FREE ACCESS

Exploring bacteria–diatom associations using single-cell whole genome amplification

Lydia J. Baker*, Paul F. Kemp

Department of Oceanography, University of Hawai'i at Manoa, 1950 East West Road,
Center for Microbial Oceanography: Research and Education (C-MORE), Honolulu, Hawai'i 96822, USA

ABSTRACT: Diatoms are responsible for a large fraction of oceanic and freshwater biomass production and are critically important for sequestration of carbon to the deep ocean. As with most surfaces present in aquatic systems, bacteria colonize the exterior of diatom cells, and they interact with the diatom and each other. The ecology of diatoms may be better explained by conceptualizing them as composite organisms consisting of the host cell and its bacterial associates. Such associations could have collective properties that are not predictable from the properties of the host cell alone. Past studies of these associations have employed culture-based, whole-population methods. In contrast, we examined the composition and variability of bacterial assemblages attached to individual diatoms. Samples were collected in an oligotrophic system (Station ALOHA, 22° 45′ N, 158° 00′ W) at the deep chlorophyll maximum. Forty eukaryotic host cells were isolated by flow cytometry followed by multiple displacement amplification, including 26 *Thalassiosira* spp., other diatoms, dinoflagellates, coccolithophorids, and flagellates. Bacteria were identified by amplifying, cloning, and sequencing 16S rDNA using primers that select against chloroplast 16S rDNA. Bacterial sequences were recovered from 32 of 40 host cells, and from parallel samples of the free-living and particle-associated bacteria. Bacterial assemblages varied substantially even among closely related host cells. Host cells and the free-living and particle-associated samples can be placed into distinct groups based on the phylogenetic relatedness of their associated bacteria, rather than the identity of the host cell. As yet, the functional implications of these groups are unknown.

KEY WORDS: Bacterioplankton diversity · Marine diatoms · Algal–bacterial interaction · Microenvironment · Multiple displacement amplification · Single-cell

INTRODUCTION

Distinct groups of bacteria have adapted to living on surfaces in the open ocean (DeLong et al.1993, Mitchell et al.1995, Blackburn et al.1998, Grossart et al. 2006, 2007); associating with surfaces may confer an evolutionary advantage in exchange for the genetic load and metabolic cost of expressing genes associated with a surface-attached lifestyle. Even in the presence of adequate nutrients, some bacteria prefer surface colonization and invest in the production of antibacterial compounds to prevent competition with other species for the same surface (Yan et al. 2002, Gram et al. 2010). Bacterial morphology and metabolism change significantly to facilitate attachment to a surface, for example by producing large extracellular glycolipids and glycoproteins (Desai & Banat 1997).

As with other surfaces in aquatic environments, bacteria are found on the surface of diatoms. Early studies assumed that bacteria remain at some distance from healthy diatoms, possibly because diatoms are capable of producing antimicrobial agents, but that bacteria would attach to dead diatoms (Waksman & Butler 1937, Cole 1982, Azam et al.

1983). Later work has shown that living, healthy diatoms and bacteria are closely associated, with possible implications for diatom surfaces being an area of high bacterial metabolic rates (Grossart 2010). Marine diatom–bacterial interactions have been studied most often using cultured diatoms (e.g. Kogure et al. 1981, Grossart 1999, Schäfer et al. 2002, Grossart et al. 2005, 2007, Kaczmarska et al. 2005, Sapp et al. 2007a). A few studies of diatom–bacterial interactions have used natural populations of diatoms (Kaczmarska et al. 2005, Rooney-Varga et al. 2005, Amin et al. 2012), and in one case a mesocosm study of a simulated diatom bloom (Smith et al.1995). Consistent associations between specific bacterial and diatom taxa have been found (e.g. Schäfer et al. 2002, Amin et al. 2012), although other work suggests that the composition of diatom-associated bacterial assemblages shifts over weeks to months in culture (Sapp et al. 2007a,b).

The existence of consistent patterns in the association of bacteria with diatoms leads inevitably to asking about the possible consequences of such associations. Most past studies have approached this question through the conceptual framework of symbiosis, focusing on cells acting in one of several possible modes of a classic symbiotic relationship, e.g. mutualism, commensalism, or parasitism. Much of this past work has focused on the benefits and costs to either the diatom host or its associated bacteria. As one example, vitamin B_{12} is thought to limit phytoplankton growth in at least some areas of the global ocean (Karl 2002, Bertrand et al. 2011). Croft et al. (2005) proposed that bacteria in the muciferous layer of *Thalassiosira pseudonana* provide the diatom vitamin B_{12} and, in return, bacteria have a secure source of carbon, thereby forming a mutualistic relationship. More commonly, commensalism has been suggested wherein diatoms are unaffected while bacteria have access to a secure carbon source (Rosowski 1992, Droop 2007). Bacteria have also been shown to act as parasites, producing enzymes that can cause dissolution of the organic matrix on which diatom frustules are built (Bidle & Azam, 1999, Mayali & Azam 2004), and some diatoms have been shown to produce antibiotics to ward off parasitic bacteria (Grossart 1999, Desbois et al. 2008).

An intriguing, alternative perspective is to view the host cell and its bacterial associates as a composite organism with collective properties that may not be predictable from knowledge of the diatom alone. In one well-studied example, the collective properties of diatom–bacteria associations determine the production of domoic acid by *Pseudo-nitzschia* diatoms;

domoic acid is an amino acid that causes amnesiac shellfish poisoning in humans. Production of domoic acid appears to require direct contact between the diatom host and bacteria (Kobayashi et al. 2009). Specific bacteria are associated with different *Pseudo-nitzschia* species (Stewart et al. 1997) and bacterial associates appear to vary with the toxigenicity of the host diatom (Guannel et al. 2011).

Ecological interactions of diatoms and attached bacteria may have biogeochemical consequences. Any assessment of the biogeochemical role of diatoms necessarily incorporates the net result of diatom–bacterial associations, but a better understanding of such associations may improve predictions. The vitamin B_{12} example mentioned above is one example: understanding how attached bacteria may alleviate B_{12} deficiency could lead to better predictions of diatom biomass production. Diatoms are globally important in biomass production and particularly in the export of biomass to the deep ocean (Karl et al. 2012). However, different diatoms have very different propensities to aggregate and sink, and diatom–bacteria interactions may influence export by influencing the relative success of different diatom species. Thus, the collective properties of diatom–bacteria associations may have global implications.

Many diatom–bacteria interactions involve direct, cell-to-cell encounters, and it is likely that bacteria living in close proximity on a single host cell interact with one another as well as with the host. However, nearly all past studies have employed population or community-level methods that obscure cell-to-cell interactions. Little is known regarding the makeup or interactions of bacterial assemblages attached to single host cells. The sole exception appears to be a few microscopy studies describing the diverse morphology of bacteria attached to individual diatoms (Kaczmarska et al. 2005).

In the present study, we examined variability in the bacterial assemblage attached to single diatom host cells. We regard this as an essential step toward understanding diatom–bacteria interactions at an appropriate physical scale. This approach allowed us to evaluate interactions that cannot be studied with population and community-based methods. For example, we assessed whether (1) the bacteria associated with diatoms exhibited host specificity; (2) attached bacteria behaved as recurring consortia; or, conversely, (3) there was evidence of non-overlapping distributions of bacterial phylotypes among host cells, such as might result from the exclusion of one bacterium by another (e.g. through competition). In short, there is much to be learned from the study of

diatom-associated bacteria at the single cell level. We are not aware of any comparable application of a single-cell approach.

To the best of our knowledge, this study was also the first to examine cell-to-cell interactions of diatoms and bacteria in the oligotrophic open ocean. Previous studies have mainly focused on eutrophic systems, both marine and fresh water. In oligotrophic systems, nutrient limitation may lead to a greater importance of bacterial–diatom interactions, which may influence the maintenance of a pertinacious species, the relative success of different diatom species, the initiation and success of summertime blooms, and the fate of diatom biomass.

MATERIALS AND METHODS

Study site

Our study site was within the North Pacific Subtropical Gyre (NPSG), the largest contiguous biome on Earth (Karl 1999). The NPSG is an oligotrophic system between 15° N to 35° N and 135° E to 135° W; samples were collected from Station ALOHA (22° 45′ N, 158° 00′ W). Diatom populations in this system vary throughout the year in both species diversity and the abundance of individual species. Highest abundance of some species occurs in the summer months, especially in July (Scharek et al. 1999). Typical diatom species seen in blooms from June through September include *Rhizosolenia*, *Hemiaulus*, and *Mastogloia* (Dore et al. 2008). The causes of blooms in this system remain enigmatic (Karl 1999).

The deep chlorophyll maximum (DCM) is a feature found in many systems, including the NPSG (Cullen 1982). The position of the DCM varies at Station ALOHA but it is typically located at approximately 100 m depth. The DCM has been found to contain distinct diatom populations that have a high fucoxanthin to cell ratio. Diatoms in the DCM are primarily smaller pennate forms and appear less likely to sink out of the euphotic zone than the larger chain-forming diatoms found in the mixed layer (Scharek et al. 1999). The DCM is also an area of high nutrient flux (Huisman et al. 2006), which may lead to higher diatom diversity (Huisman & Weissing 1999). The DCM was selected for this study because chlorophyll content per cell is typically higher and diatoms are typically smaller in the DCM; both factors are useful for isolating diatom cells by flow cytometry, as described below.

Test of the cell concentration protocol

The abundance of eukaryotic cells in the oligotrophic waters of Station ALOHA is relatively low, and requires an initial cell concentration step for effective flow cytometric sorting. This concentration step could result in loss of attached bacterial cells. To evaluate this possibility, the bacteria attached to diatoms were counted in 3 treatments, using 4′,6-diamidino-2-phenylindole (DAPI)-staining and epifluorescence microscopy. The first treatment consisted of cells picked by micropipetting from 2 non-axenic cultures of diatoms (one pennate and one centric) that had been kept in f/2-Si media (Guillard & Ryther 1962, Guillard 1975) at 26°C and 27 µmol photons m^{-2} s^{-1}. The second treatment consisted of cells from the same cultures collected on 25 mm diameter, 5 µm pore-size Nuclepore™ polycarbonate membrane filters (Whatman). The third treatment consisted of cells from the same cultures collected on Nuclepore filters and subsequently rinsed with 0.5 l of 0.2 µm filter-sterilized seawater. The bacteria on 30 diatom host cells were counted for each treatment.

Sample collection and concentration

The final protocol used for field sampling was nearly identical to the test of the cell concentration protocol, with the exception that a larger volume of seawater was processed. We included both particle-associated and free-living bacteria into our study to compare the results of the single-cell approach to the more commonly used community-level approach, with comparable sequence effort.

Samples were collected from Station ALOHA from July 8 to 10, 2010, during Hawaii Ocean Time-series (HOT) cruise 223 (RV 'Kilo Moana' cruise 1012). The DCM was sampled once per day for 3 d, and these samples were pooled in later analyses. On each day, 3 replicates of 3 l volumes were collected by gentle peristaltic pump filtration onto a 25 mm diameter, 5 µm pore-size Nuclepore™ filter. While still moist, filters were rinsed with 0.5 l of 0.2 µm filter-sterilized seawater. One filter from each sampling period was set aside for community-level analysis of bacteria associated with particles. Subsequently, the 5 µm filtrate was captured on a 25 mm diameter, 0.2 µm pore-size Nuclepore™ filter, representing the free-living bacterial assemblage. Each day, 3 replicates of 1 l volumes were collected. Filters were immediately placed in

RNAlater® buffer (Qiagen), kept at room temperature overnight to saturate cells before being stored at −20°C (as recommended in the RNAlater® protocol). RNAlater® preserves samples against degradation and provides high quality DNA for at least 6 wk (Vink et al. 2005).

Cell sorting and whole genome amplification

Host cells were gently resuspended from the collection filter into RNAlater® prior to being sent to the Bigelow Laboratory Single Cell Genomics Center (SCGC) for cell sorting and subsequent genomic amplification. The SCGC is a specialized facility that maintains a DNA-free clean room to minimize the possibility of contamination during sorting and amplification; the method for decontamination is as described in Stepanauskas & Sieracki (2007).

At the SCGC facility, host cells were separated from the RNAlater® buffer by gravity filtration through a 10 µm mesh-size cell strainer (Becton Dickinson). They were resuspended in seawater collected from Station ALOHA that had been sterilized by tangential flow filtration followed by UV-treatment. Single host cells were sorted into wells containing 0.6 µl Tris-EDTA (TE) buffer by fluorescence-activated cell sorting (FACS) using a MoFlo (Beckman Coulter) flow cytometer with a 488 nm argon laser for excitation, a 200 µm nozzle orifice, and a CyClone™ robotic arm for droplet deposition into microplates. The cytometer was triggered on side scatter and was set to 'single 1 drop' mode to maximize sort purity. Cells were kept at −80°C until further processing.

Because the Bigelow facility had not previously sorted cells within the size range expected for diatoms, preliminary testing was conducted to establish sorting parameters that would select in favor of diatoms from Station ALOHA. Final gating of cells for FACS was based on strong chlorophyll *a* signals, and forward scatter indicating cell volumes between 20 and 100 µm. Multiple displacement amplification (MDA) proceeded as described by Swan et al. (2011). Two 384 well plates were sorted, each including 62 negative control wells and 3 positive control wells. A real-time PCR screen using 18S rDNA primers Euk528F and EukB followed MDA (Medlin et al. 1988, Zhu et al. 2005), to identify wells that contain eukaryotic rDNA. Wells that were 18S-positive were Sanger-sequenced to identify the host cells, and the 18S rDNA sequences were provided by the SCGC.

Bulk environment samples

As described earlier, particle-associated and free-living bacteria were obtained from the same source water as was used for single cell amplification. Both of these samples were extracted using a guanidinium-based lysis buffer and adsorption to a silica spin column (DNeasy Blood and Tissue Kit, Qiagen®). Thereafter, they were processed as described for MDA-amplified material, starting from the point of 16S rDNA cloning and sequencing. The primary intent was not to fully characterize either free-living or particle-associated bacteria, but instead to provide a broad comparison to host cell libraries at a comparable sequencing effort. Libraries constructed from these 2 samples were similar in size to the host-cell libraries and do not provide exhaustive coverage of bacterial diversity.

Amplification and cloning of 16S rDNA

The MDA-amplified DNA includes mitochondrial and chloroplast 16S rDNA associated with the host cell. Based on similarities between chloroplast and other bacteria-derived 16S rDNA sequences, we expected PCR amplification of bacterial 16S rDNA might be overwhelmed by the host cell's chloroplast 16S rDNA. When field-collected diatom samples were amplified using conventional 16S rDNA primers, 95% of the sequences were identified as chloroplast 16S rDNA (data not shown). Hodkinson & Lutzoni (2009) identified an 895F primer sequence that permitted amplification of bacterial 16S rDNA present in lichen (a fungal/algal symbiosis) without interference from chloroplast rDNA. The primer strongly discriminates against chloroplast 16S rDNA, and because of the evolutionary relationship of chloroplasts and cyanobacteria, the primer also discriminates against cyanobacteria. The effectiveness of the 895F primer in marine systems was examined using Primer Prospector (Walters et al. 2011), and its effectiveness with select families was further investigated using ARB software (Ludwig et al. 2004).

To increase the concentration of target DNA, a nested PCR protocol was developed using a first round of amplification with the 8F/1513R primer pair (Weisburg et al. 1991, Turner et al. 1999), followed by the 895F/1391R primer pair. Because nested PCR can increase the chances of contamination, a negative control was included for both rounds of amplification to assess sample purity; plates containing negative controls that proved

positive were removed from further analysis. The master mixes for both amplifications were similar to the recommendations outlined by the Platinum® Taq Polymerase users' manual (Invitrogen), with the exception of increasing the concentration of MgCl$_2$ to 2.5 µM. The DNA was further diluted 1:50 in TE buffer and approximately 0.02 ng of DNA was added to the PCR mix. Amplification of the first round followed procedures in the Platinum® *Taq* Polymerase users' manual, i.e. 95°C for 3 min followed by 30 cycles of 94°C for 1 min, 55°C for 1 min, 72°C for 1 min, and a final extension step of 72°C for 7 min. The first-round PCR product was diluted 1/500 in sterile water and reamplified as described by Hodkinson & Lutzoni (2009) using the 895F/1391R primer pair. The PCR amplification was initiated by a 3 min denaturation step at 94°C, followed by 24 cycles that proceeded as follows: 94°C for 30 s, 55°C for 30 s (decreasing by 0.4°C with each cycle) and 1 min at 72°C (increasing by 2 s with each cycle). This was followed by 12 cycles of 94°C for 30 s; 45°C for 30 s; 72°C for 120 s, increasing by 3 s with each cycle; and a final extension step of 10 min at 72°C.

The PCR product was separated on a 1.3% agarose gel in 0.5× Tris-acetate-EDTA buffer and the product of the correct size was excised and purified using the PureLink™ PCR Purification Kit (Invitrogen). The product DNA was cloned using a TOPO TA Cloning Kit® (Invitrogen) and unidirectional sequences were obtained via Sanger sequencing on an ABI 3730XL at the Advanced Studies of Genomics, Proteomics and Bioinformatics Sequencing Services located at the University of Hawaii at Manoa.

Initial data obtained by microscopy counts, performed as is outlined in 'Test of the cell concentration protocol' above, indicated that the number of bacterial cells per diatom could range widely. Published reports of bacterial diversity per diatom cell based on electron microscopy suggest that sequence diversity could be much lower (Kaczmarska et al. 2005, 6 to 9 bacterial morphotypes per host cell). Studies of bacterial diversity associated with diatom cultures also suggest low diversity (e.g. 8 and 14 bacterial phylotypes associated with 2 *Pseudo-nitzschia* species; Guannel et al. 2011). We chose an initial sequencing effort of thirty-two 16S rDNA sequences for host cell libraries, and subsequent analyses confirmed that this size was sufficient to represent bacterial diversity well. The particle-associated and free-living libraries were based on a comparable level of effort, although slightly larger.

Evaluation of cell sorting and MDA

The quality of the sorting process was tested upon return of the sorted, MDA-processed samples to the University of Hawaii. A set of wells in which no 18S rDNA had been detected was tested for the presence of bacterial 16S rDNA. These included 40 wells that were intended to receive a host cell but had not resulted in successful recovery of 18S rDNA, and 15 wells that were intended as negative controls and were not expected to contain host cells. Samples from each were amplified using the PCR protocol described above. 16S rDNA was amplified in one-half of the wells tested. Subsets of these 16S rDNA-positive wells were chosen for cloning and sequencing to determine the identity of the bacteria. Eight clones were selected and sequenced from each of 6 of the intended host-cell wells, and from each of 7 of the intended negative-control wells. Sample sequences were removed from analysis if they matched any sequence obtained from these control wells, using a 98% similarity threshold. Sequences were also removed from analysis if their nearest neighbors in the ARB SILVA database (Pruesse et al. 2007) did not include any sequences known to be of environmental origin.

Data analysis

18S rDNA sequences were evaluated using the NCBI database BLAST (Altschul et al. 1990) and identities were assigned based on the highest sequence identity. Sequences with a percent identity below 98% were only identified to genus level. 16S rDNA sequences were evaluated and edited using Geneious® software (Drummond et al. 2012) before being aligned using the SILVA INcremental Aligner (SINA). The SINA alignment was refined further in ARB based on agreement with the consensus and correct molecular folding, to remove non-target sequences including chimeras that are expected from MDA processing (Lasken & Stockwell 2007). Sequences identified as being of mitochondrial or chloroplast origins were also excluded. The remaining sequences were identified using ARB, with a filter limited to the positions amplified (26989 to 42549; *E. coli* SSU 16S rDNA positions 880 to 1408). Sequences were clustered into operational taxonomic unit (OTU) groups at 98% similarity using Mothur (Schloss et al. 2009), and assigned to taxonomic phylotypes based on comparisons to nearest database neighbors. These phylotypes were used in a NodeXL

network analysis; NodeXL (Smith et al. 2009) is a visualization tool to examine how communities are interconnected.

Additional analyses were based on sequence relatedness in a maximum likelihood tree, independently of taxonomic assignments and 98% clustering. Sequences that had been aligned in ARB, in addition to 128 nearest neighbors and an archaeal outgroup (*Geoglobus* sp. SBH6, ACC: FJ216404), were used to generate a phylogenetic tree using RAxML (Stamatakis 2006). The GTRMIX rate distribution model and an ECOLI filter with the positions restricted to 26989 to 42549 (528 base positions total) were employed. ARB default conditions were otherwise employed.

The phylogenetic tree produced by RAxML was used as input data for analyses using the UniFrac package (Lozupone et al. 2006), which provides a set of tools to compare microbial communities based on phylogenetic information. A P test was performed to determine whether any non-stochastic variation existed in the distribution of bacterial sequences among host cells. UniFrac's Principal Coordinates Analysis (PCoA) and environment clustering modules were used to evaluate similarities or differences among bacterial sequence libraries derived from eukaryotic host cells, and the 2 libraries representing free-living and particle-associated bacteria. PCoA is used to assess causal relationships by placing samples in orthogonal, multidimensional space, where each dimension identifies variability in the data in order of most important to least. Environmental clustering is an alternative means to show relationships among environments based on the phylogenetic relatedness of their microbial assemblages. In this case, environments consisted of individual host cells, or the free-living and particle-associated samples.

Repeated sequences

In many analytical approaches used to compare libraries derived from different samples, the composition of each library can be weighted by the number of times a particular sequence occurs. Except for the free-living and particle-associated libraries, sequence data in this study were derived from MDA followed by 2 rounds of PCR. We are not confident that the number of times a given sequence was found in a clone library has any relationship to the number of times the corresponding bacterium appeared on its host cell. Sequence abundance was, therefore, ignored in the community level analyses and the unweighted UniFrac algorithm, which does not

account for sequence abundance, was employed. All analyses described herein were based solely on the presence or absence of either phylotypes or specific sequences in libraries.

RESULTS

Effect of filtration on the numbers of attached bacteria

The numbers of bacteria remaining attached through the filtration process were assessed using DAPI stained cultures of both pennate (from Station ALOHA, test 1) and centric (from Kaneohe Bay, Oahu HI, test 2) diatoms. For test 1, 25 diatoms collected directly from culture (i.e. there was no filtration step) were compared to 30 diatoms that had undergone filtration followed by washing with 500 ml of 0.2 µm filtered seawater. Diatoms lost 55% of their associated bacteria during filtering and washing (*t*-test, p < 0.01) (Fig. 1A). Test 2 (Fig. 1B) examined the effect of washing the filtered diatoms
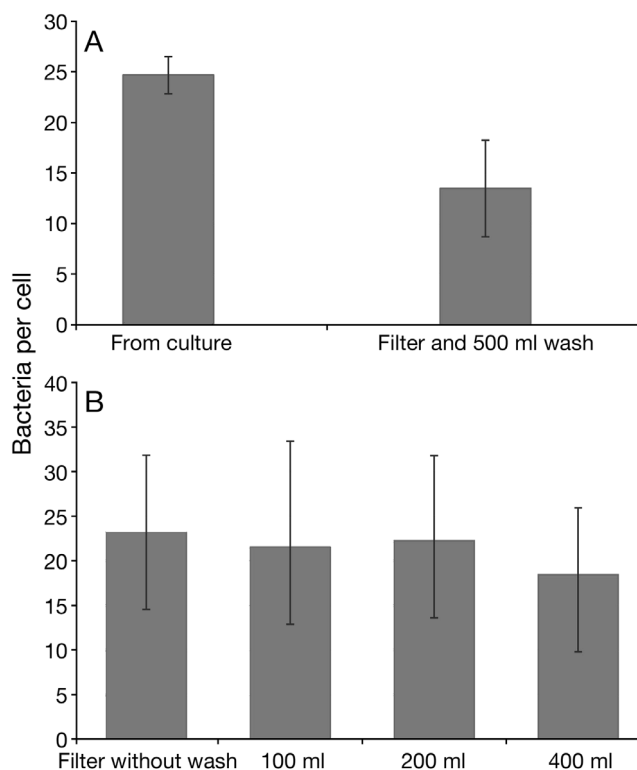

Fig. 1. Effect of filtering and rinsing on the bacteria associated with diatom cells. (A) Test 1: compares cells taken directly from their environment (culture) to cells collected on a filter, then washed. (B) Test 2: compares the effect of additional washing steps following initial capture on a filter

to remove non-attached bacteria from the filter. Although there was a small reduction in the mean number of bacteria per host cell at the highest wash volume tested (400 ml), compared to bacteria present immediately after the filtration process, the loss was not statistically significant ($F = 1.41$, $p = 0.24$).

Because there was no significant effect of washing on bacteria per cell, the 4 treatments can be used to assess the frequency distribution of bacterial cells per host cell (Fig. 2). From 5 to 61 bacterial cells were attached per diatom (mean = 24, $\sigma = 12.66$, n = 61). Visual inspection of cells before and after filtration indicated that the lost cells were probably loosely associated. The bacteria that remained after the initial filtration remain attached even after repeated rinsing. Washing was effective at removing unattached bacteria; few or no bacteria were observed on the filters following washing.

18S DNA sequences

Of the 40 host cells isolated and amplified using FACS and MDA, 31 were centric diatoms (26 *Thalassiosira* spp., 3 *Chaetoceros* spp., 1 *Guinardia sostherfothii* sp., 1 *Leptocylindrus* sp.). Two were pennate diatoms (*Pseudo-nitzschia* sp., *Delphineis* sp.). The remaining host cells included 2 dinoflagellates (*Dinophyceae* sp., *Prorocentrum triestinum*), 2 coccolithophorids (*Calcidiscus leptoporus*, *Emiliania huxleyi*), and 3 flagellates (*Bicosoeca vacillans*, *Isochrysis* sp., *Solenicola setigera*) (GenBank: JX536764– JX536803).

Investigation of 895F primer

The 895F primer was selected for use in this study because of its preferential amplification of non-chloroplast sequences. Hodkinson & Lutzoni (2009) reported that when 895F is compared to sequences in the RDP-II Probe Match analysis, there is a 66.68% sequence coverage for a strict consensus and 91.07% coverage if a single mismatch is allowed. We used Primer Prospector (Walters et al. 2011) to investigate the primer match to the SILVA database. The 895F primer was found to have a strict consensus with 62.01% of the bacterial sequences, 63.54% of the archaeal sequences, and <0.01% of the eukaryal sequences. In all sequences that included the 895F primer, substantial variability occurred at the primer's degenerate positions.

Most bacterial phyla that were found in previous work with diatom hosts would be amplified using the 895F primer under stringent conditions allowing zero mismatches to one of the sequences included in the degenerate 895F primer (Fig. 3A). At the class level, most of the classes reported as associated with diatoms also would be amplified, with the exception of *Sphingobacteria* and *Flavobacteria* (Fig. 3B), which may be underrepresented. At a more specific level, the 895F primer again was found to match to genera and families seen in previous studies of diatom-associated bacteria (*Pseudoalteromonas*, *Alteromonas*, *Flexibacteriaceae*, *Hyphomonas*, *Campylobacter*, and *Roseobacter*), although not all of these were recovered in the present samples. The 895F primer will amplify most of these groups with greater than 98% coverage using a strict consensus, with the exception of the *Flexibacteriaceae* group, for which only 2.2% of the database sequences would be recovered (Fig. 3C). This does not preclude amplification of *Flexibacteriaceae* from our samples, but it may make it less likely. Although the 895F primer does not cover all groups found in previous studies of bacteria associated with diatoms, it does cover the majority.

Cell sorting and MDA evaluation

The SCGC is designed to provide minimally contaminated, amplified material from single cells sorted from environmental samples. However, fol-
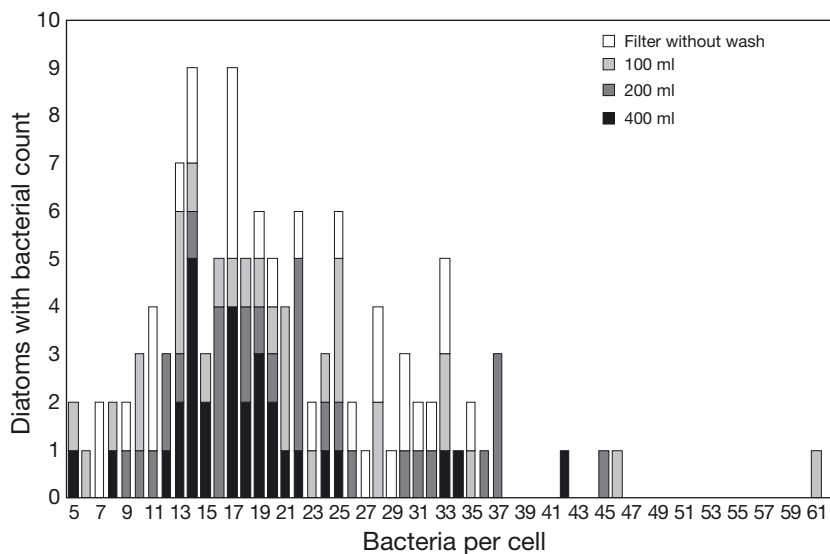

Fig. 2. Frequency distribution of the number of bacterial cells per host diatom cell; the same individual cells as in Fig. 1B. Original treatments are indicated
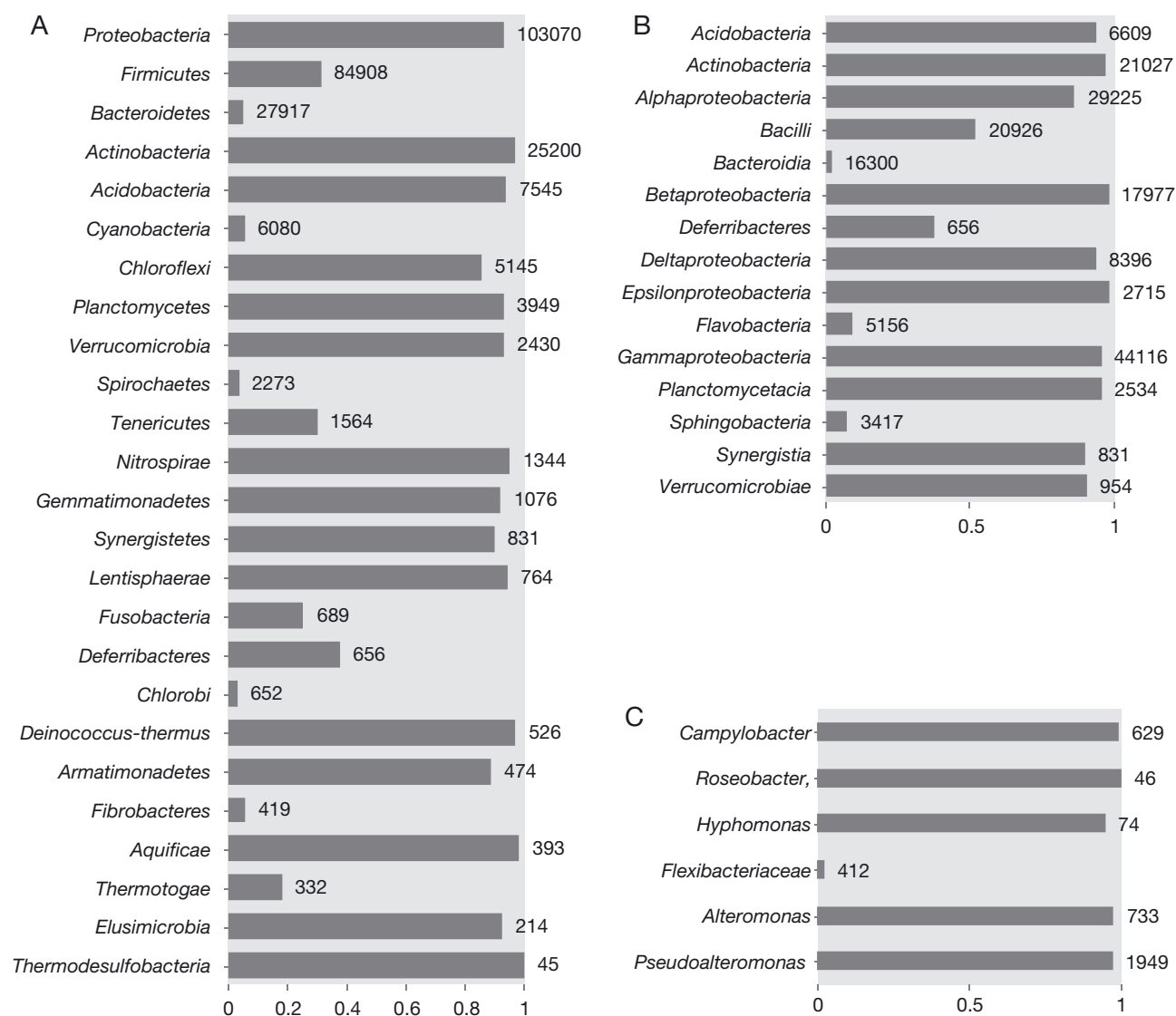
Fig. 3. Primer Prospector output for 895F when compared to the SILVA database using a strict consensus, for (A) phyla, (B) classes, and (C) genera (with the exception of *Flexibacteriaceae*) that were obtained in previous studies of diatom−bacteria interactions. The percent coverage of each group is shown on the *x*-axis and the number of sequences in each group is indicated to the right of each bar

lowing single cell isolation and whole genome ampli-fication, nested PCR presented further opportunities for the inclusion of non-host derived bacterial se-quences. We found *Propionibacterium*, *Staphylococ-cus*, and *Streptococcus* sequences in some negative control wells, all well known to be commonly associ-ated with human skin. No sequence recognized as being environmentally derived was found in nega-tive control wells. As a conservative precaution, all suspected human-associated bacteria were removed prior to any further analysis.

Identity of 16S sequences

A total of 1329 16S rDNA sequences were recov-ered from 40 host cells, 1 sample of particle-associ-ated bacteria, and 1 sample of free-living bacteria (42 libraries in total). Sequences removed prior to statistical analysis included 89 chloroplast, 496 mitochondria, 75 identified in the database as Eukarya, and 102 possible contaminants (sequences with no near database neighbors known to be of environmental origin). An additional 104 sequences

were removed because the amplified fragment matched a 16S rDNA sequence that did not lie within the intended target region positions (i.e. an alternative 16S rDNA target site existed for this primer pair) or because sequences could not be aligned with the ARB database at all (i.e. probable amplification of non-target, non-16S rDNA). The remaining 463 sequences (GenBank: JX536804– JX537266) were closely related to sequences of environmental origin (Baker 2012). When clustered at the 98% similarity level, 95 different OTUs (phylotypes) were identified representing 65 unique taxonomic identities, typically matching existing sequences at the genus level. An arbitrary designation of type 1 and type 2 was used to separate sequences that were grouped into different 98% OTUs under the same taxonomic name in ARB.

A total of 62 different OTUs were identified in the host cell libraries (not including the free-living or particle-associated libraries). No bacterial sequences were recovered from 8 host cells (6 *Thalassiosira*, 1 *Chaetoceros*, and 1 *Emiliania huxleyi*). The 32 remaining host-cell libraries contained 1 to 11 different bacterial phylotypes. The host cell with the highest number of bacterial phylotypes was identified as *Thalassiosira delicatula* (Fig. 4). The total possible diversity of bacteria on a given host is limited by the low number of bacteria per host cell (see 'Effect of filtration on the numbers of attached bacteria'). We assessed whether our sequencing effort was sufficient to retrieve most of the unique bacterial OTUs present in the MDA and PCR-amplified material derived from host cells. Sequence coverage was estimated to be 75% using the S_{ACE} metric and 90% using the Good's C metric (Kemp & Aller 2004). These estimates only represent coverage of the sequences present in the amplified material, not coverage of the diversity of diatom-attached bacteria in the field.

The particle-associated library contained 22 different phylotypes, and 15 different phylotypes were present in the free-living library. We note again that we chose to use a sequencing effort similar to that applied to the host cell libraries, for the purpose of
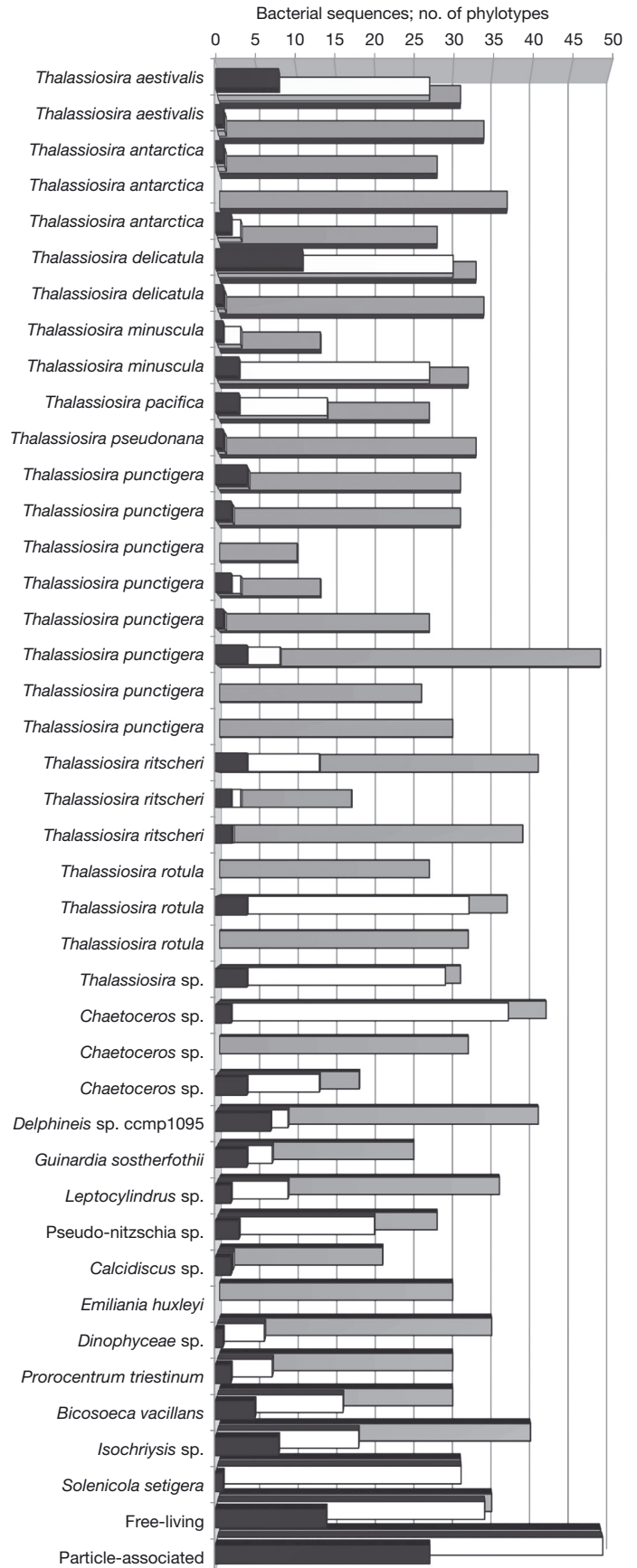

Fig. 4. Bacterial phylotypes per library. From top to bottom: *Thalassiosira* host cell libraries; other diatom host cell libraries; non-diatom host cell libraries; free-living bacterial library; particle-associated bacterial library. Shown are sequence effort per library = grey (back), the number of sequences that remained after non-marine bacterial sequences were removed = white (middle), and total phylotypes per library = dark grey (front)

making a broad comparison; it was not intended to yield an exhaustive coverage of bacterial diversity in these 2 libraries.

Phylotypes shared in different host cell libraries

Network analysis provides a useful visualization of the degree to which the bacterial assemblages on cells are connected (Fig. 5), and is shown for the 34 libraries containing valid bacterial sequences. Of the 34 libraries, 11 contained only a single bacterial phylotype; these host cells included 8 *Thalassiosira* spp., and one of each of *Chaetoceros* sp., *Dinophyceae*, and *Solenicola*. Three *Thalassiosira*-derived libraries did not have any bacterial phylotypes in common with another host cell. All other cells had at least 1 shared phylotype. Two *Arthrobacter* phylotypes were found associated with several host cells (Fig. 5, cluster of cells at center left). Only 1 phylotype was found in both the free-living and particle-associated libraries, and only 3 phylotypes were found in a host cell library and either the free-living or particle-associated library.
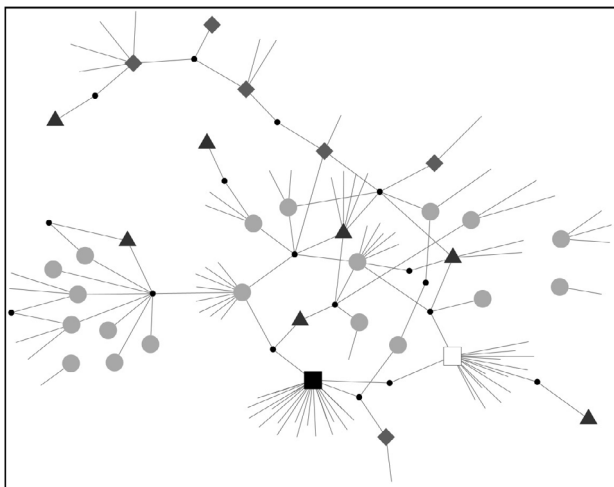

Fig. 5. Network visualization. Libraries are presented by symbol-coded nodes and associated bacterial phylotypes by radiating lines. Lines that touch at their respective ends indicate phylotypes found in 2 or more libraries; the connection points are denoted by small black circles. Libraries with no connections have no bacterial phylotypes shared with another library. The distance between libraries and the length of lines have no meaning. The network plot was generated using NodeXL (http://nodexl.codeplex.com/), an open source software tool to allow visualization of basic relationships between subjects. Circles = *Thalassiosira* host cells; diamonds = other diatoms host cells; triangles: non-diatom host cells; filled square = particle-associated library; open square = free-living bacterial library

There was little evidence at either the phylotype (last ARB ID, Table 1) or class level (Table 2) of specific associations between a bacterial group and a host cell type; most bacterial groups occurred on >1 host cell type. *Arthrobacter* is a possible exception, as it appeared on 10 of the 27 *Thalassiosira* cells and on only 1 other type of host cell (*Prorocentrum triestinum*).

Microbial community comparison using UniFrac

The relationship between host cells and bacteria was found to be non-stochastic (Unifrac P test significance, p < 0.01). We, therefore, proceeded to examine these interactions further using PCoA. The PCoA results suggested that the libraries fall into 3 distinct groups with 1 relatively distant outlier (Fig. 6). Axes 1 through 3 together explain 53% of all variability in library composition. PCoA groupings are often evaluated by relating the most significant axes to contextual information obtained for each sample (e.g. temperature, light, primary production). In the present case, all libraries were drawn from the same water sample and the only contextual information available to distinguish these PCoA groups is the host cell identity and the composition of libraries.

Groups 1 and 3 were primarily distinguished by the presence/absence of *Proteobacteria*: Group 1 by the frequent presence of *Alpha-*, *Beta-*, and *Gammaproteobacteria* and Group 3 by the presence of *Gammaproteobacteria* but absence of *Alpha-* and *Betaproteobacteria*. Group 2 lacked most of the *Proteobacteria* and appeared to be defined by the presence of *Actinobacteria*, particularly *Arthrobacter*. One *Thalassiosira* host cell library (f07) included 1 phylotype each of *Actinobacteria* (*Arthrobacter*) and *Beta-*, *Delta-* and *Gammaproteobacteria*, a combination that did not occur in any other library, and consequently it was placed distant from all other libraries.

All 3 PCoA groups contain *Thalassiosira*-host libraries as well as non-diatom host cells, and 2 groups also contained non-*Thalassiosira* diatoms. Despite having only 1 identical phylotype in common (see Fig. 5), the particle-associated library (Fig. 6, filled square) and the free-living library (Fig. 6, open square) were placed in close proximity in the largest group of libraries, indicating that their constituent bacteria were similar to each other and to host cell libraries in that group, although not identical.

Table 1. Bacterial phylotypes recovered from different host cell types. Two different 98% operational taxonomic units are identified as *Arthrobacter* in ARB, and are treated as separate phylotypes by the notation of type 1 and type 2

Class	Last ARB-ID	*Thalassiosira* spp.	Other diatoms	Other host cells	Particle	Free	Total
Actinobacteria	*Arthrobacter* (type 1)	8	0	1	0	0	9
Actinobacteria	*Arthrobacter* (type 2)	2	0	0	0	0	2
Actinobacteria	*Brachybacterium*_1	3	0	2	0	0	5
Alphaproteobacteria	*Caulobacteraceae*	2	2	2	0	0	6
Alphaproteobacteria	*Mesorhizobium*	0	2	0	0	0	2
Betaproteobacteria	*Delftia*	5	1	1	0	0	7
Betaproteobacteria	*Massilia*_1	0	1	1	0	0	2
Deltaproteobacteria	*Nannocystaceae*	1	0	1	1	0	3
Deltaproteobacteria	*Peredibacter*	1	0	1	0	0	2
Flavobacteria	*Cryomorphaceae*_2, NS7 marine group	0	0	0	1	1	2
Flavobacteria	*Tenacibaculum*_2	0	3	0	0	0	3
Gammaproteobacteria	*Acinetobacter*	1	0	1	0	0	2
Gammaproteobacteria	*Alteromonadaceae*	2	0	0	0	0	2
Gammaproteobacteria	SAR86 clade	1	0	1	0	1	3
Gammaproteobacteria	*Vibrionaceae*	1	1	0	1	0	3
Sphingobacteria	*Sphingobacterium*_1	2	0	0	0	0	2

Table 2. Bacteria at the level of class as recovered from different host cell types

Bacteria (Class ID)	*Thalassiosira* spp.	Other diatoms	Other host cells	Particle	Free	Total
Actinobacteria	12	0	3	0	1	16
Alphaproteobacteria	5	4	2	1	1	13
Gammaproteobacteria	6	1	2	1	1	11
Betaproteobacteria	4	2	2	0	0	8
Deltaproteobacteria	2	1	2	1	1	7
Flavobacteria	0	3	0	1	1	5
Cytophagia	1	0	0	1	0	2
Clostridia	1	0	1	0	0	2
Sphingobacteria	2	0	0	0	0	2
Bacilli	1	1	0	0	0	2
Acidobacteria	0	0	0	0	1	1
Lentisphaeria	0	0	0	1	0	1
Planctomycetacia	0	0	0	1	0	1
Synergistetes	1	0	0	0	0	1
Verrucomicrobiae	0	0	0	1	0	1

The defining characteristic of the PCoA groupings is that libraries within a group contain sequences that are more closely related within the group, than they are to sequences in other groups. This was recapitulated in the environmental clustering analysis (Fig. 7), which demonstrated that the observed PCoA groups can be recreated based on the phylogenetic relatedness of their bacterial assemblages.

DISCUSSION

Because this study was exploratory in nature, no attempt was made to manipulate or predetermine the host species present in the field sample. The sampling was not designed to be representative of all host populations nor representative of the diatom populations most often seen at Station ALOHA or elsewhere in the world. Our data set was dominated by a single diatom genus (*Thalassiosira*), and, therefore, is well suited for examining variability in the bacterial assemblages attached to closely related host cells. Specifically, we compared *Thalassiosira*-associated bacterial assemblages to those associated with other diatoms and other eukaryotes that were less prevalent in the sample, as well as to the free-living bacteria and particle-associated bacterial assemblages recovered from the same water sam-
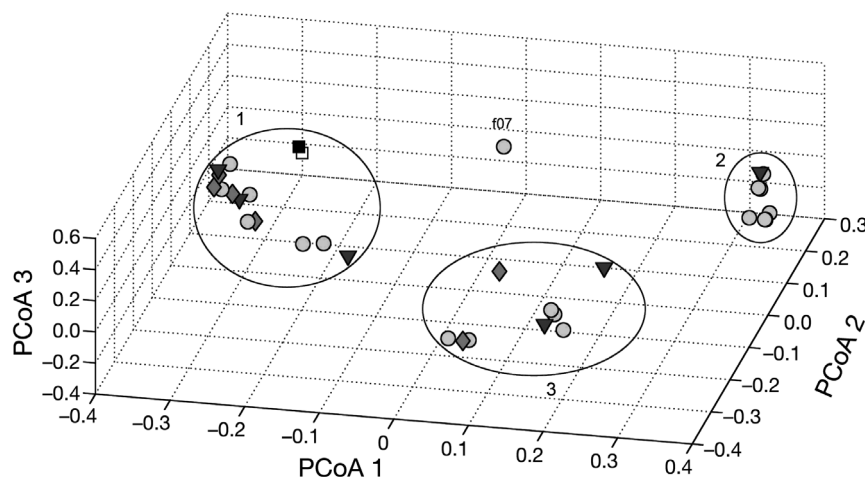
Fig. 6. Principal Coordinates Analysis. 53% of variation is explained by the first 3 axes, which divide the 34 libraries into 3 groups with one outlier. Of the 3 axes, 33% of variation is explained by PCoA axis 1, 11% by PCoA axis 2, and 9% by PCoA axis 3. Host cell f07 (*Thalassiosira deliculata*) does not fall into any of the 3 groups. See Fig. 5 for explanation of symbols

The unusual nature of the sample material, i.e. bacterial 16S rDNA in the presence of abundant host-derived plastid 16S rDNA, required a novel amplification strategy using a 16S rDNA primer developed for a very different environmental context (Hodkinson & Lutzoni 2009). Using this primer, the representation of bacterial 16S rDNA in sequences recovered from MDA-amplified material increased to 51% from 5% using conventional 16S rDNA primers.

Host cell libraries contained from 1 to 11 different bacterial phylotypes, with the majority of host cells containing >1 bacterial phylotype. The number of different phylotypes in a host-cell library is a minimum estimate of the number of bacterial cells that occurred on the host cell, but with the methods used there was no direct evidence for how many bacterial cells were attached to each of the host cells examined. As a specific example, no bacterial phylotypes were recovered from 8 eukaryotic cells, but we can-

ples. We were also able to test whether these ob-served similarities or differences were non-random, and whether they were related to host cell identity. Finally, we assessed whether any repeated consortia of bacteria occurred, and conversely whether any bacteria occurred to the exclusion of others.
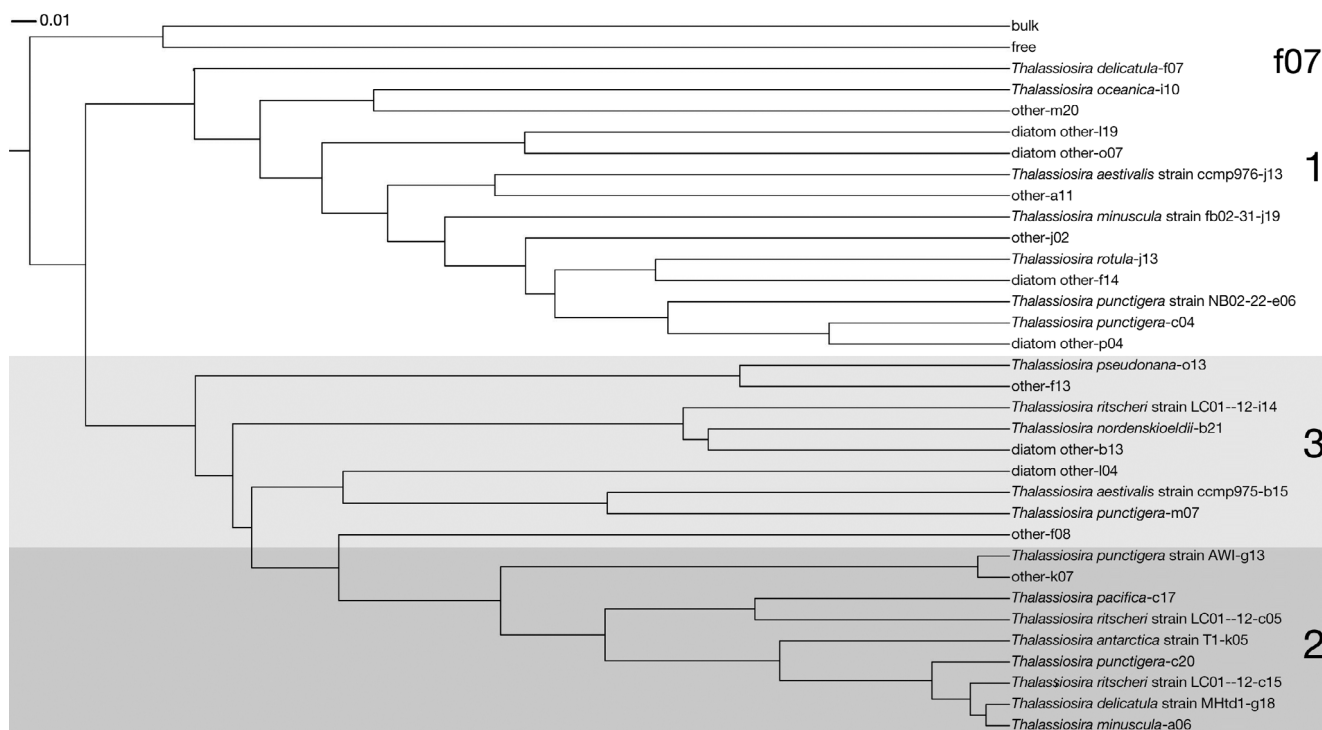

Fig. 7. Clustering of libraries based on phylogenetic relatedness of the bacterial phylotypes they contain, calculated via a jack-knife resampling procedure using the UniFrac metric. Groups 1, 2, and 3 as defined by the PCoA analysis are marked, as is a single host cell (f07) that did not fall into groups 1–3. The scale bar represents 0.01 branch length unit

not say with certainty that no bacteria were attached to those cells. We can say that the number of bacterial phylotypes recovered was not correlated with the sequencing effort per host cell library (data not shown); therefore, a lack of phytotypes associated with some host cells was not merely the result of a lack of effort.

Coverage estimates indicate that recovery of amplified sequences was good (75% or 90%, depending on the metric used) but was not complete. While each host cell library was given similar sequencing effort, the 895F primer does not amplify all bacterial groups equally; we identified a few bacterial taxa that are likely to be underrepresented. These caveats are considered in the following discussion of taxa present or absent in our sample.

Some of the bacteria we observed in association with host cells are similar to those found in previous studies. At the level of classes, our results are very similar to previous work by Grossart et al. (2005), e.g. diatoms hosted a variety of *Flavobacteria*, *Sphingobacteria*, *Alphaproteobacteria*, and *Gammaproteobacteria* (the latter 2 were found on non-diatom hosts as well). As noted earlier, *Flavobacteria* and *Sphingobacteria* may be underrepresented in libraries created with the 895F primer. *Vibrio* were found on diatoms and in the particle-associated library, and have been noted in previous studies (Bidle & Azam 2001). The consistency of our results with previous reports suggests some similarities between culture-based data and those obtained from our single-cell based analysis of a natural population.

Although many of the bacterial orders and classes found in previous studies were observed in our sample, we did not detect others. In a recent review, Amin et al. (2012) summarized bacterial taxa that have been reported in association with diatoms (their Fig. 2), primarily from studies in which the total sequencing effort was comparable to or smaller than in the present study. We did not find members of the *Rhodobacteraceae*, which have been observed in association with diatoms (e.g. Guannel et al. 2011). We did not observe members of the *Flexibacteriaceae* that were previously found in cultures of *Thalassiosira rotula* (Grossart et al. 2005) grown in a relatively high-nutrient medium. *Pseudoalteromonas* and *Alteromonas* have also been shown to be associated with diatoms (Bidle & Azam 2001), but were not found in this study. We also did not find *Campylobacter*, an *Epsilonproteobacteria* found in association with dinoflagellates (Riemann et al. 2000), or *Cytophagales* found in association with *Thalassiosira* (Riemann et al. 2000, Bidle et al. 2003). Amin et al.

(2012) identified *Bacteroidetes* as one of the main heterotrophic bacterial phyla to be associated with diatoms. We observed *Bacteroidetes* on 1 *Thalassiosira* host cell and in the particle-associated library, but not in 32 other libraries.

The absence of *Flexibacteriaceae* and cyanobacteria, and the observed low incidence of *Bacteroidetes* might have been caused by primer bias (see Fig. 4), but the rest of the aforementioned groups were readily amplified by the 895F primer. Given that our sample included similar diatom hosts and that we applied a comparable or greater total sequencing effort, the most parsimonious explanation is that these remaining bacterial groups were not as prevalent in our sample as in previous published reports. This may be due to differences in the host-cell microenvironment (e.g. physiological status of the host cells), in bacterial seed populations at different sampling sites, or to a fundamental difference in the results obtained with bulk versus single-cell approaches.

Previous studies have commented on the differences between bacteria associated with host cells, and those present in the free-living and particle-associated bacterial assemblages. We found only a few phylotypes occurred in host cell libraries and also in the free-living or the particle-associated libraries. However, free-living and particle-associated libraries fell within the same PCoA group and environmental cluster (Group 1) as a number of host cell libraries, indicating that their associated bacteria were phylogenetically related although not identical. Our data probably underrepresents the diversity of free-living and particle-associated bacteria, but that does not obviate our finding that the composition of these libraries resembled that of the largest group of host-cell libraries, given a similar level of effort.

Previous studies have suggested that there are consistent associations between bacterial taxa and diatom hosts, which imply that some degree of host specificity may occur in cultures, and perhaps in the field (e.g. Schäfer et al. 2002, Amin et al. 2012). Rather than host specificity, the present study demonstrated strong variations in the bacterial associates of congeneric host cells obtained from the same water sample. The one possible exception is the association between *Arthrobacter* and host cells assigned to PCoA Group 2, which consisted of 8 *Thalassiosira* cells and a single dinoflagellate (*Prorocentrum triestinum*). The present data offer no insight into the possible functional causes or consequences of the observed differences in the bacterial assemblages attached to host cells. However, given the

nearly complete separation of *Arthrobacter* from other bacteria on host cells, we can speculate that some bacteria may be able to colonize host cells to the exclusion of other bacteria. The remainder of the *Thalassiosira* spp. cells varied greatly with respect to their bacterial associates, and *Thalassiosira* spp. appeared in all 3 PCoA groups. Host cells in PCoA Groups 1 and 3 were much more diverse in their bacterial assemblages, but were distinguished by the absence of *Alphaproteobacteria* and *Betaproteobacteria* from one PCoA group and their presence in the other. *Alphaproteobacteria*, *Betaproteobacteria*, and *Arthrobacter* are obvious candidates for further study regarding the construction of the bacteria assemblages attached to diatom host cells.

The single-cell microenvironment is unlike other marine environmental samples and required considerable method development and testing. The following provides a very brief summary of lessons learned in the present study. (1) The combination of shipboard sampling, filtering, and flow cytometry to obtain single host cells was effective, but required numerous negative controls to recognize and avoid contamination. (2) Bacteria resident on host cells appeared to be firmly attached, and simple washing on filters was sufficient to remove non-attached bacteria. (3) MDA processing resulted in a relatively high proportion of chimeras. Recognizing this possibility, sequencing effort should be adjusted to accommodate discarding chimeric sequences. (4) 16S rDNA primers amplify chloroplast and mitochondrial 16S rDNA. In the unique environment of a photosynthetic eukaryotic host cell, the amplified plastid 16S rDNA may overwhelm the valid bacterial sequences. The 895F primer strongly selects against plastid sequences and yields an approximately 10-fold increase in the number of valid bacterial sequences, but non-target sequences must still be filtered out. (5) A thorough examination of primer bias indicated that the 895F primer recovers most bacterial taxa well, but it selects against a few taxa including cyanobacteria. We have not attempted to modify the published 895F primer, but it may be possible to adjust it to decrease or otherwise manipulate these biases to advantage. (6) Sequence coverage in our study was relatively high, in part because the diversity of bacteria attached to host cells was low. If a more exhaustive coverage is required, we would recommend increasing the library size per host cell, or moving to multiplexed high-throughput sequencing. The latter may be cost- and time-effective in comparison to creating numerous, larger libraries.

The next steps are to (1) expand upon our data set to better characterize patterns in bacterial assemblages associated with single diatom cells, and (2) examine whether bacterial assemblages attached to a host cell are truly microbial communities sensu Clements (1916), who defined a community as possessing 'a well-defined level of organization with tight interactions among organisms that comprise a causal system and gives rise to emergent properties' (paraphrasing by Konopka 2009, p. 1223). In our view, if diatom–bacteria associations display such properties, they may be considered metaorganisms (Bosch & McFall-Ngai 2011). Our long-term goal is to understand whether diatoms and their bacterial associates operate as metaorganisms, by identifying characteristic diatom-associated microbial communities, examining their functional interactions and understanding their emergent properties.

LITERATURE CITED

Amin SA, Parker MS, Armbrust EV (2012) Interactions between diatoms and bacteria. Microbiol Mol Biol Rev 76:667–684

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Azam F, Fenchel T, Field JG, Gray JS, Meyer-Reil LA, Thingstad F (1983) The ecological role of water-column microbes in the sea. Mar Ecol Prog Ser 10:257–263

Baker LB (2012) Exploring the bacteria-diatom metaorganism using single-cell whole genome amplification. MS Thesis, University of Hawaii at Manoa, HI

Bertrand EM, Saito MA, Jeon YJ, Neilan BA (2011) Vitamin B_{12} biosynthesis gene diversity in the Ross Sea: the identification of a new group of putative polar B_{12} biosynthesizers. Environ Microbiol 13:1285–1298

Bidle KD, Azam F (1999) Accelerated dissolution of diatom silica by marine bacterial assemblages. Nature 397:508–512

Bidle KD, Azam F (2001) Bacterial control of silicon regeneration from diatom detritus: significance of bacterial ectohydrolases and species identity. Limnol Oceanogr 46:1606–1623

Bidle KD, Brzezinski MA, Long RA, Jones JL, Azam F (2003) Diminished efficiency in the oceanic silica pump caused

by bacteria-mediated silica dissolution. Limnol Oceanogr 48:1855–1868

Blackburn N, Fenchel T, Mitchell J (1998) Microscale nutrient patches in planktonic habitats shown by chemotactic bacteria. Science 282:2254–2256

Bosch TCG, McFall-Ngai MJ (2011) Metaorganisms as the new frontier. Zoology 114:185–190

Clements FE (1916) Plant succession, an analysis of the development of vegetation. Carnegie Institute, Washington, DC, Publ. no. 242

Cole JJ (1982) Interactions between bacteria and algae in aquatic ecosystems. Annu Rev Ecol Evol Syst 13:291–314

Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG (2005) Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature 438:90–93

Cullen JJ (1982) The deep chlorophyll maximum comparing vertical profiles of chlorophyll *a*. Can J Fish Aquat Sci 39: 791–805

DeLong EF, Franks DG, Alldredge AL (1993) Phylogenetic diversity of aggregate-attached versus free-living marine bacterial assemblages. Limnol Oceanogr 38:924–934

Desai JD, Banat IM (1997) Microbial production of surfactants and their commercial potential. Microbiol Mol Biol Rev 61:47–64

Desbois AP, Lebl T, Yan L, Smith V (2008) Isolation and structural characterization of two antibacterial free fatty acids from the marine diatom, *Phaeodactylum tricornutum.* Appl Microbiol Biotechnol 81:755–764

Dore JE, Letelier RM, Church MJ, Lukas R, Karl DM (2008) Summer phytoplankton blooms in the oligotrophic North Pacific Subtropical Gyre: historical perspective and recent observations. Prog Oceanogr 76:2–38

Droop MR (2007) Vitamins, phytoplankton and bacteria: symbiosis or scavenging? J Plankton Res 29:107–113

Drummond AJ, Ashton B, Buxton S, Cheung M and others (2012) Geneious v56, available from www.geneious.com

Gram L, Melchiorsen J, Bruhn JB (2010) Antibacterial activity of marine culturable bacteria collected from a global sampling of ocean surface waters and surface swabs of marine organisms. Mar Biotechnol (NY) 12:439–451

Grossart HP (1999) Interactions between marine bacteria and axenic diatoms (*Cylindrotheca fusiformis*, *Nitzschia laevis*, and *Thalassiosira weissflogii*) incubated under various conditions in the lab. Aquat Microb Ecol 19:1–11

Grossart HP (2010) Ecological consequences of bacterioplankton lifestyles: changes in concepts are needed. Environ Microbiol Rep 2:706–714

Grossart HP, Levold F, Allgaier M, Simon M, Brinkhoff T (2005) Marine diatom species harbor distinct bacterial communities. Environ Microbiol 7:860–873

Grossart HP, Czub G, Simon M (2006) Algae-bacteria interactions and their effects on aggregation and organic matter flux. Environ Microbiol 8:1074–1084

Grossart HP, Tang KW, Kiørboe T, Ploug H (2007) Comparison of cell-specific activity between free-living and attached bacteria using isolates and natural assemblages. FEMS Microbiol Lett 266:194–200

Guannel ML, Horner-Devine MC, Rocap G (2011) Bacterial community composition differs with species and toxigenicity of the diatom *Pseudo-nitzschia.* Aquat Microb Ecol 64:117–133

Guillard RRL (1975) Culture of phytoplankton for feeding marine invertebrates. In: Smith WL Chanley MH (eds) Culture of marine invertebrate animals. Plenum Press, New York, NY, p 26–60

Guillard RRL, Ryther JH (1962) Studies of marine planktonic diatoms. I. *Cyclotella nana* Hustedt and *Detonula confervacea* Cleve. Can J Microbiol 8:229–239

Hodkinson BP, Lutzoni F (2009) A microbiotic survey of lichen-associated bacteria reveals a new lineage from the Rhizobiales. Symbiosis 49:163–180

Huisman J, Weissing FJ (1999) Biodiversity of plankton by species oscillations and chaos. Nature 402:407–410

Huisman J, Nga NPT, Karl DM, Sommeijer B (2006) Reduced mixing generates oscillations and chaos in the oceanic deep chlorophyll maximum. Nature 439:322–325

Kaczmarska I, Ehrman JM, Bates SS, Green DH, Leger C, Harris J (2005) Diversity and distribution of epibiotic bacteria on *Pseudo-nitzschia multiseries* (Bacillariophyceae) in culture, and comparison with those on diatoms in native seawater. Harmful Algae 4:725–741

Karl DM (1999) A sea of change: biogeochemical variability in the North Pacific Subtropical Gyre. Ecosystems 2: 181–214

Karl DM (2002) Nutrient dynamics in the deep blue sea. Trends Microbiol 10:410–418

Karl DM, Church MJ, Letelier RM, Mahaffey C (2012) Predictable and efficient carbon sequestration in the North Pacific Ocean supported by symbiotic nitrogen fixation. Proc Natl Acad Sci USA 109:1842–1849

Kemp PF, Aller JY (2004) Bacterial diversity in aquatic systems: what can we learn from 16S rDNA libraries? FEMS Microbiol Ecol 47:161–177

Kobayashi K, Takata Y, Kodama M (2009) Direct contact between *Pseudo-nitzschia multiseries* and bacteria is necessary for the diatom to produce a high level of domoic acid. Fish Sci 75:771–776

Kogure K, Simidu U, Taga N (1981) Bacterial attachment to phytoplankton in sea water. J Exp Mar Biol Ecol 56: 197–204

Konopka A (2009) What is microbial community ecology? ISME J 3:1223–1230

Lasken RS, Stockwell TB (2007) Mechanism of chimera formation during the Multiple Displacement Amplification reaction. BMC Biotechnol 7:19

Lozupone C, Hamady M, Knight R (2006) UniFrac — an online tool for comparing microbial community diversity in a phylogenetic context. BMC Bioinformatics 7:371

Ludwig W, Strunk O, Westram R, Richter L and others (2004) ARB: a software environment for sequence data. Nucleic Acids Res 32:1363–1371

Mayali X, Azam F (2004) Algicidal bacteria in the sea and their impact on algal blooms. J Eukaryot Microbiol 51: 139–144

Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71:491–499

Mitchell JG, Pearson L, Bonazinga A, Dillon S, Khouri H, Paxinos R (1995) Long lag times and high velocities in the motility of natural assemblages of marine bacteria. Appl Environ Microbiol 61:877–882

Pruesse E, Quast C, Knittel K, Fuchs BM and others (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 35: 7188–7196

Riemann L, Steward GF, Azam F (2000) Dynamics of bacterial community composition and activity during a mesocosm diatom bloom. Appl Environ Microbiol 66: 578–587

Rooney-Varga JN, Giewat MW, Savin MC, Sood S, LeGres-
ley M, Martin JL (2005) Links between phytoplankton
and bacterial community dynamics in a coastal marine
environment. Microb Ecol 49:163–175

Rosowski JR (1992) Specificity of bacterial attachment sites
on the filamentous diatom *Navicula confervacea* (Bacil-
lariophyceae). Can J Microbiol 38:676–686

Sapp M, Schwaderer S, Wiltshire KH, Hoppe HG, Gerdts G,
Wichels A (2007a) Species-specific bacterial communi-
ties in the phycosphere of microalgae? Microb Ecol 53:
683–699

Sapp M, Wichels A, Gerdts G (2007b) Impacts of cultivation
of marine diatoms on the associated bacterial commu-
nity. Appl Environ Microbiol 73:3117–3120

Schäfer H, Abbas B, Witte H, Muyzer G (2002) Genetic
diversity of 'satellite' bacteria present in cultures of mar-
ine diatoms. FEMS Microbiol Ecol 42:25–35

Scharek R, Latasa M, Karl DM, Bidigare RR (1999) Temporal
variations in diatom abundance and downward vertical
flux in the oligotrophic North Pacific gyre. Deep-Sea Res
I 46:1051–1075

Schloss PD, Westcott SL, Ryabin T, Hall JR and others (2009)
Introducing mothur: open-source, platform-indepen-
dent, community-supported software for describing and
comparing microbial communities. Appl Environ Micro-
biol 75:7537–7541

Smith DC, Steward GF, Long RA, Azam F (1995) Bacterial
mediation of carbon fluxes during a diatom bloom in a
mesocosms. Deep-Sea Res II 42:75–97

Smith M, Shneiderman B, Milic-Frayling N, Rodrigues EM
and others (2009) Analyzing (social media) networks
with NodeXL. Proc 4th Int Conf Communities and Tech-
nologies, University Park, PA, Jun 25–27, 2009, ACM,
New York, NY, p 255–264

Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-
based phylogenetic analyses with thousands of taxa and
mixed models. Bioinformatics 22:2688–2690

Stepanauskas R, Sieracki ME (2007) Matching phylogeny
and metabolism in the uncultured marine bacteria,
one cell at a time. Proc Natl Acad Sci USA 104:
9052–9057

Stewart JE, Marks LJ, Wood CR, Risser SM, Gray S (1997)
Symbiotic relations between bacteria and the domoic
acid producing diatom *Pseudo-nitzschia multiseries* and
the capacity of these bacteria for gluconic acid/glucono-
lactone formation. Aquat Microb Ecol 12:211–221

Swan BK, Martinez-Garcia M, Preston CM, Sczyrba A and
others (2011) Potential for chemolithoautotrophy among
ubiquitous bacteria lineages in the dark ocean. Science
333:1296–1300

Turner S, Pryer KM, Miao VPW, Palmer JD (1999) Investi-
gating deep phylogenetic relationships among cyano-
bacteria and plastids by small subunit rRNA sequence
analysis. J Eukaryot Microbiol 46:327–338

Vink CJ, Thomas SM, Paquin P Hayashi CY, Hedin M (2005)
The effects of preservatives and temperatures on arach-
nid DNA. Invertebr Syst 19:99–104

Waksman SA, Butler MR (1937) Relation of bacteria to
diatoms in sea water. J Mar Biol Assoc UK 22:359–373

Walters WA, Caporaso JG, Lauber CL, Berg-Lyons D, Fierer
N, Knight R (2011) PrimerProspector: de novo design and
taxonomic analysis of PCR primers. Bioinformatics 27:
1159–1161

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S
ribosomal DNA amplification for phylogenetic study.
J Bacteriol 173:697–703

Yan L, Boyd KG, Burgess JG (2002) Surface attachment
induced production of antimicrobial compounds by mar-
ine epiphytic bacteria using modified roller bottle culti-
vation. Mar Biotechnol 4:356–366

Zhu F, Massana R, Not F, Marie D, Vaulot D (2005) Mapping
of picoeucaryotes in marine ecosystems with quantitative
PCR of the 18S rRNA gene. FEMS Microbiol Ecol 52:
79–92