environmental microbiology

Environmental Microbiology (2018) 20(8), 3012–3030 doi:10.1111/1462-2920.14344

Identifying labile DOM components in a coastal ocean through depleted bacterial transcripts and chemical signals

Alexey Vorobev,[1†] Shalabh Sharma,[1] Mengyun Yu,[2] Juhyung Lee,[2] Benjamin J. Washington,[2] William B. Whitman,[3] Ford Ballantyne IV,[4] Patricia M. Medeiros[1] and Mary Ann Moran[1*]

[1]*Department of Marine Sciences, University of Georgia, Athens, GA, USA.*

[2]*Department of Statistics, University of Georgia, Athens, GA, USA.*

[3]*Department of Microbiology, University of Georgia, Athens, GA, USA.*

[4]*Odum School of Ecology, University of Georgia, Athens, GA, USA.*

Summary

Understanding which compounds comprising the complex and dynamic marine dissolved organic matter (DOM) pool are important in supporting heterotrophic bacterial production remains a major challenge. We eliminated sources of labile phytoplankton products, advected terrestrial material and photodegradation products to coastal microbial communities by enclosing water samples *in situ* for 24 h in the dark. Bacterial genes for which expression decreased between the beginning and end of the incubation and chemical formulae that were depleted over this same time frame were used as indicators of bioavailable compounds, an approach that avoids augmenting or modifying the natural DOM pool. Transport- and metabolism-related genes whose relative expression decreased implicated osmolytes, carboxylic acids, fatty acids, sugars and organic sulfur compounds as candidate bioreactive molecules. FT-ICR MS analysis of depleted molecular formulae implicated functional groups ~ 30–40 Da in size cleaved from semi-polar components of DOM as bioreactive components. Both gene expression and FT-ICR MS analyses indicated higher lability of compounds with sulfur and nitrogen heteroatoms. Untargeted methodologies able to integrate biological and chemical perspectives can be effective strategies for characterizing the labile microbial metabolites participating in carbon flux.

Received 4 April, 2018; accepted 27 June, 2018. *For correspondence. E-mail mmoran@uga.edu; Tel. +706-542-6481; Fax: +706-542-5888. †Present address: CEA - Institut de Biologie François Jacob, Genoscope, 2 Rue Gaston Crémieux, 91057 Evry, France.

Introduction

Marine dissolved organic matter (DOM) is a highly diverse mixture of molecules that serves as a major reservoir of the global carbon cycle, amounting to 660 Gt of carbon (Hansell *et al*., 2012; Hansell, 2013). Heterotrophic bacteria are the primary mediators of carbon flux through this reservoir, taking up organic matter for respiration and secondary production (Azam *et al*., 1983). However, not all DOM molecules participate equally in bacterially mediated carbon cycling because of inherent differences in their biological reactivity.

The rapidity with which molecules are processed from the DOM pool by marine bacteria falls along a continuum from minutes to millennia (Hansell, 2013; Follett *et al*., 2014). Typically, molecules are classified into three broad categories of biological lability based on the time between production and consumption. 'Labile' DOM consists of the compounds that are rapidly consumed by heterotrophic bacteria, having half-lives on the order of minutes to days after production. 'Semi-labile' DOM is less biologically reactive and persists in the surface ocean for weeks to years before biological consumption (Hansell and Carlson, 1998; Hansell, 2013). 'Refractory' DOM is the least biologically reactive and circulates through the ocean for thousands of years (Williams and Druffel, 1987; Follett *et al*., 2014). The specific molecules that fall into these three bulk categories are not yet well described, in large part because of the extreme chemical diversity of the DOM pool. Current data suggest there may be hundreds of thousands of distinct molecules in the seawater reservoir (Kim *et al*., 2003; Hertkorn *et al*., 2006). Further, there is limited knowledge and considerable debate about the chemical properties that determine their biological persistence (Jiao *et al*., 2010; Arrieta *et al*., 2015; Dittmar, 2015; Lechtenfeld *et al*., 2015). Finally, the methodological

difficulty of separating organic molecules from a seawater matrix is a challenge that remains particularly troublesome for the analysis of small polar compounds (Simjouw *et al*., 2005; Moran *et al*., 2016).

For the molecules that make up the most biologically labile components of the DOM pool, an additional impediment to their identification is that rapid bacterial utilization draws down their concentrations to extremely low levels. For example, dissolved monomers such as amino acids and sugars have concentrations in the range of a few nanomoles per litre (Mopper *et al*., 1992; Kaiser and Benner, 2012), just at or below the limits of chemical detection, yet they may represent a substantial fraction of labile DOM taken up by bacteria (Hodson *et al*., 1981; Hollibaugh and Azam, 1983). Kept at very low concentrations by efficient bacterial uptake, these biologically labile molecules blend into the complex chemical background of seawater and may change little in concentration despite rapid fluxes.

One approach to measuring the production and consumption of low concentration/high flux compounds is to use members of natural bacterial communities as sensors for the molecules that make up labile DOM. A recent application of this strategy leverages bacterial expression of transport and metabolism genes to suggest the identity of biologically labile compounds (McCarren *et al*., 2010; Poretsky *et al*., 2010; Gifford *et al*., 2013; Bergauer *et al*., 2018). Here, we made a modification to this gene expression approach based on the idea that genes whose expression decreases rapidly following isolation from a DOM source indicate the compounds that are most quickly depleted without ongoing inputs. A benefit of this approach is that the natural DOM pool is not modified or augmented.

For two tidal stages in each of two seasons in southeastern U.S. coastal waters, we characterized bacterial gene expression immediately upon sample collection and again following an *in situ* incubation of the sample for 24 h. During the incubation, samples were contained in 20 L carboys in the dark, which eliminated sources of new DOM derived from recent photosynthate or advection of terrestrial material, as well as photochemical processes. Biologically labile molecules were operationally defined as those for which transport or metabolism transcripts decreased significantly during the dark incubation (Fig. 1). It was expected that the bacterial communities were also processing molecules from more refractory categories of DOM during the assay, but gene expression changes involving these compounds should be minor compared with those for highly labile molecules because they are less affected by cessation of new inputs. It was also expected that microbial lysis, grazing and excretion activities were contributing labile molecules, but the supply of these compounds should not be interrupted by a dark incubation (Fig. 1).

Transcriptomic data were complemented with Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR MS) analyses of depleted molecular formulae in the same incubations. This analysis provided information on biologically-labile components of semi-polar DOM that adhere to solid phase extraction (SPE) resins. We characterized DOM by FT-ICR MS immediately upon sample collection and again following the 24 h dark incubation, and identified depleted formulae. Here, labile components

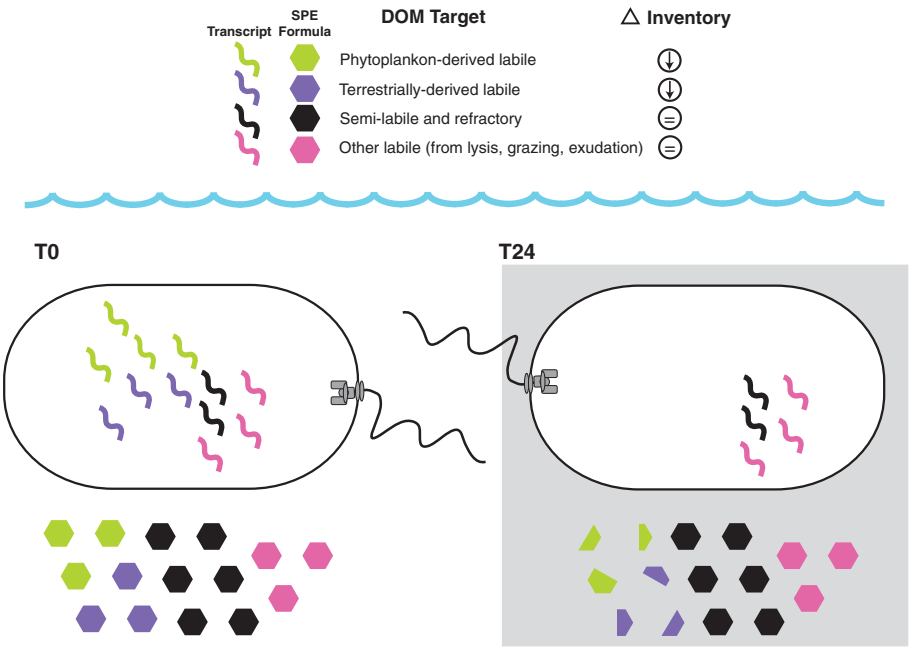

Fig. 1. Conceptual basis for interpreting depleted bacterial transcripts and molecular formulae as indicators of biologically labile components of marine DOM. T0, initial sample; T24, incubated sample; SPE, solid-phase extraction.

Table 1. Chemical and biological data for samples from Sapelo Island, GA, USA in July and October 2014.

Sample	Date	Salinity	DOC (μM)	Transcripts L^{-1}	Bacteria L^{-1}	Transcripts $cell^{-1}$	Number of Formulae by FT-ICR MS
July LT T0	7/23/2014	30	249 ± 8	$4.6 \times 10^{10} \pm 9.9 \times 10^{8}$	$2.9 \times 10^{9} \pm 5.1 \times 10^{7}$	15.9 ± 0.0	5370 ± 42
July LT T24	7/24/2014		251 ± 9	$6.0 \times 10^{10} \pm 1.1 \times 10^{9}$	$2.3 \times 10^{9} \pm 2.9 \times 10^{7}$	26 ± 0.8	5396 ±21
July HT T0	7/24/2014	33	194 ± 7	2.5×10^{11}	2.7×10^{9}	88.6	5471 ± 52
July HT T24	7/25/2014		190 ± 7	$1.5 \times 10^{11} \pm 6.9 \times 10^{10}$	$2.4 \times 10^{9} \pm 1.1 \times 10^{8}$	61.5 ± 26.4	5472 ± 47
October LT T0	10/14/2014	27	325 ± 4	$7.2 \times 10^{10} \pm 3.7 \times 10^{9}$	$1.7 \times 10^{9} \pm 1.9 \times 10^{8}$	42.1 ± 6.9	5555 ± 115
October LT T24	10/15/2014		309 ± 9	$4.9 \times 10^{10} \pm 9.2 \times 10^{8}$	$1.3 \times 10^{9} \pm 7.0 \times 10^{6}$	36.5 ± 0.9	5617 ± 58
October HT T0	10/14/2014	28	198 ± 6	$5.0 \times 10^{10} \pm 2.5 \times 10^{10}$	$1.7 \times 10^{9} \pm 4.0 \times 10^{7}$	28.9 ± 13.9	5630 ± 39
October HT T24	10/15/2014		193 ± 3	$1.1 \times 10^{11} \pm 4.6 \times 10^{10}$	$1.4 \times 10^{9} \pm 3.1 \times 10^{7}$	81.3 ± 34.4	5648 ± 40

LT, low tide; HT, high tide. n = 3 for chemical data, n = 2 for biological data except for July HT for which n = 1. Standard deviations indicate within-treatment variability.

were defined as molecular formulae that decreased significantly in relative abundance during the 24 h incubations (Fig. 1).

Results and discussion

Our sampling site on the southeastern U.S. coast represents a marine ecosystem with ecological influences from both terrestrial and oceanic environments. DOM sources include inputs from river flow, salt marshes, phytoplankton and heterotrophic microbes, with the relative importance of these sources strongly modulated by river discharge (Medeiros *et al*., 2015a). The terrigenous character of the DOM is greater during spring compared with late summer/fall, and greater at low tide compared with high tide (Medeiros *et al*., 2015a, 2017).

At two sampling dates in July and October, 2014, gene expression and chemical patterns in natural bacterial communities were analysed at both high and low tide to identify short term changes indicative of highly labile DOM components. Whole water samples collected during daylight hours were processed immediately after collection (T0) and again after dark incubation *in situ* for a 24 h period (T24) for analysis by both metatranscriptomic (in duplicate) and FT-ICR MS (in triplicate) methodologies.

The abundance of bacterial and archaeal transcripts in the initial seawater samples (0.2–3.0 μm size fraction) varied from 4.6×10^{10} to 2.5×10^{11} L^{-1}. Transcript inventories averaged ~ 50 transcripts $cell^{-1}$ (Table 1), which is ~ 30-fold lower than reported for laboratory-grown *Escherichia coli* [1380 mRNAs per $cell^{-1}$ in Neidhardt and Umbarger (1996) and 1780 $cell^{-1}$ in Taniguchi *et al*. (2010)]. Our low transcript inventories are comparable to previous estimates for free-living bacterioplankton (~ 100 transcripts $cell^{-1}$; Satinsky *et al*., 2017) and likely attributable to smaller genomes and slower growth rates of environmental bacteria compared with exponentially growing laboratory cultures. The FT-ICR MS analysis identified an average of 5500 molecular formulae in the initial DOM pools (Table 1). This level of chemical diversity is consistent with previous analyses at the site (Medeiros *et al*., 2017). Differences in bulk dissolved organic carbon (DOC) concentration after the 24 h incubation were within the deviation of the replicate samples for three out of four incubations (Table 1). Significant decreases in DOC concentration have been observed with longer incubations of samples from this site (Medeiros *et al*., 2017).

Metatranscriptome taxonomic composition

Based on homology searches against marine microbial genomic data, more than 600 taxa were represented in the metatranscriptomes. The 50 highest-recruiting taxonomic bins in each sample accounted for 54%–66% of the total transcripts (Supporting Information Table S1). Transcripts binned most often to genomes of the Gammaproteobacteria OM60/NOR5 clade, with the genomes of strains HTCC2148 and HTCC2080 together recruiting up to 11% of the reads in some datasets (Fig. 2). Alphaproteobacteria representatives of the SAR11 clade (Alphaproteobacterium HIMB5, *Candidatus* Pelagibacter sp. HTCC7211, Alphaproteobacterium HIMB114) comprised 2%–10% of the metatranscriptomes and representatives of the Roseobacter group comprised up to 8% of the metatranscriptomes. Genomes from the Verrucomicrobia and Thaumarchaeota (*Nitrosopumilus* spp.) also recruited a substantial proportion of transcripts, while Actinobacteria, Planctomycetes, Bacteroidetes and uncultured Marine Group II Euryarchaeota contributed appreciably but less consistently (Fig. 2). Together, the genome bins depicted transcriptionally active bacterial and archaeal communities that were taxonomically diverse and largely consistent with previous studies at this location (Gifford *et al*., 2014).

Changes in gene expression after 24 h

The goal of the metatranscriptome analysis was to document changes in gene expression over a time frame in which the most labile components of DOM would be depleted. To do this, potential protein encoding reads

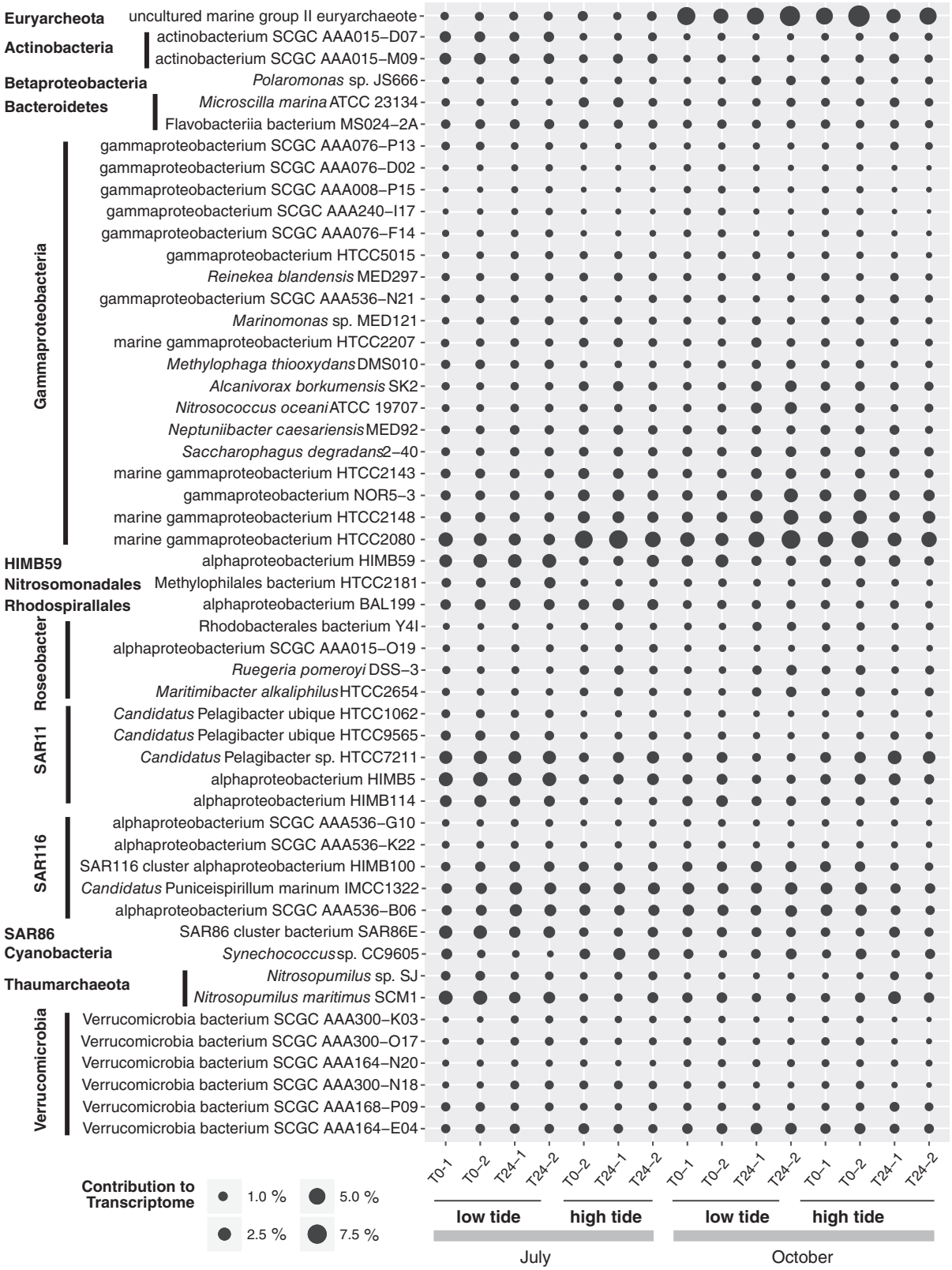

Fig. 2. Percent contribution to the total transcriptome for the 50 highest transcript-recruiting reference genomes. Sample notation: T0, initial sample; T24, incubated sample; dashed line followed by a number, replicate designation.

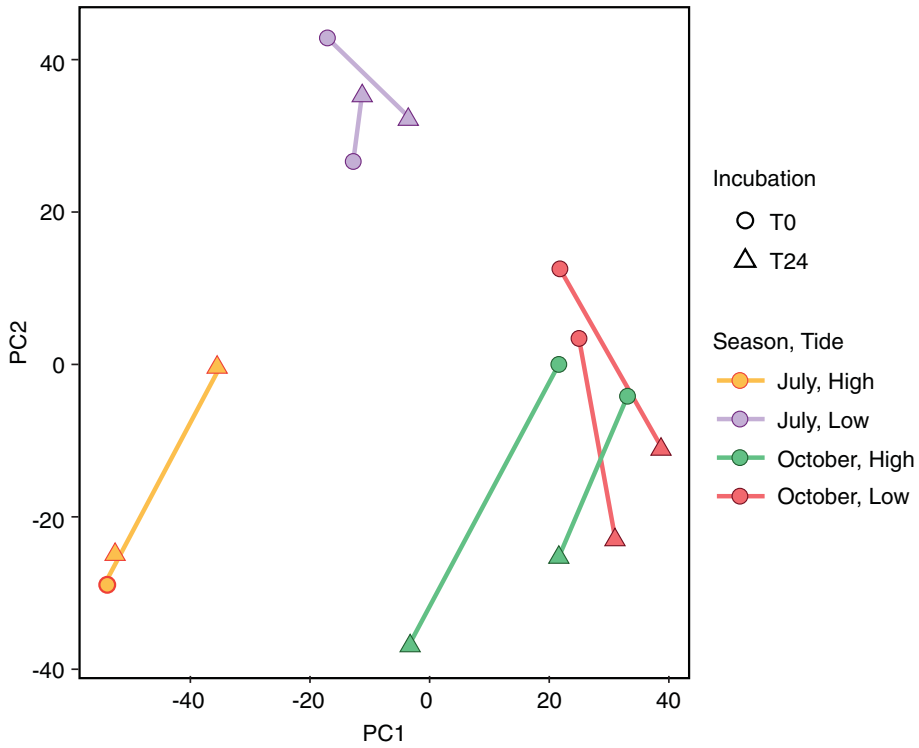

Fig. 3. Principal component analysis of gene expression data. Scores of the first two principal components are colour coded by each season and tide combination. Circles and triangles represent T0 and T24 respectively. The first and second principal components explained 20.6% and 14.2% of the variability.

were annotated by a homology search using RAPSearch2 against MarineRef II, an in-house database (http://ssharma.marsci.uga.edu/Lab/MarineRef2/) consisting of marine bacterial, archaeal, viral and eukaryotic genome sequences, and eukaryotic cDNAs from the MMETSP project (Keeling *et al*., 2014) annotated against SFams, Pfam, NCBI Protein Clusters, PhyloDB and TIGRFAMS databases (see Experimental Procedures). All of these databases have uncertainties associated with annotations, and our assignment of gene function reflects current best judgement. One replicate T0 sample from the July high tide collection had contaminating sequences in the final library and was eliminated from subsequent analyses.

At the time of initial sampling (T0), we assumed that heterotrophic microbes were processing DOM components that broadly spanned the lability spectrum and that the fluxes of organic compounds between the DOM and microbial pools were close to steady state. Following the 24 h dark incubation (T24) during which samples were not resupplied with organic matter from recent photosynthate or advected terrigenous material, we expected that the more bioreactive compounds fueling growth from these DOM pools would be depleted. One prediction that follows is that the growth rate of bacteria should decrease between T0 and T24 as some labile components of the DOM are depleted and not replaced. To test this, we compared transcription levels of genes encoding ribosomal RNA proteins, as expression of these genes has been shown to track with overall growth rates in both bacterial isolates and natural communities (Wei *et al*., 2001; Gifford *et al*., 2013). Out of the 120 ribosomal protein clusters identified in a pooled dataset from the four experiments, the only significant changes between T0 and T24 were negative (Wald tests with Benjamini and Hochberg adjusted-values, $p < 0.10$) (Supporting Information Fig. S1), supporting our expectation of substrate depletion that was manifested as slower growth rates. We also checked whether a 'bottle effect' shift in active taxa occurred by determining whether the relative contributions of the 50 highest recruiting taxonomic groups changed significantly and in a consistent direction. Changes in transcript relative abundance between T0 and T24 were typically small (1.2-fold average change for taxa that decreased, 1.4-fold average change for those that increased). There were significant changes between T0 and T24 for 28% of cases (DESeq2, $p < 0.05$) but the magnitude and direction of changes within a taxon varied across the different experiments (Supporting Information Fig. S2). The short time frame, the *in situ* incubation, and the volume of water incubated (20 L) likely reduced bottle effects.

Changes in gene expression emerging during the 24 h incubations were nested within more dominant differences linked to season and to some extent to tidal stage, a consequence of dissimilarities in the starting communities and initial DOM composition for the experiments (Fig. 3). Therefore, our functional analyses compared expression independently for each

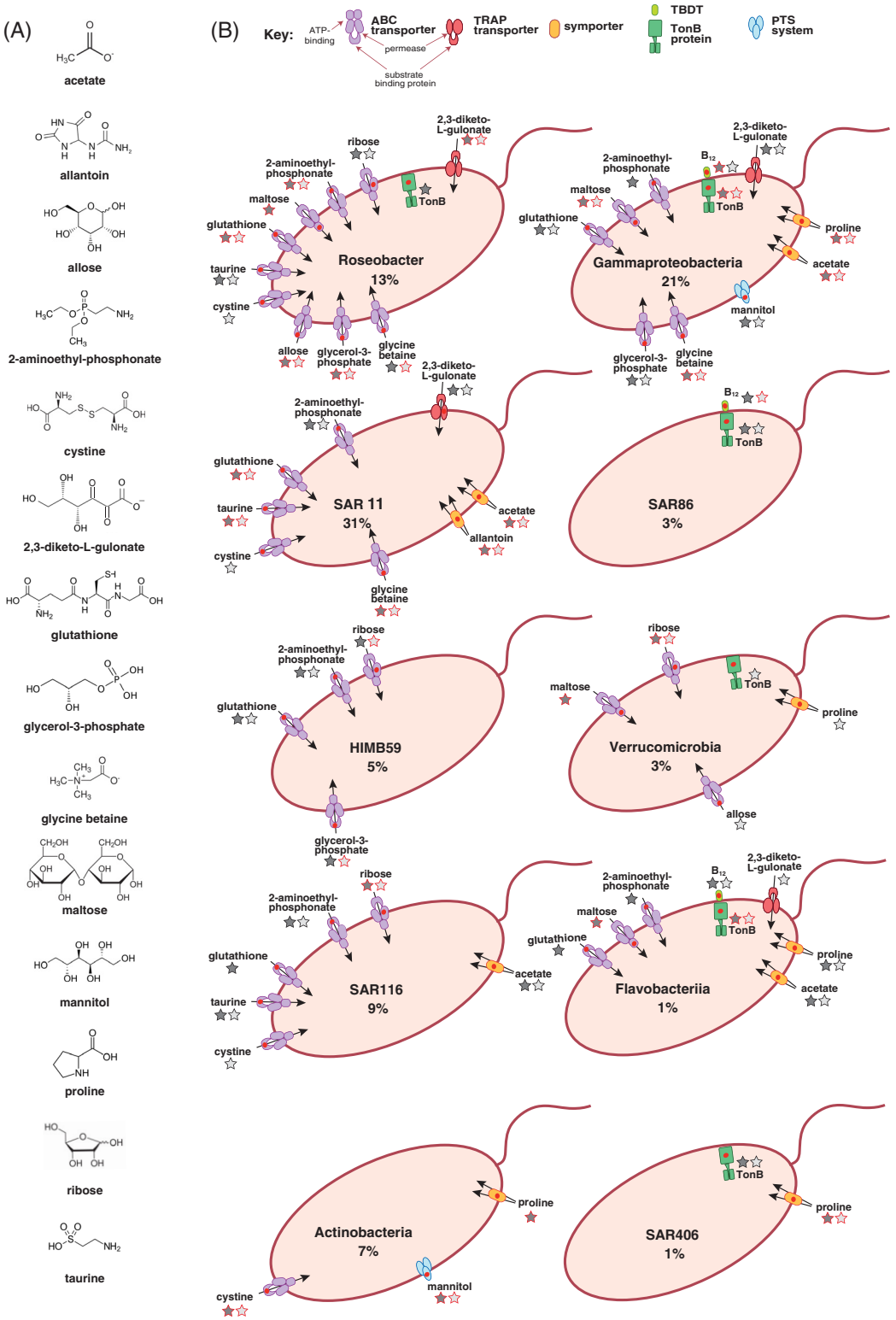

Fig. 4. Bacterial transport functions that decreased in relative expression following a 24 h incubation in the absence of fresh DOM inputs. A. Compounds annotated as the target substrates for the 14 significantly depleted transporter genes.B. Eight major taxonomic groups with the highest contributions to transcripts from significantly depleted transporters; percent contribution is indicated in each cell. SAR86 is a member of the Gammaproteobacteria but is shown separately and not included in the Gammaproteobacteria expression profile. A red dot on multi-protein transporters indicates which component was significantly depleted. Dark grey and light grey stars indicate transporters depleted in the July and October experiments, respectively; red outlines on stars mark taxa that contributed ≥ 20% of the transcripts for that transporter. Data for individual reference genomes making up each of major the taxonomic groups are given in Supporting Information Table S4. TBDT = TonB dependent transporters.

replicated T0/T24 pair to isolate incubation effects from other factors. We focused on gene expression patterns of transporters and metabolic proteins whose annotation suggested a role in processing organic compounds.

Transporter gene expression

During the July incubations, significantly depleted transcript inventories (DESeq2; adjusted $p < 0.10$) were found for transporters annotated for the uptake of osmolytes (glycine betaine, taurine, proline), sugars (allose and mannitol), a carboxylic acid (acetate), a component of phospholipids (glycerol-3-phosphate), an antioxidant (glutathione), an organic sulfur compound (cystine), a purine degradation product (allantoin) and an organic phosphorus compound (2-aminoethylphosphonate) (Fig. 4). During the October incubations, the transporters with significantly decreased expression were annotated for sugar and sugar acid uptake (ribose, maltose, allose and 2,3-diketo-L-gulonate) and acetate uptake (Table 2; Fig. 4).

Most of the transporter genes decreasing in expression during the incubations were members of the ATP-binding cassette (ABC) transporter class that use ATP hydrolysis to translocate solutes across cell membranes. The gene encoding the periplasmic substrate-binding component of these ABC transporters was most often the component that decreased significantly (6 of the 9 ABC transporter genes), consistent with previous observations in environmental samples (Poretsky *et al*., 2010) and suggesting that the transcription of substrate binding proteins may be a well-regulated step in DOM uptake. A number of permeases/symporters that rely on concentration gradients, a TonB transport protein, a phosphotransferase system (PTS) component, and a tripartite ATP-independent periplasmic (TRAP) transporter component rounded out the classes of significantly depleted transporters.

Taxonomic recruitment of transcripts to reference genomes suggested variability among the dominant bacterioplankton groups with regard to the compounds being assimilated. Reference genomes in the Gammaproteobacteria and Roseobacter groups recruited transcripts from the highest diversity of transporters, with transcripts mapping to 11 significantly depleted transporters in each group (Fig. 4b). The other taxa may utilize fewer substrates or regulate expression of their transporter genes to a lesser extent. Substrates that appear to be most broadly used by the bacterioplankton community were glutathione (transcripts recruited to 15 different major taxa), acetate (14 taxa) and 2-aminoethylphosphonate (10 taxa) (Supporting Information Table S2). On the other hand, signature substrates (defined here as ≥ 50% of transcripts mapping to a single taxonomic group in both seasons) included allose, 2,3-diketo-L-gulonate and glycerol-3-phosphate (signature substrates for Roseobacter), allantoin and taurine (signature substrates for SAR11), and cystine and mannitol (signature substrates for Actinobacteria).

A signal potentially related to vitamin B_{12} availability over the 24 h incubation emerged in the July experiment, supported by the depletion of transcripts for both a vitamin B_{12}/cobalamin outer membrane transporter and a TonB inner membrane protein; these proteins are part of the vitamin B_{12} acquisition system in many bacteria (Schauer *et al*., 2008) (Table 2; Fig. 4b). Although TonB proteins are involved in the uptake of compounds other than B_{12} (e.g., metal complexes, carbohydrates; Noinaj *et al*., 2010), the strong one-to-one correspondence between genomes recruiting transcripts for the B_{12}/cobalamin transporter gene and those recruiting transcripts for the TonB inner membrane protein, at least in the case of Gammaproteobacteria and Flavobacteriia, suggested that these two transporter components may indeed represent parts of a B_{12} acquisition system. The Verrucomicrobia and SAR406 reference genomes also recruited transcripts for TonB with depleted expression during the incubation but had no corresponding hits to a B_{12}/cobalamin outer membrane transporter (Supporting Information Table S2).

Metabolism gene expression

Expression changes in metabolism genes are less definitive transcriptional indicators of labile DOM than expression changes in transporters because of the difficulty in distinguishing exogenously acquired substrates from existing cellular metabolites. Nonetheless, diminished relative expression among metabolism genes during the July incubations suggested decreases in fatty acids and acetate availability (acyl-CoA dehydrogenase, isocitrate lyase, fatty acid oxidation complex fadJ, acetyl-CoA acetyltransferase), sugar and sugar alcohol degradation (phosphopentomutase, glycosyltransferase, mannose-1-phosphate guanyltransferase), peptide/amino acid metabolism (protease 2, dihydroorotase) and pyrimidine catabolism (phenylhydantoinase) (Table 2; Fig. 5). In two cases, a depleted metabolism gene was part of a metabolic pathway that began with the substrate of a depleted transporter gene (mannitol and glycine betaine), providing potential evidence of a biochemical link from decreased transporter expression to decreased downstream metabolism (Fig. 5).

During the October incubations, multiple genes involved in the metabolism of cysteine and methionine decreased in relative abundance. These included *metY* (O-acetylhomoserine sulfhydrylase) and *metZ* (O-succinylhomoserine sulfhydrylase), which are involved in H_2S transfer to form homocysteine; and *metE* (5-methyltetrahydropteroyltriglutamate-homocysteine

Table 2. Transporter and metabolism genes for which transcripts were significantly depleted in the metatranscriptome after a 24 h dark incubation.

Sample	Protclust accession	Protclust description	Log_2 fold change
July HT	PRK15419	Proline:sodium symporter, PutP	–1.65
	PRK11375	Allantoin permease	–1.53
	PRK09463	Acyl-CoA dehydrogenase	–1.70
	PRK06498	Isocitrate lyase	–1.26
July LT	PRK10974	Glycerol-3-phosphate ABC transporter substrate-binding protein	–1.08
	PRK11119	Glycine betaine ABC transporter substrate-binding protein	–1.04
	PRK11480	Taurine ABC transporter substrate-binding protein	–1.15
	PRK15046	2-Aminoethylphosphonate ABC transporter substrate-binding protein	–1.13
	PRK15413	Glutathione ABC transporter substrate-binding protein, GsiB	–1.14
	PRK11260	Cystine ABC transporter substrate-binding protein	–0.89
	PRK10819	Tonb system transport protein	–0.90
	PRK10641	Vitamin B12/cobalamin outer membrane transporter	–1.14
	PRK09701	D-allose ABC transporter permease	–1.26
	PRK15083	Pts mannitol transporter subunit IICBA	–1.66
	PRK11375	Allantoin permease	–0.86
	PRK12488	Acetate permease	–0.82
	PRK11154	Fatty acid oxidation complex subunit alpha FadJ	–0.72
	PRK09463	Acyl-CoA dehydrogenase	–2.29
	PRK06498	Isocitrate lyase	–2.82
	PRK06025	Acetyl-CoA acetyltransferase	–1.79
	PRK05362	Phosphopentomutase	–2.18
	PRK10073	Glycosyltransferase	–1.00
	PRK15460	Mannose-1-phosphate guanyltransferase	–0.94
	PRK10115	PROTEASE 2	–0.83
	PRK05451	Dihydroorotase	–1.44
	PRK09357	Dihydroorotase	–0.92
	PRK13805	Acetaldehyde-CoA/alcohol dehydrogenase	–0.83
	PRK08323	Phenylhydantoinase	–1.30
	PRK01202	Glycine cleavage system protein H	–0.90
October HT	PRK09512	Ribose ABC transporter permease protein	–0.74
	PRK11000	Maltose/maltodextrin transporter ATP-binding protein	–1.72
	PRK09877	2,3-Diketo-L-gulonate TRAP transporter small permease YiaM	–1.81
	PRK06084	*metY*, O-acetylhomoserine sulfhydrylase	–1.21
	PRK08248	*metY*, O-acetylhomoserine sulfhydrylase	–2.64
	PRK07812	*metY,* O-acetylhomoserine sulfhydrylase	–1.57
	PRK07504	*metZ*, O-succinylhomoserine sulfhydrylase	–1.84
	PRK08133	*metZ*, O-succinylhomoserine sulfhydrylase	–1.74
	PRK05222	*metE*, 5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	–3.31
	PRK08645	*metH*, homocysteine S-methyltransferase/methylenetetrahydrofolate reductase	–2.01
	PRK15033	Tricarballylate utilization protein TcuB	–2.10
	PRK07067	Sorbitol dehydrogenase	–2.75
	PRK15448	Ethanolamine catabolic microcompartment shell protein EutN	–2.72
	PRK10524	Propionate CoA ligase	–1.98
	PRK11274	Glycolate oxidase iron–sulfur subunit	–1.52
	PRK11892	Pyruvate dehydrogenase complex E1 component subunit beta	–1.19
	PRK09279	Pyruvate, phosphate dikinase	–1.17
	PRK15425	Glyceraldehyde-3-phosphate dehydrogenase A	–1.01
	PRK00451	Glycine dehydrogenase subunit 2	–1.65
	PRK04366	Glycine dehydrogenase subunit 1	–1.76
	PRK07028	Hexulose-6-phosphate synthase/ribonuclease regulator	–1.73
	PRK09078	Succinate dehydrogenase flavoprotein subunit	–0.98
	PRK07573	Succinate dehydrogenase flavoprotein subunit	–0.82
	PRK08937	Adenylosuccinate lyase	–2.15
	PRK03739	2-Isopropylmalate synthase	–0.93
	PRK05479	Ketol-acid reductoisomerase	–0.92
	PRK09192	Acyl-CoA synthetase	–1.80
	PRK07868	Acyl-CoA synthetase	–1.53
	PRK12824	Acetoacetyl-CoA reductase	–1.33
	PRK14360	N-acetylglucosamine-1-phosphate uridyltransferase/ glucosamine-1-phosphate acetyltransferase	–1.78
	PRK08322	Acetolactate synthase large subunit	–1.11
	PRK05862	Enoyl-CoA hydratase	–1.08

Table 2. *cont.*

Sample	Protclust accession	Protclust description	Log_2 fold change
	PRK09284	Phosphomethylpyrimidine synthase ThiC	–0.91
	PRK05820	Thymidine phosphorylase	–1.06
	PRK12810	Glutamate synthase subunit beta	–0.79
	PRK13581	Phosphoglycerate dehydrogenase	–0.59
	PRK09105	Putative aminotransferase	–1.39
	PRK14874	Aspartate-semialdehyde dehydrogenase	–0.79
	PRK05398	Formyl-CoA transferase	–0.85
	PRK07474	Sulfur oxidation protein SoxY	–1.60
	PRK06854	Adenylylsulfate reductase	–1.98
	PRK14990	Anaerobic dimethyl sulfoxide reductase subunit A	–2.26
October LT	PRK09395	Acetate permease	–1.05
	PRK09512	Ribose ABC transporter permease protein	–0.83
	PRK09701	D-allose ABC transporter permease	–0.89
	PRK08248	O-acetylhomoserine aminocarboxypropyltransferase	–3.76
	PRK07812	O-acetylhomoserine/o-acetylserine sulfhydrylase	–2.60
	PRK07810	O-succinylhomoserine sulfhydrylase	–2.15
	PRK07504	O-succinylhomoserine sulfhydrylase	–2.32
	PRK00175	Homoserine O-acetyltransferase	–1.27
	PRK05222	5-Methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase	–4.00
	PRK07534	Methionine synthase I	–1.75
	PRK08645	Homocysteine S-methyltransferase/methylenetetrahydrofolate reductase	–2.54
	PRK07067	Sorbitol dehydrogenase	–3.29
	PRK10524	Propionate CoA ligase	–1.36
	PRK11892	Pyruvate dehydrogenase complex E1 component subunit beta	–0.91
	PRK05479	Ketol-acid reductoisomerase	–0.73
	PRK07573	Succinate dehydrogenase flavoprotein subunit	–1.08
	PRK09406	NADP-dependent succinic semialdehyde dehydrogenase	–1.96
	PRK09064	5-Aminolevulinate synthase	–1.13
	PRK11259	N-Methyltryptophan oxidase	–0.72
	PRK06854	Adenylylsulfate reductase	–1.44

Dark grey, transporter; light grey, metabolism.

S-methyltransferase) and *metH* (homocysteine S-methyltransferase/5,10-methylenetetrahydrofolate reductase), which are involved in methyl group transfer to form methionine (Supporting Information Fig. S3). Transcripts were also depleted for genes annotated for catabolism of tricarballylate, sorbitol, ethanolamine, propionate, glycolate, pyruvate, glyceraldehyde-3-phosphate and glycine (Fig. 5).

Changes in DOM composition after 24 h

FT-ICR MS analysis showed that SPE-captured DOM components decreasing significantly in relative abundance during the 24 h dark incubation had average molecular mass differences of 30–40 Da compared with those increasing significantly. This small average mass difference suggested that microbial transformations of the SPE-captured components of DOM were dominated by removal of functional groups to form slightly smaller molecules; for example, demethylation would produce a compound smaller by a mass of 14.01565 Da through replacement of a CH_3 by an H. We made pairwise comparisons of all molecular formulae whose relative abundance significantly decreased to those that significantly increased (as in Medeiros *et al*., 2017). Although it is not possible to link a specific formula that decreased with a specific formula that increased, this exercise generated a list of the most common potential transformations in elemental composition, most of which were found in both the July and October experiments (Table 3). Among these were changes consistent with demethylations (net loss of CH_2), dehydrogenations (net loss of H_2), demethoxylations (net loss of CH_2O), dehydrations (net loss of H_2O), deacetylations (net loss of $C_2H_2O_2$), decarboxylations (net loss of CO_2) and potential combinations of these (Table 3).

Formulae depleted after 24 h were also biased toward S-containing compounds (average of 1.6-fold bias) and N-containing compounds (2.0-fold bias), and biased against CHO-only compounds (3.5-fold bias) compared with the element distribution of molecular formulae for the T0 samples (Fig. 6). A substantial fraction of depleted compounds fell into the category of highly unsaturated and unsaturated aliphatic compounds with low O/C ratios (Šantl-Temkiv *et al*., 2013), which could potentially include unsaturated fatty acids (Table 4). Using the expected formulae for marine- versus terrestrially-derived organic matter determined previously for this ecosystem (Medeiros *et al*., 2015a), ~ 30% of preferentially consumed molecules were marine, averaged across seasons and tidal stage, and

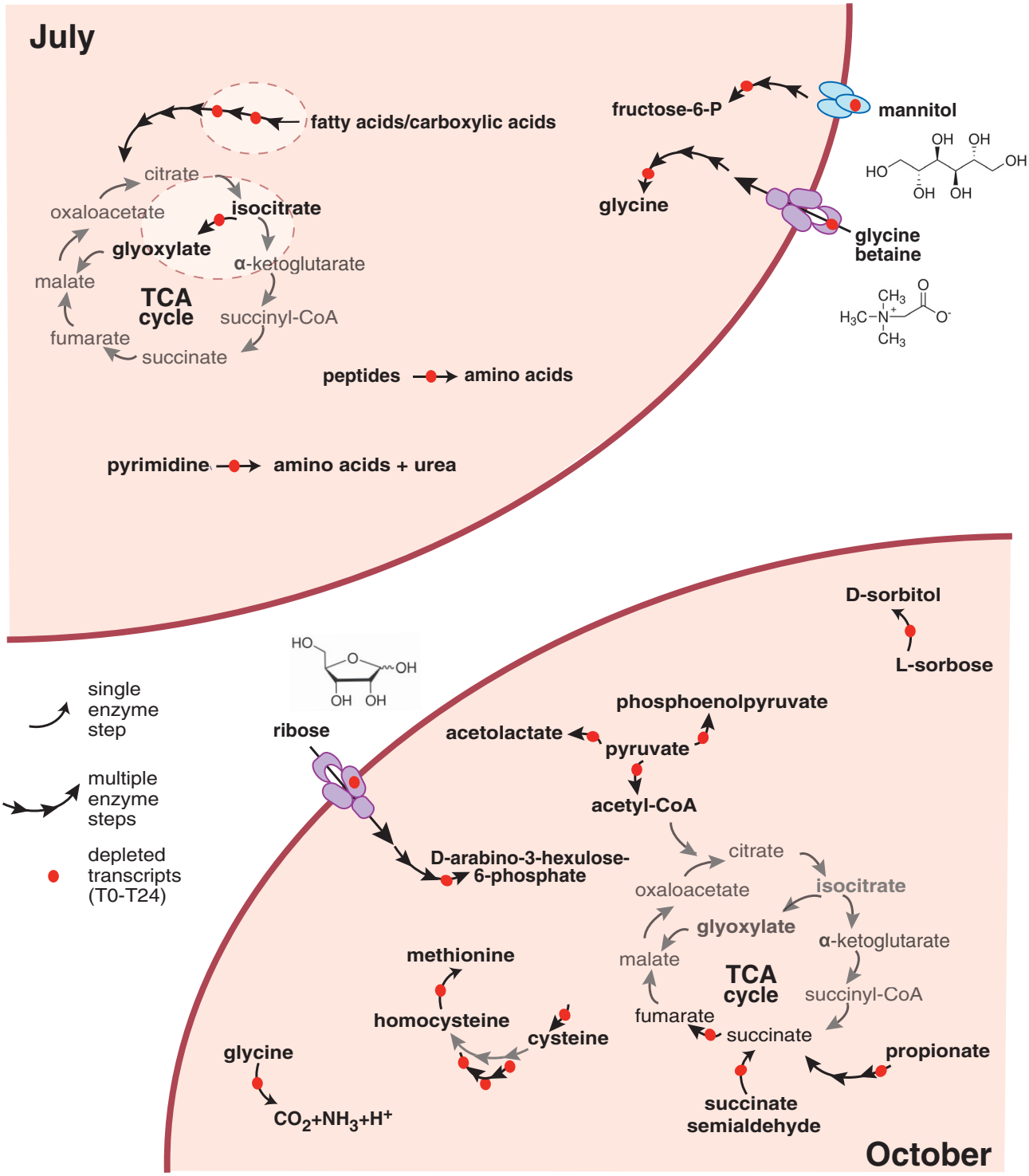

Fig. 5. Bacterial metabolism functions that decreased in relative expression following a 24 h isolation from new sources of DOM. See Fig. 4 for transporter symbol key.

~ 3%–10% were terrigenous; the remainder were ambiguous with regard to origin. Changes in bulk DOC concentrations were within the error of the analytical method for most of the incubations (Table 1).

Integration of biological and chemical data

Metatranscriptomics and FT-ICR MS analysis are complementary approaches for untargeted investigations of labile DOM components. Metatranscriptome sequencing

Table 3. Most frequent potential transformations in elemental composition of DOM after 24 h incubations, obtained by calculating mass differences between all enriched compounds and all depleted compounds.

Average relative frequency	Possible atoms removed	Possible mass lost (Da)	Possible enzymatic activity
July			
0.94	CH_2	−14.0157	1 demethylation
0.93	C_2H_2	−26.0157	
0.92	C_2H_4	−28.0313	2 demethylations
0.85	C	−12.0000	
0.83	H_2	−2.0157	1 dehydrogenation
0.80	CH_2O	−30.0106	1 demethoxylation
0.76	CO	−27.9949	
0.76	C_2H_2O	−42.0106	
0.74	C_3H_4	−40.0313	
0.73	CH_4	−16.0313	1 demethylation and 1 dehydrogenation
0.72	C_2	−24.0000	
0.69	C_2H_4O	−44.0262	1 demethylation and 1 demethoxylation
0.67	C_3H_4O	−56.0262	
0.64	C_2O	−39.9949	
0.60	C_3H_6*	−42.0470	3 demethylations
0.58	C_4H_6O*	−70.0419	
0.57	H_2O	−18.0106	1 dehydration
0.54	C_3H_2	−38.0157	
0.53	C_2H_6O	−46.0419	
0.53	H_4	−4.0313	2 dehydrogenations
0.53	C_4H_6*	−54.0470	
0.50	C_3H_6O*	−58.0419	
0.50	O	−15.9949	
0.50	CH_4O	−32.0262	1 demethoxylation and 1 dehydrogenation
0.48	CH_6	−18.0470	1 demethylation and 2 dehydrogenation
0.47	C_4H_4*	−52.0313	
0.43	C_3H_8*	−44.0626	
0.42	C_2H_6	−30.0469	2 demethylation and 1 dehydrogenation
0.41	H_4O	−20.0262	
0.38	C_3O*	−51.9949	
October			
0.79	CH_2	−14.0157	1 demethylation
0.75	H_2	−2.0157	1 dehydrogenation
0.70	CH_2O	−30.0106	1 demethoxylation
0.68	C_2H_4	−28.0313	2 demethylations
0.68	CH_4	−16.0313	1 demethylation and 1 dehydrogenation
0.65	CO	−27.9949	
0.62	C	−12.0000	
0.61	O	−15.9949	
0.61	C_2H_2O	−42.0106	
0.57	C_2H_2	−26.0157	
0.53	CH_6	−18.0469	1 demethylation and 2 dehydrogenations
0.52	H_4	−4.0313	2 dehydrogenations
0.52	C_2O	−39.9949	
0.51	C_2	−24.0000	
0.51	H_2O	−18.0106	1 dehydration
0.49	C_2H_4O	−44.0262	1 demethylation and 1 demethoxylation
0.47	C_3H_4	−40.0313	
0.46	CH_4O	−32.0262	1 demethoxylation and 1 dehydrogenation
0.44	C_2H_6	−30.0469	2 demethylation and 1 dehydrogenation
0.42	C_2H_8*	−32.0626	2 demethylation and 2 dehydrogenation
0.40	C_3	−36.0000	
0.40	C_3H_2*	−38.0157	
0.39	O_2	−31.9898	
0.39	C_3O	−51.9949	
0.36	H_6*	−6.0470	3 dehydrogenations
0.35	C_3H_4O	−56.0262	
0.35	H_4O	−20.0262	
0.34	CH_2O_2*	−46.0055	1 decarboxylation and 1 dehydrogenation
0.32	$C_2H_2O_2$*	−58.0055	1 deacetylation
0.32	CO_2*	−43.9898	1 decarboxylation

Samples within a season were combined for the analysis, and the 30 most frequent potential transformations are given for each season. Average relative frequency represents the percent of times a particular transformation was calculated. Transformations not common to both July and October are marked with an asterisk.

provides a window into the low molecular weight, typically polar molecules that are substrates for bacterial transporters, while FT-ICR MS methodology can survey the larger, semi-polar molecules of marine DOM. Previous studies employing metatranscriptomics to identify bioreactive DOM have typically amended seawater with known sources of organic matter and tracked transcript enrichment in response to the addition (McCarren *et al*., 2010; Poretsky *et al*., 2010). Here, we focused instead on transcript depletion in unamended marine DOM. Three caveats relevant to interpreting these gene expression data include: gene annotations assigned in this study are based largely on homology to known genes, and thus are more robust for genes in central metabolic pathways and more tentative for those associated with environmental substrates and microbial interactions; some marine bacteria express transporters constitutively (Cottrell and Kirchman, 2016), in which case a change in substrate availability would not invoke a change in transcription; and expression can be regulated post-transcriptionally, modified instead at the protein translation or enzyme activity stage and therefore leaving no signal in the transcriptome (Moran *et al*., 2013).

Caveats associated with the FT-ICR MS-based technique include: molecular structure information is unavailable because a single formula can represent multiple structures (Longnecker *et al*., 2015); the normalization method chosen during data analysis will influence the inferred changes in FT spectra; and differences in composition of sample extracts can influence the ionization process. A recent study of 90 common intracellular metabolites dissolved in a seawater matrix found that 58 had < 5% retention on an SPE resin (Johnson *et al*., 2017), including all five that

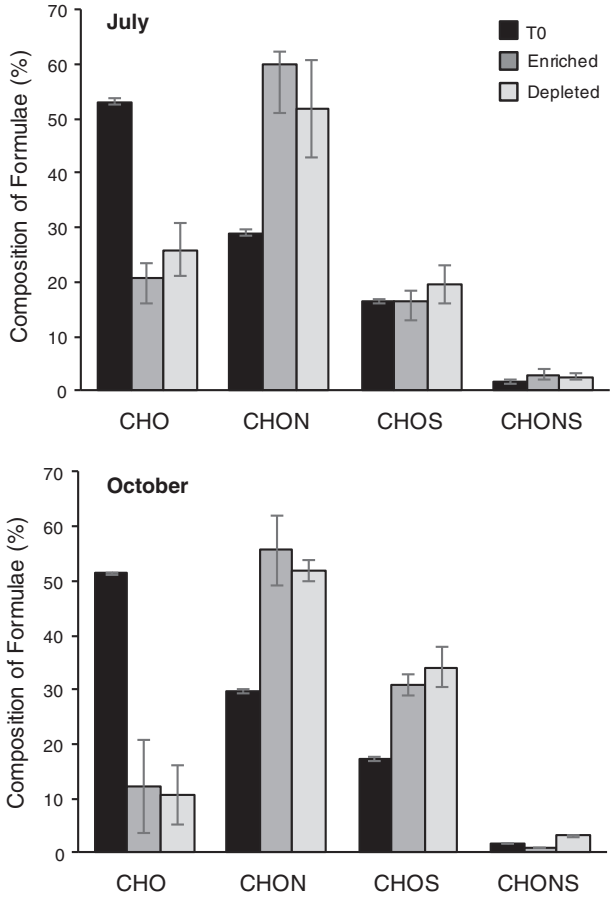

Fig. 6. Elemental composition of compounds in the initial DOM (T0) and those changed in relative abundance (depleted or enriched) after a 24 h incubation in July (top) and October (bottom) experiments. Values shown are the fraction of the molecular formulae in each category (T0, enriched or depleted) that contained only CHO, CHON, CHOS or CHONS. The sum of each category equals 100%. n = 3.

overlap with predicted substrates for depleted transporters in our study (glycine betaine, glutathione, proline, taurine and cyanocobalamin). Thus, the two methodological approaches used here primarily characterize different subsets of the marine DOM pool.

Both the metatranscriptome and FT-ICR MS analyses point to carboxylic acids as important components of labile DOM, consistent with findings for the open ocean (Bergauer *et al*., 2018). Transcripts from acetate transporters accounted for 30% of all significantly depleted transporter transcripts and mapped to a diversity of genomes (Supporting Information Table S2). Congruent with this finding, the loss of $C_2H_2O_2$, which can correspond to the removal of an acetyl moiety, was among the most frequent potential transformations in elemental composition identified by FT-ICR MS analysis (Table 3). Acetate has been identified in phytoplankton exometabolomes (Hellebust, 1974) and is also known to have a photochemical source in marine surface waters (Mopper *et al*., 1991; Wetzel *et al*., 1995). This low molecular weight carboxylic acid might also be derived from degradation of lipids; fatty acids, sterols and sulfo- and glycolipids can comprise up to 10% of extracellular release from phytoplankton (Hellebust, 1965; Billmire and Aaronson, 1976). Elemental ratios of the depleted molecular formulae are also consistent with fatty acids being well represented in the labile DOM pool (Table 4).

The sugars and sugar acids predicted to be components of labile DOM based on transcript data (ribose, maltose, 2,3-diketo-L-gulonate, sorbose; Figs. 2 and 3) are consistent with earlier studies suggesting that monosaccharides can account for up to half of bacterial secondary production in marine systems (Rich *et al*., 1996; Kirchman *et al*., 2001). Phytoplankton cells are ~ 25% carbohydrate by mass (Darley, 1977), and phytoplankton exometabolites can include both polysaccharides and monosaccharides (Ittekkot *et al*., 1981).

FT-ICR MS data showed that 22% and 37% of the SPE-captured compounds depleted in July and October, respectively, contained a sulfur atom, compared with 18% and 19% representation in the initial DOM (Fig. 6). Further, 5 of the 6 formulae exhibiting the greatest percent depletion during the incubation (from 54% to 24%

Table 4. Comparison of percent of formulae assigned to DOM molecular compound groups (based on Šantl-Temkiv *et al*., 2013) among those depleted, enriched or unaltered following the 24 h dark incubation.

	Polycyclic aromatics	Highly aromatic w/aliphatic side chains	Highly unsaturated	Unsaturated aliphatic with high O/C	Unsaturated aliphatic with low O/C	Unsaturated aliphatic containing N
July high tide						
Depleted	1.1	14.8	69.2	5.5	8.2	1.1
Enriched	11.7	8.9	74.2	0.5	2.8	1.9
Unaltered	7.0	18.8	63.8	6.7	2.1	1.6
July low tide						
Depleted	0.6	10.2	61.7	7.8	18.0	1.8
Enriched	10.6	13.8	69.3	1.1	4.2	1.1
Unaltered	6.9	18.8	63.8	6.8	2.0	1.6
October high tide						
Depleted	0.6	4.9	85.3	1.8	6.7	0.6
Enriched	2.3	15.3	75.6	1.1	3.4	2.3
Unaltered	7.7	19.3	62.7	6.5	2.2	1.7
October low tide						
Depleted	0.7	11.2	76.3	2.0	7.2	2.6
Enriched	1.7	19.3	64.8	4.5	5.1	4.5
Unaltered	7.8	18.5	63.9	6.3	2.0	1.5
Average						
Depleted	0.7 ± 0.2	10.3 ± 4.1	73.1 ± 10.1	4.3 ± 2.9	10.0 ± 5.3	1.5 ± 0.9
Enriched	6.6 ± 5.3	14.3 ± 4.3	71.0 ± 4.9	1.8 ± 1.9	3.9 ± 1.0	2.4 ± 1.5
Unaltered	7.3 ± 0.5	18.9 ± 0.3	63.3 ± 0.6	6.6 ± 0.2	2.1 ± 0.1	1.6 + 0.1

Each row sums to 100%.

relative loss) were sulfur compounds (Supporting Information Table S3). The sulfur signal in the FT-ICR MS data could derive from sulfolipids, since these are abundant components of photosynthetic membranes in marine phytoplankton (Van Mooy *et al*., 2009) and both the sulfur-containing (sulfoquinovose) and fatty acid moieties of sulfolipids are readily broken down by bacteria (Fujita *et al*., 2007; Denger *et al*., 2014; Durham *et al*., 2015). A recent analysis of SPE-captured DOM concluded that oceanic dissolved organic sulfur compounds are cycled more rapidly than the bulk carbon pool (Ksionzek *et al*., 2016), implying that they are more biologically labile on average. Transcription data agreed that organic sulfur compounds are important in DOM turnover (Figs. 4 and 5; Supporting Information Fig. S4). Similarly for nitrogen heteroatoms, FT-ICR MS data showed that 52% of depleted high molecular weight compounds contained nitrogen compared with only 29% in the initial samples (Fig. 6). Transcript data indicated labile substrates containing one (2-aminoethyl-phosphonate, glycine betaine, proline, taurine), two (cystine), three (glutathione) or four (allantoin) nitrogen atoms (Fig. 4a). Two of the predicted transporter substrates could also serve as a source of organic phosphorus (2-aminoethyl phosphonate and glycerol-3-phosphate; Fig. 4b).

Approximately 40% of the carbon fixed by marine phytoplankton is processed through heterotrophic bacteria within hours to days of fixation (Cole *et al*., 1988; Ducklow and Carlson, 1992). For coastal bacteria, the terrestrial landscape contributes additional organic carbon compounds for heterotrophic processing (Cole *et al*., 2007; Fichot and Benner, 2014). This study improves understanding of the molecules that support this link in the carbon cycle by using gene expression patterns in coastal bacterioplankton communities as biosensors of the most rapidly consumed low molecular weight substrates and FT-ICR MS analysis for molecular-level information on changes in semi-polar DOM. Osmolytes, carboxylic acids, fatty acids, sugars, organic sulfur compounds and organic nitrogen compounds, along with functional groups cleaved from larger molecules, emerged as important components of labile DOM in this ecosystem. Continued development and integration of untargeted approaches that query DOM at the level of individual compounds can advance understanding of one of the most important biological carbon fluxes on a global scale.

Experimental procedures

Near-surface water samples were collected from coastal waters near Sapelo Island, Georgia (31° 25′ 4.08″ N, 81° 17′ 43.26″ W), ~ 6 km from the mouth of Doboy Sound during July 2014 and October 2014. Samples were collected in daylight between 7:30 and 18:30. Six 20-L carboys were filled with water and wrapped in black plastic. Three were processed immediately, while the remaining three were returned to Doboy Sound for a 24 h dark incubation at 0.5 m depth before processing by an identical protocol. This experimental scheme was performed twice during each sampling event, once at high tide and once at low tide. To process the samples, 3 L were passed through a 3 μm pore-size filter to remove eukaryotic cells

(Capsule Pleated 3 μm Versapor Membrane; Pall Life Sciences, Ann Arbor, MI, USA) and then a 0.22 μm pore-size filter (Supor polyethersulfone; Pall Life Sciences) to collect bacterial cells. Subsamples for cell counts were collected prior to each filtration. The 0.22 μm filters were placed in Whirl-Pak® plastic bags (Nasco, Fort Artkinson, WI, USA) and immediately flash-frozen in liquid N_2. The 0.22 μm filtrate was used for analysis of dissolved organic matter.

Metatranscriptomics

RNA was extracted from two randomly-selected replicate filters from the initial (T0) and 24 h (T24) samples (Table 1). The 50 ml lysis tubes contained 7.5 ml CTAB Extraction Solution (Teknova, Hollister, CA), 7.5 ml phenol:chloroform:isoamyl alcohol solution (25:24:1, pH 6.7), 750 μl 10% SDS, 200 μl 1% proteinase-K, 3 g RNA PowerSoil beads (Mo-Bio, Carlsbad, CA) and internal standards (described below). Frozen filters were broken into pieces inside the Whirl-Pak bags using a rubber mallet and transferred to the lysis tubes. Tubes were vortexed for 10 min, centrifuged and the aqueous phase was transferred to a new tube. The extraction was performed again with chloroform, followed by centrifugation. RNA was precipitated from the aqueous phase with isopropanol, washed with cold 70% ethanol, and dissolved in 100 μl nuclease-free water. RNA was purified using an RNeasy Kit (Qiagen, Valencia, CA, USA) followed by two successive treatments with the Turbo DNA-free kit (Invitrogen, Carlsbad, CA, USA) to completely remove residual DNA. Samples were tested for residual DNA by a 40-cycle PCR targeting the 16S rRNA gene.

Ribosomal RNA (rRNA) was selectively removed using biotinylated rRNA probes prepared by amplification of bacterial and archaeal 16S and 23S rRNA genes and eukaryotic 18S and 28S rRNA genes using DNA from the same sample (Stewart *et al*., 2010). Probe-bound rRNA was removed via hybridization to streptavidin-coated magnetic beads (New England Biolabs, Ipswich, MA, USA), and removal of rRNA was checked using a 2200 TapeStation (Agilent Technologies, Santa Clara, CA, USA).

rRNA-depleted samples were linearly amplified using the MessageAmp II-Bacteria Kit (Applied Biosystems, Austin, TX, USA), and amplified mRNA was converted into cDNA using the Superscript III First Strand synthesis system (Invitrogen, Carlsbad, CA, USA) with random primers, followed by the NEBnext mRNA second strand synthesis module (New England Biolabs). Synthesized cDNA was purified using the PureLink® PCR Micro Kit (Invitrogen) followed by ethanol precipitation and resuspension in 70 μl TE buffer, and stored at −80°C until library preparation. The cDNA was sheared to ~ 300 bp with an E210 ultrasonicator (Covaris, Woburn, MA, USA) and TruSeq libraries (Illumina Inc., San Diego, CA, USA) were constructed. Libraries were sequenced on the Illumina HiSeq2500 platform to obtain 250 bp single end reads.

Internal standard addition

During sample processing, known copy numbers of two artificial internal mRNA standards ~ 1000 nt in length were added to each sample prior to cell lysis (Satinsky *et al*., 2014; and dx.doi.org/10.17504/protocols.io.ffwbjpe). Standards were synthesized using custom templates that were transcribed *in vitro* to RNA (Satinsky *et al*., 2012).

Bioinformatic processing

Reads from the libraries were compared with a custom database containing small and large subunit rRNAs (derived from the SILVA database; www.arb-silva.de) and the internal standard sequences (Gifford *et al*., 2011, 2013) using BLASTn. Reads with a bit score > 50 to either database were removed from further analysis. Hits to the internal standard recovered in the sequence library were tallied and used to calculate per-cell and per-litre transcript inventories. The remaining potential protein encoding reads were annotated by a homology search using RAPSearch2 against MarineRef II (http://ssharma.marsci.uga.edu/Lab/MarineRef2/), an in-house database consisting of marine bacterial, archaeal, viral, and eukaryotic genome sequences, and eukaryotic cDNAs from the MMETSP project (Keeling *et al*., 2014) that have been annotated based on SFams (Sharpton *et al*., 2012), Pfam (https://pfam.xfam.org), NCBI Protein Clusters (https://www.ncbi.nlm.nih.gov/proteinclusters), TIGRFAMS (http://www.jcvi.org/cgi-bin/tigrfams/Listing.cgi) and PhyloDB (https://drive.google.com/drive/folders/0B-BsLZUMHrDQfldGeDRIUHNZMEREY0g3ekpEZFhrTDlQSjQtbm5heC1QX2V6TUxBeFlOejQ). One replicate T0 sample from July high tide collection had contaminating sequences in the final library and was eliminated from subsequent analyses.

To identify genes whose expression changed significantly over the course of the 24 h incubations, we fit moderated negative binomial generalized linear models (Anders and Huber, 2010; Love *et al*., 2014) using the DESeq2 package in R with season, tide, incubation time and their two-way interactions to predict the $\log_2$ fold changes for each gene. DESeq2 internally normalizes for library size differences. Negative binomial models are widely used in gene differential expression analysis for modelling count data, and this distribution also enabled us to account for over-dispersion. The method corrects low dispersion estimates by modelling the dependence of the dispersion on the average expression strength (scaled counts) over all samples. We were most interested in incubation effects, and therefore analyses

emphasized expression differences between T0 and T24. Because the incubation effect was different across the four different season-tide combinations, inferences were made independently for each sample date. More specifically, the raw count data were modelled as independent negative binomial observations, and $\log_2$ expression levels of each gene (scaled by a normalization factor) were fit to a linear model using the following design:

$$\begin{aligned}\log_2 q_{ij} = {} & \mu + \alpha I(S = Oct) + \beta I(Ti = low) + \gamma I(T = 24) \\ & + \omega_1 I(T = 24, S = Oct) + \omega_2 I(T = 24, Ti = low) \\ & + \omega_3 I(T_i = low, S = Oct),\end{aligned}$$

where, q_{ij} is the mean expression level of the *i*th gene in the *j*th sample; *Ti* is the tide; *T* is the time or incubation; and *S* is the season as a categorical variable. The normalization factors are the same for all genes within a sample, and they are estimated by using the median-of-ratios method (Anders and Huber, 2010). To test for incubation effects, we constructed four Wald tests:

$$\begin{gathered}H_{0_1} : \gamma = 0 \\ H_{0_2} : \gamma + \omega_1 = 0 \\ H_{0_3} : \gamma + \omega_2 = 0, \text{and} \\ H_{0_4} : \gamma + \omega_1 + \omega_2 = 0,\end{gathered}$$

that correspond to a general incubation effect across all tide and season combinations; distinct incubation effects in different seasons; distinct incubation effects at high and low tide; and different expression levels at different tide stages in different seasons, respectively. Tests were performed for each null hypothesis for all genes. To control for type I errors, we applied the Benjamini and Hochberg adjustment for *p*-values (Benjamini and Hochberg, 1995).

Separately, we performed a principal component analysis (PCA), using centred and scaled gene expression data to capture the most variability as the coordinate axes were rotated. The data were normalized within each time/season/tide/replicate combination, with the mean of the sample used to centre the data.

The mRNA sequences used in the expression analysis are available at the NCBI SRA website under Bioproject PRJNA419903.

Cell counts

Samples of 0.6 ml were mixed with 0.6 ml of glutaraldehyde (2% final concentration) and stored at –80°C. Once thawed, 500 µl were amended with 5 µm fluorescent beads (Spherotech, Lake Forest, IL, USA), stained for 20 min with SYBR® Green I (final concentration 0.75×, Life Technologies, Waltham, MA, USA), and analysed on a CyAn Flow Cytometer (Beckman Coulter, Brea, CA, USA) with a 488 nm laser. SYBR Green fluorescence (bacteria) was detected using a FL1-530/30 bandpass filter and chlorophyll *a* fluorescence (phytoplankton) using a FL4-680/30 bandpass filter. Data were analysed in FlowJo.

Chemical analysis

The first 500–1000 ml of the 0.22 µm filtrates were discarded to avoid possible contamination from material leaching from the filters, although this is not known to be a problem. Aliquots of filtrate from the three replicate T0 and T24 treatments were stored frozen for DOC analysis. The remaining filtrates (2 L) were acidified to pH 2 with HCl, and dissolved organic matter was extracted using solid phase extraction (SPE) cartridges (Agilent Bond Elut PPL) as in Dittmar *et al*. (2008). DOC concentration in filtrates and extracts from cartridges (dried and redissolved in ultrapure water) was measured with a Shimadzu TOC-L_{CPH} analyser. Analytical accuracy and precision were tested against the Consensus Reference Material (Hansell, 2005) and were better than 5%. SPE extraction efficiency across all samples was 71% ± 4% of the DOC, which is typical for DOM with a more terrigenous signature and higher than extraction efficiencies obtained for marine DOM (Medeiros *et al*., 2015a, 2017). The molecular composition of DOM extracts (200 mg C L^{-1} in methanol) was analysed using a 9.4 T Fourier transform ion cyclotron resonance mass spectrometer with electrospray ionization (ESI; negative mode) at the National ICR Users' Facility at the National High Magnetic Field Laboratory (NHMFL, Florida State University, Tallahassee, FL, USA). Each *m/z* spectrum was internally calibrated with respect to an abundant homologous alkylation series whose members differ in mass by integer multiples of 14.01565 Da (mass of a CH_2 unit) confirmed by isotopic fine structure (Savory *et al*., 2011), achieving a mass error of < 0.4 ppm. Before each sample set, blank checks with methanol and ultrapure water were measured.

Samples were infused into the ESI interface at 400 nl min^{-1}. Molecular formulae were calculated in the mass range between 150 and 850 Da by applying the following restrictions: $^{12}C_{1–130}{}^{1}H_{1–200}O_{1–150}{}^{14}N_{0–4}S_{0–2}$. Assignment of molecular formulae was performed by Kendrick mass defect analysis (Wu *et al*., 2004) with PetroOrg software (Corilo, 2015) considering a maximum mass error of 0.5 ppm and using the criteria described by Rossel *et al*. (2013). Only mass peaks with a signal-to-noise ratio of 6 or higher were used in the analysis to eliminate inter-sample variability from peaks close to the limit of detection. FT-ICR MS data evaluation was based on normalized peak magnitudes, calculated as mass peak magnitude (i.e., intensity) divided by the summed magnitude of all mass peaks in a respective spectrum (Flerus *et al*., 2012; Lechtenfeld *et al*., 2014). The

aromaticity index (Koch and Dittmar, 2006, 2016) was combined with element ratios of molecular formulae to group molecules according to their molecular structure (Šantl-Temkiv *et al*., 2013). We emphasize that this characterization is not unambiguous and alternative structures may exist for a given molecular formula. However, categories provide a helpful overview of the likely structures behind the identified molecular formulae (Medeiros *et al*., 2015b). Patterns of DOM transformation in the 24-h incubations were identified via PCA, with statistical significance determined following Overland and Preisendorfer (1982). The patterns of transformations observed during the incubations were statistically different than what could be explained by instrument variability (i.e., repeated analysis of a single sample). To identify potential transformations in elemental compositions during the incubations, we followed the approach of Medeiros *et al*. (2017) and computed the difference in elemental composition between all molecular formulae associated with compounds whose relative abundance significantly increased and decreased during the experiment. We then sorted those differences to identify how often a potential change in elemental composition appeared. The rationale behind this calculation is that a microbial community transforming DOM via modification of specific functional groups will produce new formulae that are lower in molecular weight by an amount equivalent to the modification. Additional details are presented in Medeiros *et al*. (2017). FT-ICR data are available at BCO-DMO (https://www.bco-dmo.org/dataset/735751).

Acknowledgements

We appreciate the assistance of A. McKenna, H. Chen and Y. Corilo at the Ion Cyclotron Resonance Mass Spectrometry Users' Facility at the National High Magnetic Field Laboratory, supported by National Science Foundation Cooperative Agreement No. DMR-1157490 and the State of Florida; Roger Nilsen at the Georgia Genomics and Bioinformatics Core; and the University of Georgia's Georgia Advanced Computing Resource Centre. This research was funded by NSF grants OCE-1356010, OCE-1237140 and IOS-1656311, and by The Gordon and Betty Moore Foundation grant 5503. This is contribution 1070 of the University of Georgia Marine Institute.

Conflict of interest

The authors declare that they have no conflicts of interest related to this work.

References

Anders, S., and Huber, W. (2010) Differential expression analysis for sequence count data. *Genome Biol* **11**: R106.

Arrieta, J. M., Mayol, E., Hansman, R. L., Herndl, G. J., Dittmar, T., and Duarte, C. M. (2015) Dilution limits dissolved organic carbon utilization in the deep ocean. *Science* **348**: 331–333.

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

Benjamini, Y., and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Series B Methodol* **57**: 289–300.

Bergauer, K., Fernandez-Guerra, A., Garcia, J. A., Sprenger, R. R., Stepanauskas, R., Pachiadaki, M. G., *et al*. (2018) Organic matter processing by microbial communities throughout the Atlantic water column as revealed by metaproteomics. *Proc Natl Acad Sci* **115**: E400–E408.

Billmire, E., and Aaronson, S. (1976) The secretion of lipids by the freshwater phytoflagellate *Ochromonas danica*. *Limnol Oceanogr* **21**: 138–140.

Cole, J. J., Findlay, S., and Pace, M. L. (1988) Bacterial production in fresh and saltwater ecosystems: a cross-system overview. *Mar Ecol Prog Ser* **43**: 1–10.

Cole, J. J., Prairie, Y. T., Caraco, N. F., McDowell, W. H., Tranvik, L. J., Striegl, R. G., *et al*. (2007) Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems* **10**: 172–185.

Corilo, Y. (2015) *PetroOrg Software*. Tallahassee, FL: Florida State University.

Cottrell, M. T., and Kirchman, D. L. (2016) Transcriptional control in marine copiotrophic and oligotrophic bacteria with streamlined genomes. *Appl Environ Microbiol* **82**: 6010–6018.

Darley, W. M. (1977) Biochemical composition. In *The Biology of Diatoms*. Werner, D. (ed). Berkeley: University of California Press, pp. 198–223.

Denger, K., Weiss, M., Felux, A.-K., Schneider, A., Mayer, C., Spiteller, D., *et al*. (2014) Sulphoglycolysis in *Escherichia coli* K-12 closes a gap in the biogeochemical sulphur cycle. *Nature* **507**: 114–117.

Dittmar, T. (2015) Reasons behind the long-term stability of dissolved organic matter. In *Biogeochemistry of Marine Dissolved Organic Matter*. Hansell, D. A., and Carlson, C. A. (eds). Amsterdam: Elsevier, pp. 369–388.

Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G. (2008) A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnol Oceanogr Methods* **6**: 230–235.

Ducklow, H. W., and Carlson, C. A. (1992) Oceanic bacterial production. In *Advances in Microbial Ecology*. vol. 12. Marshall, K. C. (ed). Boston: Springer, pp. 113–181.

Durham, B. P., Sharma, S., Luo, H., Smith, C. B., Amin, S. A., Bender, S. J., *et al*. (2015) Cryptic carbon and sulfur cycling between surface ocean plankton. *Proc Natl Acad Sci* **112**: 453–457.

Fichot, C. G., and Benner, R. (2014) The fate of terrigenous dissolved organic carbon in a river-influenced ocean margin. *Global Biogeochem Cycles* **28**: 300–318.

Flerus, R., Lechtenfeld, O., Koch, B. P., McCallister, S., Schmitt-Kopplin, P., Benner, R., *et al*. (2012) A molecular perspective on the ageing of marine dissolved organic matter. *Biogeosciences* **9**: 1935–1955.

Follett, C. L., Repeta, D. J., Rothman, D. H., Xu, L., and Santinelli, C. (2014) Hidden cycle of dissolved organic carbon in the deep ocean. *Proc Natl Acad Sci* **111**: 16706–16711.

Fujita, Y., Matsuoka, H., and Hirooka, K. (2007) Regulation of fatty acid metabolism in bacteria. *Mol Microbiol* **66**: 829–839.

Gifford, S. M., Sharma, S., Rinta-Kanto, J. M., and Moran, M. A. (2011) Quantitative analysis of a deeply sequenced marine microbial metatranscriptome. *ISME J* **5**: 461–472.

Gifford, S. M., Sharma, S., Booth, M., and Moran, M. A. (2013) Expression patterns reveal niche diversification in a marine microbial assemblage. *ISME J* **7**: 281–298.

Gifford, S. M., Sharma, S., and Moran, M. A. (2014) Linking activity and function to ecosystem dynamics in a coastal bacterioplankton community. *Front Microbiol* **5**: 185.

Hansell, D. A. (2005) Dissolved organic carbon reference material program. *Eos* **86**: 318–318.

Hansell, D. A. (2013) Recalcitrant dissolved organic carbon fractions. *Ann Rev Mar Sci* **5**: 421–445.

Hansell, D. A., and Carlson, C. A. (1998) Net community production of dissolved organic carbon. *Global Biogeochem Cycles* **12**: 443–453.

Hansell, D. A., Carlson, C. A., and Schlitzer, R. (2012) Net removal of major marine dissolved organic carbon fractions in the subsurface ocean. *Global Biogeochem Cycles* **26**: GB1016.

Hellebust, J. (1965) Excretion of some organic compounds by marine phytoplankton. *Limnol Oceanogr* **10**: 192–206.

Hellebust, J. (1974) Extracellular products; algal physiology and biochemistry. *Bot Monogr* **10**: 838–865.

Hertkorn, N., Benner, R., Frommberger, M., Schmitt-Kopplin, P., Witt, M., Kaiser, K., *et al*. (2006) Characterization of a major refractory component of marine dissolved organic matter. *Geochim Cosmochim Acta* **70**: 2990–3010.

Hodson, R. E., Azam, F., Carlucci, A., Fuhrman, J. A., Karl, D. M., and Holm-Hansen, O. (1981) Microbial uptake of dissolved organic matter in McMurdo sound, Antarctica. *Mar Biol* **61**: 89–94.

Hollibaugh, J., and Azam, F. (1983) Microbial degradation of dissolved proteins in seawater. *Limnol Oceanogr* **28**: 1104–1116.

Ittekkot, V., Brockmann, U., Michaelis, W., and Degens, E. (1981) Dissolved free and combined carbohydrates during a phytoplankton bloom in the northern North Sea. *Mar Ecol Prog Ser* **4**: 299–305.

Jiao, N., Herndl, G. J., Hansell, D. A., Benner, R., Kattner, G., Wilhelm, S. W., *et al*. (2010) Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nat Rev Microbiol* **8**: 593–599.

Johnson, W. M., Kido Soule, M. C., and Kujawinski, E. B. (2017) Extraction efficiency and quantification of dissolved metabolites in targeted marine metabolomics. *Limnol Oceanogr Methods* **15**: 417–428.

Kaiser, K., and Benner, R. (2012) Organic matter transformations in the upper mesopelagic zone of the North Pacific: chemical composition and linkages to microbial community structure. *J Geophys Res: Oceans* **117**: 1023.

Keeling, P. J., Burki, F., Wilcox, H. M., Allam, B., Allen, E. E., Amaral-Zettler, L. A., *et al*. (2014) The marine microbial eukaryote transcriptome sequencing project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol* **12**: e1001889.

Kim, S., Kramer, R. W., and Hatcher, P. G. (2003) Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the van Krevelen diagram. *Anal Chem* **75**: 5336–5344.

Kirchman, D. L., Meon, B., Ducklow, H. W., Carlson, C. A., Hansell, D. A., and Steward, G. F. (2001) Glucose fluxes and concentrations of dissolved combined neutral sugars (polysaccharides) in the Ross Sea and polar front zone, Antarctica. *Deep Sea Res Part II: Topic Stud Oceanogr* **48**: 4179–4197.

Koch, B., and Dittmar, T. (2006) From mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Comm Mass Spectrom* **20**: 926–932.

Koch, B., and Dittmar, T. (2016) Erratum: from mass to structure: an aromaticity index for high-resolution mass data of natural organic matter. *Rapid Comm Mass Spectrom* **30**: 250.

Ksionzek, K. B., Lechtenfeld, O. J., McCallister, S. L., Schmitt-Kopplin, P., Geuer, J. K., Geibert, W., and Koch, B. P. (2016) Dissolved organic sulfur in the ocean: biogeochemistry of a petagram inventory. *Science* **354**: 456–459.

Lechtenfeld, O. J., Kattner, G., Flerus, R., McCallister, S. L., Schmitt-Kopplin, P., and Koch, B. P. (2014) Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean. *Geochim Cosmochim Acta* **126**: 321–337.

Lechtenfeld, O. J., Hertkorn, N., Shen, Y., Witt, M., and Benner, R. (2015) Marine sequestration of carbon in bacterial metabolites. *Nat Commun* **6**: 6711.

Longnecker, K., Futrelle, J., Coburn, E., Soule, M. C. K., and Kujawinski, E. B. (2015) Environmental metabolomics: databases and tools for data analysis. *Mar Chem* **177**: 366–373.

Love, M. I., Huber, W., and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 1.

McCarren, J., Becker, J. W., Repeta, D. J., Shi, Y., Young, C. R., Malmstrom, R. R., *et al*. (2010) Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. *Proc Natl Acad Sci* **107**: 16420–16427.

Medeiros, P. M., Seidel, M., Dittmar, T., Whitman, W. B., and Moran, M. A. (2015a) Drought-induced variability in dissolved organic matter composition in a marsh-dominated estuary. *Geophys Res Lett* **42**: 6446–6453.

Medeiros, P. M., Seidel, M., Powers, L. C., Dittmar, T., Hansell, D. A., and Miller, W. L. (2015b) Dissolved organic matter composition and photochemical transformations in the northern North Pacific Ocean. *Geophys Res Lett* **42**: 863–870.

Medeiros, P. M., Seidel, M., Gifford, S. M., Ballantyne, F., Dittmar, T., Whitman, W. B., and Moran, M. A. (2017)

Microbially-mediated transformations of estuarine dissolved organic matter. *Front Mar Sci* **4**: 69.
Mopper, K., Zhou, X., Kieber, R. J., Kieber, D. J., Sikorski, R. J., and Jones, R. D. (1991) Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycle. *Nature* **353**: 60–62.
Mopper, K., Schultz, C. A., Chevolot, L., Germain, C., Revuelta, R., and Dawson, R. (1992) Determination of sugars in unconcentrated seawater and other natural waters by liquid chromatography and pulsed amperometric detection. *Environ Sci Technol* **26**: 133–138.
Moran, M. A., Satinsky, B., Gifford, S. M., Luo, H., Rivers, A., Chan, L.-K., *et al*. (2013) Sizing up metatranscriptomics. *ISME J* **7**: 237–243.
Moran, M. A., Kujawinski, E. B., Stubbins, A., Fatland, R., Aluwihare, L. I., Buchan, A., *et al*. (2016) Deciphering ocean carbon in a changing world. *Proc Natl Acad Sci* **108**: 13612.
Neidhardt, F. C., and Umbarger, H. (1996)Chemical composition of *Escherichia coli*. In *Escherichia Coli and Salmonella: Cellular and Molecular Biology*, Neidhardt, F. C. (ed). Washington, D.C.: American Society for Microbiology, pp. 13–16.
Noinaj, N., Guillier, M., Barnard, T. J., and Buchanan, S. K. (2010) TonB-dependent transporters: regulation, structure, and function. *Annu Rev Microbiol* **64**: 43–60.
Overland, J. E., and Preisendorfer, R. (1982) A significance test for principal components applied to a cyclone climatology. *Mon Wea Rev* **110**: 1–4.
Poretsky, R. S., Sun, S., Mou, X., and Moran, M. A. (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environ Microbiol* **12**: 616–627.
Rich, J. H., Ducklow, H. W., and Kirchman, D. L. (1996) Concentrations and uptake of neutral monosaccharides along 14 W in the equatorial Pacific: contribution of glucose to heterotrophic bacterial activity and the DOM flux. *Limnol Oceanogr* **41**: 595–604.
Rossel, P. E., Vähätalo, A. V., Witt, M., and Dittmar, T. (2013) Molecular composition of dissolved organic matter from a wetland plant (*Juncus effusus*) after photochemical and microbial decomposition (1.25 yr): common features with deep sea dissolved organic matter. *Org Geochem* **60**: 62–71.
Šantl-Temkiv, T., Finster, K., Dittmar, T., Hansen, B. M., Thyrhaug, R., Nielsen, N. W., and Karlson, U. G. (2013) Hailstones: a window into the microbial and chemical inventory of a storm cloud. *PLoS One* **8**: e53550.
Satinsky, B. M., Gifford, S. M., Crump, B. C., and Moran, M. A. (2012) Use of internal standards for quantitative metatranscriptome and metagenome analysis. *Methods Enzymol* **531**: 237–250.
Satinsky, B. M., Crump, B. C., Smith, C. B., Sharma, S., Zielinski, B. L., Doherty, M., *et al*. (2014) Microspatial gene expression patterns in the Amazon River plume. *Proc Natl Acad Sci* **111**: 11085–11090.
Satinsky, B. M., Smith, C. B., Sharma, S., Landa, M., Medeiros, P. M., Coles, V. J., *et al*. (2017) Expression patterns of elemental cycling genes in the Amazon River plume. *ISME J* **11**: 1852–1864.
Savory, J. J., Kaiser, N. K., McKenna, A. M., Xian, F., Blakney, G. T., Rodgers, R. P., *et al*. (2011) Parts-per-billion Fourier transform ion cyclotron resonance mass measurement accuracy with a "walking" calibration equation. *Anal Chem* **83**: 1732–1736.
Schauer, K., Rodionov, D. A., and de Reuse, H. (2008) New substrates for TonB-dependent transport: do we only see the 'tip of the iceberg'?. *Trends Biochem Sci* **33**: 330–338.
Sharpton, T. J., Jospin, G., Wu, D., Langille, M. G., Pollard, K. S., and Eisen, J. A. (2012) Sifting through genomes with iterative-sequence clustering produces a large, phylogenetically diverse protein-family resource. *BMC Bioinform* **13**: 264.
Simjouw, J.-P., Minor, E. C., and Mopper, K. (2005) Isolation and characterization of estuarine dissolved organic matter: comparison of ultrafiltration and C18 solid-phase extraction techniques. *Mar Chem* **96**: 219–235.
Stewart, F. J., Ottesen, E. A., and DeLong, E. F. (2010) Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J* **4**: 896–907.
Taniguchi, Y., Choi, P. J., Li, G.-W., Chen, H., Babu, M., Hearn, J., *et al*. (2010) Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**: 533–538.
Van Mooy, B. A., Fredricks, H. F., Pedler, B. E., Dyhrman, S. T., Karl, D. M., Koblizek, M., *et al*. (2009) Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* **458**: 69–72.
Wei, Y., Lee, J.-M., Richmond, C., Blattner, F. R., Rafalski, J. A., and LaRossa, R. A. (2001) High-density microarray-mediated gene expression profiling of *Escherichia coli*. *J Bacteriol* **183**: 545–556.
Wetzel, R. G., Hatcher, P. G., and Bianchi, T. S. (1995) Natural photolysis by ultraviolet irradiance of recalcitrant dissolved organic matter to simple substrates for rapid bacterial metabolism. *Limnol Oceanogr* **40**: 1369–1380.
Williams, P. M., and Druffel, E. R. (1987) Radiocarbon in dissolved organic matter in the central North Pacific Ocean. *Nature* **330**: 246–248.
Wu, Z., Rodgers, R. P., and Marshall, A. G. (2004) Two-and three-dimensional van Krevelen diagrams: a graphical analysis complementary to the Kendrick mass plot for sorting elemental compositions of complex organic mixtures based on ultrahigh-resolution broadband Fourier transform ion cyclotron resonance mass measurements. *Anal Chem* **76**: 2511–2516.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Fig. S1. Changes in relative abundance for 120 ribosomal proteins between T0 and T24, pooled across the four incubation experiments. Ribosomal proteins plotting to the left side of the figure were depleted in metatranscriptomes after 24 h; those plotting to the right side were enriched. Seven ribosomal proteins with statistically significant changes are indicated with stars.

Fig. S2. Average $\log_2$ fold-change in relative contribution to the community transcriptome (T24–T0) for the 50 highest recruiting reference genomes in July (light and dark green) and October (light and dark orange). Bars plotting to the left

of centre represent taxa with higher contributions at T0 compared with T24; bars plotting to the right represent taxa with higher contributions at T24 compared with T0. Asterisks indicate significant differences between T0 and T24 (DESeq2, $p < 0.05$)

Fig. S3. Metabolism genes related to organic sulfur cycling that were significantly depleted during a 24 h dark incubation. Red dots indicate genes with significantly lower relative expression after incubation; blue asterisks indicate depletion in the July experiment, orange asterisks indicate depletion in October.

Table S1. Contribution to the total transcriptome of the 50 highest recruiting taxonomic bins. Number of taxa indicates the number of different reference genomes recruiting reads from the sample.

Table S2. T0 counts by PRK for transporter transcripts significantly depleted after a 24 h dark incubation, by season and taxon. (to be submitted as an Excel file after review)

Table S3. Mass and elemental compositions of selected formulae with the largest proportional decrease between T0 and T24. Δ Relative Abundance = average proportional decrease in the normalized peak area. Only formulae with significant decreases in all incubations are considered.

Table S4. T0 counts by taxon for transporter PRK transcripts significantly depleted after a 24 h dark incubation. (to be submitted as an Excel file after review)