups endemic to the Chesapeake Bay under colder conditions. Interestingly, no sequences from other estuaries (Crump et al. 1999, Sekiguchi et al. 2002) belong to Chesapeake *Roseobacter* Clades II to V, suggesting they may represent habitat-specific populations adapted to the bay or other large estuaries with long residence times. However, BLAST searches indicated that an anaerobic anoxygenic phototrophic strain (BS110, DQ659412) isolated from 62 m depth in the Bosphorus Strait (Oz et al. 2005) is a member of the ChesII clade. Finally, clones from the Chesapeake Bay were also associated with the previously described *Roseobacter* lineages AS26 and DG1128 (Buchan et al. 2005), GAI-37 (Gonzalez & Moran 1997) and with strains isolated from the Arctic (strain ARK9990; Brinkmeyer et al. 2003) or the North Atlantic continental slope (slope strain DI4; Teske et al. 2000), indicating a more ubiquitous distribution of these clades.

The phylogenetic reconstruction of *Rhodospirillales* was in good agreement with a previous description (Suzuki et al. 2001). One clone (CB22G09) from the Chesapeake Bay and 1 clone from the Sargasso Sea (Venter et al. 2004) fell in Subclade III. In addition, 2 clones (CB22C04 and CB22D08), along with 27 closely related unpublished clones associated with the marine sponge *Halichondria okadai* (i.e. AB054135; I. Okano et al. unpubl. data), formed a novel group we called SAR116-IV (data not shown). Clone CB22H11 fell in a novel *Rhodospirillales* clade we called SPOT-SAUG01_5m94, along with sequences from coastal southern California (Brown & Fuhrman 2005), the Sargasso Sea metagenomic database (Venter et al. 2004) and the Aegean Sea (Moeseneder et al. 2005).

Members of the *Rhodobacter* clade represented a significant fraction of clones in the March 2003 libraries in the northern and mid-bay (Kan et al. 2007). The vast majority of these clones were closely related to *Pseudorhodobacter ferrugineum*, a bacterium with psychrophilic features, isolated from sediments of the northeastern Atlantic Ocean (Ruger & Höfle 1992).

***Gammaproteobacteria* and SAR86**

High diversity was observed for Chesapeake Bay *Gammaproteobacteria*. This group was represented by many unique clones that were closely related to described strains or belonged to previously defined clades (Suzuki et al. 2004). The ubiquitous marine bacterioplankton group SAR86 (Giovannoni & Rappé 2000) represented the most abundant *Gammaproteobacteria* group, particularly in September 2002 (Kan et al. 2007). The phylogenetic reconstruction of

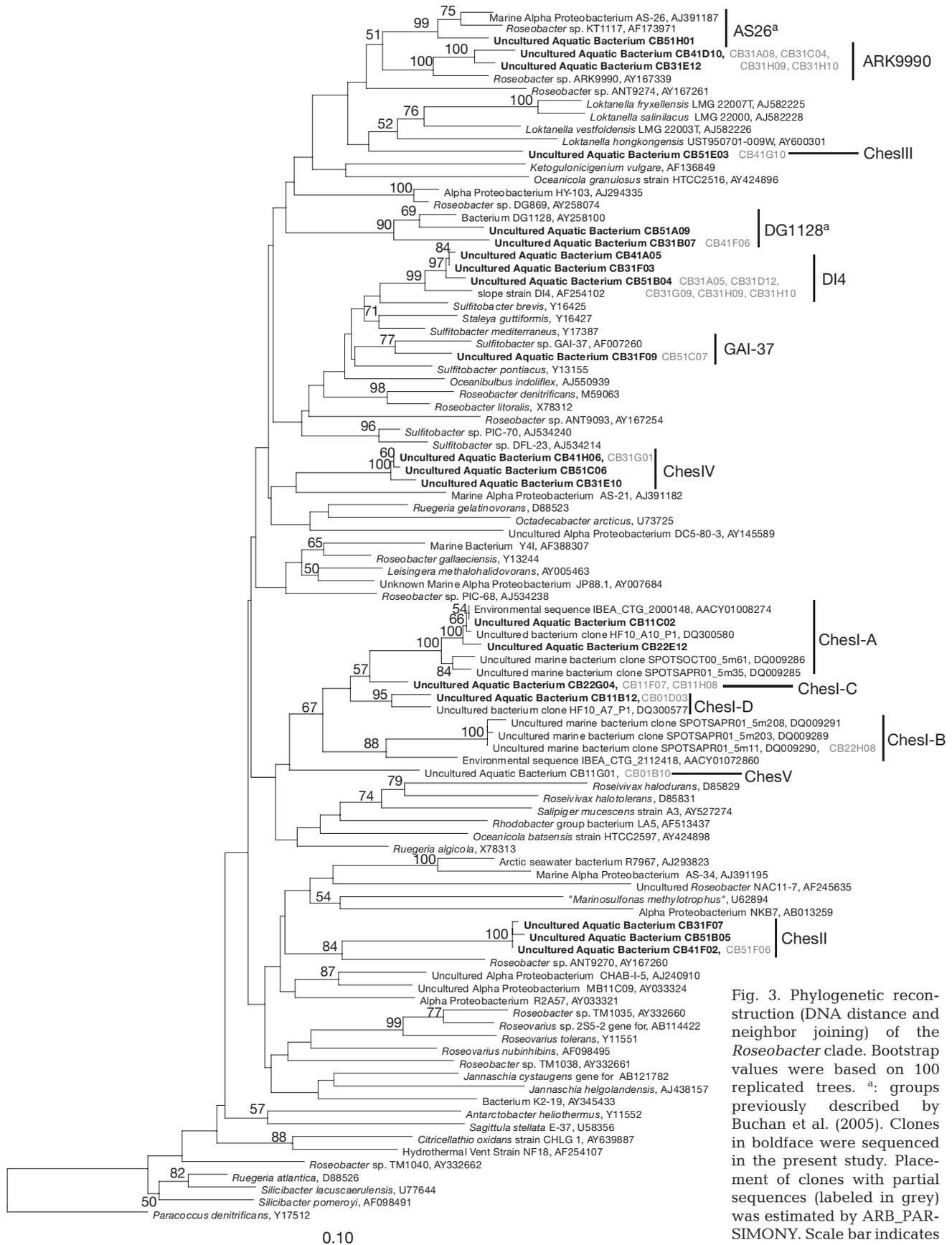


Fig. 3. Phylogenetic reconstruction (DNA distance and neighbor joining) of the *Roseobacter* clade. Bootstrap values were based on 100 replicated trees. ^a: groups previously described by Buchan et al. (2005). Clones in boldface were sequenced in the present study. Placement of clones with partial sequences (labeled in grey) was estimated by ARB_PAR-SIMONY. Scale bar indicates 10% estimated sequence divergence

this group and of closely related clades was similar to what was previously described (Suzuki et al. 2001), but in addition, 3 Chesapeake Bay clones and sequences from the Sargasso Sea (IBEA_CTG_2134086; Venter et al. 2004), Arabian Sea (A313002; Fuchs et al. 2005) and coral mucus (PDM-OTU2; Bourne & Munn 2005) formed a unique novel monophyletic group (SAR86-IV) that could further be split into 2 distinct subclades (Fig. 4). Subclade SAR86-IVA might have an estuarine distribution, as BLAST searches indicated that only 3 very closely related partial sequences from a Georgia salt marsh belong to this clade (i.e. DQ421675; Mou et al. 2007). Subclade SAR86-IVB appeared to have oceanic distribution, and included unpublished partial sequences from the East China Sea (AY663955; N. Jiao et al. unpubl. data), the West Pacific Gyre (AY664118; N. Jiao et al. unpubl. data) as well as partial sequences from 10 and 70 m at Stn ALOHA in the Subtropical Pacific gyre (DQ300651 and DQ300869; DeLong et al. 2006). Finally, Clone CB22H04 belonged to yet another novel clade that we called RedeBAC7D11 (Fig. 4), related to SAR156 (Suzuki et al. 2004) and containing environmental sequences from the Red Sea, Sargasso Sea (Venter et al. 2004) and the southern California coast (Brown & Fuhrman 2005).

Actinobacteria

Clones affiliated with *Actinobacteria* showed a remarkable seasonal variation among the libraries (Kan et al. 2007). Clones associated with the ubiquitous 'marine *Actinobacteria*' clade (Giovannoni & Rappé 2000) were prevalent in the September 2002 libraries. Typical 'freshwater *Actinobacteria*' (Warnecke et al. 2004) were retrieved in all 6 clone libraries, but were more abundant within clone libraries in March 2003 than in September 2002 (Kan et al. 2007). The preliminary placement of *Actinobacteria* clones by ARB-PARSIMONY showed that, except for marine *Actinobacteria*, clones belonged to 4 previously defined distinct phylogenetic clades of freshwater *Actinobacteria* (acI, II, III and IV; Warnecke et al. 2004) (Fig. 5).

Clade acI

Only 3 clones retrieved from the northern and middle bay were affiliated with the previously described Subclades acI-B and acI-C, and we did not retrieve any clones in Subclade acI-A (Fig. 5A). This was not particularly surprising, as these clades are overwhelmingly represented by sequences from freshwater environments (Warnecke et al. 2004), and thus were likely

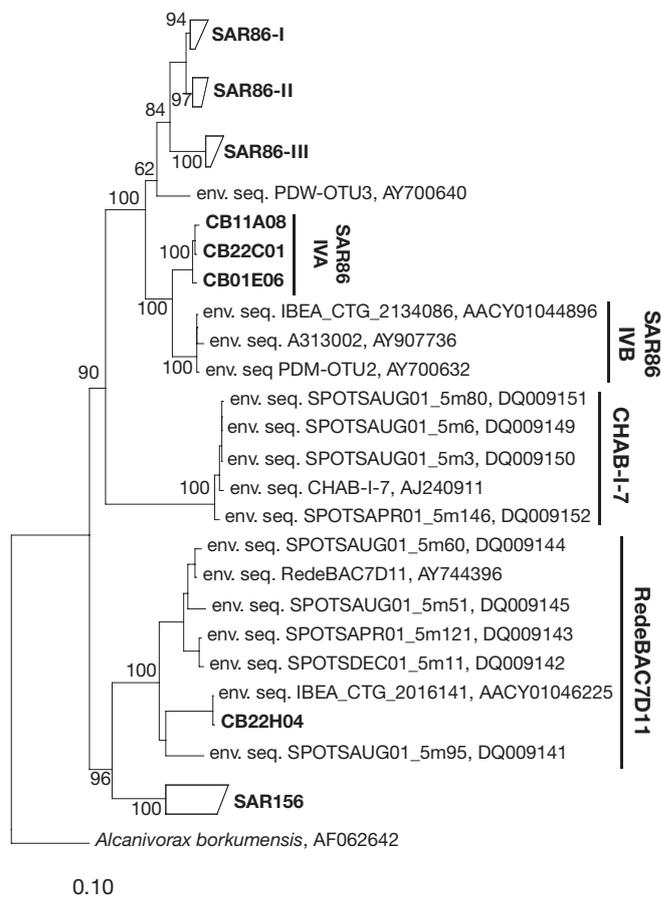


Fig. 4. Phylogenetic reconstruction (DNA distance and neighbor joining) of the SAR86 clade. Bootstrap values were based on 100 replicated trees. Clones in boldface were sequenced in the present study. Scale bar indicates 10% estimated sequence divergence

transported into the bay by riverine input. Interestingly, several sequences from Chesapeake Bay and a variety of environments formed a novel clade with high bootstrap support, which we called acI-D. Several sequences belonging to this clade originated from other saline samples such as Gulf of Alaska octocorals (DQ396268 and DQ396342; Penn et al. 2006), and Lake Bonney (DQ015810; Glatz et al. 2006). Furthermore, the fact that sequences in acI-D were retrieved in all but one Chesapeake Bay library (CB2) suggests that this clade might be adapted to estuarine conditions and not just transported into the bay with freshwater input.

Clades acII and acIII

The phylogenetic reconstruction of *Actinobacteria* Clades acII and acIII was similar to that by Warnecke et

al. (2004), although bootstrap analyses showed little support for the monophyly of Clade acII. Subclade acII-D was separated from Clades acII-A, acII-B and acII-C, and the grouping of the latter 3 clades had very high bootstrap support (Fig. 5B). Clade acII-B had low bootstrap support and might in fact be polyphyletic. Chesapeake Bay clones fell into Subclades acII-A and acII-B and Clade acIII (Fig. 5B). Similarly to the acI clade, freshwater Clade acII are relatively rare in rRNA gene libraries from coastal and open ocean environments, and so the fact that some members of these groups were found at salinities as high as 27 ppt in Chesapeake Bay was quite remarkable. Clones in Clade acII-B appear to be of riverine origin, since they were only retrieved in the upper bay in March 2003, and since close relatives of these clones were also mainly retrieved from freshwater habitats, i.e. AJ565417 (Hahn et al. 2004), AJ575510 and AJ575514 (Warnecke et al. 2004). Clones in Clade acII-A had a wider distribution in the libraries, but they were not retrieved in the southern bay in either season. Other sequences in acII-A were previously retrieved from the saline lakes Sælenvannet (e.g. AJ575515, AJ575519, AJ575520, AJ575521 and AJ575522; Warnecke et al. 2004) and Bonney (e.g. DQ015794, DQ015823 and DQ015847; Glatz et al. 2006), and the Oregon coast (DQ372844; Morris et al. 2006), indicating a possible adaptation of acII-A to estuarine environments (Fig. 5B). Similarly, clones in the acIII clade were retrieved in all Chesapeake Bay libraries, and other sequences in this group were also retrieved from saline — Lake Sælenvannet: AJ575518, AJ575523, AJ575524, AJ575525, AJ575526, AJ575527 (Warnecke et al. 2004); Lake Kauhako: AY344420 (Donachie et al. 2004); Lake Bonney: DQ015775, DQ015786, DQ015813 (Glatz et al. 2006) — or hypersaline — Mono lake: e.g. AF445364, AF447767, AF447768, AF448168, AF448170, AF448171, AF448177, AF448180 (Humayoun et al. 2003) — environments. Finally, BLAST search results corroborate the association of acIIB and acIII with saline environments, as short sequences from the Plum Island estuary (AY580343, AY580353 and AY580355; Acinas et al. 2004) belong to acIIB and clones from the Ría de Aveiro estuary (AY499436; Henriques et al. 2004) and the saline Lake Qinghai (AM182283 and AM182284; Wu et al. 2006) belong to the acIII clade.

Clade acIV

The *Actinobacteria* Clade acIV is predominantly constituted of sequences from freshwater and estuar-

ine environments (Zwart et al. 2002). In our phylogeny reconstruction of this clade, Subclades acIV-A and acIV-B were in agreement with the phylogeny proposed by Warnecke et al. (2004), and both clades were predominantly represented by sequences retrieved from freshwater environments (Fig. 5C). Clone CB11A12 was basal to acIV-A and -B, and might represent a novel clade, but currently no other near full-length sequence is closely related to this clone. In addition, 3 Chesapeake Bay clones, including 2 fully sequenced 16S rRNA genes formed a novel clade (acIV-C; Fig. 5C) related to sequences from a variety of saline environments, including the hypersaline Mono Lake (AF454303; Humayoun et al. 2003), the saline lake Kauhako (AY344421; Donachie et al. 2004) and Arctic pack ice (AF468297; Brinkmeyer et al. 2003). However, sequences retrieved from soil and fresh-water were also in this clade, and thus the association of acIV-C to high-salinity conditions is tenuous. Finally, 2 additional clones (CB01C05 and CB11H06) formed a second novel subclade that was named acIV-D (Fig. 5C). Currently the only near full-length sequence closely associated with these clones is a eutrophic lake sequence (DQ520164; Wu et al. 2007).

Other bacterial groups

Betaproteobacteria clones were more commonly retrieved in the March 2003 than in the September 2002 libraries (Kan et al. 2007), which is consistent with previous studies in the freshwater region of the bay (Bouvier & del Giorgio 2002). Subgroup GKS98, typical of freshwater environments (Zwart et al. 2002) was the most frequently retrieved *Betaproteobacteria* group and only occurred in Libraries CB3 and CB4, suggesting riverine origin. *Deltaproteobacteria* were also found in the Chesapeake Bay and contained clones closely related to the SAR324 clade (Wright et al. 1997). Unique and diverse Marine Cluster B *Synechococcus* were found in Chesapeake Bay and a detailed phylogeny based on 16S rRNA gene and the ITS and gene sequences have been discussed elsewhere (Chen et al. 2006).

High diversity was observed for the *Bacteroidetes*, and clone sequences in this group were members of many distinct phylogenetic groups (Kan et al. 2007; phylogeny not shown). In the warm season (i.e. Libraries CB0, CB1 and CB2), the 29 *Bacteroidetes* clones retrieved were closely related to the *Flavobac-*

Fig. 5. Phylogenetic reconstruction (DNA distance and neighbor joining) of *Actinobacteria*: (A) Clade acI, (B) Clades acII and acIII, and (C) Clade acIV. Bootstrap values were based on 100 replicated trees. Clones in boldface were sequenced in the present study. *: short sequences added to the original tree by ARB_PARSIMONY. Scale bar indicates 10% estimated sequence divergence

teriaceae Clone UC1 (AY080916; Radajewski et al. 2002), *Fluviicola taffensis* RW262 (AF493694; O'Sullivan et al. 2005), Clone OM273 (U70709; Rappé et al. 1997), the AGG58 cluster (AY354892; O'Sullivan et al. 2004), or were either unclassified or unidentified clones unique to the Chesapeake Bay. In contrast, 9 clones associated with Antarctic bacterium R-9033 (AJ441001; Van Trappen et al. 2002), ATAM173_A3 (AF359540; Hold et al. 2001), *Cellulophaga* sp. (AY274838; Kirchman et al. 2003), Antarctic bacterium R-9286 (AJ441012; Van Trappen et al. 2002) and TM18_28 (DQ279364; Barbieri et al. 2007) were only present in Libraries CB3, CB4 and CB5. It should be pointed out that a previous comparison between 16S rRNA diversity in a PCR-based clone library to that in a metagenomic library from the same water sample indicated that the *Bacteroidetes* phylum is likely underrepresented in PCR-based clone libraries (Cottrell et al. 2005); therefore, our study likely provides a somewhat incomplete snapshot of the diversity and phylogeny of *Bacteroidetes* in the Chesapeake Bay.

CONCLUSIONS

Our results indicated that the Chesapeake Bay contains bacteria likely originating from freshwater and also from the adjacent coastal ocean. On the other hand, several previously undescribed groups were also found in the bay. These 'novel' subclades were found in the SAR11, *Roseobacter* and SAR86, and *Actinobacteria* clades, some of which could possibly be adapted to the estuarine conditions and might be indigenous to the bay or other large temperate estuaries. Recently, a thorough metagenomic study covering the transect from the North Atlantic towards the South Pacific revealed more bacterial diversity and a more complex ocean than might have been thought (Rusch et al. 2007). With 7.7 million sequencing reads (>6 billion base pairs of sequences), in no case was it possible to assemble a full genome from ubiquitously abundant groups (including *Synechococcus*, *Prochlorococcus* and *Pelagibacter ubique*), indicating that a remarkably high diversity exists in aquatic environments. Keeping that in mind, the present study provides a starting point for exploration of the bacterioplankton diversity in the Chesapeake Bay and other large estuaries.

Acknowledgements. We thank the crews of RV 'Cape Henlopen' for sample collections. We also acknowledge the funding support from the National Science Foundation, Microbial Observatories Program (MCB-0132070, MCB-0238515 and MCB-0537041) to F.C. and Biological Oceanography (OCE-0550547) to M.T.S.

LITERATURE CITED

- Acinas SG, Klepac-Ceraj V, Hunt DE, Pharino C, Ceraj I, Distel DL, Polz MF (2004) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 430: 551–554
- Bano N, Hollibaugh JT (2002) Phylogenetic composition of bacterioplankton assemblages from the Arctic Ocean. *Appl Environ Microbiol* 68:505–518
- Barbieri E, Guidi C, Bertaux J, Frey-Klett P and others (2007) Occurrence and diversity of bacterial communities in *Tuber magnatum* during truffle maturation. *Environ Microbiol* 9:2234–2246
- Bidle KD, Fletcher M (1995) Comparison of free-living and particle-associated bacterial communities in the Chesapeake Bay by stable low-molecular-weight RNA analysis. *Appl Environ Microbiol* 61:944–952
- Bourne DG, Munn CB (2005) Diversity of bacteria associated with the coral *Pocillopora damicornis* from the Great Barrier Reef. *Environ Microbiol* 7:1162–1174
- Bouvier TC, del Giorgio PA (2002) Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. *Limnol Oceanogr* 47:453–470
- Brinkmeyer R, Knittel K, Jurgens J, Weyland H, Amann R, Helmke E (2003) Diversity and structure of bacterial communities in arctic versus antarctic pack ice. *Appl Environ Microbiol* 69:6610–6619
- Brown MV, Fuhrman JA (2005) Marine bacterial microdiversity as revealed by internal transcribed spacer analysis. *Aquat Microb Ecol* 41:15–23
- Brown MV, Schwalbach MS, Hewson I, Fuhrman JA (2005) Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series. *Environ Microbiol* 7:1466–1479
- Buchan A, Gonzalez JM, Moran MA (2005) Overview of the marine *Roseobacter* lineage. *Appl Environ Microbiol* 71: 5665–5677
- Chen F, Wang K, Kan J, Suzuki MT, Wommack KE (2006) Diverse and unique picocyanobacteria in Chesapeake Bay, revealed by 16S–23S rRNA internal transcribed spacer sequences. *Appl Environ Microbiol* 72:2239–2243
- Cottrell MT, Waidner LA, Yu L, Kirchman DL (2005) Bacterial diversity of metagenomic and PCR libraries from the Delaware River. *Environ Microbiol* 7:1883–1895
- Crump BC, Armbrust EV, Baross JA (1999) Phylogenetic analysis of particle-attached and free-living bacterial communities in the Columbia River, its estuary, and the adjacent coastal ocean. *Appl Environ Microbiol* 65:3192–3204
- DeLong EF, Franks DG, Alldredge AL (1993) Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol Oceanogr* 38:924–934
- DeLong EF, Preston CM, Mincer T, Rich V and others (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311:496–503
- Donachie SP, Hou S, Lee KS, Riley CW and others (2004) The Hawaiian Archipelago: a microbial diversity hotspot. *Microb Ecol* 48:509–520
- Felsenstein J (1989) PHYLIP-phylogeny inference package (v3.5). *Cladistics* 5:164–166
- Field KG, Gordon D, Wright T, Rappe M, Urbach E, Vergin K, Giovannoni SJ (1997) Diversity and depth-specific distribution of SAR11 cluster rRNA genes from marine planktonic bacteria. *Appl Environ Microbiol* 63:63–70
- Fuchs BM, Woebken D, Zubkov MV, Burkill P, Amann R (2005) Molecular identification of picoplankton popula-

- tions in contrasting waters of the Arabian Sea. *Aquat Microb Ecol* 39:145–157
- Giovannoni SJ, Rappé M (2000) Evolution, diversity and molecular ecology of marine prokaryotes. In: Kirchman D (ed) *Microbial ecology of the oceans*. John Wiley & Sons, New York, p 47–84
- Glatz RE, Lepp PW, Ward BB, Francis CA (2006) Planktonic microbial community composition across steep physical/chemical gradients in permanently ice-covered Lake Bonney, Antarctica. *Geobiology* 4:53–67
- Gonzalez JM, Moran MA (1997) Numerical dominance of a group of marine bacteria in the alpha-subclass of the class Proteobacteria in coastal seawater. *Appl Environ Microbiol* 63:4237–4242
- Hahn MW, Stadler P, Wu QL, Pockl M (2004) The filtration-acclimatization method for isolation of an important fraction of the not readily cultivable bacteria. *J Microbiol Methods* 57:379–390
- Henriques I, Almeida MA, Cunha MA, Correia A (2004) Molecular sequence analysis of prokaryotic diversity in the middle and outer sections of the Portuguese estuary Ria de Aveiro. *FEMS Microbiol Ecol* 49:269–279
- Hewson I, Fuhrman JA (2004) Richness and diversity of bacterioplankton species along an estuarine gradient in Moreton Bay, Australia. *Appl Environ Microbiol* 70:3425–3433
- Hold GJ, Smith EA, Rappe MS, Maas EW and others (2001) Characterization of bacterial communities associated with toxic and non-toxic dinoflagellates: *Alexandrium* spp. and *Scrippsiella trochoidea*. *FEMS Microbiol Ecol* 37:161–173
- Huber T, Faulkner G, Hugenholtz P (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* 20:2317–2319
- Humayoun SB, Bano N, Hollibaugh JT (2003) Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California. *Appl Environ Microbiol* 69:1030–1042
- Jukes TH, Cantor CR (1969) Evolution of protein. In: Munro HN (ed) *Mammalian protein metabolism*. Academic Press, New York, p 21–132
- Kan J (2006) Bacterioplankton in the Chesapeake Bay: genetic diversity, population dynamics and community proteomics. PhD thesis, University of Maryland, College Park, MD (available at <https://drum.umd.edu/dspace/handle/1903/4081>)
- Kan J, Suzuki MT, Wang K, Evans SE, Chen F (2007) High temporal but low spatial heterogeneity of bacterioplankton in the Chesapeake Bay. *Appl Environ Microbiol* 73:6776–6789
- Kirchman DL, Yu LY, Cottrell MT (2003) Diversity and abundance of uncultured *Cytophaga*-like bacteria in the Delaware Estuary. *Appl Environ Microbiol* 69:6587–6596
- Kisand V, Wikner J (2003) Combining culture-dependent and -independent methodologies for estimation of richness of estuarine bacterioplankton consuming riverine dissolved organic matter. *Appl Environ Microbiol* 69:3607–3616
- Lane DJ (1991) 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M (eds) *Nucleic acid techniques in bacterial systematics*. John Wiley & Sons, New York, p 115–175
- Ludwig W, Strunk O, Westram R, Richter L and others (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* 32:1363–1371
- Maidak BL, Olsen GJ, Larsen N, Overbeek R, McCaughey MJ, Woese CR (1997) The RDP (Ribosomal Database Project). *Nucleic Acids Res* 25:109–111
- Moeseneder MM, Arrieta JM, Herndl GJ (2005) A comparison of DNA- and RNA-based clone libraries from the same marine bacterioplankton community. *FEMS Microbiol Ecol* 51:341–352
- Morris RM, Longnecker K, Giovannoni SJ (2006) Pirellula and OM43 are among the dominant lineages identified in an Oregon coast diatom bloom. *Environ Microbiol* 8:1361–1370
- Mou X, Hodson RE, Moran MA (2007) Bacterioplankton assemblages transforming dissolved organic compounds in coastal seawater. *Environ Microbiol* 9:2025–2037
- Nixon SW, Ammerman JW, Atkinson LP, Berounsky VM and others (1996) The fate of nitrogen and phosphorus at the land sea margin of the North Atlantic Ocean. *Biogeochemistry* 35:141–180
- O'Sullivan LA, Fuller KE, Thomas EM, Turley CM, Fry JC, Weightman AJ (2004) Distribution and culturability of the uncultivated 'AGG58 cluster' of the Bacteroidetes phylum in aquatic environments. *FEMS Microbiol Ecol* 47:359–370
- O'Sullivan LA, Rinna J, Humphreys G, Weightman AJ, Fry JC (2005) *Fluviicola taffensis* gen. nov., sp. nov., a novel freshwater bacterium of the family Cryomorphaceae in the phylum 'Bacteroidetes'. *Int J Syst Evol Microbiol* 55:2189–2194
- Ozz A, Sabehi G, Koblížek M, Massana R, Oded B (2005) *Roseobacter*-like bacteria in Red and Mediterranean Sea aerobic anoxygenic photosynthetic populations. *Appl Environ Microbiol* 71:344–353
- Penn K, Wu D, Eisen JA, Ward N (2006) Characterization of bacterial communities associated with deep-sea corals on Gulf of Alaska seamounts. *Appl Environ Microbiol* 72:1680–1683
- Pinhassi J, Simo R, Gonzalez JM, Vila M and others (2005) Dimethylsulfoniopropionate turnover is linked to the composition and dynamics of the bacterioplankton assemblage during a microcosm phytoplankton bloom. *Appl Environ Microbiol* 71:7650–7660
- Radajewski S, Webster G, Reay DS, Morris SA and others (2002) Identification of active methylotroph populations in an acidic forest soil by stable-isotope probing. *Microbiology* 148:2331–2342
- Rappé MS, Kemp PF, Giovannoni SJ (1997) Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shelf off Cape Hatteras, North Carolina. *Limnol Oceanogr* 42:811–826
- Ruger HJ, Höfle MG (1992) Marine star-shaped-aggregate-forming bacteria: *Agrobacterium atlanticum* sp. nov.; *Agrobacterium meteori* sp. nov.; *Agrobacterium ferrugineum* sp. nov., nom. rev.; *Agrobacterium gelatinovorum* sp. nov., nom. rev.; and *Agrobacterium stellulatum* sp. nov., nom. rev. *Int J Syst Bacteriol* 42:133–143
- Rusch DB, Halpern AL, Sutton G, Heidelberg KB and others (2007) The Sorcerer II global ocean sampling expedition: Northwest Atlantic through eastern tropical Pacific. *PLoS Biol* 5:0398–0431
- Sekiguchi H, Watanabe M, Nakahara T, Xu B, Uchiyama H (2002) Succession of bacterial community structure along the Changjiang River determined by denaturing gradient gel electrophoresis and clone library analysis. *Appl Environ Microbiol* 68:5142–5150
- Selje N, Brinkhoff T, Simon M (2005) Detection of abundant bacteria in the Weser estuary using culture-dependent and culture-independent approaches. *Aquat Microb Ecol* 39:17–34
- Simu K, Hagström Å (2004) Oligotrophic bacterioplankton with a novel single-cell life strategy. *Appl Environ Microbiol* 70:2445–2451
- Smith DE, Leffler M, Mackiernan G (1992) Oxygen dynamics

- in the Chesapeake Bay—a synthesis of recent research. Maryland and Virginia Sea Grant College Programs, College Park, MD
- Suzuki MT, Taylor LT, DeLong EF (2000) Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol* 66: 4605–4614
- Suzuki MT, Beja O, Taylor LT, DeLong EF (2001) Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton. *Environ Microbiol* 3: 323–331
- Suzuki MT, Preston CM, Beja O, de la Torre JR, Steward GF, DeLong EF (2004) Phylogenetic screening of ribosomal RNA gene-containing clones in bacterial artificial chromosome (BAC) libraries from different depths in Monterey Bay. *Microb Ecol* 48:473–488
- Teske A, Brinkhoff T, Muyzer G, Moser DP, Rethmeier J, Jannasch HW (2000) Diversity of thiosulfate-oxidizing bacteria from marine sediments and hydrothermal vents. *Appl Environ Microbiol* 66:3125–3133
- Van Trappen S, Mergaert J, Eygen SV, Dawyndt P, Cnockaert MC, Swings J (2002) Diversity of 746 heterotrophic bacteria isolated from microbial mats from ten Antarctic lakes. *Syst Appl Microbiol* 25:603–610
- Venter JC, Remington K, Heidelberg JF, Halpern AL and others (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304:66–74
- Warnecke F, Amann R, Pernthaler J (2004) Actinobacterial 16S rRNA genes from freshwater habitats cluster in four distinct lineages. *Environ Microbiol* 6:242–253
- Wright TD, Vergin KL, Boyd PW, Giovannoni SJ (1997) A novel delta-subdivision proteobacterial lineage from the lower ocean surface layer. *Appl Environ Microbiol* 63: 1441–1448
- Wu QL, Zwart G, Schauer M, Agterveld MPKv, Hahn MW (2006) Bacterioplankton community composition along a salinity gradient of sixteen high-mountain lakes located on the Tibetan Plateau, China. *Appl Environ Microbiol* 72: 5478–5485
- Zubkov MV, Fuchs BM, Archer SD, Kiene RP, Amann R, Burkill PH (2002) A population of the alpha-proteobacteria dominate the bacterioplankton and dimethylsulphoniopropionate uptake after an algal bloom in the North Sea. *Deep-Sea Res II* 49:3017–3038
- Zwart G, Crump BC, Kamst-van Agterveld MP, Hagen F, Han SK (2002) Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquat Microb Ecol* 28:141–155

*Editorial responsibility: Dittmar Hahn,
San Marcos, Texas, USA*

*Submitted: October 3, 2007; Accepted: January 21, 2008
Proofs received from author(s): April 1, 2008*