High-Throughput Isotopic Analysis of RNA Microarrays to Quantify Microbial Resource Use

Xavier Mayali[1], Peter K Weber[1], Eoin L Brodie[2], Shalini Mabery[1], Paul D Hoeprich[1], and Jennifer Pett-Ridge[1]

[1]Physical and Life Science Directorate, Lawrence Livermore National Laboratory, Livermore, CA, USA and [2]Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Most microorganisms remain uncultivated, and typically their ecological roles must be inferred from diversity and genomic studies. To directly measure functional roles of uncultivated microbes, we developed Chip-stable isotope probing (SIP), a high-sensitivity, high-throughput SIP method performed on a phylogenetic microarray (chip). This approach consists of microbial community incubations with isotopically labeled substrates, hybridization of the extracted community rRNA to a microarray and measurement of isotope incorporation—and therefore substrate use—by secondary ion mass spectrometer imaging (NanoSIMS). Laboratory experiments demonstrated that Chip-SIP can detect isotopic enrichment of 0.5 atom % ^{13}C and 0.1 atom % ^{15}N, thus permitting experiments with short incubation times and low substrate concentrations. We applied Chip-SIP analysis to a natural estuarine community and quantified amino acid, nucleic acid or fatty acid incorporation by 81 distinct microbial taxa, thus demonstrating that resource partitioning occurs with relatively simple organic substrates. The Chip-SIP approach expands the repertoire of stable isotope-enabled methods available to microbial ecologists and provides a means to test genomics-generated hypotheses about biogeochemical function in any natural environment.

Subject Category: microbial ecology and functional diversity of natural habitats
Keywords: biogeochemistry; marine; microarray; microbial function; NanoSIMS; stable isotope probing

Introduction

Microorganisms, most of which remain uncultivated, control the biogeochemical cycles of carbon and nitrogen in all of earth's biomes, from the human gut to the deep subsurface (Whitman et al.,1998). Identifying the microbes responsible for elemental cycling processes, and in particular the factors affecting C and N cycling rates, is a critical precursor to the development of predictive models of microbial responses to environmental perturbations (such as, pollution, nutrient inputs or climate change). The advent of 16S rRNA sequencing and environmental genomics has revolutionized microbial ecology by identifying the genetic diversity (Rappé et al., 1997) and functional potential (DeLong et al., 2006) of a community without the need for cultivation. In particular, the comparative "omics" strategy, relying on sequence comparisons to infer biogeochemical activity (Beja et al., 2001; Treusch et al., 2005; Poretsky et al., 2010), has been enormously fruitful in providing information about ecosystem function for entire ecosystem function communities. Yet it is clear that although molecular surveys that describe the natural history of complex microbial systems can intimate which functions specific microorganisms or populations may perform, hypothesis-driven manipulative experiments are a critical follow-up step to enable quantitative associations between function and identity.

A suite of "function-identity" methods has been developed in the past decade to interrogate microbes that cannot be isolated or cultivated (Gutierrez-Zamora and Manefield, 2010). Many of these culture-independent approaches have ideal qualities such as high sensitivity or *in situ* resolution (for example, ^{13}C-PLFA (phospholipids fatty acid; Boschker et al., 1998); EL FISH (element labeling fluorescent in situ hybridization; Behrens et al., 2008), FISH MAR (FISH microautoradiography; Ouverney and Fuhrman, 1999), isotope arrays (Adamczyk et al., 2003)). Nucleic acid (NA) stable isotope probing (SIP) techniques (Radajewski et al., 2000; Manefield et al., 2002) are the most widely used of these means to directly connect substrate utilization to microbial identity. For SIP, natural microbial communities are incubated in the presence of a substrate enriched in a rare stable isotope (~~typically~~ ^{13}C or ^{15}N). Select organisms incorporate the substrate and they, and their NAs,

Correspondence: X Mayali or J Pett-Ridge, Physical and Life Science Directorate, Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA.
E-mail: mayali1@llnl.gov or pettridge2@llnl.gov

become isotopically enriched over time. Ultracentrifugation is used to separate isotopically enriched DNA or RNA from lighter, unenriched NAs for subsequent molecular analysis. In the past decade, SIP approaches have generated many advances in the understanding of microbial pollutant degradation, plant–microbe interactions and food web dynamics (for a comprehensive review, see Murrell and Whiteley (2010)). RNA-SIP is a particularly attractive approach, as it targets active, although not necessarily dividing populations, and may allow for faster label incorporation (Manefield *et al.*, 2002). However, as with any approach, there are advantages and disadvantages for SIP experiments; potential concerns include fertilization effects and risk of community cross-feeding, as well as RNA behavior during gradient separation (Lueders *et al.*, 2004). In some experiments, high substrate concentrations are required to meet the sensitivity threshold of density gradient separation: >20% ^{13}C and >40% ^{15}N for DNA-SIP (Uhlík *et al.*, 2009).

We sought to develop and apply a variation of traditional RNA-SIP, one that would combine its ideal qualities with higher throughput, sensitivity, taxonomic resolution and allow combined and quantitative ^{15}N and ^{13}C analyses. In commonly used SIP techniques, NAs are separated first by their isotopic composition and then characterized phylogenetically. With Chip-SIP, we use a high-density microarray to phylogenetically sort the sample into individual rRNAs, and then determine their isotopic enrichment with quantitative secondary ion mass spectrometry imaging using a Cameca NanoSIMS 50. Direct NanoSIMS analysis is made possible by implementing a new surface chemistry for synthesis of DNA on conductive material. Adopting this approach, thousands of unique phylogenetic probes assaying hundreds of taxa can be quickly analyzed from a single sample.

Materials and methods

Growth of single strains and incubation of field samples

Strains of *Pseudomonas stutzeri* ATTC 11607, *Vibrio cholerae* ATCC 14104 and *Bacillus cereus* D17 were grown from −80 °C frozen stocks in Luria-Bertani broth at 37 °C, transferred to the M9 minimal medium and then inoculated into M9 enriched in ^{13}C glucose (99% ^{13}C, uniformly labeled, Omicron Biochemicals, Inc., South Bend, IN, USA) and/or ^{15}N ammonium (98% ^{15}N, Sigma-Aldrich, St Louis, MO, USA) at isotope concentrations ranging from 0.1% to 100%. An enrichment of 10% ^{13}C indicates that 10% of the media's glucose was 99% enriched in ^{13}C, whereas 90% was natural abundance carbon (1.1% ^{13}C and 98.9% ^{12}C). Cells were centrifuged, washed and frozen at −80 °C. Bulk measurements (by isotope ratio mass spectrometry) showed that *Pseudomonas* cells grown in full ^{13}C glucose were enriched between 680 000 and 900 000 per mil, equivalent to 88–91 atm%.

For field experiments, surface water was collected from the San Francisco (SF) Bay at the public pier in Berkeley (CA, USA) (37°51′46.67″N, 122°19′3.23″W) and immediately brought back to the laboratory in a cooler. Glass bottles (500 ml) were filled without air space and dark incubated at 14 °C. In this first set of estuarine water experiments, samples were simultaneously incubated with relatively high substrate concentrations: 50 μM ^{13}C glucose and 200 μM ^{15}N ammonium. Microbial communities were harvested after 24 h by filtration through a 0.22-μM polycarbonate filter that was then immediately frozen at −80 °C. Background concentrations of ammonium in the SF Bay range from 1 to 14 μM (Dugdale *et al.*, 2007); typical estuarine glucose concentrations are 5–100 nM (Hanson and Snyder, 1980). In a second SF Bay experiment, water samples were incubated as described above, but for less time (12 h) and lower labeled substrate concentrations. Incubations were performed with either 8 μM mixed amino acids (AAs) (a mixture of 14 AAs, 99 atm % ^{15}N labeled; Cambridge Isotope Laboratories, Andover, MA, USA), 500 μg l^{-1} fatty acids (FAs) (99 atm % ^{13}C; Cambridge Isotope Laboratories) or 50 μg l^{-1} NAs (90 atm % ^{13}C; RNA extract from fully labeled ^{13}C *P. stutzeri* cultures). Water samples were then filtered as described above and frozen at −80 °C. Substrate additions were designed to approach what is typically measured in estuarine environments: 2–7 μM AAs (Evens and Braven, 1988), 25 μg l^{-1} FAs (Stauffer and Macintyre, 1970) and 10 μg l^{-1} DNA (DeFlaun *et al.*, 1987).

RNA extraction and labeling

RNA from pelleted cells (laboratory cultures) and filters (field samples) was extracted using the Qiagen RNEasy kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. RNA samples were split: one fraction saved for fluorescent labeling (see below) and the other kept unlabeled for NanoSIMS microarray analysis. This procedure was used because the fluorescent labeling protocol introduces substantial ^{12}C background carbon that dilutes the ^{13}C signal of the target RNA. Alexa Fluor 532 labeling was performed using the Ulysis kit (Life Technologies, Carlsbad, CA, USA) for 10 min at 90 °C (1 μg RNA, 10 μl labeling buffer, 2 μl Alexa Fluor reagent), followed by fragmentation. All RNAs (fluorescently labeled or not) were fragmented before hybridization using 1X fragmentation buffer (Affymetrix, Cleveland, OH, USA) for 10 min at 90 °C. Alexa Fluor-labeled RNA was purified using a Spin-OUT minicolumn (Millipore, Billerica, MA, USA), and all RNA were concentrated by isopropanol precipitation to a final concentration of 500 ng μl^{-1}.

Target taxa selection and de novo *probe design*

To identify taxa present in our SF Bay samples and to design a corresponding microarray probe set,

RNA extracts were treated with DNAse I and reverse-transcribed to produce cDNA using the Genechip Expression 3′ amplification one-cycle cDNA synthesis kit (Affymetrix). The cDNA was PCR amplified with universal bacterial and archaeal primers (Brodie *et al.*, 2006), fragmented, fluorescently labeled and hybridized to a G2 PhyloChip to establish 16S rDNA community taxonomic composition. A taxon was considered present if $>90\%$ of probes for that taxon were responsive (signal of perfect match probe >1.3 times the signal from mismatch probe, and the intensity difference between the two (perfect match probemismatch probe) >130 times the squared noise value). From ~ 1500 positively identified taxa, we targeted a subset of 100 taxa commonly found in marine environments. We chose to exclude taxa previously identified from soil, sewage and bioreactors. Using the Greengenes database (DeSantis *et al.*, 2006) implemented in ARB (Ludwig *et al.*, 2004), we designed 25 non-redundant probes per taxon (25-bp long), to create a distinct probe set for each (Supplementary Table S1). Probes were designed to minimize cross-hybridization potential. Probes for single laboratory strains (*P. stutzeri*, *B. cereus* and *V. cholerae*) were also designed using ARB (Supplementary Table S1). Probe sets were imported into ArrayScribe software (NimbleGen Systems Inc., Madison, WI, USA) and arranged in a densely packed formation.

Microarray synthesis and hybridization

Standard microarrays are not conductive which results in charge buildup during SIMS analysis. Thus, for Chip-SIP, we used glass slides coated with 150–300 Å of indium-tin oxide (ITO; Sigma cat no. 576352), and then coated with either Super Epoxy 2 silane (Arrayit Corporation, Sunnyvale, CA, USA) or an alkyl phosphonate hydroxy-linker (Pett-Ridge *et al.*, 2011) to provide a substrate for DNA synthesis. Using these slides and the ArrayScribe probe design files, microarrays (spot size $=17\,\mu$M) were synthesized using a photolabile deprotection strategy (Singh-Gasson *et al.*, 1999) on a Nimblegen Maskless Array Synthesizer (Roche, Madison, WI, USA). For quality control, slides were hybridized with complimentary *Arabidopsis* calmodulin protein kinase 6 labeled with Cy3 (Integrated DNA Technologies, Coralville, IA, USA), which hybridizes to fiducial marks (probe spots with the complementary sequence synthesized throughout the array area). For array hybridization, 1 μg RNA samples in 12 μl 1X hybridization buffer (Roche) were placed in Nimblegen X4 mixer slides and incubated inside a Maui hybridization system (BioMicro Systems, Salt Lake City, UT, USA) for 18 h at 42 °C, washed according to manufacturer's instructions (Roche) and dried. Arrays with fluorescently labeled RNA were imaged using a Genepix 4000B fluorescence scanner (Molecular Devices, Sunnyvale, CA, USA) at pmt $=650$ Units. On non-Alexa Fluor RNA arrays, corner areas of the array were marked with a diamond pen before imaging using the fluorescence scanner. This reference image allows navigation to the hybridized locations in the NanoSIMS, as both the diamond pen scratches and the array's fiducial probe spots (outlining the entire array area) are evident in the fluorescence image.

NanoSIMS analyses

To avoid dilution or loss of hybridized RNA, arrays were not metal coated before NanoSIMS analysis, but were trimmed with a glass scribe, and mounted in a $1'' \times 1.5''$ rectangular aperture within a stainless steel holder. NanoSIMS analysis of the ITO microarrays hybridized with ^{13}C and/or ^{15}N rRNA was performed at LLNL with a Cs+ primary ion beam in a Cameca NanoSIMS 50 (Cameca, Gennevilliers, France). Carbon and nitrogen isotopic ratios were determined by electrostatic peak switching on electron multipliers in pulse counting mode, alternately measuring $^{12}C^{14}N^-$ and $^{12}C^{15}N^-$ simultaneously for the $^{15}N/^{14}N$ ratio, and then simultaneously measuring $^{12}C^{14}N^-$ and $^{13}C^{14}N^-$ for the $^{13}C/^{12}C$ ratio. We used this strategy because the secondary ion count rate for the CN^- species in these samples is typically 5–10 times higher than that for the other carbon species (such as, C^-, CH^-, C_2^-), and therefore higher precision was achieved. If only one isotopic ratio was of interest, peak switching was not performed. On the basis of total counts for the analyzed cycles, we achieved precision of 2–3% for $^{13}C^{14}N$ and 1–4% for $^{15}N^{12}C$, depending on enrichment and hybridization intensity (see Supplementary Information for additional NanoSIMS analysis details).

A single microarray analysis of ~ 2500 probes, with an area of 0.75 mm^2 and the acquisition of three hundred 50-μM^2 images, was carried out using the Cameca software automated chain analysis in 16 h. Ion images were stitched together and processed to generate isotopic ratios with custom software (LIMAGE, L Nittler, Carnegie Institution of Washington). Hybridization locations were selected by hand or with the auto-region of interest function, and ratios were calculated for each selected region over all cycles. Isotopic ratios were converted to delta values using $\delta = ((R_{meas}/R_{standard}) - 1) \times 1000$, where R_{meas} is the measured ratio and $R_{standard}$ the standard ratio (0.00367 for $^{15}N/^{14}N$ and 0.011237 for $^{13}C/^{12}C$). Data were corrected for natural abundance ratios measured in unhybridized locations of the ITO microarray.

Data analyses

For each taxon, isotopic enrichment of individual probe spots was plotted versus probe fluorescence, and a linear regression slope was calculated with the

y intercept constrained to natural isotope abundances (0 per mil for ^{15}N data and -20 per mil for ^{13}C data). Spots with saturated fluorescent signal were removed before these calculations. This calculated slope (per mil/fluorescence), which we refer to as hybridization-corrected enrichment (HCE), is a metric that can be used to compare the relative incorporation of a given substrate by different taxa. Owing to the different natural concentrations of ^{13}C and ^{15}N, and more importantly, different background contributions from the microarray surface, HCEs for ^{15}N substrates and ^{13}C substrates are not comparable. However, the HCE values for multiple taxa exposed to multiple ^{13}C substrates may be compared for statistically significant differences.

With the results from our multiple substrate addition experiment in SF Bay water, we constructed a network diagram of substrate use using Cytoscape software (Cline *et al.*, 2007). Only taxa that showed HCEs with s.e. not overlapping with zero and that had >30 per mil enrichment were considered. For analyses of marine bacterial genomic information, genomes of marine bacterial isolates were selected from the DOE Joint Genome Institute's Integrated Microbial Genomes (Markowitz *et al.*, 2010) Database and gene annotation fields were searched for terms related to AA, FA and nucleoside transporters and extracellular nucleases. For phylogenetic relationships, the Greengenes 16S rRNA database (DeSantis *et al.*, 2006) was imported into ARB (Ludwig *et al.*, 2004) and all taxa except targets of the array analysis were removed with the taxon pruning function. To test whether substrate utilization had a phylogenetic signal, each qualitative trait (AA, FA, NA utilization) was randomized (1000 replicates) onto the phylogeny and parsimony scores were calculated for each replicate. These analyses were performed using the Mesquite software package (Maddison and Maddison, 2010). The parsimony score of the real data was compared with this null distribution and was considered significantly different (with 95% confidence) if >2 s.d. away from the random distribution mean.

Results and Discussion

Controlled laboratory experiments

As an initial test of the Chip-SIP approach, we grew a single bacterial strain (*P. stutzeri*) in a minimal medium with ^{13}C-glucose as the sole carbon source and extracted its RNA. After fluorescent labeling, the RNA was hybridized to a microarray probe set consisting of 60 sequences targeting different regions of the *P. stutzeri* 16S rRNA gene. Isotopic enrichment of these probe spots (Figure 1a) strongly depended on the efficiency of target RNA hybridization, as quantified by fluorescence (Figure 1b). This correlation is the result of dilution of the target RNA isotopic signal by the background of unenriched oligonucleotide probes. Thus, if less target RNA hybridizes to the array surface, higher dilution results in a lower isotopic enrichment measurement. Relative isotopic enrichment of RNA from an organism can be quantified based on the slope of the enrichment:fluorescence relationship for a single probe set (HCE; Figure 1c).

We further investigated the reproducibility of the Chip-SIP method by comparing results from multiple independent analyses of *P. stutzeri* RNA. First, we analyzed replicate regions of the same hybridization experiment, revealing good reproducibility ($R^2=0.88$, $P<0.0001$; Supplementary Figure 1a). This shows that RNA from the same extraction,

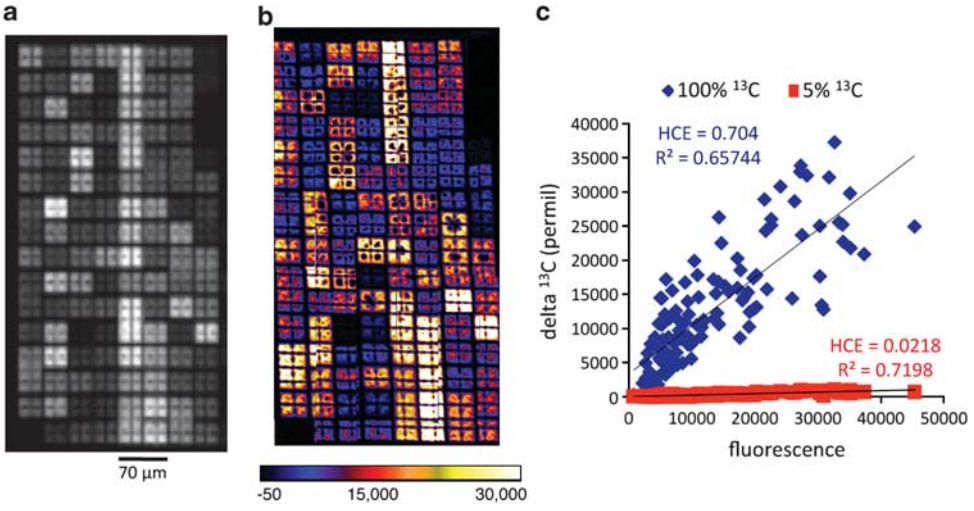

Figure 1 (**a**) Fluorescence image and (**b**) NanoSIMS ^{13}C isotope ratio image montage of an ITO microarray hybridized with RNA from a single bacterial strain (*Pseudomonas stutzeri*) grown on ^{13}C glucose; (**c**) the relationship between fluorescence and enrichment for each probe spot (referred to as 'hybridization corrected enrichment', HCE) is significant and different between RNA from cultures with 100% ^{13}C (blue diamonds)- versus 5% ^{13}C-enriched cultures (red squares). Units for the NanoSIMS color scale are per mil (‰), and range from natural abundance (black) to highly enriched (white).

hybridized simultaneously on the same replicate array, but analyzed on a different date by NanoSIMS, resulted in statistically similar data. We also compared hybridizations of different batches of *P. stutzeri* cultures, grown at 5% and 100% levels of ^{13}C, also analyzed by NanoSIMS on different dates, and found similarly good reproducibility ($R^2 = 0.78$, $P < 0.0001$; Supplementary Figure 1b). To show that enrichment with different isotopes does not affect hybridization kinetics or detection, we compared Chip-SIP data from ^{13}C- and ^{15}N-labeled *P. stutzeri* RNA, again, finding a similar level of reproducibility ($R^2 = 0.75$, $P < 0.0001$; Supplementary Figure 1c).

Before applying Chip-SIP to natural communities, we sought to test its sensitivity and ability to discriminate a mixture of differentially labeled bacterial taxa in controlled mixtures. Bacterial strains *P. stutzeri*, *V. cholerae* and *B. cereus* were grown separately to ^{15}N and ^{13}C isotopic enrichments ranging from 0.1% to 100%; mixtures of their combined RNA were then hybridized to ITO arrays with probe sets targeting each organism. Both ^{13}C (Figure 2a) and ^{15}N (Figure 2b) enrichment can easily be distinguished for each of two taxa based on their respective HCE values, and are significantly different (analysis of covariance; $P < 0.0001$). These experiments showed that by integrating the results from each organism's probe set (10–20 probes per taxon), HCE values allow the direct comparison of isotopic incorporation between two or more taxa on a single array. Positive detection was defined in which the HCE slope (minus its s.e.) was >0. Notably, we successfully detected isotopic enrichments as low as 0.5% ^{13}C RNA (half of background ^{13}C) and 0.1% ^{15}N RNA (one-third of background ^{15}N), enrichment levels that are substantially lower than those in previously reported SIP experiments (Uhlík *et al.*, 2009). In Chip-SIP tests with unenriched bacterial RNA, HCE values (minus s.e.) were $\leqslant 0$ (Figure 2c). A final experiment, combining two bacterial strains from the same phylum, showed Chip-SIP's ability to discriminate 5% from 20% ^{13}C enrichment (Figure 2d).

Method validation experiments

In a second set of experiments, we tested the viability of Chip-SIP for a diverse natural community, using a sample from the SF Bay, a eutrophic estuary. The bay water was incubated in the dark with excess (high micromolar) concentrations of ^{15}N ammonium and ^{13}C glucose for 24 h, a timescale long enough to ensure detectable isotopic labeling of the dominant active community. We expected the most active taxa to incorporate these substrates, as they are of low molecular weight, do not require extracellular breakdown before uptake and directly feed into central carbon and nitrogen metabolic pathways. The Chip-SIP array consisted of 2500

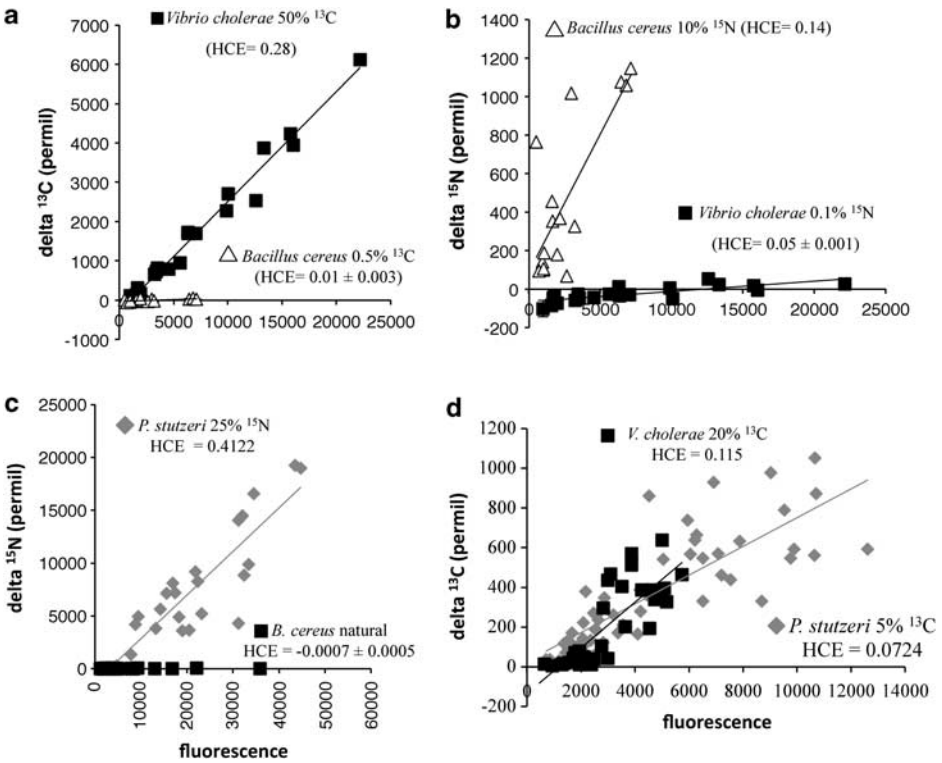

Figure 2 Chip-SIP analysis of mixtures of two bacterial strains labeled with differential isotopic enrichment; HCE, hybridization-corrected enrichment (slope of per mil enrichment and fluorescence). Each point represents an individual probe spot's fluorescence intensity value plotted against its isotopic enrichment. Comparisons included dual labeled *V. cholera* and *B. cereus* (**a**, **b**), ^{15}N labeled *P. stutzeri* and unlabeled *B. cereus* (**c**), and differentially ^{13}C labeled *V. cholerae* and *P. stutzeri* (**d**).

probes targeting 100 microbial taxa selected from a PhyloChip analysis of the same sample (Brodie *et al.*, 2006) (Supplementary Table 1). On the basis of RNA fluorescence, we positively detected 73 taxa. Although the substrate concentrations used may have favored copiotrophs, microbes adapted to high substrate concentrations (Suttle *et al.*, 1990), we also detected isotopic incorporation by the model oligotroph *Pelagibacter* (Rappe *et al.*, 2002) (Supplementary Figure 2). This result demonstrates that oligotrophic organisms retained a presence and detectable biogeochemical activity in this eutrophic environment, even with added substrate concentrations two orders of magnitude greater than what is typically measured in this environment.

As in the experiments with laboratory cultures, the relationship between fluorescence and isotopic incorporation for each taxon was positive and linear for both ^{15}N and ^{13}C (Figures 3a and b) validating the HCE approach for complex natural communities. We found that within the same bacterial family, different taxa often incorporated different levels of ^{15}N from ammonium or ^{13}C from glucose. This shows that different microbial populations, as identified by 16S rRNA gene sequence, represented physiologically distinct components of the community. When comparing taxa from different families, substrate incorporation patterns were not broadly phylogenetically conserved: members of the same bacterial families were scattered throughout the HCE distribution (Supplementary Figure 2). For example, at the family level, substrate incorporation by the Flavobacteriaceae was not distinguishable from the Rhodobacteraceae, due to the high degree of within-family variance (Figure 3c).

A useful feature of Chip-SIP is the detection of ^{13}C and ^{15}N in the same assay to uncover physiological diversity, based on the relative incorporation of two substrates incubated simultaneously. Our ability to measure taxon-specific substrate incorporation allowed us to reveal that the relationship between ammonium and glucose incorporation was positive and linear: organisms with high ammonium incorporation (high ^{15}N HCEs) also exhibited high glucose

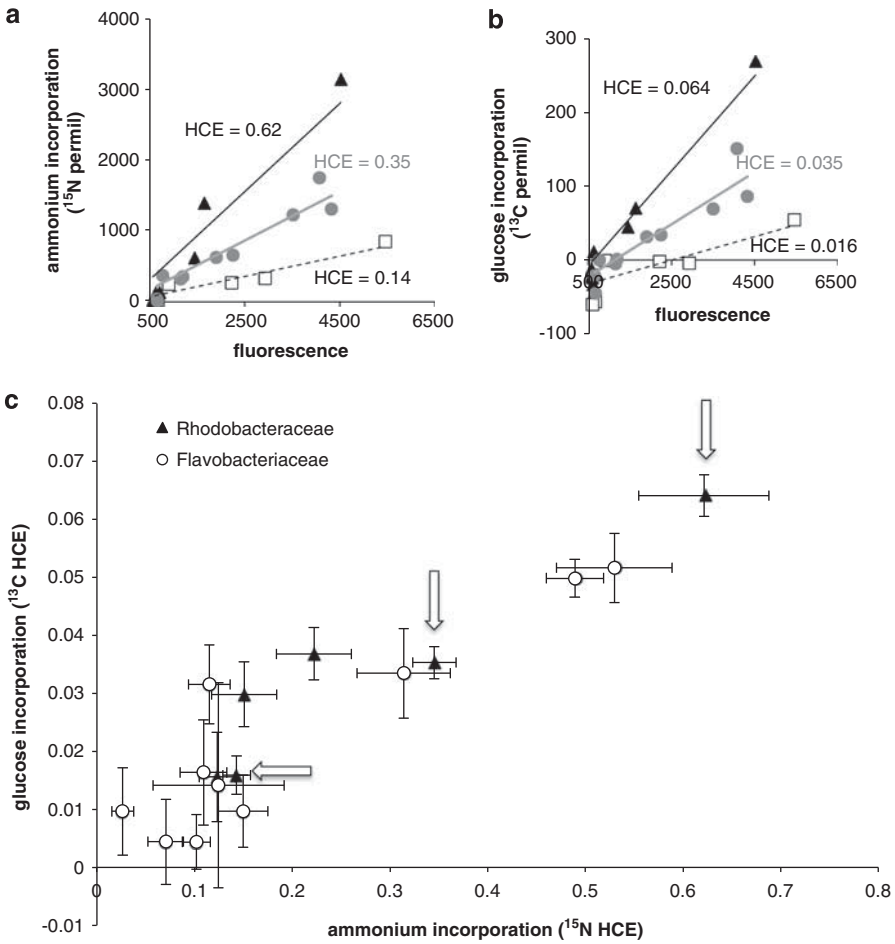

Figure 3 Substrate incorporation detected by Chip-SIP for a SF Bay estuarine microbial community incubated with ^{15}N ammonium and ^{13}C glucose for 24 h; (**a**) ammonium and (**c**) glucose incorporation for 3 taxa within the family Rhodobacteriaceae; each point is derived from a single probe spot's isotopic enrichment value plotted against fluorescence; (panel **c**) relationship between ammonium and glucose incorporation for 16 taxa from 2 bacterial families; HCE, hybridization-corrected enrichment; arrows indicate taxa plotted in panels **a** and **b**.

incorporation (high ^{13}C HCEs) and *vice versa* (Figure 3c, Supplementary Figure 2). This pattern indicates a physiological constraint for marine bacterial C and N resource use; in both our analyses of individual taxa and previous low-resolution bulk measurements (Goldman and Dennett, 1991), community-wide C/N assimilation was constant.

Field experiments with multiple substrates

In an example of the type of experiment that could eventually lead to more realistic models of marine food web structure (Pomeroy *et al.*, 2007), we examined incorporation of three organic substrates by a diverse marine community. In this case, we applied Chip-SIP to a second set of SF Bay samples, collected on a different date and incubated separately with isotopically labeled AAs, NAs and FAs. These substrates make up a significant proportion of the photoautotrophic biomass (Fernandez-Reiriz *et al.*, 1989) that provides fixed carbon for the marine microbial food web. Unlike the previous method validation experiments described above, added substrate concentrations were similar to what has been measured in this environment. In contrast, most traditional SIP experiments use substrate concentrations much higher than ambient, often in the millimolar range (Dumont and Murrell, 2005).

To compare genome-predicted biogeochemical activity with our measured substrate incorporation data, we examined the presence of genes involved in extracellular degradation or transport of AAs, NAs and FAs in the sequenced genomes of marine bacterial isolates (Supplementary Table 2). A comparison of 17 genomes with closely related Chip-SIP targeted taxa shows mixed results (Table 1). Genomic potential and detected activity agreed for 5 taxa, whereas for the 12 others there was less concordance between genome-predicted function and measured metabolic activity. This is perhaps not surprising, as the genome sequences available likely do not reflect the organisms being analyzed from SF Bay.

We detected isotopic enrichment from at least 1 of the 3 added substrates in 52 out of the 81 taxa with positive RNA hybridization (Supplementary Figures 3–5). A network diagram, based on the measured HCE values, illustrates the relative uptake of substrates by microbial taxa, and clearly indicates generalists that incorporated all three substrates versus specialist consumers of only one substrate (Figure 4). Generalists and specialists were not necessarily distinguishable based on 16S phylogeny, even within individual bacterial families. Such an analysis, which includes quantitative information visualized by the thickness of the lines connecting substrates to taxa, is potentially a substantial step forward in our understanding of organic matter flow in the microbial loop.

Incorporation of leucine is routinely used as a proxy for bacterial production in aquatic systems (Kirchman *et al.*, 1985), and meta-transcriptomic evidence suggests that most marine bacterial taxa incorporate AAs (Poretsky *et al.*, 2010). As nearly all sequenced genomes of marine bacteria (111 out of 115 genomes available on Integrated Microbial Genomes) possess annotated AA transporters, we expected active microbes in SF Bay would incorporate them. Although AAs were the most commonly incorporated out of the 3 substrates tested with Chip-SIP (46 of 52 taxa), >10% of active taxa (6 of 52) did not incorporate ^{15}N-labeled AAs, suggesting

Table 1 Comparison of Chip-SIP substrate incorporation profiles for target taxa with closely related genomes available ($N=17$)

Genome name	*Percentage similarity*	*Taxon*	*Amino acids*		*Fatty acids*		*Nucleic acids*		*Matches*
			P	*D*	*P*	*D*	*P*	*D*	
Marine gamma-proteobacterium HTCC2080	99.9	Gammaproteobacteria; OM60	+	+	+	+	+	+	3
Alteromonas macleodii ATCC 27126	99.5	Gammaproteobacteria; Alteromonadales	+	+	+	−	+	−	1
Vibrio alginolyticus 12G01	99.3	Gammaproteobacteria; Vibrionaceae	+	+	+	−	+	−	1
Thalassospira sp. TrichSKD10	98.9	Alphaproteobacteria; Verorhodospirilla	+	+	−	−	−	−	3
A. macleodii Deep ecotype, DSM 17117	98.2	Gammaproteobacteria; Alteromonadales	+	+	+	−	+	+	2
Psychromonas sp. CNPT3	98.1	Gammaproteobacteria; Alteromonadales	+	+	−	−	+	+	3
Synechococcus sp. WH 5701	97.4	Cyanobacteria; Prochlorales	+	+	−	−	−	−	3
Pedobacter sp. BAL39	97.3	Bacteroidetes; Pedobacter	+	+	−	−	+	−	2
Vibrio parahaemolyticus 16	97.3	Gammaproteobacteria; Vibrionaceae	+	+	−	−	+	−	2
Polaribacter sp. MED152	97	Bacteroidetes; Flavobacteriales	+	+	+	+	−	+	2
Marinobacter algicola DG893	96.7	Gammaproteobacteria; Marinobacter	+	+	+	−	+	−	1
Oceanibulbus indolifex HEL-45	96	Alphaproteobacteria; Rhodobacterales	+	+	+	+	+	+	3
Loktanella vestfoldensis SKA53	95.9	Alphaproteobacteria; Rhodobacterales	+	+	−	+	−	+	1
Rhodobacterales bacterium HTCC2150	95.3	Alphaproteobacteria; Rhodobacterales	+	+	+	−	−	+	1
Roseobacter sp. AzwK-3b	95.3	Alphaproteobacteria; Rhodobacterales	+	+	−	+	−	+	1
Pseudoalteromonas sp. TW-7	95	Gammaproteobacteria; Alteromonadales	+	+	+	−	+	−	1
Bacillus sp. NRRL B-14911	94.5	Firmicutes; Bacillus	+	+	+	−	−	+	1

Abbreviations: Chip-SIP, Chip-stable isotope probing; D, incorporation detected by Chip-SIP; P, incorporation predicted by genome annotation. 'Matches' indicate the number of predictions detected (maximum of 3).

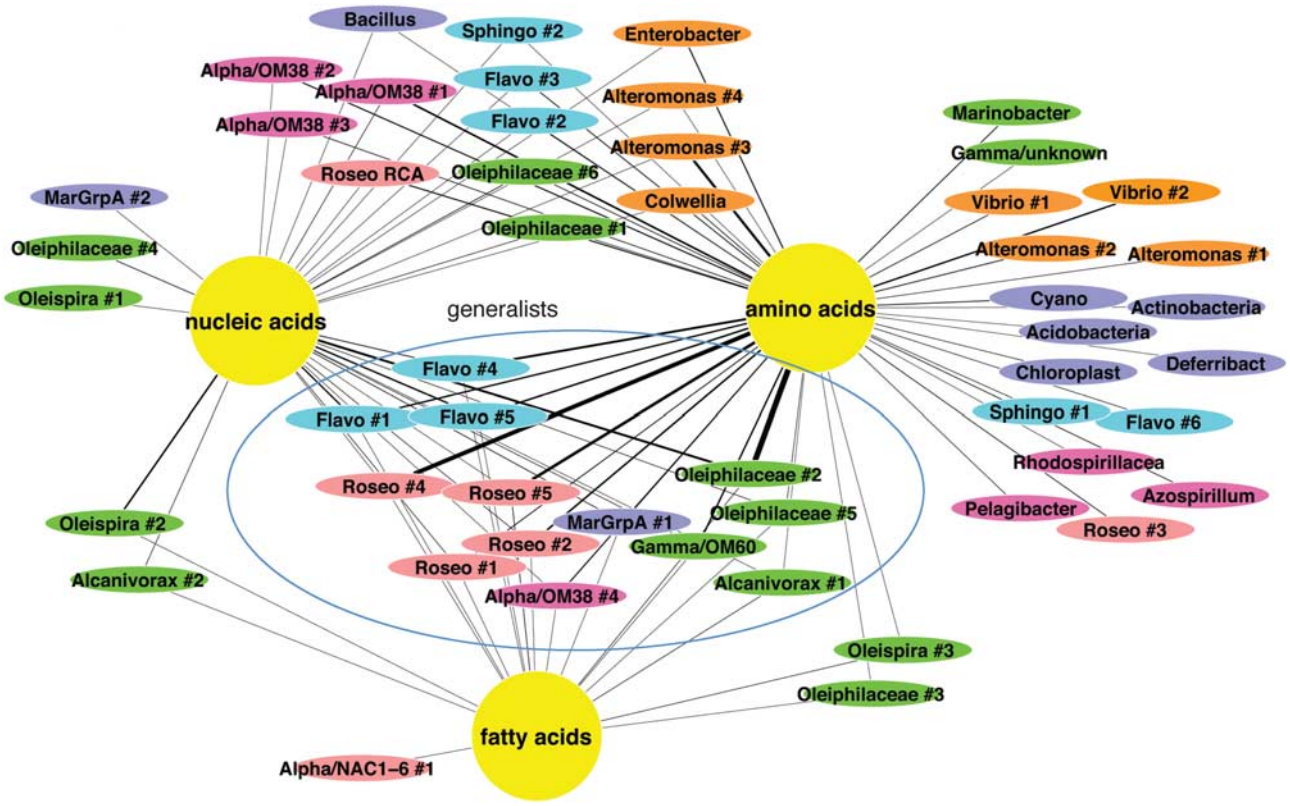

Figure 4 Chip-SIP analysis of the uptake patterns of three organic substrates by different bacterial taxa in SF Bay, identifying substrate specialists and generalists; the thicknesses of the lines are proportional to the substrate incorporation rates based on HCE calculations (Flavo, Flavobacteriaceae; Roseo, Roseobacter; MarGrpA, Marine Group A; Sphingo, Sphingobacteria; Cyano, Cyanobacteria).

that AAs may not be ubiquitous substrates for bacteria. Indeed, if rates of marine bacterial carbon production based on leucine incorporation are underestimates, this could have significant implications for global carbon modeling efforts.

Bacterial uptake of single nucleosides (for example, thymidine) is ubiquitous and used to measure rates of growth (Fuhrman and Azam, 1980), but few studies have examined longer NA molecules as a carbon or nitrogen source (Lennon, 2007). As half (56 of 115) of the sequenced marine bacterial genomes contain at least 1 nucleoside transporter or extracellular nuclease, we expected NA incorporation would be a common phenomenon in the environment. We found that 32 of 52 taxa incorporated ^{13}C-labeled NAs into their RNA, revealing that bacteria commonly incorporate carbon (and presumably nitrogen) from external NA sources. This complements previous work that identified NAs as a source of phosphorus for marine bacteria (Ammerman and Azam, 1985). NAs have C to N ratios lower than phytoplankton-derived particulate organic matter and AAs (C/N of particulate organic matter = 6.6, AA = 3.6, NA = 2.5). This makes NAs an ideal resource for bacteria that have relatively high nitrogen requirements.

Bacterial isolates often reveal high lipase activity (Martinez *et al.*, 1996), although only 41 of 115 sequenced bacterial genomes contained identified lipid transporters. In addition, comparative genomics has shown that oligotrophic marine bacterial genomes contain significantly more lipid metabolism and FA degradation genes than do copiotrophic genomes (Lauro *et al.*, 2009). If oligotrophs favor FA incorporation, we hypothesized that it would be less common than AA incorporation in our samples as a eutrophic estuary should favor copiotrophs. Our data concur with this idea: we found only 18 of 52 taxa incorporated ^{13}C-labeled FAs. FAs were less commonly incorporated than either AAs or NAs, although we did identify one taxon (uncultivated Alphaproteobacterial clade NAC1–6) that incorporated this substrate but not the others. Such measurements of taxon-specific substrate incorporation within complex communities, along with data gleaned from genomic sequencing, could clearly be useful for attempts to isolate previously uncultured microbial taxa.

Phylogenetic distribution of biogeochemical activity

During the last decade, there has been mounting evidence to suggest 16S rRNA-derived phylogeny is not a consistent proxy for functional capabilities (Doolittle and Zhaxybayeva, 2009). Chip-SIP allowed us to test this hypothesis by matching functional resource use to 16S rRNA phylogenetic relationships. We found that although AA

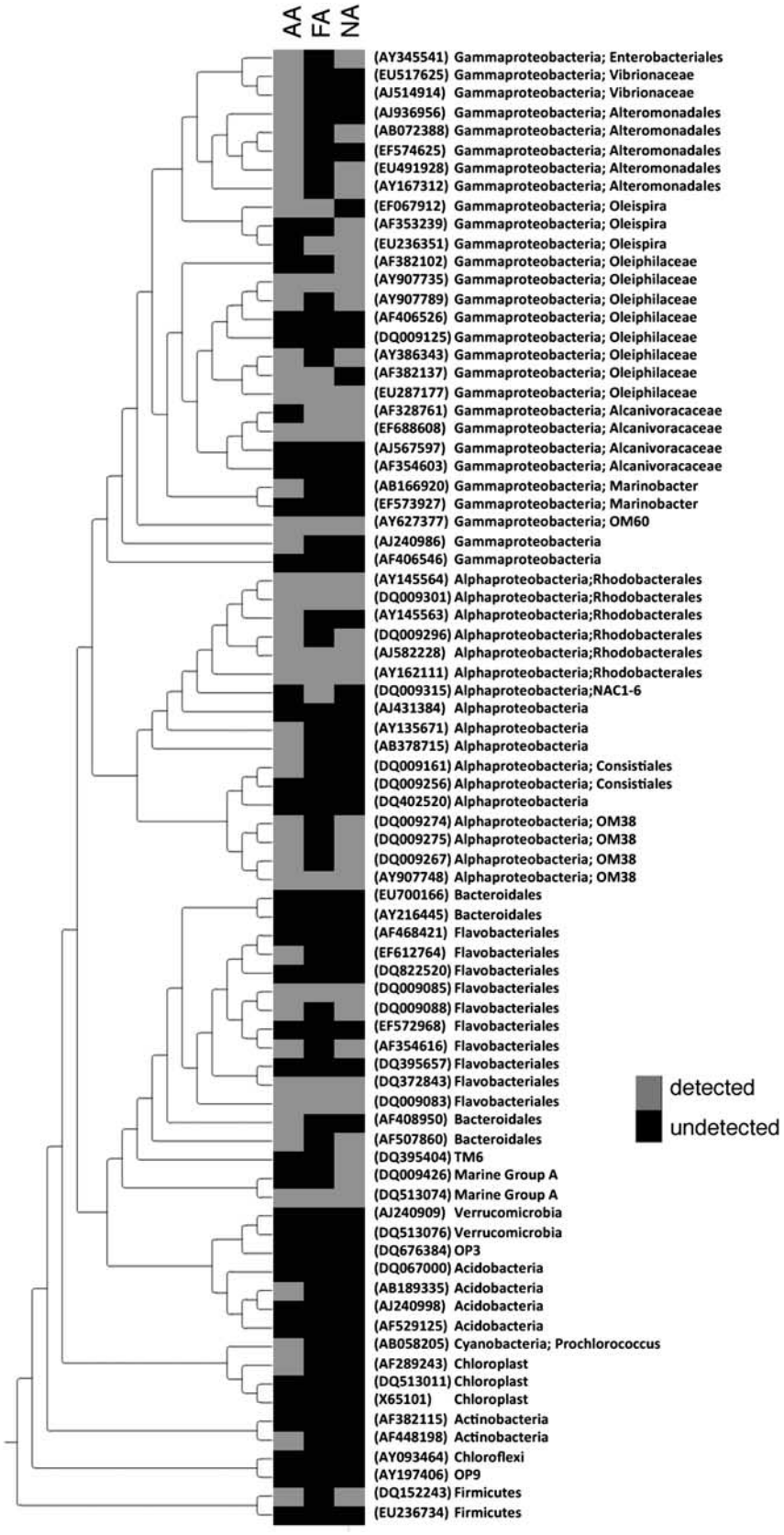

Figure 5 Relationship between substrate incorporation and 16S rRNA phylogeny for SF Bay taxa (AA, amino acids; FA, fatty acids; NA, nucleic acids); GenBank Accession numbers for targeted taxa are included in parentheses.

utilization was randomly distributed on the 16S phylogeny, FA and NA utilization were clustered (Supplementary
Figure 6), implying that these two traits have a phylogenetic signal and some evolutionary process is limiting their selection.

We also examined whether related taxa (for example, members of a bacterial family) exhibited the same substrate utilization pattern. Our data suggest that decoupling between phylogenetic and functional prediction can occur even within very closely related taxonomic groups. For example, three taxa in the Oleispira group (Gammaproteobacteria) all exhibited different substrate incorporation patterns (Figure 5): one incorporated AAs and FAs, the second incorporated only NAs, whereas the third incorporated both FAs and NAs. On the basis of these data, it would be impossible to predict the resource use of any additional *Oleispira* taxa. At the same time, for the well-studied *Vibrio* and *Alteromonas* genera, patterns of resource use did match 16S phylogeny: these taxa incorporated AAs, and several *Alteromonas* taxa incorporated NAs, whereas no taxon incorporated FAs (Figure 5). In this case, 16S-based phylogeny did correlate with resource use. Depending on the taxa examined, our data support both coupling and uncoupling between 16S phylogeny and function. This further highlights the now well-established notion that functional interpretation of 16S data should be made with great caution, as the concept that prokaryotes form ecologically distinct clusters remains controversial (Doolittle and Zhaxybayeva, 2009).

Techniques that directly and simultaneously measure the biogeochemical activities of hundreds of microbial taxa in their native environment are critical to further our mechanistic understanding of how microorganisms control biogeochemical cycles, and enable functional trait-based approaches within microbial ecology. Each technique necessarily has both drawbacks and benefits. As described here, Chip-SIP reports taxon-specific relative isotopic incorporation, based on calculating HCE values that measure isotopic enrichment normalized to fluorescence. Although this relative quantification approach is useful, further work will be necessary to allow absolute quantification. Some experimentalists may also be concerned about access; currently, there are a limited number of NanoSIMS instruments worldwide. We consider the NanoSIMS the optimum analytical tool for ITO-microarray analysis, because of its combination of imaging sensitivity, high-resolution, high ion transmission, high mass resolving power and dynamic analysis. In our proof-of-concept experiments, these attributes were critical for locating and distinguishing probe spots, maximizing ion counts, enabling precise isotopic measurements at masses with multiple isobaric interferences and reducing analysis time. Together, these traits enable the highest throughput with the lowest levels of isotopic enrichment. However, with some compromise in performance, other SIMS instruments (such as time-of-flight SIMS, standard magnetic sector SIMS) can almost certainly perform similar analyses, and with some technological development, automation could potentially be achieved with an ITO-microarray 'isotope reader'. Another concern is whether Chip-SIP can identify metabolic activity in unexpected taxa. In our initial experiments, the Chip-SIP array consisted of 2500 probes targeting 100 previously known microbial taxa selected from a PhyloChip analysis of the same sample. Ideally, Chip-SIP could also be used as a discovery tool, if selection of probe targets was achieved through previous analysis of community sequence data (for example, clone library, pyrotag, metagenome, metatranscriptome).

Where Chip-SIP may be an especially attractive option is for experiments in which high sensitivity and high throughput is important. Other benefits include a moderately rapid sample processing time ($\sim$3 days), and simultaneous ^{13}C and ^{15}N labeling. In our demonstrations of Chip-SIP's potential, we revealed patterns of resource utilization in an estuarine community with critical implications for our understanding of carbon cycling in marine environments. These data considerably expand upon previous studies that have identified marine bacterial resource partitioning based on seasonal and small-scale spatial habitat use (Hunt *et al.*, 2008) by adding relative rates of substrate utilization as a critical component of the bacterial niche.

Acknowledgements

This research was funded by the DOE OBER Genomic Sciences research program and the LLNL Laboratory Directed Research and Development (LDRD) program. This study was performed under the auspices of the US Department of Energy at the Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Part of this study was performed by the US Department of Energy at the Lawrence Berkeley National Laboratory under the auspices of the University of California Contract Number DE-AC02-05CH11231. We thank Larry Nittler for software development and Clark Santee for PhyloChip assistance. Additional thanks to M Firestone, J Fuhrman, R Mueller and three anonymous reviewers for providing constructive comments on earlier versions of this manuscript.

References

Adamczyk J, Hesselsoe M, Iversen N, Horn M, Lehner A, Nielsen PH *et al.* (2003). The isotope array, a new tool that employs substrate-mediated labeling of rRNA for determination of microbial community structure and function. *Appl Environ Microbiol* **69**: 6875–6887.

Ammerman JW, Azam F. (1985). Bacterial 5-nucleotidase in aquatic ecosystems: a novel mechanism of phosphorus regeneration. *Science* **227**: 1338–1340.

Behrens S, Losekann T, Pett-Ridge J, Weber PK, Ng W, Stevenson BS *et al.* (2008). Linking microbial

phylogeny to metabolic activity at the single-cell level by using enhanced element labeling-catalyzed reporter deposition fluorescence *in situ* hybridization (EL-FISH) and NanoSIMS. *Appl Environ Microbiol* **74**: 3143.

Beja O, Spudich EN, Spudich JL, Leclerc M, DeLong EF. (2001). Proteorhodopsin phototrophy in the ocean. *Nature* **411**: 786–789.

Boschker HTS, Nold SC, Wellsbury P, Bos D, de Graaf W, Pel R *et al.* (1998). Direct linking of microbial populations to specific biogeochemical processes by ^{13}C-labelling of biomarkers. *Nature* **392**: 801–805.

Brodie EL, DeSantis TZ, Joyner DC, Baek SM, Larsen JT, Andersen GL *et al.* (2006). Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl Environ Microbiol* **72**: 6288–6298.

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C *et al.* (2007). Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* **2**: 2366–2382.

DeFlaun MF, Paul JH, Jeffrey WH. (1987). Distribution and molecular weight of dissolved DNA in subtropical estuarine and oceanic environments. *Marine Ecol Prog Ser* **38**: 65–73.

DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard N-U *et al.* (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496–503.

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K *et al.* (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069–5072.

Doolittle WF, Zhaxybayeva O. (2009). On the origin of prokaryotic species. *Genome Res* **19**: 744–756.

Dugdale RC, Wilkerson FP, Hogue VE, Marchi A. (2007). The role of ammonium and nitrate in spring bloom development in San Francisco Bay. *Estuarine Coastal Shelf Sci* **73**: 17–29.

Dumont MG, Murrell JC. (2005). Stable isotope probing: linking microbial identity to function. *Nat Rev Microbiol* **3**: 499–504.

Evens R, Braven J. (1988). A seasonal comparison of the dissolved free amino acid levels in estuarine and English Channel waters. *Sci Total Environ* **76**: 69–78.

Fernandez-Reiriz MJ, Perez-Camacho A, Ferreiro MJ, Blanco J, Planas M, Campos MJ *et al.* (1989). Biomass production and variation in the biochemical profile (total protein, carbohydrates, RNA, lipids and fatty acids) of seven species of marine microalgae. *Aquaculture* **83**: 17–37.

Fuhrman JA, Azam F. (1980). Bacterioplankton secondary production estimates for coastal waters of British Columbia, Canada, Antarctica, and California, USA. *Appl Environ Microbiol* **39**: 1085–1095.

Goldman JC, Dennett MR. (1991). Ammonium regeneration and carbon utilization by marine bacteria grown on mixed substrates. *Marine Biol* **109**: 369–378.

Gutierrez-Zamora M-L, Manefield M. (2010). An appraisal of methods for linking environmental processes to specific microbial taxa. *Rev Environ Sci Biotechnol* **9**: 153–185.

Hanson RB, Snyder J. (1980). Glucose exchanges in a salt marsh estuary: biological activity and chemical measurements. *Limnol Oceanogr* **25**: 633–642.

Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ, Polz MF. (2008). Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**: 1081–1085.

Kirchman D, K'nees E, Hodson R. (1985). Leucine incorporation and its potential as a measure of protein synthesis by bacteria in natural aquatic systems. *Appl Environ Microbiol* **49**: 599–607.

Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S *et al.* (2009). The genomic basis of trophic strategy in marine bacteria. *Proc Natl Acad Sci USA* **106**: 15527–15533.

Lennon JT. (2007). Diversity and metabolism of marine bacteria cultivated on dissolved DNA. *Appl Environ Microbiol* **73**: 2799–2805.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar *et al.* (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

Lueders T, Manefield M, Friedrich MW. (2004). Enhanced sensitivity of DNA- and rRNA-based stable isotope probing by fractionation and quantitative analysis of isopycnic centrifugation gradients. *Environ Microbiol* **6**: 73–78.

Maddison WP, Maddison DR. (2010). Mesquite: a modular system for evolutionary analysis. 2.73, http://mesquiteproject.org.

Manefield M, Whiteley AS, Griffiths RI, Bailey MJ. (2002). RNA stable isotope probing, a novel means of linking microbial community function to phylogeny. *Appl Environ Microbiol* **68**: 5367–5373.

Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Grechkin Y *et al.* (2010). The integrated microbial genomes system: an expanding comparative analysis resource. *Nucleic Acids Res* **38**: D382–D390.

Martinez J, Smith DC, Steward GF, Azam F. (1996). Variability in ectohydrolytic enzyme activities of pelagic marine bacteria and its significance for substrate processing in the sea. *Aquatic Microbial Ecol* **10**: 223–230.

Murrell JC, Whiteley AS. (2010). *Stable Isotope Probing and Related Technologies*. ASM Press: Washington, DC.

Ouverney CC, Fuhrman JA. (1999). Combined microautoradiography-16S rRNA probe technique for determination of radioisotope uptake by specific microbial cell types *in situ*. *Appl Environ Microbiol* **65**: 1746–1752.

Pett-Ridge J, Hoeprich P, Weber P, Brodie E. (2011). *Devices, Methods and Systems for Target Detection*. US Patent and Trademark Office. Lawrence Livermore National Security: USA.

Pomeroy LR, Williams PJl, Azam F, Hobbie JE. (2007). The microbial loop. *Oceanography* **20**: 28–33.

Poretsky RS, Sun S, Mou X, Moran MA. (2010). Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environ Microbiol* **12**: 616–627.

Radajewski S, Ineson P, Parekh NR, Murrell JC. (2000). Stable-isotope probing as a tool in microbial ecology. *Nature* **403**: 646–649.

Rappe MS, Connon SA, Vergin KL, Giovannoni SJ. (2002). Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* **418**: 630–633.

Rappé MS, Kemp PF, Giovannoni SJ. (1997). Phylogenetic diversity of marine coastal picoplankton 16S rRNA genes cloned from the continental shalf off Cape Hatteras, North Carolina. *Limnol Oceanogr* **42**: 811–826.

Singh-Gasson S, Green RD, Yue Y, Nelson C, Blattner F, Sussman MR et al. (1999). Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. Nat Biotechnol 17: 974–978.

Stauffer TB, Macintyre WG. (1970). Dissolved fatty acids in the James River estuary, Virginia, and adjacent ocean waters. Chesapeake Sci 11: 216–220.

Suttle C, Fuhrman JA, Capone DG. (1990). Rapid ammonium cycling and concentration-dependent partitioning of ammonium and phosphate: implications for carbon transfer in planktonic communities. Limnol Oceanogr 35: 424–433.

Treusch AH, Leininger S, Kletzin A, Schuster SC, Klenk H-P, Schleper C. (2005). Novel genes for nitrite reductase and Amo-related proteins indicate a role of uncultivated mesophilic crenarchaeota in nitrogen cycling. Environ Microbiol 7: 1985–1995.

Uhlík O, Jecná K, Leigh MB, Macková M, Macek T. (2009). DNA-based stable isotope probing: a link between community structure and function. Sci Total Environ 407: 3611–3619.

Whitman WB, Coleman DC, Wiebe WJ. (1998). Prokaryotes: the unseen majority. Proc Natl Acad Sci USA 95: 6578–6583.

DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor The Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or The Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or The Regents of the University of California.

Ernest Orlando Lawrence Berkeley National Laboratory is an equal opportunity employer.