Mini-review

Monitoring complex bacterial communities using culture-independent molecular techniques: application to soil environment

Lionel Ranjard, Franck Poly, Sylvie Nazaret*

Laboratoire d’écologie microbienne, UMR CNRS 5557, université Claude Bernard, Lyon I, 69622 Villeurbanne cedex, France

(Submitted 20 October 1999; accepted 20 December 1999)

Abstract — Over the last decade, important advances in molecular biology led to the development of culture-independent approaches to describing bacterial communities. These new strategies, based on the analysis of DNA directly extracted from environmental samples, circumvent the steps of isolation and culturing of bacteria, which are known for their selectivity leading to a non-representative view of the extent of bacterial diversity. This review provides an overview of the potentials and limitations of some molecular approaches currently used in microbial ecology. Examples of applications to the study of indigenous soil microbial community illustrate the feasibility and the power of such approaches. © 2000 Éditions scientifiques et médicales Elsevier SAS

bacterial community / soil / molecular biology / DNA / genetic diversity / genetic structure

1. Introduction

Bacteria are an important part of the soil microflora because of their abundance (up to 10^9 cells per gram of soil [3]), their species diversity (a minimum of 4,000–7,000 different bacterial genomes per gram of soil [54]) and the multiplicity of their metabolic activities. They play a key role in the biogeochemical cycles of the main elements (carbon, nitrogen, sulphur, etc.) and of trace elements (iron, nickel, mercury, etc.) and are therefore heavily implicated in energy and nutrient exchanges within the soil. They also have the potential to reflect the past history of a given environment. It is therefore essential to understand the interrelationships between bacteria and their environment by studying the structural and functional diversity of soil bacterial communities and how they respond to various natural or man-made disturbances.

The diversity of soil bacterial communities has been investigated for many years using methods based on isolating and culturing the micro-organisms. Such techniques are known for their selectivity and are not considered representative of the extent of the bacterial community. The proportion of cells which can be cultured is estimated to be 0.1% or at most 10% of the total population [1] and few data are available concerning how closely they reflect the actual composition of these communities. Recent advances in the field of molecular biology (extraction of nucleic acids, polymerase chain reaction (PCR) amplification, DNA cloning, DNA sequencing) have made it possible to develop techniques which no longer require the isolation and culture of bacteria and thus reduce the bias associated with it. These methods involve i) direct lysis of bacterial cells in soil, ii) the extraction of the nucleic acids from the matrix, and iii) the analysis of targeted
sequences or of the whole body of genetic information.

Two main types of molecular technique are available to study bacterial communities using DNA directly extracted from the soil (figure 1):

— molecular approaches which usually investigate parts of this information by focusing on genome sequences which are targeted and amplified by PCR that are called ‘partial community DNA analysis’;

— molecular approaches which try to investigate all the genetic information in the extracted DNA and that are called ‘whole community DNA analysis’.

The remainder of this mini-review will focus on various approaches suitable for describing the genetic diversity and structure of soil bacterial communities, together with the most important original information obtained and the main technical and conceptual limitations.

2. The ‘partial community DNA analysis’ approaches

These approaches consist in the analysis of PCR-amplified sequences. The most commonly used target sequences are the genes of the ribosomal operon, and particularly the rrs gene (16S rDNA) and the spacer between the rrs and rrl genes (intergenic spacer (IGS) 16S-23S). These methods include:

— PCR fragment cloning followed by restriction and/or sequencing analysis, which enable assessment of the diversity of the community in terms of the number of different species and, to a lesser extent, the relative abundance of these species;

— ‘genetic fingerprinting’, which provides a global picture of the genetic structure of the bacterial community.

2.1. PCR fragment cloning and characterization

This approach is used to investigate the diversity of bacterial communities by producing a library of clones from rrs sequences obtained by PCR from DNA extracts. Cloning is used to separate the sequences so that they can be characterized individually using PCR/RFLP (polymerase chain reaction/restriction fragment length polymorphism) and/or by sequencing. Sequencing allows a fine identification of uncultured bacteria as well as an estimation of their relatedness to known culturable species. The substantial richness found within the soil bacterial community and the number of clones per library (usually about 100 clones) preclude the application of diversity measurements in terms of species evenness. However, such a sampling size can be informative for estimating diversity, both richness and evenness, by considering higher level phylogenetic groups (i.e. subdivisions within proteobacteria, high and low G + C firmicutes, Holophaga/Acidobacterium group, etc.).

This approach was used to compare the genetic diversity between soils differing in their geographical location, agricultural practices and physico-chemical characteristics. Species common to these different environments were found and the diversity of some eubacterial and archaeal taxons was enlarged (for a review see [23], table I). Other studies have also used functional genes involved in the nitrogen cycle, heavy metal resistance or antibiotic metabolism, as molecular markers for assessing the distribution and the diversity of corresponding functional populations in soil communities (table I).

The main limitations and drawbacks of the PCR-cloning approach are essentially technical:

— it is time-consuming and cumbersome since it is necessary to sample a large number of clones in order to obtain a good diversity estimation of the amplified sequences;

— it requires very expensive equipment (sequencer);

— there exists bias due to the PCR (i.e. choice of primers, annealing temperature and numbers of cycles, inhibition of the enzyme by humic compounds, formation of chimaeric PCR products) or the cloning strategy used, both of which have a direct impact on the sampling of the cloned 16S rDNA sequences and their identification (for a review see [45, 61]). Consequently, the image of diversity of bacterial communities may not be totally exact. However, despite the
numerous biases that can occur in the PCR-cloning method, a recent study by Dunbar et al. [10] showed that 16S rDNA cloning coupled to RFLP analysis and partial sequencing was as
valid as plate cultivation for investigating diversity in soil environments.

2.2. Genetic fingerprint techniques

These techniques are also based on PCR amplification, but do not require a clone library. They are based on the principle of resolving the diversity of the amplified sequences simply by differential electrophoretic migration on agarose or polyacrylamide gels, which depend on their size (ARDRA, t-RFLP, RISA, RAPD) or sequence (DGGE, TGGE). The genetic fingerprints provide complex band profiles (i.e. a high number of different bands per profile) which yield a representative of the genetic structure of the community as a whole or of a section of it, defined by the selected primers. These profiles are not directly interpretable in terms of richness and evenness, since one band may originate from different species and one cell may be represented by several bands. This is particularly true with ribosomal sequences, since a single bacterial species may have several different rRNA sequences [8, 37] and different bacterial species may have closely related rDNA sequences [18]. Consequently, the genetic fingerprint approaches are especially suitable for comparing bacterial communities and various methods have been developed based on this principle. The profile data can then be analysed in terms of similarities and relationships can be represented by a dendrogram of similarities [49]. Multivariate analyses such as principal component analysis were also shown to be useful tools for easily visualizing and comparing changes in data tables derived from modifications in the fingerprint profile due to shifts in genetic structure [60]. Furthermore, this type of statistical analysis enabled correlation of these changes with various environmental characteristics or events [43, 44].

2.2.1. Electrophoretic separation using a denaturing gradient

This method is used to separate fragments of DNA which have the same size but different sequences. It was originally developed to detect specific mutations within the human genome, and has been adapted to analyse bacterial communities [35]. It involves the separation of amplicons on polyacrylamide gels containing a linear gradient of a DNA denaturing agent. This denaturing agent may be chemical (urea or formamide), in which case the method is known as DGGE (denaturing gradient gel electrophoresis) or it may be a physical factor such as the temperature, in which case the name TGGE (temperature gradient gel electrophoresis) is used [13, 35]. The differences between the sequences of the amplified samples influence their denaturation. Consequently, amplified samples with different sequences are separated during electrophoretic migration yielding complex profiles representing the diversity of the fragments amplified.

Compared to ARDRA, RISA or RAPD (see below), DGGE and TGGE require laborious technical optimization including calibration of the linear gradient of DNA denaturants (chemical or physical) and improvement of the PCR primers with the insertion of a GC clamp to obtain better electrophoretic separation of the fragments. Nevertheless, DGGE and TGGE of

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characterization of bacterial community, detection of major phylogenetic</td>
<td>[5, 10, 31, 33, 50, 55, 63]</td>
</tr>
<tr>
<td>groups and identification of genus and species</td>
<td></td>
</tr>
<tr>
<td>Diversity investigation of the most active bacterial populations</td>
<td>[13, 14]</td>
</tr>
<tr>
<td>Identification of novel clusters having no known cultured members</td>
<td>[24, 25, 27, 29, 30, 46, 57]</td>
</tr>
<tr>
<td>Influence of environmental disturbance on soil genetic diversity</td>
<td>[4, 31, 33, 39, 51]</td>
</tr>
<tr>
<td>Analysis of the diversity of particular functional groups in soil (i.e.</td>
<td>[7, 17, 48, 51, 59]</td>
</tr>
<tr>
<td>ammonium oxidizers, methanogenic bacteria, nitrogen-fixing bacteria, heavy</td>
<td></td>
</tr>
<tr>
<td>metal-resistant bacteria, antibiotic-producing bacteria)</td>
<td></td>
</tr>
</tbody>
</table>
16S rDNA are commonly used to compare the genetic structure of bacterial communities from different soil types or submitted to perturbation (Table II). In a restricted case, where the bacterial diversity is low, DGGE profiles can be used to define the number of different operational taxonomic units (OTUs) present, enabling determination of the complexity of the community. DGGE or TGGE of PCR products obtained after amplification of ribosomal copy DNA of 16S rRNA (rcDNA) by reverse transcriptase (RT)-PCR would indicate the structure or the complexity of the active bacterial populations in a given soil [13, 14]. Furthermore, individual bands may be excised, reamplified, and sequenced, or challenged with a range of probes for the taxonomic identification of specific populations within a community.

The specific limitation of this approach lies in the size of the amplified sequences under investigation since they can no longer be efficiently separated beyond 500 base pairs. This limit restricts the precision of the subsequent identification of the bands based on sequence comparison and the choice of the regions of the rrs gene to define the PCR primers or probes for further identification. In addition, the resolution of the profiles obtained, in terms of band numbers, is not always sufficient to illustrate the considerable bacterial diversity in indigenous soil communities, and some studies have shown that fragments of different sequences might migrate at the same position [56].

### 2.2.2. ARDRA (amplified ribosomal DNA restriction analysis) and T-RFLP (terminal-RFLP).

ARDRA involves digesting the 16S rDNA sequences obtained by PCR using restriction enzymes and separating the restricted fragments on agarose or polyacrylamide gels. The main advantage of this method is its convenience, because it does not require any particular development. This technique has been used to demonstrate changes in the genetic structure of bacterial communities following changes in environmental conditions or exposure to exogenous toxic compounds (Table III). Like DGGE, it can be used for the taxonomic identification of bacterial populations by hybridization of the profile using specific probes.

The main limitation of this method lies in the choice of restriction enzymes, which is crucial for obtaining optimum resolution. Preliminary tests for the enzyme choice must be performed to ensure the highest resolution in detecting changes in communities. In this method, the complexity of the profile is great and no information can be deduced in detecting and quantifying a specific ribosomal pattern. An alternative has recently been found using fluorochrome-labelled primers. This makes it possible to detect only the terminal fragments and it is therefore known as T-RFLP or ‘terminal-RFLP’. The profiles obtained are simpler in terms of the numbers of bands. A data bank has been produced of the sizes of the 16S rDNA terminal restriction fragment for a given

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study of the genetic structure of the total bacterial community</td>
<td>[13, 14, 40]</td>
</tr>
<tr>
<td>Study of the genetic structure of a particular phylogenetic group or functional group</td>
<td>[20, 26, 52]</td>
</tr>
<tr>
<td>Spatial distribution of bacterial populations</td>
<td>[13]</td>
</tr>
<tr>
<td>Impact of environmental disturbance on soil community</td>
<td>[11, 12, 41]</td>
</tr>
</tbody>
</table>

### Table II. Overview of the applications of DGGE and TGGE in soil bacterial community studies.

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study of the genetic structure of the total bacterial community</td>
<td>[13, 14, 40]</td>
</tr>
<tr>
<td>Study of the genetic structure of a particular phylogenetic group or functional group</td>
<td>[20, 26, 52]</td>
</tr>
<tr>
<td>Spatial distribution of bacterial populations</td>
<td>[13]</td>
</tr>
<tr>
<td>Impact of environmental disturbance on soil community</td>
<td>[11, 12, 41]</td>
</tr>
</tbody>
</table>

### Table III. Overview of the applications of ARDRA in soil bacterial community studies.

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis of genetic similarities between different soils</td>
<td>[58]</td>
</tr>
<tr>
<td>Study of the genetic structure of a particular phylogenetic group or functional group</td>
<td>[6, 42, 58]</td>
</tr>
<tr>
<td>Impact of environmental disturbance on soil community</td>
<td>[9, 49, 53]</td>
</tr>
</tbody>
</table>
enzyme on a large set of bacterial species and this makes it possible to identify the bands present in a community by a profile to profile comparison [32]. Running such profiles on acrylamide gel with laser detection (sequencing system) increases the resolution and sensitivity power and enables the interpretation of peak height and area which can be further interpreted in terms of number and abundance of OTUs. Such an approach has also been successfully adapted to functional genes such as the mercury resistance determinants (mer genes [6]) to assess the complexity and the distribution of functional groups in soil matrix.

2.2.3. RISA (ribosomal intergenic spacer analysis)

RISA involves the analysis of the length polymorphism of the spacer between the rrs and rrl genes (IGS) whose size varies from 50 bp to more than 1.5 kb depending on the species. The primers target regions within the adjacent genes and can be defined so that part of the rrs gene is co-amplified. The subsequent sequencing of certain bands can then allow taxonomic identification of specific populations within a community. The various amplified IGSs are directly separated on polyacrylamide gels on the basis of size [4]. As mentioned for T-RFLP, fluorescent primers can be used to further run RISA profiles on sequencing gel systems and then to increase the resolution power. The high size variability of this spacer can then allow detection of slight changes in the genetic structure of a community, and the high sequence variability [36] may provide a finer taxonomic identification of bands than can be obtained using 16S rDNA sequences (as in DGGE or ARDRA). However, to date, the IGS sequence databank is not large enough to permit such efficient identification (in September 1999, 26 191 16S rDNA sequences and 561 IGS sequences were available in the data bank).

Although RISA is easy to use and does not require any particular development, it has not been used to any significant extent at the community level (table IV). Several works have demonstrated the reproducibility and the capacity of this approach to detect modifications in community structure between different soil types and even between closely related communities such as those associated with various microenvironments of a soil [43, 44].

2.2.4. RAPD (random amplified polymorphic DNA)

The RAPD technique uses short random primers (about 10 bp) which anneal at different places on the genomic DNA, generating PCR products of various lengths further resolved on agarose or acrylamide gel. This technique was demonstrated to be rapid and sensitive for revealing differences between similar complex prokaryotic genomes [19]. Though some works have evidenced the lack of reproducibility of this technique owing to its great sensitivity to PCR artefacts [19], recent studies conducted at the community level demonstrated its reproducibility in comparing the genetic structure of complex bacterial communities [60]. The main limitation is that it cannot provide phylogenetic information about the bacterial composition of the community.

The only study conducted on the soil bacterial community using this method was developed by Xia et al. [62]. The authors assessed the molecular genetic response of different soil communities to the application of 2,4-D and demonstrated the absence of a pesticide effect.

2.2.5. Limitations of genetic fingerprint techniques

Initially, genetic fingerprints were developed to discriminate and to identify pure bacterial strains. When cultured populations were found to be poorly representative of the community, in qualitative and quantitative terms [1], these

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comparative study of the genetic structure from various soil microenvironments</td>
<td>[43]</td>
</tr>
<tr>
<td>Impact of environmental perturbation on genetic structure of soil community</td>
<td>[4, 44, 47]</td>
</tr>
</tbody>
</table>
methods were then adapted for whole communities which are made up of a complex mixture of target sequences. Compared to the PCR-cloning/RFLP/sequencing strategies, genetic fingerprint methods are easier to perform and less time-consuming allowing any investigations which involve a large number of differently treated samples. However, PCR is also involved in these approaches and may lead to a distorted view of bacterial community structure owing to the previously mentioned biases (formation of chimaeric sequences, differential PCR, etc.; see part 2.1. and [61]). Furthermore, technical problems inherent in electrophoresis still limit these approaches. In some cases, DGGE has been reported not to be able to separate amplicons harbouring different sequences [56]. Similarly, resolution of different-size bands in ARDRA and RISA profiles may not always be possible especially for high molecular weight fragments. Another limitation is the complexity of such genetic profiles, which leads to difficulties in data analysis since they exhibit several hundred bands of different intensities.

To date, only a few works have compared the efficiency of different fingerprinting approaches for analysing bacterial community structure. The works of Ranjard et al. [43] demonstrated that ARDRA and RISA techniques gave the same clustering for the bacterial communities associated with the various microenvironments of a soil. Recently, Moeseneder et al. [34], by comparing the complexity of different bacterial communities in terms of OTUs with T-RFLP and DGGE of 16S rDNA, demonstrated the higher sensitivity of the former. To date, it is difficult to conclude as to the better resolution or sensitivity of a fingerprinting technique since it seems to be dependent on the type of bacterial community studied and the type of molecular markers (taxonomic, functional or random). The choice of a technique for a given community had to be deduced from a testing procedure of several approaches.

3. The ‘whole genomic community DNA analysis’ approaches

3.1. Total genomic cross-DNA hybridization

The technique involves DNA extraction and purification from environmental samples and cross-hybridization between DNA from one sample with that from the other [28]. The DNA from one sample is radioactively labelled and is used as a template. The extent to which the radiolabelled probe anneals to the filter-bound target DNA reflects the similarity of probe and target and consequently the extent to which the population structure of bacterial communities is similar. Applications to the communities of soil differing in texture, land use and location and to the response of communities exposed to pesticides are reported (table V).

This technique might not be appropriate for detecting differences between communities with relatively high DNA sequence similarity, since a high degree of variability of cross-hybridization has sometimes been observed [62].

3.2. Thermal denaturation and reassociation of whole extracted DNA

This technique involves the heat denaturation of DNA followed by reassociation of the homologous single strands. The proportion of DNA renatured is generally expressed as the function of the product (Cot) of the concentration of nucleotides (moles/L) and the reaction time. Under specified conditions, the Cot1/2 value, corresponding to one half of the reassociation reaction, is directly proportional to the size of the genome or the complexity of the

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis of genetic similarities between different soils</td>
<td>[15, 62]</td>
</tr>
<tr>
<td>Impact of chemical pollutants on soil genetic diversity</td>
<td>[62]</td>
</tr>
</tbody>
</table>

Table V. Overview of the applications of cross-DNA hybridization in soil bacterial community studies.
DNA. The Cot1/2 value is used as an index of the genetic diversity in a bacterial community based on DNA extracted from the soil [54]. Though this technique makes it possible to estimate the genetic diversity in terms of the number of different genomes (richness), it does not provide any details about the relative abundance of individual genomes (evenness).

Using this method, Torsvik and collaborators have estimated that the number of different bacterial genomes present per gram of soil ranged from 350 to 10 000 according to soil types [40, 54]. They have also suggested that most of this diversity is present in the fraction of bacteria which cannot be cultured. This technique can also be used to evaluate the impact of exogenous compounds on a soil community. Atlas et al. [2] demonstrated that soil amendment with a pesticide (2,4,5-T) decreases the genetic diversity of the bacterial community estimated by thermal denaturation and reassociation kinetics of the extract DNA. The overview of the applications using this approach to study soil bacterial community are listed in table VI.

3.3. Extracted whole DNA fractionation by base composition using density gradient

This method is based on the fact that prokaryotic DNAs vary in G + C content from ca. 24 to 76% [21]. It uses an intercalating agent, bisbenzimidazole, which is preferentially bound to A + T base pairs, thus amplifying the differences in the gravity of the DNA according to the specific G + C content. The extracted DNA is centrifuged through a caesium chloride density gradient in the presence of bisbenzimidazole. This yields a DNA community profile of the distribution of %G + C values providing a molecular fingerprint of the overall community structure. Furthermore, since %G + C of chromosomal DNA is generally characteristic for phylogenetic groups of bacteria at the genus level, the relative abundance of DNA at each %G + C is an indication of the relative abundance of phylogenetic groups in the bacterial community. Holben and Harris [21] showed that the majority of DNA extracted from cultivated soil corresponds to the %G + C of range 55–73 which includes bacterial genera known to be abundant in soil communities under field conditions, e.g. *Agrobacterium* (G + C = 59–64%), *Alcaligenes* (G + C = 55–61%), *Arthrobacter* (G + C = 63–69%), and *Pseudomonas* (G + C = 58–66%).

This method provides a low level of taxon resolution, since several taxonomic groups can share the same G + C range. Nevertheless, it led to detection of modifications in the composition of the soil community due to the influence of carbon amendment or vegetation cover [21, 39]. An overview of applications using this approach for the study of soil bacterial communities is given in table VII. This method can be extended by using analyses providing a more accurate identification of some bacterial populations by hybridization with specific probes. In situations where the diversity of the community is considered too complex to be tackled as a whole, this method can be used to select a part of the community based on its %G + C. This sample can then be subjected to more detailed molecular analysis (PCR-cloning-sequencing) to identify its constituent populations [38, 39].

The three previously described methods exhibit common disadvantages:

— the need for a large amount of DNA (≥ 50 μg);

<table>
<thead>
<tr>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation of species numbers in a complex community</td>
<td>[54]</td>
</tr>
<tr>
<td>Analysis of genetic similarities between different soils</td>
<td>[40]</td>
</tr>
<tr>
<td>Comparative study of genetic diversity between culturable and nonculturable bacteria</td>
<td>[40, 54]</td>
</tr>
<tr>
<td>Impact of chemical pollutants on soil genetic diversity</td>
<td>[2, 41]</td>
</tr>
<tr>
<td>Impact of root exudates on soil genetic diversity</td>
<td>[16]</td>
</tr>
</tbody>
</table>

**Table VI. Overview of the applications of reassociation kinetics in soil bacterial community studies.**
— the need for high DNA quality in terms of purity and size integrity;
— the co-extraction of DNA from live bacterial cells and of contaminant DNA (i.e. extracellular DNA, DNA from dead cells, DNA originating from fungi or other soil eukaryotes) that can bias the analysis;
— expensive equipment (ultracentrifuge for DNA fractionation technique).

4. Concluding remarks

These ‘molecular ecology’ methods have enabled us to extend our understanding of microbial ecology and will continue to do so. However, due to the various biases or limitations for each of them, they are not a substitute for more conventional methods (cultured population studies, measurement of activity, etc.) but they must be viewed as complementary methods for use in investigating the ecology of bacteria in their natural habitats. Along this line, future prospects in microbial ecology must be of a polyphasic nature, combining and selecting a range of molecular biological and microbiological techniques in order to understand the relationship between microorganisms and their environment.

These methods have primarily been used to assess the composition of microbial communities (identification of genus, species, or phylogenetic groups) and to monitor, over space and time, changes due to environmental disturbances (figure 1) using as a target ribosomal genes and/or function-specific genes. Data focussing on the activity of microbial populations and on gene expression and regulation in situ are still scarce, especially in a soil environment. Development of tools for addressing these points will be an important task in the coming years.

Acknowledgments

The authors thanks Benoit Cournoyer, Philippe Normand, and Jacques Balandreau for their comments on the manuscript.

References

Ranjard et al.


