Integrated metagenomic and metaproteomic analyses of marine biofilm communities

Dagmar H. Leary, Robert W. Li, Leila J. Hamdan*, W. Judson Hervey IV, Nikolai Lebedev, Zheng Wang, Jeffrey R. Deschamps, Anne W. Kusterbeck and Gary J. Vora*

*Corresponding author. Email: gary.vora@nrl.navy.mil

Metagenomic and metaproteomic analyses were utilized to determine the composition and function of complex air–water interface biofilms sampled from the hulls of two US Navy destroyers. Prokaryotic community analyses using PhyloChip-based 16S rDNA profiling revealed two significantly different and taxonomically rich biofilm communities (6,942 taxa) in which the majority of unique taxa were ascribed to members of the Alphaproteobacteria, Gammaproteobacteria and Clostridia. Although metagenomic sequencing indicated that both biofilms were dominated by prokaryotic sequence reads (> 91%) with the majority of the bacterial reads belonging to the Alphaproteobacteria, the Ship-1 metagenome harbored greater organisational and functional diversity and was comparatively enriched for sequences from Cyanobacteria, Bacteroidetes and macroscopic eukaryotes, whereas the Ship-2 metagenome was enriched for sequences from Proteobacteria and microscopic photosynthetic eukaryotes. Qualitative liquid chromatography-tandem mass spectrometry metaproteome analyses identified 678 unique proteins, revealed little overlap in species and protein composition between the ships and contrasted with the metagenomic data in that ~80% of classified and annotated proteins were of eukaryotic origin and dominated by members of the Bacillariophyta, Cnidaria, Chordata and Arthropoda (data deposited to the ProteomeXchange, identifier PXD000961). Within the shared metaproteome, quantitative $^{18}$O and iTRAQ analyses demonstrated a significantly greater abundance of structural proteins from macroscopic eukaryotes on Ship-1 and diatom photosynthesis proteins on Ship-2. Photosynthetic pigment composition and elemental analyses confirmed that both biofilms were dominated by photosynthetic processes. These data begin to provide a better understanding of the complex organismal and biomolecular composition of marine biofilms while highlighting caveats in the interpretation of stand-alone environmental ‘-omics’ datasets.

Keywords: biofouling; diatoms; metagenome; metaproteome; quantitative proteomics; ship hull

Introduction

The recent advent of large-scale and culture-independent ‘-omic’ measurements have contributed to overcoming the current inability to culture the vast majority of microbial species (Amann et al. 1995; Woese 1996), and as a result, have become powerful tools for understanding the composition, potential and function of microbial assemblages. In the marine environment, stand-alone metagenomic (DeLong et al. 2006; Rusch et al. 2007), metatranscriptomic (Hewson et al. 2010; Poretsky et al. 2009) and metaproteomic (Morris et al. 2010; Sowell et al. 2011) analyses of bacterioplankton communities have fundamentally changed the view of prokaryotic taxonomy, biogeography and biogeochemical activity in the photic zone and have only been bolstered by further elaborations using integrated ‘-omic’ datasets (Frias-Lopez et al. 2008; Gilbert et al. 2008; Shi et al. 2011; Grzymski et al. 2012; Williams et al. 2012). These findings, combined with the previous absence of tools capable of deconvoluting the inherent complexity of physically associated multi-species communities (Camps et al. 2014), suggest that the application of ‘-omics’ methods for the interrogation of marine biofilms can now also aid in providing a better biomolecular understanding of these assemblages while simultaneously assessing the strengths and limitations of current analytical technologies.

While marine biofilm communities are not well studied in general, those that form on artificial substrata with antifouling coatings are particularly poorly understood (Yebra et al. 2006; Salta et al. 2013; Camps et al. 2014). This is somewhat surprising given the economic and operational impact of microbial soft fouling communities (Schultz et al. 2011; Lejars et al. 2012) and appears to be largely due to the use of culture-dependent methods that have constrained these analyses and vastly underestimated the organismal composition and function within these biofilms (Camps et al. 2014).

In this study culture-independent methods such as PhyloChip microbial profiling, metagenomic sequencing, qualitative and quantitative nano-flow liquid chromatography-tandem mass spectrometry (LC-MS/MS) metaproteomic analyses, elemental analyses and fluorescence
spectroscopy were utilized to begin to determine the composition and function of ship hull air–water interface marine biofilms formed in different environments. Despite being harvested from identical antifouling-coated substrata, the findings (1) elucidate the true biological complexity of mature soft fouling communities from operational vessels; (2) identify qualitative and quantitative differences in their organismal and biomolecular composition; and (3) suggest exercising caution when utilizing single ‘-omic’ datasets for the description of complex environmental microbial assemblages.

Materials and methods

Sample collection

A total of ~ 44 g of wet marine biofilm samples were harvested from the hulls of two Arleigh Burke-class destroyers [USS Laboon (DDG-58, Ship-1) and USS Bainbridge (DDG-96, Ship-2)] at Norfolk Naval Station, Norfolk, VA, USA in August 2010. The entire underwater hulls of both ships were coated with Intertuf 262 – KH Series/ KHA062 epoxy anticorrosive and controlled depletion polymer Interspeed 640 – BRA series (BRA640/BRA642) polishing antifouling (International Paint LLC, San Diego, CA, USA) and were sampled after a seven-month deployment (January–July 2010). While both deployments started and concluded in Norfolk, VA, Ship-1 traveled to ports in the North and Baltic Seas whereas Ship-2 traveled to Rota, Spain. Multiple biofilm samples with a Fouling Rating of 20 (ie ‘advanced slime’ – Naval Ships’ Technical Manual 2006) were scraped from the air–water interface (starboard, mid-ship) using sterile costar cell lifters (Corning, Tewksbury, MA, USA) (Figure 1a, b). Each sample was collected at a distance of ~ 1 m from the previous sample and immediately snap frozen in sterile 50 ml conical tubes using an EtOH-dry ice bath. Surface water samples (1 l from a depth of 0.1 m) were also collected at a distance of ~ 1 m from the previous sample and immediately snap frozen in sterile 50 ml conical tubes using an EtOH-dry ice bath. Surface water samples (1 l from a depth of 0.1 m) were also collected at a distance of ~ 5 m from the hulls and all planktonic material was captured on 0.22 μm UltraClean Water DNA Filters (MO BIO Laboratories, Inc., Carlsbad, CA, USA) and snap frozen. Upon returning to the laboratory, the collected biofilm samples were used as the source material for all genomic, proteomic, element and pigment analyses (Figure 1c).

PhyloChip analysis of biofilm 16S rDNA

Metagenomic DNA was extracted from five biofilm samples (0.15 g aliquots in duplicate using two processing variations of the PowerBiofilm™ DNA Isolation Kit (MO BIO Laboratories, Inc., Carlsbad, CA, USA)) and two ship proximal filtered surface water samples and shipped to Second Genome Inc. (San Bruno, CA, USA) for prokaryotic 16S rDNA profiling via PhyloChip microarray (DeSantis et al. 2007; Hazen et al. 2010). Bacterial 16S rRNA genes were amplified by PCR using the degenerate forward primer: 27F 5′-AGGTTTATCMTTGCCTCA-G-3′ and the non-degenerate reverse primer: 1492R 5′-GTTACCTTGTTACGACTT-3′ (Lane 1991). All amplified products were quantified by electrophoresis using an Agilent 2100 Bioanalyzer (Agilent, Palo Alto, CA, USA) and then fragmented, biotin labeled and hybridized to the PhyloChip™ Array, version G3. Each array was washed, stained and scanned using a GeneArray® scanner (Affymetrix Inc., Santa Clara, CA, USA) and the hybridization values and fluorescence intensity for each taxon were calculated as a trimmed average. An operational taxonomic unit (OTU) was defined by 99% similarity. The PhyCA-Stats™ analysis software package was used for multivariate statistical analysis of the data.

Biofilm metagenome sequencing

Three metagenomic DNA extractions from each ship’s hull were pooled, subjected to integrity testing using a 2100 Bioanalyzer (Agilent) and quantified. Approximately 1.0 μg of high-quality metagenomic DNA was processed using an Illumina TruSeq DNA sample prep kit (Illumina, Inc., San Diego, CA, USA) and final individual libraries were validated and pooled based on their respective 6-bp adaptors and sequenced at 100 bp/sequence read using an Illumina HiSeq 2000 sequencer. Raw sequence reads were first trimmed using SolexaQA (Cox et al. 2010) and the mean sequence read length after trimming was reduced to 92.7 bp (Ship-1) and 91.8 bp (Ship-2). Reads of mammalian origin (potential contaminants) were removed using alignment tools against the sequenced genomes of humans and rodents. The resultant quality sequences were assembled de novo using SOAPdenovo (Li et al. 2009) and Velvet software. Open reading frames (ORFs) were predicted from all contiguous sequences > 200 bp using FragGeneScan (v.1.14) (Rho et al. 2010) and annotation was carried out using BLASTP against the SwissProt database (2011_07). Finally, functional annotations were performed according to Clusters of Orthologous Groups of proteins (COG), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Protein families (Pfam) databases as previously described (Li et al. 2012). Both whole genome shotgun projects have been deposited at DDBJ/EMBL/GenBank under National Center for Biotechnology Information (NCBI) Sequence Read Archive Accession #SRA058536, experiments SRX185752, SRX185753 and samples SRS361508, SRS361509.

Sample preparation and qualitative metaproteomics

Biofilm samples (equaling ~ 10 mg of protein as determined by amino acid analyses (AAA)) were ground by mortar and pestle in the presence of liquid nitrogen and
mixed with extraction buffer (100 mM Tris HCl buffer pH 8.8, 10 mM EDTA, 5 mM DTT, 0.9 M sucrose) and placed in a water bath sonicator for 10 min. Proteins were then extracted using the modified phenol protein extraction method (Leary et al. 2012). The protein pellets were dissolved in 2% SDS and aliquots of the resulting protein solutions were analyzed by AAA. Extracted proteins (50 μg) were separated by 1D-SDS-PAGE, digested in-gel using trypsin and analyzed by LC-MS/MS as previously described (Leary et al. 2012). Briefly, gel bands were destained, treated with trypsin overnight and peptides from each gel band were extracted with acetonitrile acidified by formic acid (FA) and dried by centrifugal evaporation. Samples were reconstituted in solvent A (0.1% FA, 2% acetonitrile in high performance liquid chromatography (HPLC) water) and analyzed using a TempoMDLC nano-flow liquid chromatography system coupled to a Q-Star Elite mass spectrometer (AB Sciex, Foster City, CA, USA). Each biofilm was sampled and analyzed in triplicate. The resulting spectra from each band in the same sample lane were merged and searched against the in-house biofilm database as well as UniMES (cluster_100, July 2008) combined with UniRef100 (March 2012) using Mascot (version 2.4, Matrix Science, London, UK). X! Tandem (The GPM, thegpm.org; version 2007.01.01.1) was used to validate spectral assignments. All search results were validated in Scaffold (version 3.0, Proteome Science Inc., Portland, OR, USA). The Mascot and X! Tandem search engines were searched with a fragment ion mass tolerance of 0.20 Da and a parent ion tolerance of 0.20 Da. Deamidation of asparagine and glutamine, oxidation of methionine, acetylation of the N-terminus and iodoacetamide derivatization of cysteine were specified in Mascot and X! Tandem as variable modifications. Proteins identified by ≥1 peptide (protein probability 80%, peptide probability 95%) were retained in the dataset for further data analysis. SwissProt accession numbers of annotated biofilm proteins were submitted for analysis (www.uniprot.org) to retrieve taxonomic hierarchy, gene ontology (GO) annotations, GO term numbers and protein families.

An independent analysis of the identified peptide sequences was also performed using Unipept 2.3 multi-peptide analysis (Mesuere et al. 2012) which used the UniProt database (www.uniprot.org) and NCBI taxonomy. Briefly, the sequences of all identified peptides were submitted to the Unipept web application (http://unipept.ugent.be/) using the following settings to calculate the lowest common ancestors: multi-peptide analysis, peptides were deduplicated, isoleucine and leucine residues were equated, and advanced missed cleavage handling was applied. The resulting Unipept plots were created using the Biofilm database search results.

All mass spectrometry proteomics data from this study have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD000961 and DOI 10.6019/PXD000961. Further experimental details provided in

Figure 1. (a) Air–water interface biofilms were harvested from the hulls of two US Navy destroyers. (b) Starboard, mid-ship biofilms with a Fouling Rating of 20 were scraped and immediately snap frozen. (c) Study overview. Individual biofilm samples were utilized for PhyloChip™ Array 16S microbial profiling and pooled biofilm samples (all samples collected from the same hull) were utilized for every other aspect of the study.
the Supplementary information include: quantitative metaproteomic methods, carbon and nitrogen isotope elemental analyses, spectroscopic analyses of the photosynthetic pigments and chemicals utilized. [Supplementary information is available via a multimedia link in the online article webpage.]

**Results**

**Prokaryotic community analyses**

PhyloChip-based 16S rDNA profiling identified a total of 8,278 OTUs from 10 biofilm samples (five biofilms from different hull locations extracted using two methods) and two proximal surface water samples. Within the biofilm samples, the majority of OTUs from both ships assigned to members of the Gammaproteobacteria, Alphaproteobacteria and Bacteroidetes. Ship-1 demonstrated greater archaeal and bacterial subfamily richness than Ship-2 (Figure S1). A comparison of the 6,942 OTUs that were present in at least one biofilm sample from both ships indicated that of the three variables tested (biofilms from different hulls, different biofilm metagenomic DNA extraction methods, biofilms sampled from different locations on the same hull) only biofilms from both ships indicated that of the three variables tested (biofilms from different hulls, different biofilm metagenomic DNA extraction methods, biofilms sampled from different locations on the same hull) only biofilms sampled from different hulls had a statistically significant association with community structure, a trend observed using both binary [Adonis p-values, different hulls = 0.012] and abundance [different hulls = 0.004] metrics. Abundance-filtered data of 2,024 taxa revealed three clusters with significant community structure differences (Ship-1 biofilms, Ship-2 biofilms and surrounding surface water bacterioplankton communities) (Figure 2). Finally, a comparison of the OTU richness from each ship revealed distinct inter-ship differences in the phylum, class and family (eg Pseudomonadaceae, Enterobacteriaceae, Clostridiaceae) levels (Table S1).

**Biofilm metagenomes**

To explore the representative accuracy of the acquired PhyloChip profiles and biological potential of each biofilm community, pooled biofilm metagenomic DNA samples from each hull were subjected to whole metagenome sequencing. A total of 49,426,099 and 33,516,399 raw sequence reads were assembled into contiguous sequences resulting in the identification of 243,146 and 183,173 ORFs from which 89,504 and 76,123 ORFs could be annotated by SwissProt from Ship-1 and Ship-2, respectively (Figure S2a, b). Analyses of the relative read abundance demonstrated that both biofilm communities were dominated by bacteria (Figure S2c) and were diverse, with 906 bacterial genera identified as common constituents and another 95 genera that were unique to either hull community. The greatest species richness and sequence read abundance belonged to the phylum Proteobacteria (specifically Alphaproteobacteria and Gammaproteobacteria; Table 1) and this was not surprising as the dominance of Proteobacteria in marine biofilm communities has previously been described (Dang & Lovell 2000; Lee et al. 2008; Chung et al. 2010; Dobretsov et al. 2013). Notable compositional read abundance differences were observed between the two communities on the phylum (eg Cyanobacteria), class (eg Flavobacteria) and genus (eg Erythrobacter) levels (Table 1). Interestingly, DNA reads from putative numerically dominant upper euphotic zone bacterioplankton were either absent (SAR86, SAR92) or found in abundances far lower than those described from filtered surface water bacterioplankton communities (Candidatus *Pelagibacter* < 0.06%, Candidatus *Puniceispirillum* < 0.09%, data not shown) (Poretsky et al. 2009; Shi et al. 2011; Grzymski et al. 2012). Although the Eukaryota were relatively minor constituents of the biofilm metagenomes (Figure S2c), marked compositional differences were also observed between the two ships (Table 2). At the phylum level, 65.94% of the eukaryotic sequences from Ship-2 were associated with the microscopic Bacillariophyta (diatoms, 4× the number found in Ship-1 biofilms). In contrast, Ship-1 had a greater number of sequences from macroscopic eukaryotic phyla (eg Cnidaria, Chordata, Arthropoda).

An annotation of sequence composition for both metagenomes using the COG database identified 3,897
Table 1. Biofilm metagenome composition comparison: percentage of bacterial DNA reads from the most abundant phyla, classes and genera.

| Phylum/a | Ship-1 (%) | Ship-2 (%) | Class/b | Ship-1 (%) | Ship-2 (%) | Genus | Ship-1 (%) | Ship-2 (%) |
|----------|------------|------------|---------|------------|------------|-------|------------|------------|
| Proteobacteria | 65.48 | 77.36 | Alphaproteobacteria | 50.46 | 58.12 | Roseobacter | 9.25 | 7.31 |
| Bacteroidetes | 21.94 | 14.13 | Gammaproteobacteria | 15.41 | 17.27 | Erythrobacter | 7.70 | 21.31 |
| Cyanobacteria | 5.14 | 0.97 | Flavobacteria | 15.26 | 8.61 | Marinibacter | 3.17 | 1.06 |
| Planctomycetes | 1.92 | 2.35 | Deltaproteobacteria | 3.44 | 2.93 | Ruegeria | 3.10 | 3.49 |
| Actinobacteria | 1.54 | 1.52 | Cytophaga | 2.66 | 1.53 | Plesioctys | 2.58 | 1.97 |
| Verrucomicrobia | 1.28 | 1.12 | Planctomycetacia | 2.36 | 2.74 | Polaribacter | 1.78 | 0.89 |
| Firmicutes | 1.06 | 0.97 | Betaproteobacteria | 2.35 | 2.09 | Cyanothoe | 1.66 | 0.25 |
| Acidobacteria | 0.29 | 0.34 | Actinobacteria | 1.89 | 1.77 | Cellulophaga | 1.63 | 0.66 |
| Chloroflexi | 0.28 | 0.25 | Bacteroidia | 0.82 | 0.53 | Planctomyce | 1.40 | 1.40 |
| Chlorobi | 0.14 | 0.12 | Sphingobacteria | 0.87 | 0.57 | Roseovarios | 1.40 | 1.36 |
| Lentisphaerae | 0.13 | 0.12 | Opitutae | 0.74 | 0.28 | Rhodopirellula | 1.37 | 2.11 |
| Deinococcus-Thermus | 0.11 | 0.11 | Clostridium | 0.58 | 0.51 | Robinigintalea | 1.36 | 0.64 |
| Spirochaetes | 0.10 | 0.10 | Bacilli | 0.53 | 0.46 | Marinobacter | 1.32 | 2.29 |
| Gemmatimonadetes | 0.07 | 0.07 | Verrucomicrobiaceae | 0.53 | 0.66 | Coralimargarita | 1.25 | 0.34 |

*a18 phyla with < 0.07% abundance not shown.
*b38 classes with < 0.20% abundance not shown.
*c926 genera with < 1.10% abundance not shown.
*d892 additional genera with < 0.8% abundance not shown.

Four genera were within the top 14 genera by abundance on Ship-2 (*Citomicrobium*: 1.25%, *Sulfotuber*: 1.05%, *Pseudomonas*: 0.83%, *Shewanella*: 0.81%) but not Ship-1 (0.00%, 0.70%, 0.75%, 0.73%, respectively).

Table 2. Biofilm metagenome and metaproteome composition comparison: percentage of eukaryotic DNA reads and annotated proteins.

| Phylum or Kingdom/a | DNA reads/b | UniRef100 + UniMES annotated proteins/c | Biofilm database annotated proteins/d | DNA reads/e | UniRef100 + UniMES annotated proteins/f | Biofilm database annotated proteins/g |
|---------------------|-------------|----------------------------------------|-------------------------------------|-------------|----------------------------------------|--------------------------------------|
| Unassigned          | 29.98       | 7.86                                   | 0.00                                | 19.06       | 7.26                                   | 0.00                                 |
| Bacillariophyta     | 16.48       | 2.83                                   | 7.73                                | 65.94       | 36.87                                  | 32.16                                |
| Cnidaria            | 15.05       | 13.21                                  | 7.18                                | 0.38        | 2.23                                   | 1.17                                 |
| Chordata            | 9.26        | 20.13                                  | 45.86                               | 2.27        | 8.94                                   | 11.11                                |
| Unclassified        | 6.77        | 0.00                                   | 0.00                                | 2.89        | 0.00                                   | 0.00                                 |
| Arthropoda          | 6.68        | 20.44                                  | 11.05                               | 1.71        | 10.06                                  | 4.09                                 |
| Ascomycota          | 3.55        | 5.66                                   | 3.87                                | 2.08        | 5.59                                   | 10.53                                |
| Streptophyta        | 3.35        | 4.09                                   | 6.08                                | 1.66        | 3.35                                   | 18.13                                |
| Nematoda            | 2.01        | 3.14                                   | 6.08                                | 1.26        | 3.91                                   | 1.17                                 |
| Chlorophyta         | 1.96        | 1.89                                   | 2.76                                | 0.16        | 0.12                                   | 0.17                                 |
| Echinodermata       | 1.21        | 1.26                                   | 2.76                                | 0.16        | 0.00                                   | 0.00                                 |
| Basidiomycota       | 0.46        | 1.57                                   | 0.00                                | 0.30        | 1.68                                   | 1.17                                 |
| Placozoa            | 0.28        | 1.57                                   | 0.00                                | 0.07        | 0.00                                   | 0.00                                 |
| Platychelminthes    | 0.25        | 1.57                                   | 1.10                                | 0.06        | 1.12                                   | 0.58                                 |
| Mollusca            | 0.15        | 2.83                                   | 2.21                                | 0.03        | 1.12                                   | 1.17                                 |
| Porifera            | 0.06        | 1.89                                   | 0.00                                | 0.01        | 0.00                                   | 0.00                                 |
| Annelida            | 0.03        | 1.26                                   | 1.10                                | 0.02        | 0.56                                   | 0.00                                 |
| Alveolata           | 0.00        | 3.77                                   | 1.10                                | 0.00        | 7.82                                   | 1.17                                 |
| Amoeboboa           | 0.00        | 0.94                                   | 0.55                                | 0.00        | 1.12                                   | 5.85                                 |
| Cryptophyta         | 0.00        | 0.63                                   | 0.00                                | 0.00        | 0.56                                   | 1.17                                 |
| Rhodophyta          | 0.01        | 0.94                                   | 0.00                                | 0.00        | 3.35                                   | 1.17                                 |
| Stramenopiles       | 0.00        | 1.57                                   | 2.21                                | 0.00        | 1.68                                   | 5.26                                 |

*Proteins annotated using Kingdom designations are unique and do not belong to any of the presented phyla. Only those Phyla or Kingdoms with > 1% representation in any category are presented.
*bPercentage of the 6.23% of the total number of annotated Eukaryota reads from the Ship-1 biofilm metagenome.
*cPercentage based on a total of 318 identified and annotated eukaryotic proteins.
*dPercentage based on a total of 181 identified and annotated eukaryotic proteins.
*ePercentage of the 7.88% of the total number of annotated Eukaryota reads from the Ship-2 biofilm metagenome.
*fPercentage based on a total of 179 identified and annotated eukaryotic proteins.
*gPercentage based on a total of 171 identified and annotated eukaryotic proteins.
clusters with skewed read abundances in three COG classes (carbohydrate transport and metabolism; inorganic ion transport and metabolism; secondary metabolites biosynthesis, transport and catabolism) for Ship-1 and two COG classes (replication, recombination and repair; translation, ribosomal structure and biogenesis) for Ship-2 (Table S2, COG classes with > 1,000 identifications). The COG annotations were combined with results generated using the Pfam (5,553 matches), KEGG (7,240 hits) and GO (1,484 terms) databases and revealed potentially unique and skewed functional potential between the biofilm communities. For example, Ship-1 contained genes encoding phycobilisome, transposase, rhodopsin, circadian clock and cadherin proteins that were not found in Ship-2 (Table S3). Conversely, Ship-2 contained genes encoding heat shock transcription factors and myelin proteins that were not found in Ship-1. Large sequence skews (> 10×) were also observed. While genes encoding an EGF-like domain, cadherin domain, cell adhesion, G-protein coupled receptor signaling pathway, sodium channel and dynein complex proteins were far more abundant in Ship-1, genes encoding HSF-type DNA binding proteins, heat shock and chromatin remodeling transcription factors were more abundant in Ship-2. Finally, from the functional categories with the greatest sequence representation, Ship-1 was enriched (as percentage abundance) for bacterial sensor histidine kinases, oxidoreductases, iron complex outer membrane receptors, ABC transport permease proteins, dehydrogenases, cation transport ATPases and AcrB/AcrD/AcrF family proteins.

Biofilm metaproteomes

Matched qualitative LC-MS/MS metaproteomic analyses identified 678 unique proteins from both hull biofilm communities with limited overlap in composition (Figure 3a). While these analyses were performed using the matched metagenomic sequences as the only searchable database (Biofilm database), the number of proteins identified in each biofilm depended on the database used to search the acquired MS/MS spectra. For a comparison, publicly available databases from the UniProt knowledgebase (UniProt Reference Clusters [UniRef100] + all UniProt Metagenomic and Environmental Sequences [UniMES]) were also searched. The use of the UniRef100 + UniMES database resulted in the identification of 578 proteins using the same probability settings (Figure 3a). While a greater number of proteins were identified from Ship-2 with the Biofilm database compared to the UniRef100 + UniMES database, fewer were identified from Ship-1.

Based on the dominance of prokaryotic sequences in both metagenomes, the expectation was for the majority of identified proteins in the matched metaproteomes to also be of prokaryotic origin. Instead, the metaproteomic findings contrasted with the metagenomic findings in that the majority of identified and annotated proteins were of eukaryotic origin (Figure 3b). Taxonomic analyses of the proteins of eukaryotic origin revealed a similar trend in that most proteins were from phyla that were most abundant in the matched metagenomes. Yet, there were also stark differences in the percentage of metagenomic DNA reads and the percentage of phylum-matched proteins (eg Bacillariophyta, Chordata, Arthropoda) (Table 2). Overall, Ship-1 was enriched with proteins from Metazoa while Ship-2 was enriched for proteins from Bacillariophyta (Figure 3c, d, Table 2). A similar analysis of the proteins of prokaryotic origin revealed that the majority of identifiable proteins were derived from members of the Proteobacteria, Firmicutes, Cyanobacteria, Bacteroidetes and Actinobacteria (Figure 3e, f) which reflected the 16S rDNA (Table S1) and metagenome (Table 1) abundance rankings. The largest inter-ship bacterial protein differences were within the Proteobacteria, Cyanobacteria and Firmicutes. Notably, archaean and viral proteins combined were < 1% of total identified proteins (Figure 3b).

Further partitioning into biological processes revealed that the biofilm metaproteome from Ship-1 demonstrated an enrichment of proteins involved in cell adhesion, DNA binding, calcium binding and eukaryotic cytoplasmic components and structures, while the Ship-2 biofilms were enriched for proteins involved in all aspects of photosynthesis, carbon fixation, membrane transport, photosynthesis and electron transport (Table S4). Proteins generally associated with housekeeping functions (eg ATP-binding, GTP-binding, ribosomes, translation) and the plasma membrane were detected at comparable levels in both biofilms.

Quantitative LC-MS/MS analyses

Stable isotope labeling of these samples proved challenging and resulted in only 5% of the identified proteins being successfully quantified in at least two replicate samples using both labeling methods (complete quantitation data with experimental replicates; Table S5). The 18O labeling method yielded 31 unique quantitated proteins whereas iTRAQ generated five unique quantitated proteins (20 overlapping proteins, 36 unique proteins total). The relative quantitation of proteins common to both biofilms demonstrated significant abundance differences. Ship-1 harbored a greater relative abundance of eukaryotic cytoskeletal proteins from macroscopic marine eukaryotes (Hydra vulgaris, Fucus vesiculosus, and Paracentrotus lividus) and insects (Lepidoglyphus destructor, Spodoptera frugiperda) (Table 3). In contrast, Ship-2 revealed a greater relative abundance of
Figure 3. (a) Venn diagrams depicting the total number of proteins and peptides identified from both biofilm metaproteomes using the Biofilm database or UniREF100 + UniMES database. (b) Superkingdom-level classification of all identified proteins from both biofilm metaproteomes using two databases. The total number of proteins identified and classified using each database can be found under each column in parenthesis. Percentage of Archaea + virus proteins identified (asterisks) – (Ship-1) UniREF100 + UniMES (0.51/0.25), Biofilm database (0.28/0.00); (Ship-2) UniREF100 + UniMES (0.41/0.41), Biofilm database (0.00/0.00). (c) Independent classification of the identified proteins using Unipept 2.3, multi-peptide analysis (Mesuere et al. 2012). Ship-1, 252 out of 499 peptides were matched. (d) Ship-2, 366 out of 709 peptides were matched. (e) Unipept classification of Ship-1 bacterial proteins (expansion of Bacteria sector from (c)). (f) Unipept classification of Ship-2 bacterial proteins (expansion of Bacteria sector from (d)). Only matched peptides were used for the plots.
| Biofilm DBID<sup>a</sup> | SwissProt annotation<sup>b</sup> | Organism | Mean ratio<sup>c</sup> | t-test<sup>d</sup> | N | Mean ratio<sup>c</sup> | t-test<sup>d</sup> | N |
|------------------------|------------------------|---------|----------------|----------------|-----|----------------|----------------|-----|
| 652351.1               | Photosystem II CP47 chlorophyll a apoprotein | *Odontella sinensis* | 0.13 | 0.02 | <0.001 | 5 | 0.49 | 0.20 | 0.014 | 4 |
| 474075.1               | Ribulose bisphosphate carboxylase large chain | *Phaeodactylum tricornutum* | 0.08 | 0.01 | <0.001 | 4 | 0.18 | 0.23 | 0.006 | 4 |
| 602495.1               | Ribulose bisphosphate carboxylase large chain | *Detonula confervacea* | 0.11 | 0.02 | <0.001 | 5 | 0.06 | 0.08 | 0.25 | 3 |
| 678959.2               | Transketolase | *Dictyospermum discoideum* | 0.06 | 0.01 | <0.001 | 5 | 0.03 | 0.00 | 0.02 | 4 |
| 640637.1               | Ribulose bisphosphate carboxylase small chain | *Cylindrotheca sp.* | 0.16 | 0.10 | <0.001 | 5 | 0.03 | 0.00 | 0.02 | 4 |
| 681869.1               | Glyceraldehyde-3-phosphate dehydrogenase, chloroplastic | *Guillardia theta* | 0.24 | 0.29 | n/a<sup>e</sup> | 4 | 0.19 | 0.14 | 0.010 | 3 |
| 579651.1               | Photosystem I P700 chlorophyll a apoprotein A2 | *Phaeodactylum tricornutum* | 0.05 | 0.01 | <0.001 | 4 | 0.06 | 0.08 | 0.025 | 3 |
| 673799.1               | Photosystem I P700 chlorophyll a apoprotein A1 | *Phaeodactylum tricornutum* | 0.10 | 0.08 | 0.003 | 3 | 0.67 | 0.09 | 0.03 | 3 |
| 454573.1               | Ribulose bisphosphate carboxylase large chain | *Phaeodactylum tricornutum* | 0.48 | 0.31 | 0.043 | 4 | 0.10 | 0.09 | 0.003 | 3 |
| 454575.1               | Ribulose bisphosphate carboxylase large chain | *Cylindrotheca sp.* | 0.26 | 0.15 | 0.014 | 3 | 0.67 | 0.09 | 0.03 | 3 |
| 674045.1               | Photosystem I reaction center subunit III | *Odontella sinensis* | 0.37 | 0.29 | 0.023 | 4 | 0.50 | 0.18 | 0.008 | 3 |
| 603721.1               | Cysteine synthase | *Synechocystis sp.* | 0.31 | 0.43 | 0.108 | 3 | 0.19 | 0.14 | 0.010 | 3 |
| 672857.1               | Photosystem I reaction center subunit II | *Odontella sinensis* | 0.31 | 0.13 | 0.002 | 4 | 0.11 | 0.03 | <0.001 | 3 |
| 675795.1               | Photosystem II CP43 chlorophyll a apoprotein | *Odontella sinensis* | 0.61 | 0.38 | n/a<sup>e</sup> | 3 | 0.15 | 0.08 | 0.003 | 3 |
| 622951.1               | No annotation | | 0.43 | 0.11 | 0.013 | 3 | 0.19 | 0.14 | 0.010 | 3 |
| 641261.1               | No annotation | | 0.22 | 0.15 | 0.012 | 3 | 0.42 | 0.17 | 0.003 | 3 |
| 682907.1               | No annotation | | 0.61 | 0.38 | n/a<sup>e</sup> | 3 | 0.19 | 0.14 | 0.010 | 3 |
| 742725.1               | No annotation | | 0.15 | 0.08 | 0.003 | 3 | 0.20 | 0.19 | 0.018 | 3 |
| 677425.1               | No annotation | | 0.43 | 0.11 | 0.013 | 3 | 0.19 | 0.14 | 0.010 | 3 |
| 551533.1               | No annotation (C-terminal peptide) | | 0.61 | 0.38 | n/a<sup>e</sup> | 3 | 0.19 | 0.14 | 0.010 | 3 |
| 448275.1               | No annotation (C-terminal peptide) | | 0.19 | 0.14 | 0.010 | 3 | 0.20 | 0.19 | 0.018 | 3 |
| 1070468.1              | Actin, non-muscle 6.2 | *Hydra vulgaris* | 9.69 | 5.24 | 0.021 | 5 | 4.15 | 1 |
| 994040.1               | Actin, non-muscle 6.2 | *Hydra vulgaris* | 7.22 | 3.90 | 0.025 | 5 | 3.59 | 1.16 | 0.061 | 3 |
| 1068352.1              | Tubulin beta chain | *Sus scrofa* | 7.14 | 5.09 | 0.054 | 5 | 2.00 | 1 |
| 965876.1               | Actin, non-muscle 6.2 | *Hydra vulgaris* | 12.21 | 7.96 | 0.067 | 4 | 9.89 | 13.11 | 0.361 | 3 |
| 845049.1               | Protein disulfide-isomerase | *Oryctolagus cuniculus* | 13.02 | 11.33 | 0.077 | 5 | 10.43 | 9.80 | 0.237 | 3 |
| 659673.1               | Histone H4 | *Mastigamoeba balamuthi* | 9.30 | 5.27 | 0.112 | 3 | 0.67 | 0.09 | 0.03 | 3 |
| 973968.1               | Tubulin alpha chain | *Lepidoglypus destructor* | 10.43 | 9.80 | 0.237 | 3 | 0.67 | 0.09 | 0.03 | 3 |
| 663908.1               | ATP synthase subunit beta, mitochondrial | *Rattus norvegicus* | 7.36 | 3.81 | n/a<sup>e</sup> | 5 | 0.19 | 0.14 | 0.010 | 3 |
| 674325.1               | 14-3-3-like protein | *Fucus vesiculosus* | 12.23 | 12.00 | 0.247 | 3 | 2.50 | 1 |
| 1059130.1              | Actin, non-muscle 6.2 | *Paracentrotus lividus* | 30.15 | 34.49 | n/a<sup>e</sup> | 5 | 1.12 | 1 |
| 664050.1               | 40S ribosomal protein S4 | *Spodoptera frugiperda* | 9.71 | 13.25 | n/a<sup>e</sup> | 3 | 0.19 | 0.14 | 0.010 | 3 |

<sup>a</sup>Protein accession numbers from the matched in-house Biofilm database.

<sup>b</sup>SwissProt annotation of proteins from the Biofilm database.

<sup>c</sup>Mean 18O ratios, standard deviations (SD) were calculated from protein ratios, average ratios were analyzed by t-test to determine if they were different from 1, the p-value of the test is shown. N represents the number of experiments in which the given protein ratio was observed (Nmax = 5).

<sup>d</sup>iTRAQ ratios, standard deviations (SD) were calculated from protein ratios, average ratios were analyzed by t-test to determine if they were different from 1, the p-value of the test is shown. N represents the number of experiments in which the given protein ratio was observed (Nmax = 4).

<sup>e</sup>The data were not normally distributed, therefore a t-test was not performed and the p-value could not be calculated.
photosystem and carbon fixation proteins from marine diatoms (*Odontella sinensis*, *Phaeodactylum tricornutum*, *Detonula confervacea* and *Cylindrotheca* sp.). Although the $^{18}$O labeling method loses precision when applied to samples with larger differences in individual protein abundances, thus yielding high variance in the calculated relative protein ratios and higher than usual $p$-values, the independent confirmation of approximately half of these protein abundance measurements via iTRAQ-labeling increased the authors’ confidence in these quantitative measurements (Table 3). Further evidence of the accuracy of these measurements was provided by the iTRAQ-based quantitation of two proteins from the *Odontella sinensis* photosystem II (chlorophyll apoproteins CP43 and CP47) which were found to be twice as abundant on Ship-2 but also quantitated at a 1:1 ratio which agrees with the known stoichiometric ratio of these proteins in a functional photosystem II complex.

**Spectroscopic photosynthetic pigment analyses and elemental analyses**

The relative pigment composition was estimated from a Gaussian fitting and deconvolution of fluorescence and excitation spectra of biofilm acetone extracts from both ships (Figure S3). In general, the similarity of the excitation spectra confirmed the presence of chlorophylls (chl) $a$, $b$ and $c$ and pheophytin $a$ (pheo $a$) in both biofilms. However, the intensities of the individual bands suggested that the pigment amounts differed considerably. When the results were normalized to chl $a$, the analyses revealed comparatively lower levels of chl $b$ and pheo $a$ on Ship-2 and lower levels of chl $c$ on Ship-1 (Table S6). Overall, both samples demonstrated low amounts of chl $b$ which could be attributed to the presence of non-chl $b$ containing photosynthetic organisms, specifically, chl $c$-containing species. The strong intensity of chl $c$ fluorescence at 630 nm confirmed this assertion (Figure S3a, c, h). The nearly twofold increase of chl $c$ in Ship-2 was indicative of a population enriched by photosynthetic diatoms and dinoflagellates, a finding that was in agreement with the metagenomic and metaproteomic results for Ship-2.

Elemental analyses revealed that total nitrogen and carbon (wt%) were significantly elevated in Ship-1 compared to Ship-2 (Table S7). The increased nitrogen content in Ship-1 was likely due to the increased protein concentration as indicated by the AQA results (data not shown) that demonstrated a 67.6% greater protein content in Ship-1 vs Ship-2. Stable carbon isotope ratios for total organic carbon (TOC) ($\delta^{13}$C$_{\text{TOC}}$) averaged $-18.4 \%$ (± 0.17) for Ship-1 and $-19.9 \%$ (± 0.21) for Ship-2. As $\delta^{13}$C$_{\text{TOC}}$ for marine phytoplankton varies from $-24$ to $-19\%$ (Peterson & Fry 1987) and values for marine phototrophic biofilms have been observed between $-30$ and $-18\%$ (Staal et al. 2007), the observed $\delta^{13}$C$_{\text{TOC}}$ were indicative of enzymatic fractionation of dissolved inorganic carbon in seawater during planktonic photosynthesis (Peterson & Fry 1987).

**Discussion**

The composition of marine biofilms is dependent on the physical, chemical, seasonal and geographic features of the local environment (Lejars et al. 2012; Salta et al. 2013). However, unlike biofilm communities from relatively stable marine environments (eg open ocean, sediments), communities that are formed on ships’ hulls are more likely to be highly complex and dynamic as they are not only subject to strong physical selection pressures (eg insolation, oxidative stress, predation, shearing, and antifouling biocides) but also accumulate, and potentially interchange, members from the surrounding environment during transit and over time. In this study, the use of large-scale, culture-independent biomolecular measurements was combined with more traditional biochemical analyses to begin to understand the complexity and function of two such communities. Overall, the results demonstrated that the biofilms formed on the hulls of vessels that initiated and concluded their deployment at the same geographic site, but traveled to different geographic locations in the interim, differed in both microorganism composition and biological potential.

To begin to understand the level of prokaryotic diversity in these communities, PhylloChip-based 16S rDNA metagenomic profiling was used (DeSantis et al. 2007; Hazen et al. 2010). Of the ~60,000 taxa that could be identified using the PhylloChip Array, the analyses identified a total of 8,278 taxa from the biofilms and surrounding water as bacterioplankton and 6,942 taxa from the biofilms alone. The prokaryotic richness observed was far greater than that observed in a previous PhylloChip analysis of ≤ 12 day old subtidal marine biofilms that identified 158 total taxa (Chung et al. 2010). This increase in species richness was likely due to the age of the ship hull biofilms (~ 7 months), as marine biofilm community richness is known to increase with immersion time, and the mobility of the substrata that likely aided in the accumulation of consortium members from different geographic environments. Interestingly, both biofilm communities were dissimilar to their matched proximal surface water bacterioplankton communities, an observation that has been previously noted in marine environments (Bengtsson & Ovres 2010; Shikuma & Hadfield 2010; Briand et al. 2012).

Given the major role fouling microalgae are known to play in marine biofilms, especially on man-made surfaces (Molino & Wetherbee 2008; Briand et al. 2012), perhaps the most unexpected finding from the metagenomes was that > 91% of the total sequence reads belonged to
bacteria. Collectively, Gammaproteobacteria and Alphaproteobacteria accounted for the majority of unique OTUs and metagenome sequence reads. However, while the Gammaproteobacteria contained ~ 2.4× more OTUs than the Alphaproteobacteria, the Alphaproteobacteria generated > 3.2× more sequence reads than the Gammaproteobacteria, suggesting that the biofilm Gammaproteobacteria exhibited greater species richness but were numerically less abundant than the Alphaproteobacteria.

On the genus level, the greatest number of Ship-1 reads (9.25% of all bacterial reads) and second greatest number of Ship-2 reads (7.31%) belonged to members of the genus Roseobacter. This was not particularly surprising considering their numerical abundance in marine environments and coastal biofilm communities (Dang & Lovell 2000, 2002). In general, members of the Roseobacter clade (Brinkhoff et al. 2008) are known for their surface attachment and colonization success in the formation of marine biofilms (Slightom & Buchan 2009; Chen et al. 2013). This was certainly true of the sampled communities as 19.28% (Ship-1) and 17.62% (Ship-2) of the bacterial metagenomic sequences were ascribed to members of the Roseobacter clade. Sequencing reads from the aerobic anoxygenic phototrophic genus Erythrobacter were found to be the most abundant bacterial reads from Ship-2 (21.31%); an amount that was ~ 2.8× greater than that found on Ship-1. In hindsight, this finding was also not particularly unexpected as Erythrobacter species are most frequently found in nutrient-rich coastal surface waters (Shi et al. 2011), an environment in which Ship-2 spent appreciably more time than Ship-1.

Previous ‘-omic’ analyses have also indicated the importance of viruses and viral infections in aquatic communities (Frias-Lopez et al. 2008; Morris et al. 2010; Sowell et al. 2011) by demonstrating the presence and abundance of viral sequences (1–10% of total sequence reads (DeLong et al. 2006; Williamson et al. 2008) and proteins (8% of all peptide spectra; Sowell et al. 2011). However, the metagenomic and metaproteomic analyses reported here suggest that the role of viruses in marine biofilm communities may be comparatively diminished (0.06% and 0.02% viral sequence reads in Ship-1 and Ship-2, respectively). As the metagenomic methods employed were not adapted for the sequencing of RNA molecules, the possibility exists that the viral load presented here has been underestimated given the presence and productivity of marine eukaryotes in these biofilms and the fact that marine RNA viruses almost exclusively infect eukaryotes (Steward et al. 2013). Similarly, archaea were also comparatively poorly represented (0.18% and 0.17% sequence reads in Ship-1 and Ship-2, respectively) and demonstrated markedly low diversity. This finding however was not completely unexpected as archaea have not been found to colonize surfaces in coastal biofilm communities (Dang & Lovell 2002).

Overall, the 16S rDNA profiling and metagenomic sequencing data suggested that both communities were dominated by prokaryotic organisms. Ship-1 harbored greater bacterial species richness and functional diversity, which may have been due in part to a higher organismal density, a contention which was supported by the observed macroscopic properties, protein concentration and carbon and nitrogen bulk analyses. Even though members of the Eukaryota were relatively minor constituents based on read abundance, the findings also revealed an interesting dichotomy between the types of eukaryotic inhabitants found on both ships. The observed differences in the compositions of microscopic and macroscopic eukaryotes were likely due not only to differences in geography, but also the maturity of each biofilm community. Combined, the TOC wt% and δ13C_TOC data supported the assertion that Ship-1 harbored a more mature hull biofilm than Ship-2, but that δ13C_TOC in both resulted from phototrophic processes and were suggestive of communities whose productivity was driven by photautotrophic microorganisms which in turn supported an array of heterotrophic coinhabitants. However, both the metagenomic and metaproteomic analyses revealed that the underlying phototrophic nature of the biofilms differed and was imparted primarily by photosynthetic prokaryotes on Ship-1 and photosynthetic microeukaryotes that are known to be major components of microbial slimes on Ship-2 (Molino & Wetherbee 2008; Zargiel et al. 2011; Briand et al. 2012).

Although the depth of the metaproteomic measurements described do not rival those of the metagenomic information obtained from the same samples, the metaproteomic content more closely reflects how active (or previously active) community members, regardless of numerical abundance, function and affect the biofilm community. As such, the fact that the majority of identified proteins were eukaryotic in origin could suggest that eukaryotic organisms dominate productivity in these ship hull communities. Based on the metagenomic analyses alone, a greater percentage of the identified metaproteome would be expected to belong to the bacterial constituents of these communities given that the vast majority of DNA sequencing reads (>90%) were of bacterial origin. Three facts may aid in explaining the observed results. (1) A eukaryotic protein skew in truly mixed communities is often expected due to the larger size of their encoded proteomes and greater dynamic range of protein expression. (2) It is possible that at the time of sampling the bacterial community members were abundant but not active. This has previously been observed in the numerically abundant Pelagibacter residing in the deep chlorophyll maximum layer (Shi et al. 2011) and recent efforts have been made to elucidate the relationship between abundance and specific activity in ocean surface waters (Frias-Lopez et al. 2008; Hunt...
employed heavily in 2013. The results also highlight considerations for the analysis of highly complex environmental microbial communities by providing data to assess the strengths and limitations of current technologies. These data prompted a number of general questions that are currently difficult to answer yet warrant further examination as they determine what is interpreted as a ‘true’ representation of the biology (vs those aspects relegated as artifacts introduced by deficiencies in experimental design, measurement sensitivity or the bioinformatic methods employed). For example, are the discordant metagenomic and metaproteomic datasets an accurate representation of the composition and activity within these biofilms (eg less abundant organisms that are more metabolically active)? Or, do the different methods of biomolecule extraction, processing and varying biology introduce unique biases that account for the observed discordance (Morgan et al. 2010; Koid et al. 2012; Yuan et al. 2012; Leary et al. 2013). When dealing with complex environmental samples, the experimental, analytical and statistical choices employed heavily influence the biological conclusions drawn in metaproteomic analyses (Dowd 2012) and most often limit them to those proteins that are most abundant or easiest to access (extraction bias), amenable to the biochemistry and biophysics employed (sample processing bias) and have previously been sequenced (sequence bias) and characterized (bioinformatic database bias) (Leary et al. 2013). The latter clearly impacted the findings of this study as 29.98% (Ship-1) and 19.06% (Ship-2) of the eukaryotic DNA sequences went unassigned, only 4.2% and 7.6% of the total acquired MS/MS spectra were assigned to in silico translated sequences from the UniRef+UniMES and Biofilm databases, respectively, and 36% and 42% of the proteins identified from Biofilm database remained unannotated and for this reason could not be classified in any biologically informative manner.

Conclusions
Despite the known impact of hull fouling biofilms (Schultz et al. 2011), investigation of these largely non-culturable microbial assemblages has been limited by a lack of analytical tools capable of measuring their inherent biological complexity and potential. In the absence of this type of information, the antifouling coatings industry has been left to develop solutions for a recalcitrant problem whose scope and specificity are not yet known. The data described in this study provide a baseline for understanding the microbial composition of true marine biofilms that form on ships’ hulls coated with antifouling paint operating in dynamic conditions. While the findings reveal the tremendous complexity and varying functionality of the biofilms, they are also obviously incomplete and present caveats and considerations for the interpretation of singular environmental ‘-omics’ datasets. A commonly used approach to improve analytical completeness is to reduce sample complexity but this is not a tenable strategy when studying physically associated communities such as marine biofilms. Part of this challenge can be countered by increasing measurement depth. However, the large remainder will rely upon continued advances in biomolecular separations and extraction science, high throughput DNA sequence read length, and the expanded sequencing of marine prokaryote and eukaryote genomes. Improvement in these areas will best complement the existing experimental ‘-omics’ suite and markedly enhance the analyses of marine biofilms at a depth and breadth that were previously unattainable. The broad applicability of this experimental platform should thus be exploited to interrogate several aspects of marine biofilms (eg formation, maturation, recruitment of macrofouling species, microbially influenced corrosion, contribution of important novel or unculturable species, influence of antifouling coatings, effect of environmental variables) to better understand these economically and operationally important microbial communities.

Acknowledgements
The authors thank Drs Todd DeSantis, Christel Chehoud and the scientists at Second Genome, Inc. for helpful discussions. They would also like to acknowledge DDG CLASS Advocate Brian D. McClain for coordinating the sampling efforts, the Executive Officers of the USS Laboon and USS Bainbridge for their cooperation and assistance and Jeffrey Brown SURF-MEPP CORCON SME. Metaproteomic data deposition to the ProteomeXchange Consortium was facilitated by the PRIDEx Team, EMBL-EBI. This work was supported by the Office of Naval Research via US Naval Research Laboratory core funds under grant number 69-P113-11. The opinions and assertions contained herein are those of the authors and are not to be construed as those of the US Navy, military service at large or the US Government.

References
Amann RI, Ludwig W, Schleifer KH. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev. 59:143–169.
Bengtsson MM, Ovreas L. 2010. Planctomycetes dominate biofilms on surfaces of the kelp Laminaria hyperborea. BMC Microbiol. 10:261.

Briand JF, Djeridi I, Jamet D, Coupe S, Bressy C, Molmeret M, Le Berre B, Rimet F, Bouchez A, Blache Y. 2012. Pioneer marine biofilms on artificial surfaces including antifouling coatings immersed in two contrasting French Mediterranean coast sites. Biofouling. 28:453–463.

Brinkhoff T, Giebel HA, Simon M. 2008. Diversity, ecology, and genomics of the Roseobacter clade: a short overview. Arch Microbiol. 189:531–539.

Camps M, Barani A, Gregori G, Bouchez A, Le Berre B, Bressy C, Blache Y, Briand J-F. 2014. Antifouling coatings influence both abundance and community structure of colonizing biofilms: a case study in the northwestern Mediterranean Sea. Appl Environ Microbiol. 80:4821–4831.

Chen C-L, Maki JS, Rittschof D, Teo SLM. 2013. Early marine bacterial biofilm on a copper-based antifouling paint. Int Biodeterior Biodegrad. 83:71–76.

Chung HC, Lee OO, Huang YL, Mok SY, Kolter R, Qian PY. 2010. Bacterial community succession and chemical profiles of subtidal biofilms in relation to larval settlement of the polychaete Hydrophides elegans. ISME J. 4:817–828.

Cox MP, Peterson DA, Biggs PJ. 2010. SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinf. 11:485. doi:10.1186/1471-2105-11-485.

Dang H, Lovell CR. 2000. Bacterial primary colonization and early succession on surfaces in marine waters as determined by amplified 16S rRNA gene restriction analysis and sequence analysis of 16S rRNA genes. Appl Environ Microbiol. 66:467–475.

Dang H, Lovell CR. 2002. Numerical dominance and phylo-type diversity of marine Rhodobacter species during early colonization of submerged surfaces in coastal marine waters as determined by 16S ribosomal DNA sequence analysis and fluorescence in situ hybridization. Appl Environ Microbiol. 68:496–504.

DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard NU, Martinez A, Sullivan MB, Edwards R, Brito BR, Chishom SW, et al. 2006. Community genomics among stratified microbial assemblages in the ocean’s interior. Science. 311:496–503.

DeSantis TZ, Andersen GL, Zubieta IX, Piceno YM, Andersen GL. 2007. High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Microb Ecol. 53:371–383.

Dobretsov S, Abed RM, Voolstra CR. 2013. The effect of surface colour on the formation of marine micro and macrofouling communities. Biofouling. 29:617–627.

Dowd WW. 2012. Challenges for biological interpretation of environmental proteomics data in non-model organisms. Integr Comp Biol. 52:705–720.

Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, Delong EF. 2008. Microbial community gene expression in ocean surface waters. Proc Natl Acad Sci USA. 105:3805–3810.

Gilbert JA, Field D, Huang Y, Edwards R, Li W, Gilna P, Joint I. 2008. Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. PLoS One. 3:e3042.

Grzymski JJ, Riesenfeld CS, Williams TJ, Dussaquer AM, Ducklow H, Erickson M, Cavichiolli R, Murray AE. 2012. A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. ISME J. 6:1901–1915.

Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N, Jansson JK, Probst A, Borglin SE, Fortney JL, Stringfellow WT, et al. 2010. Deep-sea oil plume enriches indigenous oil-degrading bacteria. Science. 330:204–208.

Hewson I, Poretsky RS, Tripp HJ, Montoya JP, Zehr JP. 2010. Spatial patterns and light-driven variation of microbial population gene expression in surface waters of the oligotrophic open ocean. Environ Microbiol. 12:1940–1956.

Hunt DE, Lin Y, Church MJ, Karl DM, Tringe SG, Izzo LK, Johnson ZI. 2013. Relationship between abundance and specific activity of bacterioplankton in open ocean surface waters. Appl Environ Microbiol. 79:177–184.

Klein DA. 2011. Bulk extraction-based microbial ecology: three critical questions. Microbe. 6:377.

Koid A, Nelson WC, Mraz A, Heidelberg KB. 2012. Comparative analysis of eukaryotic marine microbial assemblages from 18S rRNA gene and gene transcript clone libraries by using different methods of extraction. Appl Environ Microbiol. 78:3958–3965.

Lane DJ. 1991. 16S/23S rRNA sequencing. In: Stackebrandt E, Goodfellow M, editors. Nucleic acid techniques in bacterial systematics. New York: John Wiley and Sons; p. 115–175.

Leary DH, Hervey WJ 4th, Li RW, Deschamps JR, Kusterbeck AW, Vora GJ. 2012. Method development for metaproteomic analyses of marine biofilms. Anal Chem. 84:4006–4013.

Leary DH, Hervey WJ 4th, Deschamps JR, Kusterbeck AW, Vora GJ. 2013. Which metaproteome? The impact of protein extraction bias on metaproteomic analyses. Mol Cell Probes. 27:193–199.

Lee JW, Nam JH, Kim YH, Lee KH, Lee DH. 2008. Bacterial communities in the initial stage of marine biofilm formation on artificial surfaces. J Microbiol. 46:174–182.

Lejars M, Margailan A, Bressy C. 2012. Fouling release coatings: a nontoxic alternative to biocidal antifouling coatings. Chem Rev. 112:4347–4390.

Li RW, Connor EE, Li C, Baldwin Vi RL, Sparks ME. 2012. Characterization of the rumen microbiota of pre-ruminant calves using metagenomic tools. Environ Microbiol. 14:129–139.

Li R, Yu C, Li Y, Lam TW, Yu SM, Kristiansen K, Wang J. 2009. SOAP2: An improved ultrafast tool for short read alignment. Bioinformatics. 25:1966–1967.

Mesuere B, Devreese B, Debyser G, Aerts M, Vandamme P, Dawyndt P. 2012. Unipept: tryptic peptide-based biodiversity analysis of metaproteome samples. J Proteome Res. 11:5773–5780.

Molino PJ, Wetherbee R. 2008. The biology of biofouling diatoms and their role in the development of microbial slimes. Biofouling. 24:365–379.

Morgan JL, Darling AE, Eisen JA. 2010. Metagenomic sequencing of an in vitro-simulated microbial community. PLoS One. 5:e10209.

Morris RM, Nunn BL, Frazier C, Goodlett DR, Ting YS, Rocap G. 2010. Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. ISME J. 4:673–685.

Peterson BJ, Fry B. 1987. Stable isotopes in ecosystem studies. Annu Rev Ecol Syst. 18:293–320.

Poretsky RS, Hewson I, Sun S, Allen AE, Zehr JP, Moran MA. 2009. Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. Environ Microbiol. 11:1358–1375.
Rho M, Tang H, Ye Y. 2010. FragGeneScan: predicting genes in short and error-prone reads. Nucleic Acids Res. 38:e191.

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S, Wu D, Eisen JA, Hoffman JM, Remington K, Beeson K, et al. 2007. The Sorcerer II global ocean sampling expedition: Northwest Atlantic through eastern tropical Pacific. PLoS Biol. 5:e77.

Salta M, Wharton JA, Blache Y, Stokes KR, Briand JF. 2013. Marine biofilms on artificial surfaces: structure and dynamics. Environ Microbiol. 15:2879–2893.

Schultz MP, Bendick JA, Holm ER, Hertel WM. 2011. Economic impact of biofouling on a naval surface ship. Biofouling. 27:87–98.

Shi Y, Tyson GW, Eppley JM, DeLong EF. 2011. Integrated metatranscriptomic and metagenomic analyses of stratified microbial assemblages in the open ocean. ISME J. 5:999–1013.

Shikuma NJ, Hadfield MG. 2010. Marine biofilms on submerged surfaces are a reservoir for Escherichia coli and Vibrio cholerae. Biofouling. 26:39–46.

Slightom RN, Buchan A. 2009. Surface colonization by marine roseobacters: integrating genotype and phenotype. Appl Environ Microbiol. 75:6027–6037.

Sowell SM, Abraham PE, Shah M, Verberkmoes NC, Smith DP, Barošky DF, Giovannoni SJ. 2011. Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. ISME J. 5:856–865.

Staal M, Thor R, Kuhl M, van Loosdrecht MCM, Wolf G, de Brouwer JFC, Rijstenbil JW. 2007. Different carbon isotope fractionation patterns during the development of photoautotrophic freshwater and marine biofilms. Biogeosciences. 4:613–626.

Steinberger RE, Holden PA. 2005. Extracellular DNA in single- and multiple-species unsaturated biofilms. Appl Environ Microbiol. 71:5404–5410.

Steward GF, Culley AI, Mueller JA, Wood-Charlson EM, Belcast M, Poisson G. 2013. Are we missing half of the viruses in the ocean? ISME J. 7:672–679.

Whitchurch CB, Tolker-Nielsen T, Ragas PC, Mattick JS. 2002. Extracellular DNA required for bacterial biofilm formation. Science. 295:1487.

Williams TJ, Long E, Evans F, Demaere MZ, Lauro FM, Raftery M, Ducklow H, Grzymski JJ, Murray AE, Cavicchioli R. 2012. A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. ISME J. 6:1883–1900.

Williamson SJ, Rusch DB, Yooseph S, Halpern AL, Heidelberg KB, Glass JI, Andrews-Pfannkoch C, Fadrosh D, Miller CS, Sutton G, Frazier M, et al. 2008. The Sorcerer II global ocean sampling expedition: metagenomic characterization of viruses within aquatic microbial samples. PLoS One. 3:e1456.

Woese CR. 1996. Whither microbiology? Phylogenetic trees. Curr Biol. 6:1060–1063.

Yebra DM, Kiil S, Weinnell CE, Dam-Johansen K. 2006. Presence and effects of marine microbial biofilms on biocide-based antifouling paints. Biofouling. 22:33–41.

Yuan S, Cohen DB, Ravel J, Abdo Z, Fomye LJ. 2012. Evaluation of methods for the extraction and purification of DNA from the human microbiome. PLoS One. 7:e33865.

Zargiel KA, Coogan JS, Swain GW. 2011. Diatom community structure on commercially available ship hull coatings. Biofouling. 27:955–965.