Title

Experimental metatranscriptomics reveals the costs and benefits of dissolved organic matter photo-alteration for freshwater microbes

Authors

Sarah G. Nalven[1], Collin P. Ward[2], Jérôme P. Payet[1], Rose M. Cory[3], George W. Kling[3], Thomas J. Sharpton[1], Christopher M. Sullivan[1], Byron C. Crump[1]*

Affiliations

[1] Oregon State University
[2] Woods Hole Oceanographic Institution
[3] University of Michigan

***Corresponding Author**

College of Earth, Ocean, and Atmospheric Sciences
Oregon State University
104 CEOAS Admin Bldg.
Corvallis, OR 97331-5503
Phone: 541-737-4369
bcrump@coas.oregonstate.edu

Running Title

Microbial metabolism of sun-exposed organic matter

Originality-Significance Statement

This paper presents three novel findings about the interaction of sunlight and microbes during the degradation of terrigenous dissolved organic matter (DOM) in aquatic systems. First, it shows that sunlight can replace the function of certain microbial enzymes, which provides an energy savings for microbes and increases their investment in growth. Second, differential gene expression provides a new, genomics-based line of evidence for the mechanisms of photo-alteration of DOM and demonstrates the effects of these mechanisms on microbes and specific metabolic functions (through metatranscriptomics). Third, it provides the first evidence of how decarboxylation of DOM by sunlight impacts microbes.

Summary

Microbes and sunlight convert terrigenous dissolved organic matter (DOM) in surface waters to greenhouse gases. Prior studies show contrasting results about how biological and photochemical processes interact to contribute to the degradation of DOM. In this study, DOM leached from the organic layer of tundra soil was exposed to natural sunlight or kept in the dark, incubated in the dark with the natural microbial community, and analyzed for gene expression and DOM chemical composition. Microbial gene expression (metatranscriptomics) in light and dark treatments diverged substantially after 4 hours. Gene expression suggested that sunlight exposure of DOM initially stimulated microbial growth by (a) replacing the function of enzymes that degrade higher molecular weight DOM such as enzymes for aromatic carbon degradation, oxygenation, and decarboxylation, and (b) releasing low molecular weight compounds and inorganic nutrients from DOM. However, growth stimulation following sunlight exposure of DOM came at a cost. Sunlight depleted the pool of aromatic compounds that supported microbial growth in the dark treatment, ultimately causing slower growth in the light treatment over 5 days. These first measurements of microbial metatranscriptomic responses to photo-alteration of DOM provide a mechanistic explanation for how sunlight exposure of terrigenous DOM alters microbial processing and respiration of DOM.

Introduction

Inland waters, despite their modest surface area, process a substantial fraction of terrestrial carbon (C) and release this carbon to the atmosphere as CO_2 (~1-2 Pg C y^{-1}) (Cole *et al.*, 2007; Raymond *et al.*, 2013). Much of the CO_2 emitted from inland waters comes from microbial respiration of terrigenous dissolved organic matter (DOM) that is flushed from soils to streams and lakes (Cory and Kaplan, 2012; Mann *et al.*, 2012; Ward *et al.*, 2013; Sleighter *et al.*, 2014). Thus, DOM is a critical intermediate between soil organic carbon and CO_2 in the atmosphere. Understanding the controls on DOM conversion to CO_2 in inland waters is necessary to constrain local to global carbon budgets, and to forecast CO_2 emission from inland waters under future climate conditions. For example, thawing permafrost soils may release tremendous stores of DOM to inland waters of the Arctic, but we know too little about the controls on microbial respiration of this carbon to predict whether it will end up in the atmosphere as CO_2 or in the oceans as DOM (Vonk and Gustafsson, 2013; Herlemann *et al.*, 2014).

The controls on microbial respiration of DOM in surface waters are poorly defined, but likely include the *interactions* between (a) initial DOM chemistry, (b) modification of DOM chemistry by sunlight (i.e., photo-alteration), and (c) the genomic potential of the microbial community (Cory and Kling, 2018). Terrestrial DOM consists of thousands of organic molecules derived from plant and soil matter, many of which are high molecular weight (HMW), aromatic

molecules derived from lignin and tannins (Stenson *et al.*, 2003). Although these molecules are costly for microbes to degrade (Wetzel *et al.*, 1995; Vallino *et al.*, 1996; Buchan *et al.*, 2000), recent evidence suggests that HMW, aromatic, and carboxylic-acid rich DOM likely fuels the majority of microbial respiration in inland waters (Cory and Kaplan, 2012; Mann *et al.*, 2012; Ward *et al.*, 2013; Sleighter *et al.*, 2014).

These same HMW, aromatic, and carboxylic-acid rich molecules that fuel microbial respiration of DOM also absorb sunlight, and as a result can undergo photochemical mineralization to CO_2 or photochemical alteration to new compounds. For example, photo-alteration of terrestrial DOM breaks down HMW, aromatic molecules through the destabilization and cleavage of aromatic rings, oxidation of DOM, and removal of carboxyl groups (Gonsior *et al.*, 2009; Ward and Cory, 2016). This process produces lower molecular weight acids and alcohols (Kieber and Mopper, 1987; Bertilsson and Tranvik, 1998; Cory *et al.*, 2007; Gonsior *et al.*, 2009, 2014; Ward and Cory, 2016), and may release nutrients bound to DOM such as phosphorus or iron (Cotner and Heath, 1990). In turn, photochemical production of nutrients and low molecular weight (LMW) acids and alcohols can stimulate microbial growth and respiration (Wetzel *et al.*, 1995). However, stimulation of microbial activity due to production of labile carbon and nutrients may come at a cost to microbes, given that photo-alteration of DOM can also remove HMW, aromatic substrates that fuel respiration (Kaiser and Sulzberger, 2004). Thus, it follows that microbial responses to photo-alteration of DOM may be the net of the costs and benefits associated with the photochemical production and removal of important DOM substrates for microbes (Tranvik and Bertilsson, 2001; Cory *et al.*, 2013; Ward *et al.*, 2017). Consistent with this idea, prior studies have found both positive and negative effects on microbial activity when microbes are fed photo-altered DOM (Bertilsson and Tranvik, 2000; Judd *et al.*, 2006; Cory *et al.*, 2010; Reader and Miller, 2014; Ward *et al.*, 2017).

This balance of costs and benefits to microbes when DOM is altered by light is also determined by the interaction between DOM chemistry and microbial genomic potential. Genomic potential, or the gene pool of a microbial community, determines the enzymatic reactions that a microbial community is capable of performing. Interactions between genomic potential and altered DOM can be categorized into two main types: rapid changes in metabolic gene expression as individual cells retool their metabolic machinery (McCarren *et al.*, 2010; Beier *et al.*, 2015), and longer-term changes in community composition that shift the genomic potential of a community as certain populations gain selective advantage (Judd *et al.*, 2007; Logue *et al.*, 2016; Cory and Kling, 2018). To understand the controls on microbial respiration of DOM in inland waters, it is necessary to track both shorter-term transcriptomic responses and longer-term community composition responses caused by sunlight-driven changes to DOM.

Our previous research detailed the chemical changes to terrigenous DOM caused by photochemical and microbial degradation (Ward and Cory, 2016; Ward *et al.*, 2017). This research showed that sunlight broke down HMW DOM to lower molecular weight compounds, degraded aromatics, and decarboxylated and oxidized DOM (Ward and Cory, 2016), and that these photo-alterations of DOM impacted microbial activity and community composition (Ward *et al.*, 2017). However, these studies provided no evidence for the mechanisms linking photochemical and microbial degradation of DOM. Here we use new high-resolution chemical and genomic analyses of samples from this previous experiment to provide mechanistic explanations for *why* photo-alteration of DOM to lower molecular weight, less aromatic, less carboxylated, and more oxidized formulas affects microbial activity rates and community composition. We also show that photo-alteration of DOM caused global shifts in microbial

118 community gene expression and active taxonomic groups. These shifts suggest that photo-alteration of DOM initially stimulated microbial growth by replacing key steps in DOM
120 metabolism pathways, but ultimately suppressed microbial activity by removing DOM compounds that native microbial communities were adapted to use.
122

Results and Discussion

124 We conducted a replicated experiment in which natural microbial communities were incubated with DOM that had been exposed to 24 h of natural sunlight or kept in the dark (i.e.,
126 light and dark treatments). Gene expression is highly responsive to the chemical composition of DOM (McCarren *et al.*, 2010; de Menezes *et al.*, 2012; Shi *et al.*, 2012), and microbial
128 communities in this experiment adjusted to photo-altered DOM with both rapid changes in gene expression and longer-term changes in community composition. Short-term changes, assessed
130 with metatranscriptomics after 4 h incubations, showed that 27% of genes annotated to KEGG orthologs (KOs) were differentially expressed (False discovery rate or FDR < 0.05) between
132 light and dark treatments (Fig. 1A, Dataset S1). These differences caused metatranscriptomes from each treatment to cluster separately on a principle coordinates diagram (Adonis
134 PERMANOVA, $p = 0.1$; Fig. S1), and caused lower Shannon alpha diversity of metatranscriptomes in the light treatment (Fig. S1). Photo-altered DOM also caused a shift in the
136 composition of taxa with active expression (Fig. 1B), most notably causing increased expression by Gammaproteobacteria and reduced expression by Bacteroidetes. These changes in expression
138 were consistent with small initial shifts (4 h) in microbial community composition (assessed with 16S rRNA gene amplicon sequencing), and may have led to substantial longer-term shifts in
140 community composition after 5 d (Fig. S2; Dataset S6) (Ward *et al.*, 2017).

The differences in microbial community responses between light and dark treatments were
142 driven by changes in DOM chemistry (see Ward and Cory, 2016, for a description of photochemical processes that affected DOM composition). For example, sunlight exposure
144 removed about 5% of DOC by converting it to CO_2 via photo-decarboxylation or other photo-oxidation processes (Fig. S5; Ward and Cory, 2016). Figure 2 shows that of the 375 formulas
146 removed by sunlight (region II), 20% were aromatic, consistent with photochemical removal of aromatic C (quantified by ^{13}C-NMR; Ward and Cory, 2016), and 74% were classified as tannin-
148 like (corresponding with the high average molecular weight and O/C ratio of these formulas shown in region II, Fig. 2). Tannin-like formulas represent a fraction of DOM likely rich in
150 carboxylic acid functional groups (Ritchie and Perdue, 2003; Ward and Cory, 2016). Photochemical removal of carboxyl carbon (detected by ^{13}C NMR as shown in Fig. S5)
152 concurrent with removal of tannin-like DOM is consistent with photo-decarboxylation of DOM (Ward and Cory, 2016). Compared to the formulas removed by sunlight, the formulas produced
154 by sunlight were of lower molecular weight and less oxygen rich (region III, Fig. 2). These results demonstrate that sunlight degraded HMW, aromatic, and carboxyl-containing DOM into
156 LMW, less aromatic, and less carboxylated DOM (Ward and Cory, 2016). Although the formulas produced by sunlight were less oxygen rich on average compared to formulas removed
158 by sunlight (region III vs. region II in Fig. 2; consistent with decarboxylation of oxygen-rich formulas), a substantial fraction of formulas produced by sunlight exposure of DOM are
160 oxidation products (i.e., produced by photo-oxidation of DOM; Cory *et al.*, 2010; Ward and Cory, 2020). Thus, sunlight exposure simultaneously degraded HMW and aromatic DOM via
162 decarboxylation and oxidation pathways (Ward and Cory, 2016, 2020; Fig. S5; Fig. 2).

This DOM produced and removed by sunlight was similar in composition to DOM degraded by microbes in the dark treatment. For example, of the formulas produced by sunlight, 98 of these were identical to formulas degraded by microbes (region IV in Fig. 2). Of the formulas removed by sunlight, 148 of these were identical to formulas degraded by microbes (region V in Fig. 2). Thus, of the total 383 formulas degraded by microbes (region I, Fig. 2), there was substantial overlap of chemical formulas between the pool of DOM altered by sunlight and the pool of DOM degraded by microbes in the dark treatment (98 formulas produced by sunlight plus 148 removed by sunlight results in 246 formulas identical to those degraded by microbes; thus 64% of the 383 formulas degraded by microbes were identical to those altered by sunlight; Fig. 2). This result suggests that sunlight produced and removed formulas that microbes were metabolically equipped to degrade in the dark (Ward *et al.*, 2017). To identify which sunlight-driven changes to DOM chemistry are relevant to microbial communities, and to understand the mechanisms by which microbes adjust to these changes, we investigated differential expression in light and dark treatments of (a) major gene categories, (b) specific DOM metabolism genes, and (c) membrane transport genes responsible for the supply of external nutrients and substrates.

Global Gene Expression Patterns

First, expression patterns across all major gene categories suggested that in the dark treatment, microbial communities were invested in scavenging – that is, finding and using DOM and inorganic nutrients. In contrast, light-treatment microbial communities were more invested in growth. Microbial communities incubated with dark treatment DOM had greater expression of genes for metabolism, motility, and resource transport than did microbial communities incubated with photo-altered DOM (Fig. 3). These patterns in differential gene expression were also detected for Bacteroidetes, Gammaproteobacteria, and Betaproteobacteria when analyzed individually (Fig. S3, Datasets S2-S4). This finding suggests that microbial communities incubated with dark treatment DOM were allocating energy towards degrading organic compounds and searching for resources. Across all taxa, carbohydrate, lipid, amino acid, and xenobiotic metabolism were elevated in the dark treatment (Fig. 3C), as were many signal transduction and cell motility genes (Fig. 3D), including several chemotaxis and pilus related proteins (e.g., *mcp*, *cheR*, *pilI*, *pilG*), enzymes and transporters involved in phosphorus and nitrogen acquisition (e.g., *phoR, phoD*, *ntrY, nifA*), and flagellin (e.g., *fliC*) (Dataset S1). Elevated expression of these genes in the dark versus light treatment indicates that communities incubated with dark-treatment DOM experienced comparatively poor conditions for growth, because cells allocate more energy toward moving and scavenging when resources are scarce (Soutourina and Bertin, 2003). This interpretation of gene expression in the dark treatment is consistent with current understanding of microbial metabolism of terrigenous DOM. For example, HMW and aromatic carbon in terrestrial DOM fuels a substantial fraction of microbial respiration (Cory and Kaplan, 2012; Ward *et al.*, 2013; Sleighter *et al.*, 2014) despite being energetically costly for microbes to degrade (Wetzel *et al.*, 1995; Buchan *et al.*, 2000).

In contrast to communities in the dark treatment, communities incubated with photo-altered DOM had significantly higher expression of genes involved in transcription and translation (Fig. 3B), suggesting that photo-alteration of DOM caused microbial communities to invest more in growth and less in scavenging. These patterns of differential gene expression were also detected for Bacteroidetes, Gammaproteobacteria, and Betaproteobacteria when analyzed individually (Fig. S3, Datasets S2-S4). Across all taxa, genes in these categories that were expressed more in the light treatment than the dark treatment included two highly expressed RNA polymerase

genes (*rpo*A, *rpo*B), and all 42 differentially expressed ribosomal protein genes (Dataset S1). Cells tightly regulate transcription and translation, and allocate more resources to these processes when growing exponentially (log phase) (Nomura *et al.*, 1984; Kraakman *et al.*, 1993; Scott *et al.*, 2010) or preparing to do so (lag phase) (Rolfe *et al.*, 2011; Madar *et al.*, 2013). Moreover, overwhelming evidence from transcriptional studies demonstrates a positive relationship between expression of genes for transcription and translation and bacterial cell growth (Franchini and Egli, 2006; Lahtvee *et al.*, 2011; Harke and Gobler, 2013; Matsumoto *et al.*, 2013; Gifford *et al.*, 2016). Increased expression of genes involved in transcription and translation suggests that photo-alteration of DOM produced new compounds that induced microbial growth. Exact overlap in a subset of DOM formulas produced by sunlight (photo-altered DOM) and degraded by microbes (region IV in Fig. 2) suggests these are the kinds of compounds that induced microbial growth. These formulas were on average relatively LMW, aliphatic, and classified predominately as lignin-like (region IV in Fig. 2). Production of compounds that induce microbial growth is consistent with prior work suggesting that photo-alteration of DOM creates compounds that are more labile to microbes compared to DOM used by microbes in the dark (Wetzel *et al.*, 1995; Moran and Zepp, 1997; Bertilsson and Tranvik, 1998; Cory *et al.*, 2010; Satinsky *et al.*, 2017).

Taxonomic binning of ribosomal protein transcripts revealed that photo-alteration of DOM increased ribosomal expression by Gammaproteobacteria and Betaproteobacteria and decreased ribosomal expression of Bacteroidetes relative to their 16S rRNA gene abundances (Fig. S4). Together, Gammaproteobacteria and Betaproteobacteria were responsible for 90% of ribosomal protein transcripts in the light treatment (62% and 28%, respectively), and only 49% in the dark treatment (30% and 19%, respectively). The ratio of ribosomal gene expression to 16S gene abundance for Gammaproteobacteria and Betaproteobacteria was greater in the light treatment (90% / 60% = 1.5) than in the dark treatment (49% / 51% = 0.96), suggesting that growth by these groups was favored by DOM photo-alteration more than other groups. Conversely, photo-alteration of DOM decreased ribosomal expression by Bacteroidetes from 36% in the dark treatment to 6% in the light treatment. Moreover, the ratio of ribosomal gene expression to 16S gene abundance for Bacteroidetes was lower in the light treatment (6% / 4% = 1.5) than in the dark treatment (36% / 13% = 2.78), suggesting that Bacteroidetes growth was not favored by DOM photo-alteration. Although these phyla contain a diverse array of organisms that may not all react similarly to environmental changes, differences in ribosomal protein expression by these groups at 4 h provide a mechanistic explanation for how and why community composition between light and dark treatments shifted over the longer-term (5 d) (Ward *et al.*, 2017). That is, taxonomic groups with high ribosomal expression at 4 h were more dominant members of the community at 5 d, indicating that one way in which microbial communities may have adapted to photo-altered DOM is through selection for growth of certain populations, which in turn results in longer-term changes in community composition. These results provide evidence for a mechanism of community change, hypothesized by others (Judd *et al.*, 2007; Ward *et al.*, 2017; Cory and Kling, 2018), that greater ribosomal expression of taxa precedes increased abundance of the same taxa.

DOM Metabolism Genes

The second category of genes we investigated was DOM metabolism genes. Differential gene expression of specific DOM metabolism genes, taken together with sunlight-induced changes in DOM chemistry, suggested that sunlight replaced the function of key genes in

microbial DOM degradation pathways. Sunlight broke down HMW DOM into LMW DOM, decreased the aromatic content of DOM, oxidized DOM, and decarboxylated DOM (Fig. 2) (Ward and Cory, 2016). Most differentially expressed genes involved in these same processes, specifically, the aromatic degradation genes, oxygenase genes, and decarboxylase genes, were significantly less expressed by microbes incubated with photo-altered DOM, even when expression was calculated as a percentage of total Metabolism gene expression (a KEGG tier II category; Fig. 4A, Dataset S1). Expression of these genes was also reduced in the light treatment for Gammaproteobacteria and Betaproteobacteria when analyzed separately (Datasets S3-S4), and showed no pattern with treatment for Bacteroidetes and other taxa (Datasets S4-S5). Further evidence that microbes and sunlight used the same processes to degrade DOM was the strong overlap in number and chemical composition of the formulas degraded by both microbes and sunlight (Fig. 2). Of the 148 formulas degraded by both microbes and sunlight, 90% were tannin-like formulas (region V, Fig. 2) previously shown to be decarboxylated to smaller and less aromatic formulas by sunlight (Ward and Cory, 2016).

Genes for aromatic degradation, oxygenation, and decarboxylation encode costly enzymes that destabilize and break down large and complex carbon compounds (Cavin *et al.*, 1998; Bugg *et al.*, 2011; Fuchs *et al.*, 2011; Gulvik and Buchan, 2013). Given that microbes carefully regulate gene expression to produce costly enzymes only when necessary (Browning and Busby, 2004), reduced expression of these gene categories in the light treatment suggests that sunlight destabilized and cleaved aromatic rings, oxidized DOM, and removed carboxyl groups from DOM, minimizing the need for enzymes that perform these functions. These findings are consistent with the global gene expression patterns discussed above, which suggest that after 4 h, the light-treatment microbial communities were investing in growth. These findings provide an explanation for this investment in growth, suggesting that sunlight initially stimulated microbial growth by replacing enzymatic functions and transforming DOM into more easily consumed products. In this way, photo-alteration of DOM relieved microbes of energetically expensive needs and allowed microbial communities to put energy and resources towards activities such as transcription and translation.

This differential expression of specific metabolic genes not only reveals *biological* mechanisms of DOM degradation after photo-alteration, but also provides a new, genomics-based line of evidence for the effects and mechanisms of *photochemical* alteration of DOM. These metatranscriptomic data demonstrate that sunlight alters microbial metabolism of DOM by breaking down aromatic compounds and oxidizing DOM. This evidence is consistent with numerous studies showing that sunlight cleaves aromatic rings (Strome and Miller, 1978; Stubbins *et al.*, 2010) and oxidizes organic matter (Cory *et al.*, 2010; Gonsior *et al.*, 2014, Ward and Cory, 2020), likely in part through the photochemical production of reactive oxygen species from DOM (Cory *et al.*, 2010; Page *et al.*, 2014). We also provide metatranscriptomic evidence suggesting that sunlight decarboxylates DOM. Reduced expression of decarboxylase genes in response to photo-altered DOM suggests that decarboxylation is an important pathway for microbial metabolism of carboxylic acids within terrigenous DOM (an abundant fraction of DOM in inland waters) (Ritchie and Perdue, 2003; Ward and Cory, 2015). Photo-decarboxylation of DOM has been inferred (Faust and Zepp, 1993; Xie *et al.*, 2004), but in contrast to aromatic degradation and oxidation of DOM, direct evidence of photo-decarboxylation of terrigenous DOM is limited (Ward and Cory, 2016). The metatranscriptomic data presented here are in strong agreement with chemical evidence supporting decarboxylation as a driving pathway of terrigenous DOM photo-alteration (Ward and Cory, 2016).

Differential expression of specific metabolic genes also highlights the similarities between photo- and bio-degraded fractions of DOM. Reduced expression of aromatic degradation genes, oxygenase genes, and decarboxylase genes in response to photo-altered DOM indicates a substantial competition between sunlight and microbes to degrade similar types of DOM (Bowen *et al.*, 2019).

Membrane Transport Genes

The third category of genes we investigated was the ABC transporter genes. Differential expression of these transporter genes, which accounted for the bulk of membrane transport expression, suggested that photo-alteration of DOM increased the availability of sugars and alcohols, and changed the chemical forms of phosphorus, sulfur, and ferric iron (Fe(III)) available to microbes. Expression of total ABC transporter genes was similar across treatments (Fig. 3D), but 65 genes representing subunits for 35 transporters had significantly different expression across treatments when normalized to total ABC Transporter expression (a KEGG tier IV category; Fig. 4B, Dataset S1).

All differentially expressed sugar and polyol transporter genes were expressed more in the dark treatment than in the light treatment (Fig. 4B), except for the sn-glycerol 3-phosphate transporter genes (*ugp*, i.e., K05813, K05814, K05815, and K05816). Expression of these genes was also greater in the dark treatment for Gammaproteobacteria, Betaproteobacteria, and Bacteroidetes when analyzed separately (Datasets S2-S4). Greater expression of sugar and polyol transporter genes in the dark treatment suggests that sugars and polyols were limiting, because suboptimal levels of these substrates induce expression of their transporter proteins (Ferenci, 1999). In turn, lower expression of sugar and polyol transporter genes in the light treatment suggests that photo-alteration of DOM increased availability of sugars and alcohols. Sugars and alcohols contain oxidized functional groups within DOM, and there is some evidence that exposure to sunlight can increase the abundance of these and other oxidized functional groups (Gonsior *et al.*, 2014; Ward and Cory, 2016, Ward and Cory, 2020).

Transporter genes for organic phosphorus, such as sn-glycerol 3-phosphate (G3P; *ugp*A, *ugp*B, *ugp*C, *ugp*E) and phosphonate (*phn*E) transporter genes, were more expressed in the light treatment across all taxa (Fig. 4B, Dataset S1), and for Gammaproteobacteria and Betaproteobacteria when analyzed separately (Datasets S3-S4). In contrast, expression of these genes was greater in the dark treatment for Bacteroidetes and other taxa (Datasets S5-S6). The increase in phosphorus transporter expression in the light treatment for Gammaproteobacteria and Betaproteobacteria is consistent with an increase in microbial demand for phosphorus for activities such as biosynthesis, and may also indicate a shift in the available forms of phosphorus. G3P is a degradation product of plant cell phospholipids which, along with phosphonates, are abundant in a variety of soils (Tate and Newman, 1982; Turner *et al.*, 2004). Organophosphates like G3P serve diverse cellular functions, and uptake may be preferable to inorganic phosphate when cells carry out certain functions. For example, the first step in the biosynthesis of phospholipid membranes is the synthesis of G3P (Cronan, Jr. and Rock, 2008), which is probably less energetically costly to import than it is to synthesize (Ames, 1986). Elevated expression of G3P and phosphonate transporter genes may reflect a greater need for phosphorus for biosynthesis by the rapidly growing light-treatment microbial community. In addition, expression of G3P transporter genes is induced when inorganic phosphate is limiting and G3P is available (Brzoska *et al.*, 1994; Vershinina and Znamenskaya, 2002; León-Sobrino *et al.*, 2019), so elevated expression of G3P and phosphonate transporter genes may indicate that

photo-alteration of DOM makes inorganic phosphorus less available or organic phosphorus more available.

Genes for sulfate transporters were also more expressed in the light treatment across all taxa (Fig. 4, Dataset S1) and for Gammaproteobacteria and Betaproteobacteria when analyzed separately (Datasets S2-S3) but showed no pattern for Bacteroidetes and other taxa (Datasets S4-S5). Similar to expression of organic phosphorus transporter genes, this elevated expression of sulfate transporter genes is consistent with an increase in microbial demand for sulfur, and may also indicate that organic sulfur sources were less available after photo-alteration of DOM. Sulfate assimilation is an energy consuming process; energy is required to transport this ion across membranes and to reduce sulfur from an oxidation state of +6 to -2 for incorporation into cystine. Consequently, sulfate transporter genes are only expressed when sulfur is required and when favorable organic sulfur compounds are unavailable (Piłsyk and Paszewski, 2009; Campanini *et al.*, 2014). Therefore, elevated expression of sulfate transporter genes suggests that first, sulfur was needed by the light-treatment microbial communities, perhaps because these communities were growing, and second, photo-degradation of DOM made organic sulfur compounds less available than sulfate. Consistent with this interpretation, all differentially expressed genes for organic sulfur catabolic enzymes were more expressed in the dark treatment across all taxa (Dataset S1) and for Gammaproteobacteria and Betaproteobacteria when analyzed separately (Datasets S2-S3) but were not differentially expressed by Bacteroidetes and other taxa (Datasets S4-S5). These genes included sulfatases (*bet*C*, asl*A*,* K01138), a sulfotransferase (*rax*ST), an alkanesulfonate monooxygenase *(ssu*D*),* and a sulfoxide reductase *(msr*P*)* (Dataset S1).

Changes in DOM composition do not clearly support a decrease in availability of organic sulfur compounds in the light treatment. In the dark treatment, 7% of the formulas degraded by microbes contained S (Fig. 2). Of the formulas produced by light and degraded by microbes (group IV in Fig. 2), 9% contained S (Fig. 2). These results suggest that photo-alteration of DOM led to a 2% increase in availability of organic sulfur compounds compared to the dark treatment. Small differences in S-containing formulas between the dark and light treatments should be interpreted with caution due to the bias of FT-ICR MS against the detection of heteroatom containing DOM (e.g., Hockaday *et al.*, 2009) especially considering the low organic sulfur content of DOM in this watershed (Cory *et al.*, 2007).

Results from other studies support increased expression for sulfate transporter genes in the light treatment. For example, one study demonstrated substantial photo-oxidation of organic sulfur within DOM (Gomez-Saez *et al.*, 2017). Another study showed that organic sulfur within terrigenous DOM is readily mineralized to sulfate by sunlight (Ossola *et al.*, 2019).

Like the expression of phosphorus and sulfur transporter genes, expression of Fe(III) transporter genes also suggests that photo-alteration of DOM caused changes to the chemical forms and availability of iron. However, Fe(III) transporter expression indicated both light and dark treatments were actively scavenging iron (Cornelis *et al.*, 2009; Noinaj *et al.*, 2010). The main difference between the treatments was that in the dark treatment, genes involved in TonB-dependent Fe(III) transport were more expressed (*tonB, exbB, fecA, fecR)*, while in the light treatment, TonB-independent Fe(III) transporter genes were more expressed (*fbpA, fbpB, fbpC;* Fig. 4B, Dataset S1). This pattern in expression occurred across all taxa, and for Bacteroidetes (*tonB, exbB, exbD* in dark; *fbpA* in light), Gammaproteobacteria (*exbB, fecA, fecR, feoA, fhuE, hemR* in dark; *fbpA* in light), Betaproteobacteria (*fhuF* in dark; *fbpA* in light), and other taxa (*fbpA* in light) when analyzed separately (Datasets S1-S5). The inner membrane proteins tonB

and exbB (along with exbD) transduce proton motive force to TonB-dependent transporters in the outer membrane that bind and transport chelated Fe(III) including siderophore-bound Fe(III). In contrast, TonB-independent systems transport unchelated Fe(III) across the inner membrane after it crosses the outer membrane either passively or via some unknown outer membrane system (Wyckoff *et al.*, 2006; Zhang *et al.*, 2018). This suggests that dark-treatment microbes were importing chelated Fe(III) and light-treatment microbes were importing unchelated Fe(III). Transcription of these different Fe(III) transport systems is regulated in part by characteristics of Fe(III) chelating molecules (Zhang *et al.*, 2018; Dong *et al.*, 2019), suggesting that this shift in expression is driven by changes in the available forms of Fe(III). Consistently, sunlight is thought to break down chelated Fe(III) and release free Fe(III) (Voelker *et al.*, 1997), which can then precipitate or become loosely bound to other organic compounds. Therefore, it is likely that microbes in the light treatment were actively transporting Fe(III) that had been recently released from photo-degraded siderophores or other chelating molecules, such as carboxylic acids (Fujii *et al.*, 2014).

Unlike transporter genes for phosphorus, sulfur, and Fe(III), no pattern was apparent among amino acid transporter genes. Of the 20 differentially expressed amino acid transporter genes, 12 were more expressed in the dark treatment and eight were more expressed in the light treatment (Fig. 4B). Overall, however, photo-alteration of DOM caused lower expression of genes for transporting sugars and alcohols, higher expression of genes for transporting organic phosphorus and sulfate, and a shift in expression of genes for Fe(III) transport from chelated to unchelated Fe(III). Thus, ABC transporter expression suggests that photo-alteration of DOM makes many LMW organic compounds more available and changes the chemical state of several essential nutrients (i.e., phosphorus, sulfur, Fe(III)).

Photo-alteration of DOM Caused Transient Growth

The transcriptional responses outlined above (Fig. 5) suggest greater initial growth in the microbial communities incubated with photo-altered DOM. First, greater expression of ribosomal proteins and RNA polymerase genes in the light treatment suggests microbial investment in gene products that support growth. Second, lower expression of aromatic degradation genes, oxygenases, and decarboxylases in the light treatment suggests that photo-alteration of DOM replaced the function of these genes by producing compounds that were more accessible to microbes and thus capable of stimulating growth. Third, lower expression of sugar and polyol transporters in the light treatment and higher expression of organic phosphorus, sulfate, and Fe(III) transporters suggests that photo-alteration of DOM released LMW compounds that could have also stimulated growth and altered the availability of several nutrients.

Interestingly, these indicators of growth in the light treatment communities at 4 h, suggested by metatranscriptomic data, did not result in higher rates of microbial activity over the full 5 d incubation. We measured rates of respiration (O_2 consumption and CO_2 production) and new cell production (direct cell counts) over 5 d, and the rate of biomass production after 5 days (determined as the rate of incorporation of ^{14}C leucine), and all measures were lower for microbes in the light treatment compared to the dark treatment (Fig. 6). Moreover, approximately twice as much DOC was consumed by microbes in the dark treatment compared to the light treatment over 5 d (Fig. S5). This contrast between 4 h and 5 d suggests that photo-alteration of DOM supported only a transient burst of growth. Given the relatively short duration of the photo-exposure (~ 24 h natural sunlight), the production of photo-altered DOM compounds may

have been too small to boost microbial activity over the entire 5 d incubation. It is possible that a top-down control, such as viral lysis, could have caused slower microbial activity in the light treatment incubations at 5 d, but this is unlikely considering the triple-replication of the experiment and the lack of evidence of top-down control in the dark treatments or in similar studies (Judd *et al.*, 2006, 2007). It is more likely that the lower activity of microbes at 5 d in the light versus dark treatment was due to a combination of (1) too little production of beneficial compounds during the short photo-exposure of DOM to stimulate growth for 5 full days, and (2) the photochemical removal of DOM compounds that communities had been equipped to degrade prior to photo-exposure (Fig. S5) (Ward *et al.*, 2017). In other words, the benefit provided to microbes by photo-production of growth-stimulating compounds came at the cost of losing many of the DOM compounds that supported microbial activity in the dark treatment.

These results make clear that the microbial response to photo-altered DOM can be dynamic, with different effects in the short-term (hours) and long-term (days). In fact, several studies have found transient responses, especially those that include sampling after 24 hours or less of microbial incubation with photo-altered DOM (Kaiser and Sulzberger, 2004; Judd *et al.*, 2007; Gareis and Lesack, 2018). Two studies (Kaiser and Sulzberger, 2004; Judd *et al.*, 2007) found that photo-exposure of DOM initially (1-3 h) inhibited microbial activity, but over time (5-19 d) the activity of microbes growing on photo-altered DOM caught up to or exceeded activity of dark-treatment microbes. Another study (Gareis and Lesack, 2018) found the opposite response, that photo-exposure of DOM initially stimulated cell-specific bacterial production after 24 h before declining to lower bacterial production rates than the dark-treatment microbes. Regardless of positive or negative response, the transient effect on microbial activity in these studies was inferred to be caused by changes in DOM composition, consistent with our interpretation. However, in our experiment, the microbial response was more similar to the response seen by Gareis and Lesack (2018); we found that photo-alteration appeared to initially support a growth response through replacement of key steps in enzymatic pathways, and to later slow microbial activity most likely due to depletion of microbially-favored substrates (Ward *et al.*, 2017). These results and previous studies highlight the transient nature of microbial responses to photo-exposed DOM and demonstrate the dramatic impact of short- and long-term microbial adjustments to changes in DOM chemistry.

Conclusions

The fate of carbon in inland waters depends on the mechanisms by which sunlight and microbes transform terrigenous DOM. Understanding these mechanisms is especially critical in the Arctic because (a) arctic lakes and streams contain high amounts of terrigenous DOM (Cory *et al.*, 2007; Caplanne and Laurion, 2008; Gareis *et al.*, 2010), (b) terrigenous DOM is an important carbon source for microbial communities in these oligotrophic systems (Crump *et al.*, 2003; Mann *et al.*, 2012), (c) export of terrigenous DOM to sunlit waters is expected to increase as the Arctic's vast stores of soil carbon thaw (Rowland *et al.*, 2010), and (d) arctic lakes and streams are generally unshaded and shallow, making photo-alteration a critical control of DOM processing relative to more shaded and deeper aquatic ecosystems (Cory *et al.*, 2014).

Our results suggest that aromatic degradation, oxidation, and decarboxylation are important mechanisms by which both sunlight and microbes independently break down DOM. Consequently, the photo-alteration of DOM by sunlight functionally replaces key steps in microbial DOM degradation pathways and can stimulate microbial growth. In addition, our findings suggest that photo-alteration of DOM releases LMW compounds and changes the

availability of nutrients, which may also contribute to greater activity and growth. However, these benefits of DOM photo-alteration came at a cost to microbes in this experiment because sunlight removed many of the DOM compounds that microbial communities were metabolically equipped to use (Ward *et al.*, 2017). Loss of these compounds reduced microbial activity over the 5 d incubation, likely after growth-stimulating photo-products were depleted.

These contrasting responses of microbes in the short- and long-term (4 h and 5 d) show that sunlight can both produce and eliminate compounds that are useful to microbial communities, but that the net effect of DOM photo-alteration depends on (a) which compounds are produced and eliminated and (b) the timescales that microbes require to adjust to photo-altered DOM. Our results suggest that when DOM moves from soils to sunlit lakes and streams, microbial communities can quickly shift gene expression to benefit from materials released from photo-altered DOM. Over longer timescales (e.g., days to weeks), shifts in community composition can also allow microbes to benefit from DOM photo-alteration (Judd *et al.*, 2007; Cory *et al.*, 2013). However, in our study, shifts in community composition did not appear to benefit microbes, suggesting that in this case (but after only 5 days), longer-term microbial adjustments could not overcome the loss of bioavailable material due to photo-alteration. Thus, the longer-term benefits of DOM photo-alteration depend on whether the material produced by photo-alteration is derived from compounds that microbes are equipped to degrade, or from relatively refractory compounds that microbes cannot access. If photo-altered DOM is primarily derived from the former, as in this study, then the cost to microbes in loss of resources may outweigh the benefits, despite short- and longer-term adjustments by microbial communities.

As the Arctic warms, there may be increases in export of terrigenous DOM to sunlit surface waters due to thawing permafrost (Rowland *et al.*, 2010) and greater UV exposure of DOM due to longer ice-free seasons (Šmejkalová *et al.*, 2016). However, the photo-reactivity and biological availability of this DOM will likely differ from the DOM currently draining to arctic lakes and streams (e.g., Cory *et al.*, 2013; Ward and Cory, 2016; Stubbins *et al.*, 2017). Therefore, to forecast the fate of this DOM, it is increasingly important to understand the interactions between photochemical and biological DOM degradation. Our measurements of microbial metatranscriptomic responses to photo-alteration of DOM, paired with high-resolution DOM composition data, provide mechanistic explanations for how photo-alteration of DOM in inland waters affects rates of microbial activity and thus DOM fate and arctic carbon cycling.

Experimental Procedures

In this study we used new genomics results generated from a previously published experiment (Ward and Cory, 2015, 2016; Ward *et al.*, 2017) to provide mechanistic explanations for why photo-alteration of DOM affects microbial activity rates and community composition. Below, we first briefly summarize the previously published experimental design and methods, then we summarize the new genomic techniques and methodology.

Summary of experimental design and methods

Figure S6 is a schematic of the experimental design from which microbial communities from tundra soil leachates were incubated in triplicate with both light-exposed and dark-exposed soil-derived DOM for analysis in this study. The preparation and treatment of soil-derived DOM for these incubations was described previously (Ward and Cory, 2015). Briefly, soil samples from the organic layer of three adjacent pits were collected at 5-15 cm depth on June 15, 2013 in the Imnavait Creek watershed on the North Slope of Alaska (68.62° N, 149.28° W; elevation ~ 900 m). Soil was collected in plastic bags, immediately transferred to coolers, and within hours

placed in freezers at Toolik Field Station. Leachate was made by adding an equal mass of soil from each of three replicate pits for a total of 3600 g of soil and 15 L of deionized water, followed by filtration through 0.45 μm high-capacity cartridge filters (Geotech Environmental Equipment, Inc., Denver, CO). Filtration with a larger pore size (GF/F filters, nominal pore size 0.7 μm) has previously been shown to substantially reduce microbial contamination, reducing bacterial production by 93 ± 2% (mean ± 1 SD) compared to unfiltered lake water (Ward *et al.*, 2017).

Each of the three replicates of DOM leachate was split into a light treatment and dark treatment. Both treatments were placed in UV-transparent Whirlpak bags (5 L) and exposed to 24 h of natural sunlight at Toolik Field Station on June 24 and 25, 2013; dark treatments were wrapped in aluminum foil. Details of the sunlight exposure experiment are reported in Ward and Cory (2016).

Before and after sunlight exposure, subsamples were collected for DOM chemical characterization by Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR MS) and ^{13}C nuclear magnetic resonance spectroscopy (^{13}C NMR). The chemical composition of DOM by FT-ICR MS and ^{13}C-NMR used in this study was previously reported in Ward and Cory (2015, 2016) and in Ward *et al*. (2017), along with the methodology used to analyze and interpret DOM composition. Briefly, DOM was extracted using 5 g PPL solid-phase (SPE) cartridges for FT-ICR MS analysis. DOC recovery was 57 ± 1% (± 1 SE; Ward and Cory, 2015). Methanol SPE eluates were diluted to ~50 mg C per L prior to introduction to the electrospray ionization source of a 12T Bruker SolariX FT-ICR mass spectrometer. All spectra were acquired in negative mode. Formula assignment criteria have previously been described in detail (Ward and Cory, 2015, 2016). Aromatic or aliphatic character of formulas produced or degraded by sunlight or microbes was determined using the aromaticity index (AI_{MOD}; Koch and Dittmar, 2006). Formulas were assigned to compound classes using the following criteria: Tannin-like: $0.6 \leq O/C \leq 1.2$, $0.5 \leq H/C \leq 1.5$, $AI_{MOD} < 0.67$; Lignin-like: $0.1 < O/C < 0.6$, $0.5 \leq H/C \leq 1.7$, $AI_{MOD} < 0.67$ (Ward *et al.*, 2017 and therein).

Formulas were categorized as degraded versus produced by sunlight if their intensity decreased or increased after light exposure, respectively, using the 95% confidence interval of the mean of experimental replicates (N =3) to determine if a change in peak intensity was significantly different from zero (Ward and Cory, 2016). Formulas were categorized as consumed by or resistant to microbes if their intensity decreased or remained unchanged after incubation with the native microbial community, respectively (Ward *et al.*, 2017). The 95% confidence intervals calculated across experimental replicates were used to determine if a change in formula intensity after incubation with microbes was significantly greater than zero (N = 3; Ward *et al.*, 2017). This approach to analyze FT-ICR MS spectra accounts for experimental variability (e.g., natural variability in microbial respiration between incubations, extraction efficiency of PPL cartridges) and instrumental variability (e.g., ionization efficiency).

Whirlpak bags were leached with laboratory grade water (DI water) under the same experimental conditions as the soil leachates (e.g., water volume, dark and light treatments, and experiment time) to determine whether these containers added organic carbon contamination to the soil leachates. DOC concentrations were slightly higher in Whirlpak bags leached with the DI water compared to DI blanks, demonstrating that Whirlpak bags leached on average 1.4 ± 0.1% (± 1 SE) of the initial DOC concentration of the filtered soil leachates. There was no significant difference in the DOC leached from DI-water filled Whirlpak bags exposed to sunlight compared to dark controls.

Following sunlight exposure, DOM was inoculated with microbes and both the light and dark treatments were incubated in the dark at 6-7 °C. The inoculum was comprised of a mixture of leachates from the three organic layer soil pits and had been passed through GF/C filters (Whatman GE Life Sciences, Freiburg, Germany). This community was assumed to be adapted to growth on leached soil DOM because it had grown in leached soil DOM for >48 hours. Once added to DOM, the inoculum was equivalent to 20% of total DOM leachate volume. After 4 h, one subsample from each replicate was filtered and preserved for DNA and RNA analysis (see below). After 5 d, subsamples were collected for DOM characterization with FT-ICR MS. Subsamples were also incubated separately over 5 d for respiration measurements (O_2 consumption and CO_2 production), taken at 0 and 5 d time-points for cell counts, and taken at 5 d for bacterial production measured by leucine incorporation (Ward *et al.*, 2017).

Metatranscriptome methods

Metatranscriptome sequences were generated from subsamples filtered onto 0.22-µm polyethersulfone (Supor) membrane filters (Pall Corp., New York, NY), preserved with RNA*later*™ RNA Stabilization Reagent (Qiagen, Hilden, Germany), and extracted and purified as described by Satinsky et al. (2015) with some modifications (details in Supplementary Information; SI). Ribosomal RNA removal, cDNA synthesis, and Illumina HiSeq sequencing were performed at the Joint Genome Institute (JGI) in Walnut Creek, CA, using either standard or low-input RNASeq protocols, both of which involve rRNA removal using the Ribo-Zero™ rRNA Removal Kit for Bacteria (Epicentre, Madison, WI), and cDNA library generation with Illumina Truseq Stranded RNA LT kit (Illumina, Inc., San Diego, CA) (details in SI).

RNA sequences (average 9.2 x 10^7 per sample; Table S1) were trimmed and quality-filtered with the BBDuk algorithm from BBMap v38.57 (Bushnell, 2015), and assembled using MEGAHIT (Li *et al.*, 2015) (details in SI). Coding sequences (CDS) were predicted with Prodigal (Hyatt *et al.*, 2010), and annotated to the Kyoto Encyclopedia of Genes and Genomes database (KEGG) (Kanehisa *et al.*, 2008) and a custom phylogenetic database (Annotations provided in Dataset S7), according to the JGI's Standard Operating Procedure (Huntemann *et al.*, 2015). Quality-controlled reads were mapped to CDS using BBMap, and SAMtools was used to extract counts, CDS lengths, and alignment lengths from BBMap output (Li *et al.*, 2009). On average, 82% of reads mapped to KEGG-annotated CDS sequences. Counts per annotation were normalized to transcripts per million (TPM) (Wagner *et al.*, 2012) to reduce biases associated with library size, CDS length, and read alignment length, and to express all counts as a portion of one million. Collectively, this workflow assembled metatranscriptomic reads into CDS and functionally annotated the translated CDS. Prior work has demonstrated that metatranscriptomic data can be directly annotated (versus mapping reads to whole genome sequences) with high accuracy, even in the case of partial transcripts, and that this accuracy improves as the length of the annotated sequence increases (Nayfach *et al.*, 2015).

Several KEGG gene categories were curated for analysis, including an aromatic degradation category, defined as the KEGG pathway for aromatic degradation; an oxygenase category, defined as KEGG Orthologs (KOs) associated with Enzyme Commission (EC) numbers 1.13 or 1.14; and a decarboxylase category, defined as KOs associated with EC number 4.1.1. Transcript abundances of genes within these categories were then normalized to total Metabolism expression (KEGG tier II category). Several ATP Binding Cassette (ABC) transporter categories were also created, including categories for the transport of sugars, polyols (i.e., alcohols), phosphorus, iron, other ions, and amino acids. The polyol category did not include transporters of phosphate-containing polyols; these transporters were instead included in the phosphorus

transporter category. Genes within these categories were normalized to total ABC Transporter expression (KEGG tier IV pathway). All curated categories are defined in Dataset S1.

Differential gene expression (DGE) between treatments was determined using the *exactTest* function within edgeR based on the TPM-normalized KO dataset, after setting *calcnormfactors* to "none" to avoid default TMM (trimmed mean of M values) normalization by edgeR. We selected this setting because edgeR was developed with the assumption that most genes are not differentially expressed in typical model organisms (Robinson and Oshlack, 2010), which may be questionable for highly diverse (Weiss *et al.*, 2017) and transcriptionally-responsive microbial communities.

DGE was determined for TPM-normalized datasets of all KOs, ABC transporter KOs, and Metabolism KOs, the final of which included the curated groups aromatic degradation KOs, oxygenase KOs, and decarboxylase KOs. A KO was considered differentially expressed if the false discovery rate (FDR; Benjamini and Hochberg, 1995) value of p was < 0.05. Datasets of all KOs were also compared at a basic level by summing TPM values within KEGG gene categories and using paired t-tests in R (R Core Development Team, 2011) to determine significant differences in gene expression between treatments ($\alpha = 0.05$). These analyses were conducted for the complete metatranscriptomic dataset, and for datasets binned into the major taxonomic groups Bacteroidetes, Betaproteobacteria, Gammaproteobacteria, and other taxa.

Bacterial community composition was determined with PCR amplicon sequencing of 16S ribosomal RNA genes using DNA collected and extracted as previously described (Crump *et al.*, 2003, 2013). PCR amplicon sequencing followed the Earth Microbiome Project protocol (http://www.earthmicrobiome.org/emp-standard-protocols/16s) (details in SI). Samples were sequenced at Oregon State University's Center for Genome Research and Biocomputing with Illumina MiSeq 2x150 bp paired-end reads. Amplicon sequences were analyzed using tools from the MOTHUR (v.1.32.1) (Schloss *et al.*, 2009), QIIME (Caporaso *et al.*, 2010), and USEARCH (v.7.0.1001_i86linux64) (Edgar, 2013) software packages (details in SI).

Microbial cell concentrations, respiration, and bacterial production were quantified as described in Ward and Cory (2015). Briefly, cell concentrations were quantified at 0 and 5 d using epifluorescence microscopy (glutaraldehyde-fixed samples; Crump *et al.*, 1998) (glutaraldehyde-fixed samples), and cells produced per day was calculated by subtracting initial cell concentrations from final cell concentrations, and then dividing this number by days of incubation. Respiration was measured over 5 d incubations as CO_2 production and O_2 consumption relative to killed controls (1% $HgCl_2$). Membrane inlet mass spectrometry was used to measure O_2, and a DIC analyzer (Apollo Sci Tech, LLC, Newark, DE) was used to measure CO_2. Bacterial production was determined on the fifth day of incubations by measuring ^{14}C-labeled L-leucine incorporation into cells in two subsamples and one TCA-killed control incubated for 2-4 h at 6 °C in the dark (Crump *et al.*, 2003).

16S rRNA gene amplicon sequences have been deposited in the NCBI Sequence Read Archive (SRA) under the bioproject accession number PRJNA356108 (https://www.ncbi.nlm.nih.gov). Metatranscriptome sequences and assembled contigs are publicly available via IMG under GOLD study ID Gs0114298 (https://img.jgi.doe.gov).

Acknowledgements

We thank J. Dobkowski, K. Harrold, M. Stuart, and researchers, technicians, and support staff of the Arctic LTER project and Toolik Lake Field Station for assistance with fieldwork. We also thank T. Glavina del Rio for conducting RNA sequence analyses and annotation at JGI (a

DOE Office of Science National Facility in Walnut Grove, CA) and K. Roscioli and S. Burton for assisting with mass spectrometry and NMR analyses at EMSL (a DOE Office of Science User Facility in Richland, WA). Additional assistance with DOM analyses came from L. Treibergs and A. Clinger and with RNA extraction methods from B. Satinsky. We are very grateful to A. Thurber, M. Graw, and other Oregon State faculty and graduate students for insightful conversations about data analysis and presentation of this research. Support for this work came from NSF grants DEB-0639790/1147378/1147336/1347042, DEB-1637459, DEB-1754835, OPP-1023270/1022876, DEB-1026843, and CAREER 1351745, and from the Camille and Henry Dreyfus Foundation Postdoctoral Program in Environmental Chemistry. DNA sequencing and chemical analyses were funded by the DOE JGI-EMSL Collaborative Science Initiative (CSP 1782).

References

Ames, G.F.L. (1986) Bacterial Periplasmic Transport Systems: Structure, Mechanism, and Evolution. *Annu Rev Biochem* **55**: 397–425.

Beier, S., Rivers, A.R., Moran, M.A., and Obernosterer, I. (2015) The transcriptional response of prokaryotes to phytoplankton-derived dissolved organic matter in seawater. *Environ Microbiol* **17**: 3466–3480.

Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* **57**: 289–300.

Bertilsson, S. and Tranvik, L.J. (2000) Photochemical transformation of dissolved organic matter in lakes. *Limnol Oceanogr* **45**: 753–762.

Bertilsson, S. and Tranvik, L.J. (1998) Photochemically produced carboxylic acids as substrates for freshwater bacterioplankton. *Limnol Oceanogr* **43**: 885–895.

Bowen, J.C., Kaplan, L.A., and Cory, R.M. (2019) Photodegradation disproportionately impacts biodegradation of semi-labile DOM in streams. *Limnol Oceanogr*.

Browning, D.F. and Busby, S.J.W. (2004) The regulation of bacterial transcription initiation. *Nat Rev Microbiol* **2**: 57–65.

Brzoska, P., Rimmele, M., Brzostek, K., and Boos, W. (1994) The pho regulon-dependent Ugp uptake system for glycerol-3-phosphate in Escherichia coli is trans inhibited by Pi. *J Bacteriol* **176**: 15–20.

Buchan, A., Collier, L.S., Neidle, E.L., and Moran, M.A. (2000) Key aromatic-ring-cleaving enzyme, protocatechuate 3,4-dioxygenase, in the ecologically important marine Roseobacter lineage. *Appl Environ Microbiol* **66**: 4662–4672.

Bugg, T.D.H., Ahmad, M., Hardiman, E.M., and Singh, R. (2011) The emerging role for bacteria in lignin degradation and bio-product formation. *Curr Opin Biotechnol* **22**: 394–400.

Bushnell, B. (2015) BBMap short-read aligner, and other bioinformatics tools.

Campanini, B., Pieroni, M., Raboni, S., Bettati, S., Benoni, R., Pecchini, C., et al. (2014) Inhibitors of the Sulfur Assimilation Pathway in Bacterial Pathogens as Enhancers of Antibiotic Therapy. *Curr Med Chem* **22**: 187–213.

Caplanne, S. and Laurion, I. (2008) Effect of chromophoric dissolved organic matter on epilimnetic stratification in lakes. *Aquat Sci* **70**: 123–133.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

714 Cavin, J.F., Dartois, V., and Diviès, C. (1998) Gene cloning, transcriptional analysis, purification, and characterization of phenolic acid decarboxylase from Bacillus subtilis.
716 *Appl Environ Microbiol* **64**: 1466–1471.
Cole, J.J., Prairie, Y.T., Caraco, N.F., McDowell, W.H., Tranvik, L.J., Striegl, R.G., et al. (2007)
718 Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems* **10**: 171–184.
720 Cornelis, P., Matthijs, S., and Van Oeffelen, L. (2009) Iron uptake regulation in Pseudomonas aeruginosa. *BioMetals* **22**: 15–22.
722 Cory, R.M., Crump, B.C., Dobkowski, J.A., and Kling, G.W. (2013) Surface exposure to sunlight stimulates CO2 release from permafrost soil carbon in the Arctic. *Proc Natl*
724 *Acad Sci U S A* **110**: 3429–3434.
Cory, R.M. and Kaplan, L.A. (2012) Biological lability of streamwater fluorescent dissolved
726 organic matter. *Limnol Oceanogr* **57**: 1347–1360.
Cory, R.M. and Kling, G.W. (2018) Interactions between sunlight and microorganisms influence
728 dissolved organic matter degradation along the aquatic continuum. *Limnol Oceanogr Lett* **3**: 102–116.
730 Cory, R.M., McKnight, D.M., Chin, Y.P., Miller, P., and Jaros, C.L. (2007) Chemical characteristics of fulvic acids from Arctic surface waters: Microbial contributions and
732 photochemical transformations. *J Geophys Res Biogeosciences* **112**: G04S51–G04S51.
Cory, R.M., McNeill, K., Cotner, J.B., Amado, A.M., Purcell, J.M., and Marshall, A.G. (2010)
734 Singlet oxygen in the coupled photo- and biochemical oxidation of dissolved organic matter. *Environ Sci Technol* **44**: 3683–3689.
736 Cory, R.M., Ward, C.P., Crump, B.C., and Kling, G.W. (2014) Sunlight controls water column processing of carbon in arctic fresh waters. *Science* **345**:.
738 Cotner, J.B. and Heath, R.T. (1990) Iron redox effects on photosensitive phosphorus release from dissolved humic materials. *Limnol Oceanogr* **35**: 1175–1181.
740 Cronan, Jr., J.E. and Rock, C.O. (2008) Biosynthesis of Membrane Lipids. *EcoSal Plus* **3**: 612–636.
742 Crump, B.C., Baross, J.A., and Simenstad, C.A. (1998) Dominance of particle-attached bacteria in the Columbia River estuary, USA. *Aquat Microb Ecol* **14**: 7–18.
744 Crump, B.C., Kling, G.W., Bahr, M., and Hobbie, J.E. (2003) Bacterioplankton community shifts in an Arctic lake correlate with seasonal changes in organic matter source. *Appl*
746 *Environ Microbiol* **69**: 2253–2268.
Dong, Y., Geng, J., Liu, J., Pang, M., Awan, F., Lu, C., and Liu, Y. (2019) Roles of three TonB
748 systems in the iron utilization and virulence of the Aeromonas hydrophila Chinese epidemic strain NJ-35. *Appl Microbiol Biotechnol* **103**: 4203–4215.
750 Edgar, R.C. (2013) UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**: 996–998.
752 Faust, B.C. and Zepp, R.G. (1993) Photochemistry of aqueous iron(III)- polycarboxylate complexes: Roles in the chemistry of atmosphere. *Env Sci Technol* **27**: 2517–2522.
754 Ferenci, T. (1999) Regulation by nutrient limitation Ferenci. *Curr Opin Microbiol* **2**: 208–213.
Franchini, A.G. and Egli, T. (2006) Global gene expression in Escherichia coli K-12 during
756 short-term and long-term adaptation to glucose-limited continuous culture conditions. *Microbiology* **152**: 2111–2127.
758 Fuchs, G., Boll, M., and Heider, J. (2011) Microbial degradation of aromatic compounds- From one strategy to four. *Nat Rev Microbiol* **9**: 803–816.

760 Fujii, M., Imaoka, A., Yoshimura, C., and Waite, T.D. (2014) Effects of Molecular Composition of Natural Organic Matter on Ferric Iron Complexation at Circumneutral pH. *Environ Sci* 762 *Technol* **48**: 4414–4424.
Gareis, J.A.L. and Lesack, L.F.W. (2018) Photodegraded dissolved organic matter from peak 764 freshet river discharge as a substrate for bacterial production in a lake-rich great Arctic delta. *Arct Sci* **4**: 557–583.
766 Gareis, J.A.L., Lesack, L.F.W., and Bothwell, M.L. (2010) Attenuation of in situ UV radiation in Mackenzie Delta lakes with varying dissolved organic matter compositions. *Water* 768 *Resour Res* **46**: 1–14.
Gifford, S.M., Becker, J.W., Sosa, O.A., Repeta, D.J., and DeLong, E.F. (2016) Quantitative 770 transcriptomics reveals the growth- and nutrient- dependent response of a streamlined marine methylotroph to methanol and naturally occurring dissolved organic matter. *mBio* 772 **7**: 1–15.
Gomez-Saez, G.V., Pohlabeln, A.M., Stubbins, A., Marsay, C.M., and Dittmar, T. (2017) 774 Photochemical Alteration of Dissolved Organic Sulfur from Sulfidic Porewater. *Environ Sci Technol* **51**: 14144–14154.
776 Gonsior, M., Hertkorn, N., Conte, M.H., Cooper, W.J., Bastviken, D., Druffel, E., and Schmitt-Kopplin, P. (2014) Photochemical production of polyols arising from significant photo- 778 transformation of dissolved organic matter in the oligotrophic surface ocean. *Mar Chem* **163**: 10–18.
780 Gonsior, M., Peake, B.M., Cooper, W.T., Podgorski, D., D'Andrilli, J., and Cooper, W.J. (2009) Photochemically Induced Changes in Dissolved Organic Matter Identified by Ultrahigh 782 Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environ Sci Technol* **43**: 698–703.
784 Gulvik, C.A. and Buchan, A. (2013) Simultaneous Catabolism of Plant-Derived Aromatic Compounds Results in Enhanced Growth for Members of the Roseobacter Lineage. *Appl* 786 *Environ Microbiol* **79**: 3716–3723.
Harke, M.J. and Gobler, C.J. (2013) Global Transcriptional Responses of the Toxic 788 Cyanobacterium, Microcystis aeruginosa, to Nitrogen Stress, Phosphorus Stress, and Growth on Organic Matter. *PLoS ONE* **8**: e69834.
790 Herlemann, D.P.R., Manecki, M., Meeske, C., Pollehne, F., Labrenz, M., Schulz-Bull, D., et al. (2014) Uncoupling of bacterial and terrigenous dissolved organic matter dynamics in 792 decomposition experiments. *PLoS ONE* **9**: e93945.
Hockaday, W.C., Purcell, J.M., Marshall, A.G., Baldock, J.A., and Hatcher, P.G. (2009) 794 Electrospray and photoionization mass spectrometry for the characterization of organic matter in natural waters: a qualitative assessment: Characterization of organic matter in 796 natural waters. *Limnol Oceanogr Methods* **7**: 81–95.
Huntemann, M., Ivanova, N.N., Mavromatis, K., James Tripp, H., Paez-Espino, D., Palaniappan, 798 K., et al. (2015) The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v.4). *Stand Genomic Sci* **10**: 1–6.
800 Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC* 802 *Bioinformatics* **11**: 119.
Judd, K.E., Crump, B.C., and Kling, G.W. (2007) Bacterial responses in activity and community 804 composition to photo-oxidation of dissolved organic matter from soil and surface waters. *Aquat Sci* **69**: 96–107.

Judd, K.E., Crump, B.C., and Kling, G.W. (2006) Variation in dissolved organic matter controls bacterial production and community composition. *Ecology* **87**: 2068–2079.

Kaiser, E. and Sulzberger, B. (2004) Phototransformation of riverine dissolved organic matter (DOM) in the presence of abundant Iron: Effect on DOM bioavailability. *Limnol Oceanogr* **49**: 540–554.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., et al. (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Res* **36**: D480–D484.

Kieber, D.J. and Mopper, K. (1987) Photochemical formation of glyoxylic and pyruvic acids in seawater. *Mar Chem* **21**: 135–149.

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

Kraakman, L.S., Griffioen, G., Zerp, S., Groeneveld, P., Thevelein, J.M., Mager, W.H., and Planta, R.J. (1993) Growth-related expression of ribosomal protein genes in Saccharomyces cerevisiae. *Mol Gen Genet MGG* **239**: 196–204.

Lahtvee, P.J., Adamberg, K., Arike, L., Nahku, R., Aller, K., and Vilu, R. (2011) Multi-omics approach to study the growth efficiency and amino acid metabolism in Lactococcus lactis at various specific growth rates. *Microb Cell Factories* **10**: 1–12.

León-Sobrino, C., Ramond, J.-B., Maggs-Kölling, G., and Cowan, D.A. (2019) Nutrient Acquisition, Rather Than Stress Response Over Diel Cycles, Drives Microbial Transcription in a Hyper-Arid Namib Desert Soil. *Front Microbiol* **10**:.

Li, D., Liu, C.M., Luo, R., Sadakane, K., and Lam, T.W. (2015) MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Logue, J.B., Stedmon, C.A., Kellerman, A.M., Nielsen, N.J., Andersson, A.F., Laudon, H., et al. (2016) Experimental insights into the importance of aquatic bacterial community composition to the degradation of dissolved organic matter. *ISME J* **10**: 533–545.

Madar, D., Dekel, E., Bren, A., Zimmer, A., Porat, Z., and Alon, U. (2013) Promoter activity dynamics in the lag phase of Escherichia coli. *BMC Syst Biol* **7**: 136.

Mann, P.J., Davydova, A., Zimov, N., Spencer, R.G.M., Davydov, S., Bulygina, E., et al. (2012) Controls on the composition and lability of dissolved organic matter in Siberia's Kolyma River basin. *J Geophys Res Biogeosciences* **117**: 1–15.

Matsumoto, Y., Murakami, Y., Tsuru, S., Ying, B., and Yomo, T. (2013) Growth rate-coordinated transcriptome reorganization in bacteria. *BMC Genomics* **14**: 808.

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.

de Menezes, A., Clipson, N., and Doyle, E. (2012) Comparative metatranscriptomics reveals widespread community responses during phenanthrene degradation in soil. *Environ Microbiol* **14**: 2577–2588.

Moran, M.A. and Zepp, R.G. (1997) Role of photoreduction in the formation of biogically labile compounds from dossolved organic matter. **42**: 1307–1316.

Nayfach, S., Bradley, P.H., Wyman, S.K., Laurent, T.J., Williams, A., Eisen, J.A., et al. (2015) Automated and Accurate Estimation of Gene Family Abundance from Shotgun Metagenomes. *PLoS Comput Biol* **11**: 1–29.

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.

Nomura, M., Gourse, R., and Baughman, G. (1984) Regulation of the synthesis of ribosomes and ribosomal components. *Annu Rev Biochem* **53**: 75–117.

Ossola, R., Tolu, J., Clerc, B., Erickson, P.R., Winkel, L.H.E., and McNeill, K. (2019) Photochemical Production of Sulfate and Methanesulfonic Acid from Dissolved Organic Sulfur. *Environ Sci Technol* **53**: 13191–13200.

Page, S.E., Logan, J.R., Cory, R.M., and McNeill, K. (2014) Evidence for dissolved organic matter as the primary source and sink of photochemically produced hydroxyl radical in arctic surface waters. *Env Sci Process Impacts* **16**: 807–822.

Piłsyk, S. and Paszewski, A. (2009) Sulfate permeasesphylogenetic diversity of sulfate transport. *Acta Biochim Pol* **56**:.

R Core Development Team (2011) R: A Language and Environment for Statistical Computing, Vienna, Austria.

Raymond, P.A., Hartmann, J., Lauerwald, R., Sobek, S., McDonald, C., Hoover, M., et al. (2013) Global carbon dioxide emissions from inland waters. *Nature* **503**: 355–359.

Reader, H.E. and Miller, W.L. (2014) The efficiency and spectral photon dose dependence of photochemically induced changes to the bioavailability of dissolved organic carbon. *Limnol Oceanogr* **59**: 182–194.

Ritchie, J.D. and Perdue, E.M. (2003) Proton-binding study of standard and reference fulvic acids, humic acids, and natural organic matter. *Geochim Cosmochim Acta* **67**: 85–96.

Robinson, M.D. and Oshlack, A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**: R25.

Rolfe, M.D., Beek, A.T., Graham, A.I., Trotter, E.W., Asif, H.M.S., Sanguinetti, G., et al. (2011) Transcript profiling and inference of Escherichia coli K-12 ArcA activity across the range of physiologically relevant oxygen concentrations. *J Biol Chem* **286**: 10147–10154.

Rowland, J.C., Jones, C.E., Altmann, G., Bryan, R., Crosby, B.T., Geernaert, G.L., et al. (2010) Arctic Landscapes in Transition: Responses to Thawing Permafrost. *Eos Trans Am Geophys Union* **91**: 229–236.

Satinsky, B.M., Fortunato, C.S., Doherty, M., Smith, C.B., Sharma, S., Ward, N.D., et al. (2015) Metagenomic and metatranscriptomic inventories of the lower Amazon River, May 2011. *Microbiome* **3**: 39.

Satinsky, B.M., Smith, C.B., Sharma, S., Ward, N.D., Krusche, A.V., Richey, J.E., et al. (2017) Patterns of bacterial and archaeal gene expression through the lower Amazon River. *Front Mar Sci* **4**: 253.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Scott, M., Mateescu, E.M., Zhang, Z., and Hwa, T. (2010) Interdependence of Cell Growth Origins and Consequences. *Science* **330**: 1099–1102.

Shi, Y., McCarren, J., and Delong, E.F. (2012) Transcriptional responses of surface water marine microbial assemblages to deep-sea water amendment. *Environ Microbiol* **14**: 191–206.

Sleighter, R.L., Cory, R.M., Kaplan, L.A., Abdulla, H.A.N., and Hatcher, P.G. (2014) A coupled geochemical and biogeochemical approach to characterize the bioreactivity of dissolved organic matter from a headwater stream. *J Geophys Res G Biogeosciences* **119**: 1520–1537.

Šmejkalová, T., Edwards, M.E., and Dash, J. (2016) Arctic lakes show strong decadal trend in earlier spring ice-out. *Sci Rep* **6**:.

Soutourina, O.A. and Bertin, P.N. (2003) Regulation cascade of flagellar expression in Gram-negative bacteria. *FEMS Microbiol Rev* **27**: 505–523.

Stenson, A.C., Marshall, A.G., and Cooper, W.T. (2003) Exact Masses and Chemical Formulas of Individual Suwannee River Fulvic Acids from Ultrahigh Resolution Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectra. *Anal Chem* **75**: 1275–1284.

Strome, D.J. and Miller, M.C. (1978) Photolytic changes in dissolved humic substances. *SIL Proc 1922-2010* **20**: 1248–1254.

Stubbins, A., Mann, P.J., Powers, L., Bittar, T.B., Dittmar, T., McIntyre, C.P., et al. (2017) Low photolability of yedoma permafrost dissolved organic carbon. *J Geophys Res Biogeosciences* **122**: 200–211.

Stubbins, A., Spencer, R.G.M., Chen, H., Hatcher, P.G., Mopper, K., Hernes, P.J., et al. (2010) Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. *Limnol Oceanogr* **55**: 1467–1477.

Tate, K.R. and Newman, R.H. (1982) Phosphorus fractions of a climosequence of soils in New zealand tussock grassland. *Soil Biol Biochem* **14**: 191–196.

Tranvik, L.J. and Bertilsson, S. (2001) Contrasting effects of solar UV radiation on dissolved organic sources for bacterial growth. *Ecol Lett* **4**: 458–463.

Turner, B.L., Baxter, R., Mahieu, N., Sjögersten, S., and Whitton, B.A. (2004) Phosphorus compounds in subarctic Fennoscandian soils at the mountain birch (Betula pubescens)—tundra ecotone. *Soil Biol Biochem* **36**: 815–823.

Vallino, J.J., Hopkinson, C.S., and Hobbie, J.E. (1996) Modeling bacterial utilization of dissolved organic matter: Optimization replaces Monod growth kinetics. *Limnol Oceanogr* **41**: 1591–1609.

Vershinina, O.A. and Znamenskaya, L.V. (2002) The Pho Regulons of Bacteria. **71**: 15.

Voelker, B.M., Morel, F.M.M., and Sulzberger, B. (1997) Iron redox cycling in surface waters: effects of humic substanes and light. **31**: 1004–1011.

Vonk, J.E. and Gustafsson, Ö. (2013) Permafrost-carbon complexities. *Nat Geosci* **6**: 675–676.

Wagner, G.P., Kin, K., and Lynch, V.J. (2012) Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci* **131**: 281–285.

Ward, C.P. and Cory, R.M. (2020) Assessing the prevalence, products, and pathways of dissolved organic matter partial photo-oxidation in arctic surface waters. *Environ Sci Process Impacts*.

Ward, C.P. and Cory, R.M. (2015) Chemical composition of dissolved organic matter draining permafrost soils. *Geochim Cosmochim Acta* **167**: 63–79.

Ward, C.P. and Cory, R.M. (2016) Complete and Partial Photo-oxidation of Dissolved Organic Matter Draining Permafrost Soils. *Environ Sci Technol* **50**: 3545–3553.

940 Ward, C.P., Nalven, S.G., Crump, B.C., Kling, G.W., and Cory, R.M. (2017) Photochemical alteration of organic carbon draining permafrost soils shifts microbial metabolic
942 pathways and stimulates respiration. *Nat Commun* **8**: 772.
Ward, N.D., Keil, R.G., Medeiros, P.M., Brito, D.C., Cunha, A.C., Dittmar, T., et al. (2013)
944 Degradation of terrestrially derived macromolecules in the Amazon River. *Nat Geosci* **6**: 530–533.
946 Weiss, S., Xu, Z.Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., et al. (2017) Normalization and microbial differential abundance strategies depend upon data
948 characteristics. *Microbiome* **5**:.
Wetzel, R.G., Hatcher, P.G., and Bianchi, T.S. (1995) Natural photolysis by ultraviolet
950 irradiance of recalcitrant dissolved organic matter to simple substrates for rapid bacterial metabolism. *Limnol Ocean* **40**: 1369–1380.
952 Wyckoff, E.E., Mey, A.R., Leimbach, A., Fisher, C.F., and Payne, S.M. (2006) Characterization of Ferric and Ferrous Iron Transport Systems in Vibrio cholerae. *J Bacteriol* **188**: 6515–
954 6523.
Xie, H., Zafiriou, O.C., Cai, W.J., Zepp, R.G., and Wang, Y. (2004) Photooxidation and its
956 effects on the carboxyl content of dissolved organic matter in two coastal rivers in the southeastern United States. *Environ Sci Technol* **38**: 4113–4119.
958 Zhang, Y., Gao, J., Wang, L., Liu, S., Bai, Z., Zhuang, X., and Zhuang, G. (2018) Environmental Adaptability and Quorum Sensing: Iron Uptake Regulation during Biofilm Formation by
960 *Paracoccus denitrificans*. *Appl Environ Microbiol* **84**:.

962 **Figure Legends**

Figure 1. Characterization of gene expression and taxa conducting expression in the two
964 treatments. (A) MA-plot of the $\log_2$ fold-changes (logFC, y-axis) in the light and dark treatments versus the mean of normalized counts (transcripts per million; TPM) of each KEGG Ortholog
966 (KO; i.e., gene; gray symbols). Colored symbols represent differentially expressed KOs (FDR < 0.05) with greater expression in the dark treatment (blue symbols) and the light treatment
968 (orange symbols). The horizontal lines (black) indicate when absolute $\log_2$FC values are ≥ 1 between treatments. (B) Relative taxonomic composition of the whole community (16S
970 amplicons) compared to the active community (metatranscriptomes; metaT) in light and dark treatments.
972

Figure 2. Average chemical characteristics of DOM formulas that significantly increased or
974 decreased in abundance at the 95% confidence interval after exposure to sunlight or incubation with microbes. Cross-comparison between the list of formulas in each category yielded the data
976 in the Venn diagram (A) and the table (B). The table provides average molecular weight (MW); molar ratios of oxygen to carbon (O/C) and hydrogen to carbon (H/C) for each group of
978 formulas; percent aromatic, tannin-like or lignin-like formulas; and the percent of formulas that contained N or S.
980

Figure 3. Expression (summed transcripts per million) of KEGG tier II (A) and tier III (B-E)
982 categories in the dark treatment (blue) and light treatment (orange). Asterisks (*) represent significant differences according to paired t-tests ($p \leq 0.05$). Error bars indicate standard error of
984 the mean.

Figure 4. Heatmap of differential gene expression comparing light and dark treatments. The heatmap portrays standard deviations from the mean (Z-scores) of transcript abundances for each of the three replicates of all differentially expressed (A) DOM metabolism KOs (i.e., genes) in aromatic degradation, oxygenase, and decarboxylase categories, and (B) ABC transporter KOs in sugar, polyol, phosphorus, sulfur, iron, other ions, and amino acid categories. Transcript abundances were calculated as percentages of total expression within the KEGG categories Metabolism (A) and ABC Transporters (B), which are KEGG tier II and IV categories, respectively, and then converted to Z-scores. An asterisk (*) indicates KOs that are shown twice because they fall into more than one gene category.

Figure 5. Diagram of selected functions of genes that were more expressed by the microbial communities in the light treatment (top) and dark treatment (bottom). Differences between the two diagrams describe shifts in gene expression that resulted from photo-alteration of dissolved organic matter in the light treatment.

Figure 6. Microbial activity in the dark treatment (blue) and light treatment (orange), as measured by respiration (O_2 consumption and CO_2 production), new cell production measured over the 5-day incubation, and bacterial (biomass) production rate measured at the end of 5 days. Error bars represent standard error of the mean. Measurements were lower in the light treatment, but the differences between treatments for each measurement were not statistically significant (paired t-tests $p > 0.05$). Raw data for this figure were published previously in Ward *et al.*, (2017).

Figure 1

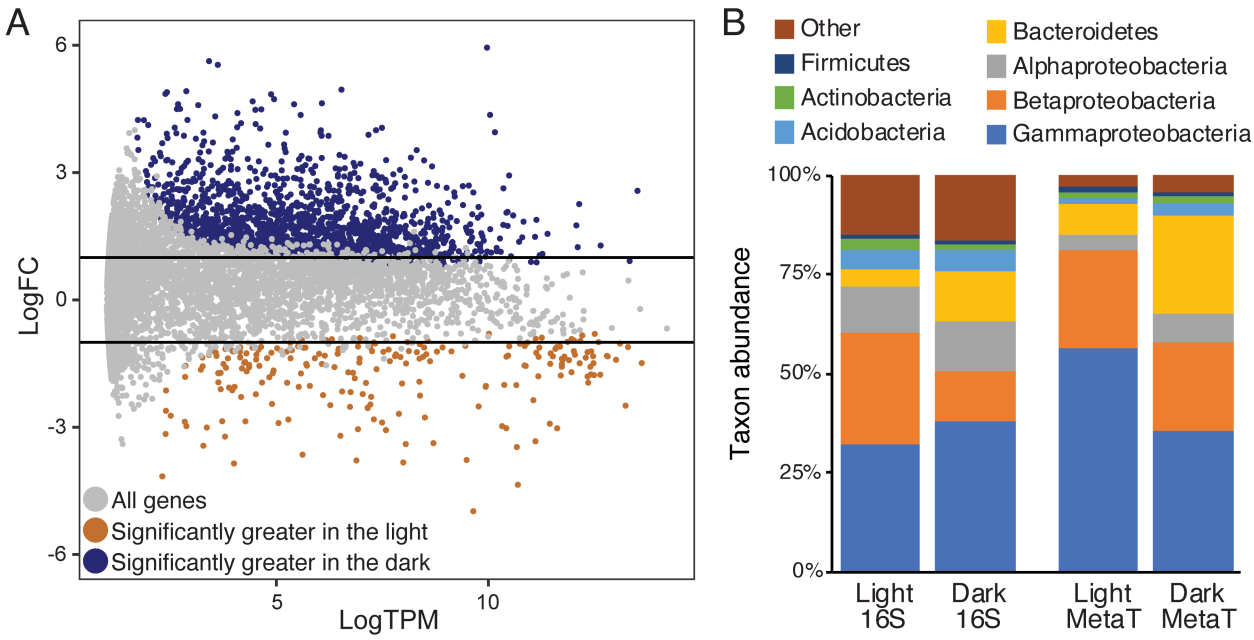
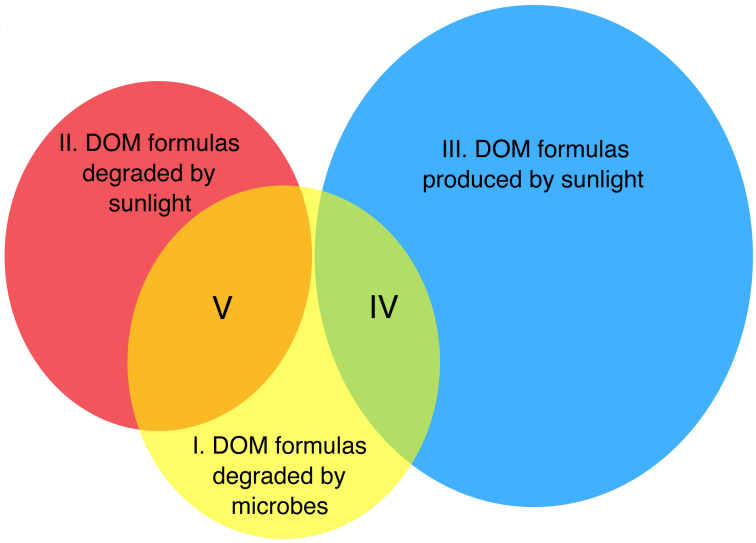

A
6
3
0
-3
-6
LogFC
5
10
LogTPM
All genes
Significantly greater in the light
Significantly greater in the dark
B
Other
Firmicutes
Actinobacteria
Acidobacteria
Bacteroidetes
Alphaproteobacteria
Betaproteobacteria
Gammaproteobacteria
100%
75%
50%
25%
0%
Taxon abundance
Light 16S
Dark 16S
Light MetaT
Dark MetaT

Figure 2

A



B

Region	Description	# of Formulas	MW	O/C	H/C	Aromatic (%)	Tannin-like (%)	Lignin-like (%)	N (%)	S (%)
I	Degraded by Microbes	383	539	0.61	0.90	39	64	27	3	7
II	Degraded by Sunlight	375	566	0.64	0.88	20	74	21	4	1
III	Produced by Sunlight	784	460	0.52	0.96	23	24	71	3	3
IV	Produced by Sunlight & Degraded by Microbes	98	454	0.52	0.99	34	38	51	4	9
V	Degraded by Sunlight & Microbes	148	570	0.70	0.80	43	90	7	2	1

Figure 3

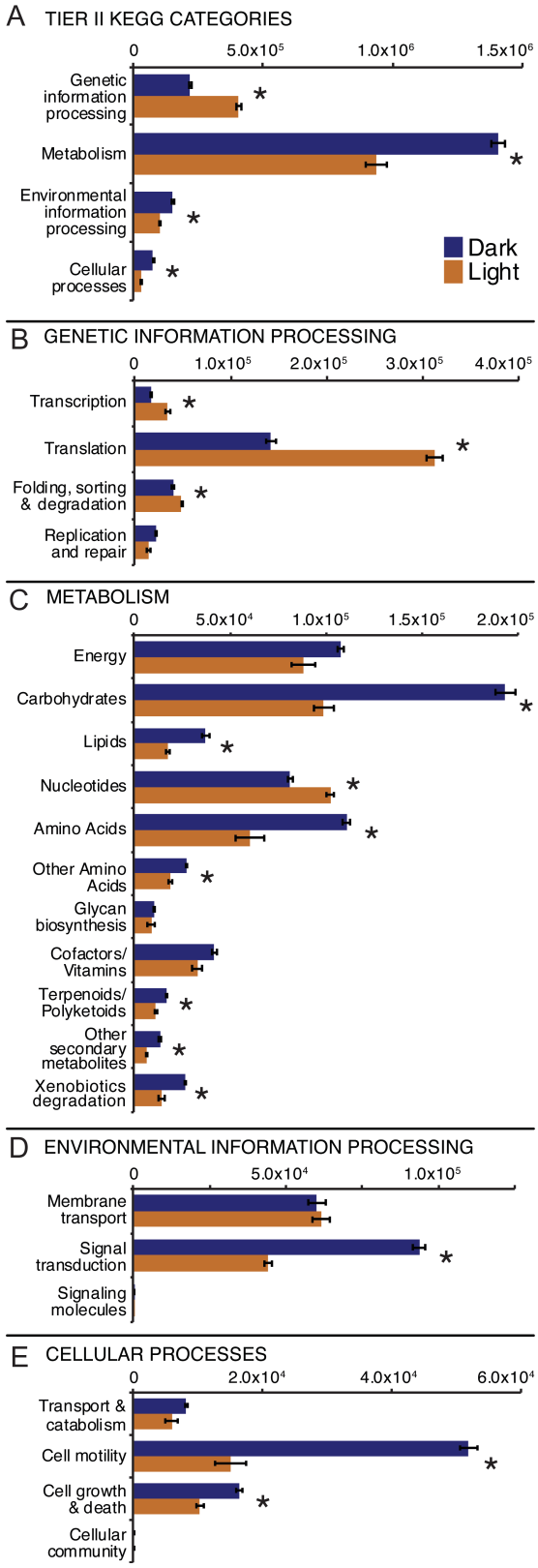

A TIER II KEGG CATEGORIES
0
5.0x10^5
1.0x10^6
1.5x10^6
Genetic information processing
Metabolism
Environmental information processing
Cellular processes
Dark
Light
B GENETIC INFORMATION PROCESSING
0
1.0x10^5
2.0x10^5
3.0x10^5
4.0x10^5
Transcription
Translation
Folding, sorting & degradation
Replication and repair
C METABOLISM
0
5.0x10^4
1.0x10^5
1.5x10^5
2.0x10^5
Energy
Carbohydrates
Lipids
Nucleotides
Amino Acids
Other Amino Acids
Glycan biosynthesis
Cofactors/ Vitamins
Terpenoids/ Polyketoids
Other secondary metabolites
Xenobiotics degradation
D ENVIRONMENTAL INFORMATION PROCESSING
0
5.0x10^4
1.0x10^5
Membrane transport
Signal transduction
Signaling molecules
E CELLULAR PROCESSES
0
2.0x10^4
4.0x10^4
6.0x10^4
Transport & catabolism
Cell motility
Cell growth & death
Cellular community

Figure 4

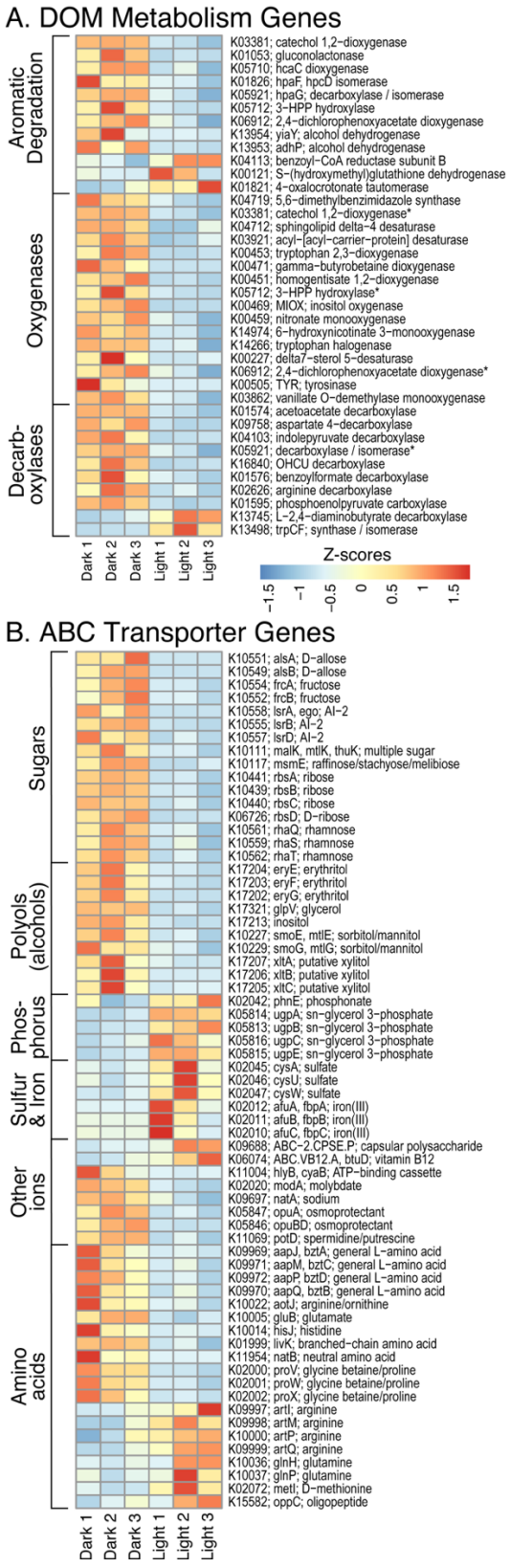

A. DOM Metabolism Genes
Aromatic Degradation
Oxygenases
Decarb-oxylases
K03381; catechol 1,2−dioxygenase
K01053; gluconolactonase
K05710; hcaC dioxygenase
K01826; hpaF, hpcD isomerase
K05921; hpaG; decarboxylase / isomerase
K05712; 3−HPP hydroxylase
K06912; 2,4−dichlorophenoxyacetate dioxygenase
K13954; yiaY; alcohol dehydrogenase
K13953; adhP; alcohol dehydrogenase
K04113; benzoyl−CoA reductase subunit B
K00121; S−(hydroxymethyl)glutathione dehydrogenase
K01821; 4−oxalocrotonate tautomerase
K04719; 5,6−dimethylbenzimidazole synthase
K03381; catechol 1,2−dioxygenase*
K04712; sphingolipid delta−4 desaturase
K03921; acyl−[acyl−carrier−protein] desaturase
K00453; tryptophan 2,3−dioxygenase
K00471; gamma−butyrobetaine dioxygenase
K00451; homogentisate 1,2−dioxygenase
K05712; 3−HPP hydroxylase*
K00469; MIOX; inositol oxygenase
K00459; nitronate monooxygenase
K14974; 6−hydroxynicotinate 3−monooxygenase
K14266; tryptophan halogenase
K00227; delta7−sterol 5−desaturase
K06912; 2,4−dichlorophenoxyacetate dioxygenase*
K00505; TYR; tyrosinase
K03862; vanillate O−demethylase monooxygenase
K01574; acetoacetate decarboxylase
K09758; aspartate 4−decarboxylase
K04103; indolepyruvate decarboxylase
K05921; decarboxylase / isomerase*
K16840; OHCU decarboxylase
K01576; benzoylformate decarboxylase
K02626; arginine decarboxylase
K01595; phosphoenolpyruvate carboxylase
K13745; L−2,4−diaminobutyrate decarboxylase
K13498; trpCF; synthase / isomerase
Dark 1
Dark 2
Dark 3
Light 1
Light 2
Light 3
Z-scores
−1.5
−1
−0.5
0
0.5
1
1.5
B. ABC Transporter Genes
Sugars
Polyols (alcohols)
Phos-phorus
Sulfur & Iron
Other ions
Amino acids
K10551; alsA; D−allose
K10549; alsB; D−allose
K10554; frcA; fructose
K10552; frcB; fructose
K10558; lsrA, ego; AI−2
K10555; lsrB; AI−2
K10557; lsrD; AI−2
K10111; malK, mtlK, thuK; multiple sugar
K10117; msmE; raffinose/stachyose/melibiose
K10441; rbsA; ribose
K10439; rbsB; ribose
K10440; rbsC; ribose
K06726; rbsD; D−ribose
K10561; rhaQ; rhamnose
K10559; rhaS; rhamnose
K10562; rhaT; rhamnose
K17204; eryE; erythritol
K17203; eryF; erythritol
K17202; eryG; erythritol
K17321; glpV; glycerol
K17213; inositol
K10227; smoE, mtlE; sorbitol/mannitol
K10229; smoG, mtlG; sorbitol/mannitol
K17207; xltA; putative xylitol
K17206; xltB; putative xylitol
K17205; xltC; putative xylitol
K02042; phnE; phosphonate
K05814; ugpA; sn−glycerol 3−phosphate
K05813; ugpB; sn−glycerol 3−phosphate
K05816; ugpC; sn−glycerol 3−phosphate
K05815; ugpE; sn−glycerol 3−phosphate
K02045; cysA; sulfate
K02046; cysU; sulfate
K02047; cysW; sulfate
K02012; afuA, fbpA; iron(III)
K02011; afuB, fbpB; iron(III)
K02010; afuC, fbpC; iron(III)
K09688; ABC−2.CPSE.P; capsular polysaccharide
K06074; ABC.VB12.A, btuD; vitamin B12
K11004; hlyB, cyaB; ATP−binding cassette
K02020; modA; molybdate
K09697; natA; sodium
K05847; opuA; osmoprotectant
K05846; opuBD; osmoprotectant
K11069; potD; spermidine/putrescine
K09969; aapJ, bztA; general L−amino acid
K09971; aapM, bztC; general L−amino acid
K09972; aapP, bztD; general L−amino acid
K09970; aapQ, bztB; general L−amino acid
K10022; aotJ; arginine/ornithine
K10005; gluB; glutamate
K10014; hisJ; histidine
K01999; livK; branched−chain amino acid
K11954; natB; neutral amino acid
K02000; proV; glycine betaine/proline
K02001; proW; glycine betaine/proline
K02002; proX; glycine betaine/proline
K09997; artI; arginine
K09998; artM; arginine
K10000; artP; arginine
K09999; artQ; arginine
K10036; glnH; glutamine
K10037; glnP; glutamine
K02072; metI; D−methionine
K15582; oppC; oligopeptide
Dark 1
Dark 2
Dark 3
Light 1
Light 2
Light 3

Figure 5

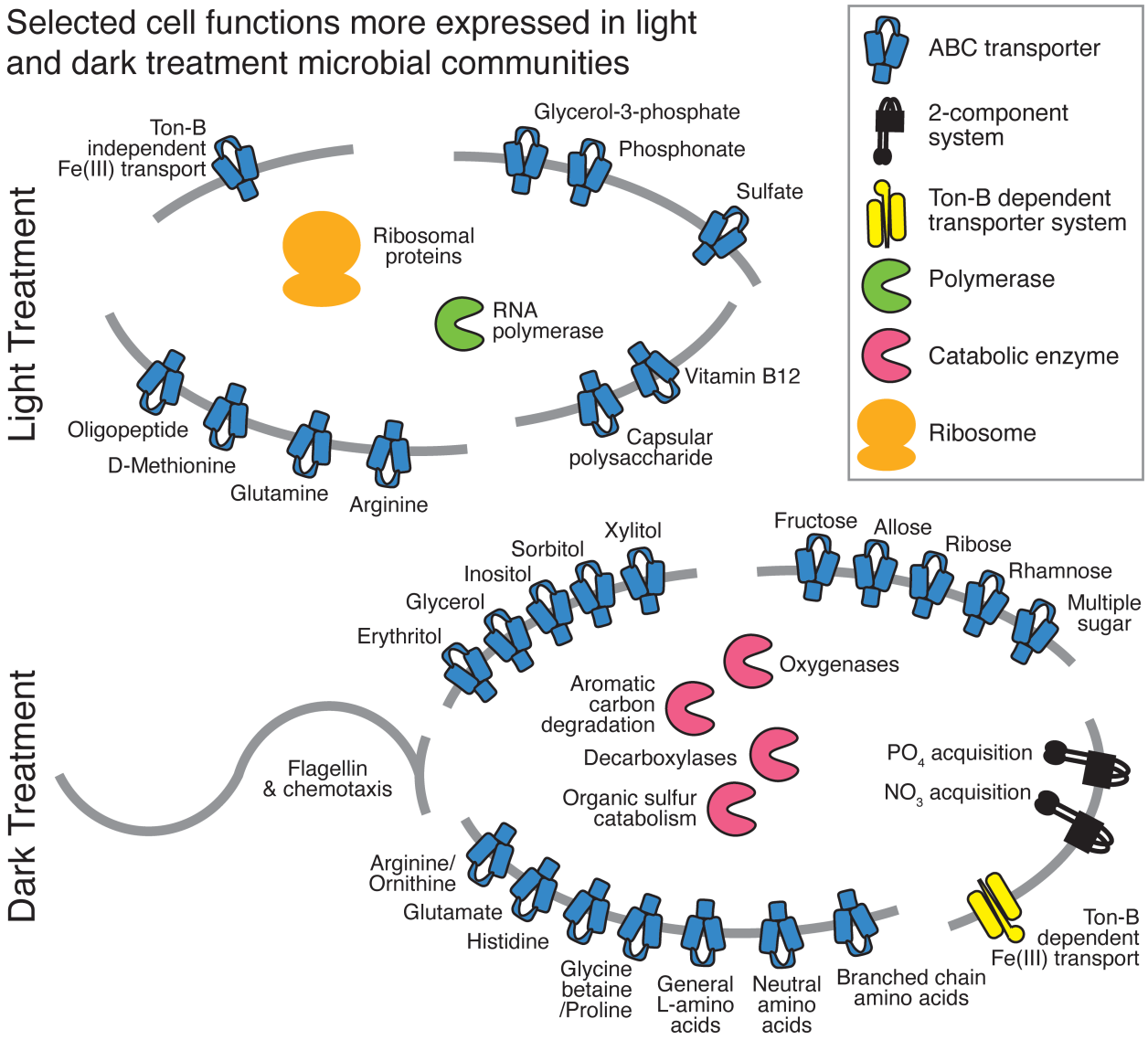
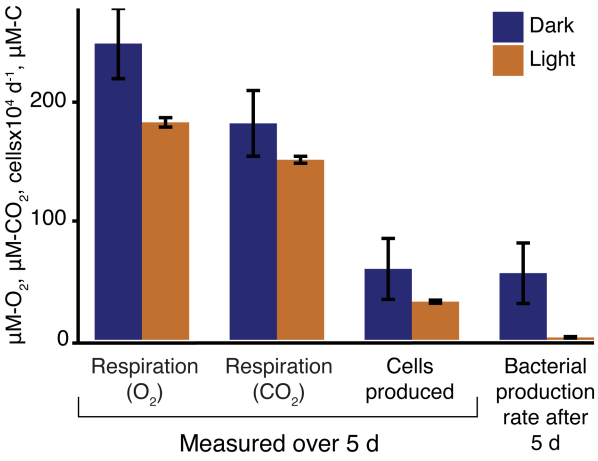

Selected cell functions more expressed in light and dark treatment microbial communities
Light Treatment
Dark Treatment
Ton-B independent Fe(III) transport
Glycerol-3-phosphate
Phosphonate
Sulfate
Ribosomal proteins
RNA polymerase
Vitamin B12
Capsular polysaccharide
Oligopeptide
D-Methionine
Glutamine
Arginine
Xylitol
Sorbitol
Inositol
Glycerol
Erythritol
Fructose
Allose
Ribose
Rhamnose
Multiple sugar
Oxygenases
Aromatic carbon degradation
Decarboxylases
Organic sulfur catabolism
Flagellin & chemotaxis
PO4 acquisition
NO3 acquisition
Arginine/ Ornithine
Glutamate
Histidine
Glycine betaine /Proline
General L-amino acids
Neutral amino acids
Branched chain amino acids
Ton-B dependent Fe(III) transport
ABC transporter
2-component system
Ton-B dependent transporter system
Polymerase
Catabolic enzyme
Ribosome

Figure 6


Dark
Light
µM-O2, µM-CO2, cellsx10^4 d^-1, µM-C
200
100
0
Respiration (O2)
Respiration (CO2)
Cells produced
Bacterial production rate after 5 d
Measured over 5 d

Supplementary Information for

Experimental metatranscriptomics reveals costs and benefits of DOM photo-alteration for freshwater microbes in the Arctic.

Sarah G. Nalven[1], Collin P. Ward[2], Jérôme P. Payet[1], Rose M. Cory[3], George W. Kling[3], Thomas J. Sharpton[1], Christopher M. Sullivan[1], Byron C. Crump[1*]

Affiliations
[1] Oregon State University
[2] Woods Hole Oceanographic Institution
[3] University of Michigan

*Corresponding Author
College of Earth, Ocean, and Atmospheric Sciences
Oregon State University
104 CEOAS Admin Bldg.
Corvallis, OR 97331-5503
Phone: 541-737-4369
Byron.Crump@oregonstate.edu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Supplementary Methods

RNA extraction. RNA samples were filtered onto 0.22-μm polyethersulfone (Supor) membrane filters (Pall Corp.) and preserved with RNAlater™ RNA Stabilization Reagent (Qiagen). RNA extraction and DNA removal was carried out as described by Satinsky et al. (Satinsky *et al.*, 2015) with some modifications. This protocol modifies the RNeasy mini kit protocol (Qiagen) for RNA extraction and uses the TURBO DNA-free kit (Ambion) for DNA removal. Filters were removed from RNAlater, sliced into small pieces with a sterile scalpel, and combined with 2 mL sterilized zirconium beads, 10 mL Ambion denaturation solution, and 500 μl of Plant RNA Isolation Aid (Ambion) in a 50 mL conical tube. Tubes were agitated on a vortex adapter for 10 minutes on high and centrifuged for 1 minute at 3500 x g. Lysate was transferred to a 15 mL conical tube and centrifuged for 5 minutes at 3500 x g. Lysate was then washed twice with 3.5 mL of saturated phenol (pH 4.3), and once with 5 mL chloroform:isoamyl alcohol (24:1) in 50 mL conical tubes. Lysate was then combined with an equal volume of 100% ethanol, homogenized by passing it through a syringe needle 3-4 times, and then passed through a Direct-Zol RNA Kit column (Zymo Research) using a vacuum manifold. RNA was purified and resuspended following the Direct-Zol RNA Kit (Zymo Research) manufacturer's instructions and eluted twice with 50 μL nuclease-free water. DNA was removed using the TURBO DNA-free Kit (Ambion).

rRNA removal, cDNA synthesis and Illumina HiSeq sequencing. Ribosomal RNA removal, cDNA synthesis and Illumina HiSeq sequencing were performed at the Joint Genome Institute in Walnut Creek, CA, using either standard or low-input RNASeq protocols. In both protocols, rRNA was removed from 1 μg or 100 ng of total RNA (regular vs. low-input protocol, respectively) using the Ribo-Zero™ rRNA Removal Kit for Bacteria (Epicentre). Stranded cDNA libraries were generated using the Truseq Stranded RNA LT kit (Illumina). Resulting rRNA depleted-RNA was fragmented and reverse transcribed using random hexamers and SSII (Invitrogen) followed by second strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 8 or 10 cycles of PCR (regular vs. low-input protocol, respectively). The quantified libraries were then multiplexed into pools of 4 or 3 libraries (regular vs. low-input protocol, respectively), and the pool was then prepared for sequencing on the Illumina HiSeq sequencing platform using a TruSeq paired-end cluster kit, v4, and Illumina's cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flowcell was performed on the Illumina HiSeq2500 sequencer using HiSeq TruSeq SBS sequencing kits, v4, following a 2x150 indexed run recipe.

RNA sequence trimming and quality control. RNA sequences were trimmed, quality-controlled, and assembled by the Joint Genome Institute's assembly team. Raw reads were first quality-trimmed to Q10 and adapter-trimmed using BBDuk (Bushnell, 2015; options: ktrim=r, k=25, mink=12, tpe=t, tbo=t, qtrim=r, trimq=10, maq=10, maxns=3, minlen=50). Reads were then filtered for process artifacts using BBDuk (options: k=16). Ribosomal RNA reads were removed with BBMap (Bushnell, 2015) by mapping against a trimmed version of the Silva 119 database (options: fast=t minid=0.90 local=t). BBMap was also used to remove human reads. Metatranscriptomes were assembled with trimmed and quality-controlled reads using MEGAHIT (Li *et al.*, 2015) (version 0.2.0; options: --cpu-only -m 100e9 --k-max 123 -l 155).

Statistical analyses and metatranscriptome exploration tools. Pairwise similarities among metatranscriptomes were calculated using Bray-Curtis similarity values (Legendre and Legendre, 2012) and visualized with principle coordinates analysis using Primer-e software (www.primer-e.com). The difference between treatments was assessed with PERMANOVA. Alpha diversity of each KEGG annotated metatranscriptome was assessed by the Shannon index using the 'diversity' function in the R package "vegan" (v.2.5-5) (Oksanen et al. 2019). (Oksanen *et al.*, 2019). MEtaGenome ANalyzer

(MEGAN) (Huson *et al.*, 2011) and ShotMAP (Nayfach *et al.*, 2015) software were used to explore abundances of transcripts within KEGG categories, pathways, and individual KOs.

16S rRNA gene amplification. PCR 16S rRNA amplicon sequencing of extracted DNA followed the Earth Microbiome Project protocol (http://www.earthmicrobiome.org/emp-standard-protocols/16s/): Primers focused on the V4 region of the 16S rRNA gene (515F, GTGCCAGCMGCCGCGGTAA and 806R, GGACTACHVGGGTWTCTAAT), and were combined at 250 nM with template DNA, sterile water and HotMasterMix (5Prime) under the following conditions: 94°C for 3 min; 30 cycles of 94°C for 45 sec, 50°C for 60 sec, 72°C for 90 sec; 72°C for 10 min. Triplicate amplifications were pooled, quantified with Picogreen, combined in equimolar concentrations, and cleaned using the Ultraclean PCR Clean-Up Kit (MoBio).

16S rRNA gene analysis. 16S amplicon sequences were paired using make.contigs (Schloss *et al.*, 2009) from MOTHUR v. 1.32.1, and converted to QIIME format with split.groups from MOTHUR and add_qiime_labels.py from QIIME (Caporaso *et al.*, 2010). Sequences were quality filtered with an expected error rate of 0.5, dereplicated (derep_fulllength), and abundance sorted (sortbysize) using USEARCH v.7.0.1001_i86linux64 (Edgar, 2013). Singleton sequences were removed in the latter step to prevent them from seeding clusters when clustering OTUs. Reads were then clustered (cluster_otus) at 97% similarity and chimeras were removed via the de novo chimera check inherent in the cluster_otus in addition to reference-based chimera filtering (uchime_ref) with the Gold Database (www.genomesonline.org) as reference. Reads (including singletons) were subsequently mapped back to OTUs using UPARSE (usearch_global) and an OTU table was created. Taxonomy of the representative sequences was assigned in QIIME (assign_taxonomy.py) using the RDP classifier trained to the SILVA database (v.111 database clustered to 97% OTUs). OTUs classified as Eukaryote, Archaea, Chloroplast, Mitochondria, and Unknown were removed, and the OTU table was rarefied to 3800 sequences per sample, and beta-diversity calculated as Bray-Curtis similarity. The raw OTU table is provided in Dataset S6.

Taxonomic binning of KEGG annotated RNA transcripts. Dominant taxa from KEGG annotated metatranscriptomes were classified to the phylum or class level for analysis. Taxa that did not make dominant contributions to the metatranscriptomes were classified as "other."

DOM characterization. The percent of DOM that was mineralized to CO_2 was determined by dividing the amount of CO_2 produced by the initial amount of DOM, measured as dissolved organic carbon (DOC). The percent of DOM consumed or altered by sunlight was determined by adding mineralized DOM and partially oxidized DOM (Ward and Cory, 2015, 2016). Fourier transform ion cyclotron mass spectra (FT-ICR MS) were analyzed according to Ward & Cory (Ward and Cory, 2016). Formulas were classified as aromatic, aliphatic, tannin-like, lignin-like or highly-oxidized on the basis of their chemical composition (Ward and Cory, 2015, 2016; Ward *et al.*, 2017). Formulas were categorized as photo-degraded or photo-produced if their intensity decreased or increased, respectively, after sunlight exposure, and categorized as bio-degraded if their intensity decreased after the microbial incubation period. 95% confidence intervals calculated across experimental replicates were used to determine if a change in formula intensity due to photo-degradation, photo-production, or bio-degradation was significantly greater than zero (Ward and Cory, 2015; Ward *et al.*, 2017). Quantification of functional groups such as carboxyl carbon was determined from ^{13}C NMR spectroscopy following (Ward and Cory, 2016).

Additional explanation of DOM characterization for those not familiar with FT-ICR MS. The classification of DOM chemical composition used in this experiment is described in detail in (Ward and Cory, 2015, 2016; Ward *et al.*, 2017), and is based on literature conventions that have been described in detail including in many reviews on this topic (Sleighter *et al.*, 2008; Hockaday *et al.*, 2009; Stubbins *et al.*, 2010). Briefly, every ion detected by FT-ICR MS with a unique mass is assigned a formula (e.g., a

mass of 332.30474 Da is assigned a formula of $C_{17}H_{16}O_7$). Formulas are then grouped into broad categories of biomolecules such as tannins or lignin by comparing the atomic ratio of each formula (e.g., H/C, O/C, N/C) with the atomic ratios of major biomolecules expected to contribute to the DOM pool. DOM formulas that overlap with the range of atomic ratios of major biomolecules, such as tannins and lignin, share chemical properties such as number of double bonds, degree of saturation, and oxidization state, with these biomolecules. Whether a formula is aromatic or aliphatic is unambiguously determined based on the number of C, H, O, N, S and P in the formula and the double bond equivalence (DBE) of each formula (Koch and Dittmar, 2006). This sharing of chemical properties between DOM formulas and major groups of organic molecules forms the expectations for DOM compounds as being more labile or more recalcitrant to degradation.

Table S1. RNA sequencing statistics for the six metatranscriptomes.

IMG Genome ID	3300005640	3300005644	3300006426	3300005650	3300006423	3300005646
Crump Lab number	2013-052	2013-053	2013-054	2013-055	2013-056	2013-057
Treatment	light	dark	light	dark	light	dark
Replicate	1	1	2	2	3	3
Total number of reads	73,629,230	78,208,760	120,499,524	121,268,126	79,264,904	79,188,270
Number of contigs	610,923	796,077	960,066	1,446,774	493,929	992,332
RNA genes	15,039	13,957	95,512	19,738	77,669	12,863
Protein genes (CDS)	703,933	926,550	1,030,181	1,661,625	500,885	1,151,205
Genes annotated to KO	354,037	427,150	488,296	721,583	208,627	527,077
Reads mapped to CDS	62,465,501	66,390,431	95,739,719	98,052,903	62,500,740	64,198,959
Reads mapped to annotated CDS	30,280,392	26,417,344	43,034,810	39,872,640	19,766,388	29,593,071
Reads mapped to prokaryotic CDS	30,174,590	26,375,653	42,720,850	39,793,557	19,517,390	29,541,901

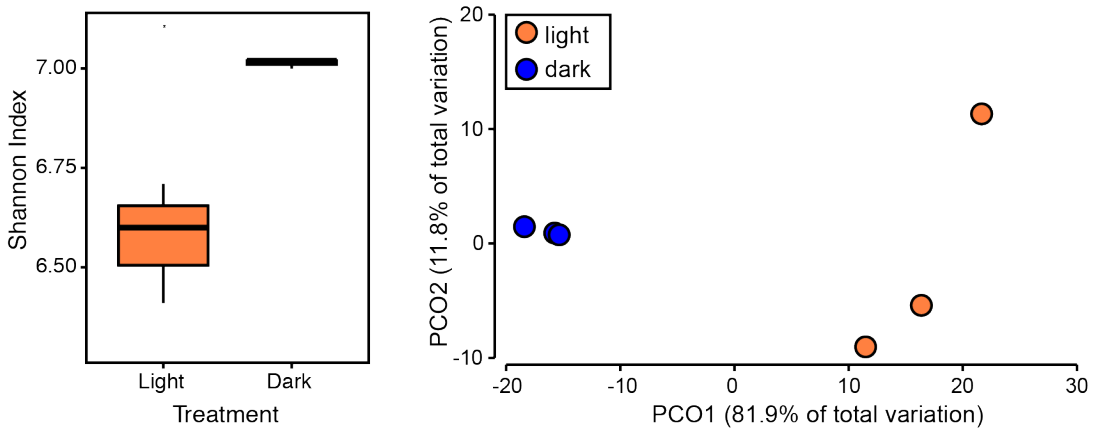


Figure S1. A boxplot (left) representing Shannon alpha diversity index of KEGG annotated metatranscriptomes for each treatment (paired t-test, P = 0.033), based on raw gene counts. Boxes represent the interquartile range (IQR); the line inside each box represents the median. Upper whiskers on boxes represent the smaller of the maximum value or quartile 3 + (1.5 x IQR). Lower whiskers on boxes represent the larger of the minimum value or quartile 1- (1.5 x IQR). Principle coordinates diagram (right) based on Bray-Curtis similarity of TPM normalized KEGG annotated metatranscriptomes.

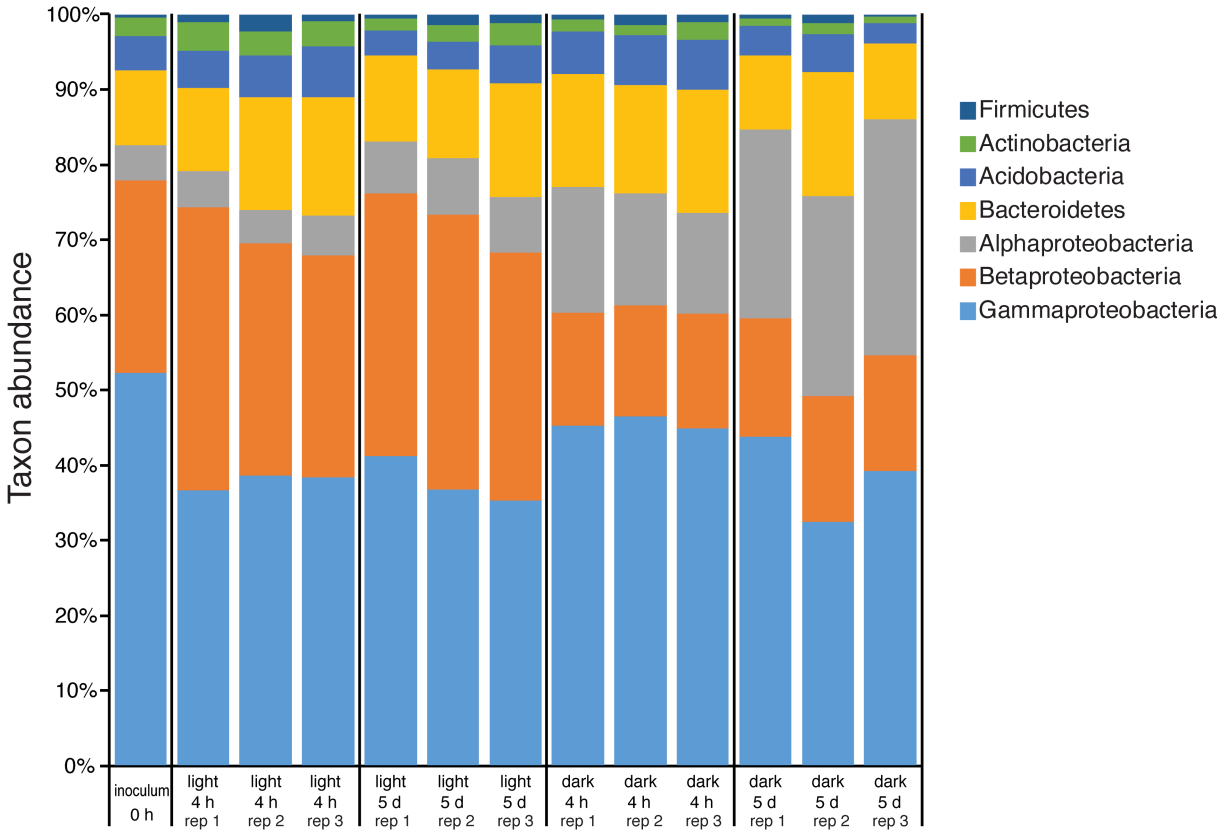


Figure S2. Relative taxonomic composition of the whole community (16S amplicons) in the inoculum and in all treatment replicates at 4 hours and 5 days incubation.

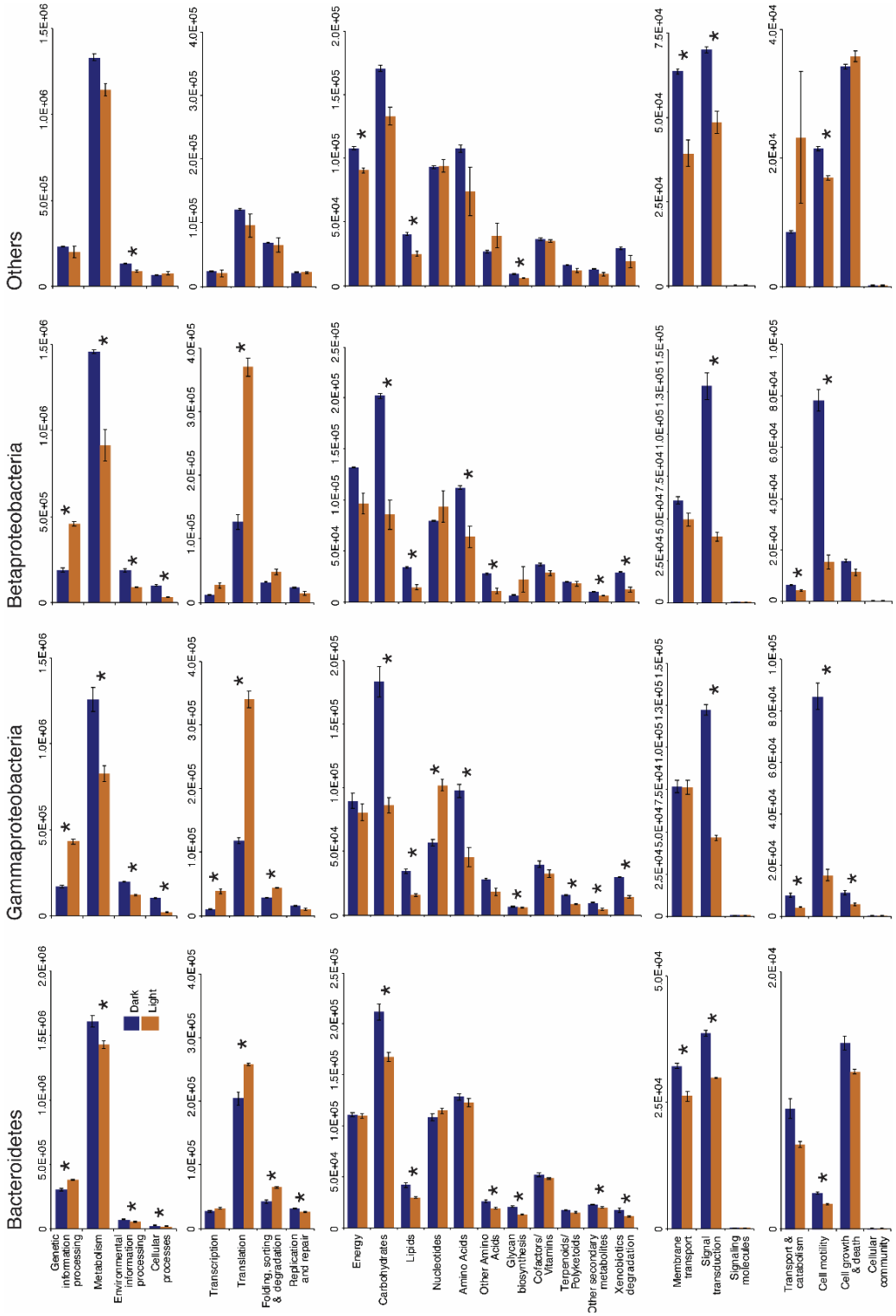


Figure S3. Expression (summed TPM) of KEGG tier II (A) and tier III (B-E) categories in the dark treatment (blue) and light treatment (orange) for transcripts classified to Bacteroidetes, Gammaproteobacteria, Betaproteobacteria, and other bacteria. Asterisks represent significant differences according to paired t-tests ($p \leq 0.05$). Error bars indicate standard error of the mean.

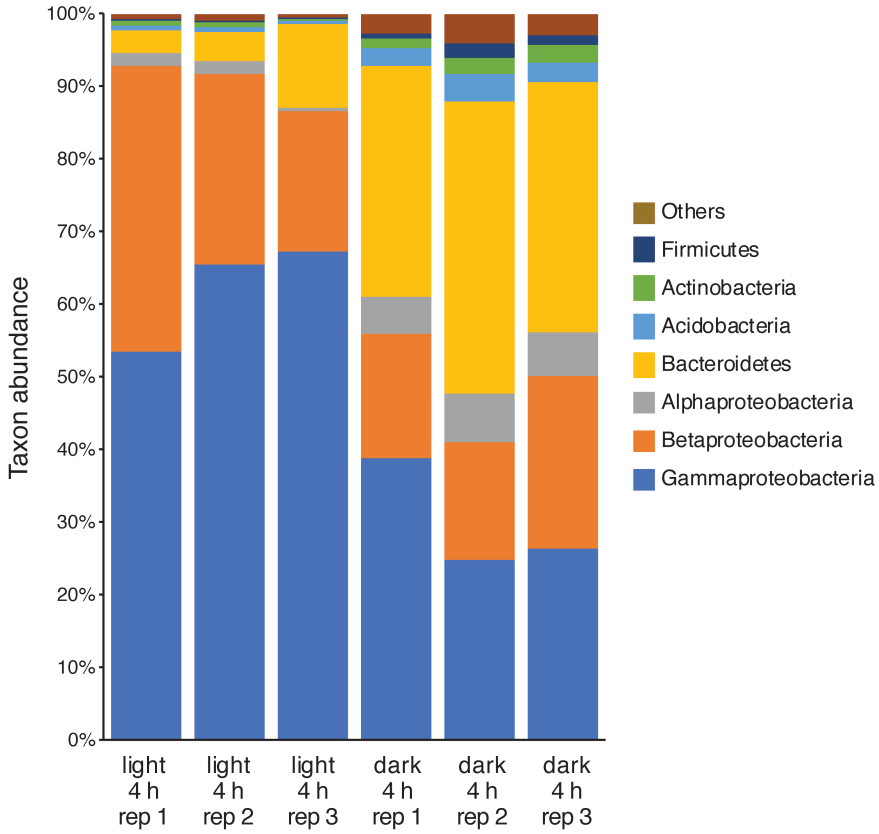


Figure S4. Relative taxonomic composition of expressed genes belonging to the Ribosome KEGG pathway in all treatment replicates at 4 hours.

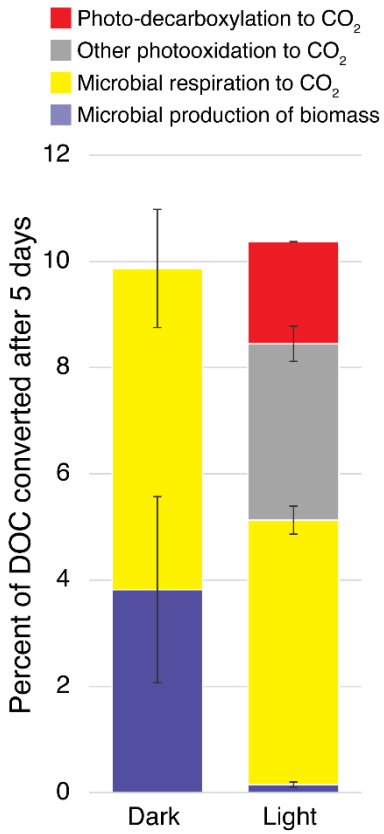


Figure S5. Percent of dissolved organic carbon (DOC) oxidized to CO_2 by sunlight and microbes or converted to microbial biomass over 5 days of incubation in the two treatments. Error bars indicate standard error of the mean.

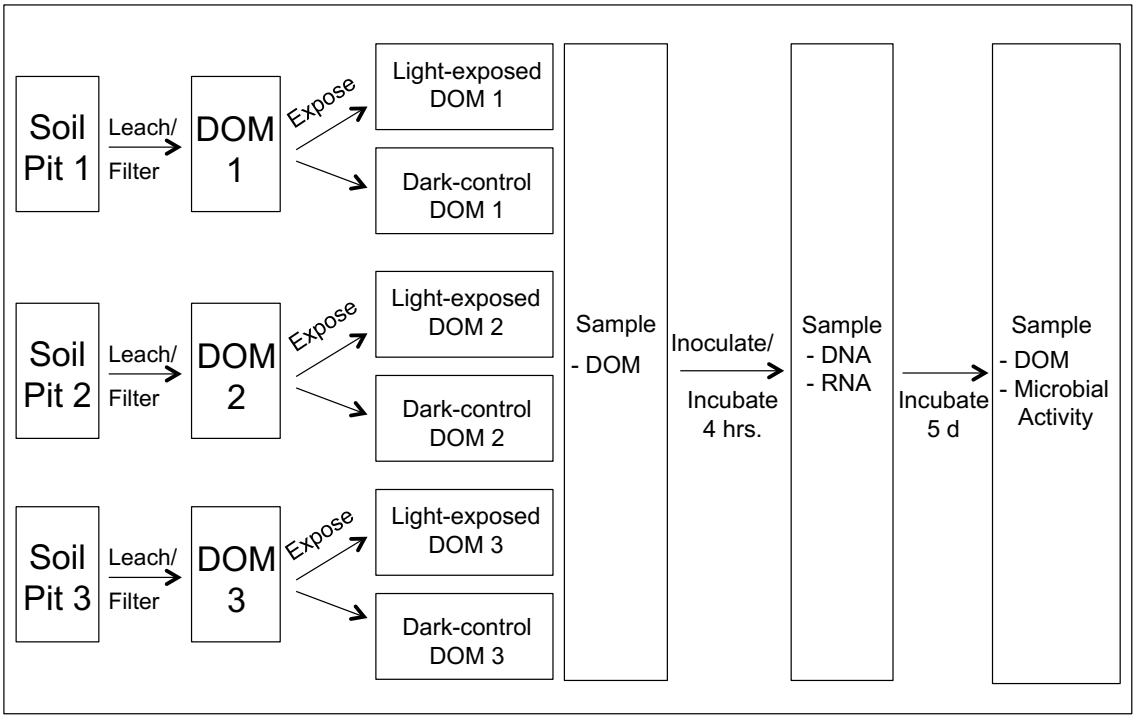


Figure S6. Schematic of experimental design.

Dataset S1
Excel spreadsheet providing transcript abundances, expressed as number of reads mapped (Rg) and as transcripts per million (TPM), for genes annotated to the KEGG database. The first tab includes all KEGG Orthologs (KOs) found in samples. The second tab includes all KOs categorized as KEGG tier IV ABC Transporters (ABC transporter genes). The third tab includes all KOs categorized as KEGG tier II Metabolism genes. The fourth, fifth, and sixth tabs include subsets of the Metabolism category and defined as aromatic degradation genes, oxygenase genes, decarboxylase genes, respectively. Statistics from EdgeR analysis using TPM normalization are provided including $\log_2$ fold change (logFC) between treatments, the average log2 counts per million (logCPM), p-value, and false discovery rate (FDR).

Dataset S2
Excel spreadsheet providing transcript abundances, expressed as number of reads mapped (Rg) and as transcripts per million (TPM), for genes annotated to the KEGG database and taxonomically classified to Bacteroidetes. The first tab includes all KEGG Orthologs (KOs) found in samples. The second tab includes all KOs categorized as KEGG tier IV ABC Transporters (ABC transporter genes). The third tab includes all KOs categorized as KEGG tier II Metabolism genes. The fourth, fifth, and sixth tabs include subsets of the Metabolism category defined as aromatic degradation genes, oxygenase genes, decarboxylase genes, respectively. Statistics from EdgeR analysis using TPM normalization are provided including $\log_2$ fold change (logFC) between treatments, the average log2 counts per million (logCPM), p-value, and false discovery rate (FDR).

Dataset S3
Excel spreadsheet providing transcript abundances, expressed as number of reads mapped (Rg) and as transcripts per million (TPM), for genes annotated to the KEGG database and taxonomically classified to Gammaproteobacteria. The first tab includes all KEGG Orthologs (KOs) found in samples. The second tab includes all KOs categorized as KEGG tier IV ABC Transporters (ABC transporter genes). The third tab includes all KOs categorized as KEGG tier II Metabolism genes. The fourth, fifth, and sixth tabs include subsets of the Metabolism category defined as aromatic degradation genes, oxygenase genes, decarboxylase genes, respectively. Statistics from EdgeR analysis using TPM normalization are provided including $\log_2$ fold change (logFC) between treatments, the average log2 counts per million (logCPM), p-value, and false discovery rate (FDR).

Dataset S4
Excel spreadsheet providing transcript abundances, expressed as number of reads mapped (Rg) and as transcripts per million (TPM), for genes annotated to the KEGG database and taxonomically classified to Betaproteobacteria. The first tab includes all KEGG Orthologs (KOs) found in samples. The second tab includes all KOs categorized as KEGG tier IV ABC Transporters (ABC transporter genes). The third tab includes all KOs categorized as KEGG tier II Metabolism genes. The fourth, fifth, and sixth tabs include subsets of the Metabolism category defined as aromatic degradation genes, oxygenase genes, decarboxylase genes, respectively. Statistics from EdgeR analysis using TPM normalization are provided including $\log_2$ fold change (logFC) between treatments, the average log2 counts per million (logCPM), p-value, and false discovery rate (FDR).

Dataset S5
Excel spreadsheet providing transcript abundances, expressed as number of reads mapped (Rg) and as transcripts per million (TPM), for genes annotated to the KEGG database and taxonomically classified to organisms other than Bacteroidetes, Gammaproteobacteria, and Betaproteobacteria. The first tab includes all KEGG Orthologs (KOs) found in samples. The second tab includes all KOs categorized as KEGG tier IV ABC Transporters (ABC transporter genes). The third tab includes all KOs categorized as KEGG tier II Metabolism genes. The fourth, fifth, and sixth tabs include subsets of the Metabolism category defined

as aromatic degradation genes, oxygenase genes, decarboxylase genes, respectively. Statistics from edgeR analysis using TPM normalization are provided including $\log_2$ fold change (logFC) between treatments, the average log2 counts per million (logCPM), p-value, and false discovery rate (FDR).

Dataset S6

Excel spreadsheet providing the raw 16S rRNA gene amplicon OTU table, and a list of taxonomic identities and TPM for all gene transcripts identified in metatranscriptomes. The metadata tab provides information about each sample.

Dataset S7

Tab-delimited file of the taxonomic (phylodist) and KEGG (KO) annotations for each CDS identified in this study.

References

Bushnell, B. (2015) BBMap short-read aligner, and other bioinformatics tools.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Edgar, R.C. (2013) UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* **10**: 996–998.

Hockaday, W.C., Purcell, J.M., Marshall, A.G., Baldock, J.A., and Hatcher, P.G. (2009) Electrospray and photoionization mass spectrometry for the characterization of organic matter in natural waters: a qualitative assessment: Characterization of organic matter in natural waters. *Limnol Oceanogr Methods* **7**: 81–95.

Huson, D.H., Mitra, S., Ruscheweyh, H.J., Weber, N., and Schuster, S.C. (2011) Integrative analysis of environmental sequences using MEGAN4. *Genome Res* **21**: 1552–1560.

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

Legendre, P. and Legendre, L. (2012) Numerical Ecology, Amsterdam: Elsevier Science B.V.

Li, D., Liu, C.M., Luo, R., Sadakane, K., and Lam, T.W. (2015) MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676.

Nayfach, S., Bradley, P.H., Wyman, S.K., Laurent, T.J., Williams, A., Eisen, J.A., et al. (2015) Automated and Accurate Estimation of Gene Family Abundance from Shotgun Metagenomes. *PLoS Comput Biol* **11**: 1–29.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., et al. (2019) vegan: Community Ecology Package. R package version 2.5-5.

Satinsky, B.M., Fortunato, C.S., Doherty, M., Smith, C.B., Sharma, S., Ward, N.D., et al. (2015) Metagenomic and metatranscriptomic inventories of the lower Amazon River, May 2011. *Microbiome* **3**:.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Sleighter, R.L., McKee, G.A., Liu, Z., and Hatcher, P.G. (2008) Naturally present fatty acids as internal calibrants for Fourier transform mass spectra of dissolved organic matter: DOM mass spectra calibration with fatty acids. *Limnol Oceanogr Methods* **6**: 246–253.

Stubbins, A., Spencer, R.G.M., Chen, H., Hatcher, P.G., Mopper, K., Hernes, P.J., et al. (2010) Illuminated darkness: Molecular signatures of Congo River dissolved organic matter and its photochemical alteration as revealed by ultrahigh precision mass spectrometry. *Limnol Oceanogr* **55**: 1467–1477.

Ward, C.P. and Cory, R.M. (2015) Chemical composition of dissolved organic matter draining permafrost soils. *Geochim Cosmochim Acta* **167**: 63–79.
Ward, C.P. and Cory, R.M. (2016) Complete and Partial Photo-oxidation of Dissolved Organic Matter Draining Permafrost Soils. *Environ Sci Technol* **50**: 3545–3553.
Ward, C.P., Nalven, S.G., Crump, B.C., Kling, G.W., and Cory, R.M. (2017) Photochemical alteration of organic carbon draining permafrost soils shifts microbial metabolic pathways and stimulates respiration. *Nat Commun* **8**: 772.