

# Evidence for tropical endemicity in the Deltaproteobacteria Marine Group B/SAR324 bacterioplankton clade

Mark V. Brown<sup>1,2,\*</sup>, Stuart P. Donachie<sup>2</sup>

<sup>1</sup>NASA Astrobiology Institute, PSB211, 2565 McCarthy Mall, and <sup>2</sup>Department of Microbiology, Snyder Hall, 2538 McCarthy Mall, University of Hawaii, Honolulu, Hawaii 96822, USA

**ABSTRACT:** The fine-scale phylogeny of the deeply branching Deltaproteobacteria Marine Group B/SAR324 (MGB/SAR324) bacterioplankton clade was analyzed using partial 16S rRNA gene sequence and internal transcribed spacer (ITS) sequences. Both sets of analyses revealed considerable sequence variation, which, along with bootstrap calculations, strongly support the discrimination of 3 'species' level clusters (i.e. displaying <98% within-cluster and >97% between-cluster 16S rRNA gene sequence identity) within this clade, which are here designated as MGB/SAR324 clade groups I and II and the MGB/SAR276 clade. The biogeographical ranges of these 3 clades were analyzed using newly acquired 16S and ITS sequences from Stn ALOHA (22.45° N, 158° W), along with sequences available in the public domain. MGB/SAR324 clade group I and the MGB/SAR276 clade display a clearly restricted distribution, occurring only in tropical and subtropical waters at depths <250 m. These groups represent the first reported species- and genus-level phylotypes displaying a 'tropical-only' global distribution. MGB/SAR324 clade group II, which has been previously described as ubiquitous, occurs across the latitudinal gradient, but is generally restricted to deep, colder waters in the lower latitudes. Enzyme sequences associated at least with MGB/SAR324 clade group II appear to indicate a role for this group in dissolved organic phosphate cycling.

**KEY WORDS:** Marine Group B · SAR324 · Biogeography · Internal transcribed spacer · ITS

*Resale or republication not permitted without written consent of the publisher*

## INTRODUCTION

Molecular markers, such as 16S rRNA genes, offer a variety of levels of discrimination, and are beginning to provide evidence for defined global distribution patterns within marine microbial communities. The small cell size and large populations of microorganisms should be expected to facilitate widespread distribution of organisms throughout the marine ecosystem. Given this premise, the presence of an organism at detectable levels in a given environment will be determined by its ability to reproduce within the regime of the environment in question (Dolan 2005), i.e. 'Everything is everywhere (Beijerinck 1913), the environment selects' (Baas-Becking 1934). The global marine environment is far from homogenous in its physical,

chemical, and biological make up, and mounting evidence suggests that whole microbial community composition differs significantly across temporal (e.g. Morris et al. 2005), latitudinal (e.g. Baldwin et al. 2005), and depth (e.g. Karner et al. 2001) gradients, as well as between similar environments, e.g. subtropical oligotrophic environments (Delong et al. 2006). Despite these compositional differences, many phylogenetic clusters appear in clone libraries from throughout the marine system and appear to have a cosmopolitan distribution (Pommier et al. 2005). However, the level of environmental change sufficient to cause selection between closely related organisms is still not well understood. Few reports have highlighted differential spatial distributions for closely related marine taxa. Of those that have, taxonomic differentiation generally

\*Email: mbrown@ifa.hawaii.edu

correlates with depth, e.g. within *Alteromonas macleodii* ecotypes (Lopez-Lopez et al. 2005), temperature, e.g. within the *Rosobacter* RCA clade (Selje et al. 2004), or both, e.g. within the SAR11 clade (Field et al. 1997, Garcia-Martinez & Rodriguez-Valera 2000, Brown & Fuhrman 2005) and the *Prochlorococcus* (Johnson et al. 2006). Using a 98% sequence similarity 'contig' approach, Pommier et al. (2005) identified 11% of the common marine bacterial phylotypes occurring in tropical, temperate, and polar ecosystems, i.e. they displayed a truly cosmopolitan distribution at the level of analysis used. Notably no 'tropical only' phylotypes were identified in their analysis. One clade identified as 'cosmopolitan' was the Marine Group B/SAR324 (MGB/SAR324) clade. Herein, we use the newly acquired 16S rRNA gene and internal transcribed spacer (ITS) sequences from the Hawaiian Ocean Time-Series (HOT) in conjunction with GenBank sequences originating from 35 separate samples across 18 studies to refine the phylogeny of the MGB/SAR324 clade and analyze the distribution of the resulting phylotypes to identify biogeographic patterns.

## MATERIALS AND METHODS

Seawater from 10, 800 and 4400 m depths was collected at the HOT Stn ALOHA (22.45°N, 158°W) during HOT Cruise 169. Seawater (4 l) from each depth was pre-filtered through a 0.8 Whatman glass fiber GFA before passage through a 0.22 µm Durapore filter. Filters were transported in liquid nitrogen to the laboratory and stored at -80°C until DNA extraction, using the UltraClean Soil DNA Kit (MO BIO Laboratories). Clone libraries were constructed from polymerase chain reaction (PCR)-amplified products using the primers 27f (5'-AGAGTTTGATCMTGGCTCAG-3', *Bacteria* domain-specific 16S rDNA; Lane 1991) and 23Sr (5'-GGGTTBCCCCATTCRG-3', *Bacteria* domain-specific 23S rDNA; Fisher & Triplett 1999). The 100 µl reaction mixture contained a final concentration of 1 × PCR AmpliTaq Gold buffer (Applied Biosystems), 3.5 mM MgCl<sub>2</sub> (Applied Biosystems), 350 µM of each dNTP (Promega), 800 nM of each primer, 0.4 ng BSA (bovine serum albumin) µl<sup>-1</sup>, 5 U AmpliTaq Gold, and 0.1 ng template DNA µl<sup>-1</sup>. The reaction mixture was held at 94°C for 10 min followed by 24 cycles of amplification at 94°C for 40 s, 55°C for 40 s, and 72°C for 3 min, with a final step at 72°C for 7 min. Amplification products were purified using a MinElute PCR purification kit (QIAGEN) and ligated into pGEM-T Easy (Promega Corporation) plasmid vectors. Ligation reactions were transformed into JM109 High Efficiency Competent Cells (Promega Corporation), and colorimetric screening was used to determine which clones

contained recombinant vectors. Sequencing reactions were carried out using ABI PRISM BigDye Terminator chemistry (Applied Biosystems) and 5 pmol of primer 23Sr, and were run on an ABI3730XL capillary-based DNA sequencer at the Center for Genomics, Proteomics and Bioinformatics Research Initiative (CGP-BRI), University of Hawaii, Manoa. Electropherograms were visualized and annotated using the program Chromas, Version 1.45 ([www.mb.mahidol.ac.th/pub/chromas/chromas.htm](http://www.mb.mahidol.ac.th/pub/chromas/chromas.htm)). Sequences originating from this project were deposited in GenBank under the Accession Numbers EF370451 to EF370471.

Representative 16S rRNA gene sequences and ITS sequences from the 3 MGB/SAR324 clades were compared to the NCBI nucleotide (nr) and environmental trace (env\_nt) databases using the nucleotide-nucleotide BLAST (BLASTn) tool, with the number of descriptions and alignments for return set at 500. All sequences displaying phylogenetic affiliations within the MGB/SAR276 clade or the MGB/SAR324 clade were subsequently included in the analysis. All sequences were initially aligned using ClustalW, and alignments were manually curated in BioEdit ([www.mbio.ncsu.edu/BioEdit/bioedit.html](http://www.mbio.ncsu.edu/BioEdit/bioedit.html)). In our 16S rRNA gene sequence alignments in which multiple sequences originated from a single sample, 1 representative sequence was retained and the average pairwise similarity of those removed was reported. Distance matrices of representative sequences were generated using the Jukes-Cantor distance method in SplitsTree IV (Huson & Bryant 2006), which was also used to create neighbor-joining phylogenetic reconstructions, as well as for bootstrap analysis (1000 replicates). Sequence clustering and bootstrap analysis were verified using the programs SeqBoot, DNAdist (with the Jukes-Cantor distance method), Neighbor (with the neighbor-joining method), and Consensus (with 1000 replicate samplings in the PHYLIP package) (Felsenstein 2004). The Intergenic Transcribed Spacers Word Count-based System Analyzer (IWOCs; <http://egg.umh.es/iwocs/>) was used to determine tRNA-associated regions within ITS sequences.

Where temperature data were not available in original reports, representative data were acquired from the World Ocean Circulation Experiment database or the SCICEX database.

## RESULTS

16S rRNA gene and associated ITS sequences associated with the uncultured MGB/SAR324 clade were obtained from 3 depths at Stn ALOHA. As many of the related 16SrRNA gene sequences contained in GenBank constituted fragments from various regions of the

molecule, phylogenetic analysis was carried out on 3 portions of the molecule where it was deemed that most representatives contained sequence data of alignable quality. Thus, Fig. 1 represents phylogenies

of available sequences over the clone *Escherichia coli* bp positions (1) 87 to 535, (2) 536 to 1126, and (3) 1144 to 1530, in total representing 1355 base pairs of 1509 bp. This method also allows the incorporation of

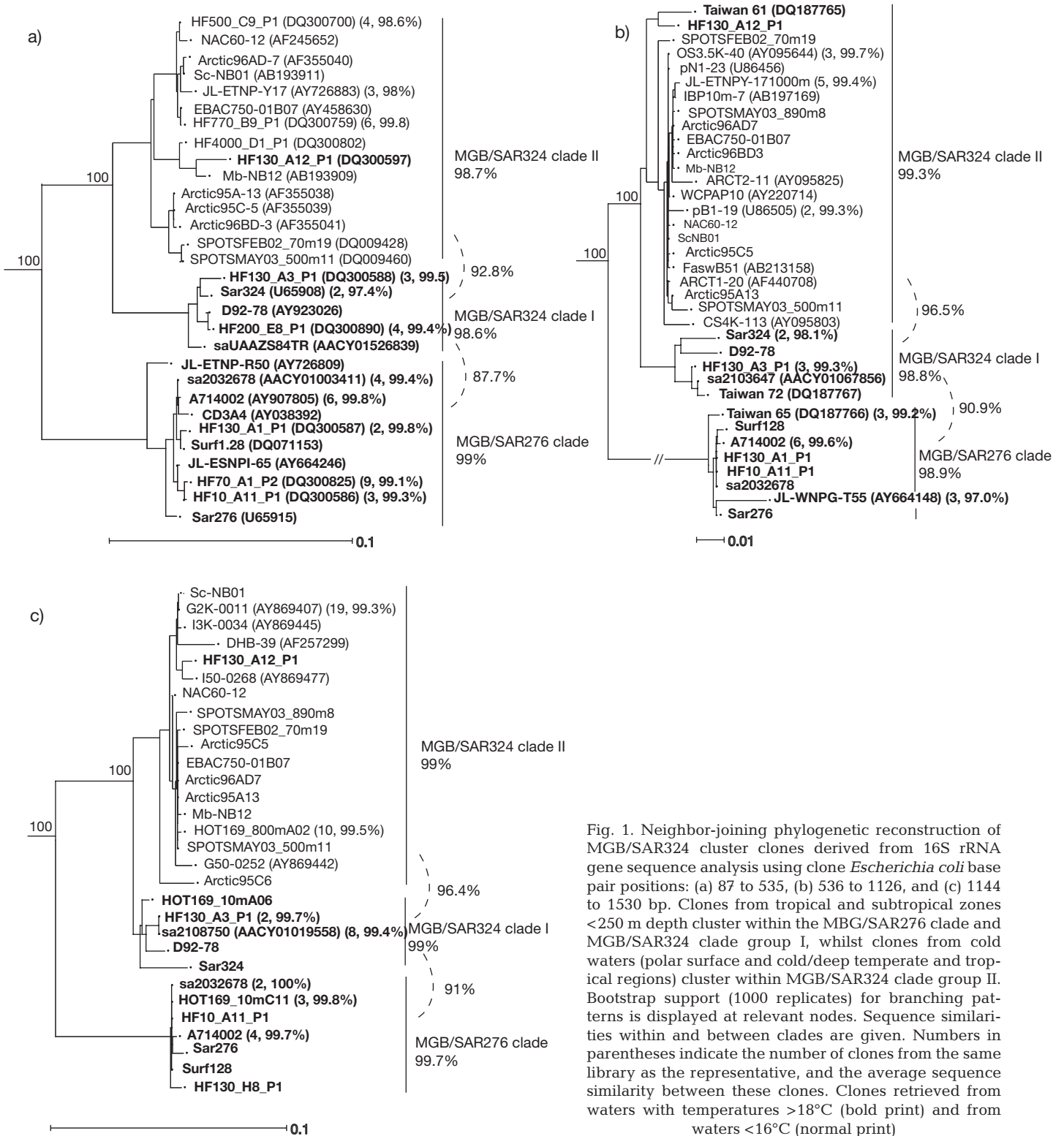


Fig. 1. Neighbor-joining phylogenetic reconstruction of MGB/SAR324 cluster clones derived from 16S rRNA gene sequence analysis using clone *Escherichia coli* base pair positions: (a) 87 to 535, (b) 536 to 1126, and (c) 1144 to 1530 bp. Clones from tropical and subtropical zones <250 m depth cluster within the MBG/SAR276 clade and MGB/SAR324 clade group I, whilst clones from cold waters (polar surface and cold/deep temperate and tropical regions) cluster within MGB/SAR324 clade group II. Bootstrap support (1000 replicates) for branching patterns is displayed at relevant nodes. Sequence similarities within and between clades are given. Numbers in parentheses indicate the number of clones from the same library as the representative, and the average sequence similarity between these clones. Clones retrieved from waters with temperatures >18°C (bold print) and from waters <16°C (normal print)

partial chimeric sequence information. For example, clone IBP10m-7 from the Mid-Okinawa Trough was used as the outgroup in Fig. 1a, but clusters within MGB/SAR324 clade group II when only the 3' portion

of the sequence is analyzed (Fig. 1b). The 3 clones from Stn ALOHA depth 4400 m contained a 62 bp insert at MGB/SAR324 bp position 1417. Due to this, the length of alignable 16S rRNA gene sequence from these 3

Table 1. Sampling locations, depths, temperatures, and reference information for representative MGB/SAR324 clade sequences used in this study

Clone prefix	Latitude (° N)	Longitude (° W)	Depth (m)	Temp. (°C)	Region	Source
<b>MGB/SAR324 clade group I</b>						
HOT169_10m	22.45	158	10	25.4	Tropical North Pacific	Present study
sa	32	64	10	20	Sargasso Sea	Venter et al. (2004)
Taiwan	25	172	100	24	Kuroshio Current Northeast Taiwan	Unpubl.
HF130	22.45	158	130	22.19	Tropical North Pacific	Delong et al. (2006)
D92	32	64	200	19.2 <sup>b</sup>	Subtropical North Atlantic	Morris et al. (2005)
HF200	22.45	158	200	18.53	Tropical North Pacific	Delong et al. (2006)
Sar324	31	64	250	18.4 <sup>b</sup>	Subtropical North Atlantic	Wright et al. (1997)
<b>MGB/SAR324 clade group II</b>						
G50	72.4	-4	50	3.9	Greenland Sea	Zaballos et al. (2006)
I50	36.5	-15.5	50	15.2	Mediterranean Sea	Zaballos et al. (2006)
Arctic95A	75.47	-206.32	55	-2.1 <sup>a</sup>	Arctic Ocean	Bano & Hollibaugh (2002)
Arctic95C	72.34	-204.13	55	-2.2 <sup>a</sup>	Arctic Ocean	Bano & Hollibaugh (2002)
Arct2	79.5	-167	58	-2 <sup>a</sup>	Arctic Ocean	Unpubl.
Arct1	85	-176.5	58	-2.1 <sup>a</sup>	Arctic Ocean	Unpubl.
SPOTSFEBO2_70m	33.55	118.4	70	13.6	Southern California Coast	Brown et al. (2005)
Taiwan	25	172	100	24	Kuroshio Current Northeast Taiwan	Unpubl.
HF130	22.45	158	130	22.19	Tropical North Pacific	Delong et al. (2006)
Arctic96AD	88.25	-356.11	131	-1.5 <sup>a</sup>	Arctic Ocean	Bano & Hollibaugh (2002)
Arctic96BD	86.17	-51.45	131	-1.6 <sup>a</sup>	Arctic Ocean	Bano & Hollibaugh (2002)
SPOTSMAY03_500m	33.55	118.4	500	6.4	Southern California Coast	Brown et al. (2005)
NAC60	59	21	500	7.48 <sup>b</sup>	North Atlantic	Gonzalez et al. (2000)
HF500	22.45	158	500	7.25	Tropical North Pacific	Delong et al. (2006)
EBAC750	35.5	124.89	750	4.57 <sup>b</sup>	Central California Coast	Unpubl.
HF770	22.45	158	770	4.78	Tropical North Pacific	Delong et al. (2006)
HOT169_800m	22.45	158	800	4.57	Tropical North Pacific	Present study
SPOTSMAY03_890m	33.55	118.4	890	4	Southern California Coast	Brown et al. (2005)
pB1	32.17	64.3	1000	6.9 <sup>b</sup>	Subtropical North Atlantic	Fuhrman & Davis (1997)
WCPAP	49	-16	1000	8.8 <sup>b</sup>	Porcupine Abyssal Plain	Gallagher et al. (2004)
JL-TNP	19	158	1000	3.88 <sup>b</sup>	Tropical North Pacific	Unpubl.
IBP10m	27.47	-126	1033	4.56 <sup>b</sup>	Tropical North Pacific	Unpubl.
Sc-NB01	28.34	-140.3	1360	3.15 <sup>b</sup>	Suiyo seamount	Unpubl.
G2K	72.4	-4	2000	-0.8	Greenland Sea	Zaballos et al. (2006)
I3K	36.5	-15.5	3000	13.8	Mediterranean Sea	Zaballos et al. (2006)
pn1	32.3	121	3000	1.6 <sup>b</sup>	Northwest Pacific	Fuhrman & Davis (1997)
DHB	-59	55	3000	1.29 <sup>b</sup>	Antarctic Polar Front	Lopez-Garcia et al. (2001)
OS3.5K	32.5	120	3500	1.55 <sup>b</sup>	North East Pacific	Unpubl.
HF4000	22.45	158	4000	1.46	Tropical North Pacific	Delong et al. (2006)
CS4k	-16	-166	4000	1.41 <sup>b</sup>	Coral Sea	Unpubl.
HOT169_4.4K	22.45	158	4400	1.14	Tropical North Pacific	Present study
Fasw	17.5	143	5000	1.49 <sup>b</sup>	Mariana Trough	Unpubl.
Mb-NB12	17.5	143	5000	1.5 <sup>b</sup>	Mariana Trough	Unpubl.
<b>MGB/SAR276 clade</b>						
CD3A4	12.18	69	2	28.9 <sup>b</sup>	Curacao, Caribbean	Frias-Lopez et al. (2002)
Surf			2		Gulf stream ring	Unpubl.
HOT169_10m	22.45	158	10	25.4	Tropical	Present study
A7	19	-67	10	26.86 <sup>b</sup>	Arabian Sea	Fuchs et al. (2005)
sa	32	64	10	20	Sargasso Sea	Venter et al. (2004)
HF70	22.45	158	70	24.93	Tropical North Pacific	Delong et al. (2006)
JL-ETNP	19	158	100	25.2 <sup>b</sup>	East Tropical North Pacific	Unpubl.
Taiwan	25	172	100	24	Kuroshio Current Northeast Taiwan	Unpubl.
HF130	22.45	158	130	22.19	Tropical North Pacific	Delong et al. (2006)
Sar276	31	64	250	18.4 <sup>b</sup>	Subtropical North Atlantic	Wright et al. (1997)

<sup>a</sup> Temperature estimated from SCICEX database

<sup>b</sup> Temperature estimated from WOCE database

clones was determined to be insufficient for inclusion in the 16S rRNA gene alignment, but a full-length ITS sequence was obtained from these clones and was included in the ITS and biogeographical analyses.

Partial sequence analysis over all 3 segments of the 16S rRNA gene resulted in a consistent pattern showing that the MGB/SAR324 clade consists of 3 distinct clusters, always with >98% within-cluster sequence similarity, but displaying <97% between-cluster sequence similarity (Fig. 1). These 3 clusters are designated MGB/SAR324 clade group I, which contains the original clone from 250 m depth in the Sargasso Sea; MGB/SAR324 clade group II, which contains the majority of available 16S rRNA gene sequences associated with this clade; and the MBG/SAR276 clade, a clearly divergent group displaying <92% average sequence identity to the MGB/SAR324 clade clones. These designations are all supported by 100% bootstrap values at the relevant nodes. The shading in Fig. 1 represents the actual or derived water temperature of the region from which the clones were retrieved (Table 1), with bold print representing water >18°C and normal print representing water <16°C. MGB/SAR276 clade group I contains sequences from tropical and subtropical waters above 250 m depth. These clones originated from waters with an average temperature of 24°C, with a minimum and maximum of 18.4 and 28.9°C, respectively (Table 1). Similarly, MGB/SAR324 clade group I contains clones originating from waters with an average temperature of 21.3°C, with a minimum and maximum of 18.4 and 25.4°C, respectively. These clones are also from tropical and subtropical waters above 250 m. Conversely, MGB/SAR324 clade II mainly contains sequences originating from the near surface (50 to 131 m depths) and deep waters (2000 and 3000 m) in permanently cold polar regions and in deep (>500 m) colder waters in temperate and tropical regions (Table 1). The average water temperature of origin for these clones is 4.8°C, and the minimum is -2.2°C. There are 2 instances of MGB/SAR324 clade group II sequences occurring at or above 130 m in tropical areas (130 m at Stn ALOHA and 100 m in the Kuroshio Current) and 2 instances of it occurring in temperate regions (70 m at SPOTS and 50 m in the Mediterranean Sea); hence, the maximum temperature of origin is 24°C, 20°C above the average.

ITS sequence analysis strongly supports the distinction between these 3 clades, again with 100% bootstrap support at the relevant nodes (Fig. 2a). As with the 16S rRNA gene sequences, the ITS database contains some relevant clones that are not full-length sequences. The longest full-length ITS sequence associated with this group was 578 bp (clone HOT169\_800mF01). Relative to this clone, the first 407 bp was used in ITS sequence analysis (Fig. 3) to ensure quality

alignable sequences for most representatives. Although the number of environments from which ITS sequences have been derived is much smaller than that from which 16S rRNA gene sequences have been obtained, the same pattern of clones from tropical and subtropical surface waters associating with the MBG/SAR276 clade and MGB/SAR324 clade group I and from cold/deep water sequences associating with MGB/SAR324 clade group II is maintained (Table 1). Interestingly, the ITS sequence associated with the Mediterranean surface clone I50268 does not fall within the MGB/SAR324 clade group II cluster in the ITS analysis. The average inter-cluster ITS sequence similarity of 75.9% between MGB/SAR324 clade groups I and II is similar to the level of ITS sequence similarity shown by these clades to the MBG/SAR276 clade (72.3 and 75.9%, respectively; Fig. 1c), which, as described above, branches well below the 'species' level (97%) in 16S rRNA gene sequence analysis. All ITS sequences contain 2 highly conserved regions, corresponding to tRNA Ile and Ala (Fig. 3). The ITS sequence associated with the 50 m sample from the Mediterranean Sea, whose 16S rRNA gene sequence clusters within MGB/SAR324 clade group II, falls well outside that group in this analysis. The majority of available ITS sequences are associated with MGB/SAR324 clade group II. These sequences fall into 2 broad categories, those containing a ~90 bp insert at Position 240 bp (relative to clone SPOTSFE02\_70m19) (Fig. 3b) and those without the insert (Fig. 3c). Unlike the 16S rRNA gene insert, which only occurred in sequences from 4400 m depth at Stn ALOHA, there was no clear depth association for the ITS insert. However, the clones from the Greenland Sea (G50, G2K), the only polar clones analyzed, generally clustered together.

A graphical representation of the latitudinal versus depth parameters of all sequences examined (16S rRNA gene and ITS sequences) provides a picture of the global distributions of the 3 groups described (Fig. 4). MGB/SAR324 clade group I and the MBG/SAR276 clade are restricted to <250 m in tropical and subtropical regions. MGB/SAR324 clade group II clones appear widespread in both the high and low latitudes, appearing at depths to 5000 m, but are restricted to depths >130 m in tropical and subtropical regions.

## DISCUSSION

The MGB/SAR324 clade, a deeply branching cluster within the Deltaproteobacteria, was first reported originating from depths of 500 and 3000 m in the north-eastern Pacific Ocean, from 1000 m in the subtropical

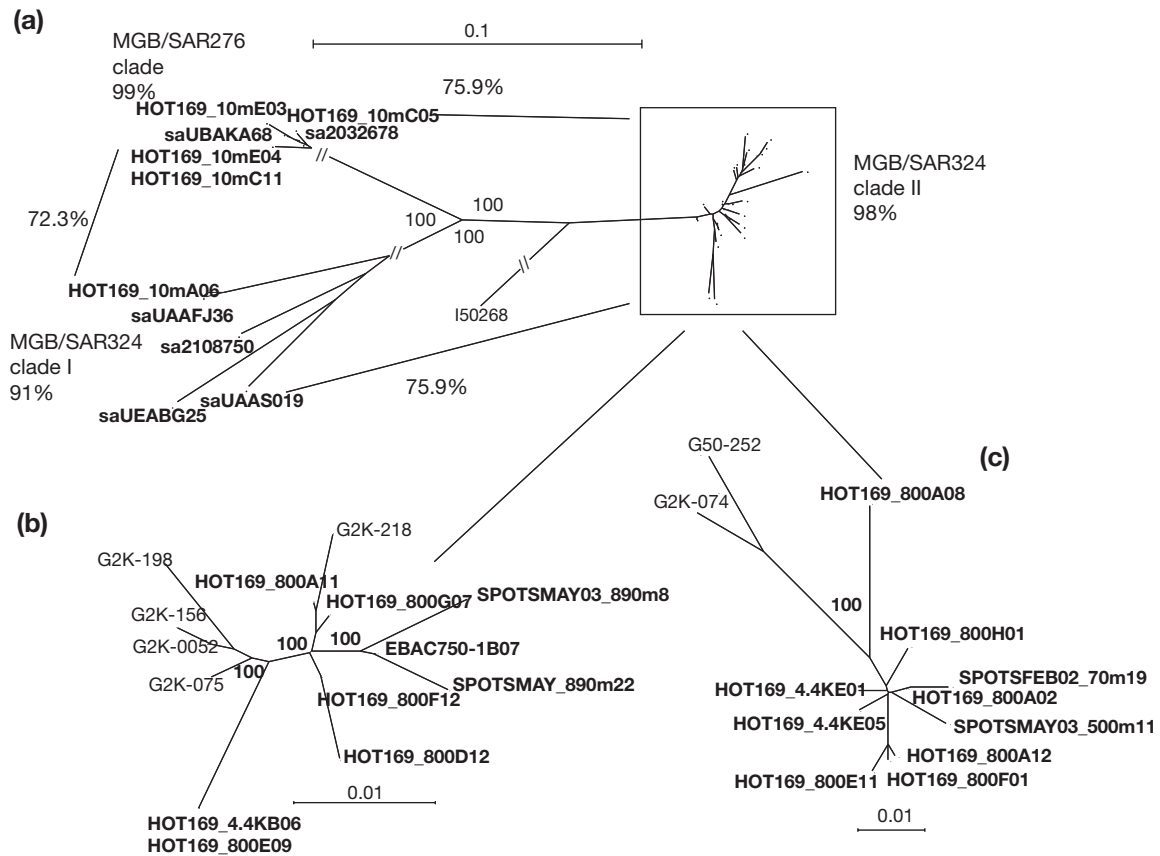


Fig. 2. (a) Neighbor-joining equal-angle reconstructions of MGB/SAR324 clade clones derived from 407 bp of the ITS sequence. Bootstrap support values are displayed at relevant nodes. Clones retrieved from waters with temperatures  $>20^{\circ}\text{C}$  (bold print) and from waters  $<16^{\circ}\text{C}$  (normal print). Clones from MGB/SAR324 clade group II were further analyzed in 2 parts: (b) those containing a ~90 bp insert and (c) those without insert

Atlantic Ocean (Fuhrman & Davis 1997), and from 250 m in the Sargasso Sea (Wright et al. 1997). Although no direct measure of absolute abundances are available for these clades, evidence suggests that MGB/SAR324 clade group II represents a relatively greater proportion of the bacterioplankton community in its niche than MGB/SAR324 clade group I and the MBG/SAR276 clade. Wright et al. (1997) used dot blot hybridization of clade-specific probes to amplified rRNA and rDNA to characterize the increased abundance of the group at depth over subtropical Atlantic Ocean and sub-arctic North Pacific Ocean water column profiles. At its maximum, around 200 m, the group accounted for 18% of the total amplified bacterial rRNA gene content. Furthermore, Gonzalez et al. (2000) detected increasing abundances of DNA in fragment lengths associated with this clade using terminal restriction fragment length polymorphism (TRFLP) analysis of waters to 500 m during a North Atlantic algal bloom. Neither the probes used by Wright et al. (1997), nor the TRFLP protocol used by Gonzalez et al. (2000) were able to discriminate

between the 3 species-level phylotypes—MGB/SAR324 clade groups I and II and the MBG/SAR276 clade identified here. Therefore, a clear picture of the distributions of these phylotypes has never been elucidated. However, given the restricted tropical and subtropical distributions of MGB/SAR324 clade group I and the MBG/SAR276 clade, it is likely that only the analysis of subtropical Atlantic Ocean waters by Wright et al. (1997) would have been influenced by the presence of these groups. Another potential indication that MGB/SAR324 clade group II is more abundant than those clades found in surface waters comes from a comparative metagenomics analysis of the microbial communities retrieved from 7 depths at the HOT site (Delong et al. 2006). Analysis of high-scoring protein hits (HSPs) from each depth revealed that between 0.6 and 1% of proteins from depths of 200, 500 and 770 m most closely matched those from a large genome fragment (EBAC750-01B07) containing a 16S rRNA gene sequence belonging to the MGB/SAR324 clade group II retrieved from 750 m in Monterey Bay, California. Less than 0.1% of the proteins found in samples

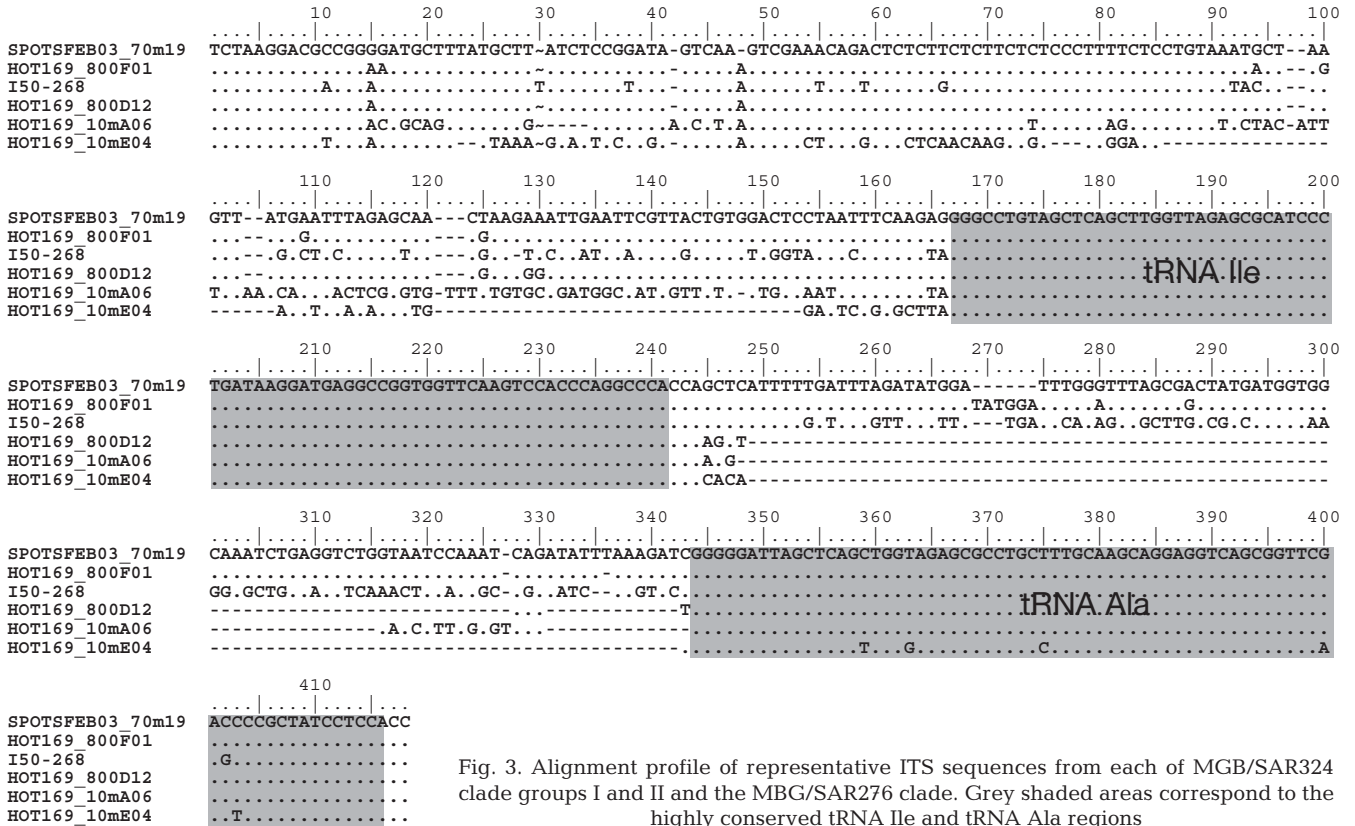


Fig. 3. Alignment profile of representative ITS sequences from each of MGB/SAR324 clade groups I and II and the MBG/SAR276 clade. Grey shaded areas correspond to the highly conserved tRNA Ile and tRNA Ala regions

from 10, 70, and 130 m most closely matched this fragment. However, this result could also be a function of differing gene complements in the surface-related clades.

It has been suggested previously that thermal tolerance may define the environmental distribution of this group (Bano & Hollibaugh 2002). Although this suggestion was made on the incorrect assumption of phylogenetic coherence between MGB/SAR324 clade groups I and II, it does appear to be accurate. Using the definition of Pommier et al. (2005), the MBG/SAR276 clade and MGB/SAR324 clade group I comprise, respectively, the first identification of 'genus'- and 'species'-level marine organisms to display a 'tropical-only' distribution. The fact that MGB/SAR324 clade group I and the MBG/SAR276 clade co-exist in low-latitude, upper water columns indicates they have evolved to perform different functions within the community. However, the apparent replacement of MGB/SAR324 clade group II by the species-level relative MGB/SAR324 clade group I within the same environment suggests that these organisms perform a similar function and that a niche separation in temperature or a correlating environmental factor (such as different availability of nitrogen species) drives their distribution, and could have been the selective pressure which led to this speciation event. Unlike the

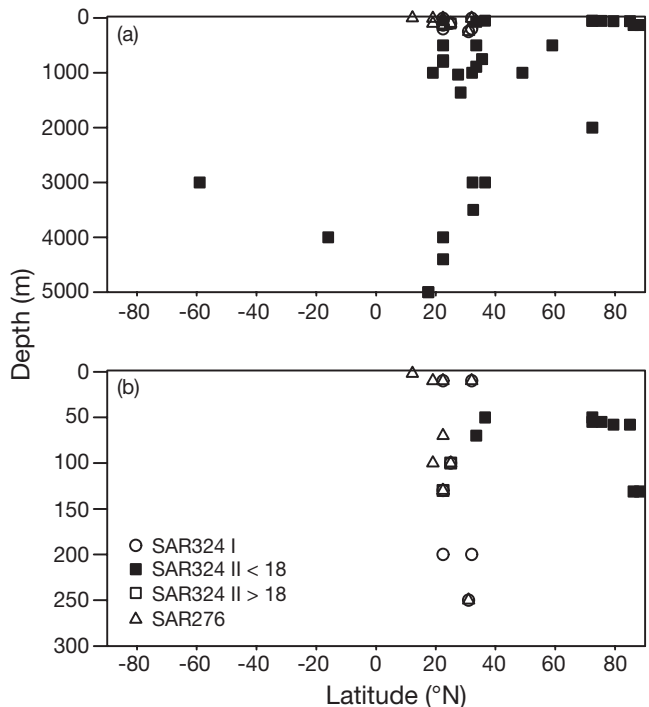


Fig. 4. (a) Depth and latitude distributions of all MGB/SAR324 clade groups I and II and MBG/SAR276 clade sequences currently available, showing the restricted distributions of all 3 species-level phylotypes. (b) A higher resolution image of depths above 300 m

SAR11 clade, in which different subpopulations inhabit deep waters compared to cold surface waters (Field et al. 1997, Garcia-Martinez & Rodriguez-Valera 2000, Brown & Fuhrman 2005), current data suggest that the deep representatives of MGB/SAR324 clade group II are very closely related to those in cold surface water samples. This suggests that a different selection process is occurring for these organisms. The distribution of the SAR11 clade may be related to light regimes rather than actual temperature regimes, whilst temperature may be the actual selection pressure for the MGB/SAR324 clades. However, whilst more ITS data are required to resolve the fine-scale branching patterns, preliminary evidence (Fig. 3b,c) suggests that there may be some variation in phylotypes occurring at the poles.

The number of clusters displaying sequence differentiation in the 16S rRNA gene or ITS sequence when compared between tropical, temperate, and polar regions, or between surface and deep ecosystems, indicates a substantial degree of niche-specific adaptation occurring in these oceanic zones. Whilst the actual physical movement of microbial cells between these temperature zones may not be a barrier, as evidenced by the mixing of MGB/SAR324 clade groups I and II at 130 m at the HOT station and at 100 m in the Kuroshio Current, it appears adaptations have occurred which support the maintenance of prevalent ecotypes, thus supporting the view that reproductive isolation may not require geographic isolation (Dolan 2005). Relatively small-scale environmental differences may lead to adaptation and competition relegating closely related imported strains, which perform a similar function, to background levels (i.e. at approximately the level of importation). At such levels, these imported organisms would have little or no effect on gene flow or biogeochemical cycling in the system, a situation that may eventually lead to microbial speciation events.

Our findings add to the growing body of evidence (e.g. Brown & Fuhrman 2005, Morris et al. 2005, Pommer et al. 2005) supporting the assertions of Selje et al. (2004) that the global distribution of major marine bacterioplankton components is related to oceanic water masses and controlled by their environmental and biogeochemical properties. Given the niche-specific distributions now reported here for the MGB/SAR324 clade and previously reported for the SAR11 clade (Field et al. 1997, Garcia-Martinez & Rodriguez-Valera 2000, Brown et al. 2005), the *Alteromonas macleodii* clade (Lopez-Lopez et al. 2005), the *Roseobacter* clade (Selje et al. 2004), and the *Prochlorococcus* (e.g. Johnson et al. 2006), it appears there are few, if any, truly ubiquitous marine bacteria, i.e. those which have not undergone some degree of niche specialization (Fuhrman et al. 2006).

The environmental roles of the 3 clades described here are difficult to determine from current knowledge. The DNA sequence of EBAC750-01B07 reveals a complement of 40 proteins, half of which are hypothetical. One putative enzyme product (GenBank No. AAR37470) is identified by Blast homology as a phosphonate monoester hydrolase. A nearly identical gene fragment (ASXB2404.b2 HF500\_10-06-02), sharing 96% sequence similarity over 1047 bp, was recovered from 500 m at the HOT site (Delong et al. 2006). Although the latter fragment is not linked to a 16S ribosomal gene, its sequence similarity and depth of retrieval are consistent with it belonging to MGB/SAR324 clade group II. The characterized protein showing best homology to these fragments is a phosphonate monoester hydrolase from *Burkholderia caryophylli* PG2982 (Dotson et al. 1996). The most catalytically efficient substrate for this enzyme was found to be bis-(p-nitrophenyl) phosphate, suggesting that the enzyme may function *in vivo* as a phosphodiesterase (Dotson et al. 1996). This may implicate the MGB/SAR324 clade in cycling of the complex dissolved organic phosphorous pool by degradation of phosphate mono- and diester compounds, which may not be readily accessible by many phytoplankton species (e.g. Yamaguchi et al. 2005). Protein homologues also appear in whole genome sequences of several marine Alphaproteobacteria thought to be involved in sulfur cycling, such as *Silicibacter pomeroyi* (Moran et al. 2004), *Roseovarius nubinhibens* (EAP75110), and *Roseobacter* sp. MED193 (EAQ43397). Furthermore, the role of phosphonate monoester hydrolase from *B. caryophylli* PG2982 in the degradation pathway of glyceryl-glyphosate (Dotson et al. 1996), a potent herbicide, suggests that MGB/SAR324 may provide a natural source of bioremediation for xenobiotic organophosphates in marine systems, and also suggests a potentially powerful selective substrate for isolation of the group.

*Acknowledgements.* We thank C. Mahaffey for assistance with the manuscript, Norbert Schorghofer for many discussions, and the crew of the RV 'Kilo Moana' and the Hawaiian Ocean Time-Series for sample collection. This work was funded by support from the NASA Astrobiology Institute under Cooperative Agreement NNA04CC08A at the Institute for Astronomy (University of Hawaii-Manoa).

#### LITERATURE CITED

- Baas-Becking LGM (1934) Geobiologie of Inleiding Tot de Milieukunde. Van Stockum and Zoon, The Hague
- Baldwin AJ, Moss JA, Pakulski JD, Catala P, Joux F, Jeffrey W (2005) Microbial diversity in a Pacific Ocean transect from the Arctic to Antarctic circles. *Aquat Microb Ecol* 41: 91–102



- Bano N, Hollibaugh JT (2002) Phylogenetic composition of bacterioplankton assemblages from the Arctic Ocean. *Appl Environ Microbiol* 68:505–518
- Beijerinck MW (1913) De infusies en de ontdekking der bacterien. In: *Jaarboek van de Koninklijke Akademie v Wetenschappen*. Muller, Amsterdam (Reprinted in *Verzamelde geschriften van M. W. Beijerinck*, vijfde deel, p 119–140, Delft, 1921)
- 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 ARISA to show marine microbial diversity; development and application to a time series. *Environ Microbiol* 7:1466–1479
- Delong EF, Preston CM, Mincer T, Rich V and 8 others (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311:496–503
- Dolan JR (2005) Biogeography of aquatic microbes. *Aquat Microb Ecol* 44:39–48
- Dotson SB, Smith CE, Ling CS, Barry GF, Kishmore GM (1996) Identification, characterization, and cloning of a phosphate monoester hydrolase from *Burkholderia caryophylli* PG2982. *J Biol Chem* 271:25754–25761
- Felsenstein J (2004) PHYLIP (Phylogeny Inference Package) Version 3.6. Distributed by the author, Department of Genome Sciences, University of Washington, Seattle
- Field KG, Gordon D, Wright T, Rappe M, Urback 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
- Fisher MM, Triplett EW (1999) Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Appl Environ Microbiol* 65:4630–4636
- Frias-Lopez J, Zerkle AL, Bonheyo GT, Fouke BW (2002) Partitioning of bacterial communities between seawater and healthy, black band diseased, and dead coral surfaces. *Appl Environ Microbiol* 68:2214–2228
- Fuchs BM, Woebken D, Zubkov MV, Burkill P, Amann R (2005) Molecular identification of picoplankton populations in contrasting waters of the Arabian Sea. *Aquat Microb Ecol* 39:145–157
- Fuhrman JA, Davis AA (1997) Widespread *Archaea* and novel *Bacteria* from the deep sea as shown by 16S rRNA gene sequences. *Mar Ecol Prog Ser* 150:275–285
- Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV, Naeem S (2006) Annually reoccurring bacterial communities are predictable from oceanic conditions. *Proc Natl Acad Sci* 103:13104–13109
- Gallagher JM, Carton MW, Eardly DF, Patching JW (2004) Spatio-temporal variability and diversity of water column prokaryotic communities in the eastern North Atlantic. *FEMS Microbiol Ecol* 47:249–262
- Garcia-Martinez J, Rodriguez-Valera F (2000) Microdiversity of uncultured marine prokaryotes: the SAR11 cluster and the marine archaea of group I. *Mol Ecol* 9:935–948
- Giovannoni SJ, Britschgi TB, Moyer CL, Field KG (1990) Genetic diversity of Sargasso Sea bacterioplankton. *Nature* 345:60–63
- Gonzalez JM, Simo R, Massana R, Covert JS, Casamayor EO, Pedros-Alio C, Moran MA (2000) Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl Environ Microbiol* 66:4237–4246
- Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. *Mol Biol Evol* 23:254–267
- Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EMS, Chisholm SW (2006) Niche partitioning amongst *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* 311:1737–1740
- Karner MB, DeLong EF, Karl DM (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409:507–510
- Lopez-Garcia P, Lopez-Lopez A, Moreira D, Rodriguez-Valera F (2001) Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. *FEMS Microb Ecol* 36:193–202
- Lopez-Lopez A, Bartual SG, Stal L, Onyshchenko O, Rodriguez-Valera F (2005) Genetic analysis of housekeeping genes reveals a deep-sea ecotype of *Alteromonas macleodii* in the Mediterranean Sea. *Environ Microbiol* 7:649–659
- Moran MA, Buchan A, Gonzalez JM, Heidelberg JF and 31 others (2004) Genome sequence of *Silicibacter pomeroyii* reveals adaptations to the marine environment. *Nature* 432:910–913
- Morris RM, Vergin KL, Cho JC, Rappé MS, Carlson CA, Giovannoni SJ (2005) Temporal and spatial response of bacterial lineages to annual convective overturn at the Bermuda Atlantic time-series study site. *Limnol Oceanogr* 50:1687–1696
- Pommier T, Pinhassi J, Hagström Å (2005) Biogeographic analysis of ribosomal RNA clusters from marine bacterioplankton. *Aquat Microb Ecol* 44:79–89
- Selje N, Simon M, Brinkhoff T (2004) A newly discovered *Roseobacter* cluster in temperate and polar oceans. *Nature* 427:445–448
- Venter JC, Remington K, Heidelberg JF, Halpern AL and 19 others (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304:66–74
- 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
- Yamaguchi H, Yamaguchi M, Fukami K, Adachi M, Nishijima T (2005) Utilization of phosphate diester by the marine diatom *Chaetoceros ceratosporus*. *J Plankton Res* 27:603–606
- Zaballos M, Lopez-Lopez A, Galan Bartual S, D'Auria G and 5 others (2006) Comparison of prokaryotic diversity at offshore oceanic locations reveals a different microbiota in the Mediterranean Sea. *FEMS Microb Ecol* 56:389–405

Editorial responsibility: Jed Fuhrman,  
Los Angeles, California, USA

Submitted: August 1, 2006; Accepted: November 6, 2006  
Proofs received from author(s): January 20, 2007