Chapter from the book *Gel Electrophoresis - Principles and Basics*
Downloaded from: http://www.intechopen.com/books/gel-electrophoresis-principles-and-basics

Interested in publishing with InTechOpen?
Contact us at book.department@intechopen.com
Electrophoretic Techniques in Microbial Ecology

Elena González-Toril, David Lara-Astiaso, Ricardo Amils, and Angeles Aguilera

1 Centro de Astrobiología (INTA-CSIC), Spain
2 Centro de Biología Molecular (UAM-CSIC), Spain

1. Introduction

Classical microbial ecology analysis is limited by the unavoidable need for isolation of the microorganisms prior to their characterization. Although it is obvious that isolation of microorganisms is indispensable for their full characterization, it is now well recognised among microbiologists that only a small fraction of all bacteria have been isolated and characterised (Ward et al., 1992). Comparison of the percentage of culturable bacteria with total cell counts from different habitats showed enormous discrepancies (summarised by Amann et al., 1995). The introduction of molecular biology methods (such as fluorescent in situ hybridization, denaturing gradient gel electrophoresis or cloning) has enabled a significant advance in microbial ecology (Amann et al., 1995), especially in the study of extreme environments such as acidic habitats, in which conventional methods are severely limited, and some may even lead to equivocal conclusions, with occasionally grievous economic consequences.

In recent years the use of ribosomal RNAs (rRNAs) and their genes have produced an authentic revolution in microbial ecology (Akkermans et al., 1994). The sequencing of these genes has allowed a whole range of microorganisms, mainly prokaryotes, to be studied without running into selective enrichment and isolation problems. Most molecular ecology techniques are based on these genes (Akkermans et al., 1994; Amann et al., 1995).

Among various 16S rRNA gene-based genotyping approaches, terminal restriction fragment length polymorphism (T-RFLP) (Liu et al., 1997) and denaturing gradient gel electrophoresis (DGGE) (Muyzer et al., 1993) are the most commonly used techniques to either compare microbial communities of various samples or to capture a single microbial community structure of an individual sample. Pulsed Field Gel Electrophoresis (PFGE) is not a technique habitual in microbial ecology for this reason we decided not include it in this chapter. However, this electrophoretic technique is very usefully for a genomic differentiation between strains isolated from natural environments (García-Moyano et al., 2008). With these techniques, DNA fragments of similar length but with different base-pair sequences can be electrophoretically resolved. In addition, functional genomic approaches, such as proteomics, greatly enhance the value of genome sequences by providing a global level
assessment of which genes are expressed, when genes are expressed and at what cellular levels gene products are synthesized. To purify, reduce the complexity or visualize individual proteins, cellular extracts or sub-cellular fractions can be separated using gel-electrophoresis methods. Separation by molecular weight (1 dimensional gel electrophoresis; 1DE) can help to remove compounds that inhibit peptide digestion or Mass Spectrophotometry. Two dimensional gel electrophoresis (2DE) separates proteins based on their isoelectric point (pI) through Isoelectrofocusing (IEF), and molecular weight through polyacrylamide gel electrophoresis (PAGE). The technique can further help to reduce sample complexity, and allows snapshot (one sample) and differential (comparison of more than one sample) protein profiles to be generated.

2. Denaturing gradient gel electrophoresis (DGGE)

DGGE is a fingerprinting technique, whereby DNA fragments of similar length but with different base-pair sequences can be resolved (Muyzer et al., 1996). Separation by DGGE is based on the electrophoretic mobility of a partially melted double stranded DNA molecule in polyacrylamide gels containing a linear gradient of DNA denaturants (urea and formamide), with less mobility than the completely helical form of the molecule. The DNA molecules run through the gel until denaturation occurs as a consequence of exposure to the denaturing gradient. Denaturation of double stranded DNA the denaturation of the double strand occurs in a sharpe manner not in a zipper like mode, and is dependent on its melting temperature ($T_m$), and the $T_m$ is a consequence of the sequence, so the retention of DNA molecules in the gel is determined by its sequence. Thus DNA fragments of the same length and with different sequences can be separated (Muyzer et al., 1996). The technique was designed to locate single base mutations in genes, and it has been very useful in molecular pathology (Myers et al., 1987). In 1993 Muyzer et al. applied this technique for the first time to the study of an ecosystem (Muyzer et al., 1993).

Currently, the most useful gene for prokaryotic identification is the 16S rRNA gene. From an environmental sample we can carry out a DNA and/or RNA extraction followed by amplification of the 16S rRNA genes through the Polymerase Chain Reaction (PCR), using universal primers for Bacteria or Archaea domains. DGGE is useful to evaluate the number of 16S rRNA genes corresponding to different microorganisms present in the sample. Spatial or temporal variations of microbial populations can be studied using these fingerprints. Furthermore, each band corresponding to a different class of 16S rRNA genes can be excised and the corresponding DNA sequenced. Sequences can be compared with a sequence database and the closest microorganisms identified. In addition, a preliminary phylogenetic study with the retrieved sequences could be performed (for instance using parsimony), but we must bear in mind that the sequences are too short for a real phylogenetic study. The newly generated sequences can also be used to design specific hybridization probes for the microorganisms present in the sample. Using this approach, around 50% of the sequences can be detected in DNA fragments of up to 500 bp. This percentage can be increased to nearly 100% by the attachment of GC-rich sequences to the DNA fragment, which will then act as a high temperature melting domain (Sheffield et al., 1989).

This genetic diversity also means that it is very unlikely that the whole range of diversity can be detected in a single sample, especially where communities are complex. Attempts to
Electrophoretic Techniques in Microbial Ecology

estimate the total bacterial community diversity have used partial analysis of the total community (that is clone library screening) combined with theoretical models. This approach, however, does not reveal the identity of the less abundant components of the assemblage. Increasing the number of clones per clone library has been successful in detecting novel bacterial clades (for example, Chouari et al., 2005) or comparing different environments in terms of bacterial community composition (for example, Rappe et al., 2000; Zaballos et al., 2006). However, despite the decreasing costs for nucleotide sequencing, the success of this approach is still limited because of the huge scale of bacterial diversity — perhaps as many as $2 \times 10^9$ different species just in the oceans (Curtis et al., 2002). The recent introduction of pyrosequencing of 16S rRNA tags (Sogin et al., 2006) may represent an alternative because of the much lower cost per sequence; but pyrosequencing still does not allow analyses and comparison of the bacterial diversity in different environments on a routine basis.

Alternative approaches, such as denaturing gradient gel electrophoresis (DGGE), are routinely used to determine diversity because they avoid large scale sequencing efforts. However, these are also likely to detect only a small subfraction of the total diversity. The use of bacterial or archaeal PCR primers is likely to miss minor fractions of the microbial community because most of the PCR product will be composed of the more abundant species. Faint DNA bands on DGGE gels are unlikely to be detected or their identity determined. To overcome this limitation and to detect less abundant sequence clones, Holben et al. (2004) fractionated 16S rRNA gene sequences from a microbial community according to their G+C content before DGGE analysis. However, with the possible exception of the high G+C-containing Actinobacteria, this method has limited application and does not separate bacteria by phylogeny. Combining bromodeoxyuridine immunocapture and DGGE has been proposed to separate the DNA of the actively growing bacteria from the rest of the environmental DNA (Hamasaki et al., 2007). However, this results in the analysis of subgroups of bacteria that are not defined on phylogenetic criteria, and does not allow screening of the whole range of microbial bacterial diversity.

DGGE of PCR amplified 16S rDNA fragments have been used to profile community complexity of microbial mats and bacterial biofilms of a wide variety of ecosystems. DGGE analysis of PCR amplified rDNA fragments has been used to provide information on the microbial diversity of microbial communities associated to hydrothermal vents (Li et al., 2008; Maugeri et al., 2009), soil (Vaario et al., 2011) or extreme acidic environments (González-Toril et al., 2003) (Fig.1). PCR-DGGE has been applied to profile the distribution of microbial populations inhabiting regions with different temperatures in hot spring cyanobacterial communities (Aditiawati et al., 2009; Teske et al., 2009), to identify bacteria in a biodegraded wall painting (Rolleke et al., 1996) or historical limestone buildings (Ettenauer et al., 2011).

The particular advantage of DGGE is that the bands of interest can be excised from the gel for subsequent sequencing. This band-retrieving approach on one dimensional gel electrophoresis is only applicable to low bacterial diversity, since co-migration of DNA fragments on DGGE gels can be a problem when fingerprinting complex microbial communities (Sekiguchi et al., 2001). Furthermore, DGGE gels have low comparability due to their variability at different runs.
Microbial ecological studies often require the sampling at different points over a long periods of time. As mentioned in the introduction, cloning techniques are not suited for the analysis of many different samples. However, by using DGGE, many samples taken at different time intervals during the study can be simultaneously analysed. This makes the technique a powerful tool for monitoring community behaviour after environmental changes (Donner et al., 1996; Ferris et al., 1997).

Fig. 1. DGGE fingerprints of 16S rRNA obtained by using universal primers for members of the domain Bacteria using samples from an extreme acidic environment, Río Tinto (SW, Spain) (González-Toril et al., 2003). Lane numbers correspond to different sampling sites. Arrows label the different bands posteriorly sequenced.

Although PCR and DGGE are a widely used for describing microbial community structures and diversity based on extracted DNA, there are drawbacks with this combination of methods, both in the DNA extraction and in PCR-DGGE step, due to for example biases in primer annealing, base pair mismatches, and limitations in DGGE-resolution (Sekiguchi et al., 2001). Furthermore, these analyses only provide presence/absence data and not abundance of certain species, mainly due to the “qualitative nature” of PCR. Basically, all physical, chemical and biological steps in the diversity analyses involve risks of being biased (Nubel et al., 1999). However, fingerprinting techniques such as PCR-DGGE/TRFLP and others, are necessary when high number of samples are needed to asses ecological impact. Truly quantitative information using molecular methods can only be obtained if cell lyses and extraction efficiency, as well as biases in the PCR step are under experimental control. This is not feasible in analyses of environmental samples containing an unknown amount of cells, a wide range of cell types, inhibitors as well as DNA from non-target organisms. Thereby, presence and intensities of bands on a DGGE-gel cannot be considered to represent true changes in the microbial community. However, the use of internal standards in the DNA extraction and PCR/DGGE-steps ensures that at least introduced experimental
variations are controlled, thus making it possible to describe relative changes in abundance and diversity between samples that originate from the same type of environment (Petersen et al., 2005).

Another limitation of the DGGE is the separation of only relatively small fragments, up to 500 basepairs (Myers et al., 1987). This limits the amount of sequence information for phylogenetic inferences as well as for probe design. Besides, it has been demonstrated that it is not always possible to separate DNA fragments which have a certain amount of sequence variation. Related to the problem of resolution is the maximum number of different DNA fragments which can be separated by DGGE or Temperature Gradient Gel Electrophoresis (TGGE), in which there is also a temperature gradient across the gel. For instance, by using DNA/DNA reannealing experiments Torsvik et al. (1990) found that there might be as many as $10^4$ different genomes present in soil samples. It will be obvious to the reader that DGGE or TGGE cannot separate all of the 16S rDNA fragments obtained from such a variety of microorganisms. In general, these electrophoretic techniques will only display the rDNA fragments obtained from the predominant species present in the community (Muyzer & Smalla, 1998).

3. Terminal restriction fragment length polymorphism (T-RFLP)

Among various 16S rRNA gene-based genotyping approaches, T-RFLP and DGGE are the most commonly used techniques to either compare microbial communities of various samples or to capture a single microbial community structure of an individual sample. T-RFLP and DGGE differ mainly in comparability between runs, detection sensitivity, quantification capability and accuracy (Nocker et al., 2007). T-RFLP is a high-throughput, sensitive and reproducible method that can be used to carry out both qualitative and quantitative analyses of a particular gene in a microbial community.

This technique is based on digesting PCR mixture of a single gene by restriction enzymes and detecting the size of the terminal fragment by DNA sequencer. T-RFLP technique is used also to amplify small subunit (16S or 18S) rRNA genes from total community DNA using the polymerase chain reaction (PCR) wherein one or both of the primers used are labelled with a fluorescent dye. The resulting mixture of rRNA gene amplicons is then digested with one or more restriction enzymes that have four base-pair recognition sites. The mixture of restricted PCR products is physically separated using acrylamide sequencing gels or sequencing capillaries and the sizes and relative abundances of the fluorescently labeled TRFs are determined using an automated DNA sequencer. Since differences in the sizes of T-RFs reflect differences in the sequences of 16S rRNA genes (i.e., sequence polymorphisms), phylogenetically distinct populations of organisms can be resolved. Thus, the pattern of T-RFs is a composite of DNA fragments with unique lengths that reflects the composition of the numerically dominant populations in the community. Ideally, primers chosen for T RFLP analysis should be specific to the targeted taxonomic group yet sufficiently general so that they can amplify all bacterial populations that are of interest. There are no known primers that satisfy both of these criteria (Ursel et al., 2008). Therefore, cloning each size selected T-RF is required in order to infer the phylogenetic identities of predominant T-RFs with high confidence (Mengoni et al., 2002).
Although T-RFLP allows for highly sensitive detection of labeled DNA fragments and is compatible with high-throughput analyses, one drawback is the need for restriction digestion. Incomplete or nonspecific restriction leads to an overestimation of diversity by increasing the number of fragments. However, restriction efficiency can be tested by including the amplified product from a well-characterized individual sequence in the restriction step (Nocker et al., 2007). This control PCR product should contain a different fluorescent label from the PCR products being analyzed. The presence of more than one control signal indicates partial digestion.

The discrimination of bacterial populations by T-RFLP analysis relies on detecting 16S rRNA gene sequence polymorphisms using restriction enzymes. Typically, enzymes that have four base-pair recognition sites are used due to the higher frequency of these recognition sites. It has been shown by several groups that the use of more than one restriction enzyme facilitates the resolution of bacterial populations (Liu et al., 1997). This is due to the fact that different bacterial populations can share the same terminal restriction fragment length for a particular primer–enzyme combination but not others (Marsh et al., 2000). The ability of different restriction enzymes to resolve unique sequences has been examined in studies of gene sequence databases, communities with different richness, and iterative random sampling from a derived database of T-RFs (Engebretson & Moyer 2003). For communities with more than 50 operational taxonomic units (OTUs), none of the restriction enzymes resolved more than 70% of the total OTUs. Then T-RFLP can most efficiently be used for communities with low or intermediate richness.

Species identification by comparison of complex communities with internal and external laboratory databases is computationally challenging, but has become more straightforward through implementation of Web-based automated assignment tools (Kent et al., 2003). A prerequisite, however, is the exact determination of fragment lengths. Discrepancies in fragment length between the experimental and predicted (in silico) value for a known sequence are often observed. This problem is addressed by using tolerance ranges for length assignment (e.g., T1 or 2 bases) to allow for matching with database entries. This fact, however, increases the numbers of species associated with each fragment and complicates community predictions given the large number of T-RFs in complex samples.

Despite its limitations, T-RFLP has become a valuable method for rapidly comparing the relationships between bacterial communities in environmental samples and temporal changes. It is a valuable method for comparison of complex communities when high throughput and high sensitivity are required without the need for direct sequence information.

4. Proteomic analysis by using two dimensional gel electrophoresis (2DE)

Proteomics is the term used to describe studies that examine the global protein complement of an organism, tissue or community. The proteome consists of all proteins expressed by an organism under a given set of conditions and therefore represents the functional complement of the genome (Goodlett & Yi, 2002). The proteome represents the product of global gene expression (transcription plus translation), protein stability, protein processing and turnover. Proteomics therefore extends beyond genomic analyses, which only describe the theoretical capability of an organism or community, by providing a direct measure of which proteins are synthesized, when they are synthesized and what their cellular (or extracellular) abundance is (Pandey & Mann, 2000).
Increasingly, proteomics has emerged as a promising technique to characterize microbial activities at the molecular level. Proteomics started to develop in the 1970s when protein profiles of single organisms were analyzed by 2-DE. At that time protein identification was, if at all possible, time consuming and cost-intensive due to a lack of genomic sequence information and advanced protein sequence analyses. Since the 1990s proteomics has become much more widespread, feasible, and reliable thanks to three technical revolutions: (i) the enormous increase of genomic and metagenomic data provides a solid basis for protein identification; (ii) tremendous progress in sensitivity and accuracy of mass spectrometers enables a correct, high-throughput protein identification, relative and absolute quantification of proteins, and the determination of post-translational modifications; and (iii) formidable improvements in computing power and bioinformatics allow processing and evaluation of substantial datasets.

Protein samples derived from natural environments do not lend themselves to direct Mass Spectrophotometry analysis; rather, sample complexity has to be reduced first by gel-based or chromatographic techniques. This can be accomplished either on the protein level or on the peptide level after proteolytic degradation of sample proteins. For many years 2-D PAGE was regarded as the “gold standard” of proteomics research (Shneider & Riedel, 2010). With this method proteins are first separated along a pH gradient by IEF, followed by a second separation according to mass on SDS-PAGE gels. In this way, over a thousand proteins can be resolved on a single gel as discrete spots. Staining the gels (e.g. with silver, Coomassie blue, or fluorescent dyes) allows the relative determination of protein abundances based on protein spot size and intensity. Protein spots can be subsequently excised and digested ingel (most commonly with trypsin, see below) prior to mass spectrometric analysis. A significant improvement of this technology was introduced in the late 1990s, when it became possible to label different samples with fluorescent dyes and pool these samples before PAGE Differential Gel Electrophoresis (DIGE), thereby reducing gel-to-gel variations. This method is commonly used in combination with 2-D PAGE (2-D DIGE) (Fig. 2).

Despite having been frequently used in various environmental studies (see following paragraphs), 2-D PAGE suffers from several weaknesses. Most notably, proteins with extreme molecular masses and/or isoelectric points as well as membrane proteins are difficult to analyze, and co-migration of proteins and protein isoforms hampers accurate identification and quantification. The method is also labor-intensive and consequently hardly automatable and not suited for high-throughput analyses. In the last decade, one or multi-dimensional Liquid Chromatography (LC) coupled to Mass Spectrometry (MS) has emerged as a promising alternative to 2-D PAGE (reviewed in Lane, 2005). An experimental strategy that has proven extremely useful for the analysis of membrane proteins or highly polluted samples (where contaminants might interfere with trypsin digestion) is the separation of proteins by 1-D PAGE, followed by in-gel digestion of excised protein bands and separation of the resulting peptides by Reversed Phase Chromatography (RP-HPLC).

Extending beyond laboratory-based manipulation of axenic cultures, environmental proteomics (metaproteomics) provides the means to assess proteins that are synthesized by microbial communities. Linking metaproteomic data to environmental genomics (metagenomics) and geo-physico-chemical data provides a powerful means of inferring the roles of indigenous microbial communities in whole ecosystem function (Schneider &
Riedel, 2010). To purify, reduce the complexity or visualize individual proteins, cellular extracts or sub-cellular fractions can be separated using gel-based or gel-free methods before MS. Separation by molecular weight only (1 dimensional gel electrophoresis; 1DE) can help to remove compounds that inhibit peptide digestion or MS. Two dimensional gel electrophoresis (2DE) separates proteins based on their isoelectric point (pl) through IEF, and molecular weight through polyacrylamide gel electrophoresis (PAGE). The technique can further help to reduce sample complexity, and allows snapshot (one sample) and differential (comparison of more than one sample) protein profiles to be generated (Gygi et al., 2000). Following electrophoresis, individual proteins or groups of proteins are visualized by staining techniques, of which several are available for use. The staining techniques have different ranges and limits of detection and the choice of specific techniques should be considered based on requirements. Coomassie staining is inexpensive, easy to use, reliable and compatible with MS. However, it has low sensitivity (> 1 mg to ~100 ng). Silver staining can detect proteins at lower abundance within a narrow range of concentrations (~5 to ~80 ng), uses complicated methodology and can produce problems with MS. More recently several fluorescent stains have been developed, which are relatively simple to implement and have ranges of detection limits from ~1 ng or less to > 1 µg. However, these stains use relatively expensive reagents and require specialized equipment for visualization and spot excision. Following staining and visualization, bands containing multiple (1DE) or individual (2DE) proteins are excised and digested in-gel, and peptides subjected to RP-LC separation before MS analysis (Ting et al., 2010). When developing proteomic methods it is wise to test whether gel-based separation tends to help (e.g. removes contaminants thereby improving the quality of mass spectra) or hinders (e.g. reduces protein yield without improving the quality of mass spectra) protein identification and coverage.

Fig. 2. 2-DE preparative gels obtained for Chlamydomonas acidophila (Cid et al., 2010). The spots resolved by 2-DE from preparative gels were stained with MALDI-MS-compatible silver reagent for peptide mass fingerprinting analysis. Numbered spots marked with circle corresponded to common proteins in each pair compared and identified by MALDI-TOF MS. A)- gel obtained with cells growing under BG11/f2 artificial media at pH2. B)- gel obtained with cells growing under natural heavy metal-rich water NW/f2 at pH2.
5. Conclusion

When viewed in relation to its enormous potential, the actual output of environmental techniques appears so far to be disappointingly limited. Present studies have mainly focused on microbial communities with a relatively low diversity or dominated by a particular phylogenetic group. The main obstacles toward a comprehensive coverage seem to be (i) the irregular species distribution within environmental samples, (ii) the wide range of gene expression levels within microbial cells, and (iii) the enormous genetic heterogeneity within microbial populations. It is encouraging to note, however, that constantly improving extraction methods alongside advances in technology and a steadily growing pool of bioinformatics data might soon help to overcome the current challenges and limitations of environmental research.

6. Acknowledgment

This work has been supported by the Spanish Science and Innovation Grants CGL2008-02298/BOS, CGL2009-08648-E and CGL 2011-22540/BOS.

7. References


Most will agree that gel electrophoresis is one of the basic pillars of molecular biology. This coined terminology covers a myriad of gel-based separation approaches that rely mainly on fractionating biomolecules under electrophoretic current based mainly on the molecular weight. In this book, the authors try to present simplified fundamentals of gel-based separation together with exemplarily applications of this versatile technique. We try to keep the contents of the book crisp and comprehensive, and hope that it will receive overwhelming interest and deliver benefits and valuable information to the readers.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: