Applied Soil Ecology xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Applied Soil Ecology

journal homepage: www.elsevier.com/locate/apsoil

Short communication

Rapid analysis of photoautotroph microbial communities in soils by flow cytometric barcoding and fingerprinting

László Menyhárt[a], Szabolcs Nagy[b], Anita Lepossa[c,*]

[a] University of Pannonia, Georgikon Faculty, Department of Economic Methodology, H8360 Keszthely, Deák F. u. 16., Hungary
[b] University of Pannonia, Georgikon Faculty, Department of Animal Sciences, H8360 Keszthely, Deák F. u. 16., Hungary
[c] University of Pannonia, Georgikon Faculty, Department of Crop Production and Land Use, H8360 Keszthely, Deák F. u. 16., Hungary

ARTICLE INFO

Keywords:
Soil algae
Microbial community structure
Flow cytometry
Barcoding
Fingerprinting

ABSTRACT

The investigations of soil microbiota and soil interactions are more and more significant from several points of view, including climate studies, environmental or human health. As a rapid cultivation-independent technique to study the alterations in soil photoautotroph microbial community structures flow cytometry was tested in soil-water suspensions prepared from the fresh soil of different crop cultures. Flow cytometric data were analyzed with barcoding and fingerprinting approaches to study the soil algal (< 30 μm) community structure based on chlorophyll-a and phycoerythrin autofluorescence. Both statistical approaches revealed larger differences between crop cultures than within, indicating that both methods have potential in studying the dynamic changes of soil microalgal community structures.

1. Introduction

The investigations of soil microbiota and soil interactions are more and more significant from several points of view, including climate studies, environmental or human health. However, such studies are rather biased and focus on such taxa only which can be cultivated in the laboratory – usually about < 1% of the microbial taxa present in the given environment (De Roy et al., 2012; Koch et al., 2013). Moreover, studies focusing only on cultivable microbes are rather time consuming (De Roy et al., 2012) and unable to reveal effects of abiotic factors as such detection methods are not rapid enough to study changes in community dynamics (Koch et al., 2013).

Up to date molecular fingerprinting methods that do not require cultivation have been developed and used, like PCR-DGGE (denaturing gradient gel electrophoresis), T-RFLP (terminal restriction fragment length polymorphism) or next generation sequencing (Daquiado et al., 2016; Koch et al., 2014a; Yu et al., 2018), as well as the PLFA (phospholipid fatty acid analyses) (Dai et al., 2017; Yuan et al., 2015; Zhao et al., 2015). However, such approaches are still qualitative or semi-quantitative (i.e. no actual cell numbers can be determined), time consuming, labor intensive (De Roy et al., 2012) and rather expensive.

Recently, flow cytometry was introduced as a rapid cultivation-independent technique for studying dynamic changes in microbial community structures (De Roy et al., 2012; Koch et al., 2013, 2014a, 2014b; van Gelder et al., 2018). The advantages of flow cytometric methods are the rapid data acquisition and analysis (even under 30 min) and that the saved data files can be reanalyzed later or compared to previously collected samples (De Roy et al., 2012). Such fast measurements where data are available immediately allow quick, on site studies of microbial population dynamics even within minutes (Koch et al., 2014a).

Koch et al. (2014b) suggested four approaches to study the dynamics of complex microbial communities with flow cytometric fingerprinting methods, where fingerprints are based on the number of cell clusters, the position of these clusters and cell numbers per cluster. Two of these approaches are based on image analysis of plots generated with flow cytometric analysis softwares (Dalmatian plot and Cytometric Histogram Image Comparison, CHIC), one applies a gate template built subjectively by the user (cytometric barcoding, CyBar) and one is an operator independent approach, based on a geometrical grid (FlowFP).

Although flow cytometry has been widely used in aquatic environmental studies since the 1980s (e.g. Olson et al., 1985; Troussellier et al., 1993; Vives-Rego et al., 2000; Wang et al., 2010), its application in soil microbiology is still scarce (Frossard et al., 2016; Lentendu et al., 2013; Resina-Pelfort et al., 2003; Zhang et al., 2017).

The aim of the present pilot study was to test two of the approaches suggested by (Koch et al., 2014b), namely, CyBar and FlowFP to study the community structure of photoautotroph soil microorganisms (< 30 μm) based on their chlorophyll-a and phycoerythrin autofluorescence (Givan, 2001) reanalyzing the data of our earlier investigation (Lepossa et al., 2015).

* Corresponding author.
E-mail addresses: menyhart-l@georgikon.hu (L. Menyhárt), nagy.szabolcs@georgikon.hu (S. Nagy), lepossaa@georgikon.hu (A. Lepossa).

https://doi.org/10.1016/j.apsoil.2018.06.013
Received 6 June 2018; Accepted 16 June 2018

2. Materials and methods

2.1. Soil sample collection, preparation

Soil samples (six per culture) were collected from four crop cultures: winter wheat (*Triticum aestivum* L.); crimson clover (*Trifolium incarnatum* L.); maize (*Zea mays* L.); soybean (*Glycine* max L. Merrill.) from the top five cm soil layer in Keszthely (Hungary, N46,74°, E14,24°) in July of 2013. The soil type was the same, Eutric cambisol in each plant culture. Every soil sample was a pool of three subsamples collected in plastic bags and then subsequently transferred to the laboratory. After a thorough mixing, soil water content and pH were measured, and the fresh soil samples were stored aseptically at +4 °C in the dark overnight.

For flow cytometry, 10 g fresh soil/ 100 ml distilled water suspensions were prepared in 250 ml Erlenmeyer flasks, then homogenized in an ultrasound bath (50 Hz, 280 sec) according to our direct algal counting protocol (Lepossa and Ördög, 2006). After a 30 sec sedimentation, 900 µl suspensions were filtered through a 30 µm filter (Sysmex CellTrics disposable filters) and subsequently measured with a flow cytometer.

2.2. Flow cytometry

Measurements were done with a Beckman Coulter FC-500 flow cytometer equipped with a 488 nm Ar ion laser (20 mV) and FC-500 CXP acquisition software. Flow rate was set to "LOW" (10 µl/min) and acquisitions were stopped after 5 min. Data files were stored in standard LMD file format.

Chlorophyll-a autofluorescence was detected on FL4 (675 nm BP) detector, while phycoerythrin autofluorescence was detected on FL2 (575 nm BP) detector, in linear mode. Chlorophyll-a fluorescence was set as discriminator to exclude other events in the soil suspensions from the data files. The flow cytometer was calibrated with FlowCheck fluorospheres (6605359, Beckman Coulter); moreover, in order to monitor random drift in fluorescence intensities between samples, 10 µl FlowSet fluorosphere (6607007, Beckman Coulter) was added to every sample, serving as internal control.

2.3. Data analysis

2.3.1. Cytometric barcoding

Flow cytometric barcoding was done according to Schumann et al. (2015), using the flowCyBar package under R environment. Dual parameter chlorophyll-a vs. phycoerythrin density plots were drawn with WinMDI 2.8. software and a gate template was established manually via the visual inspection of every sample (Fig. 1a). Internal control fluorospheres were gated out from data analysis. Individual samples were then analyzed with the gate template and community structures were compared with the flowCyBar package. Event number variations (indicating algae cell abundance variations) within gates were compared between samples after data normalization and non-metric multidimensional scaling (NMDS) was applied to visualize the distances between samples (Fig. 1c) using normalized event numbers. Moreover, the flowCyBar package was used to calculate correlations between the event numbers of autofluorescent subpopulations and the measured abiotic factors (soil humidity and pH).

2.3.2. FlowFP

An additional method to analyze the flow cytometric data was the combination of probability binning (PB) and multiclass classification. For each sample the crop culture was estimated from the fingerprint, then the estimated-observed hit ratio was calculated.

In the first step the fingerprint of each sample was determined with PB using data of FL4 and FL2 channels. The process of PB is briefly as follows: all the samples are pooled and the pooled population is divided into two bins such that both bins contain the same number of events. The division is made at the median of the fluorescence parameter with the largest variance. The algorithm then repeats the process on each of the two newly-defined bins, again determining the median and variance of all parameters for each bin (Fig. 1b). After n iterative steps the number of bins is 2^n. The obtained PB model is then applied to each individual sample, which results in a feature vector of counts for each bin of the model. This vector is referred to as fingerprint (De Roy et al., 2012). Additional details on PB can be found in Roederer et al. (2001). The fingerprints were produced on the base of FL2 and FL4 with the flowFP package (Holyst and Rogers, 2009) from R/BioConductor. The number of iterative steps in PB was varied between 3 and 7.

In the next step the crop culture was estimated with regularized multinomial logistic regression. The bin counts were considered as explanatory variables, so totally 2^n explanatory variables were used. Given the large number of explanatory variables, lasso regularization was used to reduce the variance of the estimation. More details on lasso regularization are written by Tibshirani (1996). Calculations, including the lambda hyperparameter controlling the lasso regularization, were performed using H2O R package (H2O.ai team, 2017).

The accuracy of the model was assessed by the estimated-observed hit ratio, which is the proportion of the well-classified samples. It measures the separability of the samples from different crop cultures on the base of flow cytometric data. The hit ratio was estimated with leave-one-out cross-validation (LOOCV). Similarly to barcoding, non-metric multidimensional scaling (NMDS) was used to visualize distances between samples (Fig. 1d).

3. Results

3.1. Cytometric barcoding

The soil microalgal (< 30 µm) community structure was similar in the soil samples of all investigated crop cultures – the gate template contained four regions only (Fig. 1a). The subpopulation with highest event numbers showed both high chlorophyll-a and phycoerythrin autofluorescence indicating the prevalence of the phycoerythrin containing cyanobacteria. Overall event numbers were low (i.e. < 1000 events during acquisition). NMDS analysis revealed larger distances between than within crop cultures, except for crimson clover (Fig. 1c). Correlation analysis between abiotic factors and event numbers in different gates showed significant correlations between soil pH and region 3 ($r = 0.59$, $p < 0.001$), and between soil humidity and region 2 ($r = 0.42$, $p < 0.05$).

3.2. FlowFP

The largest hit ratio was obtained for n = 4. Its value was 79% ($p = 3 \cdot 10^{-7}$), that is 19 out of the 24 samples were correctly classified. All of the wheat, 5 of the maize, 4 of the crimson clover and 4 of the soybean samples were correctly classified.

4. Discussion

Both flow cytometric data analysis approaches revealed larger differences between crop cultures than within, which indicates that both methods have potential in studying soil microalgal community structures. FlowFP may be considered superior since this method is not dependent on operator experience as the method is fully automatic and does not need gates set by the operator. On the other hand, this method requires deeper knowledge in data analysis in R environment and the binning procedure does not reveal biological subcommunities (Koch et al., 2014b). However, it is suitable for detecting quick community structure changes when taxonomy is not a priority. FlowCyBar is much more dependent on operator experience as the gate template is established subjectively (De Roy et al., 2012). However, gates offer the