Mapping Changes in Soil Microbial Community Composition Signaling Bioremediation

Article *in* Bioremediation Journal · May 1998

DOI: 10.1080/10889869809351341

7 authors, including:

Jonas S Almeida
Stony Brook University Hospital
237 PUBLICATIONS **9,197** CITATIONS
SEE PROFILE

Kam Tin Leung
Lakehead University Thunder Bay Campus
87 PUBLICATIONS **5,628** CITATIONS
SEE PROFILE

Sarah Macnaughton
Isis Innovation Ltd
135 PUBLICATIONS **6,243** CITATIONS
SEE PROFILE

Cecily A Flemming
Government of Ontario, Canada
64 PUBLICATIONS **2,864** CITATIONS
SEE PROFILE

Mapping Changes in Soil Microbial Community Composition Signaling Bioremediation

Jonas S. Almeida,[1] Kam T. Leung,[2] Sarah J. Macnaughton,[4] Cecily A. Flemming,[2] Michael H. Wimpee,[4] Gregory A. Davis,[4] and David C. White[2,3]

[1]ITQB, Univ. Nova de Lisboa, P.O. Box 127, 2780 Oeiras, Portugal; [2]Center for Environmental Biotechnology, The University of Tennessee, Knoxville, TN 37932-2575; [3]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6036; [4]Microbial Insights, Inc., 2340 Stock Creek Blvd., Rockford, TN 37853-3044

Abstract: Chemical signatures of biological processes reflect their complex interrelationships. The chemical profile is rich in information but poor in content due to the complex processes underlying the chemical composition of natural biological communities. A nonlinear mapping technique, based on artificial neural networks (ANNs), was proposed to highlight information coded in lipid signatures in soil by demonstrating the biological response to hydrocarbon contamination. ANNs do not require mechanistic assumptions, and they can cope with nonlinear associations. Soil sample lipid signatures were mapped using ANNs to recover information on exposure to contamination, to assess the potential for bioremediation as assessed by polymerase chain reaction (PCR)/ deoxyribonucleic acid (DNA) gene probes, and to monitor the effects of selected inocula. A two-coordinate system was built from signature lipid biomarkers containing 64 components from which the values of target parameters (6 components) could be recovered. The map tracks bioremediation, as characterized by the target parameters, and provides information on how parameters interreact during bioremediation. Using 23 soil sample signatures, a map was built from which the 6 target parameters could be recovered with 4.7% average error. Principal component analyses and nonlinear factor analyses by autoassociative ANN were compared to the nonlinear mapping information. Although these methods provided a good description of signature shift, they did not discriminate among all target parameters.

Keywords: hydrocarbon bioremediation, mapping, soil, community composition, lipid signature, neural networks.

Introduction

The signature lipid biomarker (SLB) technique has been used with soil samples to provide a chemical signature for the entire biological community (White et al., 1996). This technique is useful because fewer than 1% of soil bacteria detected by direct count are culturable (Bakken, 1985; Skinner et al., 1952). Although some lipids are uniquely present in specific groups, most are present across a wide range of organisms (Komagata and Suzuki, 1987; Lechevalier and Lechevalier, 1988; White et al., 1996). Likewise, certain lipids have been shown to specifically change in response to physiological stress (Heipieper et al., 1992; Keift et al., 1994; Sikkema et al., 1995; White et al., 1998).

However, for different organisms and culture conditions, most lipids respond differently to the same environmental changes. Consequently, lipid signatures encode for highly complex information on identity and physiological status. Conclusions from such information can be best inferred using methods that account for the nonlinear associations among signature components, such as ANNs (Almeida et al., 1995).

1058-8337/98/$.50

To arrive at a more comprehensive assessment of the potential for in situ bioremediation, SLBs have been combined with PCR/DNA gene probe analysis (White et al., 1997). Three steps are involved: DNA extraction by a direct cell-lysing protocol; amplification by PCR; and hybridization of the amplified DNA with a specific chemiluminescent gene probe (White et al., 1997). By combining the two techniques, information on community structure, viable biomass, physiological status, and the potential for activity of relevant enzymes can be obtained. Such joint analysis provides a powerful tool to quantitatively characterize the effectiveness of bioremediation throughout the entire microbial community without the need to isolate and culture individual microorganisms.

A wealth of complementary information potentially can be recovered from characteristic biochemical signatures. However, analyzing SLB shifts using linear techniques such as principal component analysis (PCA) or clustering on the basis of Euclidean distances will face two serious drawbacks (Morris and Boddy, 1995; Sackin and Jones, 1993). First, nonlinear associations among SLB components are overlooked. Second, many causes of shifts in the SLBs are not parameters of interest, with the consequent overshadowing of relevant variables such as the potential for bioremediation, extent of bioremediation, exposure to contaminants, and persistence of selected/engineered inocula. Developing mapping techniques that will highlight variables of interest is critical for providing more information when using SLBs to monitor in situ bioremediation.

Materials and Methods

Experimental Procedure

Four sets of duplicate soil samples (sandy loam, pH 6.4, obtained from the University of Tennessee, Knoxville Agricultural Experimental Station) were subjected to combinations of two treatments: contamination by a hydrocarbon mixture and inoculation by a defined mixed culture. The hydrocarbon mixture (HC) included (dry wt soil) 100 μg/g naphthalene, 100 μg/g phenanthrene, 1000 μg/g *n*-decane, 1000 μg/g *n*-octane, 50 μg/g *o*-xylene, and 1700 μg/g acetone. The defined bacterial inoculum (Inoc) included *Pseudomonas putida* (ATCC 33015); *Pseudomonas oleovorans* (ATCC 29347); and *Sphingomonas* sp. (ATCC 39723; previously *Flavobacterium* sp.). Duplicate subsamples were sacrificed at 0, 7, and 14 days. Initial samples were taken immediately before incubation but immediately after contamination, when performed.

Sample analyses consisted of determining the SLB profile using modified Bligh and Dyer extraction followed by fractionation and methylation (Guckert et al., 1985; White et al., 1979) and PCR/DNA gene probing for selected hydrocarbon degradation genes *alkB* (alkane hydroxylase), *xylE* (catachol 2,3 dioxygenase), and *pcpC* (tetrachloro-hydroquinone dehalogenase):

- Primer *alkB* (Kok et al., 1989)
  forward 5′-GTTCTGGATTCCGCTCCAGAGTACG, base pair 640-664
  reverse 5′-GCGCTAGTCCGTTCACGATACCCAG, base pair 997-1021
- Primer *xylE* (Nakai et al., 1983)
  forward 5′-TGGCCGCGCGATCTGAAAGGTATGG, base pair 415-439
  reverse 5′-GATATCGAT[A/C]GA[T/G]GTGTCGGTCATG, base pair 699-723
- Primer *pcpC* (Orser et al., 1993)
  forward 5′-AAGATGCCTGAAGTCAGTCTC, base pair 391-411
  reverse 5′-AATGGTCCGATCCGGCGACGG, base pair 1141-1161.

The SLB profile consisted of 64 lipids, listed in Table 1. The experimental data analyzed in this report were presented, in part, by White et al. (1997), where further details on the experimental procedure can be found.

By using DNA probes for *alkB*, *xylE*, and *pcpC*, complementary information on the potential for biodegradation was obtained. The amount of amplification product detected by each probe was assumed to reflect the level of the gene in the original sample, providing a measure of the potential for bioremediation. None of the three degradation genes was detected in the uncontaminated and uninoculated microcosms after incubation. After incubation, *alkB* was detected only in the inoculated and contaminated microcosms; *xylE* was observed in all the contaminated and/or inoculated microcosms; *pcpC* was detected in the microcosms that had been inoculated independently of contamination, by the hydrocarbon mixture. The experimental data leading to these observations were reported by the authors in White et al. (1997). The information above can be used to detect both inoculation and contamination from the gene probe results alone, as summarized in the following logical statements:

$$P(alkB) = (I) \text{ and } (HC) \quad (1)$$

$$P(xylE) = (I) \text{ or } (HC) \quad (2)$$

Table 1. SLB components. Fatty acids are referred to according to the nomenclature described in Ringelberg et al. (1989).

Terminally Branched Saturates	Monoenoics	Branched Monoenoics	Mid-Chain Branched Saturates	Normal Saturates	Eukaryotes
i14:0	15:1w6c	i15:1	br15:0	14:0	18:2
i15:0	16:1w9c	br15:1a	br15:0	15:0	18:2
a15:0	16:1w7c	br15:1b	10me16:0	16:0	18:2w6
i16:0	16:1w7t	i16:1a	11me16:0	17:0	18:3w3
i17:0	16:1w5c	i16:1b	12me16:0	18:0	20:4w6
a17:0/1w8c	17:1w6c	br17:1	13me16:0		20:5w3
i19:0	cy17:0	i17:1w7c	br17:0a		i20:0
	17:1	a17:1	br17:0b		20:2w3
	18:1w9c	br19:1	10me17:0		20:1w11c
	18:1w7c		12me17:0/18:2		20:1w9c
	18:1w7t		10me18:0		20:1w7c
	18:1w5c		12me18:0		20:0
	19:1a				22:0
	19:1b				23:0
	19:1w6c				24:0
	cy19:0				

$$P(pcpC) = (I) \qquad (3)$$

where P = presence above detection limits after incubation, I = inoculation, and HC = contamination. A total of 23 samples (four combinations of two treatments in duplicate, three samples at different incubation times, one individual duplicate was lost) were analyzed for SLBs and were probed for the three target gene sequences. The samples, sorted by treatment and incubation time, were coded S01 to S23, as defined in Table 2.

Data Analyses

SLB components were normalized with respect to total lipids and expressed as mole percent. Gene probe re-

Table 2. Sample codes (S), sorted by inoculation (Inoc), contamination with hydrocarbons (HC), and days of incubation (D).

Sample Code	Microcosm	Sample Code	Microcosm
S01	Inoc HC D0	S12	Inoc No HC D0
S02	Inoc HC D0	S13	Inoc No HC D0
S03	Inoc HC D7	S14	Inoc No HC D7
S04	Inoc HC D7	S15	Inoc No HC D7
S05	Inoc HC D14	S16	Inoc No HC D14
S06	No inoc HC D0	S17	Inoc No HC D14
S07	No inoc HC D0	S18	No inoc No HC D0
S08	No inoc HC D7	S19	No inoc No HC D0
S09	No inoc HC D7	S20	No inoc No HC D7
S10	No inoc HC D14	S21	No inoc No HC D7
S11	No inoc HC D14	S22	No inoc No HC D14
		S23	No inoc No HC D14

sults were expressed as *log(fgDNA/g dry soil)*, corresponding to the equivalent total DNA of the strains from which the oligonucleotide was obtained (refer to Experimental Procedure above). PCA was performed using Statistica 5.1 (Statsoft Inc.). ANN analysis was implemented using BrainCell 2.5 (Promised Land Tech.) and Matlab 5.0 equipped with the neural network toolbox (MathWorks Inc.). The logsig equation was used as a transfer function, and error backpropagation was used to optimize connection weights. Full interconnection between neighbor layers was used.

ANNs are assemblies of parallel processing units that mimic the processing of information by natural nervous systems. Clustering (Gyllenberg and Koski, 1995) and learning from examples with no a priori knowledge of causality (Hinton, 1992) are some of the most important uses of ANNs. Refer to Hagan et al. (1996) for applied design and implementation of ANNs, to Haykin (1994) for the theoretical foundations of ANNs, and to Montague and Morris (1994) for a review of ANN contributions in biotechnology.

Nonlinear Factor Extraction by ANN

The SLB was treated by an autoassociative neural network in which the 64-component profile is associated with itself by feed-forward ANN with three hidden layers, the middle layer consisting of only two nodes. The number of nodes in the neighboring hidden layers is allowed to change to minimize association error. Nonlinear factor loadings are obtained by recording the values of the two nodes in the middle hidden layer (Figure 1). The two hidden nodes correspond to the two coordinates and can be used to recover the original SLB profile within the regression error.

Nonlinear Mapping by Feed-Forward ANN

The feed-forward ANN approach is similar to nonlinear factor extraction except that the output vector is a set of variables of interest characterizing the sample, i.e., incubation time, event of inoculation, event of contamination, and PCR/gene probe results (Figure 2). This topology was implemented to generate two-dimensional (2D) coordinates representing the nonlinear association between the chemical profile and the biologically meaningful parameters. Previously, Noble et al. (1997) had demonstrated that hidden node values could be used to assess biological variability associated with the chosen output parameters.

Results and Discussion

Analysis of the SLB data was aimed at tracking bioremediation status, using the parameters of incubation time, exposure to contamination, persistence of selected inocula, and presence of genes for degrada-

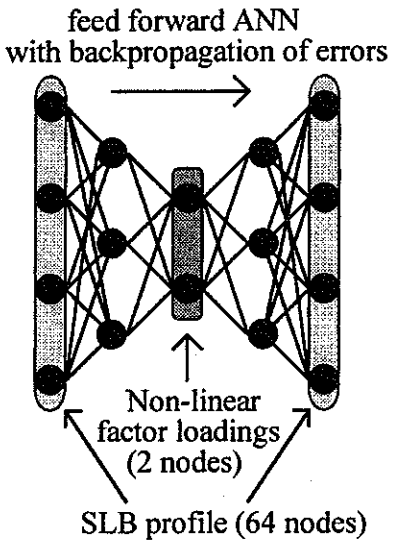

Figure 1. Simplified architecture of autoassociative ANN used to extract two nonlinear factors from a set of SLB profiles. The number of nodes in the second and third hidden layers is allowed to change in order to minimize autoassociation error.

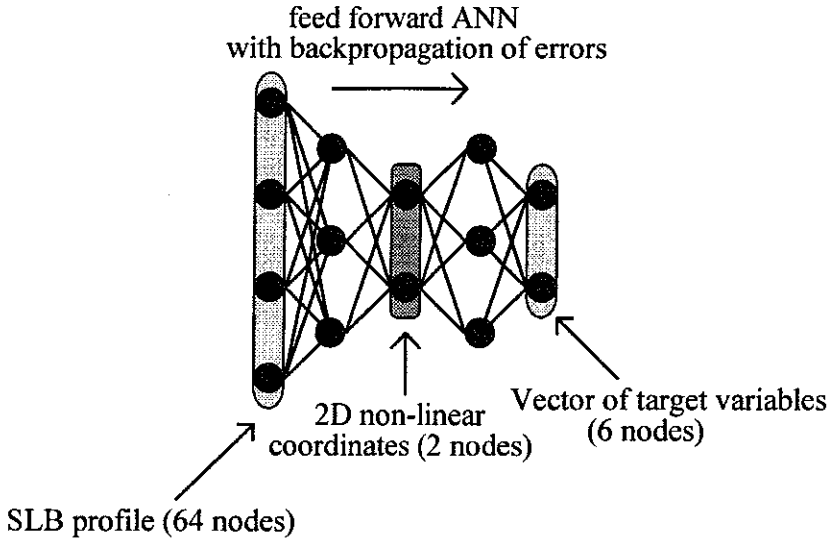

Figure 2. Simplified architecture of feed-forward ANN used to extract coordinates for a 2D nonlinear map of SLB shifts. The number of nodes in the second and third hidden layers is allowed to change to minimize autoassociation error.

tion of reference hydrocarbons. Extraction of the primary components by PCA provided a linear map of SLB shifts (Figure 3). From the factor loadings plotted in Figure 3, it is clear that most of the initial variance is associated with inoculation (compare group I to group II). This was to be expected, as SLB profiles provide a chemical signature for the entire microbial community (White et al., 1996). Conversely, shifts in the signatures of incubated microcosms (days 7 and 14) occur mostly from contamination by hydrocarbons (HC; compare group III to group IV).

As incubation proceeded, contamination overtook inoculation as the main factor determining microbial community composition shifts as assessed by the SLBs.

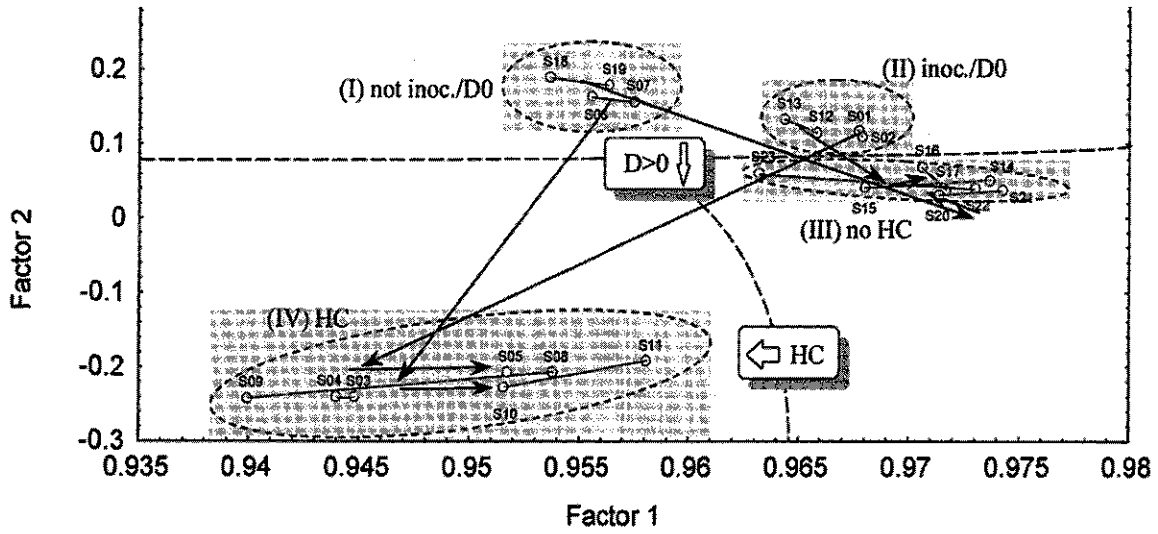

Figure 3. Factor loadings for principal component extraction of SLB. Small circles represent samples (codes defined in Table 2), and thin arrows join time series for days 0 to 7 and 14. Dashed lines cluster samples as follows: group I consists of noninoculated microcosms harvested at day 0 (contaminated or not); group II is as group I but for inoculated microcosms; group III and group IV delimit noncontaminated from contaminated microcosms, harvested at day 7 and day 14, respectively.

The variances represented by the two factors are 92% and 0.5%, respectively, which imply that, by and large, the greatest differences in lipid signatures occurred between groups I and IV compared to groups II and III.

Extracting principal components to map shifts on the SLBs overlooks dependencies among individual SLB components. Signature shifts must be coordinated to account for interdependencies among the individual lipids. However, because PCA considers each signature component independently, PCA will overrepresent effects that trigger changes in multiple signature components. To correct for signature interdependencies, nonlinear component extraction was performed by autoassociative ANN (Figure 4).

The optimized autoassociative ANN was able to recover the original signatures solely from two coordinates (the values for the two nodes of the ANN middle hidden layer, see Materials and Methods) with 4.5% error. The corresponding two-coordinate map is plotted in Figure 4. Extraction of nonlinear factors is not sensitive to redundant SLB shifts (correlated SLB component shifts), because they can be recovered from the same combination of factors. Therefore, the distance between samples is proportional not only to signature shift but particularly to uncorrelated shift, minimizing the bias introduced by redundant responses.

As before (Figure 3), contaminated microcosms sampled after 7 and 14 days of incubation (group IV) form a distinct cluster (Figure 4). Unlike mapping with principal component loadings (Figure 3), contaminated microcosms that were inoculated are now distinguishable from those that were uninoculated (groups IVa and IVb in Figure 4, respectively). Analysis of SLB variability suggests that inoculation has a measurable effect on the community composition of contaminated soils after a 2-week incubation period.

The position of an unknown sample in the PCA or nonlinear factor map is interpreted as suggesting characteristics similar to those of the closest neighbors. This analysis of map position highlights the fact that the goal of mapping signatures is to define position not with respect to signature shift, but rather with respect to shifts in the sample characteristics. Both mapping techniques represent a difference of position as a measure of signature shift. To map shifts in the community composition that are relevant with regard to a set of target characteristics, a new approach was followed. SLB profiles were associated with the set of target parameters by an ANN with three hidden layers, with two nodes on the second hidden layer (the coordinates in a 2D nonlinear map; see Materials and Methods, Figure 2).

Six target parameters were used to find the 2D coordinates: incubation time (0/7/14 days); event of inoculation (0/1); event of contamination (0/1); and the results for the three gene probes (*log(fgDNA/gSoil)* of *alkB*, *xylE*, and *pcpC*). Association was accomplished with a 4.7% error (Figure 5). Unlike previous mapping techniques (Figures 3 and 4), the position in

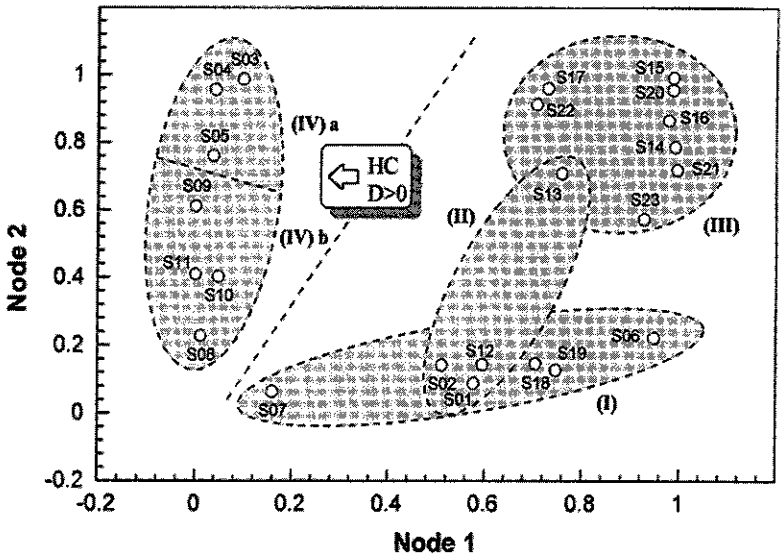

Figure 4. Nonlinear factor extraction by an autoassociative ANN (4.5% error, 11 nodes in the first and third hidden layers). The groups defined by principal component analysis (Figure 3, roman numerals) are delimited by dashed lines.

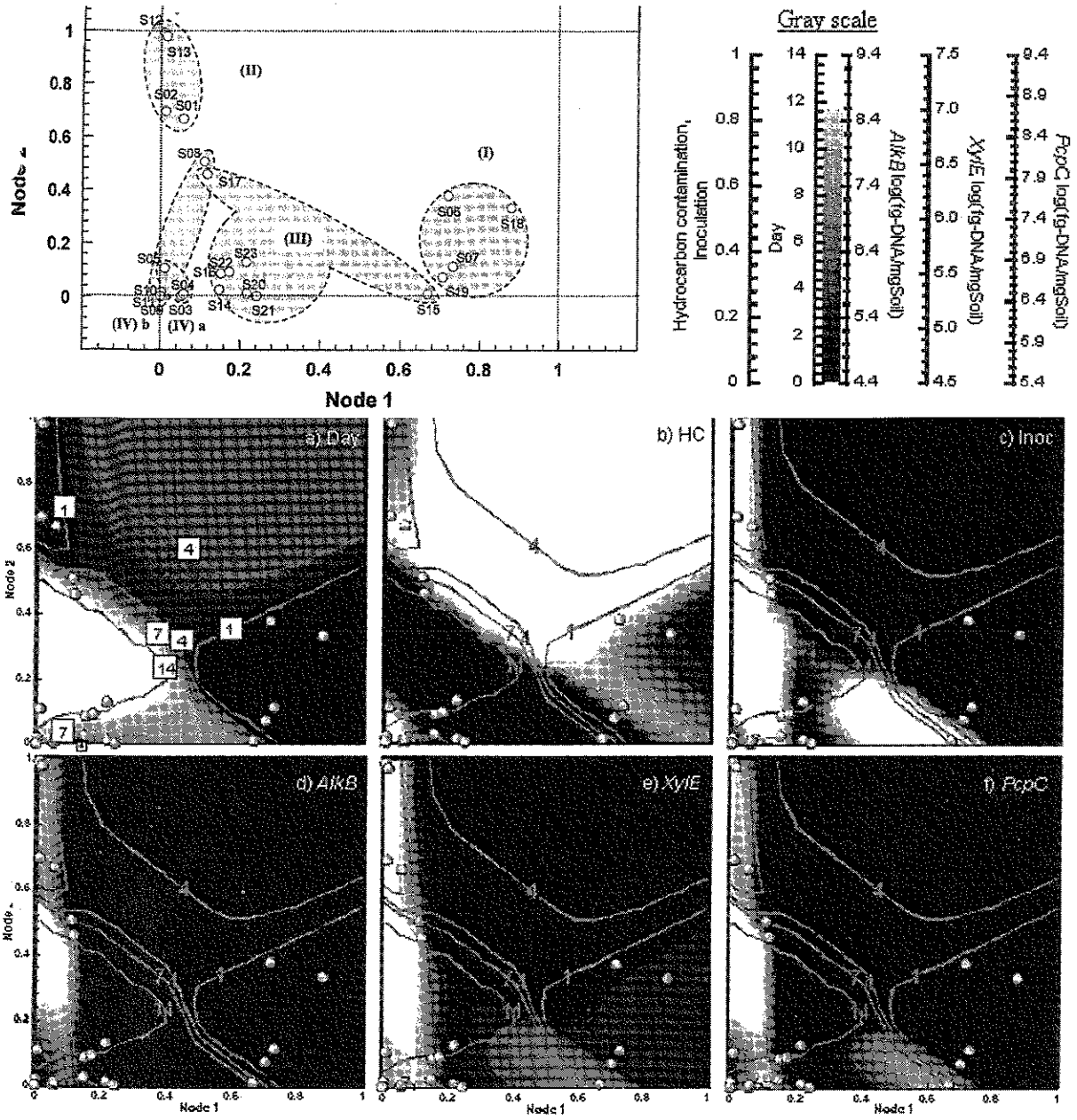

Figure 5. Map of SLB shifts with regard to six parameters. Map with identified samples and membership to groups I through IVa,b (as described in Figure 3) is presented in the upper left corner. The same map is replotted below with a gray-scale background quantifying the association with target parameters (a through f) with a 4.7% average error. The corresponding numerical scale is specified in the upper right corner. Lower boundary values for the gene probe scale correspond to the detection limit. When plotting associated time of incubation (variable day), the coordinates with values 1, 4, 7, and 14 were united as lines (a). These lines were reproduced in subsequent plots (b through f) to highlight incubation time progression from coordinates 1,0 (inoculated, not incubated) and 0,1 (uninoculated, not incubated) toward 0,0 (contaminated, incubated).

the nonlinear map (Figure 5, upper left plot) is associated with a combination of characteristics (Figures 5a-f) rather than with a combination of signature components. A successfully trained ANN will be able to correctly derive map coordinates for new samples from the new SLB profiles.

The position in the map can be read to extract a numeric prediction about the sample's target charac-

teristics. The lighter background signifies a stronger association (gray scale, Figure 5). As an example, S01 and S02 (Table 2) are located in a dark area in Figure 5a, because those signatures come from unincubated microcosms, and in a light area in Figure 5b, correctly signaling for contamination by hydrocarbons. Similarly, inoculation is signaled by a light background for the same samples in Figure 5c. As incubation of inoculated HC-treated microcosms progresses from S01 and S02 to S03 and S04 (day 7) and S05 (day 14), the corresponding points in the map correctly retain the same backgrounds for contamination and inoculation (Figures 5b and 5c). The loss of *alkB* in uncontaminated microcosms, the retention of *xylE* and *pcpC* above detection limits, and the emergence of *xylE* detection in the natural community of contaminated microcosms can be followed in Figures 5d, 5e, and 5f.

The nonlinear map in Figure 5 (upper left plot) describes SLB shifts with incubation from two starting points (I and II) converging toward two endpoints (III and IV). The spatial coordinates are such that individual samples are included in the correct domains for the six target parameters (Figure 5a through 5f). Interdependencies among target parameters can now be identified by superimposition of domains.

The analysis described above is valid for regions with experimental data for the target parameters. Signatures of new samples will be located automatically in the map by using the previously trained ANN. If a new signature is positioned far away from previous data, its interpretation is only tentative. However, if information on the target parameters is available, mapping can be recalculated to incorporate the new information, validating a previously uncharted area. In addition, the opposite analysis can be implemented, because each coordinate in the map is differently sensitive to individual SLB components. This approach can be useful to identify biomarkers for particular combinations of environmental conditions and/or particular biological community compositions.

Feed-forward ANNs were used to infer target parameter values directly from SLBs (results not shown). However, interpretation of results is not as transparent because interdependencies among target parameters are not explicit in the connection weights (Haykin, 1994).

The nonlinear mapping technique is particularly suited to monitor bioremediation in situ. In addition, it can be applied to signatures other than the SLB profile provided that it assesses the global microbial community composition, such as restriction length fragment polymorphisms (RLFPs). The implementation of the technique requires the simultaneous recording of both the signature and the parameters deemed meaningful for the specific type of contamination (i.e., assessing exposure to contaminant, rate of removal, and the potential for bioremediation). After this initial investment, an ANN is set as described above. The ANN is used to predict values for meaningful parameters based solely on sample signatures. Typically, obtaining the signatures is much faster, requires less work, and costs significantly less than assessing bioremediation directly. However, predictions made using the nonlinear map should be validated periodically by direct analysis of bioremediation parameters.

Conclusions

The proximity of signatures as measured by two coordinates does not necessarily imply similarity between signatures. Instead, it reflects the fact that they occur in similar combinations of circumstances. To accommodate all observed combinations of characteristics, a coordinate system is arranged by the ANN to minimize proximity between associated events. As a consequence, proximity in the map reflects information redundancy regarding the target parameters. As more data become available, fortuitous associations will become less frequent, rendering the map more accurate and more amenable to interpretation.

The nonlinear mapping technique was developed to monitor parameters to assess bioremediation status and the potential for bioremediation without requiring direct measurements of the parameters. The method is based on an associative memory concept relying on ANNs, an artificial learning technique. The biochemical lipid signature of biological community composition was used as a key to infer the values of the target parameters. The resulting nonlinear map represents information implicit in the experimental data documenting the association. New signatures can be positioned by using the ANN to calculate their coordinates. Additionally, the nonlinear map offers a description of how target parameters interreact during in situ bioremediation.

Acknowledgments

This work has been supported by Fundação Luso-Americana para o Desenvolvimento (FLAD), by grant INV/2.35 from Junta Nacional de Investigação Científica e Tecnológica (JNICT), and by the U.S. Department of Energy DE-Fco2-96 Er62278, for Assessment Program Leadership in the Natural and Accelerated Bioremediation (NABIR) program, administered by J. Houghton.

References

Almeida, J.S., A. Sonesson, D.B. Ringelberg, and D.C. White. 1995. "Application of artificial neural networks (ANN) to the detection of *Mycobacterium tuberculosis*, its antibiotic resistance and prediction of pathogenicity amongst *Mycobacterium* spp. based on signature lipid biomarkers." *Binary-Computing in Microbiology 7*: 53-59.

Bakken, L.R. 1985. "Separation and purification of bacteria from soil." *Appl. Environ. Microbiol. 49*: 1188-1195.

Guckert, J.B., C.P. Antworth, P.D. Nichols, and D.C. White. 1985. "Phospholipid ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments." *FEMS Microbiol. Ecol. 31*: 147-158.

Gyllenberg, M., and T. Koski. 1995. "A taxonomic associative memory based on neural computation." *Binary 7*: 61-66.

Hagan, M.T., H.B. Demuth, and M. Beale. 1996. *Neural Network Design.* Prindle Weber & Schmidt Pub., Boston, MA.

Haykin, S. 1994. *Neural Networks—A Comprehensive Foundation*, pp. 397-443. Macmillan College Pub., New York, NY.

Heipieper, H.J., R. Diffenbach, and H. Keweloh. 1992. "Conversion of *cis* unsaturated fatty acids to *trans*, a possible mechanism for the protection of phenol-degrading *Pseudomonas putida* P8 from substrate toxicity." *Appl. Environ. Microbiol. 58*: 1847-1852.

Hinton, G.E. 1992. "How neural networks learn from experience." *Sci. Am. 9*: 145-151.

Keift, T.L., D.B. Ringelberg, and D.C. White. 1994. "Changes in ester-linked phospholipid fatty acid profiles of subsurface bacteria during starvation and desiccation in a porous medium." *Appl. Environ. Microbiol, 119*: 303-308.

Kok, M., R. Oldenhuis, M.P.G. van der Linden, P. Raatjes, J. Kigma, P.H. van Lelyveld, and B. Witholt. 1989. "The *Pseudomonas oleovorans* alkane hydroxylase gene." *J. Biol. Chem. 264*:5435-5441.

Komagata, K., and K. Suzuki. 1987. "Lipid and cell-wall analysis in bacterial systematics." In: R.R. Collwell, and R. Grigorova (Eds.), *Methods in Microbiology.* Vol. 19, pp. 161-207. Academic Press Limited, London.

Lechevalier, H., and M.P. Lechevalier. 1988. "Chemotaxonomic use of lipids—An over view." In: C. Ratledge and S.G. Wilkinson (Eds.), *Microbial Lipids,* pp. 869-902. Academic Press, Oxford.

Montague, G., and J.N. Morris. 1994. "Neural network contributions in biotechnology." *Trends in Biotech. 12*: 312-324.

Morris, C. W., and L. Boddy. 1995. "Artificial neural networks in identification and systematics of eukaryotic microorganisms." *Binary 7*: 70-76.

Nakai, C., H. Kagamiyama, and M. Nozaki. 1983. "Complete nucleotide sequence of the metapyrocatechase gene on the TOL plasmid of *Pseudomonas putida* mt-2." *J. Biol. Chem. 258:* 2923-2928.

Noble, P.A, K.D. Bidle, and M. Fletcher. 1997. "Natural microbial community composition compared by a back-propagation neural network and cluster analysis." *Appl. Environ. Microbiol. 63*:1762-1770.

Orser, C.S., J. Dutton, C. Lange, P. Jablonski, L. Xun, and M. Hargis. 1993. "Characterization of a flavobacterium glutathione S-transferase gene involved in reductive dechlorination." *J. Bacteriol. 175*: 2640-2651.

Ringelberg, D.B., J.D. Davis, G.A. Smith, S.M. Pfiffner, P.D. Nichols, J.B. Nickels, J.M. Hensen, J.T. Wilson, M. Yates, D.H. Kampbell, H.W. Reed, T.T. Stocksdale, and D.C. White. 1989. "Validation of signature polar lipid fatty acid biomarkers for alkane-utilizing bacteria in soils and subsurface aquifer materials." *FEMS Microbiol. Ecology 62*: 39-50.

Sackin, M. J., and D. Jones. 1993. "Computer-assisted classifications." In: M. Goodfellow, and A.G. O'Donnell (Eds.), *Handbook of New Bacterial Systematics,* pp. 281-313. Academic Press, London.

Sikkema, J., J.A.M. deBont, and B. Poolman. 1995. "Mechanisms of membrane toxicity of hydrocarbons." *Microbiol. Rev. 59*: 201-222.

Skinner, F.A., P.C.T. Jones, and J.E. Mollison. 1952. "A comparison of a direct and a plate counting technique for the quantitative estimation of soil microorganisms." *J. Gen. Microbiol. 6*: 261-271.

White, D.C., W.M. Davis, J.S. Nickels, J.D. King, and R.J. Bobbie. 1979. "Determination of the sedimentary microbial biomass by extractable lipid phosphate." *Oecologia 40*: 51-62.

White, D.C., C.A. Flemming, K.T. Leung, and S.J. Macnaughton. 1998. "In situ microbial ecology for quantitative assessment, monitoring, and risk assessment of pollution remediation in soils, the subsurface,

the rhizosphere and in biofilms." *J. Microbiol. Methods*, in press.

White, D.C., K. Leung, S.L. Macnaughton, C. Flemming, M. Wimpee, and G. Davis. 1997. "Lipid/DNA biomarker analysis for assessment of in situ bioremediation effectiveness." In: *In Situ and On-Site Bioremediation*, Vol. 5, pp. 319-324. Battelle Press, Columbus, OH.

White, D.C., J.O. Stair, and D.B. Ringelberg. 1996. "Quantitative comparisons of in situ microbial biodiversity by signature biomarker analysis." *J. Indust. Microbiol. 17*: 185-196.