ELSEVIER

FEMS Microbiology Ecology 32 (2000) 97-102

FEMS MICROBIOLOGY Ecology

www.fems-microbiology.org

Determining replication for discrimination among microbial communities in environmental samples using community-level physiological profiles

M.B. Lowit, L.K. Blum, A.L. Mills *

Laboratory of Microbial Ecology, Department of Environmental Sciences, University of Virginia, Clark Hall, Charlottesville, VA 22903, USA

Received 17 February 2000; accepted 1 March 2000

Abstract

A statistical approach was employed to assess microbial community variability within a single 4-l water sample and among 10 independent 4-l water samples from a single location using community-level physiological profiles. Power calculations demonstrated that duplicate analyses could distinguish between two different locations at two times during the year. Variability associated with replicates from a single container or from different containers was nearly the same for the two sites examined. Duplicate assays (one from each of two independent samples) were sufficient to resolve the between-site differences at $a = 0.05$ with power exceeding 0.95 at both times of the year. Similar methods are recommended to determine the appropriate number of replicates for environments of interest. © 2000 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.

Keywords: Community structure; Physiological profile; Environmental variability; Statistical power

I. Introduction

Two types of variability are inherent in any environmental measurement. Test reproducibility is the variation associated with the analytical tool being used, i.e. the variance associated with repeated measurements on the same sample or on a series of identical samples. Environmental variation is related to the inhomogeneous distribution of the element of interest being measured. Any measure of community structure, including measures of similarity between two community samples, can vary either because of variability of the measurement tool, or because of non-uniform distribution of the community members within or among samples that are collected. If test variability approaches the level of environmental variability, the test will be insensitive to the environmental variability, and will therefore be useless for discriminating among communities.

A number of procedures for comparing components of microbial community structure have been proffered recently, including a number of molecular methods along with several physiological approaches. All such methods have the potential to contain both of the aforementioned types of variability. Here, we examine one of those methods, the community-level physiological profile (CLPP) approach [1], to determine both test reproducibility and the concomitant sensitivity to environmental variation. This is an approach that has been used for community comparisons in a variety of environments, and while a few reports dismiss the approach because it is culturally based (e.g. [2]), many comparisons have shown that it compares favorably with genetic and biochemical approaches in its ability to discriminate among communities (e.g. [3-8]).

CLPP measures of community similarity are based on response patterns obtained when environmental samples are inoculated into each of the 96 wells of a BIOLOG® plate. Each well contains a single organic compound to serve as a sole carbon, energy, and electron source for any organisms that can utilize the compound under the incubation conditions imposed. The response is measured as the reduction of colorless tetrazolium violet to the highly colored formazan. Thus, the extent of carbon utilization in each well is reflected in the intensity of the color development. Upon incubation, the pattern produced by color development in each of the wells forms a fingerprint for that community. The fingerprints represent the phys-

* Corresponding author. Tel.: +1 (804) 924-0564; Fax: +1 (804) 982-2137.

0168-6496/00/$20.00 © 2000 Federation of European Microbiological Societies. Published by Elsevier Science B.V. All rights reserved.
PII: S0168-6496(00)00017-9

iological potential of the community with respect to the substrates employed, and communities from similar or disparate environments are differentiated by some form of pattern analysis, often, principal components analysis (PCA) [1,9-15]. In the plates most commonly used, each substrate occurs only once in each plate, so that replication must be done by inoculating replicate plates.

Glimm et al. [15] expressed concerns about replicate variability and the ability of CLPP to detect true differences among communities. These authors worried that separate samples drawn from the same community could produce different response patterns on different plates, necessitating multiple replicates of each community to ensure that difference among samples is real. Although Hitzel et al. [14] recommended that all substrates used in this approach are run in at least triplicate for each sample collected to reduce or at least characterize the analytical variation, Mills and Bouma [13] suggested that analytical variation is small. In a series of tests involving five replicate plates per sample, overall analytical variability was below 5%; that is, fewer than 5% of the total tests were not in complete agreement across all replicates. Those experiments were conducted in a series of carefully controlled bioreactors operated as continuously stirred tank reactors with little fluctuation in operating conditions over relatively long periods. Furthermore, the communities in the reactors comprised only four or five different isolates growing together as gnotobiotic assemblages. It is not clear if the results of that study can be extrapolated to making determinations of community similarity in natural habitats containing diverse communities existing in a heterogeneous environment.

The present study was conducted to determine the reproducibility of the CLPP assay in two adjacent tidal creeks to compare analytical reproducibility with environmental variability. Environmental variability consisted of differences due to both spatial separation of the two creeks and temporal changes that occurred in the two creeks between summer and fall. Additionally, we sought to examine the sensitivity of the CLPP procedure as applied to these creeks to help define a strategy for determining adequate sampling allocation within each creek.

2. Materials and methods

The CLPP approach using BIOLOG® technology was used to assess microbial community similarity in two well-mixed tidal marsh creeks on the Eastern Shore of Virginia. Phillips Creek, a tidal creek of the coastal lagoon complex on the seaside of the Delmarva Peninsula, and Hungars Creek, on the Chesapeake Bay side, are separated by only 7 km of land, yet the bacterial biomass, abundance and productivity are different [16], and the communities differ genetically [17,18]. Samples were collected from each creek on two different dates (October 7, 1996 and May 22, 1997) to examine differences in the CLPP. At each site and time, closely spaced, multiple independent samples were collected and analyzed to determine both test reproducibility and spatial variability. Ten 4-l samples of surface water were collected in clean containers from the middle of each creek during a mid-day low tide. One BIOLOG® GN plate was inoculated from each of nine of the containers and 10 plates were inoculated from the 10th container. This process produced a measure of within-site (i.e. between-container) and within-container variability. The latter represents a combination of analytical variability (test reproducibility) and environmental variability at the scale of the sampling container due to the heterogeneous distribution of suspended bacteria.

In subsequent sampling (23 June, 31 July, 28 August and 1 October 1997), two containers were collected at the previously sampled sites within each creek. One BIOLOG® plate was inoculated from each container to compare differences in community similarity through time.

Inoculation of BIOLOG® plates was delayed no more than 2 h after sampling the first site. Rapid inoculation is essential so that community profiles are developed based on the physiological potentials of the in situ community, not a temporally realted product of that community. Although confinement may not cause drastic changes in species composition [19], significant changes in the abundance [20] and metabolic activity of bacterioplankton assemblages can occur [21]. Samples were inoculated on land to avoid any possible cross-well contamination resulting from boat movement [1]. All BIOLOG® plates were incubated at room temperature (ca. 25°C). Color formation in the 96 wells of each plate was read at 590 nm using a Labsystems Multiskan RC plate reader. Individual absorbance values of the 95 wells containing carbon sources were corrected by subtracting the absorbance value of the control well. Plates were normalized by calculating average well color development (AWCD) as described by Garland and Mills [1]. Although this approach is occasionally questioned (e.g. [22]), the alternative is to count the organisms and dilute the communities so that each is inoculated with the same number of organisms. This approach is flawed because of the time delays as indicated above, and because the dilution of unevenly distributed communities will result in the loss of some populations. The functional potential of individual populations cannot be summed to yield a pattern representative of the intact community, and dilution may result in the loss of activities produced by interpopulation synergism [13]. Plates were read periodically (every 6-10 h) until each plate reached a (blank-corrected) AWCD of approximately 0.5 absorbance units [23]. During the warmer months, this level was achieved in about 48 h, and in the colder months after 96 h. The data were transformed by division of each blank-corrected well by the plate AWCD to further reduce the variation among plates [1,23]. PCA was performed on the AWCD-transformed data.

The means of the PC scores for each sample site and time combination (7 October 1996 and 22 May 1997) were used to determine the power of the test given the measured variability. Power is the probability of rejecting the null hypothesis when it is, in fact, false. Once determined, power can be used to allocate sampling effort required to detect a desired difference in sample means at a desired a-level [24]. The power calculations were used to determine the number of CLPP replicates needed to discriminate (i) between the two creeks and (ii) between dates within each creek.

3. Results and discussion

CLPP successfully discriminated among the microbial communities present in Phillips and Hungars Creeks in October 1996 and May 1997 when 10 replicates at each site were used (Fig. 1), a result which agreed with those previously observed in bacterial dynamics [16] and genetic similarity [17,18] in these two creeks. In general, the variability associated with replicates from a single container or from different containers was nearly the same for each site, although in Phillips Creek, the variance associated with the single-container samples was slightly larger than that from the several containers. This implies that the patch size of the microbial community is either much larger than the area sampled by the 10 grab samples or is much smaller than that sampled by a single 4-l grab. The most plausible explanation is that the large-scale variation is a result of the actual communities being quite small, but combined together in a mosaic that cannot be resolved at the sampling scale employed.

Although few investigators report using extensive replication, Bossio and Scow [25] used multiple replication at two levels. They had four replicate treatments with three subsamples inoculated from each treatment replicate. The subsample plates were averaged and this 'average plate value' was used to represent each of the replicate treatments in analysis. This is a strong statistical approach, but it requires many replicate plates that may not add to the ability to discriminate among the communities. Use of appropriate power calculations can indicate at what levels and to what extent replication is necessary to distinguish among communities in similar or dissimilar samples and to allocate sampling effort most efficiently to optimize the analytical sensitivity [20]. In most cases, environmental variability is greater than analytical variability such that the best overall reduction of variance will be achieved by independent replication at a higher level, i.e. more samples rather than more plates.

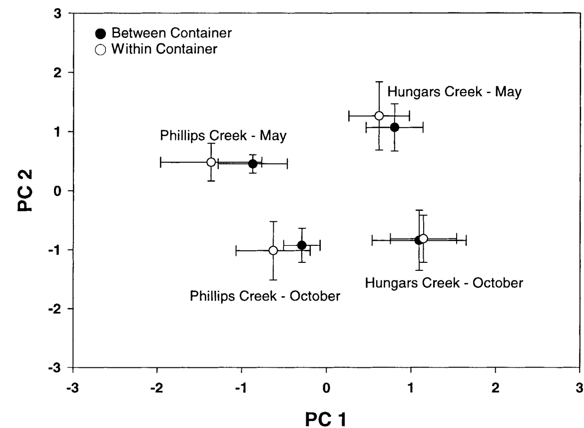

Fig. 1. Results of PCA of 'between-container' (within-site) and 'within-container' replicates in Phillips and Hungars Creeks at two sampling times. For each analysis at each site and each time, $n = 10$. Error bars represent 99% confidence intervals.

The data generated by PCA from the BIOLOG® plates (Fig. 1) were normally distributed for each principal component (data not shown), therefore, the power of CLPP to discriminate among communities could be calculated. The results of the power calculation (Table 1) indicated that for this system, use of two plates per site (one from each of two independently collected samples) was adequate to resolve the between-site differences at $a = 0.05$ with high power. Power values above 0.8 are considered adequate, greater than 0.9 good, and greater than 0.95 excellent [26]. For all of the comparisons, two plates from each site yielded a power greater than or equal to 0.95, a highly

Table 1
Number of replicate BIOLOG® plates needed to demonstrate significant differences among microbial communities at a given a-level

Site comparison	Replicates per site	Power at $a = 0.05$	Power at $a = 0.01$
Phillips Creek, 10/1996 vs. Hungars Creek, 10/1996	4	0.9999+	0.9999+
	3	0.9999+	0.994
	2	0.95	0.48
Phillips Creek, 5/1997 vs. Hungars Creek, 5/1997	4	0.9999+	0.9999+
	3	0.9999+	0.996
	2	0.97	0.52
Phillips Creek, 10/1996 vs. Phillips Creek, 5/1997	4	0.9999+	0.9999+
	3	0.9999+	0.997
	2	0.98	0.56
Hungars Creek, 10/1996 vs. Hungars Creek, 5/1997	4	0.9999+	0.9999+
	3	0.9999+	0.998
	2	0.99	0.66

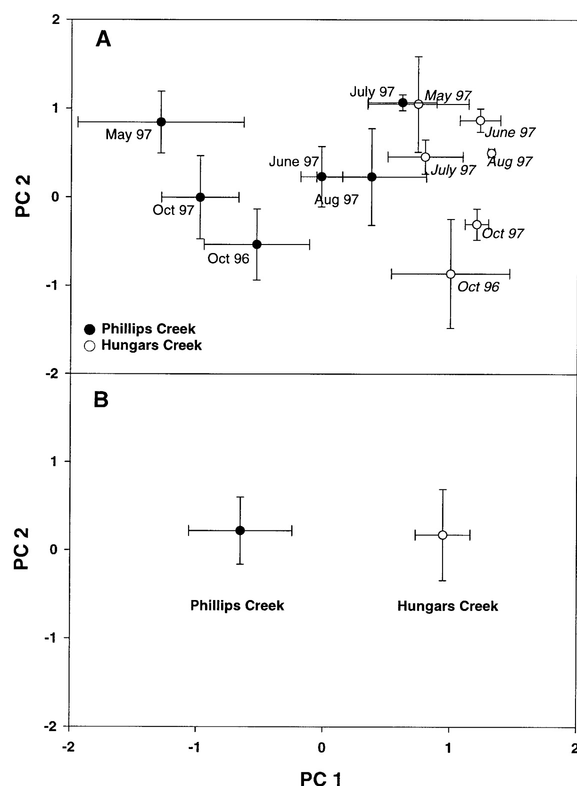

Fig. 2. PCA of temporal changes in CLPP responses in Phillips Creek (e) and Hungars Creek (0). A shows all the data separated by site and sampling times. Error bars represent 1 S.D. (n = 10 for October 1996 and May 1997 ; n = 2 for all other times). B includes all the data separated only by site; error bars represent the 99% confidence interval (n = 28 for each site).

desirable situation. Three plates would yield a power above 0.99 ; four plates would not significantly improve that value. If an a-level of 0.01 is necessary, three plates per site would be adequate to give power in excess of 0.99. Clearly, the test using duplicate plates should not fail to discriminate true differences in the communities (that can be evaluated with CLPP) due to inadequate within-site replication.

Although duplicate plates (one from each of two independent samples) provided high power for testing differences in this system, additional replicates might be necessary in other environments. If investigators question the sensitivity of their analysis, a power determination such as conducted here with an intensive reconnaissance sampling to ascertain the magnitude of the variance, followed by determination of the power at the desired level of a will provide the appropriate sample size necessary.

When examined over time using two independent replicates as suggested by the power calculations, the creeks were separated in every month sampled. The communities at the two sites were more similar in the warmer summer months than during the cooler months of sampling (Fig. 2A). Differences in community similarity are not a direct result of differences in abundance since there is a greater difference in abundance during the warmer months [16,27] when the communities were most similar. When the raw data used to generate Fig. 2A were collapsed over time, the overall site differences between the creeks were clear (Fig. 2B). Temporal changes in the communities of both creeks occurred on the second principal component with site differences separating on the first principal component (Fig. 2A).

Using CLPP procedures, Garland [10] found consistent differences among rhizosphere communities from four plant types maintained under controlled conditions. Here, we demonstrate the capability to consistently differentiate communities in uncontrolled conditions in agreement with previously published activity [16], abundance [16] and genetic [17,18] measures, which supports the robustness of the CLPP approach. Garland [10] also saw temporal shifts in soybean rhizosphere communities in response to plant developmental state. Again, the temporal shifts reported here were obtained within an uncontrolled environment. Perhaps even more interesting is that in both creeks, a seasonal cycle was suggested by the similarity in the October 1996 and October 1997 samples. A definitive conclusion about seasonality can only be reached by additional replication over several years.

Some investigators (e.g. [12,28]) have suggested that properties of in situ microbial community function might be inferred from the specific set of substrates having high loading scores on PC1 and PC2, while others [14] have proposed using a reduced set of 'ecologically relevant' substrates in CLPP assays. We examined the factor loading scores for all substrates making a significant contribution to PC1 (score > 0.5 in magnitude) using the scores from separate ordination of the within-container (10 replicates drawn from a single container) and within-site (10 replicates drawn from 10 separate containers) for CLPP. The factor loading scores obtained from ordination of these two sets of data were compared to the factor loading scores for the pooled data. Ordination of the within-container and within-site data sets resulted in clustering of the creeks and times on PC1 and PC2 (ordination plots not shown) that were similar to the ordination results for the pooled data (Fig. 1); the creeks always separated on PC1 and time always separated on PC2. However, the substrates with high factor loading scores for PC1 were different for each of the ordination plots (Table 2). For example, seven compounds gave significant factor loading scores when the pooled data were considered. For the within-container data, 15 substrates made a significant contribution to PC1, while 10 substrates made significant contributions to PC1 for the within-site data. Only three substrates (psicose, �-methylglucoside and glucuronamide) were common contributors to PC1. This lack of consistency among the factor loading scores for the same raw data strongly implies that the specific substrates separating

Table 2
Factor loading scores for all substrates contributing significantly to PC1 (score > 0.5 in magnitude) in PCA using all data and subsets of the data grouped by sampling stratum

Substrate	All data	Within-sample	Within-site
Maltose	−0.616	−0.715	−0.300
D-Melibiose	−0.528	−0.407	−0.594
�-Methyl glucoside	−0.552	−0.578	−0.603
Psicose	−0.557	−0.683	−0.499
a-Hydroxy butyric acid	0.505	0.576	0.310
a-Keto valeric acid	0.559	0.684	0.399
Glucuronamide	0.537	0.587	0.573
D,L-a-Glycerol phosphate	0.516	0.445	0.590
D-Arabitol	−0.390	−0.517	−0.324
D-Raffinose	0.368	0.536	0.024
a-Keto butyric acid	0.424	0.607	0.098
Alaninamide	0.489	0.549	0.298
D-Alanine	0.456	0.520	0.313
L-Ornithine	0.484	0.557	0.324
L-Phenylalanine	−0.481	−0.613	−0.219
2-Amino ethanol	0.447	0.564	0.467
2,3-Butanediol	0.410	0.558	0.376
D-Fructose	−0.455	−0.444	−0.551
y-Hydroxy butyric acid	−0.287	−0.134	−0.549
L-Asparagine	0.231	0.010	0.596
L-Aspartic acid	0.438	0.273	0.523
Urocanic acid	−0.370	−0.283	−0.570
Glucose-6-phosphate	0.328	0.162	0.571

Within-sample refers to the samples drawn from a single container, while within-site refers to samples drawn from the 10 separate containers.

the communities are not reflective of the function of the community in situ.

The use of PCA with BIOLOG® technology when there are fewer observations (BIOLOG® plates) than descriptors (95 carbon sources) has been met with criticism and alternative methods for analyzing the data have been suggested [14,22,29]. Hitzl et al. [14] screened for substrates most likely to discriminate among samples and then used a smaller number of substrates in high replication to increase the replication to variable ratio. Insam and Hitzl [29] then proposed a means to determine the appropriate number of replicates to arrive at a suitable number of replicates. Concerns about the ratio of samples to variables voiced by Hitzl et al. [14] and Insam and Hitzl [29] have been discussed extensively in the literature for the past quarter century. A large number of Monte-Carlo simulations of random data sets demonstrated clearly that significant eigenvalues for the factors can be obtained reliably when the percentage of variance explained by the factors exceeds a set value (which for the present study is about 20% for each of the first two factors) [30]. All factors included in the present study exceeded the values presented for their respective sample/variable ratio presented by Overland and Preisendorfer [30]. Legendre and Legendre [31] also point out that even when the sample to variable ratio is low, the first few eigenvectors, those typically used, are little affected so that the problem of low replication compared to number of variables should not lead to incorrect interpretation of ordination in reduced space. Concerns about replicate variability and the ability of CLPP to detect differences among communities are real; however, in this case, use of duplicate plates (one from each of two independent samples) was adequate to allow microbial communities from two well-mixed salt marsh creeks to be distinguished with a high degree of confidence. To achieve a similar level of confidence in other environments using the CLPP approach or molecular measures of microbial community similarity may require a greater number of replicates. Reconnaissance of the environment of interest using the technique of choice and performing power calculations like those used here provide the only means to determine an acceptable number of replicate samples.

Acknowledgements

This work was supported by the Virginia Sea Grant and National Science Foundation through the Virginia Coast Reserve LTER project.

References

[1] Garland, J.L. and Mills, A.L. (1991) Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level-sole-carbon-source utilization. Appl. Environ. Microbiol. 57, 2351-2359.

[2] Konopka, A., Oliver, L. and Turco, R.F. (1998) The use of substrate utilization patterns in environmental and ecological microbiology. Microb. Ecol. 35, 103-115.

[3] Palojarvi, A., Sharma, S., Rangger, A.M. von Lu.. tzow and Insam, H. (1997) Comparison of Biolog and phospholipid fatty acid patterns to detect changes in microbial community. In: Microbial Communities: Functional Versus Structural and Genetic Approaches (Insam, H and Rangger, A., Eds.), pp. 37-48. Springer-Verlag KG Berlin, Berlin.

[4] van Elsas, J.D., Duarte, G.F., Rosado, A.S. and Smalla, K. (1998) Microbiological and molecular biological methods for monitoring microbial inoculants and their effects in the soil environment. J. Microbiol. Methods 32, 133-154.

[5] Schneider, C.A., Mo, K. and Liss, S.N. (1998) Applying phenotypic fingerprinting in the management of wastewater treatment systems. Water Sci. Technol. 37, 461-464.

[6] Di Giovanni, G.D., Watrud, L.S., Seidler, R.J. and Widmer, F. (1999) Comparison of parental and transgenic alfalfa rhizosphere bacterial communities using biolog GN metabolic fingerprinting and enterobacterial repetitive intergenic consensus sequence-PCR (ERIC-PCR). Microb. Ecol. 37, 129-139.

[7] El Fantroussi, S., Verschuere, L., Verstraete, W. and Top, E. (1998) Effect of repeated herbicide applications on soil bacterial communities. Meded. Fac. Landbouwkd. Toegep. Biol. Wet. Univ. Gent 63, 1733-1739.

[8] Di Giovanni, G.D., Watrud, L.S., Seidler, R.J. and Widmer, F. (1999) Fingerprinting of mixed bacterial strains and BIOLOG Gram-negative (GN) substrate communities by enterobacterial repetitive intergenic consensus sequence-PCR (ERIC-PCR). Curr. Microbiol. 38, 217-223.

[9] Garland, J.L. and Mills, A.L. (1994) A community-level physiological

approach for studying microbial communities. In: Beyond the Biomass: Compositional and Functional Analysis of Soil Microbial Communities (Ritz, K., Dighton, J. and Giller, K.E., Eds.), pp. 77-83. Wiley, London.

[10] Garland, J.L. (1996) Patterns of potential C source utilization by rhizosphere communities. Soil Biol. Biochem. 28, 223-230.

[11] Garland, J.L. (1997) Analysis and interpretation of community-level physiological profiles in microbial ecology. FEMS Microbiol. Ecol. 24, 289-300.

[12] Zak, J.C., Willig, M.R., Moorhead, D.L. and Wildman, H.G. (1994) Functional diversity of microbial communities: A quantitative approach. Soil Biol. Biochem. 26, 1101-1108.

[13] Mills, A.L. and Bouma, J.E. (1997) Strain and function stability in gnotobiotic reactors. In: Microbial Communities: Functional Versus Structural and Genetic Approaches (Insam, H. and Rangger, A., Eds.), pp. 184-194. Springer-Verlag KG Berlin, Berlin.

[14] Hitzl, W., Rangger, A., Sharma, S. and Insam, H. (1997) Separation power of the 95 substrates of the BIOLOG system determined in various soils. FEMS Microbiol. Ecol. 22, 167-174.

[15] Glimm, E., Heuer, H., Engelen, B., Smalla, K. and Backhaus, H. (1997) Statistical comparisons of community catabolic profiles. J. Microbiol. Methods 30, 71-80.

[16] MacMillin, K., Blum, L.K. and Mills, A.L. (1992) Comparison of bacterial dynamics in tidal creeks of the lower Delmarva Peninsula, Virginia, USA. Mar. Ecol. Prog. Ser. 86, 111-121.

[17] Lowit, M.B. and Blum, L.K. (1997) Microbial community response to 'transplant' from an autotrophic to a heterotrophic estuary, Presented at the General Meeting Estuarine Research Federation.

[18] Franklin, R.B., Taylor, D.R. and Mills, A.L. (1999) Characterization of microbial communities using randomly amplified polymorphic DNA (RAPD). J. Microbiol. Methods 35, 225-235.

[19] Lee, S. and Fuhrman, J.A. (1991) Species compositional shift of confined bacterioplankton studied at the level of community DNA. Mar. Ecol. Prog. Ser. 79, 195-201.

[20] Kaper, J.B., Mills, A.L. and Colwell, R.R. (1978) Evaluation of the accuracy and precision of enumerating aerobic heterotrophs in water samples by the spread plate method. Appl. Environ. Microbiol. 35, 756-761.

[21] Ferguson, R.L., Buckley, E.N. and Palumbo, A.V. (1984) Response of marine bacterioplankton to differential filtration and confinement. Appl. Environ. Microbiol. 47, 49-55.

[22] Howard, P.J.A. (1997) Analysis of data from BIOLOG plates: comments on the method of Garland and Mills. Soil Biol. Biochem. 29, 1755-1757.

[23] Garland, J.L. (1996) Analytical approaches to the characterization of samples of microbial communities using patterns of potential C source utilization. Soil Biol. Biochem. 28, 213-221.

[24] Sokal, R.R. and Rohlf, F.J. (1981) Biometry. W.H. Freeman, New York.

[25] Bossio, D.A. and Scow, K.M. (1995) Impact of carbon and flooding on the metabolic diversity of microbial communities in soils. Appl. Environ. Microbiol. 61, 4043-4050.

[26] Wall, F.J. (1986) Statistical Data Analysis Handbook. McGraw-Hill Book Co., New York.

[27] Wu, K.W. (1995) Protozoan grazing effects on bacterial dynamics in estuarine tidal marsh creeks. MS Thesis, University of Virginia, Charlottesville, VA.

[28] Staddon, W.J., Duchesne, L.C. and Trevors, J.T. (1997) Microbial diversity amd community structure of postdisturbance forest soils as determined by sole-carbon-source utilization patterns. Microb. Ecol. 34, 125-130.

[29] Insam, H. and Hitzel, W. (1999) Data evaluation of community-level physiological profiles: a reply to the letter of P.J.A. Howard. Soil Biol. Biochem. 31, 1198-1200.

[30] Overland, J.E. and Preisendorfer, R.W. (1982) A significance test for principal components applied to a cyclone climatology. Mon. Weather Rev. 110, 1-4.

[31] Legendre, P. and Legendre, L. (1998) Numerical Ecology. Elsevier Scientific Publishing Co., Amsterdam.