Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea

Jay McCarren[a,b], Jamie W. Becker[a,c], Daniel J. Repeta[c], Yanmei Shi[a], Curtis R. Young[a], Rex R. Malmstrom[a,d], Sallie W. Chisholm[a], and Edward F. DeLong[a,e,1]

Departments of [a]Civil and Environmental Engineering and [e]Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; [c]Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; [b]Synthetic Genomics, La Jolla, CA 92037; and [d]Joint Genome Institute, Walnut Creek, CA 94598

This contribution is part of the special series of Inaugural Articles by members of the National Academy of Sciences elected in 2008.

Contributed by Edward F. DeLong, August 2, 2010 (sent for review July 1, 2010)

Marine dissolved organic matter (DOM) contains as much carbon as the Earth's atmosphere, and represents a critical component of the global carbon cycle. To better define microbial processes and activities associated with marine DOM cycling, we analyzed genomic and transcriptional responses of microbial communities to high-molecular-weight DOM (HMWDOM) addition. The cell density in the unamended control remained constant, with very few transcript categories exhibiting significant differences over time. In contrast, the DOM-amended microcosm doubled in cell numbers over 27 h, and a variety of HMWDOM-stimulated transcripts from different taxa were observed at all time points measured relative to the control. Transcripts significantly enriched in the HMWDOM treatment included those associated with two-component sensor systems, phosphate and nitrogen assimilation, chemotaxis, and motility. Transcripts from *Idiomarina* and *Alteromonas* spp., the most highly represented taxa at the early time points, included those encoding TonB-associated transporters, nitrogen assimilation genes, fatty acid catabolism genes, and TCA cycle enzymes. At the final time point, *Methylophaga* rRNA and non-rRNA transcripts dominated the HMWDOM-amended microcosm, and included gene transcripts associated with both assimilatory and dissimilatory single-carbon compound utilization. The data indicated specific resource partitioning of DOM by different bacterial species, which results in a temporal succession of taxa, metabolic pathways, and chemical transformations associated with HMWDOM turnover. These findings suggest that coordinated, cooperative activities of a variety of bacterial "specialists" may be critical in the cycling of marine DOM, emphasizing the importance of microbial community dynamics in the global carbon cycle.

carbon cycle | marine | bacteria | metagenomics | metatranscriptomics

Microbial activities drive most of Earth's biogeochemical cycles. Many processes and players involved in these planetary cycles, however, remain largely uncharacterized, due to the inherent complexity of microbial community processes in the environment. Cycling of organic carbon in ocean surface waters is no exception. Though marine dissolved organic matter (DOM) is one of the largest reservoirs of organic carbon on the planet (1), microbial activities that regulate DOM turnover remain poorly resolved (2).

Marine DOM is an important substrate for heterotrophic bacterioplankton, which efficiently remineralize as much as 50% of total primary productivity through the microbial loop (3–6). Though some DOM is remineralized on short timescales of minutes to hours, a significant fraction escapes rapid removal. In marine surface waters, this semilabile DOM transiently accumulates to concentrations 2–3 times greater than are found in the deep sea (7), and represents a large inventory of dissolved carbon and nutrients that are potential substrates for marine microbes. Time-series analyses of semilabile DOM accumulation in temperate and subtropical upper ocean gyres show an annual cycle in DOC inventory with net accumulation following the onset of summertime stratification, and net removal following with deep winter mixing. In addition, multiyear time-series data suggest that surface-water DOM inventories have been increasing over the past 10–20 y (8). The ecological factors behind these seasonal and decadal DOC accumulations are largely unknown. Nutrient (N, P) amendments do not appear to result in a drawdown of DOC, and other factors such as the microbial community structure and the chemical composition of semilabile DOM have been invoked to explain the dynamics of the semilabile DOC reservoir (9, 10). Whatever the cause, the balance and timing of semilabile DOM remineralization are critical factors that influence the magnitude of DOM and carbon exported to the ocean's interior through vertical mixing.

There are significant challenges associated with characterizing and quantifying complex, microbially influenced processes such as DOM cycling in the sea. These challenges include inherent phylogenetic and population diversity and variability, the complexities of microbial community metabolic properties and interactions, and those associated with measuring microbial assemblage activities and responses on appropriate temporal and spatial scales. Past approaches have included measuring the bulk response of microbial communities to nutrient addition (e.g., community substrate incorporation or respiration), following changes in total or functional group cell numbers by microscopy or flow cytometry, or monitoring changes in relative taxa abundance, typically using rRNA-based phylogenetic markers. A number of field experiments (9–13) have indicated that specific shifts in microbial community composition might be linked to surface-water carbon utilization. However, the pure compound nutrient additions (such as glucose) frequently used in such field experiments (9, 11, 14, 15) may not well approximate the environmentally relevant chemical mixtures or compound concentrations present in naturally occurring DOM.

Though complications associated with direct experimentation on natural microbial communities limit our understanding of oceanic carbon cycling to some extent, significant insight into these processes have been recently reported. For example, Carlson et al. (10) showed differences among depth-stratified microbial communities that may be related to their ability to use semilabile DOM that

Author contributions: J.M., J.W.B., D.J.R., R.R.M., and E.F.D. designed research; J.M., J.W.B., D.J.R., Y.S., and R.R.M. performed research; S.W.C. contributed new reagents/analytic tools; J.M., J.W.B., D.J.R., Y.S., C.R.Y., R.R.M., and E.F.D. analyzed data; and J.M., J.W.B., D.J.R., Y.S., C.R.Y., and E.F.D. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. SRA020733.11 and HQ012268–HQ012278).

[1]To whom correspondence should be addressed. E-mail: delong@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1010732107/-/DCSupplemental.

accumulates in ocean surface waters. In addition, phylogenetic analyses of time-series samples have identified some taxonomic groups that appear to be responsive to deep-water mixing events, which may be relevant to organic carbon cycling dynamics (16, 17).

To better define the processes and population dynamics associated with marine microbial DOM cycling in ocean surface waters, we performed controlled experiments using seawater microcosms amended with freshly prepared, naturally occurring DOM. High-molecular-weight DOM (HMWDOM, defined here as the size fraction >1,000 Da and <30,000 Da) was concentrated by ultrafiltration using a 1-nm membrane filter, followed by a second filtration step to remove viruses. Whole, unfiltered seawater was distributed into replicate microcosms (20 L each) that were incubated at in-situ temperatures and light intensities. The ambient concentration of dissolved organic carbon (DOC) in the unamended microcosms was 82 μM DOC, whereas the HMWDOM-amended microcosms contained 328 μM DOC, representing a 4-fold increase over ambient DOC concentration. Replicate control and experimental microcosms were sampled periodically over the course of a 27-h period.

The responses of microbial community members to HMWDOM addition over time were followed using flow cytometric, metagenomic, and metatranscriptomic analytical techniques. HMWDOM-induced shifts in microbial cell numbers, community composition, functional gene content, and gene expression were observed at each time point, as indicated by changes in the DOM-treated microcosms relative to an unamended control. The data indicated rapid and specific HMWDOM-induced shifts in transcription, metabolic pathway expression, and microbial growth that appear to be associated with HMWDOM turnover in ocean surface waters.

Results and Discussion

HMWDOM-Induced Cell Dynamics. Replicate microcosms were established immediately before sunrise and sampled over the course of 27 h to track the changes in microbial cell numbers, community composition, gene content, and gene expression in control vs. HMWDOM-treated microcosms. Though cell numbers in control microcosms remained constant over the time course of the experiment, the HMDOM-treated microcosm exhibited a ~50% increase in total cells within 19 h (Fig. 1*A*). Assuming a 50% growth efficiency, this HMWDOM-stimulated cell growth represents consumption of less than 1% of the total added DOC. Flow cytometry indicated that the majority (> 80%) of this increase in cells was attributable to the growth of a specific population of larger, high-DNA-content cells (Fig. 1*B*). The distinct flow cytometric signature of the HMWDOM-responsive population at the final time point allowed us to separate these large, high-DNA-content cells for further analyses (*SI Appendix*, Fig. S1). Large, high-DNA-content cells were isolated and collected via fluorescence-activated cell sorting and used to generate a SSU rRNA gene amplicon library. Near full-length rRNA gene sequences from the sorted cells recovered were all affiliated with the phylum Proteobacteria, falling into one of three clades (Fig. 1*C*). One subset of the flow-sorted cell population contained Alphaproteobacteria, closely related to *Thalassobius* isolates within the family Rhodobacteraceae. The remaining rRNA genes from the cell-sorted population were derived from Gammaproteobacteria, with one subset most closely related to *Alteromonas* isolates, and a second subset most similar to *Methylophaga* isolates within the order Thiotrichales.

Taxon-Specific Patterns of rRNA Gene and rRNA Representation in Control vs. HMWDOM-Treated Metagenomic and Metatranscriptomic Datasets. Community genomic DNA samples from T_0 and T_{27hrs} were pyrosequenced on the Roche 454 FLX platform, yielding ≈500,000 reads per sample (Table 1). Though SSU rDNA genes represent a small fraction (~1%) of the total genomic pyrosequencing reads, sufficient data (~500–750 individual reads) was available for phylogenetic analyses, which avoids PCR bias, and other artifacts associated with PCR amplicon "pyrotag" libraries (18–20). Classification of these of rRNA genes (*Methods*) provided

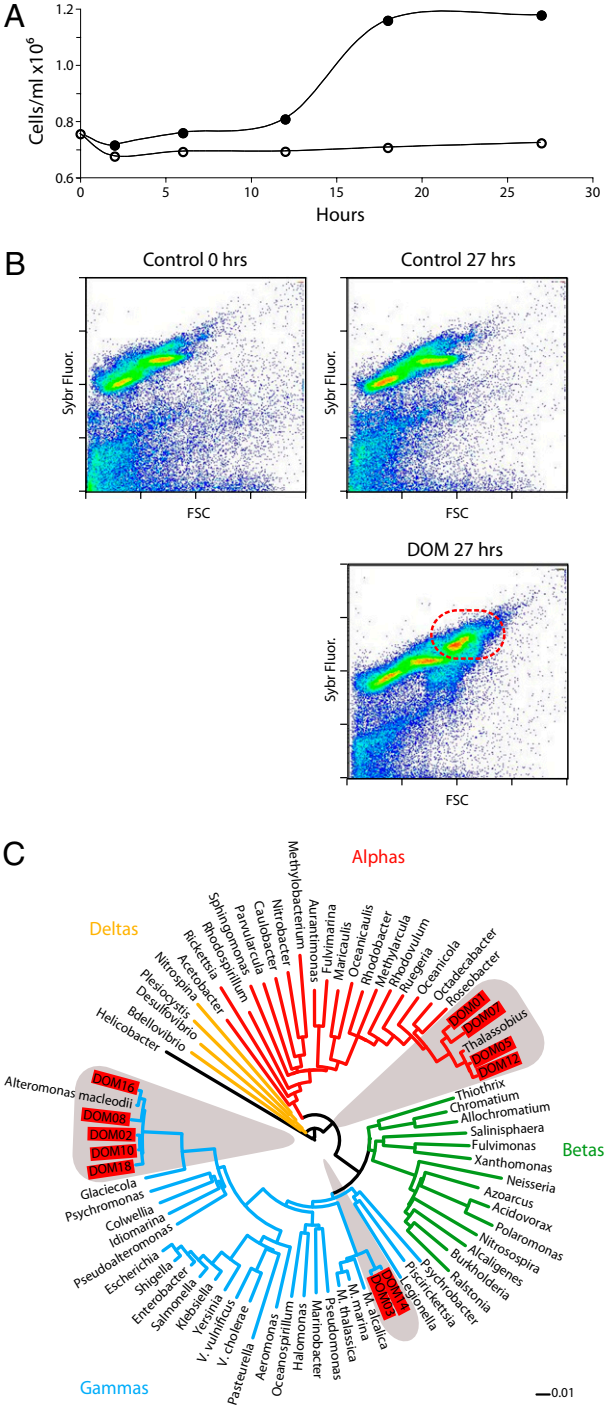

Fig. 1. Dynamics of microbial populations during 27-h microcosm incubations. (*A*) Flow cytometric counts of microbial cells from control (○) and DOM-amended (●) treatments. Samples displayed in *B* highlighted in red. (*B*) Flow cytometry scatterplots from selected samples show little change in the distribution of cell size [as measured by forward scatter (FSC)] and DNA content (SYBR fluorescence) of control samples from beginning to end of the experiment, whereas most of the increase in cell numbers observed in the DOM-amended treatment can be attributed to the appearance of larger, high-DNA-content cells (circled in red). (*C*) Weighted neighbor-joining tree of selected SSU SSU rDNA sequences from proteobacterial type strains and the sequences obtained from flow cytometric sorting of the larger, higher-DNA-content population of cells present after DOM amendment. The sequences obtained from the flow-sorted population are restricted to three specific taxonomic clades: Rhodobacteraceae, Methylophaga, and Alteromonas.

Table 1. Number of pyrosequences analyzed in control and treatment DNA and cDNA libraries

Treatment	Sample	0 h	2 h	12 h	27 h
Control	DNA	557,099	NA	NA	422,666
	cDNA	505,075	221,751	470,578*	514,670
	(non rRNA)	(18,345)	(12,658)	(12,934)	(18,078)
+DOM	DNA	NA	NA	NA	526,681
	cDNA	NA	230,376	251,690	751,284
	(non rRNA)	NA	(14,762)	(15,748)	(42,689)

*One of two technical replicate sequencing runs for this sample contained a spuriously high representation of a single sequence (~4.2% of reads) not present in the other replicate sequencing run. These nearly perfect duplicate reads (>99% nucleotide identity and read-length difference of <5 bp) were removed before subsequent analysis.

an overview of microbial community composition over the course of the experiment (Fig. 2*A*, inner rings). As expected, typically abundant planktonic bacterial taxa such as *Pelagibacter* (Rickettsiales) and *Prochlorococcus* (Cyanobacteria) were highly represented (Fig. 2*A* and *SI Appendix*, Fig. S2). The community composition of the control microcosm did not change substantially from the beginning to the end of the experiment. In contrast, the representation of several taxonomic groups increased in the HMWDOM-amended microcosm over the 27-h incubation. Three specific gammaproteobacterial groups—the families Idiomar-

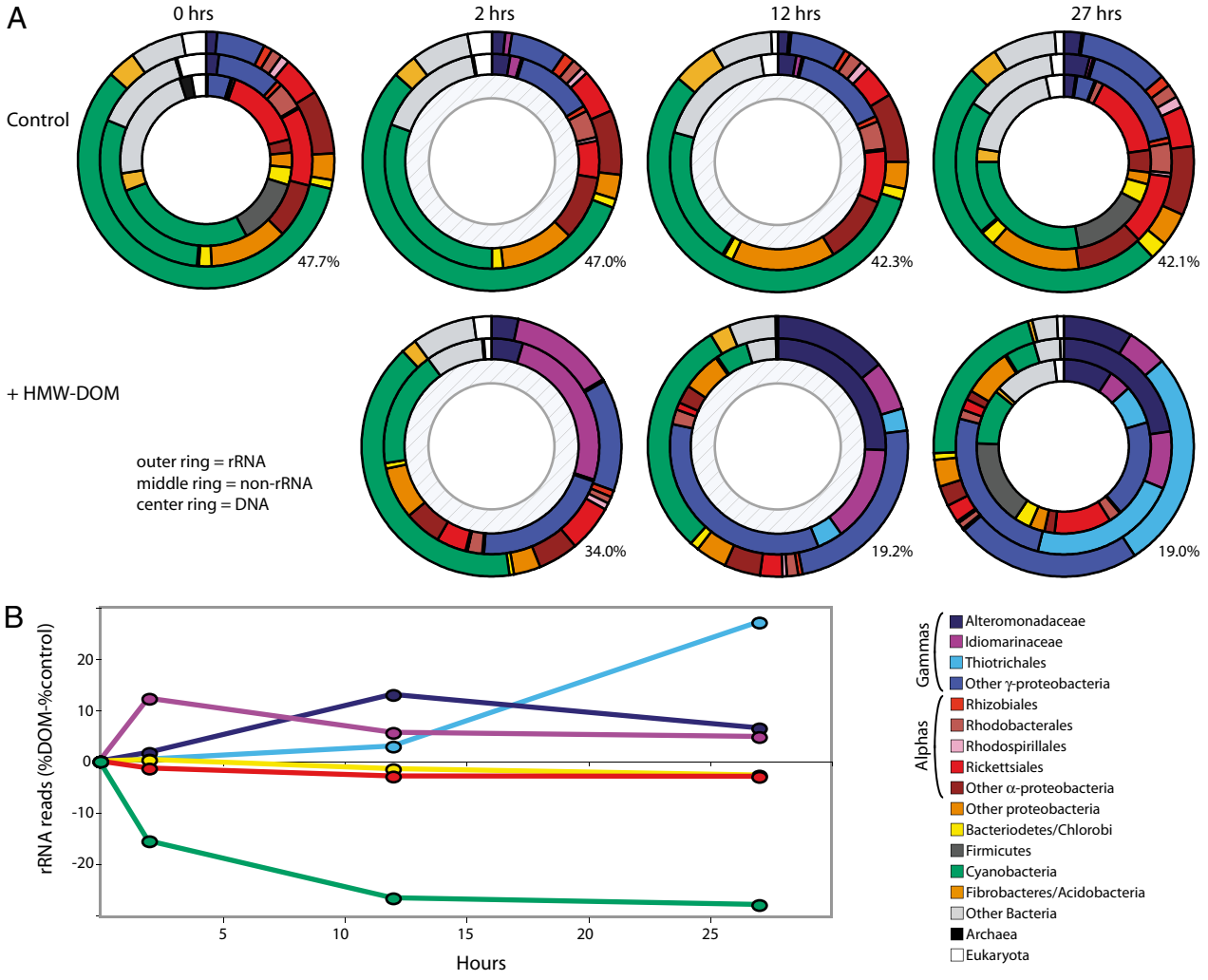

Fig. 2. Microbial community composition assessed by taxonomic classification of metagenomic and metatranscriptomic sequence reads. (*A*) SSU rRNA reads (outer ring) and non-rRNA reads (middle ring) from metatranscriptomic datasets as well as those reads from metagenomic datasets identified as SSU rDNA reads (center ring). Only taxonomic groups that represent >1% of total reads in at least one dataset have been included with all other groups binned together with unassigned reads. In some instances, reads can only be confidently assigned to broad class- and order-level taxonomic groups and are labeled as such. For mRNA datasets, some reads have no significant blast hits, the percentage of which is noted beside each sample. (*B*) Tracking the changes in community composition by comparing the difference between the DOM-amended treatment and control reveals distinct taxonomic groups responding at each time point. Only taxonomic groups showing more than ±2% change are plotted.

inaceae and Alteromonadaceae (both of which fall in the order Alteromonadales) and the order Thiotrichales—all increased in rRNA gene representation following HMWDOM amendment (Fig. 2 *A* and *B* and *SI Appendix*, Fig. S2). Two of these HMWDOM-stimulated groups (Alteromonadaceae and Thiotrichales) corresponded to the same dominant groups found in the FACS-sorted, high-DNA-containing cell populations (Fig. 1). The Rhodobacteraceae group that was recovered in the flow-sorted population did not, however, show a corresponding rRNA enrichment in the HMWDOM-treated metagenomic or metatranscriptomic datasets. These alphaproteobacteria may simply represent a background population of cells that were sorted along with the DOM-stimulated gammaproteobacteria because their flow cytometric signal overlapped with the large, high-DNA-content cell fraction.

Analyses of metagenomic sequence reads yields information on the relative representation of taxonomic groups, but not absolute cell numbers. Though cyanobacteria represented more than a quarter of all SSU rRNA genes throughout the time course of the experiment in the control microcosm, in the HMWDOM treatment they comprised only 10% of the rRNA sequence reads by 27 h. Enumeration of *Prochlorococcus* cells via flow cytometry indicated, however, that absolute *Prochlorococcus* cell numbers changed by less than 1% in the HMWDOM-amended microcosm. The changes in community composition observed in the metagenomic datasets therefore appear due to the growth of specific population members (in particular, Alteromonadaceae and Thiotrichales) and not to the disappearance of other dominant groups.

Compared with SSU rDNA reads from metagenomic DNA datasets, pyrosequencing of total community cDNA yielded orders of magnitude more total rRNA sequences that could be similarly classified taxonomically (Fig. 2*A*, outer rings). [The cDNAs in this study were not subjected to upstream rRNA subtraction procedures that have been reported in other metatranscriptomic studies (21–23).] In contrast to rRNA gene abundance in the DNA, rRNA in the cDNA pool reflects the cellular abundance of specific phylogentic groups, as well as their cellular rRNA copy numbers. For example, the rRNAs of several groups (e.g., Rickettsiales, Firmicutes, and Archaea) were less abundant in the cDNA datasets in comparison with their corresponding genes in the genomic DNA dataset (Fig. 2 and *SI Appendix*, Fig. S2). Conversely, cyanobacterial rRNAs were more highly represented in the cDNA than the corresponding rRNA genes in the DNA (Fig. 2 and *SI Appendix*, Fig. S2). Similarly, in the 27 h post-HMWDOM amendment, the Thiotrichales comprised nearly one-third of all SSU rRNA sequences in the cDNA, but represented less than 8% of all SSU rRNA genes in the DNA of the same sample.

Taxon-Specific Responses to HMWDOM Addition Inferred from Functional Gene Transcript Abundance. Taxonomic classification of non-rRNA transcripts from cDNA datasets (Fig. 2*A*, middle ring; *Methods*) generally paralleled the trends observed for rRNA taxon abundance, indicating parallel responses in both functional gene transcript and rRNAs (Fig. 2). Two exceptions to this correspondence were observed: cyanobacterial rRNA sequences were present in much greater abundance than non-rRNA cyanobacterial transcripts at all time points in both the control and the HMWDOM treatment. Conversely, Idiomarinaceae and Alteromondaceae were underrepresented in rRNAs, relative to non-rRNA transcripts present in the HMWDOM-treated microcosm cDNAs.

Distinct shifts in the cDNAs of specific subpopulations occurred in response to HMWDOM addition. Though the control remained virtually unchanged throughout the experiment, at each time point following HMWDOM addition, a different taxonomic group dominated the cDNA pool for both rRNA and non-rRNA transcripts (Fig. 2 *A* and *B*). Two hours post-HMWDOM amendment, Idiomarinaceae sequences represented nearly 13% of all rRNA sequences in the cDNAs from the HMWDOM treatment, though they remained less than 1% of the total rRNA sequences in all control cDNAs. By 12 h, the abundance of Idiomarinaceae rRNA sequences in the HMWDOM treatment receded closer to control values, whereas Alteromonadaceae rRNA sequences in the transcript pool rose to 15% of the total rRNAs relative to the control (Fig. 2*B*). Similarly, by the end of the experiment, Alteromonadaceae rRNA sequences decreased in relative abundance compared with earlier time points, when Thiotrichales-like rRNA represented the most abundant rRNAs. Strikingly, though Thiotrichales-like rRNAs represented approximately one-third of the total rRNA sequences in cDNA at the final HMWDOM-treated time point, Thiotrichales never represented more than 0.04% of in any of the controls at all time points.

Idiomarinaceae and Alteromonadaceae are closely related families within the order Alteromonadales (24). Because these closely related taxa were differentially represented at two different time points in the HMWDOM treatment, we searched for potential differences in their functional gene transcript representation at different times. All sequence reads having a best match to the full genome sequence of these two dominant taxa [*Idiomarina loihiensis* (25) and *Alteromonas macleodii* (26)] were analyzed separately for each taxonomic bin (*SI Appendix*, Tables S1 and S2). There were many similarities in the distribution of cDNA reads of functional gene categories between the two taxa. Examination of the 2-h and 12-h HMWDOM microcosm time points for Idiomarinaceae and Alteromonadaceae, respectively, indicated that transcript representation for many nutrient acquisition genes were similarly abundant within both taxonomic groups at the two different time points. An outer membrane receptor for a TonB-associated iron transporter was among the most abundant transcripts for both Idiomarinaceae and Alteromonadaceae. Similarly, the three genes require for the glutamine synthase cycle involved in nitrogen assimilation were abundant in each taxonomic bin. Genes involved in fatty acid catabolism were abundant in both Idiomarinaceae and Altermonadaceae bins (*SI Appendix*, Tables S1 and S2). Additionally, the two enzymes specific for the glyoxylate cycle (isocitrate lyase and malate synthase), which could use acetyl-CoA output by the β-oxidation of fatty acids, were abundant in both bins. One striking difference between the two different Alteromonadales cDNA bins was the high representation of one gene, triacylglycerol lipase (10-fold more abundant in treatment than control), found only among Idiomarinaceae-like reads. Interestingly, triacylglycerol lipase reads were virtually absent from reads assignable to the Alteromonadaceae bin.

The taxonomic groups that appeared most responsive to HMWDOM addition comprised only a small fraction of the starting microbial community. In contrast, transcripts from typically more dominant taxa such as *Pelagibacter* and *Prochlorococcus* decreased in relative abundance in the HMWDOM treatment over time. Additionally, because the differences in transcript abundance between control and treatment were small for *Prochlorococcus* and *Pelagibacter*, our sequencing depth allowed the detection of only a few significantly different transcripts between controls and treatments (*SI Appendix*, Figs. S3 and S4). Only seven *Pelagibacter* ORFs were identified as having statistically significant changes in transcript abundance ($P < 0.001$; *Methods*) in the HMWDOM-treated sample vs. the control (*SI Appendix*, Fig. S3). This small number of transcriptionally responsive ORFs (within our detection limits) was consistent with the hypothesis that *Pelagibacter* has a relatively small genome and streamlined regulatory network (27) and so may be less responsive to large fluctuations in ambient nutrient concentrations. The absolute *Pelagibacter* cell numbers appear to have increased slightly over the course of incubation in the treatment relative to the control, as evidenced by its higher gene abundances in the treatment relative to *Prochlorococcus* (whose absolute cell numbers remained constant as determined by flow cytometry; Fig. 2). The enrichment of transcripts encoding DNA-directed RNA polymerase and methionine biosynthesis protein (*SI Appendix*, Fig. S3) may indicate some utilization of some fraction of HMWDOM by *Pelagibacter*

cells to obtain reduced sulfur for the biosynthesis of sulfur-containing amino acids (28). The depletion of proteorhodopsin transcripts in the treatment at the final time point (*SI Appendix*, Fig. S3) suggested a potentially diminished requirement for proteorhodopsin phototrophy, with the increase in carbon availability. For *Prochlorococcus*, most of the significantly different transcripts were depleted in the treatment relative to the control at the earlier time points, whereas a few transcripts were enriched at the final time point. Several of these treatment-stimulated *Prochlorococcus* transcripts appeared to be involved with cellular repair processes, including oxidative damage protection and protein folding (*SI Appendix*, Fig. S4).

Small RNAs. Thirty putative sRNA (psRNA) clusters comprising >100 reads were identified, 20 of which showed statistically significant differences in abundance between the treatment and control for one or more time points (*SI Appendix*, Fig. S5). Based on the Rfam 10.0 database (http://rfam.sanger.ac.uk/), five clusters were identified as transfer-messenger RNA (tmRNA), and one was RNaseP RNA. Notably, all but one tmRNA cluster was overrepresented in the treatment, in part reflecting increases in specific taxa in the treatment vs. control (Fig. 1). For instance, cluster 7 tmRNA, which was overrepresented at 2 h, was most closely related to Idiomarinaceae, whereas *Methylophaga*-like cluster 9 tmRNA was enriched at later time points. Several psRNA clusters mapped into previously reported abundant psRNA groups found in microbial community transcripts sampled from the water column at Station ALOHA (29) (*SI Appendix*, Fig. S5). Five apparently different psRNA clusters (cluster 2, 3, 4, 8, and 14) were adjacent to genes encoding class II fumarate hydratase, an enzyme that catalyzes the reversible hydration/dehydration of fumarate to S-malate in the tricarboxylic acid cycle. To test the possibility that these clusters belonged to the same group but did not merge due to stringent clustering method, we performed pairwise alignment analysis among representative sequences of these five clusters (*SI Appendix*, Fig. S6). Only cluster 3 and cluster 14 merged (based on high sequence identity in the alignment at the end of both sequences), confirming that several divergent psRNA species, all adjacent to fumarate hydratase genes, were enriched in response to HMWDOM addition.

Global trends in functional gene transcript abundances in the HMWDOM treatment vs. control. All non-rRNA cDNA sequences were compared with NCBI-nr, KEGG (30), and GOS protein clusters databases (31) using BLASTX (32). We focused in particular on quantifying KEGG ortholog abundances in the HWM DOM-treated microcosm relative to the unamended controls across all time points (*SI Appendix*, Tables S3–S6).

Among all of the controls (0 h, 2 h, 12 h, and 27 h), only a few orthologs exhibited significant changes between time points ($n = 43$; *SI Appendix*, Table S3). Among these significantly different orthologs, about half were due to differences between the initial time point (0 h) and the other controls. In contrast, a larger number of orthologs exhibited differences in abundance between the pooled controls and the HMWDOM treatment (*SI Appendix*, Tables S4–S6). At 2 h post-HMWDOM addition, 67 KEGG orthologs exhibited differences from the control, with 58 of those enriched in the treatment vs. pooled controls (detectable effect sizes of enriched orthologs: 2.0- to 550-fold change; *SI Appendix*, Table S4). At 12 h, 221 differences were apparent, and 200 of those were enriched in the treatment vs. controls (detectable effect sizes of enriched orthologs: 2.3- to 2,200-fold change; *SI Appendix*, Table S5). At 27 h, 390 differences were detected, and 311 of those orthologs were enriched in the treatment (detectable effect sizes of enriched orthologs: 1.6- to 1,100-fold change; *SI Appendix*, Table S6).

Significantly enriched transcripts in the HMWDOM treatment included those encoding enzymes in KEGG pathways for carbohydrate, nitrogen, methane, sulfur, and fatty acid metabolic genes. Numerous transcripts associated with signal transduction and membrane transport pathways were also enriched in the HMWDOM treatment. Amino acid and nucleotide metabolism were also enriched in the HMWDOM addition microcosms, as were transcripts encoding enzymes involved in transcription and translation. The effect for all of these categories, however, was much more pronounced for the 12- and 27-h post-HMWDOM treatments than for the 2-h treatment. This is apparently due to the fact that the predominant DOM-responsive taxa were initially low in numbers, but increased in both cell density and transcriptional activity over the time course of the experiment.

At 12 h in the HMWDOM microcosm a variety of two-component sensor systems and several transporters were overrepresented. Particularly abundant were genes involved in nutrient acquisition. Specifically, both the components of the phosphate two-component sensor system (phoB, phoR, phoA, and OmpR phoB) as well as all components of the ABC transporter for phosphate (pstS, pstC, pstA, and pstB) were overrepresented at 12 and 27 h post-HMWDOM addition. At 27 h post-HMWDOM addition, members of several two-component sensor systems are enriched, including those associated with glucose (BarA, UvrY, CsrA), glucose-6-P (UhpB), nitrogen (GlnL, GlnG), C4-dicarboxylate (YfhK, YfhA), redox state of the quinone pool (ArcA), misfolded proteins (CpxR), carbon storage (BarA, UvrY, CsrA), and bacterial flagellar chemotaxis (CheA, CheV, CheY). Flagellar biosynthesis-associated transcripts were also similarly enriched, with 18 of 42 KOs associated with flagellar biosynthesis more the 4-fold more abundant in the amended microcosm relative to controls.

Transcripts encoding components of the GS/GOGAT pathway (glutamine and glutamate synthesis) were also significantly enriched in the HMWDOM treatment. Nitrogen two-component systems enriched in the DOM treatment transcript pool (GlnL, GlnG) typically sense nitrogen limitation via the intracellular glutamine pool and respond to nitrogen limitation by activating glutamate metabolism (33), which is consistent with the observed elevated GS/GOGAT transcript levels. Other enzymes in the nitrogen pathway, however, appeared relatively unchanged except for aminomethyltransferase (involved in glycine synthesis), which was less prevalent in the HMWDOM treatment. [Transcripts for one specific family of Amt family ammonium transporters from *Prochlorococcus* were significantly depleted in the HMWDOM treatment (*SI Appendix*, Fig. S4)]. Similar to the signatures of nitrogen limitation, the prevalence of the OmpR family phosphate two-component system, and the enrichment of a PIT family inorganic phosphate transporter, suggested that over the course of the experiment, the HMWDOM microcosm community was experiencing nitrogen and phosphate limitation as a consequence of the elevated DOC levels relative to the control.

Transcripts associated with sulfur-metabolizing enzymes were enriched in the HMWDOM treatment at the final time point and included enzymes associated with sulfate metabolism, and serine metabolism. Serine metabolism produces acetate that potentially could be shunted into the reductive carboxylate cycle, also enriched in the DOM treatment. Transcripts encoding three enzymes of the fatty acid metabolism pathway were also enriched in the HMWDOM treatment, as well as those encoding a short-chain fatty acid transporter. Furthermore, fatty acid biosynthesis pathway transcripts were significantly depleted in the HMWDOM treatment, suggesting a potential shift to catabolic metabolism of fatty acid-like molecules in the HMWDOM treatment. At the first time point, just 2 h postamendment, the two most enriched transcripts that corresponded to KEGG orthologs were triacylglycerol lipase and acyl-CoA dehydrogenase (50-fold and 109-fold, respectively). These enzymes catalyze two early steps in the catabolism of triacylglycerols (TAGs). These signals may be the result of cell wall material copartitioning in the HMWDOM concentrate, or the tendency of lipid compounds to associate with HMWDOM concentrates (34).

Methylophaga species were the most highly represented single taxon in both rRNA and functional gene transcripts in the

HMWDOM microcosm at the final time point. Consistent with this observation, two key enzymes involved in the ribulose monophosphate (RuMP) pathway, hexulose-6-phosphate synthase and 6-phospho-3-hexuloisomerase, were also highly abundant in the amended microcosm (eighth and second most abundant, respectively) while remaining undetected in the control. The cyclical RuMP pathway is an assimilatory pathway that is widespread in bacteria, functioning as a pathway for formaldehyde fixation and detoxification. In the first two reactions in this pathway, formaldehyde is condensed with ribulose-5 phosphate, which is then isomerized to fructose-6-phosphate. Moreover, gene transcripts for the enzymes encoding many of the steps in this pathway were enriched by the end of this experiment (Fig. 3) and increased over the time course of the experiment (Poisson ANOVA; *SI Appendix*, Table S7). Though a large variety of one-carbon compounds are processed through the RuMP pathway, all methyltrophic pathways share formaldehyde as a common entry point. Formaldehyde can also be oxidized to CO_2 via several routes, and several of the enzymes involved in these dissimilatory pathways were also abundant in the amended treatment (Fig. 3), particularly those associated with the tetrahydromethanopterin-dependent pathway. In total, the data reflected the enrichment of pathways for both assimilatory and dissimilatory single-carbon compound utilization, which coincided with the appearance of an actively growing *Methylophaga* population in the HMWDOM treatment (Figs. 2 and 3).

Conclusions

Semilabile DOM may support up to 40% of marine bacterial carbon demand (35, 36), yet little is known about the specific microorganisms and metabolic pathways responsible for its degradation and transformation in the ocean's water column. There is growing evidence that microbial transformation of semilabile DOM renders DOM less and less labile, further increasing accumulation in oligotrophic gyres and ultimately leading to export as refractory DOM (36). Microbial population dynamics and metabolic processes are therefore central to understanding the cycling of DOM in the sea.

In this study, short-term incubation of bacterial populations from surface seawater with naturally occurring HMWDOM from the same environment revealed specific shifts in microbial cells, rRNAs, and DOM-responsive gene transcripts relative to unamended controls. Cell numbers nearly doubled specifically in response to HMWDOM. Flow sorting and rRNA gene and transcript abundances consistently indicated the stimulation of several phylogenetic groups within the Alteromondales (*Idiomarina* and *Alteromonas* sp.) and Thiotrichales (*Methylophaga* sp.). Analysis of microbial cDNA abundances over time via pyrosequencing revealed that 2 h after DOM addition, close relatives of *Idioma-*

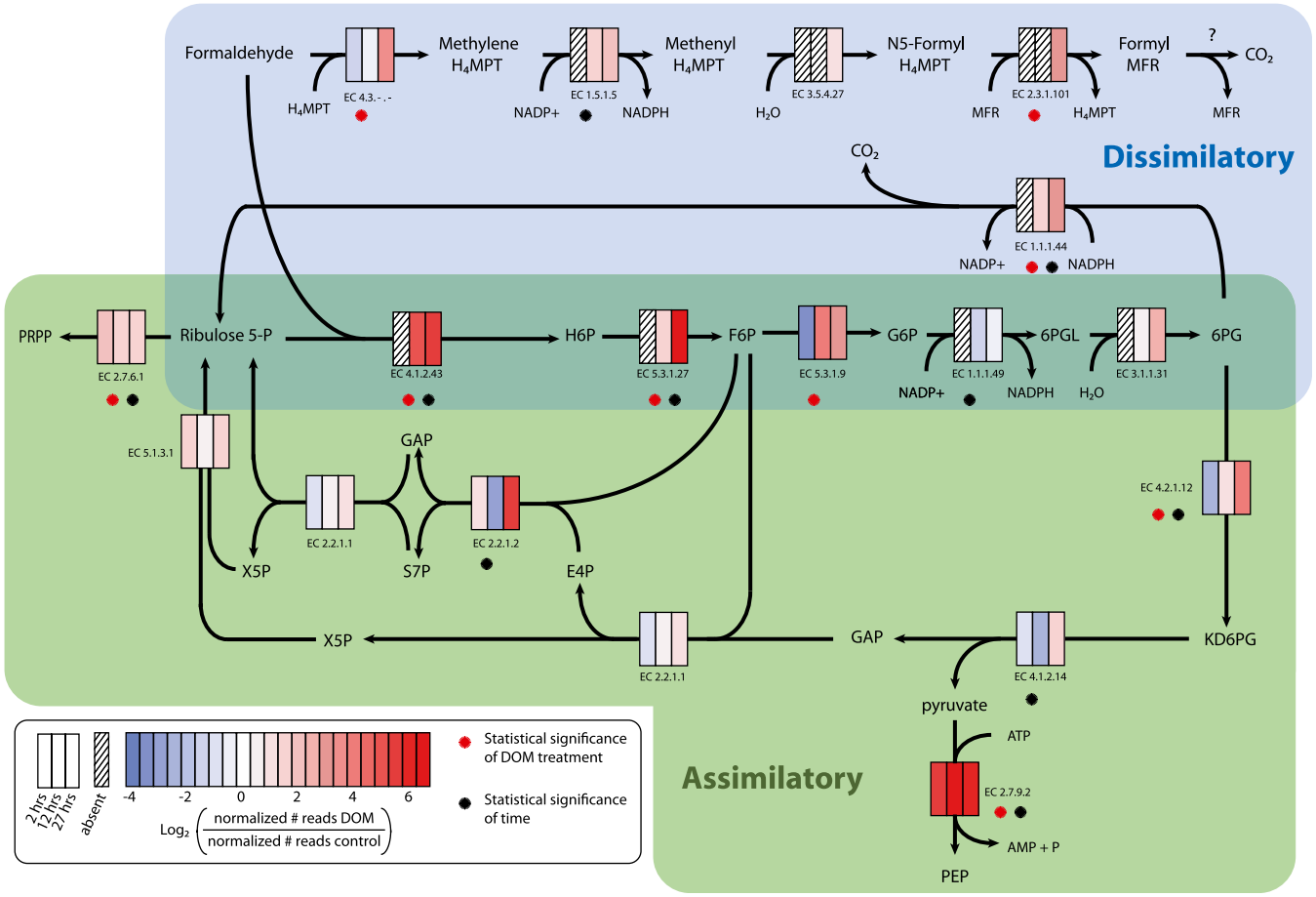

Fig. 3. Diagram of representative dissimilatory and assimilatory methylotrophic pathways and enzymes that show increased transcript abundance following DOM amendment. A KEGG ortholog-based expression ratio comparing normalized abundances of reads present in the DOM-amended treatment with those from an untreated control at 2, 12, and 27 h following DOM addition. Asterisks mark those enzymes showing statistically significantly differences in transcript abundance relative to time and/or unamended control (*SI Appendix*, Table S7). H_4MPT, tetrahydromethanopterin; MFR, formylmethanofuran; H6P, hexulose-6-phosphate; F6P, fructose-6-phosphate; 6PGL, 6-phosphogluconolactone; 6PG, 6-phosphogluconate; KD, ketodeoxy; PEP, phosphoenolpyruvate; GAP, glyceraldehyde phosphate; E4P, erythrose-4-phosphate; X5P, xylulose-5-phosphate; S7P, sedoheptulose-5-phosphate; PRPP, phosphoribosyl diphosphate.

rina sp. were stimulated by HMWDOM. In apparent microbial succession, a few hours later, *Alteromonas macleodii*-like rRNAs and mRNAs increased dramatically relative to the unamended control. After 27 h, the same indicators showed that *Methylophaga* sp. (order Thiotrichales) predominated. We interpret this succession as a specific metabolic sequence and successional cascade that reflects sequential processing and degradation of specific components within HMWDOM. Analyses also indicated that 27 h post-DOM addition, both the dissimilatory and assimilatory single-carbon compound utilization pathways were highly expressed, coincident with the appearance and high abundance of *Methylophaga* sp. at the final time point.

The data indicate several specific groups of bacteria that appear to operate in succession and synergy to catalyze the turnover of naturally occurring HMWDOM in the marine environment. These findings may reflect regular (and predictable) metabolic cascades and community succession patterns that in part regulate the transformation and turnover of naturally occurring semilabile DOM. Furthermore, our findings are suggestive of some of the chemical attributes and degradation patterns of naturally occurring DOM. In previous chemical analyses, about 15% of DOM carbohydrate has been shown to consist of methyl sugars (37, 38). Our present findings suggest that Alteromonadales (specifically, *Idiomarina* spp. and *Alteromonas macleodii*) might be metabolizing semilabile DOM methyl sugars to methanol or formaldehyde, and carbon dioxide, among other products. The methanol and/or formaldehyde produced could be further oxidized and incorporated by *Methylophaga* sp. in the terminal portion of this aerobic food chain. Such a specific carbon compound-driven syntrophy has rarely been observed in aerobic microbial consortia. Although confirmation awaits further experimentation and chemical analyses, if correct, DOM methyl sugar metabolism might provide a partial explanation for the ubiquitous presence of methylotrophs in open-ocean and coastal environments (12, 39–42).

In summary, the experimental metatranscriptomic approach described here is beginning to reveal metabolic pathways and microbial taxa involved in the chemical transformation and turnover of naturally occurring marine DOM. These techniques can be used to track a variety of microbial processes in the environment, and set the stage for future inquiries on the nature and details of microbial community environmental responses and dynamics in situ. In this study, we gained detailed perspective on microbial community dynamics and metabolism associated with the ocean carbon cycle in marine surface waters. The apparent resource partitioning of DOM by different bacterial species that was suggested by the data supports the significance of microbial community dynamics in the ocean's carbon cycle. The findings also underscore the importance of describing microbial synergistic interactions and population dynamics occurring on relatively short time-scales of hours to days.

Methods

Microcosm Setup and Biomass Sampling. Seawater for microcosm incubation experiments was collected (23°12.88′ N, 159°8.17′ W) from 75-m depth, predawn, on August 16, 2007, during the Center for Microbial Oceanography: Research and Education (C-MORE) BLOOMER Cruise. See *SI Appendix* for further details on the seawater collection and microcosm preparation.

HMW DOM Preparation. Surface seawater obtained from the uncontaminated underway system of the R/V *Kilo Moana* was filtered to remove microbes and small particles using a clean (10% HCl overnight soak), 0.2-μm Whatman Polycap TC polyether sulfone capsule filter. HMWDOM was concentrated using a custom-built ultrafiltration apparatus equipped with a stainless-steel membrane housing and centripetal pump along with a fluorinated high-density polyethylene reservoir. The system was plumbed with Teflon tubing and PVDF valves, and fitted with a dual thin-film ultrafiltration membrane element (Separation Engineering). The membrane has a 1-nm pore size that nominally retains organic matter of a molecular weight greater than 1,000 Da (>98% rejection of vitamin B_{12}). Membranes were precleaned with 0.01 mol L^{-1} hydrochloric acid (overnight wash) and 0.01 mol L^{-1} sodium hydroxide (overnight wash), and rinsed with copious amounts of distilled water until the pH returned to neutral. Membranes were flushed with 100 L of seawater for 45 min just before sample collection. Surface seawater (2,000 L) was concentrated 100-fold over a period of 24 h. Samples were taken for DOC quantification from the inflow and permeate during ultrafiltration, and of the concentrate upon completion. A 2-L subsample of the concentrate was prefiltered using a 0.2-μm Polycap TC filter (Whatman) before filtration through a prerinsed 30-kDa Ultracel regenerated cellulose membrane loaded in a high-output stirred cell (Millipore) to remove viral particles.

Dissolved Organic Carbon. DOC samples of 30 mL were transferred into combusted (450 °C for 8 h) glass vials and acidified with 150 mL of a 25% phosphoric acid solution before sealing with acid-washed Teflon septa and storage at 4 °C until processing. Analysis was performed using the high-temperature combustion method on a Shimadzu TOC-VCSH with platinized alumina catalyst. Sample concentrations were determined alongside potassium hydrogen phthalate standards and consensus reference materials (CRM) provided by the DOC-CRM program (http://www.rsmas.miami.edu/groups/biogeochem/CRM.html).

Flow Cytometry and Cell Sorting. At each time point, 1 mL of seawater was preserved with 0.125% glutaraldehyde (final concentration), frozen in liquid nitrogen, and stored at –80 °C for subsequent flow cytometric analysis and cell sorting using an Influx (Becton Dickinson). Before counting and sorting, samples were stained with SYBR Green (Invitrogen) for 15 min, and DNA-containing cells were identified based on fluorescence and scatter signals (43). See *SI Appendix* for further details on cell sorting and rRNA amplicon sequencing from the sorted population.

RNA Amplification and cDNA Synthesis. Metatranscriptome analyses were performed as previously described (44) with minor modifications. Briefly, 100 ng of total RNA was amplified using MessageAmp II (Ambion) following the manufacturer's instructions and substituting the T7-BpmI-$(dT)_{16}$VN oligo (44) in place of that supplied with the kit. Amplified RNA was then reverse transcribed into cDNA using SuperScript Double-Stranded cDNA Synthesis kit (Invitrogen) and random hexamer priming. Last, the cDNA was digested with BpmI and used for pyrosequencing. See *SI Appendix* for further details on pyrosequencing.

Bioinformatic Analyses. Full-length SSU rDNA amplicon sequences from flow-sorted cells were classified using both the Greengenes (45) NAST aligner and the Ribosomal Database Project (RDP) naïve Bayesian classifier (46). Resulting alignments were compared with the SILVA (47) databases using ARB (48). RDP classifier results were compared also with type strains using tools available at the RDP (49) and Interactive Tree of Life web sites (50).

cDNA datasets were parsed to separate rRNA sequences from the remaining non-rRNA sequences. rRNA sequences were identified as previously described (44) using a bit-score cutoff of 40 for BLASTN (32) searches against a custom 5S, SSU, 18S, 23S, and 28S rRNA databases. Non-rRNA sequences were compared with NCBI-nr, KEGG, and GOS protein clusters databases using BLASTX (32) for functional gene analyses as previously described (29, 44). See *SI Appendix* for further details.

Statistical Analyses. Statistical analyses were conducted on KEGG ortholog groups using the packages DegSeq (51) and ShotgunFunctionalizeR (52) in the R Statistical Package (53). In all statistical analyses, we assumed that the data (counts for a particular KEGG ortholog group) followed a Poisson sampling distribution. Analyses were conducted at the individual gene level as well as at the pathway level. See *SI Appendix* for further details on statistical analyses.

Accession Numbers. All 454 FLX pyrosequencing .sff files have been deposited in the GenBank database under accession no. SRA020733.11. Full-length SSU SSU rRNA sequences obtained from flow-sorted cells have been deposited to the GenBank/EMBL/DDJB databases under accession nos. HQ012268-HQ012278.

ACKNOWLEDGMENTS. We thank the captain and crew of the R/V *Kilo Moana* for facilitating sample collection, Chief Scientist Ricardo Letelier and all participants of the C-MORE BLOOMER cruise for help and encouragement, and Rachel Barry for pyrosequence library production and sequencing. This work was supported by the Gordon and Betty Moore Foundation (E.F.D., S.W.C., and D.J.R.), the Office of Science–Biological and Environmental Research, US Department of Energy (E.F.D and S.W.C), the National Science Foundation (D.J.R.), and National Science Foundation Science and Technology Center Award EF0424599 (to E.F.D. and S.W.C.). This article is a contribution from the National Science Foundation Science and Technology Center for Microbial Oceanography: Research and Education (C-MORE).

1. Hedges JI (1992) Global biogeochemical cycles: Progress and problems. *Mar Chem* 39: 67–93.
2. Ogawa H, Amagai Y, Koike I, Kaiser K, Benner R (2001) Production of refractory dissolved organic matter by bacteria. *Science* 292:917–920.
3. Pomeroy LR (1974) Oceans food web, a changing paradigm. *Bioscience* 24:499–504.
4. Azam F, et al. (1983) The ecological role of water-column microbes in the sea. *Mar Ecol Prog Ser* 10:257–263.
5. Azam F (1998) Microbial control of oceanic carbon flux: The plot thickens. *Science* 280: 694–696.
6. Ducklow H (1999) The bacterial component of the oceanic euphotic zone. *FEMS Microbiol Ecol* 30:1–10.
7. Benner R, Pakulski JD, McCarthy M, Hedges JI, Hatcher PG (1992) Bulk chemical characteristics of dissolved organic matter in the ocean. *Science* 255:1561–1564.
8. Church MJ, Ducklow HW, Karl DM (2002) Multiyear increases in dissolved organic matter inventories at station ALOHA in the North Pacific Subtropical Gyre. *Limnol Oceanogr* 47:1–10.
9. Carlson C, et al. (2002) Effect of nutrient amendments on bacterioplankton production, community structure, and DOC utilization in the northwestern Sargasso Sea. *Aquat Microb Ecol* 30:19–36.
10. Carlson C, et al. (2004) Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnol Oceanogr* 49:1073–1083.
11. Pinhassi J, Berman T (2003) Differential growth response of colony-forming alpha- and gamma-proteobacteria in dilution culture and nutrient addition experiments from Lake Kinneret (Israel), the Eastern Mediterranean Sea, and the Gulf of Eilat. *Appl Environ Microbiol* 69:199–211.
12. Pinhassi J, et al. (2004) Changes in bacterioplankton composition under different phytoplankton regimens. *Appl Environ Microbiol* 70:6753–6766.
13. Schafer H, et al. (2001) Microbial community dynamics in Mediterranean nutrient-enriched seawater mesocosms: Changes in the genetic diversity of bacterial populations. *FEMS Microbiol Ecol* 34:243–253.
14. Allers E, et al. (2007) Response of Alteromonadaceae and Rhodobacteriaceae to glucose and phosphorus manipulation in marine mesocosms. *Environ Microbiol* 9:2417–2429.
15. Cecilia A, Jakob P (2006) *Roseobacter* and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environ Microbiol* 8:2022–2030.
16. Hansell DA, Carlson CA (2001) Biogeochemistry of total organic carbon and nitrogen in the Sargasso Sea: Control by convective overturn. *Deep Sea Res Part II Top Stud Oceanogr* 48:1649–1667.
17. Morris RM, et al. (2005) Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-Series Study site. *Limnol Oceanogr* 50:1687–1696.
18. Quince C, et al. (2009) Accurate determination of microbial diversity from 454 pyrosequencing data. *Nat Methods* 6:639–641.
19. Kunin V, Engelbrektson A, Ochman H, Hugenholtz P (2010) Wrinkles in the rare biosphere: Pyrosequencing errors can lead to artificial inflation of diversity estimates. *Environ Microbiol* 12:118–123.
20. Turnbaugh PJ, et al. (2010) Organismal, genetic, and transcriptional variation in the deeply sequenced gut microbiomes of identical twins. *Proc Natl Acad Sci USA* 107: 7503–7508.
21. Stewart FJ, Ottesen EA, Delong EF (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J* 4: 896–907.
22. Vila-Costa M (2010) Transcriptomic analysis of a marine bacterial community enriched with dimethylsulfoniopropionate. *ISME J*, 10.1038/ismej.2010.62.
23. Gilbert JA, et al. (2008) Detection of large numbers of novel sequences in the metatranscriptomes of complex marine microbial communities. *PLoS ONE* 3:e3042.
24. Ivanova EP, Flavier S, Christen R (2004) Phylogenetic relationships among marine Alteromonas-like proteobacteria: Emended description of the family Alteromonadaceae and proposal of Pseudoalteromonadaceae fam. nov., Colwelliaceae fam. nov., Shewanellaceae fam. nov., Moritellaceae fam. nov., Ferrimonadaceae fam. nov., Idiomarinaceae fam. nov. and Psychromonadaceae fam. nov. *Int J Syst Evol Microbiol* 54:1773–1788.
25. Hou S, et al. (2004) Genome sequence of the deep-sea gamma-proteobacterium Idiomarina loihiensis reveals amino acid fermentation as a source of carbon and energy. *Proc Natl Acad Sci USA* 101:18036–18041.
26. Ivars-Martinez E, et al. (2008) Comparative genomics of two ecotypes of the marine planktonic copiotroph Alteromonas macleodii suggests alternative lifestyles associated with different kinds of particulate organic matter. *ISME J* 2:1194–1212.
27. Giovannoni SJ, et al. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309:1242–1245.
28. Tripp HJ, et al. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* 452:741–744.
29. Shi Y, Tyson GW, DeLong EF (2009) Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. *Nature* 459:266–269.
30. Kanehisa M, et al. (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 36:D480–D484.
31. Yooseph S, et al. (2007) The Sorcerer II Global Ocean Sampling expedition: Expanding the universe of protein families. *PLoS Biol* 5:e16.
32. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
33. Zimmer DP, et al. (2000) Nitrogen regulatory protein C-controlled genes of Escherichia coli: Scavenging as a defense against nitrogen limitation. *Proc Natl Acad Sci USA* 97:14674–14679.
34. Mannino A, Harvey HR (1999) Lipid composition in particulate and dissolved organic matter in the Delaware Estuary: Sources and diagenetic patterns. *Geochim Cosmochim Acta* 63:2219–2235.
35. Repeta DJ, Aluwihare LI (2006) High molecular weight dissolved organic carbon cycling as determined by natural abundance radiocarbon measurements of neutral sugars. *Limnol Oceanogr* 51:1045–1053.
36. Lou Y-W, Friedrichs MAM, Doney SC, Church MJ, Ducklow HW (2010) Oceanic heterotrophic bacterial nutrition by semilabile DOM as revealed by data assimilative modeling. *Aquat Microb Ecol* 60:273–287.
37. Panagiotopoulos C, Repeta DJ, Johnson CG (2007) Characterization of methyl sugars, 3-deoxysugars and methyl deoxysugars in marine high molecular weight dissolved organic matter. *Org Geochem* 38:884–896.
38. Quan TM, Repeta DJ (2007) Characterization of high molecular weight dissolved organic carbon using periodate over-oxidation. *Mar Chem* 105:183–193.
39. Lidstrom ME (2006) Aerobic methylotrophic prokaryotes. *The Prokaryotes*, ed Dworkin M (Springer, New York), 3rd Ed, pp 618–634.
40. Neufeld JD, Boden R, Moussard H, Schafer H, Murrell JC (2008) Substrate-specific clades of active marine methylotrophs associated with a phytoplankton bloom in a temperate coastal environment. *Appl Environ Microbiol* 74:7321–7328.
41. Neufeld JD, Chen Y, Dumont MG, Murrell JC (2008) Marine methylotrophs revealed by stable-isotope probing, multiple displacement amplification and metagenomics. *Environ Microbiol* 10:1526–1535.
42. Neufeld JD, et al. (2007) Stable-isotope probing implicates Methylophaga spp and novel Gammaproteobacteria in marine methanol and methylamine metabolism. *ISME J* 1:480–491.
43. Marie D, Partensky F, Jacquet S, Vaulot D (1997) Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. *Appl Environ Microbiol* 63:186–193.
44. Frias-Lopez J, et al. (2008) Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci USA* 105:3805–3810.
45. DeSantis TZ, et al. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–5072.
46. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267.
47. Pruesse E, et al. (2007) SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35: 7188–7196.
48. Ludwig W, et al. (2004) ARB: A software environment for sequence data. *Nucleic Acids Res* 32:1363–1371.
49. Cole JR, et al. (2009) The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37:D141–D145.
50. Letunic I, Bork P (2007) Interactive Tree of Life (iTOL): An online tool for phylogenetic tree display and annotation. *Bioinformatics* 23:127–128.
51. Wang L, Feng Z, Wang X, Wang X, Zhang X (2010) DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26: 136–138.
52. Kristiansson E, Hugenholtz P, Dalevi D (2009) ShotgunFunctionalizeR: An R-package for functional comparisons of metagenomes. *Bioinformatics* 25:2737–2738.
53. R Development Core Team (2010) *R: A language and environment for statistical computing, version 2.11.1* (R Foundation for Statistical Computing, Vienna). Available at http://www.r-project.org/.