An inter-laboratory comparison of multi-enzyme and multiple substrate-induced respiration assays to assess method consistency in soil monitoring.

MS for: Biology and Fertility of Soils

Rachel Creamer[1,4], Pat Bellamy[1], Helaina Black[2], Clare M Cameron[2], Colin D Campbell[2,5], Paul Chamberlain[3], Jim Harris[1], Nisha Parekh[3], Mark Pawlett[1], Jan Poskitt[3], Dote Stone[1] and Karl Ritz[1]

1. Natural Resources Department, School of Applied Sciences, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, U.K.; 2. The Macaulay Institute, Craigiebuckler, Aberdeen, AB15 8QH, U.K.; 3. CEH Lancaster, Lancaster Environment Centre, Bailrigg, Lancaster, LA1 4AP, U.K.; 4. Teagasc, Johnstown Castle Environment Research Centre, Wexford, Ireland. 5 Department Soil and Environment, Swedish University Agricultural Sciences, Uppsala, Sweden.

Corresponding author address: Teagasc, Johnstown Castle Environment Centre, County Wexford, Ireland. Rachel.creamer@teagasc.ie

Abstract

The use of indicators in soil monitoring schemes to detect changes in soil quality is receiving increased attention, particularly the application of soil biological methods. However, to date, the ability to compare information from different laboratories applying soil microbiological techniques in broad-scale monitoring has rarely been

taken into account. This study aimed to assess the consistency and repeatability of two techniques that are being evaluated for use as microbiological indicators of soil quality; multi-enzyme activity assay and multiple substrate induced respiration (MSIR). Data was tested for intrinsic (within-plate) variation, inter-laboratory repeatability (geometric mean regression and correlation coefficient) and land-use discrimination (principal components analysis, PCA). Intrinsic variation was large for both assays suggesting that high replicate numbers will be required. Inter-laboratory repeatability showed diverging patterns for the enzyme assay and MSIR. Discrimination of soils was significant for both techniques with relatively consistent patterns, however combined laboratory discrimination analyses for each technique showed inconsistent correspondence between the laboratories. These issues could be addressed through the adoption of reliable analytical standards for biological methods along with adequate replication. However, until the former is addressed, dispersed analyses are not currently advisable for monitoring schemes.

Key words: CLPP, multiple substrate induced respiration, repeatability, soil enzymes, soil monitoring.

Introduction

The need to define, assess and monitor soil quality is increasing with a heightened awareness of soil's fundamental role in delivering ecosystem services and therefore as part of the progressive development of international and national strategies for soil protection. Large-scale (e.g. national) monitoring of soils is increasingly being targeted

to provide information on the status of soil functioning and the potential of any such changes in soils due to pressures or drivers e.g. climate change, atmospheric pollution or land use / management practices. This requires robust indicators of soil quality that can not only provide information on the current status of a soil but can also support an accurate assessment of changes to soils over time. Soil quality is a complex and much debated concept that embodies biological, physical and chemical properties of soils at a wide range of time scales. Monitoring of soil quality requires repeat sampling at often extended, time intervals and over diverse spatial landscapes. A fundamental requirement is therefore that the methods used to assess soil quality are robust and capable, within established degrees of confidence, of producing comparable information over extended spatial and temporal scales. Whilst the degree of accuracy (i.e. the extent to which an estimate represents the true value) is fundamentally important, in a monitoring context the methods precision (i.e. reproducibility or repeatability) should be the first priority since it is equally important. Given the extensive scale and protracted nature of large-scale soil monitoring, which can generate 1000's of soil samples, it is entirely likely that soil analyses within and between monitoring periods will be carried out in different facilities, by different people or in entirely different laboratories. Therefore the question of repeatability is important to understand. There appears however to be little information available on the precision or repeatability of commonly proposed biological indicators of soil quality. In respect to broad-scale soil monitoring if data is to be adequately comparable, then it needs to be established that techniques used to estimate biological properties or processes are both accurate and reproducible in all circumstances, and that the standard operating procedures (SOPs) can produce data consistently.

Amongst the plethora of potential biological indicators of soil quality, those based on microbial activity measurements have to demonstrable utility that are comparable between laboratories and monitoring periods or surveys. The application of soil microbiological methods as indicators of soil quality is becoming increasingly accepted and tested (e.g. Breure et al. 2003; Kubat 2003; Francaviglia 2004; Black et al. 2008) since soil biological properties and processes have been proposed as good measures of ecological soil processes and interpretable in terms of maintaining ecosystem function (Bardgett et al. 2005). There are numerous international standards for some techniques such as the determination of soil microbial biomass (ISO 1997) and hand sorting and formalin extraction of earthworms (ISO 2006). However, it would appear that there is little information available on the accuracy or repeatability of commonly proposed microbiological indicators of soil quality.

Univariate gross measures such as basal respiration (ISO 2002) or dehydrogenase activity (ISO 2005), particularly when taken in the context of the size of the soil microflora (Anderson and Domsch 1978), have been shown to effectively reflect the environmental and physiological status of soils (Bastida et al. 2006; Oliveira and Pampulha 2006; Ros et al. 2006; Fließbach et al. 2007). However, multivariate soil activity profiling methods offer a considerably higher degree of sensitivity and discrimination in assessing the functional status of soils, and hence utility in a monitoring context. Such profiling techniques include those based upon the simultaneous determination of a range of enzyme activities (Marx et al. 2001) or carbon substrate utilisation rates (Degens and Harris 1997; Campbell et al. 2003). Both of these approaches are finding increasing application in understanding factors which govern soil microbial activity, and soil biotic responses to environmental factors and

management (e.g. Bending et al. 2004; Kandeler et al. 2006; Lalor et al. 2007; Winding and Hendriksen. 2007; Campbell et al. 2008; Wakelin et al. 2008). They are particularly pertinent to large-scale monitoring contexts since relatively high-throughput techniques, based upon microtitre-plate systems, are now available. However, to date there has been no published assessment of the accuracy or repeatability of these assays within and between laboratories. We assessed for both these aspects by simultaneously applying a multi-enzyme activity assay and a multiple substrate-induced respiration (MSIR) assay to a range of soil samples split between three laboratories for MSIR and two laboratories for multi-enzyme activity assay, in order to establish the precision and accuracy of these two tests and consequently review the methodological aspects of their applicability as soil quality indicators for large-scale (e.g. national) soil monitoring.

Materials and methods

Sample collection and preparation

Three contrasting land uses were sampled, viz. arable, grassland and woodland, in each of three geographical regions within Great Britain to obtain a total of nine bulked soil samples. The sampling sites were prescribed to provide a range of soil physical and chemical characteristics (Table 1) and to ensure a range and contrast in the soil microbiological properties being considered in this study. Sampling was conducted in March 2006, when five randomly distributed soil sub-samples (0-10 cm depth) were collected from each site, bulked within site, passed through a 2 mm sieve and stored at 4^{o}C prior to use.

Soil physical and chemical characteristics

Soil pH was measured on fresh soil using an electrode in a 1:2.5 soil:water slurry (BS ISO 10390 1994). Loss-on-ignition (SOM) was determined by oven drying 5 g of fresh soil at 105°C for 3 h to determine soil moisture content, followed by combustion at 450°C for 2.5 h. Presence of calcium carbonate was assessed qualitatively using the field method of applying drops of 10% HCl to determine presence/absence by effervescence (Hodgson 1997). Water-holding capacity of sieved soil was measured using a Haines-funnel system, whereby 100 ml of water were added to 50 g of fresh soil in a funnel for 30 minutes. Excess water collected from the funnel system and its volume measured (Jenkinson and Powlson 1976).

Inter-laboratory exercise

At the instigation of the inter-laboratory exercise, each soil sample was thoroughly mixed, adjusted to 40% water holding capacity and coned into 3 sub-samples. Three sets of sub-samples for all soils were packed in cool boxes and transported to each of the 3 laboratories; hereafter denoted Laboratory 1, 2 and 3. The samples were then pre-incubated at each laboratory at 25^{o}C in the dark for 7 days, prior to application of two microbial activity assays. All nine soil samples were incubated on the same day at all three laboratories to ensure that the methodologies were consistent.

Multi-enzyme activity assay

This multi-enzyme activity assay was applied by Laboratory 2 and 3 only to five soil samples only (Nos. 2, 4, 5, 7, 8) and a standard (STD), which was based upon a freeze-dried sub-sample of soil number 2. The activities of eight hydrolytic enzymes, based upon the hydrolysis of methylumbelliferone (4-MUB)-containing substrates were determined using a method based on that of Marx et al. (2001). The following enzyme activities were measured: B-cellobiohydrolase (CELL), N-acetyl-β-glucosaminidase (GLUC), β-glucosidase (GLYC), acid phosphatase (PHOS), β-galactosaminidase (GALA), β-xylosidase (XYLO), β-galactosidase (GALS) and sulfatase (SULF). These were determined based upon hydrolysis rates of 4-MUB- β—D-cellobioside, 4-MUB-N-acetyl- β-glucosaminide, 4-MUB- β-D-glucoside, 4-MUB-phosphate, 4-MUB N-acetyl-β-D-galactosaminide, 4-MUB β-D-xyloside, 4-MUB β-D-galactopyranoside and 4-MUB sulphate respectively. Sub-samples (0.5g) of incubated soil were weighed into a polypropylene centrifuge tube and 50 ml of deionised water added. The tube was shaken on an end-over-end shaker for 30 min and then soil suspension was transferred to a beaker. A homogenous suspension was obtained by vigorously stirring and aliquots of 50 μl were withdrawn and dispensed into 96 well black microplates, adopting a design that incorporated three analytical replicates per sample per substrate and two soil samples per plate. A sample blank was prepared by dispensing 50 μl of deionised water in duplicate for each substrate. 50 μl of 0.1 M MES buffer was added to each well. A 4-MUB standard curve was prepared (0, 10, 30, 50 μM) in triplicate for each sample and for the sample blank to account for the degree of fluorescence quenching. Finally 100 μl of 1 mM substrate solutions were added to the sample and blanks. The reagents were mixed for 5 s by placing the plate onto the plate reader, covered to prevent any contamination, and then incubated at 30°C for 3h. Fluorescence (excitation 360 nm;

emission 460 nm) was measured after 3 h using a Gemini EM (Molecular Devices, UK) microplate reader. Fluorescence was converted into the amount of 4-MUB and calculated as nmol g^{-1} soil h^{-1}.

Multiple substrate-induced respiration assay (MSIR)

A multiple substrate-induced respiration assay (MSIR) was carried out using the MicroResp™ technique (Campbell et al. 2003). The MicroResp™ assay was applied to all nine soils by all three laboratories. The assay utilised seven sole-C substrates, *viz*. n-acetyl glucosamine (NAGA), γ-aminobutyric acid (GABA), L-arginine, citric acid, D-glucose, α-ketoglutaric acid (AKGA), L-malic acid, plus water only to provide a basal respiration measure. Substrates were selected following the 15 suggested substrates given in Campbell et al. (2003) and providing a spectrum of substrates such as acids, basic sugars and proteins. Substrates were prepared using reagents from the same source and batch (confirmed via batch number) and then distributed between laboratories. For each substrate a concentration of 30 mg substrate g-1 soil water was prepared, except arginine, where 7.5 mg l^{-1} achieved the maximum response from the soil. These concentrations were determined by Vmax response curves of previous trial assays. After each substrate was prepared, a 25 µl measure was dispensed into each well in 96 deep-well microtitre plates (12 wells per substrate). Substrates were allocated in blocks of 3 cells (4 blocks of 3 cells each, were realised) randomly within plates, to compensate for possible edge effects. Plates were constructed so that each plate provided a full replicated suite of substrates for one soil sample. Therefore nine plates were run in total at each laboratory.

Soil (400 µl volume) was added into the deep-well plate using a filling device (Campbell et al. 2003) which ensures that the same volume of soil is added into every cell and the plate tapped twice to ensure consistent packing. This resulted in final soil water content in each well of approximately 60%. A colorimetric gel detector plate had been prepared 10 days in advance of the experiment, using indicator dye, cresol red (12.5 ppm, w/w), potassium chloride (150 mM), and sodium bicarbonate (2.5 mM), set in 150 µl of Purified agar (1%). This method uses colorimetric a microtitre plate reader to attain value of colour change in the indicator dye to indicate carbon dioxide evolved. The absorbance (A_{570}) value for each cell within the gel indicator plates was determined at 570 nm wavelength using a microtitre plate reader immediately prior to incubation and again after a 6 hour incubation period at 25^{o}C. Independent calibration curves were performed for each laboratory, to compensate for different plate readers being used at each location. Carbon dioxide concentrations were standardised for each well by dividing the absorbance (A_{570}) value by the average obtained across all wells within each plate. A respiration rate CO_2 rate (µg CO_2-C/g/h) was calculated by converting the 6 h % CO_2 to µg/g/h CO_2-C using gas constants and constants for headspace volume in the well (945 µl), fresh weight of soil per well (g), incubation time (h) and soil sample % dry weight.

Statistical methods

Respiration and fluorescence data were tested for normality using the Anderson–Darling test (significance level of $p<0.05$). Respiration data showed a non-normal distribution, but further transformations did not improve the dataset. Therefore statistical tests were applied which did not require a normal distribution. The fluorescence data showed a

normal distribution. For both tests, we expressed the variation of replicates within laboratories by calculating the coefficient of variation (CoV) as a measure of variation occurring independently of user or laboratory bias. This allows us to account for the intrinsic variation in laboratory replicates, for example, from each laboratory, the respiration data included 12 replicates for each substrate per soil sample while the fluorescence data consisted of 3 replicates for each enzyme per soil sample.

Correlation between data pairs for laboratories was assessed using the correlation coefficient. The precision and repeatability of the methods across the three laboratories were assessed by pair-wise regression between results of individual laboratories using geometric mean regression (GMR), also known as reduced major axis regression (Sokal and Rohlf 1981). The slope of each line was compared to unity (1) to test the hypothesis that there was no significant difference between the estimates from the two laboratories. This method determines the difference between absolute values for each soil sample across the laboratories. R^2 (the proportion of variance explained by the linear relationship) is used to quantify the degree of agreement between the laboratories, where absolute values may not agree, one laboratory may consistently measure higher than another laboratory, resulting in a high R^2 value.

Principal component analysis (PCA) was used to determine the consistency of discrimination pattern by the different laboratories through ordination of the samples within laboratories. The first two principal components were analyzed for each laboratory, using analysis of variance (ANOVA), to establish whether the samples were discriminated in a similar pattern by each lab and whether similar substrates or enzymes were identified as contributing to such discrimination. Fisher least significant difference

test (significant level of $p<0.05$) was applied to identify individual differences. All statistical tests were performed with Statistica 6.1 (STATISTICA 2004).

Results

Multi-enzyme activity assay

Mean enzyme response patterns across all soil samples were similar between the two laboratories for all enzymes although laboratory 2 obtained far higher enzyme responses for cellobiohydrolase, glucosidase, phosphatise and sulfatase in soil sample 8. The coefficient of variation for the enzyme assays (Table 2) showed a wide range from 1.9-89.1%, but there were few distinct trends in the magnitude of this variation either in relation to substrate or laboratory. Cellobiohydrolase appeared to show relatively high variation between soils within both laboratories, and relatively low variation was apparent for sulfatase for both laboratories.

Table 3 shows the response for each substrate across all soil samples in a pair-wise regression of laboratories. Linear correlation between laboratories for the estimated enzyme rates was only significant ($p<0.05$) for four of the substrates, and not for cellobiohydrolase, glucosidase, phosphatase or galactosidase (Table 3). However, the GMR analysis showed that in all cases, except for galactosidase, the regression slopes between the estimates for the two laboratories were not significantly different from one. This was supported by the PCA of the aggregated enzyme profile data for each laboratory, which showed similar variation, accounted for by the first two principal components and ordinations for both PC1 and PC2 with respect to soil samples (numbers; 4, 5 and 7), with the exception of two soil samples (2 and 8, Figure 1 a-b).

There was strong agreement between laboratories in terms of the loadings associated with the substrates for PC1 (Figure 1c-d), and notably that all substrates had similarly high loading values with respect to this PC. However, for PC2 the ordination of the loading values for the substrates was less similar between laboratories for glucosidase, cellobiohydrolase and phosphatise activities in particular (Fig 1c-d), notwithstanding that this component accounted for 4-fold less of the variance. When combined (Figure 2), the PCA illustrates that although most soil samples are discriminated, the same soil samples analysed in two laboratories do not coincide.

Multiple substrate-induced respiration (MSIR)

Mean carbon substrates response patterns across all soil samples were similar across the three laboratories although, where substrate responses were elevated e.g. soil samples 4 and 8, laboratory 2 generally demonstrated the higher substrate response values. Coefficients of variation for the MSIR assays ranged from 3.0 to 80%, with no distinct trends in relation to laboratory, substrate or soil origin (Table 4). Table 5 shows the response for each substrate across all soil samples in a pair-wise regression of laboratories. There was significant and strong linear correlation between estimated respiration rates in all pairwise combinations of laboratories and substrates (Table 5). However GMR data showed significant differences in absolute concentrations with four exceptions, two for n-acetylglucosamine, and one for γ-aminobutyric acid and one for basal (water) respiration (Table 5). In the PCA for each laboratory, there was general similarity in the percentage variation accounted for by the first two principal components for all three laboratories (Figure 3a-c). There was a high degree of similarity of the ordination of the soils within PC1 with loading values relating to the

substrates similar between all three laboratories for PC1 and accounting for > 77% of the overall variation across all laboratories. There was less concordance between labs with respect to PC2 (Figure 3 a-c), particularly in relation to Laboratory 1 compared to Laboratory 2 and 3, and most notably for arginine, which carried an opposite weighting in these two circumstances (Figure 2d-f). This is not unusual, given the nature of the PCA procedure and the experimental circumstances; with this design, most variation should be associated with differences between soils (they were prescribed on the basis of diversity), and then between the laboratories. It is notable that the greatest inter-laboratory disparity was associated with arginine. Similar to the enzyme profiles, when the MSIR profile data for laboratories are combined (Figure 4), the PCA illustrates that the individual soil samples analysed across the three laboratories do not coincide.

Discussion

Intrinsic variation

In general, the variation in results for each soil sample from each laboratory for both assays were relatively low although the coefficients of intrinsic variation were generally higher for the multi-enzyme assay than for MSIR, with the enzyme cellobiohydrolase exhibiting the highest CV across all substrates suggesting that this enzyme may be unsuitable for inclusion in multi-enzyme assays for the purposes of large-scale monitoring. Both techniques apply require only small quantities of soil for the technique and this results in large variation for each soil sampled.

The differences in CVs between the assays can partly be explained by the difference in the number of analytical replications adopted for each methodology (i.e. 3

for enzyme assay and 12 for MSIR). Typically 3 analytical replicates per substrate per soil have been applied to both these techniques (Marx et al. 2001; Lalor et al. 2007; Campbell et al. 2008). Of course, this parameter can be prescribed by the user, but will effectively be limited by some combination of the number of soil samples that are required for analysis and the resources available to achieve the measurements.

There was some evidence that particularly acidic substrates may cause a similar problem, predominantly with calcareous soils. There was a particularly large discrepancy between the relative ordination of Soil 4 and Laboratories 1 versus 2 and 3, associated with citric acid (Figure 2), and this soil gave a positive fizz test (Table 1) and contained notable calcareous granules. However, Soil 3 was also calcareous but only showed a weak response to the addition of HCl acid compared to soil 4. Due to small sample masses and the heterogeneity of the samples, the carbonate content within the different laboratory fractions may have influenced the assay for this particular soil. It was concluded that the large disparity between laboratories for arginine was due to the notably high pH (12) relative to the other substrates and a consequence of this extreme may be a particular sensitivity to assay conditions when applied under subtly different circumstances.

There may also be a confounding issue here with the mass of soil being analysed. Both techniques were scaled to be suitable for microtitre-plate systems which are necessary for high-throughput analyses and potential application in monitoring schemes. However, small masses of soil may lead to greater intrinsic variation which may also be reflected by soil texture: for example it is considerably more difficult to acquire homogeneous samples of small masses of clay and highly organic soils. This also questions whether the small mass of soil can be considered representative of the

soil sample/plot. However for both techniques the intrinsic variation was not consistently high for any particular soil type, suggesting that the response was enzyme – soil type specific.

Inter-laboratory repeatability

Here it is important to discriminate between the repeatability of absolute values, i.e. whether laboratories return similar ranges of concentrations, and repeatability of relative ranking of samples i.e. whether laboratories agree on the relative order of samples, or their intrinsic inter-relationships. Our results show diverging patterns for the enzyme assay and MSIR.

The enzyme assay gave reasonable correspondence in absolute repeatability between laboratories for all but one enzyme. The relative ranking of the soil samples was reproduced between the two laboratories; however there was limited discriminatory power between the different soil samples on PC2. In contrast, MSIR displayed poorer absolute repeatability across all three laboratories but did display a high degree of relative repeatability in all soil samples between laboratories. This means that, in general, that the three laboratories agreed upon the relative ranking of the samples. However, since absolute values were not comparable between laboratories, there was limited discrimination between the soil samples.

Therefore MSIR is a useful and consistent tool for identifying soil samples by treatment or origin, when considering analyses from single laboratories. However application within in a large-scale monitoring programme operated across multiple laboratories/users would require prior determination of inter-laboratory calibration to

translate absolute values into relative values. This could be accomplished by an inter-laboratory trial using a standard set of soil samples in a similar manner to that carried out here. The results for MSIR also indicate that great care must be taken when comparing absolute CO_2 respiration rates across the literature from MSIR approaches. This study suggests that the data from micro-titre assays may not be fully transportable between studies. This may be an issue associated with the small masses of soil involved in the assays, and the subtle but consequential interaction of the assay conditions in the plate with operators. Lalor et al. (2007) measured 6-fold greater respiration rates in a relatively larger-scale bottle-based assay system when simultaneously compared to the MicroResp™ system. More rigorous testing of the nature and origins of differences between MSIR assays at different scales and by different detection systems for respiration are warranted. This observation is of particular significance in relation to the potential utility of glucose-induced respiration as a microbial biomass measure (Anderson and Domsch 1976). Since glucose can be readily incorporated into the substrate suite of the MSIR procedure, it is in principle possible to use it to estimate microbial biomass within the context of the MSIR suite and thus provide an additional important biological soil metric.

Discrimination pattern

PCA analysis was applied to visualise the degree of similarity in the enzyme assay and MSIR profiles between the different soil samples. Analyses for individual laboratories indicated that both assays could significantly discriminate between soil samples on PC1 and the patterns of discrimination were found to be generally repeatable between labs. The MSIR technique, with a larger percentage variation on PC1 than that for the

enzyme assay for each laboratory, resulted in a much more distinct and relatively consistent pattern of separation between the different soils in the PCA ordinations. However the combined PCA served to illustrate that the lack of repeatability of absolute values between laboratories would produce different patterns of discrimination in both assays. This has significant implications, not only for soil monitoring but also more generally for inter-study comparisons.

A major limitation in repeatability of soil biological data is the availability and widespread utility of reliable analytical standards for biological methods. Our results suggest that this cannot rely on single soil samples or substrates, since soils across the range of soil chemical and physical properties may give rise to differing levels of repeatability between laboratories. The identification of suitable biological standards requires urgent attention if biological indicators of soil quality are to be adopted. Prescription of biological indicators in soil quality monitoring schemes must take due account of three potential sources of variation, *viz.* (i) intrinsic variation within a laboratory technique; (ii) inter-laboratory repeatability and (iii) discrimination between samples from different land-uses/provenances. The majority of studies to date largely focus on the discrimination between samples and do not account for the possibility that their conclusions may not be fully repeatable by different laboratories. For example, Degens and Vojvodic-Vukovic (1999) studied the impact of sample replication in the laboratory compared to the field in terms of broad scale soil biological monitoring and found that a greater difference was found between field replicates than between laboratory replicates, however this study only considered analysis of samples from a broad-scale monitoring project by 1 laboratory. We have demonstrated in this study that both multi-enzyme profiling based upon fluorimetric substrates, and MSIR profiling

based on MicroResp™ are prone to inter-laboratory variation such that where high inter-comparability or repeatability are important then dispersed analyses should be either discounted or additional efforts are made to calibrate and bring the laboratories into greater agreement. The latter approach is feasible by ensuring consistent facilities and equipment are used and preliminary trails such as the one described here are performed to identify the discrepancies and investigate the factors responsible.

Acknowledgements

The work was funded by the UK Department of Environment, Food and Rural Affairs (Defra). We thank Christian Poll and Ellen Kandeler, Universität Hohenheim, for liaison regarding the enzyme fluorimetry.

References

Anderson JPE, Domsch KH (1978) A physiological method for the quantitative measurement of microbial biomass in soils. Soil Biol Biochem 10: 215-221

Bardgett RD, Usher MB, Hopkins DW (2005) Biological Diversity and Function in Soils. Cambridge University Press, Cambridge

Bastida F, Moreno J, Hernández T, Garcia C (2006) Microbiological activity in a soil 15 years after its devegetation. Soil Biol Biochem 38:2503-2507

Bending GD, Turner MK, Rayns F, Marx M, Wood M (2004) Microbial and biochemical soil quality indicators and their potential for differentiating areas under contrasting agricultural management regimes. Soil Biol Biochem 36: 1785-1792

Black HIJ, Wood C, Parekh N, Chamberlain PM et al. (2008) Prioritising biological indicators of soil quality for deployment in a national-scale soil monitoring scheme. Report on Project SP0529. Defra, London

Breure AM, Mulder C, Rutgers M, Schouten AJ, de Zwart D (2003) A Biological Indicator for Soil Quality. OECD, Paris

British Standard BS7755 (1995); ISO 10390, 1994; Soil Quality, Part 3, Chemical Methods, section 3.2, Determination of pH

Campbell CD, Chapman SJ, Cameron CM, Davidson MS, Potts JM (2003) A rapid microtitre plate method to measure carbon dioxide evolved from carbon substrate amendments so as to determine the physiological profiles of soil microbial communities by using whole soil. Appl Environ Microbiol 69: 3593-3599

Campbell CD, Cameron CM, Bastias BA, Chen C, Cairney JWG (2008) Long term repeated burning in a wet sclerophyll forest reduces fungal and bacterial biomass and responses to carbon substrates. Soil Biol Biochem 40:2246-2252

Degens BP, Harris JA (1997) Development of a physiological approach to measuring the catabolic diversity of soil microbial communities. Soil Biol Biochem 29: 1309-1320

Degens BP, Vojvodic-Vukovic M (1999) A sampling strategy to assess the effects of land use on microbial functional diversity in soils. Aust J Soil Res 37: 593-601

Fließbach A, Oberholzer H-, Gunst L, Mäder P (2007) Soil organic matter and biological soil quality indicators after 21 years of organic and conventional farming. Agri, Ecosyst Environ 118: 273-84

Francaviglia R (Ed) (2004) Agricultural Soil Erosion and Soil Biodiversity: Developing Indicators for Policy Analyses. OECD, Paris

Hodgson JM (ed) (1997) Soil survey field handbook. Soil Survey Technical Monograph No. 5, Silsoe, UK

ISO (1997) ISO 14240-2:1997 Soil quality -- Determination of soil microbial biomass -- Part 2: Fumigation-extraction method

ISO (2002) ISO16072:2002 Soil quality -- Laboratory methods for determination of microbial soil respiration

ISO (2005) ISO 23753-1:2005 Soil quality -- Determination of dehydrogenase activity in soils -- Part 1: Method using triphenyltetrazolium chloride (TTC)

ISO (2006) ISO 23611-1:2006 Soil quality -- Sampling of soil invertebrates -- Part 1: Hand-sorting and formalin extraction of earthworms

Jenkinson DS, Powlson DS (1976) The effects of biocidal treatments on metabolism in soil-V. A method for measuring soil biomass. Soil Biol Biochem 8: 209-213

Kandeler E, Mosier AR, Morgan JA, Milchunas DG, King JY, Rudolph S, Tscherko, D (2006) Response of soil microbial biomass and enzyme activities to the transient elevation of carbon dioxide in a semi-arid grassland. Soil Biol Biochem 38: 2448-2460

Kubat J (2003) Incidence and activities of the soil micro-organisms as an indicator of soil biodiversity. OECD, Paris

Lalor BM, Cookson WR, Murphy DV (2007) Comparison of two methods that assess soil community level physiological profiles in a forest ecosystem. Soil Biol Biochem 39: 454-462

Marx M, Wood M, Jarvis SC (2001) A microplate fluorimetric assay for the study of enzyme diversity in soils. Soil Biol Biochem 33: 1633-1640

Molecular Devices, UK

Oliveira A, Pampulha ME (2006) Effects of long-term heavy metal contamination on soil microbial characteristics. J. of Biosci 102:157-161

Ros M, Klammer S, Knapp B, Aichberger K, Insam H (2006) Long-term effects of compost amendment of soil on functional and structural diversity and microbial activity. Soil Use Manag 22: 209-218

Sokal RR, Rohlf FJ (1981) Biometry. W.H. Freeman, New York

STATISTICA (2004) STATISTICA v6.1. StatSoft Inc, 1984-2004

Wakelin SA, Macdonald LM, Rogers SL et al. (2008) Habitat selective factors influencing the structural composition and functional capacity of microbial communities in agricultural soils. Soil Biol Biochem 40: 803-813

Winding A, Hendriksen NB (2007) Comparison of CLPP and enzyme activity assay for functional characterization of bacterial soil communities. J Soils Sediments 7: 411-417

Figure legends

Figure 1. Principal components analysis of multiple enzyme activity assay with 8 substrates tested on 5 soil samples. (left column) and associated loadings (right column). (a, b) Laboratory 2; (c, d) Laboratory 3. (a, c) First (PC1) and second (PC2) component plots; Points show means, numbers within them refer to soils; bars show s.e. (n=3). Note that in some instances error bars fall within confines of points. Percentage variation accounted for by PCs shown in parentheses on axes labels. (b, d) Loadings associated with adjacent PC plots. CELL – Cellobiohydrolase; GLUC – Glucosaminidase; GLUS – Glucosidase; PHOS – Phosphatase; GALA – Galactosaminidase; XYLO – Xylosidase; GALS – Galactosidase; SULF – Sulphatase.

Figure 2. Principal component analysis of multiple enzyme activity assay with 8 substrates tested on 5 soil samples for Laboratories 2 and 3 ordinated together. Points show means, numbers within them refer to soils; bars show s.e. (n=3). Note that in some instances error bars fall within confines of points.

Figure 3. Principal components analysis of MSIR CLPPs with 8 carbon sources tested on 9 soil samples. (a, b) Laboratory 1; (c, d) Laboratory 2; (e, f) Laboratory 3. (a, c, e) First (PC1) and second (PC2) component plots; numbers refer to soils. Points show means, numbers within them refer to soils; bars show s.e. (n=9). Note that in some instances error bars fall within confines of points. Percentage variation accounted for by PCs shown in parentheses on axes labels. (b, d, f) Loadings associated with adjacent PC

plots. GABA = γ-aminobutyric acid; NAGA = n-acetyl glucosamine; AKGA = α-ketoglutaric acid.

Figure 4. Principal components analysis (PCA) for MSIR CLPPs with 8 carbon sources tested on 9 soil samples, for all laboratories ordinated together. Points show means, numbers within them refer to soils; bars show s.e. (n=3). Note that in some instances error bars fall within confines of points.

Table 1. Summary of origins, land-use and principal characteristics for soils used in this study.

Grid Ref[a]	Land-use	Sample code	Texture	pH_{H20}	SOM (%)	WHC (%)	Carbonate present[b]
TL082356	Arable	1	Clay loam	7.1	6.4	43.6	N
SD498399	Arable	2	Sandy silt loam	6.5	5.8	40.2	N
NJ183626	Arable	3	Loamy sand	6.6	3.1	31.4	Y
TL335510	Grassland	4	Sandy silt loam	7.5	13.0	49.3	Y
SD349457	Grassland	5	Sandy silt loam	6.1	9.3	44.4	N
NO665785	Grassland	6	Silt loam	6.9	15.2	44.2	N
TL082356	Woodland	7	Sandy loam	6.7	2.2	35.6	N
SD435795	Woodland	8	Silt loam	6.9	25.3	59.6	N
NO652802	Woodland	9	Organic	4.5	44.4	49.9	N

[a]UK OS National Grid [b]HCl effervescence test; N is no; Y is yes.

Table 2. Precision within the multiple-enzyme activity assay, as denoted by mean values (nmol g-1 soil h-1) (coefficient of variation (%) in brackets) for within-plate fluorimetric measurements, in relation to laboratory, soil and substrate.

Laboratory	Soil	Enzyme							
		Cellobiohydrolase	Glucosaminidase	Glucosidase	Phosphatase	Galactosaminidase	Xylosidase	Galactosidase	Sulphatase
2	2	7 (87.3)	38 (53.6)	100 (14.7)	156 (8.8)	16 (36.0)	9 (33.8)	4 (41.2)	28 (4.6)
	4	455 (14.6)	881 (8.7)	1781 (6.0)	1883 (4.0)	411 (8.5)	248 (6.1)	109 (9.6)	353 (7.3)
	5	251 (89.1)	349 (23.3)	623 (8.3)	636 (26.2)	73 (20.4)	303 (18.8)	98 (14.1)	869 (15.7)
	7	140 (86.4)	275 (51.6)	720 (11.7)	374 (14.3)	85 (42.9)	50 (69.3)	27 (30.0)	88 (2.4)
	8	191 (3.5)	310 (26.6)	1858 (23.8)	2508 (14.4)	177 (46.8)	163 (4.6)	100 (32.5)	1152 (8.3)
	Mean	56.2	32.8	12.9	13.5	30.9	26.5	25.5	7.7
3	2	157 (75.6)	205 (8.5)	629 (39.6)	605 (7.9)	181 (1.9)	177 (3.0)	175 (3.2)	240 (17.0)
	4	271 (35.4)	720 (19.4)	1767 (33.4)	1856 (18.5)	444 (22.6)	375 (26.8)	231 (8.2)	324 (23.5)
	5	361 (26.0)	514 (8.3)	1216 (17.2)	1222 (15.7)	335 (10.8)	521 (6.8)	393 (10.7)	993 (9.6)
	7	175 (16.0)	425 (18.9)	898 (17.9)	657 (19.3)	282 (10.3)	186 (14.2)	195 (23.9)	142 (15.4)
	8	97 (25.4)	227 (22.2)	726 (34.2)	1085 (20.1)	257 (9.4)	206 (38.6)	243 (37.4)	534 (7.2)
	Mean	35.7	15.4	28.5	16.3	11.0	17.9	16.7	14.5

Table 3. Correlation coefficient (R^2) and geometric mean regression (GMR) of enzyme activity by Laboratories 2 and 3

Enzyme	R^2	P[a]	GMR	P[a]
Cellobiohydrolase	0.4152	ns	1.59 ± 0.61	Ns
Glucosaminidase	0.8154	*	1.49 ± 0.32	Ns
Glucosidase	0.3589	ns	1.59 ± 0.64	Ns
Phosphatase	0.4898	ns	1.80 ± 0.64	Ns
Galactosaminidase	0.7317	*	1.46 ± 0.38	Ns
Xylosidase	0.8631	**	0.79 ± 0.15	Ns
Galactosidase	0.5497	ns	0.49 ± 0.1	*
Sulfatase	0.6303	*	1.43 ± 0.43	Ns

[a] Ns = $p>0.05$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

Table 4. Precision within the MicroResp assay, as denoted by mean sample values (µg/g/h CO2-C) (coefficient of variation (%) in brackets) for within-plate respiration measurements, in relation to laboratory, soil and substrate.

Laboratory	Soil	Substrate							
		Water	AKGA	Arginine	Citric acid	GABA	Glucose	Malic acid	NAGA
1	1	0.4 (29.3)	2.0 (16.7)	0.7 (11.3)	1.1 (44.6)	0.5 (25.9)	1.2 (23.4)	1.3 (8.4)	0.8 (19.4)
	2	0.6 (20.8)	7.6 (5.2)	1.8 (27.4)	5.5 (14.5)	1.1 (15.5)	2.9 (6.0)	2.7 (50.5)	1.7 (8.5)
	3	0.4 (5.4)	2.4 (6.4)	1.5 (14.2)	1.1 (6.4)	0.8 (10.2)	1.6 (8.5)	1.6 (15.5)	1.1 (16.5)
	4	5.8 (33.0)	19.5 (5.6)	4.9 (50.4)	21.3 (5.1)	5.9 (15.8)	10.1 (11.6)	21.4 (8.5)	7.2 (16.8)
	5	1.1 (41.9)	3.5 (15.3)	2.9 (20.2)	2.9 (14.9)	1.8 (8.3)	3.0 (7.1)	3.0 (14.3)	1.7 (10.3)
	6	1.3 (7.6)	4.5 (15.5)	3.7 (9.8)	3.4 (22.8)	2.7 (4.2)	5.3 (5.0)	4.9 (7.1)	4.0 (3.0)
	7	0.8 (13.3)	4.5 (11.5)	1.2 (36.2)	3.7 (12.4)	1.2 (10.0)	2.9 (5.1)	3.7 (27.2)	2.0 (9.1)
	8	3.0 (21.6)	21.2 (8.3)	7.2 (22.3)	16.1 (4.3)	5.9 (18.1)	11.7 (3.8)	13.8 (8.1)	6.5 (4.6)
	9	1.8 (19.2)	5.0 (27.6)	5.1 (41.4)	4.2 (9.1)	3.5 (10.9)	6.5 (11.8)	7.7 (31.1)	3.5 (7.6)
	Mean	1.7	7.8	3.2	6.6	2.6	5.0	6.7	3.2
2	1	0.9 (14.1)	5.9 (20.5)	1.5 (24.8)	2.7 (18.6)	1.3 (17.6)	2.8 (16.0)	2.7 (10.5)	1.6 (11.5)
	2	1.5 (8.6)	10.1 (20.4)	3.0 (15.3)	7.0 (18.7)	2.1 (9.1)	4.3 (23.7)	5.1 (9.0)	2.8 (14.3)
	3	1.1 (8.8)	4.8 (9.7)	3.2 (11.0)	2.6 (6.7)	1.7 (11.5)	3.2 (7.4)	2.6 (18.8)	2.4 (8.0)
	4	12.6 (26.2)	50.7 (37.8)	7.0 (44.0)	71.1 (15.3)	11.8 (14.0)	20.4 (12.1)	61.7 (31.0)	14.6 (30.1)

	5	1.9 (7.0)	6.2 (6.6)	5.0 (17.4)	5.2 (9.0)	2.9 (8.0)	4.2 (10.3)	4.7 (7.3)	2.5 (9.0)
	6	2.6 (11.2)	8.2 (12.4)	10.5 (8.9)	6.0 (6.8)	4.5 (9.4)	9.3 (13.0)	8.1 (10.3)	6.1 (8.9)
	7	1.6 (8.8)	8.0 (16.0)	1.9 (23.8)	5.8 (16.3)	2.0 (9.7)	5.1 (14.6)	7.8 (13.1)	3.4 (10.3)
	8	6.7 (12.3)	89.7 (59.3)	16.8 (29.0)	38.0 (36.5)	9.4 (14.0)	22.4 (20.0)	34.5 (13.5)	11.9 (17.6)
	9	4.0 (6.5)	9.0 (9.9)	18.3 (17.9)	8.2 (14.0)	6.1 (14.3)	9.5 (7.8)	12.7 (9.3)	5.5 (11.0)
	Mean	3.7	21.4	7.5	16.3	4.7	9.0	15.6	5.6
3	1	0.9 (7.2)	3.6 (34.6)	1.8 (15.9)	2.2 (7.6)	1.1 (6.4)	2.1 (7.1)	2.0 (22.2)	1.5 (8.4)
	2	1.8 (5.8)	9.6 (19.8)	5.4 (8.8)	7.5 (29.7)	2.6 (22.2)	4.9 (8.8)	5.9 (7.8)	3.3 (13.4)
	3	1.3 (22.7)	4.5 (4.1)	3.3 (8.8)	2.2 (8.9)	1.8 (9.9)	3.2 (9.4)	2.8 (19.2)	2.5 (6.8)
	4	12.3 (15.4)	24.6 (7.6)	10.7 (25.4)	25.9 (5.6)	10.4 (14.1)	15.6 (11.2)	26.4 (5.6)	12.8 (20.7)
	5	2.2 (15.4)	5.6 (11.3)	5.0 (15.4)	3.7 (29.8)	2.9 (19.9)	4.2 (14.9)	5.1 (10.0)	2.9 (10.2)
	6	3.0 (15.5)	6.9 (35.4)	7.6 (36.4)	5.9 (16.0)	5.0 (10.2)	9.1 (6.3)	8.9 (23.6)	11.1 (6.0)
	7	2.0 (14.0)	5.9 (12.1)	3.0 (27.0)	5.2 (17.6)	2.5 (5.0)	4.3 (4.6)	6.3 (7.4)	3.3 (5.9)
	8	7.7 (17.8)	28.4 (11.0)	14.3 (7.7)	18.6 (36.3)	9.1 (17.3)	17.1 (6.9)	21.1 (4.5)	10.8 (11.6)
	9	7.7 (77.5)	10.9 (73.6)	12.3 (47.8)	8.0 (6.3)	7.2 (18.2)	12.5 (44.8)	14.6 (79.9)	7.3 (14.3)
	Mean	4.3	11.1	7.1	8.8	4.7	8.1	10.3	5.7

Table 5. Correlation coefficient (R^2) and geometric mean regression (GMR) of respiration rates for each substrate in MSIR assay, via pairwise comparison between laboratories.

Substrate	Laboratory comparisons	R^2	P[a]	GMR	P[a]
H_2O	1 vs 2	0.9903	***	1.20 ± 0.09	NS
	1 vs 3	0.8501	***	1.70 ± 0.52	***
	2 vs3	0.8654	***	1.42 ± 0.21	**
Arginine	1 vs 2	0.6728	**	3.78 ± 0.88	***
	1 vs 3	0.9296	***	2.67 ± 0.27	***
	2 vs3	0.9299	***	0.70 ± 0.08	*
Malic acid	1 vs 2	0.9596	***	1.66 ± 0.21	***
	1 vs 3	0.9773	***	1.25 ± 0.07	*
	2 vs3	0.9470	***	0.75 ± 0.05	*
γ-Aminobutyric acid	1 vs 2	0.9757	***	1.79 ± 0.11	***
	1 vs 3	0.9741	***	1.62 ± 0.10	***
	2 vs3	0.9725	***	0.91 ± 0.06	NS
n-Acetyl glucosamine	1 vs 2	0.9769	***	1.97 ± 0.11	***
	1 vs 3	0.9862	***	0.87 ± 0.08	NS
	2 vs3	0.9587	***	0.87 ± 0.07	NS
Glucose	1 vs 2	0.9777	***	1.99 ± 0.11	***
	1 vs 3	0.9857	***	1.49 ± 0.07	**
	2 vs3	0.9515	***	0.75 ± 0.06	*
A-Ketoglutaric acid	1 vs 2	0.8549	***	0.28 ± 0.17	***
	1 vs 3	0.9918	***	1.29 ± 0.04	*
	2 vs3	0.8303	***	4.54 ± 0.22	***
Citric acid	1 vs 2	0.8392	***	0.39 ± 0.26	***
	1 vs 3	0.9886	***	1.25 ± 0.05	*
	2 vs3	0.9893	***	3.24 ± 0.07	***

[a] NS = $p>0.05$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$

5