Molecular Diagnosis for Bacterial Identification in Fermented Food: Case Study Nata de Coco

Antonius SUWANTO

Department of Biology, Faculty of Mathematics and Natural Science, Bogor Agricultural University, Jl. Raya Pajajaran, Bogor 16144; and South East Asian Regional Center for Tropical Biology (SEAMEO-BIOTROP), Bogor, INDONESIA
E-mail: asuwanto@indo.net.id

Summary

Molecular methods for microbial identification and typing can be grouped into metabolic-, protein-, fatty acid-, and nucleic acid-based methods. Nucleic acid profile analysis includes plasmid fingerprinting, restriction endonuclease analysis (REA), restriction fragment length polymorphisms (RFLP), ribotyping, macrorestriction fragment length polymorphisms (MFLP) employing pulsed field gel electrophoresis, and nucleic acid amplification (PCR-based methods). All of these techniques can be combined with DNA hybridization and/or sequencing analysis. None of these typing methods will generate wrong results, but each will provide a different way or angle of looking at the same problem. However, many scientific publications dealing with genome analysis reported that MFLP consistently yielded more discriminative and reproducible results when compared with REA, RFLP, ribotyping, and plasmid fingerprinting. MFLP is also the only technique available to date, which allows one to construct physical and genetic maps of bacterial genome. Complete physical and genetics maps of bacterial genome will definitely facilitate more precise identification and typing. Outlines of plasmid fingerprinting, REA, ribotyping, RFLP, 16S- and 18S-rRNA sequence analysis, and MFLP for microbial profiling and their applications to monitor microbial community analysis in *nata de coco* fermentation will be presented.

Introduction

The primary purpose of all microbial classification schemes is to identify a microorganism down to the species level, which forms the basic unit of a taxonomic grouping. However, the separation and accurate recognition of subspecies or strains within a pathovar is assuming greater importance in all branches of microbiology, particularly in medical microbiology as well as other areas in which accurate microbial typing is important, such as environmental studies concerning the release of and monitoring of novel microorganisms into new habitats (1), screening of microbial products, and patenting a novel or genetically engineered microorganism.

In recent years many microbial taxonomists have accepted that molecular studies, especially nucleic acid analyses provide the best available and most reliable methods for designating species and determining relationships between different organisms. DNA sequence analyses represent the ultimate reference for recognizing sub-types within a species. Ideally, comparison among strains within a species should be conducted through DNA sequence analysis. However, long range DNA sequence analysis, at least up to date, is rather impractical for most clinical purposes. Therefore any DNA typing technique available to date relies on closely finding which technique can detect differences at the DNA level between isolates. For this reason, lack of distinguishable characters determined by a chosen typing method does not necessarily imply that the two strains that being compared are identical. Such organisms may be distinguishable when ‘viewed’ through the typing ‘window’ generated by another typing scheme.

DNA fingerprinting for bacteria may be grouped into (i) Plasmid-based fingerprinting (ii) Restriction endonuclease analysis. Any of these typing methods may incorporate DNA hybridization or DNA sequence analysis. This paper will only deal with the first two groups of nucleic acid typing for bacteria. Use of ERIC and REP primers for PCR-based typing are recently reviewed elsewhere (2).

Plasmid Fingerprinting

Many strains of bacteria harbor extra-chromosomal genetic elements designated as plasmids. Although most plasmids known to date are relatively small in size (usually less than 200-kb) and having circular DNA topology, linear plasmids as well as very large plasmids (designated megaplasmids) have also been reported (3.4). Conventional plasmid isolation using alkaline lysis procedures as well as numerous derivative methods have made plasmid isolation and analysis become routinely employed techniques in many microbiology and molecular biology kaboratories. Therefore, plasmid profiles, either as intact or digested plasmid DNA, can be adopted as a reliable molecular fingerprinting for bacterial identification.

Although in most cases plasmid profile analysis has been considered less discriminatory than other genotypic methods (5), it was reported very recently that plasmid fingerprinting was highly discriminatory compared with Pulsed Field Gel Electrophoresis (PFGE) methods or immunoblotting (6).

A major drawback in plasmid profile analysis is the possible instability of plasmids, in particular bacterial strains, and some bacteria may not naturally harbor any endogenous plasmid. In addition, many large plasmids, such as megaplasmids in *Rhizobium* or *Agrobacterium* may not be easily isolated using conventional plasmid isolation protocols.

Restriction Endonuclease Analysis

Most bacterial chromosomes comprise of a single circular double-stranded DNA molecule. The size of bacterial chromosome varies from as small as 500-kb in some species of *Mycoplasma* up to approximately 9000-kb in *myxococcus* sp. The *Escherichia coli* chromosome is approximately 4700-kb long. However, some exceptions in both chromosome number and topology have been reported in several bacteria, for example *Rhodobacter sphaeroides* 2.4.1 and *Leptospira interrogans* have two circular chromosomes (7,8), *Borrelia burgdoferi* has one linear chromosome (9), while *Agrobacterium tumefaciens* has one circular and one linear chromosome (3).

In restriction enzyme analysis (REA), restriction endonucleases with relatively frequent restriction sites are used to digest the bacterial DNA, thereby generating hundreds of fragments ranging from ~0.5-kb to 50-kb in length (5). Such fragments can be separated by size with the use of constant-field agarose gel electrophoresis, and the pattern can be detected by staining the gel with ethidium bromide and photographing it under ultraviolet light. Different strains of the same bacterial species may have different REA profiles because of variations in their DNA sequences that alter the availability or distribution of restriction sites.

Genome analysis employing REA analysis frequently faces the difficulty of interpreting the complex DNA profiles which may consist of hundreds of DNA bands that may be unresolved and overlapping. REA becomes more complicated if the organism carries endogenous plasmid/s, since the fragment DNAs generated from plasmid DNA should also be taken into consideration. Practically it is difficult to identify DNA fragments generated from the endogenous plasmid among commonly overcrowded DNA bands resulting from frequent cutting restriction endonucleases.

Restriction Fragment Length Polymorphisms (RFLP)

The difficulty in interpreting DNA banding patterns in REA analysis can be overcome by the use of Southern hybridization analysis (10). Using a labeled DNA fragment as a probe, one can detect the restriction fragment (s) containing sequences (loci) homologous to the probe variations in the number and size of these fragments are referred to as restriction fragment length polymorphisms (RFLPs) which reflect variations in both the number of loci that are homologous to the probe and the location of restriction sites within or flanking those loci. All strains carrying loci homologous to the probe are typable, and the results are highly reproducible (5).

Several different DNA fragments can be used as probes, such as (I) ribosomal DNA or RNA which form the basis of ribotyping; (ii) randomly cloned DNA fragment of unknown function (11) (such a probe is often applicable only to the species from which it was derived originally although the method itself is a general one); and (iii) cloned DNA fragment of a known gene.

The use of any of these probe types can significantly reduce the number of bands in the original chromosomal fingerprint, and thereby will facilitate comparisons between strains. The main disadvantage of this approach, as well as ribotyping methods, is that the information gained only reflects DNA sequence within specific region/s of the genome that happen to hybridize with a particular probe being used. For universal purposes, PCR techniques can be used to amplify 16-S or 18S-RNA genes. The amplification product is subsequently digested with specific restriction enzymes to reveal specific RFLP patterns. This technique, often designated as Amplified Ribosomal DNA Restriction Analysis (ARDRA), has been employed to select for unique strains among a number of isolates prior to DNA sequencing, or to conduct genetic diversity analysis of closely related microbial isolates (20).

Ribotyping

The use of rRNA-based probes for RFLP is referred to as ribotyping (12). Ribotyping has emerged as one of the most powerful of the currently available methods for investigating the epidemiology of widely diverse bacteria. As have been mentioned previously, the other probe types may be limited to a particular species or strains within species.

Many of the rRNA sequences found in bacteria, archaea, and eukarya appear to have changed a little during the course of molecular evolution (13), so that probes specific for these sequences can detect a wide range of bacteria with similar rRNA sequences.

rRNA genes in bacteria are usually organized into operons, within which the individual genes coding for 16S, 23S, and 5S RNA are often separated by non-coding spacer DNA. The use of a labeled mixture of 16S and 23S RNA as a probe results in hybridization only with the fragments in the chromosomal fingerprint that contains parts of the corresponding genes. In contrast, the use of cloned fragments of the rRNA genes themselves as a probe may result in hybridization with the corresponding parts of the rRNA genes and the spacer sequences (1). Therefore, different hybridization patterns can be obtained depending on the precise probe used (14).

16S- and 18S-rRNA Sequences for Microbial Identification

Because of the likely antiquity of the protein-synthesizing machinery and for several other reasons, ribosomal RNAs are excellent molecules for discerning evolutionary relationships among living organisms. Ribosomal RNAs are ancient molecules, functionally constant, universally distributed, and moderately well conserved in sequence across broad phylogenetic distances. Also, because the number of different possible sequences of large molecules such as ribosomal RNAs is so large, similarity in two sequences always indicates some phylogenetic relationship. However, it is the degree of similarity in ribosomal RNA sequences between two organisms that indicate their relative evolutionary relatedness. From comparative sequence analyses, molecular genealogy can be constructed leading to phylogenetic trees that show the true evolutionary position of organisms relative to one another (See Fig. 1)

There are three ribosomal RNA molecules, which in procaryotes have sizes of 5S, 16S, and 23S. The large bacterial rRNAs, 16S and 23S rRNA (approximately 1500 and 2900 nucleotides, respectively) contain several regions of highly conserved sequence useful for obtaining proper sequence variability in other regions of the molecule to serve as excellent phylogenetic chronometers.

The 5S rRNA has also been used for phylogenetic measurements but its small size (-120 nucleotides) limits the information obtainable from this molecule. Because 16S RNA is more experimentally manageable than 23S RNA, it has been used extensively to develop the phylogeny of both prokaryotes and eukaryotes (using the 18S rRNA counterpart of prokaryotic 16S rRNA counterpart of prokaryotic 16S rRNA). Because 16S rRNA originates from the small (30S) sub unit of the ribosome, the acronym SSU (for small sub unit) sequencing is synonymous with 16S- or 18S-rRNA gene sequencing.

The database of rRNA sequences in the Ribosomal Database Project (RDP) now numbers over 10,000 and can be accessed on the internet (http://www.cme.msu.edu/RDP/). Use of 16S rRNA as a phylogenetic tool was pioneered in the early 1970s by Carl Woese at the University of Illinois and the method is now widely used.

The methods for obtaining ribosomal RNA sequences and generating phylogenetic trees are now quite routine and involve a combination of molecular biology and computer analyses. Newly generated sequences are compared with sequences in the RDP and/or with sequences obtained from other databases such as GenBank (USA) or EMBL (Germany) to reveal their closely related microorganisms (for microbial identification). In addition, using a treeing program, a phylogenetic tree is generated that summarizes the evolutionary information inherent in the sequences.

Macrorestriction Fragment Length Polymorphisms (MFLPs)

RFLP-based methods, such as ribotyping require DNA hybridization analysis to identify the size and distribution of DNA fragments involved in generating polymorphism. This additional step is needed due to the difficulty in interpreting overlapping DNA bands resulting from REA. The development of Pulsed Field Gel Electrophoresis (PFGE) allows us to separate large DNA molecule up to 5000-kb (15). Digestion of total bacterial genomic DNA using rare-cutting restriction endonucleases (16) yielded 10-30 DNA fragments that are manageable for direct interpretation of REA profiles. However, a PFGE apparatus is generally needed to separate the resulting DNA fragments, generated from rare-cutting restriction enzyme digestion, which are usually 50-2000-kb in length.

Conventional gel electrophoresis methods are based on the ability of a gel matrix to sieve alternating electric fields, which force the migrating DNA molecules to their direction of migration continuously. It is assumed that larger DNA molecules change direction more slowly than the smaller molecule, such that when certain “pulse time” in one orientation is applied, the actual time for movement for larger DNA molecules is less than that for the smaller ones. To achieve separation of a broad range of DNA sizes, such as often required for typing purposes, it is a common practice to increase the pulse time gradually over the course of the experiment (termed ramping) which ensures that each size range of DNA molecules is subjected to optimal separation conditions for at least part of the running time. Therefore, the choice of appropriate pulse time is one of the most important parameters for PFGE.

Application of PFGE for Bacterial Typing

PFGE of bacterial genomic DNA digested with rare-cutting restriction endonucleases generate a unique DNA banding pattern for almost each strain of bacteria. This DNA banding pattern, which is termed schizotyping (17), provides an estimate of the degree of genomic relationship between strains that is important for most epidemiological or clinical purposes (18, 19). The discriminatory value and information generated by the technique can be increases further by the use of more than one rare-cutting enzyme, either individually or in multiple digestion. Genomic fingerprinting employing PFGE is a reliable technique that generates reproducible results, which make PFGE analysis a suitable general technique that deserves consideration for investigating the epidemiology of most microorganisms.

Case study: Nata de Coco Fermentation

Nata de coco or nata is a jelly-like dessert food widely known in Southeast Asia, especially in Indonesia, The Philippines, Malaysia, and Thailand. This sweet delicacy is usually cut into small cubicles (1.5x1.5x1.5 cm^3) as part of fruit cocktail or yogurt. The texture of nata creates a unique “bite sensation” which is one of the most important reasons why people like to enjoy it.

Nata is actually cellulose mat produced by certain strains of *Acetobacter xylinum* through traditional static fermentation. Although this bacterial cellulose has been subjected for extensive investigation, especially on the fiber microstructure, biosynthetic pathway, and production of cellulose in agitated culture; very few studies was conducted to elucidate the biochemical and microbiological aspects of fermentation in nata production for food.

In our study to screen for effective producers of bacterial cellulose from sucrose, we found that traditional fermentation of nata de coco was not supported by single bacterial strains only. In fact, it was a polymicrobial fermentation involving strains of *Acetobacter*, lactic acid bacteria, and yeast. PFGE was employed to reveal genetic diversity among strains of Acetobacter and Amplified Ribosomal DNA Restriction Analysis (ARDRA) based on 16S-rRNA gene amplification was employed to analyze the dynamics of microbial community fermentation during nata fermentation.

Our preliminary study indicated that at least two genotypes of *Acetobacter* played important roles in making nata of prime quality. In addition, we also found that certain strains of lactic acid bacteria and yeast could enhance cellulose production in static culture. PFGE consistently showed higher resolution for genetic profiling of *Acetobacter* strains when compared to ARDRA. However, ARDRA or Terminal Restriction Fragment Length Polymorphisms (T-RFLP) combined with DGGE or SSCP would be very important to study microbial community dynamics which could lead to molecular diagnostics of traditional nata de coco fermentation.

Acknowlegment

We thank Deborah Hadisusanto and Cecilia A. Seumahu for conducting preliminary experiment on microbial community analysis of nata fermentation.

References

Mieth, D. 1993. The release of microorganisms-ethical criteria. In: Wohrmann, K., Tomiuk, J. eds. Transgenic organisms: Risk Assessment of deliberate release. Birkhauser Verlag. Switzerland. 1993.

Versalovic J., T. Koeuth, and R. Lupski. 1991. Distribution of repetitive DNA sequence in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. 19: 6823-6831.

Allardet-Servent A., S. Michaux-Charachon, E. Jurnas-Bilak, L. Karayan, and M. Ramuz. 1993. Presence of one linear and one circular chromosome in the *Agrobacterium tumefaciens* C58 genome. J. Bacteriol. 175:7869-7874.

Honeycut, R.J., M. McClelland, and B.W.S. Sobral. 1993. Physical map of the genome *Rhizobium meliloti* 1021. J. Bacteriol. 175:6945-6952.

Maslow, J.N., M.E. Mulligan, and R.D. Arbeit. 1993. Molecular epidemiology: application of contemporary techniques to the typing of microorganisms. Clin. Infect. Dis. 17:153-164.

Hartstein, A.I., C.L. Phelps, R.Y.Y. Kwok, and M.E. Mulligan. 1995. In vivo stability and discriminatory power of methicillin-resistant *Staphilococcus aureus* typing by restriction endonuclease analysis of a plasmid DNA compared with those of other molecular methods. J. Clin. Microbiol. 33:2022-2026.

Suwanto, A., and S. Kaplan. 1989. Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: presence of two unique circular chromosomes. J. Bacteriol. 171:5850-5859.

Zuemer, R.L., J.L. Hermann, and I. Saint Girons. 1993. Comparison of genetic maps for two *Leptospira imterrogans* serovars provides evidence for two chromosomes and interspecies heterogeneity. J. Bacteriol. 175:5445-5451.

Ferdow, M.S., and A.G. Barbour. 1989. Megabase-size linear DNA in the bacterium *Borrelia burgdoferi*, the lyme disease agent. Proc. Natl. Acad. Sci. USA. 86:5969-5973.

Southern, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Tompkin, L.S., N. Troup, A. Labigne-Roussel, and M.L. Cohen. 1986. Cloned, random chromosomal sequences as probes to identify Salmonella species. J. Infect. Dis. 154:156-162.

Grimont, P.A.D., and F. Grimont. 1986. Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. Annales de L'Institute Pasteur. 137B: 165-175.

Woese, CR. 1987. Bacterial evolution. Microbiol. Rev. 51:221-271.

Saunders, N.A., T.G. Harrison, Haththotuwa, and A.G. Taylor. 1991. A comparison of probes for restriction fragment length polymorphisms (RFLP) typing of *Legionella pneumophilla*. J. Med. Microbiol. 35:152-158.

Burmeister, M., and L. Ulanovsky, eds. 1992. Pulsed-Field Gel Electrophoresis. Methods in Molecular Biology. Humana Press, Totowa, NJ. 481.

Suwanto, A., and S. Kaplan. 1989. Physical and genetic mapping of the *Rhodobacter sphaeroides* 2.4.1 genome: Genome size, fragment identification, and gene localization. J. Bacteriol. 171:5840-5849.

Suwanto, A., and S. Kaplan. 1992. Chromosome transfer in *Rhodobacter sphaeroides*: Hfr formation and genetic evidence for two unique circular chromosomes. J. Bacteriol. 174:1135-1145.

Suwanto, A., M. Yuhana, E. Herawaty, and S.L. Angka. 1998. Genetic Diversity of luminous *Vibrio* isolated from shrimp larvae. *In* Flegel, T.W. (ed.) Advances in shrimp biotechnology. National Center for Genetic Engineering and Biotechnology. Bangkok.

Widjaja, R., A. Suwanto, and B. Tjahjono. 1999. Genome size and macrorestriction map of *Xanthomonas campestris* pv. glycines YR32 chromosome. FEMS Microbiol. Lett. 75:59-68.

Ross, I.L., Y. Alami, P. Harvey, W. Achouak, and M.H. Ryder. 2000. Genetic diversity and biological control activity of novel species of closely related *Pseudomonas* isolated from wheat field soils in South Australia. Apl. Environ. Microbiol. 66:1609-1616.

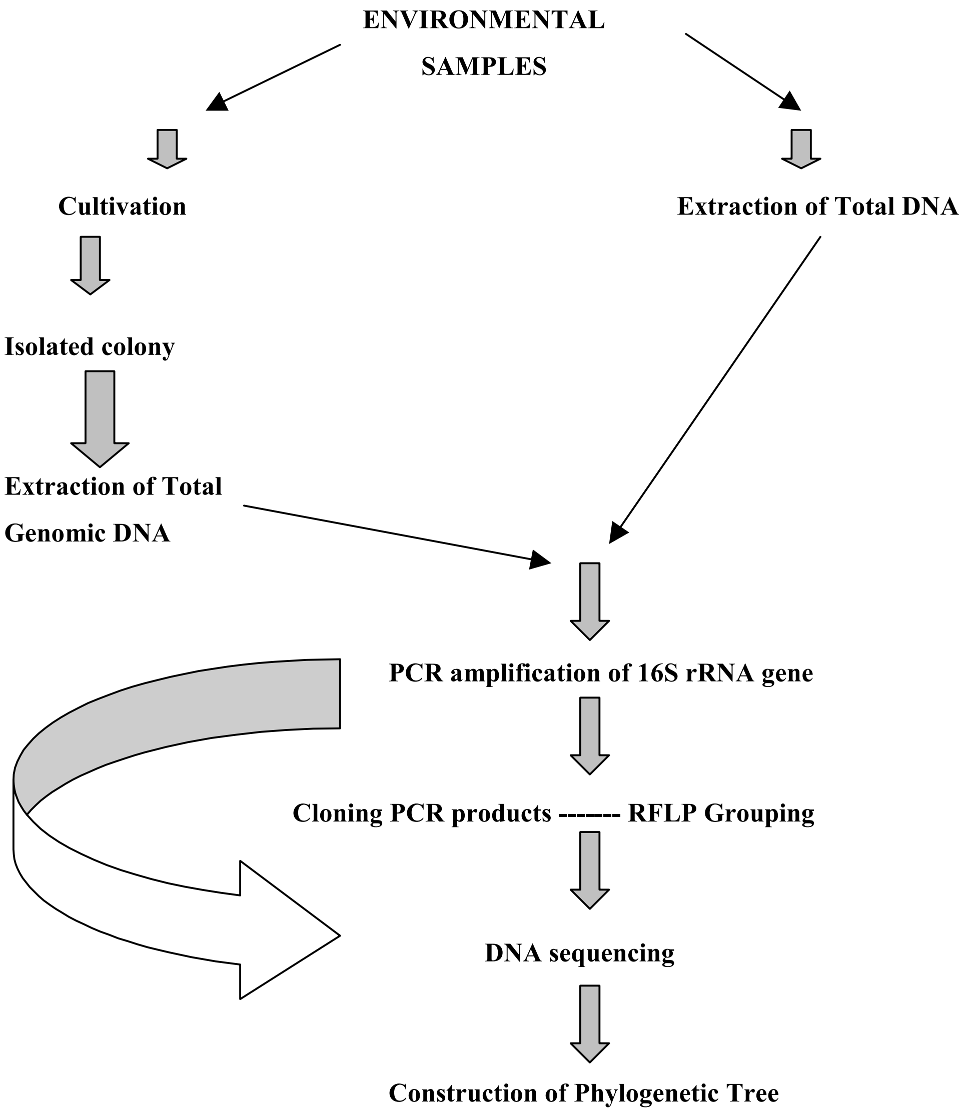

Fig. 1: Molecular diagnosis for bacterial identification