Community-Level Physiological Profiling

Chapter *in* Methods in molecular biology (Clifton, N.J.) · January 2010

Impact Factor: 1.29 · DOI: 10.1007/978-1-60761-439-5_16 · Source: PubMed

2 authors:

Kela P Weber
Royal Military College of Canada
87 PUBLICATIONS **371** CITATIONS
SEE PROFILE

Raymond Louis Legge
University of Waterloo
159 PUBLICATIONS **3,852** CITATIONS
SEE PROFILE

Chapter 16

Community-Level Physiological Profiling

Kela P. Weber and Raymond L. Legge

Abstract

Community-level physiological profiling (CLPP) is a technique which offers an easily applied protocol yielding information regarding mixed microbial community function and functional adaptations over space and time. Different communities can be compared and classified based on sole carbon source utilization patterns (CSUPs) gathered using BIOLOG™ microplates. One of the most challenging aspects associated with the CLPP method is in the data analysis. This chapter describes the relatively simple CLPP laboratory protocol and provides a detailed description of different data analysis techniques.

Key words: Community-level physiological profiling (CLPP), BIOLOG™, carbon source utilization pattern (CSUP), microbial community, microbial ecology, multivariate analysis, principle component analysis (PCA).

1. Introduction

The term community-level physiological profiling (CLPP) was first coined by Lehman et al. (1) to describe the characterization and classification of heterotrophic microbial communities based on sole carbon source utilization patterns (CSUPs). Although CLPP is considered a broad term which could cover many different types of studies undertaken using a number of different assays, currently the term CLPP is used almost exclusively in reference to data collected using BIOLOG™ microplates. BIOLOG™ microplates are 96-well plates where each well contains a different carbon source and a redox dye indicator, most often tetrazolium violet. When a mixed microbial community sample is inoculated into each of the wells, the production of NADH via cell respiration reduces the tetrazolium dye to formazan, resulting in a

S.P. Cummings (ed.), *Bioremediation*, Methods in Molecular Biology 599,
DOI 10.1007/978-1-60761-439-5_16,

colour change within each individual well, which can be detected photometrically.

There are a number of different microplates manufactured by BIOLOG™ for CLPP use, with the three most popular being the GN2, GP2 and EcoPlates. The GN2 plate is the most recent version of the GN plate and is suitable for characterizing or identifying Gram-negative bacteria. The GP2 plate is the most recent version of the GP plate, and is suited to characterizing or identifying Gram-positive bacteria. The GN2 and GP2 plates both contain 95 different carbon sources with one of the 96 wells serving as a blank. Both the GN2 and the GP2 plates were originally developed for species identification (2), but are now commonly used for CLPP. The BIOLOG™ EcoPlate contains 31 different carbon sources and a blank in triplicate. Use of triplicates allows for increased confidence in statistical analysis of the resulting plate data. The EcoPlate was developed for environmental applications, which dictated the selection of carbon sources, with at least nine substrates considered constituents of plant root exudates (3,4).

BIOLOG™ offers a number of other plates suitable for CLPP studies. SF-N and SF-P microplates are alternatives to the GN2 and GP2 plates, as they provide the same corresponding substrates, but without the tetrazolium dye. Turbidity or a different metabolic indicator can be added to assess activity. MT microplates contain the same redox chemicals as the GN2 and GP2 plates, but do not contain any substrates. These plates allow for the creation of customized plates by adding suitable substrates for specific ecological studies. FF plates, which have been recently introduced by BIOLOG™ for the study of fungi and yeasts, contain a unique set of carbon sources and use both turbidity and/or reduction of tetrazolium as activity indicators. The GN plate and its corresponding successor, GN2, have been favoured in CLPP studies although other plates may offer greater relevance and analytical options (4).

Garland and Mills (5) were the first to use BIOLOG™ plates for characterizing heterotrophic soil bacterial communities, and a number of studies have since followed (*see* (4,6) for examples). The advantage of CLPP over both classic cell culturing and molecular level RNA/DNA amplification-based techniques is its relatively simple protocol and ease of use. Both classic cell culturing and molecular level RNA/DNA amplification-based techniques can be time consuming and require specialized expertise (7).

Limitations pertaining to the CLPP approach using BIOLOG™ microplates have been discussed in the literature (4,6,7). Limitations and pitfalls pertaining to data analysis have also been recently described (8). Some of the most pertinent limitations include the bias in the technique towards rapidly growing bacteria, the need to ensure similar inoculum sample sizes in the

wells, the need to reduce time between sampling and inoculation of the microplates and the difficulties with meaningful data analysis and interpretation.

Some of the limitations surrounding the CLPP method pertain to the long incubation times, the indirect measurement of microbial activity, and the use of high substrate concentrations. The use of lower substrate concentrations would allow for less selective enrichment within the wells as growth and incubation conditions would be more akin to those from which the sample originates. To allow for the use of decreased substrate concentrations, more sensitive and direct activity measurements would then be required. Newly developed CLPP methods include those of Degens et al. (9) and Garland et al. (10). Degens et al. (9) developed a method for the detection of CO_2 generation from mixed microbial communities utilizing a range of carbon substrates. Although Degens et al. (9) refer to their measurements as microbial catabolic diversity, the basis behind the study is similar to that of CLPP. Garland et al. (10) developed a fluorescent-based method of measuring O_2 consumption for mixed microbial communities utilizing a range of carbon substrates. Using this method, Garland et al. (10) found that incubation times could be reduced to less than 24 h, and the substrate concentrations could be reduced by a factor of 10–100 when compared to BIOLOG™ plates. Currently, the term CLPP almost exclusively refers to the use of BIOLOG™ microplates. Studies such as those conducted by Garland et al. (10) and Degens et al. (9) point to the expanding usage of the term CLPP. Improvements to the CLPP method both with and without the use of BIOLOG™ plates will no doubt lead to the evolution and changing and/or broadening of the term CLPP and its associated methods in the future.

CLPP is a technique which offers an easily applied protocol yielding large amounts of information regarding mixed microbial community function and functional adaptations over space and time. Carrying out the CLPP laboratory protocol is a relatively simple process, which has lead to its recent increase in popularity. However, the data analysis aspect associated with CLPP can be challenging, often requiring a background in multivariate analysis methods. Following is a description of the CLPP laboratory protocol and an in-depth description of the data analysis procedure.

2. Materials

1. Suspended mixed microbial community sample
2. BIOLOG™ microplate(s)

3. Microplate reader equipped with a 590 nm filter

Optional:

1. Buffer solution – May be needed if (i) performing a detachment protocol from a sediment sample to generate a suspended microbial community sample or (ii) performing serial dilutions before plate inoculation (*see* **Section 3.2**).
2. Incubator – May be needed if incubating samples at a temperature other than room temperature (*see* **Section 3.3**).

3. Methods

3.1. CLPP Protocol – General Description

Each well of the BIOLOG™ plate is inoculated with 150 μL of the sample of interest and incubated at temperatures generally ranging from 20 to 30°C. Absorbance readings (590– 600 nm) are performed as necessary using a microplate reader over an incubation period ranging from 10 to 200 h. The sample should be a uniform suspension, so if sampling sediment or biofilm, an appropriate detachment and/or homogenization protocol is necessary.

3.2. Inoculation

About 150 μL of a suspended mixed microbial community sample is inoculated into each of the 96 wells of the BIOLOG™ microplate. An undiluted sample is recommended as dilution of samples containing a mixed population has been shown to affect the resulting CLPPs (11,12). Analytical methods for dealing with small differences in inoculation densities are discussed later; however, it is important to ensure similar cell densities of the samples to reduce any error in the CLPP analysis (*see* **Note 1**). Although not recommended in all cases, if cell densities in the original samples are exceedingly high, serial dilutions may be needed before microplate inoculation. If the plates are inoculated at high cell densities, colour development may proceed at a rate where capturing meaningful data is difficult. It has been suggested that formazan production does not occur until cell densities between 10^5 and 10^8 cells/mL are reached (6,12). Lastly, the time between sampling and inoculation should be kept to a minimum in order to reduce cell death or community structural changes of the sample prior to plate inoculation.

3.3. Incubation

Plates can be incubated over a range of temperatures with room temperature being the most common. Incubation periods tend to range between 10 and 200 h; and standard incubation temperatures between 20 and 30°C. Incubation temperatures similar to those from which the sample was collected are ideal. There is some debate surrounding the effect that incubation tempera-

ture has on the resulting CLPPs. Christian and Lind (13) showed that temperature had an effect on the calculated CSUPs, whereas Classen et al. (14) showed CLPPs to be relatively insensitive to incubation temperature. Room temperature incubation has not been criticized in the literature, as the resulting CLPPs have proven useful and reasonable for most published studies.

The plates do not need to be agitated during incubation due to the relatively long incubation times. Stationary incubation at room temperature is the most common method, although shaking is required prior to plate reading to ensure uniform distribution of the formazan. Incubation periods will vary for different studies, but generally range from 10 to 200 h depending on the study and the inoculation density. Choice of a specific incubation time is not obvious, and is largely dependent on the subsequent data analysis (*see* **Section 4**). **Figure16.1** is a picture of a BIOLOG^{TM} EcoPlate after a 20 h incubation period following inoculation with interstitial water from a wetland mesocosm (15).

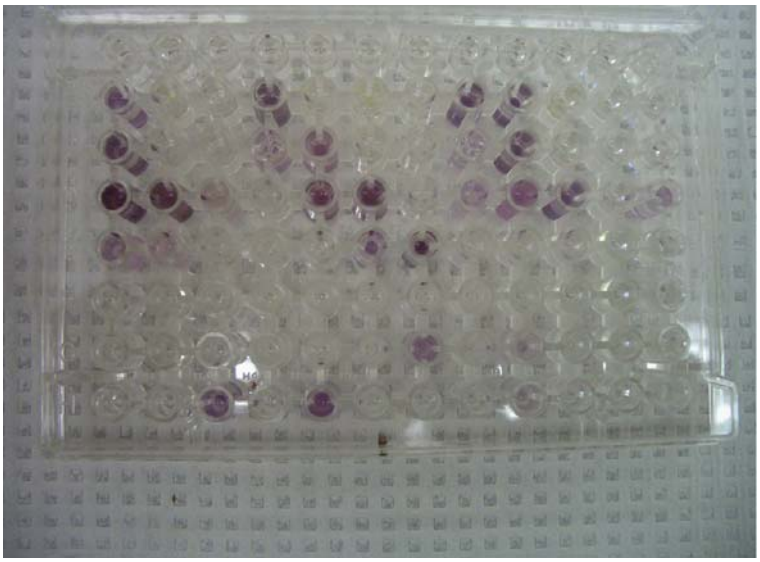

Fig. 16.1. BIOLOG^{TM} EcoPlate 20 h after inoculation with interstitial water from a wetland mesocosm system. Plate used in the study of Weber et al. (15).

3.4. Data Collection

Absorbance readings (590 nm) for all 96 wells are collected throughout the incubation period to reveal the kinetic profiles for each of the carbon sources. Reading frequency will vary based on the nature of the inoculum and the type of metric chosen for analysis (*see* **Note 2**); if data analysis does not require the fitting of kinetic profiles and a single time point is used for analysis (*see* **Section 4**), less frequent readings are reasonable. Plates should

be agitated before each reading to ensure sufficient colour distribution in each well.

An example of the type of data that can be collected when following the described protocol is provided in **Fig.16.2**. Colour development for individual wells for a single set of EcoPlate replicates is presented (31 different carbon sources and 1 blank); colour development curves show a general sigmoidal shape.

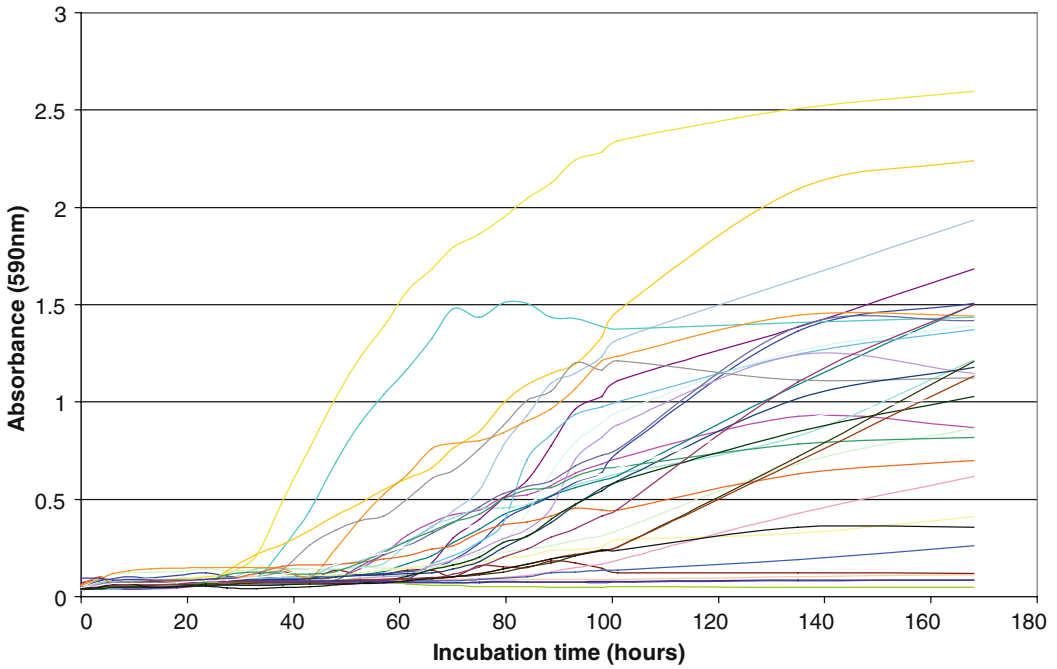

Fig. 16.2. Individual well colour development curves for a single set of EcoPlate replicates (31 carbon sources and 1 blank). Data from Weber et al. (15).

3.5. Anaerobic Protocol

A small number of earlier studies indicated that formazan is not produced in regular aerobic-usage BIOLOG™ plates incubated under anaerobic conditions (4,16). However, a number of recent studies have shown formazan production does occur under anaerobic conditions (13,17,18). BIOLOG™ manufactures an AN microplate for the identification of anaerobic bacteria although a number of anaerobic CLPP studies have favoured the use of GN or ECO microplates as these plates have been previously shown to be suitable for ecological and exploratory CLPP studies for mixed microbial systems. The anaerobic CLPP protocol is identical to the aerobic protocol described here, with a few changes focusing on minimizing the exposure to oxygen throughout the procedure (*see* **Note 3**). This can be accomplished through any number of standard anaerobic culturing and testing techniques.

3.6. Data Analysis

A large amount of data can be collected with a single BIOLOG™ plate; when numerous plates are used, the amount of data can then become overwhelming. To deal with the large amount of data and large number of variables to be processed, some form of multivariate analysis is required. Data analysis will be largely discussed with direction towards the use of principle component analysis (PCA) as this is the most commonly and easily used method for analysing and visualizing the CLPP data. A number of other techniques will also be discussed later.

3.7. Standard Analysis Method When Using Principle Component Analysis

Before reaching the point of performing a multivariate analysis technique (such as PCA), a number of steps should be followed:

(1) Decide what metric will be used for the data analysis

(2) Standardization of the data

(3) Assess heterogeneity, normality and the underlying factor structure of the data

(4) Perform a data transformation if required

3.8. Selecting a Metric

First a metric needs to be selected and extracted from the BIOLOG plate data to represent activity in each well. The three most commonly utilized metrics are:

(i) An absorbance value for each well for a specific incubation time point

(ii) An absorbance value for each well taken from a time point representing a specific average well colour development (AWCD) for that plate

(iii) Some type of logistic curve fitting value such as lag, maximum utilization rate (slope), area under the curve or an asymptote value

3.8.1. Selecting a Specific Incubation Time Point

To evaluate all plate data within a study, a specific incubation time point can be chosen as a metric, but choosing this time point may not be obvious. An increase in the differences (or variation) between well absorbance values indicates an increase in the amount of information contained within the data set. Using absorbance values taken early in the incubation time would yield little information, for at early stages of growth (for example, 10 h in **Fig.16.2**) the difference between well absorbance values is too small to yield useful information. Using absorbance values taken later during the incubation can provide more information regarding the CLPP of the microbial inoculum as long as the values are not above a value of 2. As seen in **Fig.16.2**, there is an increase in the dispersion of (or differences between) well absorbance values as the incubation proceeds. This dispersion of well absorbance values can be represented as the standard deviation calculated at each time point (**Table16.1**).

Table 16.1
Calculated AWCD, number of values above an absorbance of 2, and standard deviations for absorbance values over the 168 h incubation time for the plate shown in Fig. 16.2. Data from Weber et al. (15)

Incubation time (hours)	AWCD	# values above absorbance of 2	Standard Deviation
0	0.00	0	0.01
4	0.01	0	0.02
6	0.01	0	0.02
9	0.01	0	0.02
13	0.01	0	0.02
18	0.00	0	0.02
24	0.00	0	0.02
28	0.01	0	0.02
34	0.02	0	0.04
38	0.04	0	0.09
43	0.06	0	0.14
48	0.09	0	0.20
53	0.12	0	0.25
57	0.16	0	0.29
62	0.19	0	0.33
66	0.24	0	0.37
70	0.27	0	0.40
75	0.31	0	0.40
80	0.37	0	0.43
84	0.42	0	0.45
89	0.47	1	0.46
93	0.53	1	0.49
98	0.57	1	0.50
101	0.60	1	0.52
135	0.85	2	0.59
168	0.99	2	0.63

Absorbance readings above 2 contribute to measurement error as they are outside the linear absorbance range. An appropriate time point will be the time point that preserves the greatest variance between well responses while retaining the maximum number of wells within the linear absorbance range. For example for the data in **Table16.1**, Weber et al. (15) chose to use absorbance data from the 84 h time point for subsequent mul-

tivariate analysis (*see* **Note 4**). This study included a number of BIOLOG™ plates comparing a number of different microbial samples. The data from all plates were considered before a time point was chosen. When using this simplified method of choosing a specific time point for all analyses, similar inoculation densities for all plates is essential (12).

3.8.2. Selecting a Time Point Related to a Specific AWCD Reference Value

As recommended by Garland et al. (12), a specific AWCD value can also be chosen as a reference point for all plates analysed. Absorbance values used for subsequent multivariate analysis are extracted from each set of plate data at the specific time point where the reference AWCD occurs. Garland (7) showed that using AWCD reference values between 0.25 and 1.0 yields relatively similar CLPPs for use in community classification (*see* **Note 5**).

3.8.3. Kinetic Analysis

Kinetic analysis can also be performed on the well colour development data (for examples, *see 4, 12, 19, 20*). A number of different metrics can be chosen for use in multivariate analysis. Factors such as lag time, maximum utilization rate (slope), area under the curve or asymptote values have been used. A large amount of data needs to be acquired for a logistic curve to fit the data, and not all data are suited to a logistic fit. Deviation in absorbance readings and non-characteristic responses in some wells can have a large effect on curve fitting, making some data unfit for kinetic analysis.

Kinetic approaches have great potential, as a more detailed understanding of the nature of the color responses can be theoretically attained (4,12). However, a general lack of understanding regarding physiological or ecological bases for differences in the derived kinetic parameters limits the amount of information that can be extracted when using a kinetic approach (12). If the objective is to classify different microbial populations, using data from a single absorbance point reading may be more useful. Garland et al. (12) found that using a single absorbance point reading corrected by the AWCD was more successful than using kinetic parameters for classifying different soil bacteria populations, and remarked that the use of kinetic parameters for CLPP may provide some additional information, but only if the influence of inoculum density is carefully considered (*see* **Note 6**).

The overall incubation time for any one study is dependent upon what type of metric will be used and cannot be easily determined. It is preferable to run a number of test plates inoculated with microbial community samples similar to those to be characterized in the overall study before starting an experimental regime. Using these test plates, careful determination of which type of metric will be used and the resulting incubation time can be assessed.

3.9. Standardization of the Data

When performing a CLPP analysis, numerous plates are often used to study different mixed microbial communities in space or over time. As recommended by Garland (7), if the choice of metric is to use a single time point absorbance, and not perform a kinetic analysis of the data, an initial standardization of the data helps to reduce any bias due to inoculum density differences between samples. Standardization of the data involves correcting each absorbance value by its corresponding blank and then dividing by the AWCD for that time point. The standardized absorbance for well k can be calculated as:

$$\bar{A}_k = \frac{A_k - A_o}{\frac{1}{31}\sum_{i=1}^{31}(A_i - A_o)} \quad [1]$$

where A_i represents the absorbance reading of well i and A_0 is the absorbance reading of the blank well (inoculated, but without a carbon source). Where there is very little response in a well, negative values of standardized absorbance may occur and, since this is physically meaningless, they are coded as zeros for further analysis. Standardization of the data may not be needed when performing certain types of kinetic analyses.

3.10. Assess the Suitability of the Data Set for Multivariate Analysis

Efficient and meaningful statistical methods for dealing with a large number of interdependently correlated variables are needed when evaluating CSUPs from BIOLOG™ plates; most researchers have turned to some form of multivariate technique and, in particular, principal component analysis (PCA) (21). In performing PCA, each plate (p number of plates) is considered an object with n variables (31 for EcoPlates, 95 for GN2 or GP2 plates) giving a matrix with p rows and n columns. The transformation of BIOLOG™ plate data is an important aspect of multivariate analysis techniques such as PCA. Weber et al. (8) provide an in-depth study on data preparation techniques briefly described in the following section.

Many multivariate analysis techniques assume two fundamental properties of a data set: normality and homoscedasticity (that is, homogeneity of variance: all variables are assumed to have the same variance). In PCA, the dimensionality of the data set is reduced by extracting an orthogonal set of principal components (PCs) made up of linear subsets of the original ordinates; the extraction is designed so that the maximum amount of variance is concentrated in the first PC, with the second largest amount of variance contained in the second PC, and so on. This analysis technique is most powerful if the data have an underlying factor structure; that is, it is dependent on linear correlations between the different variables (22). Weber et al. (8) concluded that if

homoscedasticity, normality and the number of linear correlations within a data set are not evaluated and the possibility of transforming the data is not considered, erroneous analysis and misleading conclusions may arise when performing multivariate analysis on microplate data (*see* **Note 7**). Following is a short summary of useful data transformation techniques and data assessment methods used when working with BIOLOG™ microplate data.

3.10.1. Normality

Normality of BIOLOG™ microplate data can be evaluated through formal statistical tests. The kurtosis and skewness of each variable can be calculated and the standard errors found according to Equations (**15.2**) and (**15.3**) for kurtosis and skewness, respectively:

$$SE_{kurtosis} = \sqrt{\frac{24}{n}} \quad [2]$$

$$SE_{skewness} = \sqrt{\frac{6}{n}} \quad [3]$$

where $SE_{kurtosis}$ and $SE_{skewness}$ are the standard errors for kurtosis and skewness, respectively, and n is the number of observations. The corresponding z values can also be calculated as:

$$z_{kurtosis} = \frac{kurtosis}{SE_{kurtosis}} \quad [4]$$

$$z_{skewness} = \frac{skewness}{SE_{skewness}} \quad [5]$$

Either z value may be used in a formal statistical test of a null hypothesis that the data is normally distributed versus an alternative that it is not. A two-tailed test is used and the null hypothesis rejected with 95% confidence if $|z| > 1.96$.

Normality according to Weber et al. (8) can be assessed using both kurtosis and skewness, by calculating the mean value of the statistics across all variables as well as testing the individual variables. Note that the 95% significance level applies to the individual tests and not to the global set of tests of all variables; therefore, the number of significant results is considered to be indicative of the number of significantly non-normal variables tested. For a more detailed example see Weber et al. (8).

3.10.2. Homoscedasticity

Homoscedasticity is perhaps best assessed by a scatter plot of pairs of variables (a characteristic oval appearance will result for homoscedastic pairs); however, this is not feasible for so many variables, and instead a variance ratio can be calculated:

$$\text{variance ratio} = \frac{\text{highest variance}}{\text{lowest variance}} \quad [6]$$

This is adapted from the concept that a lesser degree of variation in separate variances contributed by many variables will constitute a lower ratio between the highest variance of any one variable and the lowest variance of any one variable in that data set (23). This cannot be tested formally and should only be considered indicative of the relative homogeneity of variance between data sets (lower values being relatively more homogeneous).

3.10.3. Underlying Factor Structure – Linear Correlations

The number of linearly correlated variables within a data set can be calculated by obtaining the correlation matrix and counting the number of correlation coefficients greater than Pearson's critical r value for the specified number of observations. This corresponds to a pairwise formal test of the null hypothesis of no correlation between variables versus an alternative of (positive or negative) correlation at a 95% confidence level. Again, in making multiple comparisons the global confidence level of the test (over all pairs) is lower than the nominal pairwise level, but the number of significant results can, for the sake of this assessment, be interpreted as an indicator of the suitability of the data for PCA. If a transformation significantly reduces the number of linear correlations between variables, then it can be suspected that this may cause a problem in subsequent analysis.

3.10.4. Perform a Data Transformation if Required

As presented in Weber et al. (8) two transformations commonly employed in ecological data analysis can also be used for BIOLOG™ microplate data: the Taylor power law transformation and the logarithmic transformation. The Taylor transformation (24) is commonly used to stabilize variances and make data conform to the assumptions of parametric analysis such as normality (22). It is based upon the assumption that:

$$S^2 = a\bar{y}^2 \qquad [7]$$

where S is the standard deviation of a sample variable, $\bar{y}$ is the mean of a sample variable and a is the sampling factor. This leads to

$$\log S^2 = \log a + b \log \bar{y}^2 \qquad [8]$$

where the slope, b, may be obtained by linear regression of the data for all variables. This leads to the conditional transformation:

$$y_i' = y_i^{(1-b/2)} \quad \text{for } b \neq 2 \qquad [9]$$

or,

$$y_i' = \ln(y_i) \quad \text{for } b = 2 \qquad [10]$$

where y_i' is the value of the transformed variable.

A logarithmic transformation can also often serve to normalize skewed data (22). A common logarithmic transformation used in ecological data analysis is of the form:

$$A' = \ln(\bar{A}_k + 1) \qquad [11]$$

where A′ is the value of the transformed variable.

These two simple data transformation examples are given to provide a starting point for utilizing transformations when performing multivariate data analysis of microplate data. Many other data transformations exist and may be more suitable for specific data sets. See (22) and (25) for detailed discussions regarding different data transformation techniques.

3.11. Perform PCA on the Data Set

Principle component analysis (PCA) is the most commonly employed multivariate analysis technique when working with BIOLOG™ microplate data. PCA is based on an eigenanalysis of an R-mode (between variables) variance− covariance matrix (22). In short, PCA is able to take a high dimensional space (32 dimensions in this case) and ordinate samples (objects) on a two-dimensional plane while preserving the maximum allowable amount of variance within the data set. PCA is most commonly used to visualize data plotted on the first two principle component (eigenvector) axes for interpretation. Common uses include the study of ecological shifts over time and space (22). PCA analysis can preserve varying degrees of the original variance within the first two axes; values from 40 to 80% are commonly achieved. PCA ordinations allow the CSUPs from the bacterial community samples to be grouped and differentiated. PCs are most commonly extracted from the covariance matrix of the data. Use of the covariance matrix preserves scale.

PCA has been widely adopted for analysing CLPPs based on CSUPs generated using BIOLOG™ microplates. As outlined in previous sections, attention needs to be paid to the distribution of the underlying variables and the possibility of applying a transformation to the data to improve the analysis. One of the significant advantages of PCA is that it is robust and analyses remain valid even if the assumptions of normality and homoscedasticity are not met; however, the analysis can be improved if the data can be transformed to meet these assumptions. Recent CLPP example studies utilizing PCA analysis of BIOLOG™ microplate data include (15,26,27).

Included as **Table 16.2** is a “Quick Reference Guide”, which attempts to briefly summarize the steps in the CLPP protocol within a single page for laboratory use.

Table 16.2
CLPP protocol – quick reference guide

I – Inoculation
→150 μL of a suspended mixed microbial sample into each of the 96 wells
NOTES: Time between sampling and inoculation should be minimized. Keep inoculation densities similar between plates.
II – Incubation
→ Incubate at room temperature
NOTES: Incubation period selected based on the type of metric to be used for data analysis (Step IV-A-1). Common incubation times between 10 and 200 h. Common incubation temperatures between 20 and 30°C.
III – Data Collection
→ Periodic absorbance (590 nm) readings taken for all wells during incubation period
NOTES: Plates should be shaken before each reading.
IV – Data Analysis
(A) PCA Analysis
(1) Choose a metric
(i) An absorbance value for each well from a specific incubation time point
(ii) An absorbance value for each well taken from a time point representing a specific AWCD for that plate
(iii) Some type of logistic curve fitting value such as lag, slope, area under the curve, or an asymptote value.
(2) Standardization of data if not performing a kinetic analysis
(3) Check data set for
(i) normality
(ii) homoscedasticity
(iii) underlying factor structure– liner correlations
(4) Perform a data transformation if required
(5) PCA analysis
(B) Other Analysis Methods
(1) Clustering analysis
(2) Diversity indices – substrate diversity, substrate richness, substrate evenness
(3) Alternative methods – factor analysis, PCoA, DCA, NMDS, DA, CCorA, RDA

3.12. Other Analysis Methods

In addition to PCA, a number of other analysis methods have been successfully utilized in garnering information from BIOLOG™ plates. Some of the more common methods include:

(1) Clustering analysis

(2) Diversity indices – substrate diversity, substrate richness, substrate evenness

(3) Alternative methods – factor analysis, PCoA, DCA, NMDS, DA, CCorA, RDA

3.12.1. Clustering Analysis

As recommended by Legendre and Legendre (22), clustering analysis is often performed to verify and validate results obtained using PCA. Clustering analysis allows for the CSUP similarities to be visualized in a dendogram for any given number of plates. An unweighted pair-group method using arithmetic averages (UPGMA) clustering analysis is often recommended. See (15,26,28) for recent examples of clustering analysis using CSUPs from BIOLOG™ microplates.

3.12.2. Substrate-Related Diversity Indices

BIOLOG™ plates have also been used, in a more traditional ecological sense, to calculate diversity indices based on CSUPs (29). The Shannon index or what is often called "diversity" is a common ecological metric used to track and understand shifts in communities over space and time. Using the CSUP gathered from a single BIOLOG™ plate, substrate diversity (H) can be calculated as:

$$H = -\sum p_i \ln(p_i), \qquad [12]$$

where

H – substrate diversity

p_i – ratio of the activity of a particular substrate to the sums of activities of all substrates

activity– chosen metric for analysis (absorbance value, kinetic parameter, etc.)

Two other parameters associated with substrate diversity which can be calculated using CSUPs are substrate richness (S) and substrate evenness (E). Substrate richness is a measure of the number of different substrates utilized by a microbial population. Substrate evenness is defined as the equitability of activities across all utilized substrates; substrate richness is calculated as the number of wells with a corrected absorbance greater than 0.25. Substrate evenness is calculated as:

$$E = H/H_{\max} \qquad [13]$$

Recent examples of studies utilizing the Shannon index include (15,26,27).

3.12.3. Multivariate Analysis Methods

Although PCA is the most popular approach, with proper data treatment essentially any multivariate analysis technique can be used to analyse the data matrix collected when applying the CLPP protocol to any number of mixed microbial community samples. A short list of reference studies utilising some of the less popular and/or more recently introduced multivariate methods either

examining or relating BIOLOG™ data to other data sets include: factor analysis (30), principle coordinates analysis (PCoA) (31), detrended correspondence analysis (DCA) (32,33), non-metric dimensional scaling (NMDS) (14), discriminant analysis (DA) (20), canonical correlations analysis (CCorA) (34), and redundancy analysis (RDA) (27). For an in-depth description of the mentioned multivariate methods, see Legendre and Legendre (22).

4. Notes

1. Using a minimum inoculation density of 10^5 cells/mL is the best way to reduce lag times although smaller inoculation densities can be used. True cell densities can be difficult to determine; therefore, an alternative inoculation approach is to dilute the sample to an optical density if ~0.2 at 420 nm for the suspended mixture. This inoculation approach may lead to inoculation density differences, but from a practical perspective is more easily controlled and implemented.
2. Reading frequencies can vary quite widely. For the data seen in **Fig.15.2**, a reading frequency of 4 h was used. This reading frequency provided enough data to decipher the sigmoidal shape of the colour development curves. However, if a larger inoculum density was used, a faster response would be observed, and therefore more frequent readings would be required in order to properly decipher and/or model the sigmoidal shape of the colour development curves. Preliminary trial runs using inoculum densities and bacterial communities similar to one's study samples is always a good idea. They can help one determine inoculum dilutions, reading frequencies and metric choices, which are essential in gathering meaningful data.
3. Following anaerobic inoculation, plates can be covered with non-slit silicon plate seal, or simply sealed around the edges with a generous amount of parafilm and masking tape. Both procedures have been proven effective. Microplate absorbance readings are then periodically taken without removing the plate lids. In the author's experience, overall profiles and readings have been shown to not be significantly affected by leaving the lid on during plate readings. Some plate readers require that the lids to be removed before absorbance readings can be taken. These plate readers unless equipped with or situated in a nitrogen purging area would not be suitable for anaerobic samples.
4. Each metric contains associated positives and negatives. Use of a single time point reading often guides the user towards using data points in an area where almost all carbon source

utilization curves are in a stationary (steady-state) phase. This can be useful when comparing plates over extended time periods, as the basis for comparison is relatively stable. However, in interpreting this type of data, one should be aware that the activity levels of the community on specific carbon sources is not emphasized due to the carbon utilization curves being in the stationary phase.

5. When choosing to use a reference AWCD, one should be aware that the carbon utilization curves can be in the lag, exponential growth or the stationary phase. Therefore, this method, although based on a fixed reference point, may not give stable comparison results over an extended time period for community monitoring studies. However, in comparison to a fixed time point, this method does emphasize activity in each well, which may be of interest to the user. It should also be mentioned that using an AWCD reference point may not be appropriate in studies where some plates contain a large number of unresponsive wells.
6. Kinetic analysis allows one to compare many different aspects of the carbon utilization curves, allowing the user to tune the analysis to emphasize a specific aspect of the community. However, in modelling the data sometimes poor-fit can occur, and in many instances kinetic analysis is not feasible due to time or instrument constraints.
7. The authors have found that with larger data sets (100+ objects), normality, homoscedasticity and the underlying factor structure of the data do not have as large an effect on PCA results and subsequent data interpretation when compared to smaller data sets. However, it should be emphasized that assessing the data set for normality, homoscedasticity and the underlying factor structure, and considering an initial data treatment are necessary steps in the data analysis procedure. If a large difference between PCA results from transformed and untransformed data is not observed, the data set was likely already suited for PCA or was of a size where a data transformation did not have a large effect on the PCA results.

References

1. Lehman, R. M., Colwell, F. S., Ringelberg, D. B., and White, D. C. (1995) Combined microbial community-level analyses for quality assurance of terrestrial subsurface cores. *J. Microbiol. Meth.***22**, 263–281.
2. Insam, H. (1997) A new set of substrates proposed for community characterization in environmental samples, in *Microbial Communities: Functional Versus Structural Approaches* (Insam, H., and Rangger, A., eds.), Springer-Verlag, Berlin.
3. Campbell, C. D., Grayston, S. J., and Hirst, D. J. (1997) Use of rhizosphere carbon sources in sole carbon source tests to discriminate soil microbial communities. *J. Microbiol. Meth.* **30**, 33–41.
4. Preston-Mafham, J., Boddy, L., and Randerson, P. F. (2002) Analysis of microbial community functional diversity using sole-

carbon-source utilisation profiles – a critique. *FEMS Microbiol. Ecol.* **42**, 1–14.
5. Garland, J. L., and Mills, A.L. (1991) Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. *Appl. Environ. Microbiol.* **57**, 2351–2359.
6. Konopka, A., Oliver, L., and Turco, R. F. (1998) The use of carbon substrate utilization patterns in environmental and ecological microbiology. *Microb. Ecol.* **35**, 103–115.
7. Garland, J. L. (1997) Analysis and interpretation of community-level physiological profiles in microbial ecology. *FEMS Microbiol. Ecol.* **24**, 289–300.
8. Weber, K. P., Grove, J. A., Gehder, M., Anderson, W. A., and Legge, R. L. (2007) Data transformations in the analysis of community-level substrate utilisation data from microplates. *J. Microbiol. Meth.* **69**, 461–469.
9. Degens, B. P., Schipper, L. A., Sparling, G. P., and Duncan, L. C. (2001) Is the microbial community in a soil with reduced catabolic diversity less resistant to stress or disturbance? *Soil Biol. Biochem.* **33**, 1143–1153.
10. Garland, J. L., Roberts, M. S., Levine, L. H., and Mills, A. L. (2003) Community-level physiological profiling performed with an oxygen-sensitive fluorophore in a microtiter plate. *Appl. Environ. Microbiol.* **69**, 2994–2998.
11. Franklin, R. B., Garland, J. L., Bolster, C. H., and Mills, A. L. (2001) Impact of dilution on microbial community structure and functional potential: Comparison of numerical simulations. *Appl. Environ. Microbiol.* **67**, 702–712.
12. Garland, J. L., Mills, A. L., and Young, J. S. (2001) Relative effectiveness of kinetic analysis vs single point readings for classifying environmental samples based on community-level physiological profiles (CLPP). *Soil Biol. Biochem.* **33**, 1059–1066.
13. Christian, B. W., and Lind, O. T. (2006) Key issues concerning Biolog use for aerobic and anaerobic freshwater bacterial community-level physiological profiling. *Internat. Rev. Hydrobiol.* **91**, 257–268.
14. Classen, A. T., Boyle, S. I., Haskins, K. E., Overby, S. T., and Hart, S. C. (2003) Community-level physiological profiles of bacteria and fungi: plate type and incubation temperature influences on contrasting soils. *FEMS Microbiol. Ecol.* **44**, 319–328.
15. Weber, K. P., Gehder, M., and Legge, R. L. (2008) Assessment of the changes in the microbial community in response to acid mine drainage exposure. *Water Res.* **42**, 180–188.
16. Winding, A., and Hendriksen, N. B. (1997) Biolog substrate utilization assay for metabolic fingerprints of soil bacteria: incubation effects, in *Microbial Communities: Functional Versus Structural Approaches* (Insam, H., and Rangger, A., eds.), Springer, Berlin.
17. Mills, A. L., and Garland, J. L. (2002) Application of physiological profiles to assessment of community properties, in *Manual of Environmental Microbiology 2nd Edition* (Hurst, C. J., Crawford, R. L., Knudsen, G. R., Mcinerney, M. J., and Stetzenbach, L. D., eds.), ASM Press, Washington, DC.
18. Beaumont, V. L. (2007) Investigation of Microbial Fuel Cell Performance and Microbial Community Dynamics during Acclimation and Carbon Source Pulse Tests. *MASc Thesis*, Department of Chemical Engineering, University of Waterloo. 200 pp.
19. Haack, S. K., Garchow, H., Klug, M. J., and Forney, L.J. (1995) Analysis of factors affecting the accuracy, reproducibility, and interpretation of microbial community carbon source utilization patterns. *Appl. Environ. Microbiol.* **61**, 1458–1468.
20. Mondini, C., and Insam, H. (2003) Community level physiological profiling as a tool to evaluate compost maturity: a kinetic approach. *Eur. J. Soil Biol.* **39**, 141–148.
21. Glimm, E., Heuer, H., Engelen, B., Smalla, K., and Backhaus, H. (1997) Statistical comparisons of community catabolic profiles. *J. Microbiol. Meth.* **30**, 71–80.
22. Legendre, P., and Legendre L. (1998) *Numerical Ecology* (2nd edition). Elsevier B.V. Amsterdam, Netherlands.
23. Lo, A. W., and MacKinlay, C. A., (1988) Stock market prices do not follow random walks: evidence from a simple specification test. *Rev. Financial Stud.* **1**, 41–66.
24. Taylor, L. R. (1961) Aggregation, variance, and the mean. *Nature.* **189**, 732–735.
25. Montgomery, D. C. (2001) *Design and Analysis of Experiments* (5th edition). John Wiley and Sons, Inc., New York.
26. He, X.Y., Wang, K., Zhang, W., Chen, Z., Zhu, Y., and Chen, H. (2008) Positive correlation between soil bacterial metabolic and plant species diversity and bacterial and fungal diversity in a vegetation succession on Karst. *Plant Soil.* **307**, 123–134.
27. Farnet, A.M., Prudent, P., Cigna, M., and Gros, R. (2008) Soil microbial activities in

a constructed soil reed-bed under cheese-dairy farm effluents. *Biores. Technol.* **99**, 6198–6206.
28. De Paolis, M. R., and Lippi, D. (2008) Use of metabolic and molecular methods for the identification of a Bacillus strain isolated from paper affected by foxing. *Microbiol. Res.* **163**, 121–131.
29. Zak, J. C., Willig, M. R., Moorhead, D. L., and Wildman, H. G. (1994) Functional diversity of microbial communities: a quantitative approach. *Soil Biol. Biochem.* **26**, 1101–1108.
30. Niklinska, M., Chodak, M., and Laskowski, R. (2005) Characterization of the forest humus microbial community in a heavy metal polluted area. *Soil Biol. Biochem.* **37**, 2185–2194.
31. Hackett, C. A., and Griffiths, B. S. (1997) Statistical analysis of the time-course of Biolog substrate utilization. *J. Microbiol. Meth.* **30**, 63–69.
32. Houlden, A., Timms-Wilson, T. M., Day, M. J., and Bailey, M. J. (2008) Influence of plant developmental stage on microbial community structure and activity in the rhizosphere of three field crops. *FEMS Microbiol. Ecol.* **65**, 193–201.
33. Garland, J. L. (1996) Analytical approaches to the characterization of samples of microbial communities using patterns of potential C source utilization. *Soil Biol. Biochem.* **28**, 213–221.
34. Leriche, F., Bordessoules, A., Fayolle, K., Karoui, R., Laval, K., and Dufour, E. (2004) Alteration of raw-milk cheese by *Pseudomonas spp.*: monitoring the sources of contamination using fluorescence spectroscopy and metabolic profiling.*J. Microbiol. Meth.* **59**, 33–41.