Metagenomic approach to soil microbial diversity and functions

Takeshi Fujii, Sho Morimoto, Yuko T. Hoshino, Hiroaki Okada,
Yong Wang, Haiyan Chu**, Bao Zhihua, Matushita Yuko, Seiya Tsushima
National Institute for Agro-Environmental Sciences
Kannondai 3-1-3, Tsukuba, Ibaraki 305-8604, Japan.

**Institute of Soil Science, Chinese Academy of Sciences,
East Beijing Road 71, Nanjing 210008, China
Email: ftakeshi@affrc.go.jp

Abstract

The improved methods to analyze DNA or RNA from soil enable us to study microbial diversity and functions in soil which have never been revealed by conventional culture methods. Using one of such metagenomic approaches, polymerase chain reaction-denaturing gradient gel electrophoresis (PCR-DGGE), we succeeded in monitoring the change of microbial diversity in the soil treated with a chemicals or various fertilizer. Based on these achievements, since 2006, National Institute for Agro-Environmental Sciences (NIAES) play a central role in a Japanese national project addressing development of analytical methods for soil biological properties using environmental DNA (eDNA). In the project, we have optimized and standardized the conditions for preparation of DNA from soil and PCR-DGGE using soil metagenome to compare the microbial communities in different soil samples. The obtained PCR-DGGE data accompanying the physical and chemical soil properties and information about crop productivity have been used to construct a database (eDDASs: eDNA database for agricultural soils). In the project, we also succeeded in detecting microbial gene expression in soil using RNA extracted from soil. Our final goal is to determine the relationships between the biological properties evaluated by eDNA or RNA analysis and the soil physicochemical properties as affected by cultivation practice.

Keywords: Soil extracted DNA and RNA/PCR-DGGE/eDDASs: eDNA database for agricultural soils

1. Introduction

The rapid advancement of today's molecular biological techniques brought a new frontier of science, metagenomics. The new analytical approaches using DNA extracted directly from the environment enable us to access the genome, which is called metagenome, of all microorganisms inhabiting in the environment. The information obtained from these approaches will reveal more precise figure of biodiversity and much more useful functions of microorganisms in the environment than those obtained from previously cultured microorganisms. Especially, soil metagenomics expect to give us such useful information concerning about agriculture.

It is estimated that microorganisms appeared 3 billion years ago on the earth, whereas human being appeared there only 5 million years ago. Though human beings flourish all over the world in these days, microbes still surpass us in mass and numbers. Especially, soil microbes play important roles in material cycles (carbon, nitrogen etc.) in the environment through decomposition of

organic materials derived from plants, animals and microorganisms. In other words, we live on the earth that has been constructed by microbes and depend on them completely. Soils contain large amounts of bacteria of numerous species, but more than 99% of these species remain unknown and cannot be cultured (Amman *et al*. 1995). In agricultural fields, to achieve sustainable production of high quality crops and vegetables, it is essential to maintain soil fertility and to overcome soil disease such as sickness due to continuous cropping. For this purpose, we have to evaluate not only physical and chemical properties of soils but also biological ones and to utilize their information for soil management. Especially, it is important to elucidate what kind of microbes exist in soil or how soil microbes are involved in the phenomena in agricultural fields, such as crop growth, pest infestation and suppressiveness. We need to collect the information about biodiversity and functions of soil microbes, and compare it across soil samples. Soil microbial biomass and activities are frequently used as an early indicator of changes in soil chemical and physical properties resulting from soil management and environmental stresses in agricultural ecosystems. However, our knowledge on the soil biological properties has been still far from enough to evaluate the soil biological properties for soil management. This is mainly because most of soil microorganisms cannot be cultivated with conventional methods. Useful general methods for examining soil microbial population and functions have been desired for long time. Recent methodological developments in molecular microbial ecology have made it possible to analyze the microbial communities in environmental samples without any culture steps. Today, one of such methods, denaturing gradient gel electrophoresis (DGGE) of polymerase chain reaction (PCR)-amplified genes from environmental samples, is a useful tool in environmental microbiology (Muyzer 1999). For example, the PCR-DGGE method targeting the 16S rRNA gene (16S rDNA) (Muyzer *et al*. 1993) is most widely applied for the studies of bacterial community structure in the environment, because this gene is essential to all living procaryotes and helpful in tracing phylogenetic relationships. The PCR-DGGE targeting appropriate sequences can also be used for studying the community structures of other microorganisms in the environment, such as fungi (Hoshino and Matsumoto 2007), or nematodes (Okada and Oba 2008). The method, like PCR-DGGE, analyzing nucleic acids directly extracted from the environment will bring us more wide range of information about communities and functions of microorganisms existing in soil environment.

Since we succeeded in establishment of improved DNA extraction methods using skim milk from soils that strongly adsorb DNA (Hoshino and Matsumoto 2004), we have developed new methods using nucleic acids for analyzing the structure and functions of soil microorganisms. Our several research achievements and trends of metgenomic approaches in the worlds are reviewed.

2. Establishment of improved DNA extraction methods using skim milk from soils

Direct extraction of DNA from soil is essential for the study of microbial ecology. This method facilitates the analysis of whole microbial communities, including unculturable microbes. Many protocols have been developed to yield DNA for molecular analysis. However, these methods are not satisfactory for some types of soil, especially for Andisols, which are volcanic ash soils widely distributed all over the world and especially in the circum-Pacific Ring of Fire, including Japan,

East Africa, and the Mediterranean. We devised an improved method of extracting DNA from Andisols (Hoshino and Matsumoto 2004). Seven Andisol samples from 6 agricultural fields and a forest in Japan were tested with a commercially available kit that uses bead-beating. Soil DNA could be extracted from only 2 of the samples. Pure DNA added to the soil was adsorbed to the clay particles, revealing why DNA extraction is difficult in Andisols. However, when we added skim milk to the extraction buffer at a rate of 40 mg g soil^{-1}, DNA could be detected by electrophoresis from all the soil samples (Hoshino and Matsumoto 2004). Certain molecules, such as RNA and skim milk, are known to be adsorbed to soil colloids in a similar manner to that of DNA. Soil extracts proved to be suitable for PCR when skim milk was added, rather than RNA. Addition of skim milk did not affect PCR-DGGE profiles. PCR-DGGE is a complementary tool for determining the genetic diversity of natural microbial communities in a culture-independent fashion. Our improved method is applicable to the molecular analysis of microbial communities in soils that strongly adsorb DNA.

3. Isolation of Effective 3-Chlorobenzoate-Degraders in Soil using Community Analyses by PCR-DGGE

Screening of pollutant degraders by relying only on cultivation techniques such as liquid enrichment often fail to isolate the actual degraders in the environment. In fact, we found that PCR-DGGE patterns targeting benzoate 1,2-dioxygenase alpha subunit genes (*benA*) of soil culture amended with 3-chlorobenzoate (3CB), substrate for *ben* operon, were completely different from those of liquid culture (Fig. 1). We succeeded in demonstrating that community analyses by PCR-DGGE were successfully used to detect and isolate bacteria that can degrade 3CB effectively in soil (Morimoto *et al*. 2005, Morimoto *et al*. 2008). A forest soil sample was repeatedly dosed with 3CB (500 mg Kg^{-1}) to enrich the indigenous 3CB-degraders, and changes in the bacterial community were monitored by PCR-DGGE of 16S rRNA genes and *benA*. Initially, it required about 3 weeks to degrade 3CB in the soil, whereas it took only 3 days after the third dose. With this accelerated degradation, several intensified bands appeared in the DGGE profiles of both 16S rRNA and *benA* genes. We succeeded in isolating five 3CB-degrading *Burkholderia* strains corresponding to these bands from the 3CB-dosed soil by direct plating, while most of them fail to proliferate (or grow) in liquid enrichment (Morimoto *et al*. 2008). Inoculation of the strains into the soil demonstrated that the five strains could degrade 3CB effectively in the soil (Morimoto *et al*. 2008). This study clearly shows significant bias during the liquid enrichment process and the advantage of using PCR-DGGE in screening effective degraders under environmental conditions.

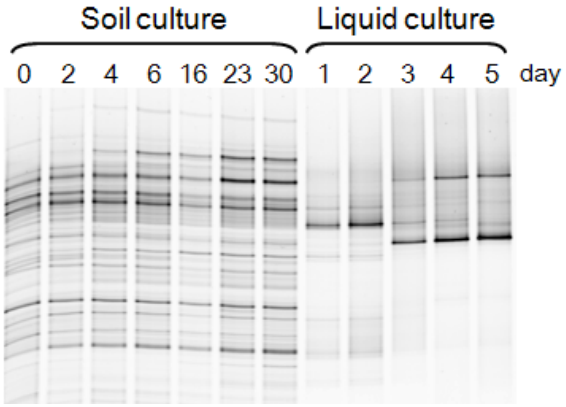

Fig. 1 Detection of 3CB-degradable bacteria in soil or liquid culture by PCR-DGGE targeting *benA*. For a liquid culture, the basal salts medium [18] (99ml) is inoculated with 1g of 3CB-dosed forest soil. At 0 day, 3CB (500 mg Kg^{-1} for soil culture or 500mgL^{-1} for liquid culture) was added.

4. Community Structure of Ammonia-oxidizing Bacteria Revealed by PCR-DGGE Analysis

PCR-DGGE analysis also reveal the effects of mineral fertilizer (NPK) and organic manure on the community structure of soil ammonia-oxidizing bacteria (AOB) in a long-term fertilizer experiment, which has been carried out for 16 years in agricultural fields in China (Chu *et al*. 2007). The experiment included seven treatments: organic manure, half-organic manure N plus half-fertilizer N, fertilizer NPK, fertilizer NP, fertilizer NK, fertilizer PK and the control (without fertilization). N fertilization greatly increased soil nitrification potential, and mineral N fertilizer had a greater impact than organic manure, while N-deficiency treatment (PK) had no significant effect (Chu *et al*. 2007-1). AOB community structure was analyzed by PCR-DGGE targeting the *amoA* gene, which encodes the α subunit of ammonia monooxygenase. DGGE profiles showed that AOB community was more diverse in N-fertilized treatments than those in the PK treatment and the control, while one dominant band observed in the control could not be detected in all fertilized treatments (Fig. 2). Phylogenetic analysis showed that the DGGE bands derived from N-fertilized treatments, belonged to *Nitrosospira* cluster 3, indicating that N fertilization resulted in the dominance of *Nitrosospira* cluster 3 in soil (Chu *et al*. 2007-1). These results demonstrate that long-term application of N fertilizers could result in the increased soil nitrification potential, and the AOB community shifts in soil. On the other hand, PCR-DGGE analysis targeting the 16S rDNA gene revealed that mineral fertilization did not affect General bacterial community structure, while specific DGGE band was observed in organic manure-fertilized soils (Chu *et al*. 2007-2). These DGGE bands could be candidates for indicators of soils biological properties for soil management.

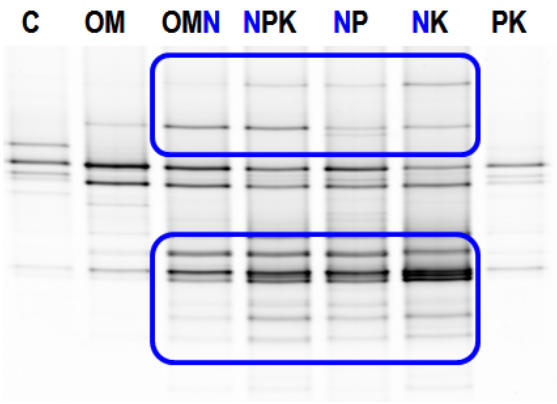

Fig. 2 DGGE analysis of *amoA* fragments retrieved from the soils under long-term (16 years) application of mineral fertilizer and organic manure. The *amoA* DGGE bands specifically found in soils treated with N (inside the squares).

5. eDNA Project

In 2006, Ministry of Agriculture, Forestry and Fisheries of Japan started a project addressing development of analysis methods for soil biological properties using environmental DNA (eDNA = metagenome) which is extracted from soil. It united researchers of national institutes, universities and prefectural agricultural experiment stations in Japan. The project consists of three subthemes.

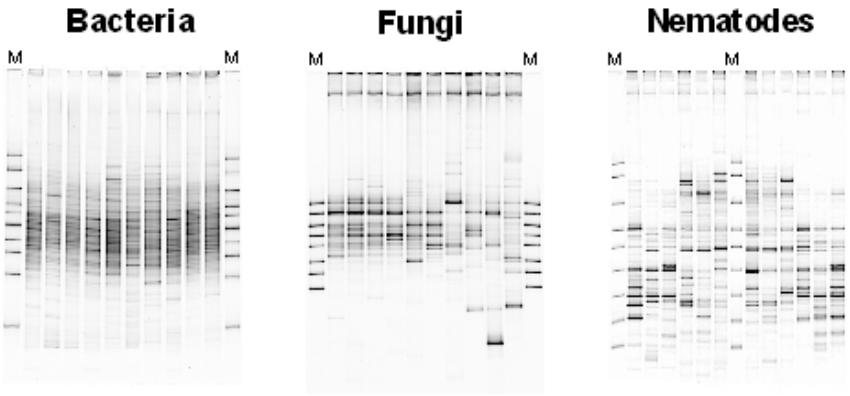

Fig. 3. Soil biological analysis by PCR-DGGE. Soil samples were prepared from various agricultural fields and DNA was extracted from them according to the methods in the manual. The extracted DNA was used to amplify bacterial 16S rDNA, fungal 18S rDNA or nematodes 18S rDNA by PCR with the specific primer sets that we had chosen. The PCR fragments were analyzed by DGGE according to the standardized method in the manual. Community structures of bacteria, fungi or nematodes in each agricultural field can be evaluated as band patterns. The PCR fragments obtained from each soil are aligned in the same order in each gel. “M” indicates DGGE marker we developed to compare band patterns across gels.

In the subtheme 1, we have optimized PCR-DGGE procedure for each of bacterial, fungal and nematode communities, and make up a manual. With the standardized and optimized DGGE procedures we developed, clear and sharp band patterns can be produced (Fig. 3). The subtheme 2 aims to develop analytical methods for elucidation of relationship between soil biodiversity and crop productivity. We have been collecting soil samples under various types of agricultural management at different localities from the northern to the southern Japan (Fig. 4), and analyzing the metagenome prepared from each soil sample by the standardized PCR-DGGE method. The subtheme 3 has been constructing a database (eDDASs, eDNA database for agricultural soils) collecting the results of soil metagenomic analysis accompanying the physical and chemical soil properties with information about crop productivity in each agricultural fields (Fig. 5). This database also provides some tools for data retrieval and analysis developed by this project and others. In the database, we can display the diversity indexes of the microorganism existing in the soil calculated from the digital image of PCR-DGGE patterns that had been obtained from various soil samples. Applications of new analytical methods using microarray or metagenome sequencing are also considered in this project. We expect to elucidate the mechanisms in which soil biota affect agricultural production, by analyzing the

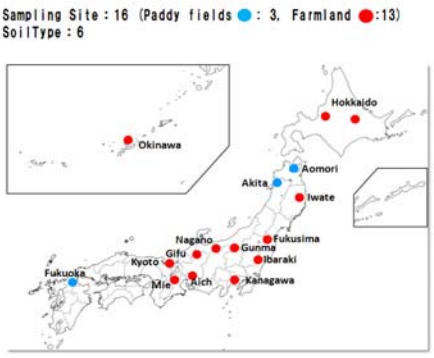

Fig. 4 Sampling place for eDNA project

Fig. 5 DDASs, eDNA database for agricultural soils

biological information revealed in our project. Our final goal is to determine the relationships between the biological properties evaluated by eDNA or RNA analysis and the soil physicochemical properties as affected by cultivation practice.

6. Trends in metagenomic approaches in the world

As it is rather rapid and cheep, the metagenomic approach using PCR-DGGE is convenient to compare total diversity of microbial community structure in different soil samples roughly, or to find specific DNA bands in each metagenome. However it provides us only a variation of DNA fragments amplified by PCR derived from one kind of gene. Conversely, metagenome-sequencing will give us complete information of biodiversity in the environment, while it requires huge budget and time. Sequencing a metagenome derived from a particular soil sample have a potential to give us genes that encode novel biocatalysts for biosynthetic or biodegradation processes such as degradation of pollutants, synthesis of biofuels and production of novel drugs. To screen such novel useful genes, more than ten of metagenomic libraries in a cloning vector have been constructed since 2000 (Elsas et al. 2008). Although these studies already have been successful in getting several novel genes and functions from metagenome, the basic work can be tedious and inefficient. Especially, hit rates of target genes can be very low if they exist in nondominant species. In this context, direct shotgun sequencing of metagenome using high-throughput sequencing methods is becoming increasingly popular. Large-scale metagenomic sequencing efforts will be necessary to provide sufficient data to understand the whole feature of soil microbial community diversity and function. However, the number of sequences to be determined is huge. It was estimated that one gram of soil contain 10^9-10^{10} microbes, and each of them contain 10^6-10^7 base pair of DNA. So far, such sequencing effort for entire metagenome cannot be undertaken by a single laboratory or even by a single country. One international effort to combine the skills of the global scientific community to focus on sequencing and annotating the soil metagenome was proposed (Vogel et al. 2009). This international project, the TerraGenome international sequencing consortium, has a primary objective: the complete sequencing of a 'reference' soil metagenome. The soil system chosen for investigation, Park Grass, is an internationally recognized agroecology field experiment that has been running for more than 150 years at the UK agricultural sciences institute, Rothamsted Research. The organizers of this project invite the international community to participate in this project, and hope to eventually expand the project to other soil sites.

7. Detection of gene expression in a Bacteria growing in soil using RNA extracted from soil

DNA extracted from soil can provide information on the structure and potential function of soil bacterial communities. However, to acquire detailed knowledge of the function of a soil community, gene expression

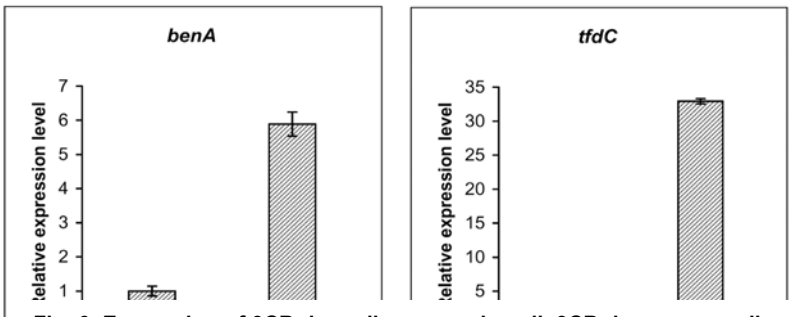

Fig. 6. Expression of 3CB degrading genes in soil. 3CB- is corresponding to Day1 soil culture without 3CB, and 3CB+ is corresponding to Day2 soil culture at the presence of 3CB. The expression levels have been normalized by 16S rRNA (an internal control).

studies using RNA extracted from soil are required. In cases of species that resist all attempts of cultivation, the direct detection of gene expression in the soil is the only choice. Thus, bacterial gene expression in soils is important in various contexts such as bioremediation, nitrogen cycling, carbon cycling, and bacterial community structure. As the first step of such a study, a method of RNA extraction from soil microorganisms has to be established. Although there are some reports describing RNA extraction from microorganisms in soil, the documented methods handled small amounts of soil only or resulted in the co-extraction of humic substances together with RNA. Humic substances interfere with many enzyme reactions, nucleic acid detection and measurement, and RNA hybridization. So, we developed a new method of extracting RNA with high purity from soil reproducibly (Wang 2008, Wang 2009). A soil RNA extraction method was improved by optimization of lysis conditions and further purification by a spin column, in terms of removing humic substances efficiently which hinder enzymatic reactions of the obtained RNA. Fluorescence spectrophotometry demonstrated that the improved method removed both humic and fulvic acids efficiently. By the improvement, the signal of gene expression detected by real-time RT-PCR increased 10-fold (Wang 2009). Using the method, we extracted RNA from a sterilized field soil, which was inoculated with *Pseudomonas putida* KT2440 transformed with a chloroaromatic degrading plasmid, in the presence or absence of 3-CB. Real-time RT-PCR performed with the extracted RNA confirmed the induction of chloroaromatic degrading genes in 3-CB amended soil (Fig. 5). These achievements will be basic technologies to analyze microbial functions in soil environments.

8. Conclusions

The improved soil-DNA extraction method and PCR-DGGE analysis for soil metagenome enabled us to screen effective degraders for a chloroaromatic compound and detect DNA sequences derived from AOB community specific for long long-term fertilizer experiments. Based on these achievements, NIAES play a central role in a national project addressing development of analytical methods for soil biological properties using eDNA (eDNA project). In the project, we have been collecting soil samples under various types of agricultural management at different localities from the northern to the southern Japan, and analyzing the metagenome extracted from each soil sample by the standardized PCR-DGGE method. The obtained data have been used to construct a database (eDDASs) collecting the results of soil eDNA analysis accompanying the physical and chemical soil properties and information about crop productivity in each agricultural fields. We succeeded in detecting microbial gene expression in soil environment using a soil-RNA extracted from soil. The data will be used to analyze the relationships between the biological properties evaluated by eDNA or RNA analysis and the soil physicochemical properties as affected by cultivation practice.

Acknowledgement

Most of these works were supported by a Grant-in-aid eDNA (Soil eDNA) from the Ministry of Agriculture, Forestry, and Fisheries of Japan.

References

Amann R.I., W. Ludwig, and K.H. Schleifer. 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbial Review 59:143–169.

Chu, H., T. Fujii, S. Morimoto, X. Lin, K. Yagi and J. Zhang. 2007. Community structure of ammonia-oxidizing bacteria under long-term application of mineral fertilizer and organic manure in a sandy loam soil. Applied and Environmental Microbiology 73: 485-491.

Chu, H, X. Lin, T. Fujii, S. Morimoto, K. Yagi, J. Hu and J. Zhang. 2007. Soil microbial biomass, dehydrogenase activity, bacterial community structure in response to long-term fertilizer management, Soil Biology and Biochemistry 39:2971-2976.

Elsas J.D., R. Costa, J. Jansson, S. Sjoling, M. Bailey, R. Nalin, T.M. Vogel and L. Overbeek. 2008. The metagenomics of diseasesuppressive soils – experiences from the METACONTROL project. Trends in Biotechnology 26:591-601

Hoshino, Y.T. and N. Matsumoto. 2004. An improved DNA extraction methods using skimmed milk from soils that strongly adsorb DNA. Microbes and Environments 19: 13-19

Hoshino, Y.T. and N. Matsumoto. 2007. Changes in fungal community structure in bulk soil and spinach rhizosphere soil after chemical fumigation as revealed by 18S rDNA PCR-DGGE. Soil Science and Plant Nutrition 53:40-55.

Morimoto, S., K. Togami, N. Ogawa, A. Hasebe and T. Fujii. 2005. Analysis of a Bacterial Community in 3-Chlorobenzoate-Contaminated Soil by PCR-DGGE Targeting the 16S rRNA Gene and Benzoate 1,2-Dioxygenase Gene (*benA*). Microbes and Environments 20:151-159.

Morimoto, S., N. Ogawa, A. Hasebe and T. Fujii. 2008. Isolation of effective 3-chlorobenzoate-degraders in soil using community analyses by PCR-DGGE, Microbes and Environments 23: 285-292.

Muyzer, G., 1999. DGGE/TGGE a method for identifying genes from natural ecosystems. Current Opinion in Microbiology 2:317-322.

Muyzer, G., E.C. de Waal and A.G. Uitterlinden. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis and polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and Environmental Microbiology 59:695–700.

Okada, H. and H. Oba. 2008. Comparison of nematode community similarities assessed by polymerase chain reaction-denaturing gradient gel electrophoresis (PCR-DGGE) and by morphological identification. Nematology 10:689-700.

Vogel, T.M., P. Simonet, J.K. Jansson, P.R. Hirsch, J.M. Tiedje, J.D. Elsas, M.J. Bailey and R. Nalin. 2009. TerraGenome: a consortium for the sequencing of a soil metagenome. Nature Reviews Microbiology 7: 252

Wang, Y., J. Shimodaira, T. Miyasaka, S. Morimoto, T. Ohmori, N. Ogawa, M. Fukuda and T. Fujii. 2008. Detection of *bphAa* gene expression of *Rhodococcus* sp. strain RHA1 in soil using a new method for RNA preparation from soil. Bioscience, Biotechnology, and Biochemistry 72: 694-701.

Wang, Y., S. Morimoto, N. Ogawa, T. Oomori and T. Fujii. 2009. An improved method to extract RNA from soil with efficient removal of humic acids. Journal of Applied Microbiology 107:1168-1177.