498

Running title: *Lipid analysis of Cr impacted microbial communities*

Section: General Microbial Ecology

Assessment of the impact of Cr^{3+} contamination on the soil microbial community via phospholipid fatty acid profile analysis and application of neural networks.

Jonas S. Almeida[1], Sarah J. Macnaughton[2*], John. R. Stephen[2Ψ], Terry L. Marsh[3], David Long[3], Ying-Don Gan[2], Gary Icopini[3], Craig C. Brandt[4], Jack C. Schryver[4], Susan M. Pfiffner[2] and David C. White[2,4]

[1]ITQB & FCT/Univ. Nova de Lisboa, R Qta Grande 6, 2780 Oeiras, Portugal

[2]Center for Environmental Biotechnology, The University of Tennessee, 10515 Research Drive, Suite 300, Knoxville, TN 37932, USA

[3] ROME Project Laboratory, Center for Microbial Ecology, Department of Microbiology, Michigan State University, East Lansing, MI 48824, USA

[4]Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge , TN 37831, USA

*Corresponding author: Mailing address: National Environmental Technology Centre AEA Technology Environment, E6 Culham Science Centre, Culham, Abingdon Oxfordshire OX14 3ED, UK. Phone: 011- 44- 1235 463974; Fax 011-44-1235 463010.

Ψ Current address: Crop and Weed Science Dept., Horticulture Research International, Wellesbourne, CV35 9EF, UK

Abstract

The accumulation of toxic metals is a major concern at industrial and defense related sites. Soil microbial community structures are known to change in the presence of toxic metals with adaptation to pollutants. Herein we have used phospholipid fatty acid (PLFA) biomarker characterization and viable bacterial counts to determine shifts in the bacterial and microeukaryote biomass/communities in soils taken from a heavily contaminated Superfund site located near Sault St. Marie, Upper Peninsula, MI. Chromium (Cr^{3+}) at this site ranges from background levels (0-50 mg kg^{-1}) to ~200,000 mg kg^{-1}. Linear and non-linear techniques were used to map changes in the microbial communities correlating with Cr^{3+} concentration. Although total biomass (from PLFA and/or viable counts) showed no correlation with Cr^{3+} concentration (P>0.05), relative proportions of PLFA indicative of sulfate reducing bacteria peaked at 10^3 mg kg^{-1} Cr^{3+}, while PLFA indicative of environmental "stress" were positively associated with the highest concentration of Cr^{3+}. The ordination of PLFA profiles together with sample characteristics by principal component analysis further revealed associations between Cr^{3+} and PLFA. However, multi-linear regression of PLFA profiles to predict Cr^{3+} highlighted the fact that the correlation was not linear ($R^2 = 0.80$). The association between PLFA profile and Cr^{3+} concentration was further investigated using artificial neural networks (ANN), an artificial intelligence technique. A predictive (cross-validated) association was found, including 11 hidden nodes. The neural network was a highly accurate predictor for levels of Cr^{3+} as low as 100 mg kg^{-1}. Furthermore, the ANN prediction was observed to depend mostly on the concentrations of PLFA components rather than other sample characteristics.

INTRODUCTION

The release of toxic metal waste into the environment at defense-related and industrial sites has resulted in widespread surface and groundwater contamination (12, 43). Unlike the majority of other toxicants, metal wastes are not biodegradable and having entered the environment, their potential toxicity is controlled to a great extent by geochemical and biological factors (48). Microbiological communities are of primary importance in the bioremediation of metal contaminated soils as they represent a malleable agent that is able to affect virtually all biogeochemical pathways. Microorganisms can alter metal chemistry and mobility through reduction, oxidation, accumulation and immobilization (2, 8, 32, and 49).

In some cases, a specific metal waste may have an imperceptible impact on the total viable biomass of the natural population, while the community structure and metabolic characteristics of the population may be drastically affected (48). Changes in microbial population structure following metal contamination can be determined using a wide range of techniques including ATP assays (4), select enzyme activity assays (7, 37), phylogenetic analysis via the polymerase chain reaction and denaturing gradient gel electrophoresis (PCR-DGGE; 28, 34), and phospholipid fatty acid analysis (3, 18, 41). Of these, only PLFA currently provides a truly direct analysis, rendering it useful for real time monitoring of the microbial population at contaminated sites. As constituents of all eukaryote and bacterial cell membranes, PLFA provide a non-selective means to assay changes in microbial communities in situ. Numerous studies have shown how PLFA analysis can aid in determining the impact of environmental change, *e.g.* exposure to hydrocarbons (13, 35) or metals (17, 18), on the microbial community structure.

However, the information contained within community PLFA profiles is often extremely complex, with typically 50-70 different PLFAs detected in any sample from a single site. Consequently, conclusions can best be drawn from such data using multivariate statistical approaches. These approaches include principal components analysis (46), and/or methods that account for non-linear associations among the lipid biomarkers, such as artificial neural networks (1, 10, 38). The latter approach is particularly suitable for the analysis of PLFA profiles because artificial neural networks (ANN) are learning tools able to identify non-linear associations without requiring assumptions about the underlying mechanisms (*e.g.* "learning from experience", 23). Artificial neural networks were originally developed to mimic nervous systems (21) and have since matured as a statistical tool (11). Their application to environmental monitoring is enjoying increasing popularity due to the development of multi-parametric multi-purpose biosensors (29). The complex nature of biological systems is a consequence of its behavior being as much a result of combined component behaviors as of the interaction between them (19). As a consequence, the distinction of signal from noise in biological associations is fundamentally irreducible and is best achieved by following a machine learning approach (5).

In this study we demonstrate the application of linear and non-linear statistical techniques to map the response of the subsurface micro-eukaryote and bacterial communities to chromium contamination at a heavily contaminated industrial Superfund site located 1.5 miles west of Sault St. Marie, Upper Peninsula, MI. Chromium waste exists in either of two oxidation states, Cr^{3+} or Cr^{6+}, of which Cr^{3+} is both the less toxic, and the only form detected in these samples. Surface samples (0-0.15 m depths) were

extracted and analyzed for PLFA, direct and viable bacterial counts, Cr^{3+}, K, P, Ca and Mg concentration and percent total organic carbon (TOC) and organic matter (TOM). From the PLFA profiles, shifts in total biomass, community structure and physiological status were quantified and compared to Cr concentration using principal components analysis. The ANN was used to establish a predictive association with a sensitivity analysis performed to quantify the contribution of individual PLFA.

MATERIALS AND METHODS

Site description and soil sampling. Samples were obtained from the Cannelton industrial site located 1.5 miles west of Sault Ste. Marie, Upper Peninsula, MI, between October 1997 and August 1998. This is a 75 acre property on the Saint Marie river front that is contaminated with Cr and other heavy metals as the result of waste disposal from a tannery (Northwestern Leather Company) that operated from approximately 1900-1958. A map of the site showing sampling sites and a qualitative environmental interpretation is presented in Figure 1A. Chromium contamination at these sites has been shown to be dominated by the relatively immobile Cr^{3+} (Figure 1B is a graphical representation of the Cr^{3+} distribution in the surface sediments at the Cannelton Tannery Superfund site). Sampling sites were chosen to represent the range of Cr^{3+} concentrations from the highest contamination (~200,000 mg kg^{-1}) to background levels (0-50 mg kg^{-1}). At each site, samples were taken from 0-0.15 m depth and stored at 4°C for viable counts/ chemical analyses or at -80°C for PLFA analysis. Samples were sent on dry ice overnight to the laboratory for subsequent PLFA analysis.

Metal Concentrations. Briefly, soil samples were extracted with nitric acid in pressurized vessels (CEM Inc., Ontario, CA) heated in a microwave (22). This procedure dissolves the more reactive fraction of the sample and leave behind the more resistant silicate minerals. The extracted liquid was diluted and analyzed for total metal concentrations using a Micromass™ Inductively Coupled Plasma-Mass Spectrometer (ICP-MS, Micromass, UK) with a hexapole collision cell. Calibration standards for analyses were prepared using distilled, deionized water and stock standards (J.T. Baker Analyzed, Phillipsburg, NJ). Standards and sample dilutions were prepared under identical solution conditions. All chemicals and reagents used were analytical metal grade or better.

Determination of total carbon. The total carbon present in soil samples was determined by dry combustion using a Leco Carbon Analyzer according to the recommended protocol of the manufacturer (30). Briefly, soil samples were ground in a ball-mill grinder to pass through a 100 mesh sieve. Calibration of the instrument was a two step process that included compensation for the amount of argon in the oxygen cylinder (two blanks containing the recommended amounts of tin and iron accelerators) followed by high (0.9%) and low (0.05%) carbon standards (Leco Corp.). Calibration was repeated until the recorded percent C was within the confidence limits of the standards. After calibration, soils samples (approximately 0.1g) and the tin and iron accelerators were placed in a Leco crucible and total carbon was determined. The percent organic matter was calculated from the total organic carbon measurement as follows:

Percent OM = (TOC)(1/0.58).

This conversion factor is derived from estimates of the percentage of carbon in humus that range from 55-60% (27, 40). All measurements were performed at the Soil and Plant Nutrient Laboratory at Michigan State University.

Viable counts. Viable counts were determined within 48 hours after collection. One gram of soil (wet weight) was placed into 3 ml of sterile 100 mM phosphate buffer (pH 7.4) and vortexed vigorously. The large soil particulates were allowed to settle for one minute, after which the supernatant was serially diluted to extinction in sterile phosphate buffer. R2A agar plates (DIFCO, Detroit, MI) were spread with 100 µl from the serial dilution tubes and incubated at 25°C. Each dilution was plated in triplicate and the plates were counted after seven days of incubation. The total viable count is calculated as the average of three to six plates from the dilution tubes providing optimal distribution of colonies.

Lipid analysis. All solvents used were of GC grade and were obtained from Fisher Scientific (Pittsburgh, PA). All glassware used was washed in a 10% (v/v) Micro cleaning solution (VWR Scientific, Pittsburgh, PA), rinsed 10 times in tap water then 10 times in deionized water. The glassware was then heated at 450°C for 4 hrs in a muffle furnace prior to use. Lipids were extracted from samples (10 g wet weight) using the modified Bligh and Dyer method as described in (50). The total lipids obtained were then fractionated into glyco-, neutral- and polar-lipids (20). The polar lipid was subjected to a sequential saponification/acid hydrolysis/esterification (36). The PLFA were separated, quantified and identified by gas chromatography-mass spectrometry (GC-MS; 52). Fatty acids were identified by relative retention times, comparison with authentic

standards (Matreya Inc., Pleasant Gap, PA) with identifications confirmed by the mass spectra (collected at an electron energy of 70 mV) (44). Fatty acid nomenclature is in the form of "A:BωC' where 'A' designates the total number of carbons, 'B' the number of double bonds , and 'C' the distance of the closest unsaturation from the aliphatic end (ω) of the molecule. The suffixes 'c' for *cis* and 't' for *trans* refer to geometric isomers. The prefixes 'i' , 'a' and 'me' refer to iso and anteiso methyl branching, and mid chain methyl branching, respectively, with cyclopropyl rings indicated by "cy". (25).

Statistical analysis. Results were expressed per gram dry weight of the substrate. Phospholipid fatty acids were analyzed both as pmole g^{-1} soil and as mole percents. Given the large number of samples, for ease of analysis Cr^{3+} concentration (mg kg^{-1}) was coded; 1 = 0-99 (N = 20), 2 = 100-999 (N = 15), 3 = 1 000-9 999 (N = 14), 4 = 10 000 = 99 999 (N = 23), 5 = >100 000 (N = 4). Analysis of variance (ANOVA) was used to determine shifts in relative proportions of specific PLFA with Cr^{3+} concentration (coded). Groupings for ANOVA were assigned *a posteriori*. The ANOVA and correlation analyses between PLFA and Cr^{3+} concentration were performed using Statistica Version 5.1 for Windows software (Statsoft Inc., Tulsa, OK). The same software was used for exploratory statistical analysis by extraction of principal components.

Artificial Neural Net analysis. The ANN analysis was developed in MATLAB 5.3 (The Mathworks, Inc., Natick, MA) environment using feedforward topologies with one hidden layer. The algorithms used are extensions to NetLab (9). The new code written for this study incorporates cross-validation for optimization of topology, bootstrapping for accurate evaluation of each topology attempted, and sensitivity analysis to quantify the

contribution of each PLFA to the prediction of Cr^{3+}. The purpose of cross-validation is to avoid over fitting by using the part of the experimental data that was not used to develop and train the ANN as a validation set. This procedure can be repeated with different validation sets in order to represent the entire data set, therefore bootstrapping each topology evaluated (15). The ANN with the best median error is selected as the best predictor. The ultimate goal of repeated cross-validation is to match the complexity of the PLFA/ Cr^{3+} association with an ANN solution of similar complexity (42). The implementation of cross-validation and bootstrapping of neural networks is not usually performed due to the associated intensive computing load (42). The ANN computations reported here were performed using a dual-Pentium III 600 MHz computer with 500 Mb RAM. The relative importance of each PLFA to predict the target values was calculated by performing sensitivity analysis on the trained ANN. The procedure is briefly outlined below for the general case of *ni* independent variables (PLFA) being used to predict *nj* dependent parameters (*nj* = 1, Cr^{3+} concentration). Due to the non-linearity of the ANN solution, the overall sensitivity results from the combination of the sensitivities were calculated for each experimental value (each individual PLFA of individual profile obtained for every sample). Consequently, the sensitivity of an output parameter $Out_{j=1,2,...nj}$ to an input parameter $In_{i=1,2,...ni}$ was defined as the normalized ratio between variations caused in Out_j by variations introduced in In_j and is represented by the following equation:

$$NS_{i,jc} = (dOut_{j,c} / d\ In_{i,c})(In_{i,c} / Out_{j,c})$$

$$S_i = [\ \Sigma_{j=1,2,\ldots nj;\ c=1,2,\ldots nc}\ (NS_{i,jc})\ ] / [\ \Sigma_{i=1,2,\ldots ni;\ j=1,2,\ldots nj;\ c=1,2,\ldots nc}\ (NS_{i,jc})\ ] \qquad \text{(eq. 1)}$$

i= 1, 2, ... ni; input index
j= 1, 2, ... nj; output index
c= 1, 2, ... nc; sample (case) index

RESULTS

Chemical analyses. The concentration of Cr^{3+} ranged between ~10 mg kg^{-1} at site B5 to between 120 000 – 263 211 mg kg^{-1} at H17. pH values at the different sampling locations (in parenthesis) ranged between 5.2 (C8) to 8.0 (C16, L21). Calcium content ranged between 350 (B9) - 13263 (H15) μg g^{-1}; Potassium was present at between 25 (P23) - 914 (J21) μg g^{-1}); Magnesium was present at 74 (P23) - 4801 (E18) μg g^{-1}; and P at 1.0 (H15) - 109 (J21) μg g^{-1}. Percent total carbon and organic matter ranged between 0.34 - 43.0 % and 0.6-66.4 %, respectively and both showed a significant positive correlation with Cr^{3+} concentration (Table 1).

Biomass. Bacterial abundance was calculated based on the amount of bacterial PLFA recovered at each site (6). As with any conversion factor, it is important to remember that the number of cells can vary by up to an order of magnitude (16). Bacterial abundance at this site as described by PLFA ranged from a minimum of ~6-7 x 10^7 bacteria to a maximum of ~1 x 10^9 in over 40 % of the remaining sites. In general, viable cell counts were approximately 1-3 orders of magnitude lower than were the cell numbers calculated from PLFA content (ranging between 3 x 10^4 cfu g^{-1} to 10^7 cfu g^{-1}). The PLFA content showed no correlation with Cr^{3+} concentration (P>0.05, Table 1; Fig. 2A). PLFA content did show a positive correlation with both percent total organic matter (TOM; P<0.001) and total organic carbon (TOC; P<0.001) (Table 1). In contrast, viable counts did not correlate with either TOC or TOM, although they showed a weak positive correlation with PLFA (P=0.05).

Community structure. Shifts in the PLFA profiles were detected for certain PLFA with the increased Cr^{3+} concentration. Only PLFA which demonstrated significant shifts in relative proportion with Cr^{3+} concentration are reported here and in Fig. 2. Compared to samples with low Cr^{3+} concentrations, the relative proportions of terminally branched chain fatty acids were significantly lower in samples containing >10,000 mg kg^{-1} Cr^{3+} (specifically i16:0, a17:0), and >100,000 mg kg^{-1} Cr^{3+} (specifically a15:0, i15:0, i17:0) (Figure 2B). The relative proportion of 10me16:0 fatty acid, indicative of the presence of sulfate reducing bacteria (14) increased in samples containing up to ~10^3 mg kg^{-1} Cr^{3+}, before decreasing again with the increased Cr^{3+} concentration (Figure 2C). Furthermore 10me18:0, a biomarker commonly used for detection of actinomycetes (17,18, 33), decreased significantly over the Cr^{3+} concentration range (Fig 2D). Finally, samples containing >100,000 mg kg^{-1} Cr^{3+} also contained significantly more normal saturate PLFA which are common to all genera (51, 53; 2E).

Physiological status. Microorganisms lipid profiles are a product of their metabolic pathways and therefore reflect the phenotypic response of the microorganism to its environment (51). In these samples, the relative proportions of 18:1ω7t (*trans*) compared to that of 18:1ω7c (*cis*) increased with increasing Cr^{3+} concentration (Fig. 2F).

PCA, ANN. The ordination of PLFA profiles together with sample characteristics by principal component analysis (PCA) revealed associations between Cr^{3+} concentration and specific PLFA (Fig. 3). The multi-linear regression of PLFA profiles to predict Cr^{3+} highlighted the fact that the correlation was not linear ($R^2 = 0.80$, results not shown). The

association between PLFA profile and Cr^{3+} concentration was further pursued using artificial neural networks (ANN), an artificial intelligence technique. A predictive (cross-validated) association was found for an ANN containing 11 hidden nodes (Fig. 4). The relative sensitivity coefficients (Table 2) assigned to each of the elements of the combined profile of PLFA and soil characteristics shows little association between the Cr^{3+} concentration and most relevant PLFA (highlighted in Fig. 3 and Table 2). This finding was to be expected given the failure to identify a predictive multilinear dependency. Therefore, the accuracy achieved by the ANN predictor was due to the ability of this technique to uncover non-linear associations. It should also be noted that the ANN made a distinction between TOM and TOC (Table 2). Such distinctions between similar parameters often occur when two similar variables are used, one of which is measured with more accuracy.

Table 2 shows the sensitivities of the different PLFA in the ANN prediction of Cr^{3+} concentration, with 20% of the variables responsible for 50% of the predictive value (shaded). Of these PLFA, only 10me16:0 is commonly used as a specific marker for sulfate/metal reducing bacteria (14; 45). Of the remaining PLFA in the list, i15:0, 16:1ω11c, and i17:0 all correlated strongly with 10me16:0 (at $P<0.001$). The PLFA 18:1ω9c (the most sensitive) correlated strongly with 18:2ω6 (indicative of microeukaryote biomass). The PLFA 18:1ω9c is generally taken to be indicative of both Gram-negative prokaryotes and microeukaryotes (31, 53).

DISCUSSION

The lack of any correlation between total biomass (PLFA and/or viable counts) and Cr^{3+} concentration is reflective of the relative lack of toxicity of the Cr^{3+} compared to Cr^{6+}. However, from the shifts in the relative proportions of specific PLFA with increasing Cr^{3+} concentration it was evident that the microbial community was impacted by the contamination. Specifically, the increase in the relative proportion of 10me16:0 fatty acid (most commonly associated with sulfate reducing (14) and iron reducing bacteria *e.g. Geobacteraceae* (45), was associated with the median Cr^{3+} concentration of ~10^3 mg kg^{-1}. Above that concentration, however, the relative proportion of 10me16:0 decreased significantly, indicating a negative impact on sulfate/metal reducing bacteria containing 10me16:0. Terminally branched saturates have been associated with both Gram-positives and anaerobic Gram negatives, (39, 53). In this case, the relative proportions of a number of terminally branched saturate PLFA correlated strongly with that of the 10me16:0 (P<0.001), suggesting that these terminally branched PLFA were generally indicative of anaerobic Gram-negatives sulfate/metal reducing bacteria.

Gram-negative bacteria make *trans* fatty acids as a response to changes in their environment (47). An increase in the ratio of *trans/cis* fatty acids (specifically 16:1ω7t/c and 18:1ω7t/c) has been suggested to be indicative of starvation (26), nutrient stress (35) and/or metal toxicity (17). However the *trans/cis* response to metal toxicity has tended to be contradictory, with inconclusive evidence often presented (18). Herein a significant increase in the proportion of 18:1ω7t to 18:1ω7c was detected when the Cr^{3+} concentration exceeded 10^4 mg kg^{-1} (Fig 2F), however, this shift was not reflected in the ratio of 16:1ω7t/16:1ω7c which stayed constant over the entire concentration range (data

that it will not respond to values below the limit, set at approximately 100 mg kg^{-1} (Fig. 4).

CONCLUSIONS

Increased levels of Cr^{3+} were correlated with the decrease in the PLFA representative of sulfate/metal reducing bacteria, and the increase in relative proportions of lipids indicative of metabolic stress. Microbial community composition, as revealed by the PLFA characterization, was non-linearly associated with Cr^{3+} concentration.

The artificial neural network was shown to be a highly accurate predictor of Cr^{3+}concentration, largely outperforming the conventional linear regression techniques. Furthermore, the ANN prediction was observed to depend mostly on the concentration of PLFA components, highlighting the biological nature of the non-linear association with Cr^{3+} concentration. These results suggest that effective monitoring of both heavy metal concentration and, conceivably, also of heavy metal immobilization activity, should be based on the biochemical composition of the microbial cell membranes. This conclusion can be rationalized by noting that it is the cell membrane that mediates interactions between biological activity and environmental conditions. Therefore, PLFA signature holds great potential as a biosensor for soil bioremediation.

ACKNOWLEDGEMENTS

This work was supported by grants from the U.S. Department of Energy, Natural and Accelerated Bioremediation Research (NABIR) program to DCW, award No. DE-FC02-96ER62278, and to CCB, JCS, and SMP award No. ERKP280. JSA also acknowledges

the financial support by grant FMRH/BSAB/60/98 Fundação para a Ciência e Tecnologia /MCT.

REFERENCES

1. **Almeida, J.S., K.T. Leung, S.J. Macnaughton, C.A. Flemming, M.H. Wimpee, G.A. Davis, and D.C. White.** 1998. Mapping changes in soil microbial community composition signaling bioremediation. Bioremediation J. **1:**255-264.
2. **Avery S.V.** 1995. Cesium accumulation by microorganisms: uptake mechanisms, cation competition, compartmentalization and toxicity. J. Ind. Microbiol. **14:** 76-84.
3. **Bååth, E., M. Diaz-Ravina, A. Frostegård, and C.D. Campbell.** 1998. Effect of metal-rich sludge amendments on the soil microbial community. Appl. Env. Microbiol. **64:** 238-245.
4. **Babich H. and G. Stotzky.** 1985. Heavy metal toxicity to microbe-mediated ecologic processes: A review and potential application to regulatory policies. Environ. Res. **36:**111-137.
5. **Baldi, P. and S. Brunak.** 1998. Bioinformatics – the machine learning approach, p 39-57 MIT Press, Cambridge, Massachusetts.
6. **Balkwill, D.L., F.R. Leach, J.T. Wilson, J.F. McNabb, and D.C. White.** 1988. Equivalence of microbial biomass measures based on membrane lipid and cell wall components, adenosine triphosphate, and direct counts in subsurface sediments. Microbial. Ecol. **16:** 73-84.
7. **Barnhart C.L.H. and J.R. Vestal.** 1983. Effects of environmental toxicants on metabolic activity of natural microbial communities. Appl. Environ. Microbiol. **46:** 970-977.

8. **Beveridge T.J.** 1989. Role of cellular design in bacterial metal accumulation and mineralization. Ann. Rev. Microbiol. **43:**147-171.

9. **Bishop, C.M.** 1995. Neural networks for pattern recognition. Clarendon Press, Oxford, UK. http://www.ncrg.aston.ac.uk/netlab/index.html.

10. **Brandt, C.C., J.C. Schryver, S.M. Pfiffner, A.V. Palumbo, and S. Macnaughton**.1999. Using artificial neural networks to assess changes in microbial communities, pp. 1-6. *In*: A.Leeson and B.C. Alleman (Eds) Bioremediation of Metals and Inorganic Compounds. Battelle Press, Columbus, OH.

11. **Cheng, B and D.M. Titterington.** 1994. Neural networks - A review from a statistical perspective. Stat. Sci. **9:** 2-30.

12. **Cornish, J.E., W.C. Golberg, R.S. Levine, and J.R. Benemann.** 1995. Phytoremediation of soils contaminated with toxic elements and radionuclides, p. 55-62. *In* R.E. Hinchee, J.L Means, and D.R. Burris (eds.), Bioremediation of Inorganics: Third International *In situ* and on-site bioreclamation Symposium, No. 10. Battelle Press, Columbus, OH.

13. **Cox, E.E., D.W. Major, D.W. Acton, T.J. Phelps, and D.C. White,** 1994. Evaluating trichloroethylene biodegradation by measuring the *in situ* status and activities of microbial populations, p. 37-49. *In* R.E. Hinchee, A. Leeson, L. Semprini, and S.K. Ong, (eds), Bioremediation of chlorinated polycyclic aromatic compounds. Lewis Publications, Ann Arbor, MI.

14. **Dowling, N.J.E., F. Widdel, and D.C. White.** 1986. Phospholipid ester-linked fatty acid biomarkers of acetate-oxidizing sulphate-reducers and other sulfide forming bacteria. J. Gen. Microbiol. **132:**1815-1825.

15. **Efron, B., and R.J. Tibshirani.** 1993. The empirical distribution function and the plug-in principle. An introduction to the bootstrap, pp. 31-37. Chapman and Hall, NY.

16. **Findlay, R.H. and F.C. Dobbs.** 1993. Quantitative description of microbial communities using lipid analysis, p. 271-284. *In* P.F. Kemp, B.F. Sherr, E.B. Sherr, and J.J. Cole (eds.), Handbook of methods in aquatic microbial ecology. Lewis Publishers, Boca Raton, FL.

17. **Frostegård, A., A. Tunlid, and E Bååth.** 1993. Phospholipid fatty acid composition, biomass, and activity of microbial communities from two soil types experimentally exposed to different heavy metals. Appl. Environ. Microbiol. **59:**3605-3617.

18. **Frostegård, A., A. Tunlid, and E. Bååth.** 1996. Changes in microbial community structure during long-term incubation in two soils experimentally contaminated with metals. Soil Biol. Biochem. **28:**55-63.

19. **Gallagher, R and T. Appenzeller.** 1999. Beyond reductionism. Science. **284:**79-79.

20. **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.

21. **Hayken, S.** 1994. Neural networks-A comprehensive foundation, pp. 397-443 Macmillan College Pub., New York, NY.

22. **Hewitt A.D. and Reynolds C.M.** 1990. Dissolution of metals from soils and sediments with a microwave-nitric digestion technique. Atomic Spectrosc. **11:** 187-192.

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

24. **Kaneda, T.** 1991. Iso- and anteiso-fatty acids in Bacteria: biosynthesis, function, and taxonomic significance.. Microbiol. Rev. **55:**288-302.

25. **Kates, M.** 1986. Techniques in lipidology: isolation, analysis and identification of lipids. Second edition, Elsevier Press, Amsterdam.

26. **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 of a porous medium. Appl. Environ. Microbiol. **60:**3292-3299.

27. **Kononova, M. M.** 1966. Soil organic matter; its nature, its role in soil formation and in soil fertility. Translated from the Russian by T.Z. Nowakowsli and A.C.D. Newman. Oxford, New York, Pergamon Press.

28. **Konopka, A., T. Zakharova, M. Bischoff, L. Oliver, C. Nakatsu and R.F. Turco.** 1999. Microbial biomass and activity in Lead-contaminated soil. Appl. Environ. Microbiol. **65:**2256-2259.

29. **Kress-Rogers, E.** 1997. Biosensors and electronic noses for practical applications, pp. 3-39. Handbook of biosensors and electronic noses for medicine food and the environment, CRC Press, NY.

30. **Leco Corporation Application Bulletin No. 203-601-224**; Carbon in Soil, Rock, Limestone and Similar Material; Leco Corporation; 3000 Lakeview Avenue; St. Joseph, MI 49085.

31. **Lösel, D.M.** 1988. Fungal lipids, p. 699-806. *In* C. Ratledge, and S.G. Wilkinson (eds.), Microbial lipids. Academic Press, London.

32. **Lovley D.R.** 1995. Bioremediation of organic and metal contaminants with dissimilatory metal reduction. J. Ind. Microbiol. **14:**85-93.

33. **Macnaughton S.J. and A.G. O'Donnell.** 1994. Tuberculostearic acid as a means of estimating the recovery (using dispersion and differential centrifugation) of actinomycetes from soil. J. Microbiol. Methods **20:**69-77.

34. **Macnaughton, S.J., J.R. Stephen, Y-J. Chang, A.P. Peacock, C.A. Flemming, K.T. Leung, and D.C. White.** 1999a. Characterization of metal-resistant soil eubacteria by polymerase chain reaction-denaturing gradient gel electrophoresis with isolation of resistant strains. Can. J. Microbiol. **45:**116-124.

35. **Macnaughton, S.J., J.R. Stephen, A.D. Venosa, Y-J. Chang, G.A. Davis, and D.C. White.** 1999b. Microbial population changes during bioremediation of an experimental oil spill. Appl. Environ. Microbiol. **65:**3566-3574.

36. **Mayberry, W.R., and J.R. Lane.** 1993. Sequential alkaline saponification /acid hydrolysis/esterification: a one tube method with enhanced recovery of both cyclopropane and hydroxylated fatty acids. J. Microbiol. Methods. **18:**21-32.

37. **Montuelle B., X. Latour, B. Volat, and A.M. Gounat.** 1994. Toxicity of heavy metals to bacteria in sediments. Bull. Environ. Contam. Toxicol. **53:**753-758.

38. **Noble, P.A., J.A. Almeida, and C.R. Lovell.** 2000. Application of neural computing models for interpreting phospholipid fatty acid profiles of natural microbial communities Appl. Environ. Microbiol., **66:**694-699.

39. **O'Leary, W.M., and S.G. Wilkinson.** 1988. Gram-positive bacteria, pp. 117-201. *In* C. Ratledge, and S.G. Wilkinson, (eds.), Microbial lipids. Academic Press, London.

40. **Paul, E.A., and F.E.Clark.** 1996. Soil Microbiology and Biochemistry. Academic Press, San Diego.

41. **Pfiffner, S.M., C.C. Brandt, J.C. Schryver , A.V. Palumbo, and. J.S. Almeida.** 1999. Using artificial neural networks to assess microbial communities. pp. 205-211. *In*: G.A Uzochukwu and G.B. Reddy (Eds.). Proceedings of the 1998 National Conference on Environmental Remediation Science and Technology, Battelle Press, Columbus, OH.

42. **Reed, R.D. and R.J. Marks.** 1999. Neural smithing – supervised learning in feedforward artificial neural networks, p.257-264. MIT Press, Cambridge, Massachusetts.

43. **Riley, R.G., and J.M. Zachara.** 1992. Chemical contaminants of DOE lands and selection of contaminant mixtures for subsurface science research. DOE/ER-0547T. National Technical Information Service, US Department of Commerce, Springfield, VA.

44. **Ringelberg, D.B., J.D. Davis, G.A. Smith, S.M. Pfiffner, P.D. Nichols, J.B. Nickels, J.M. Henson, J.T. Wilson, M. Yates, D.H. Kampbell, H.W. Reed, Stocksdale, T.T., 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. Ecol. **62:**39-50.

45. **Rooney-Varga, J.N., R.T. Anderson, J.L. Fraga, D. B. Ringelberg, and D.R. Lovley.** 1999. Microbial communities associated with anaerobic benzene degradation in a petroleum-contaminated aquifer. Appl. Environ. Microbiol. **65:**3065-3063.
46. **Rosswell T., and E. Kvillner.** 1978. Principal components and factor analysis for the description of microbial populations. Adv. Microbial Ecol. **2:**1-48.
47. **Sikkema, J., J.A.M. deBont, and B. Poolman.** 1995. Mechanisms of membrane toxicity of hydrocarbons. Microbiol. Rev. **59:**201-222.
48. **Sterritt R.M. and J.N. Lester.** 1980. Interactions of heavy metals with bacteria. Sci. Total. Environ. **14:**5-17.
49. **Valentine N.B., H. Bolton Jr., M.T. Kingsley, C.R. Drake, D.L. Balkwill, and A.I. Plymale.** 1996. Biosorption of cadmium, cobalt, nickel and strontium by a *Bacillus simplex* strain isolated from the vadose zone. J. Ind. Microbiol. **16:**189-196.
50. **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.
51. **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.

52. **White D.C. and D.B. Ringelberg.** 1998. Signature lipid biomarker analysis, pp 225-272. . *In*: Techniques in Microbial Ecology R.S. Burlage, R. Atlas, D.Stahl, G.Geesey and G. Sayler (*Eds*) Oxford University Press, New York, NY.

53. **Wilkinson, S.G.** 1988 Gram-negative bacteria, p. 299-488. *In* C. Ratledge, and S.G. Wilkinson (eds.), Microbial lipids. Academic Press, London.

FIGURES

Figure 1: (A) Qualitative map image with sample locations; (B) map of Cr^{3+} contamination (mg kg^{-1}).

Figure 2: Box and whisker plots of the specific PLFA (mole percents) with different Cr^{3+} contamination. (A) Total PLFA; (B) Terminally branched saturates; (C) 10me16:0; (D) 10me18:0; (E) normal saturates; (F) 18:1ω7t/18:1ω7c. The x axis is of Cr^{3+} concentration (encoded, 1 = 0-99 (N = 20), 2 = 100-999 (N = 15), 3 = 1 000-9 999 (N = 14), 4 = 10 000 = 99 999 (N = 23), 5 = >100 000 (N = 4)).

Figure 3: A principal components analysis using all 57 PLFA variables as well as selected soil characteristics (soil moisture (wetland), total biomass (PLFA and viable counts) Ca, P, Mg, K, TOM, TOC, pH). Factors 1 and 2 provided 21 % and 14.9 % of the variance, respectively. The highlighted PLFA correspond to the top parameters accounting for 50% of ANN predictive sensitivity (Table2).

Figure 4: ANN predictive accuracy - valid for values above the lower threshold level of ~ 100 mg kg^{-1}.

Figure 5: Relative (bars) and cumulative (solid line) sensitivities of the optimized NN to specific PLFA and chemical (Ca, P, Mg, K, TOM, TOC) and other selected parameters (moisture content (wetland), total biomass (PLFA and viable counts)). PLFA are arranged in order of their importance to the prediction of Cr^{3+} concentration, with the top

20 % of variables responsible for >50% of the predictive value (dotted lines). Table 2 shows the precise sensitivity values.

TABLE 1. Correlation table for Cr^{3+}, Biomass PLFA, viable counts, percent total organic matter and % total carbon. Significant positive correlations are indicated by *P* values.

	Cr^{3+} (mg kg^{-1})	Biomass (PLFA)	Viable counts (cfu g^{-1})	% Total organic matter	% Total carbon
Cr^{3+} (mg kg^{-1})	1.0000	0.1055	0.1552	0.3532	0.3863
		P=0.407	*P*=0.221	*P*=0.004	*P*=0.002
Biomass	0.1055	1.0000	0.3498	0.4918	0.4945
(PLFA)	*P*=0.407		*P*=0.005	*P*<0.001	*P*<0.001
Viable counts	0.1552	0.3498	1.0000	-0.0333	-0.0299
(cfu g^{-1})	*P*=0.004	*P*=0.005		*P*=0.794	*P*=0.815
% Total	0.3532	0.4918	-0.0333	1.0000	0.9297
organic matter	*P*=0.004	*P*<0.001	*P*=0.794		*P*<0.001
% Total	0.3863	0.4945	-0.0299	0.9297	1.0000
carbon	*P*=0.002	*P*<0.001	*P*=0.815	*P*<0.001	

TABLE 2. Sensitivity of ANN predictions of Cr^{3+} concentration. Twenty percent of all variables (highlighted) are responsible for 50% of the predictive value.

Variables	Average Sensitivity	Cumulative Sensitivity	Variables	Average Sensitivity	Cumulative sensitivity
18:1ω9c	6.6%	6.6%	18:2a	1.2%	84.5%
i17:0	4.8%	11.4%	br17:0a	1.0%	85.7%
18:1ω7c	4.6%	16.0%	Biomass	1.1%	86.8%
10me18:0	4.3%	20.3%	15:1	1.0%	87.8%
a17:0	4.1%	24.4%	Phosphorus	1.0%	88.8%
i15:1ω11c	4.0%	28.4%	15:1b	0.9%	89.7%
I15:1a	3.7%	35.7%	Wetland	0.9%	90.6%
16:1ω5c	3.6%	35.6%	br15:0b	0.9%	91.5%
i15:0	3.1%	38.8%	cy19:0	0.8%	92.3%
20:1ω9c	3.0%	41.9%	Magnesium	0.7%	93.0%
16:1ω11c	2.9%	44.8%	i14:0	0.7%	93.7%
10me16:0	2.5%	47.3%	18:0	0.7%	94.4%
Br18:1	2.5%	49.8%	16:1ω7c	0.7%	95.0%
a15:0	2.4%	52.2%	23:0	0.6%	95.7%
i16:0	2.4%	54.6%	br16:0	0.6%	96.3%
%TOM	2.4%	57.0%	Potassium	0.4%	96.7%
16:0	2.3%	59.3%	11me16:0	0.4%	97.1%
cy17:0b	2.2%	61.6%	20:5ω3	0.4%	97.5%
17:0	2.0%	63.6%	12me18:0	0.3%	97.8%
15:0	2.0%	65.6%	20:0	0.3%	98.2%
20:3ω6	1.9%	67.5%	cy17:0a	0.3%	98.5%
Calcium	1.8%	69.3%	i15:1c	0.3%	98.8%
21:0	1.7%	71.0%	18:2ω6	0.2%	99.0%
pH	1.7%	72.7%	14:0	0.2%	99.2%
12me16:0	1.6%	74.3%	22:0	0.2%	99.3%
16:1ω7t	1.5%	75.8%	i16:1	0.2%	99.5%
18:3ω3	1.4%	77.2%	16:2	0.1%	99.6%
%TOC	1.3%	78.5%	24:0	0.1%	99.8%
i17:1ω8	1.3%	79.7%	br15:0c	0.1%	99.9%
br15:0b	1.3%	81.0%	20:4ω6	0.1%	100%
br17:0b	1.2%	82.2%	Viable counts	0.0%	100.0%
18:1ω7t	1.2%	83.4%	18:1ω5c	0.0%	100.0%

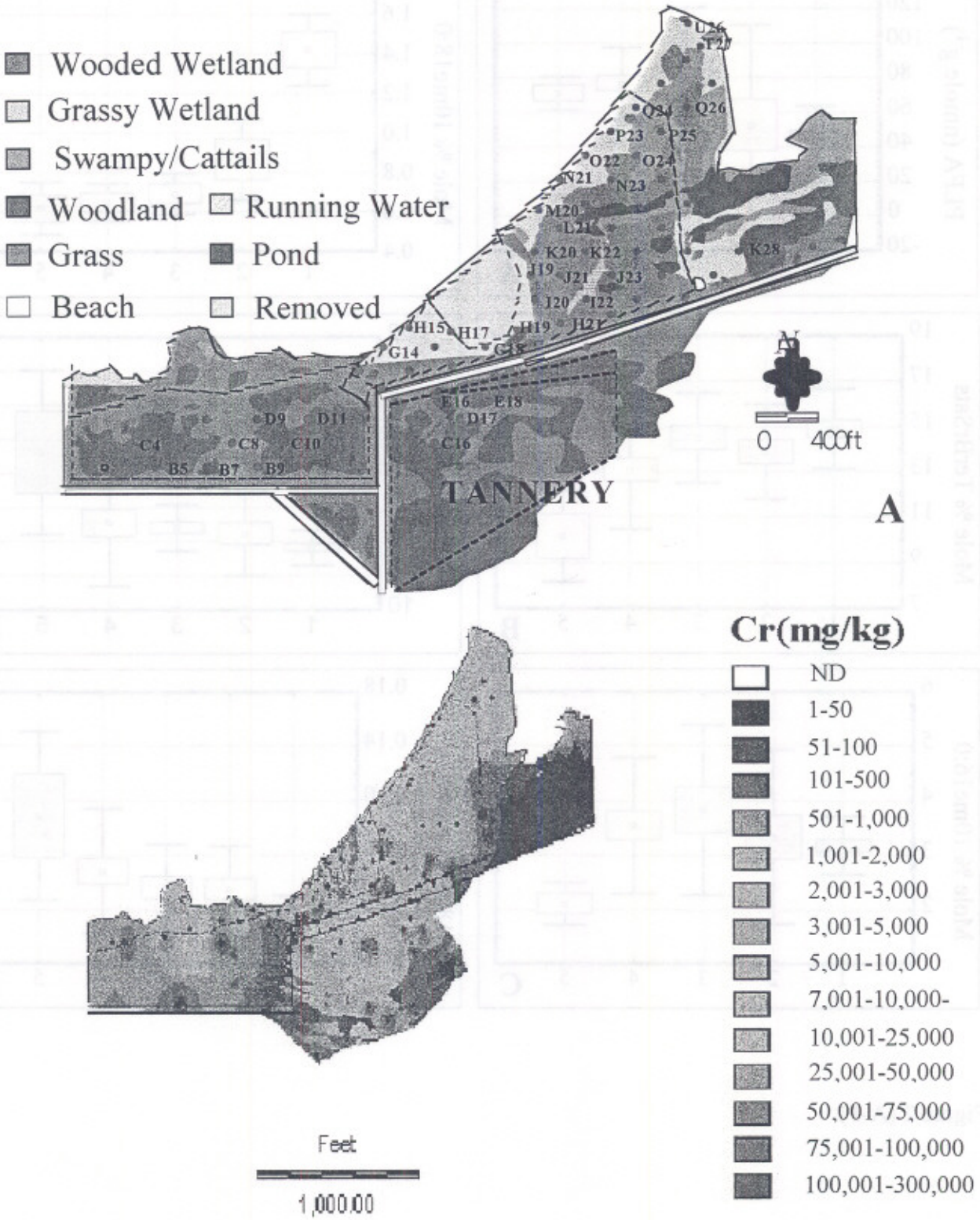

Figure 1 (A and B)

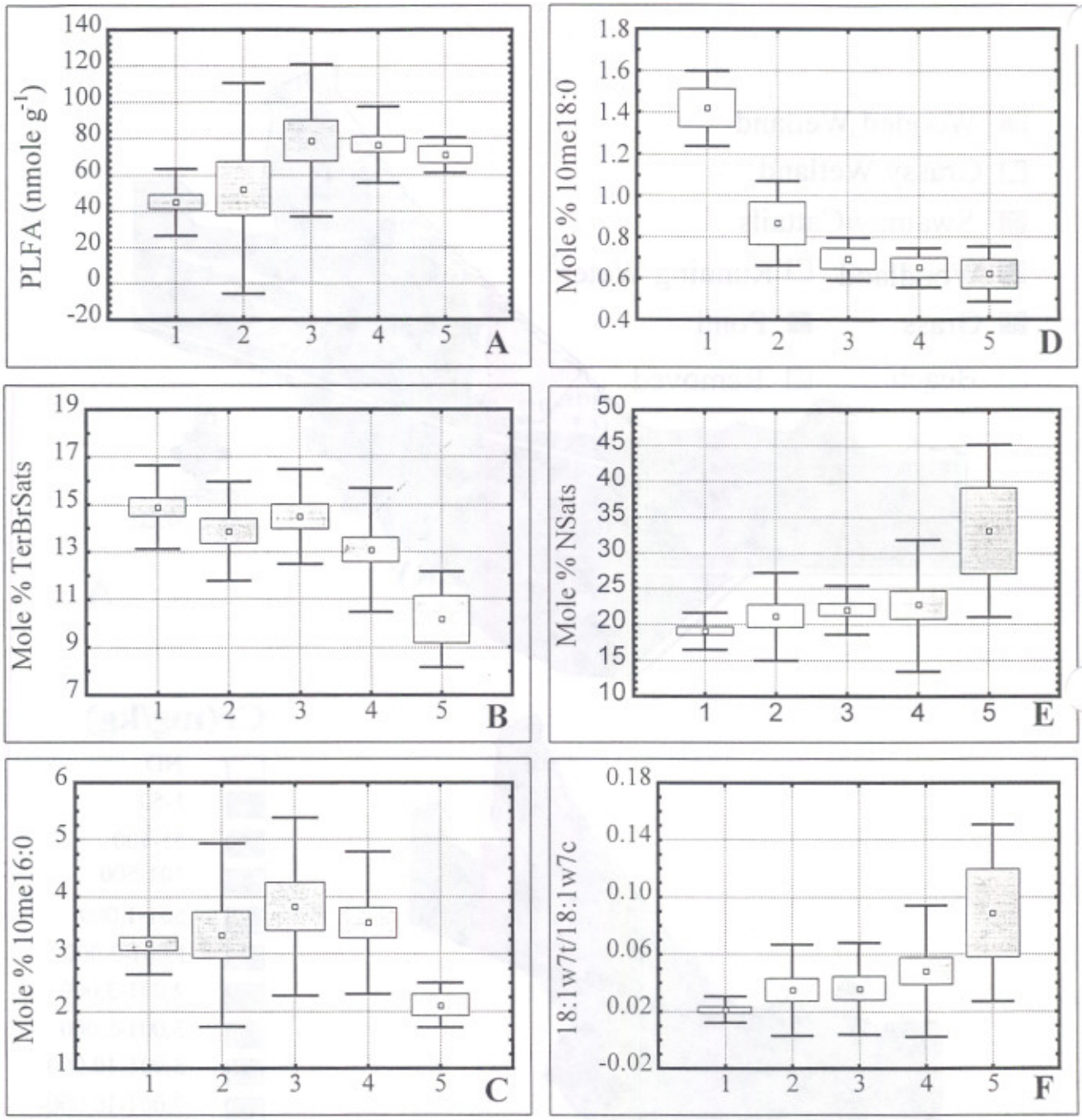
PLFA (nmole g-1)
140
120
100
80
60
40
20
0
-20
1 2 3 4 5
A
Mole % 10me18:0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
1 2 3 4 5
D
Mole % TerBrSats
19
17
15
13
11
9
7
1 2 3 4 5
B
Mole % NSats
50
45
40
35
30
25
20
15
10
1 2 3 4 5
E
Mole % 10me16:0
6
5
4
3
2
1
1 2 3 4 5
C
18:1w7t/18:1w7c
0.18
0.14
0.10
0.06
0.02
-0.02
1 2 3 4 5
F

Figure 2 (A-F)