emi_2092

Environmental Microbiology (2009) doi:10.1111/j.1462-2920.2009.02092.x

The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity

M. S. Schwalbach, H. J. Tripp,† L. Steindler, D. P. Smith and S. J. Giovannoni*

Oregon State University, Department of Microbiology, 220 Nash Hall, Corvallis, OR 97331, USA.

Summary

Bacteria in the SAR11 clade are highly abundant in marine surface waters, but currently little is known about the carbon compounds that support these large heterotrophic populations. To better understand the carbon requirements of these organisms, we conducted a multiphasic exploration of carbohydrate utilization among SAR11 isolates from the Northeast Pacific Ocean and the Sargasso Sea. A comparison of three SAR11 genomes showed they all lacked a recognizable PTS system, the oxidative portion of the pentose phosphate shunt (*zwf*⁻, *pgl*⁻), and had incomplete Embden–Meyerhoff–Parnas (*pfk*⁻, *pyk*⁻) or Entner–Doudoroff (*eda*⁻) glycolysis pathways, suggesting these isolates are limited in their ability to transport and oxidize exogenous carbohydrates. Growth assays, radioisotopes, metagenomics and microarrays were utilized to test this hypothesis. Galactose, fucose, rhamnose, arabinose, ribose and mannose could not serve as carbon sources for the isolates tested. However, differences in glucose utilization were detected between coastal and ocean gyre isolates, with the coastal isolates capable of transporting, incorporating and oxidizing glucose while the open ocean isolate could not. Subsequent microarray analysis of a coastal isolate suggested that an operon encoding a variant of the Entner–Doudoroff pathway is likely responsible for the observed differences in glucose utilization. Metagenomic analysis indicated this operon is more commonly found in coastal environments and is positively correlated with chlorophyll *a* concentrations. Our results indicated that glycolysis is a variable metabolic property of SAR11 metabolism and suggest that glycolytic SAR11 are more common in productive marine environments.

Received 14 November, 2008; accepted 10 September, 2009. *For correspondence. E-mail steve.giovannoni@oregonstate.edu; Tel. (+••) •• Fax (+••) ••. †Present address: University of California, Ocean Sciences Department, Earth and Marine Sciences Building Room A-433, Santa Cruz, CA 95064, USA.

Introduction

Marine bacteria of the SAR11 clade are the most abundant aerobic, heterotrophic bacteria in ocean surface waters (Morris *et al.*, 2002) and previous studies have indicated these cells play an important role in marine carbon cycling (Malmstrom *et al.*, 2005; Morris *et al.*, 2005). Despite their abundance, little is known about the organic compounds that native SAR11 populations use to fulfil cellular biosynthetic and energetic demands. This is due, in part, to the chemical complexity of natural marine dissolved organic carbon (DOC), which is comprised of a diverse but poorly resolved mixture of organic compounds that are produced via a multitude of biotic and abiotic processes (Moran and Zepp, 1997; Carlson *et al.*, 1998; Williams, 2000; Kujawinski *et al.*, 2004; Repeta and Aluwihare, 2006). While previous studies have indicated marine bacteria play a central role in the cycling of marine DOC (Azam *et al.*, 1983; Carlson *et al.*, 2004; Azam and Malfatti, 2007), determining which carbon compounds are utilized by which marine prokaryotes in a natural environment remains technically challenging.

Given the difficulties in cultivating marine organisms and in resolving the chemical complexity of dissolved organic matter, marine microbiologists began to investigate marine bacterial carbon utilization by examining the response of key community parameters, such as bacterial production, to the addition of defined carbon compounds or fractionated DOC (e.g. the high-molecular-weight fraction etc.). Many of these studies focused on the uptake and oxidation of dissolved free amino acids (DFAA), in part because proteins constitute a significant fraction of a bacterium's biomass. From these studies emerged a general trend that DFAA can support a significant percentage of a marine bacterial production (see review by Kirchman, 2003). However, the contribution of DFAA to bacterial production was at times found to be surprisingly low in oligotrophic environments such as the Sargasso Sea (Suttle *et al.*, 1991; Keil and Kirchman, 1999), suggesting other carbon sources are also important in these environments.

Phytoplankton are known to release significant amounts of polysaccharides into the water column (see review by Carlson, 2002), suggesting a role for carbohydrates in marine bacterial carbon cycling. Early investigations into the reactivity of neutral monosaccharides indicated these molecules are rapidly utilized by marine bacteria, with glucose concentrations at times being sufficient to directly support a significant amount (5–85%) of the measurable bacterial production in oligotrophic environments such as the equatorial Pacific Ocean and the Gulf of Mexico (Rich *et al.*, 1996; Skoog *et al.*, 1999). In addition, reports of aggressive enzymatic hydrolysis of the polysaccharides (Smith *et al.*, 1992) found in high-molecular-weight DOC (Aluwihare *et al.*, 1997; Skoog and Benner, 1997) also suggest that many marine bacteria actively use carbohydrates to meet their biological carbon demands. However, similar to early DFAA studies, carbohydrate utilization appears to vary both temporally and spatially, making general conclusions difficult.

With the advent of molecular techniques that address *in situ* activity and substrate utilization, such as microautoradiography (MAR) combined with fluorescent *in situ* hybridization (FISH), community fingerprinting and stable isotope probing (SIP), it is possible to assay the uptake of isotope-labelled compounds by specific members of a bacterial community. The recent application of MAR-FISH to a marine bacterioplankton community indicated that bacteria of the SAR11 clade accounted for nearly half of the uptake of nanomolar amounts of radiolabelled amino acids (Malmstrom *et al.*, 2004). Aggressive uptake of DFAA by SAR11 may be explained, in part, by recent observations that marine SAR11 are auxotrophic for the amino acids glycine and serine (Tripp *et al.*, 2009) and that they may rely on exogenous methionine as a sulfur source (Tripp *et al.*, 2008), but currently it is not known to what extent marine SAR11 utilize amino acids or exogenous proteins as general sources of carbon.

SAR11 carbohydrate metabolism has also been probed by MAR-FISH studies, with results varying by location. SAR11 dominated the uptake of D-glucose in MAR-FISH studies collected from the Northwest Atlantic Ocean and in samples collected from the coastal North Sea (Malmstrom *et al.*, 2005; Alonso and Pernthaler, 2006), but members of the SAR11 clade did not compete as strongly for free glucose in the Arctic Ocean or the Delaware Estuary (Elifantz *et al.*, 2005; Alonso-Saez *et al.*, 2008). Similarly, other experimental work examining the response of natural populations of bacterioplankton to glucose addition indicated that SAR11 populations are at times not responsive to glucose addition (Carlson *et al.*, unpublished). Combined, these studies suggest that SAR11 carbohydrate metabolism varies among ecotypes and differs in its extent by location, but currently the mechanism(s) responsible for this variability are not understood.

While the MAR-FISH results of glucose utilization by marine SAR11 have been variable, they do clearly indicate that some SAR11 can utilize glucose. Complete genomes provide one of the most comprehensive sources of information about microbial metabolism. Thus, it was with curiosity we noted that a complete pathway for glucose oxidation could not be reconstructed from the first SAR11 genome (Giovannoni *et al.*, 2005), and genomes that followed showed even less evidence of this central metabolic pathway (Stingl *et al.*, 2009). These observations further suggested that the variable responses of SAR11 populations to glucose that have been reported might be due to natural genetic variation in SAR11 carbohydrate metabolism. To address this question we used multiple approaches to examine carbohydrate metabolism in two high-light-adapted, SAR11 surface Ia ecotypes, HTCC1062 and HTCC7211 (99.9% 16s rRNA similarity) (Carlson *et al.*, 2009), that had been isolated from either the Oregon Coast or the Sargasso Sea respectively. Multiple lines of evidence indicated that coastal SAR11 have a greater capacity for carbohydrate oxidation than SAR11 from ocean gyres, and that at least some gyre SAR11 lack glycolysis completely and cannot oxidize glucose to CO_2.

Results

Comparative genomics

Initial annotation of the HTCC1062 genome showed that genes for glucose uptake and phosphorylation (i.e. the PTS system) were absent from this isolate (Giovannoni *et al.*, 2005). The typical Embden–Meyerhoff–Parnas (EMP) pathway of glycolysis is incomplete since typical genes for 6-phosphofructokinase (E.C. 2.7.1.11) and pyruvate kinase (E.C. 2.7.1.40) are absent from the genome. It was also observed that HTCC1062 contained a predicted homologue of the Entner–Doudoroff (ED) dehydratase gene (*edd*, E.C. 4.2.1.12), but that the genome lacked a recognizable gene encoding an ED aldolase (*eda*, 4.1.2.14) (Giovannoni *et al.*, 2005). However, the presence of two ABC transport systems (Table S1) that were annotated as 'sugar' transporters and an *edd* gene suggested that some form of glycolysis might be encoded by the genome, although the pathway remained uncertain at the time (Giovannoni *et al.*, 2005).

Additional SAR11 genome sequences from HTCC1002 and HTCC7211 have become available after the analysis of the first SAR11 genome, providing comparative data about variation in SAR11 carbohydrate metabolism. HTCC1002 differs in 16S rRNA gene sequence by one base pair from HTCC1062, and HTCC7211 differs by 15 base pairs in 16S rRNA gene sequence from the other two strains. Many of the central metabolism genes noted as

absent in the HTCC1062 genome were also lacking from the HTCC1002 and HTCC7211 genomes (Table S1). Specifically, no genes were found that encoded PTS transport (*pts*IGH), the oxidative portion of the pentose phosphate cycle (*zwf*, *pgl*), 6-phosphofructokinase (*pfk*) and pyruvate kinase (*pyk*) (Table S1). In addition to these key lesions, HTCC1002 also lacked a recognizable phosphoenolpyruvate (PEP) carboxylase (E.C. 4.1.1.31). Hence, many of the central metabolism lesions observed in HTCC1062 appear to be conserved in isolates obtained from within and across ocean basins.

Despite the observed variations, other important features of central metabolism are conserved among the three genomes. All three isolates possessed a complete TCA cycle as well as a complete gluconeogenic pathway that utilizes the alternate enzyme pyruvate orthophosphate dikinase (*ppdk*) to generate phosphoenolpyruvate as well as an alternate fructose 1,6-bisphosphatase *glp*X. All isolates have genes that encode a complete glyoxylate cycle that is regulated, in part, by a conserved glycine-binding riboswitch (Tripp *et al.*, 2009). Furthermore, all three isolates contain genes encoding PQQ-dependent aldose dehydrogenases (Table S1). Interestingly, HTCC1002 and HTCC7211 have only one of the two predicted ABC-sugar transport systems found in HTCC1062 (Table S1). The sugar transporter genes found in HTCC1002 are localized near the predicted ED genes, as is the case in HTCC1062. However, the predicted sugar transporter genes in HTCC7211 are more similar to msm-like sugar transporters of HTCC1062, which are not located near the ED genes in the HTCC1062 genome (Table S1). The genome position of the putative ABC-sugar transporters near the *edd* genes in HTCC1062 and HTCC1002 led us to speculate that this region of the genome contains an operon that encodes the transport and oxidation of glucose. Homologues of HTCC1062 genes SAR11_0769_0778 were present in the genome of HTCC1002, but were mostly lacking from the HTCC7211 genome. However, two of the genes (SAR11_0775, SAR11_0778) in the glycolysis operon of the coastal isolates were present in the HTCC7211 genome, but were not located near each other, nor next to the putative ABC-sugar transporter genes found in HTCC7211, suggesting they may have alternate functions (Table S1).

Radioisotope and cultivation assays

In accord with the predictions from genome annotation, both isotope and growth assays indicated that coastal SAR11 isolates could utilize glucose as a carbon source, but the Sargasso Sea isolate could not. Radiotracer studies with ^{14}C-D-glucose indicated the Sargasso Sea SAR11 isolate HTCC7211 was incapable of transporting, incorporating or oxidizing glucose when supplied at a final concentration of 1 μM (Fig. 1A) or 10 μM (data not shown). Conversely radiolabelled glucose supplied at a final concentration of 1 μM was transported, incorporated and oxidized to CO_2 by the Oregon coast isolate HTCC1062 (Fig. 1A). Similar glucose oxidation was also observed for a close relative of HTCC1062, SAR11 isolate HTCC1002 (data not shown). Both HTCC7211 and HTCC1062 transported, incorporated and oxidized radiolabelled ^{14}C-pyruvate, indicating all cells were active during the experiments (Fig. 1B).

Results from growth assays on glucose were consistent with results obtained from isotope studies. Cultivation assays with SAR11 isolates HTCC1062 and HTCC7211 were conducted in seawater media amended with D-glucose, pyruvate or both as the primary carbon sources and all other nutrients (ammonia, phosphorus, L-methionine, L-serine, iron and vitamins) in excess. Control treatments also contained all nutrients in excess, but were not amended with a primary carbon source. Control treatments reached similar final cell densities (mean ± SD, HTCC1062: 1.93 ± 0.01 cells × 10^6 ml^{-1}, HTCC7211: 2.79 ± 0.43 cells × 10^6 ml^{-1}) (Fig. 1C and D). Similarly both isolates reached maximum final cell densities when pyruvate was the principal carbon source (HTCC1062: 7.20 ± 0.46 cells × 10^6 ml^{-1}, HTCC7211: 8.58 ± 0.25 cells × 10^6 ml^{-1}, Fig. 1C and D). Final cell densities of HTCC7211 grown on glucose were not significantly higher than control treatments (3.79 ± 0.10 cells × 10^6 ml^{-1}, $P = 0.22$) but were significantly lower ($P = 3.5 \times 10^{-6}$) than cell concentrations obtained when grown on pyruvate (Fig. 1D). HTCC1062 grown on glucose reached a final cell density comparable to, but slightly lower than, densities obtained when grown on pyruvate (5.97 ± 0.21 cells × 10^6 ml^{-1}), but growth rate on glucose was slightly lower compared with growth rate on pyruvate (0.520 ± 0.007 day^{-1} versus 0.561 ± 0.005 day^{-1}, $P = 0.0001$) (Fig. 1A and B). Additional carbohydrate utilization patterns did not differ between isolates, with neither able to utilize the carbohydrates fucose, rhamnose, mannose, galactose, ribose and arabinose nor the disaccharides maltose and trehalose (Table 1). Finally, it was noted that neither isolate utilized gluconic acid despite both having predicted PQQ-dependent glucose dehydrogenases, but both isolates utilized taurine, lactate, acetate and oxaloacetate (Table 1).

Metagenomic analysis

The experimental data provided conclusive evidence that glucose oxidation is a variable property of SAR11 metabolism, but the low number of isolates available for study left uncertainty about how prevalent glycolysis deficiencies

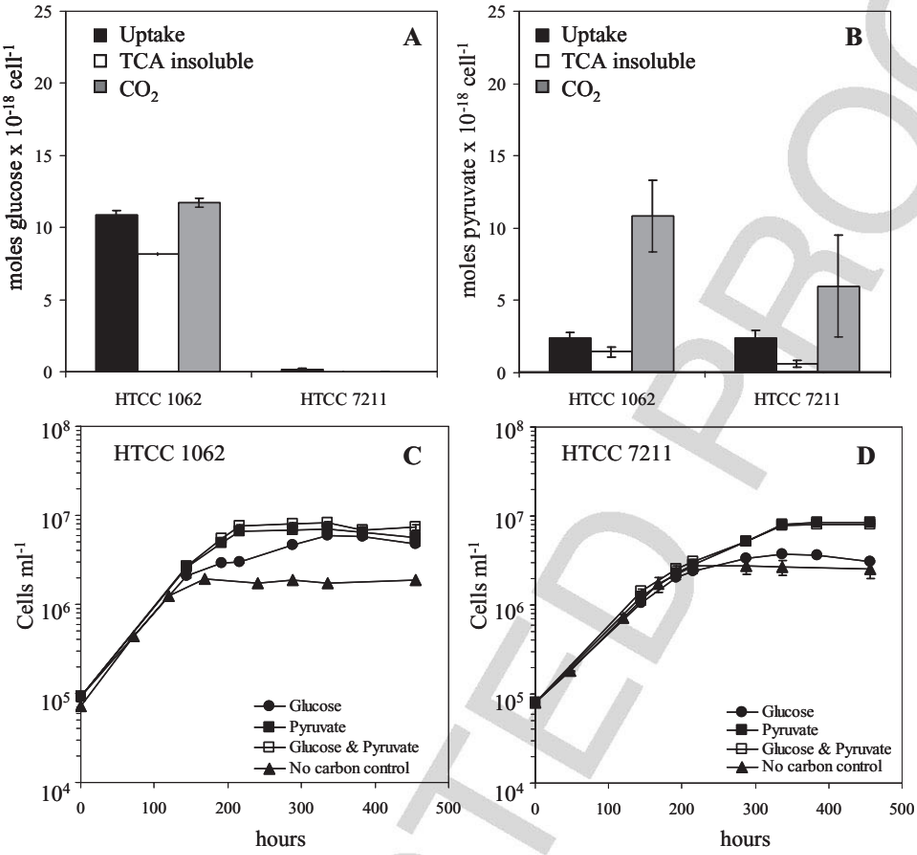

Fig. 1. A and B. Utilization of radiolabelled (A) ^{14}C-D-glucose (1 μM final concentration) and (B) ^{14}C-pyruvate (1 μM final concentration) by HTCC isolates 1062 and 7211. Assays were conducted for 24 h at room temperature (~22°C) in artificial seawater media. Error bars represent the mean ± SD of triplicate measurements.
C and D. Growth curves of SAR11 isolates HTCC1062 (C) and HTCC7211 (D) when glucose (●) or pyruvate (■) was supplied as the primary carbon source. Control treatments that lacked a carbon source are indicated with the symbol ▲ and HTCC7211 cells grown on pyruvate in the presence of 25 μM glucose are indicated with the symbol □. Data for glucose, pyruvate, and both pyruvate and glucose combined are mean ± SD of quadruplicate flasks. Control treatments represented by mean ± range of duplicate flasks.

are among marine SAR11 and whether the absence of this important trait is more common in open ocean isolates, as suggested by the data. To address these questions we used BLAST searches and a reciprocal best BLAST (RBB) test to determine the frequency of all SAR11 glycolysis genes (Fig. 2A) in metagenomic data collected during the global ocean survey (GOS) (Venter *et al.*, 2004; Rusch *et al.*, 2007). The results confirmed that the SAR11 glycolysis operon is not uniformly present among marine SAR11 bacteria but rather is variably present and its presence in SAR11 genomes is positively correlated with chlorophyll *a* concentrations ($r = 0.692$, $P = 6.3 \times 10^{-5}$) (Fig. 2B). The frequency of SAR11 r*ec*A genes, which were used to normalize the BLAST search data, indicated that SAR11 bacteria comprised on average 35 ± 7.5% of the cells in the GOS data sets examined. Among the individual GOS samples, the analysis indicated that the frequency of SAR11 cells bearing the glycolysis operon varied from 30% to 100% (Fig. 2B). The genomic arrangement of these genes is indiscernible from the current data set due to the fragmentary nature of metagenomic libraries; however, in all cases the entire operon was detected within a sample.

Microarray results

Patterns of gene expression often provide clues about metabolic pathways since genes involved in the same process are often coordinately expressed. Therefore, we used Affymetrix GeneChip oligonucleotide microarrays to examine the transcriptional response of the predicted glycolysis operon to the presence of glucose. HTCC1062 cells were grown in filtered, autoclaved seawater media with inorganic nutrients and vitamins in excess and glucose either present or absent. The resulting transcription profiles were highly similar between

Table 1. Maximum cell densities and growth rates from carbon utilization tests of HTCC1062 and HTCC7211.

	HTCC1062			HTCC7211		
Carbon compound	Result	Cells × 10^6 ml^{-1} (mean ± range)	Doubling time (days)	Result	Cells × 10^6 ml^{-1} (mean ± range)	Doubling time (days)
Pyruvate	+	7.8 ± 0.41	1.3	+	5.9 ± 0.15	1.6
L-lactate	+	6.6 ± 0.74	1.6	+	5.7 ± 0.06	1.7
Oxaloacetate	+	5.9 ± 0.72	1.3	+	6.1 ± 0.27	1.4
Acetate	+	3.1 ± 0.05	1.4	+	4.2 ± 0.23	1.8
Taurine	+	3.1 ± 0.22	1.3	+	5.3 ± 0.25	2.0
D-(+)-glucose	+	5.1 ± 0.01	1.4	–	3.4 ± 0.09	1.8
No Carbon control		1.3 ± 0.02	1.4		3.5 ± 0.20	1.9
L-(+)-fucose	–	1.5 ± 0.08	1.4	–	3.6 ± 0.18	1.6
D-(+)-ribose	–	1.8 ± 0.04	1.3	–	3.0 ± 0.80	2.1
D-(+)-mannose	–	1.8 ± 0.09	1.3	–	3.8 ± 0.46	1.7
D-(+)-maltose	–	1.7 ± 0.14	1.4	–	3.9 ± 0.03	1.6
D-(+)-galactose	–	1.7 ± 0.05	1.3	–	3.7 ± n.d.	2.1
D-(+)-arabinose	–	1.5 ± 0.04	1.4	–	2.8 ± 0.03	2.0
L-(+)-rhamnose	–	1.4 ± 0.03	1.4	–	3.6 ± 0.09	1.7
D-(+)-trehalose	–	1.5 ± 0.04	1.4	–	3.5 ± 0.07	1.6

Each carbon compound was supplied at a final concentration of 25 μM. Cultures were incubated at 20°C, in the dark. Error bars represent the range of duplicates.

treatments, with a high proportion, 1240 of the 1378 (90%) coding sequences, detected in both treatments (Fig. 3). Nearly all of the most highly expressed genes, including proteorhodopsin, did not differ between treatments (Table S2) and were similar to highly expressed proteins observed in previous proteomic studies of HTCC1062 (Sowell *et al*., 2008a). Similarly, of the 1240 detected genes, only 43 genes (3.5%) were differentially expressed (*t*-test *P*-value ≤ 0.05 and fold change ≥ 1.5) between treatments (Table S2). Averaged over the entire operon, mRNAs encoding glycolysis genes were twice as abundant (2.0 ± 0.6-fold change) when glucose was added to the cultures (Table S2). It was noted that the genes encoding the glyoxylate pathway (*ace*A, *glc*B) were highly expressed in the No Carbon treatment (Table S2), suggesting low-molecular-weight organic

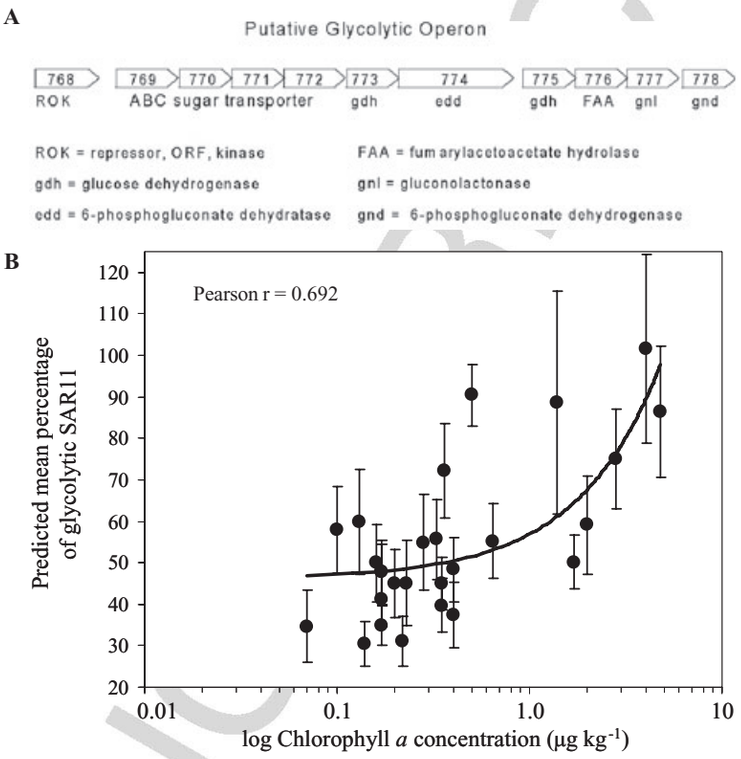

Fig. 2. A. SAR11 Entner–Doudoroff (ED) operon (SAR11_768 to SAR11_778) present in coastal isolates HTCC1062 and HTCC1002 but lacking in HTCC7211.
B. Mean ± SE percentage of SAR11 containing the ED operon derived from metagenomic sequencing libraries collected from a global ocean survey (GOS) plotted as a function of chlorophyll *a* concentrations. The abundance of the ED operon was estimated by averaging the abundance of all individual operon genes in a metagenomic data set. Thus the standard error of the mean reflects the error in whole operon abundance estimates as predicted from individual genes.

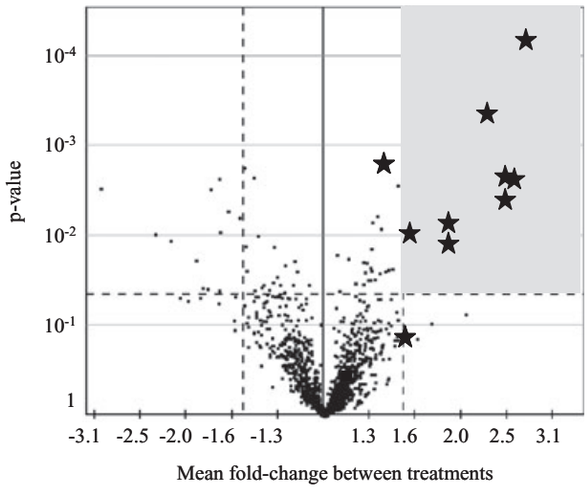

Fig. 3. Volcano plot (*P*-value versus mean fold change) of microarray results obtained from triplicate microarray experiments of HTCC1062 grown on media amended with 25 μM glucose or no primary carbon source. Positive (glucose) or negative (No Carbon) fold-change values indicate carbon source on which transcription was greater. Dashed lines indicate greater than 1.5-fold change and had *P*-values ≤ 0.05. Grey box represents significantly upregulated genes in glucose treatments. SAR11 genes found in the glycolysis operon are indicated by stars.

acids were being utilized as a carbon source in these treatments.

Discussion

Although the absence of glycolysis in aerobic heterotrophic bacteria is not unheard of, the absence of glycolysis genes in some members of the competitively successful SAR11 clade was surprising because glucose is regarded as a high-value carbon source. Radiotracer assays with isolates obtained from the Oregon coast and the Sargasso Sea confirmed that D-glucose was not utilized by the open ocean SAR11 isolate although it could be incorporated and oxidized to CO_2 by the coastal isolates, which had additional genes for glycolysis. However, even in the coastal isolates glucose was a poor carbon source as judged by its influence on the growth of cultures (Fig. 1A). Microarray studies supported the hypothesis that the function of the SAR11 ED operon is glucose oxidation by showing that it is transcribed more highly when glucose was the primary carbon source (Table S2). While the ED operon appears to be inducible, regulatory mechanisms that could explain this observation were not in evidence. Glycolytic bacteria such as *Escherichia coli* rely on a suite of global transcription factors (i.e. *arc*AB, *cra*, *crp*/*cya*A, *fnr* and *mlc*) to regulate gene expression and control carbon metabolism through the EMP and ED pathways (Perrenoud and Sauer, 2005). One of the most important is the transcriptional dual regulator CRP, which binds with cyclic adenosine monophosphate (cAMP) and regulates the expression of nearly 200 different genes (Kolb *et al.*, 1993; Zheng *et al.*, 2004). Glucose lowers cAMP levels in *E. coli*, preventing the induction of alternate carbon metabolisms until glucose is depleted, resulting in the well-known pattern of diauxie observed in batch cultures of *E. coli* and other bacteria (Monod, 1949). In addition, the PTS system also plays a central regulatory role in bacterial carbon metabolism (Deutscher *et al.*, 2006). While the PTS system has been found in other marine bacteria (Delong and Yayanos, 1987; Riemann and Azam, 2002), surprisingly, many of the above regulatory genes as well as the PTS system are lacking from the genomes of marine SAR11 (Table S1). Thus, at this time the underlying genetic mechanisms that regulate the expression of the SAR11 ED operon are unknown.

There are many variations of metabolic pathways for glycolysis, thus the results presented here address an uncertainty that appeared in the first publication describing a SAR11 genome (Giovannoni *et al.*, 2005). That genome (HTCC1062) encoded an unconfirmed variant of the ED pathway but lacked an *eda* gene and other key glycolytic enzymes, making glycolysis a difficult property to predict solely based on bioinformatics. However, our results clearly indicate some isolates of SAR11 can utilize glucose and evidence suggests that a variant of the ED is responsible. We speculate that the SAR11 ED pathway relies on the kinase activity of the ROK domain containing gene SAR11_0768 (Miller and Raines, 2004) to activate the gluconic acid. In addition, the conserved presence of a FAA hydrolase homologue (SAR11_0776) in the SAR11 ED operon suggests that this protein catalyses the cleavage of 2,4-keto-3-deoxy-6-phosphogluconate to 3-phosphoglycerate and pyruvate instead of the usual cleavage of 2-keto-3-deoxy-6-phosphogluconate to 3P-glyceraldehyde and pyruvate by the *eda* protein (Fig. S1). The ability of SAR11 strains to utilize glucose as a carbon source apparently depends on the acquisition of the ED operon, since the Sargasso Sea strain HTCC7211, which lacks this operon, was unable to oxidize glucose (Fig. 2A). In all of the SAR11 genomes we studied key genes are missing from alternative pathways of glucose oxidation – the EMP pathway or the pentose phosphate shunt (Table S1). While this study has confirmed HTCC7211 cannot utilize exogenous D-glucose or other common free carbohydrates, we cannot completely rule out a metabolic role for exogenous carbohydrates in this organism. HTCC7211 contains genes annotated as putative *msm*-like transporters and a gene encoding a putative *alpha*-amylase, suggesting that polysaccharides or other less common carbohydrates which were not examined here may be utilized by this organism. How these compounds would be subsequently oxidized though remains

uncertain given the predicted lesions in the EMP, ED and pentose phosphate pathways.

One of the most powerful uses of metagenomics has been its application to understand natural variation. Therefore, it was fortunate that large GOS databases were available to test hypotheses that emerged from the study of carbohydrate oxidation in the cultured isolates. The metagenomic data supported the hypothesis that the ED operon is indeed a variable property of SAR11 genomes, and furthermore supported the conclusion that this operon is present more frequently in SAR11 genomes from coastal ecosystems. However, even though the genes in the glycolysis operon are under-represented in open ocean metagenomic data sets, homologues of all the genes were present and expressed in the Sargasso Sea (Wilhelm *et al.*, 2007; Sowell *et al.*, 2008b), suggesting a subset of SAR11 in this environment are likely able to utilize glucose to some extent. Thus, our results lend support to the notion that there is likely a range of glucose utilization capabilities among SAR11 ecotypes and shed light onto why previous studies have found variable responses of SAR11 to glucose additions. While we speculate that genetic recombination likely influences the distribution of the glycolysis operon (Vergin *et al.*, 2007; Gilbert *et al.*, 2008), to what extent recombination influences the distribution of the operon remains to be determined.

The results of this study are now the third example of a surprising metabolic deficiency associated with genome reduction in SAR11 cells. The small sizes of SAR11 genomes have been attributed to streamlining selection favouring cell designs that use nutrient resources efficiently. Previous reports have shown that genome reduction occurs at a cost – SAR11 cells require glycine or serine and reduced forms of sulfur because they lack genes and enzymes to synthesize these compounds from precursors (Tripp *et al.*, 2008; 2009). While loss of the ED operon does not result in a unique nutrient requirement, it nonetheless represents a loss of nutritional versatility associated with genome reduction. Some studies have suggested that selection favours a generalist life-history strategy in some marine bacteria (Mou *et al.*, 2008), but our results (Table 1) and other studies (Giovannoni *et al.*, 2008) have indicated that genome streamlining can significantly reduce carbon utilization capabilities compared with marine bacteria with larger genomes. However, our results have demonstrated that SAR11 cells are adept at utilizing a variety of low-molecular-weight organic acids and are predicted to have a complete gluconeogenic pathway (Table S1). We speculate that the formation of carboxylic acids such as oxaloacetate and pyruvate via photolysis (Kieber *et al.*, 1989; Moran and Zepp, 1997; Obernosterer *et al.*, 1999) and the production of acetate and taurine by common marine phytoplankton, zooplankton or other glycolytic marine bacteria (Roberts *et al.*, 1955; Jackson *et al.*, 1992; Visscher *et al.*, 1999) may make these compounds regularly available in the illuminated surface waters of the ocean where high-light-adapted SAR11 Ia ecotypes (Carlson *et al.*, 2009; Treusch *et al.*, 2009) similar to HTCC1062 and HTCC7211 are most abundant.

Our findings raise the question: why in the course of natural selection has glucose oxidation apparently become an expendable property of SAR11 genomes? The answer is not obvious. Glucose is widely used by microbial ecologists to label natural microbial communities because of its important, central role in metabolism. The oxidation of glucose and other carbohydrates produces a relatively large yield of ATP in comparison with organic acids, and many studies have shown that a relatively large fraction of phytoplankton production is carbohydrate synthesis (Morris and Skea, 1978; Behrenfeld *et al.*, 2008). We speculate that the correlation between the frequency of the ED operon in SAR11 cells and chlorophyll content of the water reflects fundamental trends in the value of glucose and other carbohydrates as substrates in different ecosystems. While little information is available to address this issue, it is well known that productive coastal ecosystems are dominated by eukaryotic phytoplankton, whereas picocyanobacteria are the most abundant primary producers in temperate and tropical ocean gyres.

It is interesting to speculate that the decreased frequency of the SAR11 ED operon in oligotrophic environments may be related to the dominance of picocyanobacteria primary producers in those systems. Initial investigations into the composition of DOM produced by *Prochlorococcus* indicated cultures of these organisms released significant amounts of carboxylic acids into the media (Bertilsson and Jones, 2003). Recent time-series studies have reported that biotic interactions appear to play a prominent role in defining the ecological niche of individual marine microorganisms (Fuhrman and Steele, 2008). Hence, it would not be surprising if the most abundant marine autotrophs and most abundant marine heterotrophs have forged a variety of direct or indirect relationships that have yet to be fully appreciated. If this is indeed the case, then the lack of glycolysis in some open gyre SAR11 may be a genomic reflection of one such indirect species interaction. Perhaps future research can address whether it is phytoplankton composition or some other factor that leads to the increased maintenance of the ED operon in SAR11 genomes in productive ocean waters.

Experimental procedures

Genomic and metagenomic analyses

To examine the distribution of the HTCC1062 glycolysis operon in GOS data sets, metagenomic data from 46 GOS

samples (Venter *et al.*, 2004; Rusch *et al.*, 2007) were screened for all genes in the glycolysis operon using a RBB approach (Wilhelm *et al.*, 2007). Briefly, metagenomic data sets were searched for candidate SAR11 glycolysis genes using TBLASTN (*e*-value = 1×10^{-5}). Metagenomic sequences recovered by this initial search were then further screened using RBB filtering with the BLASTP algorithm (*e*-value = 10) and NCBI's non-redundant protein database (NR), which contains predicted protein sequences from genomes but not metagenomic data. Reciprocal best BLAST filtering removes sequences from the analysis that do not retrieve a SAR11 gene as their best BLASTP match. Permissive *e*-values were used in the first search (TBLASTN) to recover homologues, while the second RBB search provided specificity. Generally, RBB searches provide far greater accuracy in 'binning' genes by taxon than do expect score cut-offs, but are not used more often because of the computational demands of the procedure. Sequences retained after RBB filtering were normalized for length as previously described (Biers *et al.*, 2009) and expressed as a percentage of SAR11-like *rec*A genes in corresponding libraries (Venter *et al.*, 2004). Predicted SAR11 *rec*A genes were identified using the same RBB filtering method used to identify predicted SAR11 glycolysis genes. We represent the abundance of the operon as the mean of the individual gene frequencies for all genes in the operon, and plot the standard error of the mean to represent the variability in estimates of the operon abundance.

Growth assays

Seawater for experiments was collected at Newport Hydroline station NH-05 (latitude: 44.65°, longitude: –124.18°) from 10 m depth between September 2007 and June 2008. Water was transported to the laboratory and filtered through a 0.2 μm Supor (Millipore) membrane prior to being autoclaved. Autoclaved seawater was subsequently sparged first with 0.1 μm filtered CO_2 and then with 0.1 μm filtered air for a period of 24 h respectively. In all experiments, autoclaved, sparged seawater medium was amended with excess amounts of (final concentrations): nitrogen (NH_4, 10 μM), inorganic phosphate (KH_2PO_4, 1 μM), reduced sulfur (L-methionine, 1 μM), a glycine/serine source (L-serine, 1 μM), iron chloride ($FeCl_3$, 10 nM) and excess vitamins (final concentrations): thiamine (593 nM), niacin (227 pM), B12 (74 pM), para-amino benzoic acid (PABA, 409 pM), pyridoxine (59 nM), pantothenic acid (81 nM), inositol (555 nM) (Connon and Giovannoni, 2002). To this base seawater medium, carbon compounds were then supplied at final concentrations ranging between 20 and 50 μM as detailed below. All assays were conducted in 250 ml polycarbonate flasks and were incubated at 20°C in the dark for the duration of the experiments. Log phase cells grown on unamended seawater were used as inoculate in all experiments. Cells were quantified by staining with the DNA stain SYBR Green I (Invitrogen) for 1 h, followed by enumeration with a flow cytometer (Guava Technologies) (Tripp *et al.*, 2008).

Carbon utilization growth assays

The effects of organic compounds on cell growth were studied using axenic cultures of HTCC1062 and HTCC7211 grown in duplicate batch of sterile seawater media amended with 20 μM (final concentration) of the individual organic compounds. Cultures were incubated at 20°C in the dark for the duration of the experiments. Resulting final cell densities were compared with control treatments that were amended with inorganic nutrients and vitamins but not an organic carbon compound. Note that seawater media amended with inorganic nutrients and vitamins but no additional organic carbon contained sufficient concentrations of unidentified natural DOC compounds to support heterotrophic SAR11 growth to final cell densities similar to those observed in nature (*c.* 10^6 cells ml^{-1}; Rappe *et al.*, 2002). Hence, when final cell densities in treatments amended with organic compounds exceeded the final cell densities observed in unamended control treatments by 20% or more, the test for growth stimulation by an organic compound was scored as positive. Comparable growth rates among treatments were interpreted to indicate that the added organic compounds were not toxic at these concentrations, regardless of whether the compound stimulated the culture to attain a higher cell density.

Radio tracer experiments

SAR11 cells utilized for radio tracer assays were cultivated in base seawater media that was amended with D-glucose (50 μM, final concentration) as the carbon source. Aliquots (40 ml) of cells were harvested while cultures were in log phase and collected via centrifugation (1.5 h at 43 000 *g*) at 16°C in polycarbonate centrifuge tubes. Efficiency of cell collection ranged between 70% and 90% of initial cell concentrations following this protocol. Following centrifugation, the supernatant was poured off and cells were re-suspended in 5 ml of minimal-salt artificial seawater mix pH 7.8 (ASW) (Moore *et al.*, 2007) and transferred to 15 ml polystyrene centrifuge tubes that had been pre-rinsed in triplicate with ASW prior to use for radio assays. ASW was comprised of the following (final concentration): $NaHCO_3$ (6 mM), NaCl (481 mM), $MgSO_4$ (28 mM), $MgCl_2$ (27 mM), CaCl (10 mM), KCl (9 mM), HEPES (pH 7.5, 1.0 mM), and was autoclaved prior to use. All chemicals were obtained from Sigma-Aldrich.

Glucose-D-[^{14}C (U)] (1 μM, final concentration) and pyruvate-[1-^{14}C] (1 μM, final concentration) radio assays were conducted at room temperature (22°C) for 24 h prior to collection. Radioisotopes were obtained from American Radiolabelled Chemicals Incorporated and were supplied in sterile water. Controls were processed in an identical fashion to live samples but were preserved with formalin (20 μl ml^{-1}) for 1 h prior to the addition of the isotope to the sample. Determination of uptake of radiolabelled substrate was conducted as follows: 2 ml of cells were collected via filtration through a 25 mm GSVP 0.22 μm filter (Millipore) and filters were then rinsed six times with 2 ml aliquots of ASW. The cold TCA-insoluble fraction was determined in a similar fashion except filters were first rinsed in triplicate with 2 ml of ASW then were incubated 2 min in cold 5% TCA followed by six 2 ml rinses with 5% cold TCA. After rinsing, all filters were transferred to scintillation vials containing 5 ml of UltimaGold (Perkin-Elmer) scintillation fluid and allowed to sit overnight prior to being read on a Beckman LS-65000 liquid scintillation counter. The remaining 1 ml of sample was transferred to a

1.5 ml centrifuge tube and placed inside a glass serum vial that also contained an additional 1.5 ml centrifuge tube holding a piece of Whatman filter paper (Type 1) that had been saturated with 200 μl of 2-phenethylamine (Sigma). The sample was then acidified to a pH ~ 2.0 with hydrochloric acid (6 N), quickly sealed and allowed to stand at room temperature overnight to trap $^{14}CO_2$. After incubations, the filter paper was transferred into 5 ml of UltimaGold scintillation fluid and allowed to dissolve for 12 h prior to being read on the liquid scintillation counter. In addition, 100 μl of the acidified sample was placed directly into 5 ml of UltimaGold scintillation fluid to determine the amount of isotope remaining in the sample. Efficiency of counting ranged between 80% and 90% for all live samples and was greater than 70% for formalin-killed controls. On average, 95 ± 8% of the total isotope used was recovered following this protocol.

Microarray experiments

HTCC1062 cells used in microarray experiments were grown in batch cultures containing base seawater media. Glucose treatments were also amended with 25 μM (final concentration) of D-glucose whereas control treatments did not receive an additional carbon source and are referred to as No Carbon treatments. Glucose and No Carbon growth assays were conducted in quadruplicate for each treatment. RNA for microarrays was collected using RNeasy Mini kits (Qiagen) and amplified using the MessageAmp-II Bacteria RNA amplification kit (Ambion) per the manufacturer's instructions. Briefly, 40 ml of cells from each biological replicate were harvested in log phase (time = 192 h) via centrifugation (1 h at 43 000 *g*, 10°C), after which the supernatant was decanted and cell pellets were re-suspended in 1 ml of RNA stabilization reagent (Bacteria Protect, Qiagen) for 15 min. Following RNA stabilization, cells were collected by centrifugation (30 min at 40 000 *g*, 10°C), the supernatant was removed, and cells were then stored frozen at –80°C prior to extraction. RNA extractions began by incubating cells in T.E. lysis buffer (pH 8.0) containing 400 μg ml^{-1} lysozyme for 5 min at room temperature and then were processed via RNeasy Mini kits (Qiagen) to extract total RNA. RNA from all samples was amplified simultaneously within several days of isolation and 10–20 ng of template RNA was used per reaction. Briefly, total RNA from each sample was initially polyadenylated followed by the synthesis of cDNA utilizing a T7-oligo(dT) primer. Biotinylated, antisense RNA (aRNA) was generated from the cDNA via 14 h of *in vitro* transcription in the presence of biotin-11-CTP and biotin-16-UTP nucleotide analogues (Perkin-Elmer). Resulting aRNA was screened for length and quality using a Bioanalyzer 2100 (Agilent) and quantified utilizing a Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific). Biotinylated aRNA (5.5 μg) from triplicate samples was fractionated and then hybridized (45°C) overnight to custom '*Candidatus* Pelagibacter ubique' Affymetrix GeneChip arrays that contained probes for HTCC1002, HTCC1062 and HTCC7211 (Pubiquea520471f) using Affymetrix GeneChip Fluidics Station 450, and Affymetrix GeneChip Hybridization Oven 640. Arrays were then washed following manufacturer's instructions and the resulting images were analysed using an Affymetrix GeneChip Scanner 3000. Background corrections and raw expression values were normalized between chips via application of the robust microarray algorithm (RMA), quantile-normalization and median polish analyses (Bolstad *et al*., 2003; Irizarry *et al*., 2003a,b) in the Affymetrix Expression Console program using the AGCC software package. Resulting data files have been deposited in the NCBI GEO database, reference number GSE14974. Statistical analysis was conducted using the MultiExperiment Viewer (MeV, version 4.2.1) software package available from The Institute for Genomic Research (TIGR). Differences between treatments were deemed significant when mean RMA normalized signal intensities differed statistically (P-value ≤ 0.05) and genes exhibited a 1.5-fold or greater change between treatments.

References

Alonso, C., and Pernthaler, J. (2006) Roseobacter and SAR11 dominate microbial glucose uptake in coastal North Sea waters. *Environ Microbiol* **8:** 2022–2030.

Alonso-Saez, L., Sanchez, O., Gasol, J.M., Balague, V., and Pedros-Alio, C. (2008) Winter-to-summer changes in the composition and single-cell activity of near-surface Arctic prokaryotes. *Environ Microbiol* **10:** 2444–2454.

Aluwihare, L.I., Repeta, D.J., and Chen, R.F. (1997) A major biopolymeric component to dissolved organic carbon in surface sea water. *Nature* **387:** 166–169.

Azam, F., and Malfatti, F. (2007) Microbial structuring of marine ecosystems. *Nat Rev Microbiol* **5:** 782–791.

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

Behrenfeld, M.J., Halsey, K.H., and Milligan, A.J. (2008) Evolved physiological responses of phytoplankton to their integrated growth environment. *Philos Trans R Soc Lond B Biol Sci* **363:** 2687–2703.

Bertilsson, S., and Jones, J.B.J. (2003) Supply of dissolved organic matter to aquatic ecosystems. In *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*. Findlay, S.E.G., and Sinsabaugh, R.L. (eds). New York, USA: Academic Press, pp. 2–24.

Biers, E.J., Sun, S.L., and Howard, E.C. (2009) Prokaryotic genomes and diversity in surface ocean waters: interrogating the global ocean sampling metagenome. *Appl Environ Microbiol* **75:** 2221–2229.

Bolstad, B.M., Irizarry, R.A., Astrand, M., and Speed, T.P. (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19:** 185–193.

Carlson, C.A. (2002) Production and removal processes. In *Biogeochemistry of Marine Dissolved Organic Matter*. Hansell, D.A., and Carlson, C.A. (eds). New York, USA: Academic Press, pp. 91–151.

Carlson, C.A., Ducklow, H.W., Hansell, D.A., and Smith, W.O. (1998) Organic carbon partitioning during spring phytoplankton blooms in the Ross Sea polynya and the Sargasso Sea. *Limnol Oceanogr* **43:** 375–386.

Carlson, C.A., Giovannoni, S.J., Hansell, D.A., Goldberg, S.J., Parsons, R., and Vergin, K. (2004) Interactions among dissolved organic carbon, microbial processes, and community structure in the mesopelagic zone of the northwestern Sargasso Sea. *Limnol Oceanogr* **49:** 1073–1083.

Carlson, C.A., Morris, R., Parsons, R., Treusch, A.H., Giovannoni, S.J., and Vergin, K. (2009) Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *ISME J* **3:** 283–295.

Connon, S.A., and Giovannoni, S.J. (2002) High-throughput methods for culturing microorganisms in very-low-nutrient media yield diverse new marine isolates. *Appl Environ Microbiol* **68:** 3878–3885.

Delong, E.F., and Yayanos, A.A. (1987) Properties of the glucose transport system in some deep-sea bacteria. *Appl Environ Microbiol* **53:** 527–532.

Deutscher, J., Francke, C., and Postma, P.W. (2006) How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiol Mol Biol Rev* **70:** 939–1031.

Elifantz, H., Malmstrom, R.R., Cottrell, M.T., and Kirchman, D.L. (2005) Assimilation of polysaccharides and glucose by major bacterial groups in the Delaware Estuary. *Appl Environ Microbiol* **71:** 7799–7805.

Fuhrman, J.A., and Steele, J.A. (2008) Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquat Microb Ecol* **53:** 69–81.

Gilbert, J.A., Muhling, M., and Joint, I. (2008) A rare SAR11 fosmid clone confirming genetic variability in the '*Candidatus* Pelagibacter ubique' genome. *ISME J* **2:** 790–793.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., *et al.* (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309:** 1242–1245.

Giovannoni, S.J., Hayakawa, D.H., Tripp, H.J., Stingl, U., Givan, S.A., Cho, J.C., *et al.* (2008) The small genome of an abundant coastal ocean methylotroph. *Environ Microbiol* **10:** 1771–1782.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003a) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* **31:** e15.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003b) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **4:** 249–264.

Jackson, A.E., Ayer, S.W., and Laycock, M.V. (1992) The effect of salinity on growth and amino acid composition in the marine diatom *Nitzschia pungens*. *Can J Bot* **70:** 2198–2201.

Keil, R.G., and Kirchman, D.L. (1999) Utilization of dissolved protein and amino acids in the northern Sargasso Sea. *Aquat Microb Ecol* **18:** 293–300.

Kieber, D.J., McDaniel, J., and Mopper, K. (1989) Photochemical source of biological substrates in sea water: implications for carbon cycling. *Nature* **341:** 637–639.

Kirchman, D.L. (2003) The contribution of monomers and other low-molecular weight compounds to the flux of dissolved organic material in aquatic ecosystems. In *Aquatic Ecosystems: Interactivity of Dissolved Organic Matter*. Findlay, S.E.G., and Sinsabaugh, R.L. (eds). New York, USA: Academic Press, pp. 217–243.

Kolb, A., Busby, S., Buc, I.I., Garges, S., and Adhya, S. (1993) Transcriptional regulation by cAMP and its receptor protein. *Annu Rev Biochem* **62:** 749–797.

Kujawinski, E.B., Del Vecchio, R., Blough, N.V., Klein, G.C., and Marshall, A.G. (2004) Probing molecular-level transformations of dissolved organic matter: insights from electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. *Mar Chem* **92:** 23–37.

Malmstrom, R.R., Kiene, R.P., Cottrell, M.T., and Kirchman, D.L. (2004) Contribution of SAR11 bacteria to dissolved dimethylsulfoniopropionate and amino acid uptake in the North Atlantic ocean. *Appl Environ Microbiol* **70:** 4129–4135.

Malmstrom, R.R., Cottrell, M.T., Elifantz, H., and Kirchman, D.L. (2005) Biomass production and assimilation of dissolved organic matter by SAR11 bacteria in the Northwest Atlantic Ocean. *Appl Environ Microbiol* **71:** 2979–2986.

Miller, B.G., and Raines, R.T. (2004) Identifying latent enzyme activities: substrate ambiguity within modern bacterial sugar kinases. *Biochemistry* **43:** 6387–6392.

Monod, J. (1949) The growth of bacterial cultures. *Annu Rev Microbiol* **3:** 371–394.

Moore, L.R., Coe, A., Zinser, E.R., Saito, M.A., Sullivan, M.B., Lindell, D., *et al.* (2007) Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol Oceanogr-Methods* **5:** 353–362. 12

Moran, M.A., and Zepp, R.G. (1997) Role of photoreactions in the formation of biologically labile compounds from dissolved organic matter. *Limnol Oceanogr* **42:** 1307–1316.

Morris, L., and Skea, W. (1978) Products of photosynthesis in natural populations of marine phytoplankton from the Gulf of Maine. *Mar Biology* **47:** 303–312.

Morris, R.M., Rappe, M.S., Connon, S.A., Vergin, K.L., Siebold, W.A., Carlson, C.A., and Giovannoni, S.J. (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420:** 806–810.

Morris, R.M., Vergin, K.L., Cho, J.C., Rappe, M.S., Carlson, C.A., and Giovannoni, S.J. (2005) Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-Series Study site. *Limnol Oceanogr* **50:** 1687–1696.

Mou, X.Z., Sun, S.L., Edwards, R.A., Hodson, R.E., and Moran, M.A. (2008) Bacterial carbon processing by generalist species in the coastal ocean. *Nature* **451:** 708–U704. 13

Obernosterer, I., Kraay, G., de Ranitz, E., and Herndl, G.J. (1999) Concentrations of low molecular weight carboxylic acids and carbonyl compounds in the Aegean Sea (Eastern Mediterranean) and the turnover of pyruvate. *Aquat Microb Ecol* **20:** 147–156.

Perrenoud, A., and Sauer, U. (2005) Impact of global transcriptional regulation by *arcA*, *arcB*, *cra*, *crp*, *cya*, *fnr*, and *mlc* on glucose catabolism in *Escherichia coli*. *J Bacteriol* **187:** 3171–3179.

Repeta, D.J., and Aluwihare, L.I. (2006) Radiocarbon analysis of neutral sugars in high-molecular-weight dissolved organic carbon: implications for organic carbon cycling. *Limnol Oceanogr* **51:** 1045–1053.

Rich, J.H., Ducklow, H.W., and Kirchman, D.L. (1996) Concentrations and uptake of neutral monosaccharides along 140 degrees W in the equatorial Pacific: contribution of glucose to heterotrophic bacterial activity and the DOM flux. *Limnol Oceanogr* **41:** 595–604.

Riemann, L., and Azam, F. (2002) Widespread *N*-acetyl-D-glucosamine uptake among pelagic marine bacteria and its ecological implications. *Appl Environ Microbiol* **68:** 5554–5562.

Roberts, R.B., Abelson, P.H., Cowie, D.B., Bolton, E.T., and Britten, R.J. (1955) *Studies of Biosynthesis in* Escherichia coli. Washington, DC, USA: Carnegie Institution of Washington Publication 607.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., *et al.* (2007) The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biol* **5:** 398–431.

Skoog, A., and Benner, R. (1997) Aldoses in various size fractions of marine organic matter: implications for carbon cycling. *Limnol Oceanogr* **42:** 1803–1813.

Skoog, A., Biddanda, B., and Benner, R. (1999) Bacterial utilization of dissolved glucose in the upper water column of the Gulf of Mexico. *Limnol Oceanogr* **44:** 1625–1633.

Smith, D.C., Simon, M., Alldredge, A.L., and Azam, F. (1992) Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. *Nature* **359:** 139–142.

Sowell, S.M., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Callister, S.J., Smith, R.D., *et al.* (2008a) Proteomic analysis of stationary phase in the marine bacterium '*Candidatus* Pelagibacter ubique'. *Appl Environ Microbiol* **74:** 4091–4100.

Sowell, S.M., Wilhelm, L.J., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Barofsky, D.F., *et al.* (2008b) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* **••:** ••–••.

Stingl, U., Tripp, H.J., and Giovannoni, S.J. (2009) The genome sequence of SAR11 isolate 7211 (in preparation).

Suttle, C., Chan, A.M., and Fuhrman, J.A. (1991) Dissolved free amino-acids in the Sargasso Sea – uptake and respiration rates, turnover times and concentrations. *Mar Ecol Prog Series* **70:** 189–199.

Treusch, A.H., Vergin, K.L., Finlay, L.A., Donatz, M.G., Burton, R.M., Carlson, C.A., and Giovannoni, S.J. (2009) Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME* (in press).

Tripp, H.J., Kitner, J.B., Schwalbach, M.S., Dacey, J.W.H., Wilhelm, L.J., and Giovannoni, S.J. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* **452:** 741–744.

Tripp, H.J., Schwalbach, M.S., Meyer, M.M., Kitner, J.B., Breaker, R.R., and Giovannoni, S.J. (2009) Unique glycine-activated riboswitch linked to glycine-serine auxotrophy in SAR11. *Environ Microbiol* **11:** 230–238.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304:** 66–74.

Vergin, K.L., Tripp, H.J., Wilhelm, L.J., Denver, D.R., Rappe, M.S., and Giovannoni, S.J. (2007) High intraspecific recombination rate in a native population of *Candidatus* Pelagibacter ubique (SAR11). *Environ Microbiol* **9:** 2430–2440.

Visscher, P.T., Gritzer, R.F., and Leadbetter, E.R. (1999) Low-molecular-weight sulfonates, a major substrate for sulfate reducers in marine microbial mats. *Appl Environ Microbiol* **65:** 3272–3278.

Wilhelm, L., Tripp, H.J., Givan, S., Smith, D., and Giovannoni, S. (2007) Natural variation in SAR11 marine bacterioplankton genomes inferred from metagenomic data. *Biol Direct* **2:** 27.

Williams, P.J.L.B. (2000) Heterotrophic bacteria and the dynamics of dissolved organic material. In *Microbial Ecology of the Oceans*. Kirchman, D.L. (ed.). New York, USA: Wiley-Liss, pp. 153–200.

Zheng, D., Constantinidou, C., Hobman, J.L., and Minchin, S.D. (2004) Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res* **32:** 5874–5893.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Proposed Entner–Doudoroff pathway found in some members of the SAR11 clade. Gene names are for HTCC1062.

Table S1. Comparison of central metabolism in SAR11 isolates HTCC1002, HTCC1062 and HTCC7211. Due to predicted lesions in pentose phosphate shunt, reactions yielding NADPH are indicated with a 'NADPH'.

Table S2. Transcription profile of HTCC1062 when grown using glucose as the primary carbon source compared with No Carbon control treatments. Positive (glucose) or negative (No Carbon) fold-change values indicate carbon source on which transcription was greater. Most highly detected transcripts are also shown. Genes in the SAR11 ED operon are indicated in bold.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Toppan Best-set Premedia Limited	
Journal Code: EMI	Proofreader: Jason
Article No: 2092	Delivery date: 13 October 2009
Page Extent: 11	Copyeditor: Harry

AUTHOR QUERY FORM

Dear Author,

During the preparation of your manuscript for publication, the questions listed below have arisen. Please attend to these matters and return this form with your proof.

Many thanks for your assistance.

Query References	Query	Remark
q1	AUTHOR: A running head short title was not supplied; please check if this one is suitable and, if not, please supply a short title that can be used instead.	
q2	AUTHOR: Please provide Tel. and Fax details for corresponding author	
q3	AUTHOR: **Suttle 1991** has been changed to **Suttle et al. 1991** so that this citation matches the Reference List. Please confirm that this is correct	
q4	AUTHOR: (Carlson *et al.*, unpublished) Please provide the initial and surname of all the authors of the unpublished data	
q5	AUTHOR: In the following text 'Entner-Doudoroff' has been abbreviated as 'ED' throughout for consistency. Is this OK	
q6	AUTHOR: Should 'msm' be italicized for consistency?	
q7	AUTHOR: 'principle' has been changed to 'principal'. Is this OK?	
q8	AUTHOR: (latitude: 44.65°, longitude: QUOTE -124.18°) Please use the form like 44°56′N, 71°45′E	
q9	AUTHOR: 'dissolved organic carbon compounds (DOC)' has been changed to 'DOC compounds'. Is this OK?	
q10	AUTHOR: **Rappe et al. 2002** has not been included in the Reference List, please supply full publication details	
q11	AUTHOR: 'where as' has been changed to 'whereas'. Is this OK?	
q12	AUTHOR: Please confirm that Journal title *Limnol Oceanogr-Methods* has been abbreviated correctly	
q13	AUTHOR: Please confirm the page range in Reference Mou et al. 2008 is correct	
q14	AUTHOR: (Roberts et al. 2005) 'Carnegie Institution of Washington Publication 607' has been set as the publisher. Is this OK?	
q15	AUTHOR: (Sowell et al. 2008b) Please provide the volume number and page range	

q16	AUTHOR: (Stingl et al. 2009) Papers that have not yet been accepted for publication should not be included in the Reference List. Please update this reference if it has now been published, use 'in press' (with journal title details, if appropriate) if it has been accepted for publication, or remove it from Reference List and change to 'U. Stingl, H.J. Tripp and S.J. Giovannoni, in preparation' in the text if it has not yet been accepted.	
q17	AUTHOR: (Treusch et al. 2009) If this reference has now been published online, please add relevant year/DOI information. If this reference has now been published in print, please add relevant volume/page/year information	
q18	AUTHOR: 'no carbon' has been changed to 'No Carbon' for consistency. Is this OK?	
q19	AUTHOR: 'stdev' has been changed to 'SD'. Is this OK?	
q20	AUTHOR: 'standard deviation' has been abbreviated as 'SD' for consistency. Is this OK?	
q21	AUTHOR: 'stderr' has been changed to 'SE'. Is this OK?	
q22	AUTHOR: 'no carbon' has been changed to 'No Carbon' for consistency. Is this OK?	
q23	AUTHOR: Please confirm the format of the table is OK	
q24	AUTHOR: 'No carbon' has been changed to 'No Carbon' for consistency. Is this OK?	

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	··· under matter to remain	New matter followed by or or
Insert in text the matter indicated in the margin		
Delete	through single character, rule or underline or through all characters to be deleted	or
Substitute character or substitute part of one or more word(s)	through letter or through characters	new character or new characters
Change to italics	under matter to be changed	
Change to capitals	under matter to be changed	
Change to small capitals	under matter to be changed	
Change to bold type	under matter to be changed	
Change to bold italic	under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Change bold to non-bold type	(As above)	
Insert 'superior' character	through character or where required	or under character e.g. or
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	
Insert single quotation marks	(As above)	or and/or or
Insert double quotation marks	(As above)	or and/or or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking characters	
Insert or substitute space between characters or words	through character or where required	
Reduce space between characters or words	between characters or words affected	