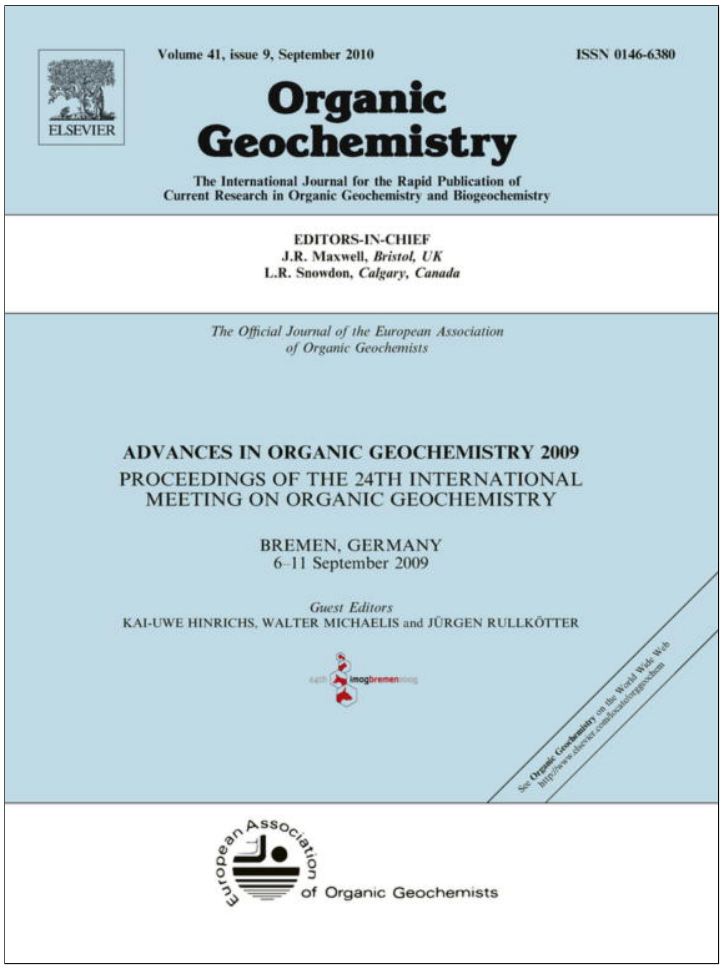

Organic Geochemistry 41 (2010) 1019–1021

Contents lists available at ScienceDirect

Organic Geochemistry

journal homepage: www.elsevier.com/locate/orggeochem

Comparison of multivariate microbial datasets with the Shannon index: An example using enzyme activity from diverse marine environments

A.D. Steen [a,b,*], K. Ziervogel [a], C. Arnosti [a]

[a] Department of Marine Sciences, University of North Carolina – Chapel Hill, NC, USA
[b] Center for Geomicrobiology, Aarhus University, Ny Munkegade 114-116, Bygning 1540, Aarhus, Denmark

ARTICLE INFO

Article history:
Received 18 September 2009
Received in revised form 9 May 2010
Accepted 14 May 2010
Available online 17 June 2010

ABSTRACT

Heterotrophic microbial communities contain substantial functional diversity, so studies of community function often generate multivariate data sets. Techniques for data reduction and analysis can help elucidate qualitative differences among sites from multivariate data sets that may be difficult to grasp intuitively from raw data. The Shannon index is one such technique, used commonly in ecological studies to quantify species evenness. Here, the Shannon index is used to compare quantitatively the extent to which complex microbial communities vary in their capability to access polysaccharides. It is maximized when hydrolysis rates for all polysaccharides are equal and minimized when the range among individual hydrolysis rates at a given site is large. Application of the technique to depth profiles of polysaccharide hydrolysis rates from four distinct pelagic marine environments indicates that, in three of four cases, surface water communities accessed substrates at a more even rate than in deeper waters. The technique could usefully be applied to other types of data obtained in studies of microbial activity and the geochemical effects.

1. Introduction

Microbial communities cycle ca. 50% of marine primary productivity (Azam, 1998), remineralizing organic carbon to CO_2 and transforming substrates to microbial biomass. Factors controlling the activity of these complex natural communities are not well understood, since only a small fraction of bacteria have been isolated in pure culture (DeLong and Pace, 2001) and many major groups lack cultured representatives (Hugenholtz et al., 1998). Genomic information provides insight into community diversity and metabolic potential (DeLong et al., 2006), but in most cases direct information about the extent to which potential is expressed is lacking. In this context, measurement of specific steps in carbon cycling pathways can provide insight into the metabolic capabilities of complex communities without the need for identification and/or isolation of specific organisms.

Measurement of rates of multiple processes in environmental samples lead to multivariate data sets. Data reduction techniques can be used to quantitatively compare these data sets and to search for patterns on larger spatial and/or temporal scales. Measurements of the activity of specific microbial extracellular enzymes are one example of the type of data. For over a decade, we have used a suite of fluorescently labeled polysaccharides to measure substrate specificity and activity of microbial extracellular enzymes in a range of pelagic environments, including temperate and Arctic sediments, riverine–seawater transects and the pelagic ocean (e.g. Arnosti, 1995, 1998, 2008; Arnosti et al., 2005; Steen et al., 2008; Ziervogel and Arnosti, 2009).

Here, we demonstrate the use of a simplified form of the Shannon index, widely used among ecologists and mathematicians, as a statistical means for comparing the evenness of enzymatic capability in complex microbial communities among different locations and/or depth in the water column. We also present an algorithm for estimating the influence of analytical error on the calculated rate and suggest that the approach can also be used to analyze rates of other microbial processes. Analysis of hydrolysis rates in depth profiles from four pelagic marine sites demonstrates the utility of the approach.

The Shannon index was developed in the context of computer science to measure the information content of data sets and is now frequently used by ecologists, among others, to determine the evenness of species distributions at a given location (Legendre and Legendre, 1998). In our approach, the index is maximized when all rates at a given site are equal (maximum evenness) and minimized when only one rate is detected (maximum unevenness). It thus provides a straightforward, direct way of comparing the extent to which microbial communities are generalists, performing a range of comparable reactions at similar rates, or are specialists, performing some reactions much faster than others.

* Corresponding author. Tel.: +45 8942 3241; fax: +45 8942 2722.
E-mail address: andrew.steen@biology.au.dk (A.D. Steen).

0146-6380/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.orggeochem.2010.05.012

Table 1
Hydrolysis rate for six polysaccharides in pelagic waters and summed rates, nM h^{-1} (see text for location and depth).[a]

	Ara	Chn	Fuc	Lam	Pul	Xyl	Sum
J surf	0	5.00 ± 0.17	2.10 ± 0.28	0.93 ± 0.06	0	0.90 ± 0.04	8.93 ± 0.20
J deep	0	7.36 ± 0.07	0.9 ± 0.72	0.39 ± 0.10	0	2.97 ± 0.86	11.7 ± 0.65
T15 surf	2.46 ± 1.42	0	0	14.2 ± 0.75	2.10 ± 0.94	6.5 ± 3.14	25.3 ± 2.11
T15 deep	1.34 ± 0.57	0	0	12.1 ± 0.21	0	0	13.4 ± 0.35
DBM surf	9.60 ± 11.3	22.4 ± 8.69	7.05 ± 1.92	20.5 ± 0.31	1.26 ± 0.10	16.5 ± 2.71	77.4 ± 8.45
DBM deep	12.77 ± 2.91	24.2 ± 1.12	22.6 ± 1.38	23.0 ± 0.88	6.17 ± 0.06	24.6 ± 1.42	113.4 ± 2.19
GS surf	7.98 ± 3.57	1.89±.08	7.24 ± 0.81	7.67 ± 0.41	0.12 ± 0.02	13.3 ± 2.36	38.2 ± 2.53
GS mid	0.20 ± 0.12	0.07 ± 0.06	0	0.53 ± 0.08	0	0	0.79 ± 0.09
GS deep	0	0.10 ± 0.06	0	1.67 ± 0.31	0	0	1.77 ± 0.18

[a] Ara = arabinogalactan; Chn = chondroitin sulfate; Fuc = fucoidan; Lam = laminarin; Pul = pullulan; Xyl = xylan; error in the summed rates is propagated.

2. Methods

2.1. Study sites and seawater collection

Surface and subsurface seawater was collected via Niskin or Go-Flo bottle at four sites: Station T15 (eastern Pacific: 23.09°S, 79.17°W, 3 and 100 m depth; described by Arnosti et al. (2005)), Station DBM (Delaware Bay mouth, 38.52°N, 75.05°W, 1 m and 14 m depth described in Ziervogel and Arnosti (2009)), Station GS (Gulf Stream, 36.43°N, 74.75°W, 1.5 m, 90 m and 146 m depth described by Steen et al. (2008)) and Station J (Svalbard; 79°43′N, 11°05′E, 2 m and 210 m depth.)

2.2. Measurement of extracellular enzymatic hydrolysis rates

Extracellular enzymatic hydrolysis rates were measured using six fluorescently labeled polysaccharides (Arnosti, 1995): pullulan, laminarin, xylan, fucoidan, arabinogalactan, and chondroitin sulfate. These polysaccharides are marine-derived, and/or enzymes hydrolyzing them have been identified in marine bacteria or in the genomes of fully-sequenced marine bacteria (Arnosti and Repeta, 1994; Glockner et al., 2003; Bauer et al., 2006; Weiner et al., 2008). The polysaccharides (from Sigma or Fluka) were fluorescently labeled with fluoresceinamine as described by Arnosti (2003) and 3.5 μmol monosaccharide-equivalent ml^{-1} of each polysaccharide was added to triplicate samples that were incubated at in situ temperature. Subsamples were withdrawn periodically, filtered through 0.2 μm pore filters and stored frozen until analysis. Polysaccharide hydrolysis was measured using a gel permeation chromatography/HPLC system with fluorescence detection, and hydrolysis rates were calculated as described by Arnosti (1995, 2003). Hydrolysis rates for T15, GS and DBM have been reported (Arnosti et al., 2005; Steen et al., 2008; Ziervogel and Arnosti, 2009); the application of the Shannon index to these data is new.

2.3. Statistical analysis of data

The modified form of the Shannon index, H, used here is given by $H = -\sum_{i=1}^{n} p_i \ln p_i$, where n is the total number of substrates and p_i the hydrolysis rate of the ith polysaccharide normalized to the sum of all absolute polysaccharide hydrolysis rates r_i, such that $p_i = \frac{r_i}{\sum_{i=1}^{n} r_i}$.

When there was no detectable hydrolysis of a specific polysaccharide, $p_i \ln p_i$ was set to zero. H is minimized at zero when only one polysaccharide is detectably hydrolyzed and maximized when all polysaccharides are hydrolyzed at the same rate, with a value given by $H_{max} = -\ln(1/n)$. Here, for the six polysaccharides, $H_{max} = 1.79$. We emphasize that comparisons of H among sites using this equation are only valid when the same number of enzyme activities are measured at each site.

The influence of variability among replicates in the underlying rates (r_i) on the final value of H was assessed using a Monte Carlo technique, implemented in Visual Basic for Applications (Microsoft Corp.; see Supplementary material).

3. Results and discussion

Enzymatic hydrolysis rate for the six polysaccharides varied greatly by substrate, depth and site (Table 1). At stations T15 and GS, the number of substrates hydrolyzed decreased with depth, but not at stations DBM or J. At station DBM, enhanced sediment resuspension may have broadened the spectrum of enzymes active in the water column (Ziervogel and Arnosti, 2009). In general, laminarin was hydrolyzed at every depth and site, but hydrolysis of fucoidan and pullulan, for example, was relatively infrequent. Comparison of hydrolysis rates at different depths and sites can provide insight into the individual enzymatic capabilities of natural microbial communities, but only a qualitative comparison of the multivariate data set in Table 1 is possible without a technique for data reduction.

Comparison of hydrolysis patterns via the Shannon index (Fig. 1) provides further information, demonstrating that subsurface waters generally have a less even distribution of hydrolysis rates than surface waters from the same location. In subsurface waters, certain substrates may be hydrolyzed comparatively

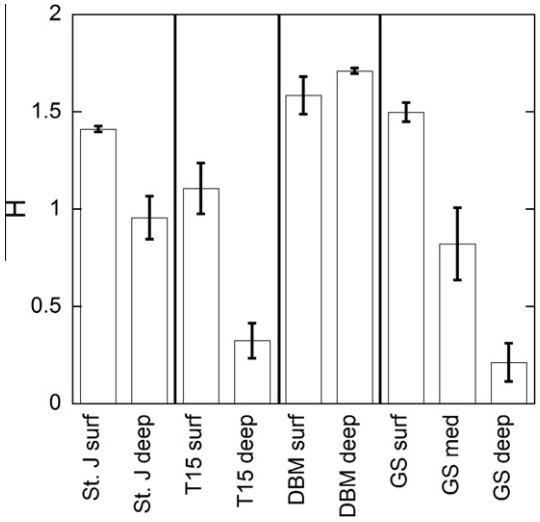

Fig. 1. Shannon index values showing evenness of hydrolysis rate at each station and depth. Stations are identified along the x axis and vertical lines separate data from different stations. Error bars represent standard deviation of values generated in a Monte Carlo simulation (see Section 2 and Supplementary material).

rapidly and others at rates considerably more slowly than in surface waters. This observation suggests that subsurface microbial communities are 'tuned' to specific substrates to a greater extent than their surface counterparts. Such an ecologic strategy might be developed in the context of the fate of sinking organic matter (OM), from which presumably the most labile components are removed rapidly in the surface ocean. As particles and aggregates sink through the water column, a more focused set of enzymes may be more useful to the subsurface pelagic community in remineralizing high molecular weight OM.

The Shannon index as applied here is a simple technique for quantifying the breadth of the ability of microbial communities to hydrolyze a structurally defined set of polysaccharides. The technique is potentially applicable to any analysis of microbial community function that provides data meeting two criteria: (i) community function must be measured along a reasonably large number of independent dimensions and (ii) measurements must be of comparable processes and must be expressed in identical units.

Other common measurements that might be amenable to this technique include uptake rates of different radiolabeled substrates (e.g. a suite of volatile fatty acids), or hydrolysis rates of different small substrate analogs for enzyme activity (e.g. the methylumbelliferyl substrates). Use of the index to analyze such data sets offers a new pathway for probing the capabilities of microbial communities among different environments.

Acknowledgements

WE thank two anonymous reviewers whose careful review prompted us to improve the work. We also thank the captains and crews of the M/S Farm, R/V Cape Henlopen, R/V Hugh R. Sharp and R/V N.B. Palmer. This work was supported by the EPA STAR graduate fellowship (A.D.S.) and NSF-OCE (C.A).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.orggeochem.2010.05.012.

Guest Associate Editor—**A. Pearson**

References

Arnosti, C., 1995. Measurement of depth- and site-related differences in polysaccharide hydrolysis rates in marine sediments. Geochimica et Cosmochimica Acta 59, 4247–4257.

Arnosti, C., 1998. Rapid potential rates of extracellular enzymatic hydrolysis in Arctic sediments. Limnology and Oceanography 43, 315–324.

Arnosti, C., 2003. Fluorescent derivatization of polysaccharides and carbohydrate-containing biopolymers for measurement of enzyme activities in complex media. Journal of Chromatography B 793, 181–191.

Arnosti, C., 2008. Functional differences between Arctic sedimentary and seawater microbial communities: contrasts in microbial hydrolysis of complex substrates. FEMS Microbiology Ecology 66, 343–351.

Arnosti, C., Repeta, D.J., 1994. Extracellular enzyme activity in anaerobic bacterial cultures: evidence of pullulanase activity among mesophilic marine bacteria. Applied and Environmental Microbiology 60, 840–846.

Arnosti, C., Durkin, S., Jeffrey, W.H., 2005. Patterns of extracellular enzyme activities among pelagic marine microbial communities: implications for cycling of dissolved organic carbon. Aquatic Microbial Ecology 38, 135–145.

Azam, F., 1998. Microbial control of oceanic carbon flux: the plot thickens. Science 280, 694–696.

Bauer, M., Kube, M., Telling, H., Richter, M., Lombardot, T., Allers, E., Wurdemann, C.A., Quast, C., Kuhl, H., Knaust, F., Woebken, D., Bischof, K., Mussmann, M., Choudhuri, J.V., Meyer, F., Reinhardt, R., Amann, R.I., Glockner, F.O., 2006. Whole genome analysis of the marine Bacteroidetes '*Gramella forsetii*' reveals adaptations to degradation of polymeric organic matter. Environmental Microbiology 8, 2201–2213.

DeLong, E.F., Pace, N., 2001. Environmental diversity of bacteria and archea. Systematic Biology 50, 470–478.

DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.-U., Martinez, A., Sullivan, M.B., Edwards, R., Brito, B.R., Chisholm, S.W., Karl, D.M., 2006. Community genomics among stratified microbial assemblages in the ocean's interior. Science 311, 496–503.

Glockner, F.O., Kube, M., Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Borzym, K., Heitmann, K., Rabus, R., Schlesner, H., Amann, R., Reinhardt, R., 2003. Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. Proceedings of the National Academy of Sciences USA 100, 8298–8303.

Hugenholtz, P., Goebel, B.M., Pace, N.R., 1998. Impact of culture-independent studies on the emerging phylogenetic view of bacterial diversity. Journal of Bacteriology 180, 4765–4774.

Legendre, P., Legendre, L., 1998. Numerical Ecology. Elsevier, Amsterdam.

Steen, A.D., Hamdan, L., Arnosti, C., 2008. Dynamics of high molecular weight dissolved organic carbon in the Chesapeake Bay: insights from enzyme activities, carbohydrate concentrations, and microbial metabolism. Limnology and Oceanography 53, 936–947.

Weiner, R.M., Taylor II, L.E., Henrissat, B., Hauser, L., Land, M., Coutinho, P.M., Rancurel, C., Saunders, E.H., Longmire, A.G., Zhange, H., Bayer, E.A., Gilbert, H.J., Larimer, F., Zhulin, I.B., Ekborg, N.A., Lamed, R., Richardson, P.M., Borovok, I., Hutcheson, S., 2008. Complete genome sequence of the complex carbohydrate-degrading marine bacterium, Saccharophagus degradans Strain 2-40T. PLoS Genetics 5, e1000087.

Ziervogel, K., Arnosti, C., 2009. Enzyme activities in the Delaware Estuary affected by elevated suspended sediment load. Estuarine, Coastal and Shelf Science 84, 253–258.