



E-ZINE OF BIOLOGICAL SCIENCES

ISSN: 2456-7264 | Issue - 20 | Published On 08/02/2022

TECHNIQUES FOR FINGERPRINTING OF MICROBES

Dr. Arijit Shome

*Assistant Professor, Dept. of Veterinary Biochemistry, College of Veterinary Sciences,
Assam Agricultural University, Khanapara, Guwahati, Assam-781022*

Introduction

The vast subject of microbiology including all its areas like environmental, industrial, medical and agricultural microbiology involves identification and classification of microbes based on their phenotypic characters. With the advances in molecular genetic techniques, our ability to identify and characterize microorganisms has immensely enhanced. Prior to this, microorganisms were detected based on morphological, physiological and cultural characteristics. These methods have one or more problems related to reproducibility and untypeability. Moreover, as these methods are based on products of gene expression, these are highly sensitive to environmental variations like growth temperature and spontaneous mutations. Due to such drawbacks, genotypic methods of microbial typing gained impetus.

All the modern methods of typing rely on DNA visualization. The DNA banding patterns or so-called fingerprints have been developed and are already in place. Most of these techniques employ PCR and restriction site analysis as their basic tools but require huge set-up cost and thus are not economical for use. Much work is going on worldwide to make them more applicable and economic. Here, we shall be discussing about different DNA fingerprinting techniques commonly used for detection of pathogens.

Restriction Endonuclease Analysis of Chromosome (REAC)

REAC, as the name suggests, uses the technique of restriction enzyme (RE) digestion of chromosomal DNA to arrive at a specific banding pattern or 'fingerprint' after agarose gel electrophoresis or polyacrylamide gel electrophoresis (PAGE). These bands are then blotted on a suitable membrane (nylon or nitrocellulose) and then stained. These banding patterns obtained after the RE digestion indicate the cutting sites for the specific enzymes present on the chromosomal DNA of that particular pathogen, which is usually specific to the pathogen. Different strains of a pathogen showing one band difference in the fingerprint are considered to be subtypes of each other. Similarly, if the isolates differ in two or more banding patterns, they are considered as separate strains altogether. REAC is considered to be a critical tool for epidemiological investigation as even point mutation, insertions, deletions, site-specific recombination and transformation can be detected using this technique. A major advantage of

this technique is that it uses the whole chromosome for RE analysis and hence all strains can be typed and no prior sequence data is required. However, interpretation of the data is a daunting task and presence of plasmid DNA in the reaction mixture can interfere with the data.

Restriction Endonuclease Analysis of Plasmid DNA (REAP)

This technique is employed for those bacteria which harbor plasmids within them. Plasmids can be isolated using an alkaline lysis procedure followed by agarose gel separation. The number and size of plasmid is taken into consideration before strain identification. It is mostly used for identification of staphylococcal isolates which carry multiple plasmids. But some selected species of *Enterobacteriaceae* carry large plasmids of the range 100- 150 kb that makes it difficult to distinguish between strains. In such cases, addition of RE digestion step would increase the discriminatory power of agarose gel electrophoresis.

Restriction Fragment Length Polymorphism (RFLP) and Ribotyping

It is a very well-known established technique that uses the principle of restriction endonucleases (RE) digestion for a defined genetic region of the pathogen followed by electrophoresis to separate the digested fragments, and immobilization of the bands by southern blotting for detection by probe hybridization. One pertinent question that might arise in our minds is that bacteria have their own RE produced *in vivo* which might attack the bacterial genomic DNA, then why does the bacterial DNA not get degraded? The answer lies within the cells' own machinery of methylating adenine or cytosine using cellular methylases thus protecting the genomic DNA from degradation. Therefore, RFLP demands a careful choice of RE to avoid any inhibitory effect of methylation in the DNA. RFLP has been used successfully for *Brucella* spp., *Pseudomonas* and *Mycobacterium tuberculosis*.

Ribotyping or the rRNA gene restriction pattern analysis is a type of RFLP in which probe's target sequence is the *rrn* operon gene. The typical bacterial *rrn* operon consists of 16S–23S–5S rRNA genes, and is highly transcribed in bacteria, especially during the exponential phase of growth. These operons exist in multiple copies and are highly conserved sequences in bacteria. Multicopy existence of this operon makes it a perfect target for intra- and inter- species discrimination. Strains differentiated by ribotyping are called 'ribotypes'. Ribotyping is used for typing of *Legionella pneumophila*, *Haemophilus influenzae*, *Clostridium* spp. etc. RFLP and ribotyping offer good typing methods and are very useful for long-term epidemiological studies. However, it is time-consuming and requires technical expertise for interpretation.

Similar to normal ribotyping, we have PCR-ribotyping which entails the use of specific intergenic spacer region of ribosomal DNA. This has been used for characterization of different serotypes of *Salmonella*, *H. influenza* and *Yersinia enterocolitica*. Amplified ribosomal DNA restriction analysis (ARDRA) is a modified method of PCR ribotyping based

on PCR amplification of the 16S rRNA gene followed by restriction digestion of the amplified products. The advantage of this technique is that it can be used either for identification of the organism in pure cultures or in microbial communities.

Random Amplification of Polymorphic DNA (RAPD)

It is an amplification of random segments of genomic DNA with a single primer of arbitrary nucleotide sequence. Several arbitrary, short, non-specific primers (8–12 nucleotides) are created, then PCR is carried out using a large template of genomic DNA hoping that the fragments will get amplified. These primers bind at various ‘best-fit’ sequences on the denatured DNA at low stringent PCR conditions and carry out extension efficiently to yield short amplicons and the subsequent cycles are carried out with high stringency to generate products of fixed lengths. The major advantage is that no prior information of sequence is needed. However, stringency and efficiency come at the cost of some drawbacks. Rigorous standardization is needed, since the primers are random and can lead to several priming events. Moreover, this technique suffers from a lack of inter-laboratory reproducibility. A little change in protocol or primers may give different results. Nevertheless, this technique has been used successfully for typing of a number of bacteria like *Campylobacter jejuni*, *C. coli*, *Listeria monocytogenes*, *Staphylococcus hemolyticus* and *Vibrio vulnificus*.

Repetitive Sequence-based PCR (rep-PCR)

This technique has been recognized as an effective method for bacterial strain typing. Bacterial chromosomes contain multiple interspersed repetitive sequences that occupy intergenic regions at sites dispersed throughout the genome. It is these sites, which are made targets for the oligonucleotide probes enabling the generation of unique DNA fingerprints for individual bacterial strains. Mostly three different repetitive sequences are used for this technique, namely, 35-40 bp repetitive extragenic palindromic sequence (REP) elements, 124-127 bp enterobacterial repetitive intergenic consensus (ERIC) sequence and the 154 bp BOX elements. All these elements are genetically stable and differ only in their number and chromosomal locations between bacterial species. REP and ERIC sequences are most widely used targets in gram-negative bacteria. These were first identified in *E. coli* and *Salmonella typhi*. Rep-PCR is becoming more popular and a method of choice for typing bacteria and has been used for identifying various strains of *L. monocytogenes*, *V. parahaemolyticus*, *Burkholderia cepacia*, *Haemophilus somnus* etc.

Pulse Field Gel Electrophoresis (PFGE)

PFGE is considered as the ‘Gold Standard’ among the molecular typing techniques because of its reproducibility and discriminatory power. PFGE involves immobilization of the bacterial cells in agarose plugs and extraction of genomic DNA from those cells by lysis.

Later the extracted DNA is subjected to rare-cutting RE digestion. A highly specialized apparatus is used to carry out electrophoresis of the digested plugs, in which the orientation and duration of electric field are periodically changed which are called the 'Pulse time' and 'Switch time', respectively. The pulse-field allows large DNA fragments (10-800 kb) to orient and reorient themselves resulting in a better resolution. Interpretation of PFGE bands is relatively straightforward. If identical patterns are obtained, then the strains are considered identical. If the patterns differ by 1-3 bands, then the strains are considered to be closely related. If there is a difference of 4-6 bands then, the strains are possibly related and their genomes differ by two independent genetic events. Strains are considered unrelated, if their patterns differ by >6 bands. Although sensitive, PFGE is a time-consuming technique.

Amplified Fragment Length Polymorphism (AFLP)

It is a method of fingerprinting bacterial genomes based on PCR amplification of a selective subset of genomic restriction fragments. A small amount of purified genomic DNA is digested with two different restriction enzymes. The fragments are then ligated with linkers or adapters, possessing a sequence corresponding to the recognition site of the restriction enzyme and a point mutation such that the initial restriction site is not restored after ligation. Ligated fragments are then subjected to PCR amplification under high stringency conditions with two primers. Each primer is complementary to either adapter molecules. The 3' end of each primer extends to one to three nucleotides beyond the restriction site, resulting in selective amplification of those restriction fragments that contain nucleotides complementary to 3' nucleotide of that primer.

Multilocus Sequence Typing (MLST)

MLST is a technique in molecular biology for typing of multiple loci using DNA sequences of internal fragments of multiple housekeeping genes to characterize isolates of microbial species. The first MLST scheme to be developed was for *Neisseria meningitidis*, the causative agent of meningococcal meningitis and septicemia. MLST has also been established for *Streptococcus pneumoniae* and is in progress for *Streptococcus pyogenes*, *Haemophilus influenzae* and *Campylobacter jejuni*. The typical workflow of MLST involves: data collection, data analysis and multilocus sequence analysis. In the data collection step, definitive identification of variation is obtained by nucleotide sequence determination of gene fragments. In the data analysis step, all unique sequences are assigned allele numbers and combined into an allelic profile and assigned a sequence type (ST). In the final analysis step of MLST, the relatedness of the isolates are determined by comparing the allelic profiles. MLST is a suitable method to study long-term and global epidemiology or evolution of microbes. Also, MLST is well-suited for the construction of global databanks accessible to different laboratories for result comparison and classification.

Whole Genome Sequencing

This is unequivocally the most sensitive method for distinguishing microbes. The nucleotide sequence is determined with an automated sequencing instrument. DNA sequencing instruments are based on the modification of dideoxynucleotide chain terminator chemistry in which the sequencing primer is labeled at 5' end with one of the four fluorescent dyes. Each of the fluorescent dyes represents one of the four nucleotides and thus four different annealing and extension reactions are performed. Finally, the four sets of reactions are combined, concentrated and loaded in a single well on a polyacrylamide gel. During electrophoresis the fluorescently labeled products are excited and the corresponding signal is automatically detected. Subsequently, the resulting data is processed into a final sequence with the aid of computer software. Although regarded as the ultimate technique for identification, DNA sequencing is highly expensive and requires a high degree of technical competency.