

## MINIREVIEW

# Principles and Applications of Methods for DNA-Based Typing of Microbial Organisms

D. MICHAEL OLIVE\* AND PAMELA BEAN

*Millennium Strategies, Madison, Wisconsin 53719*

Outbreaks of infectious disease often result from exposure to a common source of the etiologic agent. Generally, the etiologic agent causing an outbreak of infection is derived from a single cell whose progeny are genetically identical or closely related to the source organism. In epidemiological terms, the organisms involved in the outbreak are clonally related; that is, they have a common origin. Clonally related organisms are members of the same species that share virulence factors, biochemical traits, and genomic characteristics. However, there is sufficient diversity at the species level that organisms isolated from different sources at different times and in different geographical regions may be differentiated or classified into subtypes or strains.

The process of subtyping is important epidemiologically for recognizing outbreaks of infection, detecting the cross-transmission of nosocomial pathogens, determining the source of the infection, recognizing particularly virulent strains of organisms, and monitoring vaccination programs.

Subtyping or strain classification has been accomplished by a number of different approaches. All of these methods must meet several criteria in order to be broadly useful. All organisms within a species must be typeable by the method used. For some phenotypic methods, such as those based on a reaction with an antibody or the presence of a bacteriophage receptor, the particular characteristic may not be present in all members of the species. Serotypes can be assigned only if a particular serological marker is present. For example, 10 to 15% of *Mycobacterium avium* strains are untypeable because they fail to react with any available antiserum (78). Similarly, bacteriophage typing has often been applied to the characterization of *Staphylococcus aureus*, yet *S. aureus* isolates lacking the bacteriophage receptor cannot become infected with the typing phage (10, 64). In the absence of the marker, the isolates cannot be assigned a phage type.

Any subtyping method must have high differentiation power (3). It must be able to clearly differentiate unrelated strains, such as those that are geographically distinct from the source organism, and at the same time to demonstrate the relationship of all organisms isolated from individuals infected through the same source. As will be discussed in this review, molecular methods can differ widely in their ability to differentiate strains.

A third concern for subtyping methodologies is reproducibility (3). Reproducibility refers to the ability of a technique to yield the same result when a particular strain is repeatedly tested. Reproducibility is especially important for the construc-

tion of reliable databases containing all known strains within a species to which unknown organisms can be compared for classification. Variable expression of phenotypic characteristics, such as sporadic expression of virulence genes or antigens, can contribute to problems with reproducibility (3).

Many of the currently used molecular techniques for typing rely on electrophoretic separation of DNA fragments of different molecular lengths. The electrophoretic result is represented by a pattern of bands on a gel. Since these patterns may be extremely complex, the ease with which the patterns are interpreted and related is a factor in evaluating the utility of a particular typing method (3). Along with considerations related to a particular method's ease of interpretation, its ease of use is also important (3). The technical difficulty, cost, and time to obtain a result must also be evaluated in assessing the utility of a particular typing method.

The shortcomings of phenotypically based typing methods have led to the development of typing methods based on the microbial genotype or DNA sequence, which minimize problems with typeability and reproducibility and, in some cases, enable the establishment of large databases of characterized organisms. In this review, we have compared some of the most current technologies for genotypic typing with regard to the principles of the methods and their reproducibility and discrimination power, ease of use, and cost in order to give the reader some measure of the clinical practicality and utility of the molecular tools available today.

## MOLECULAR TYPING METHODS

**PFGE of whole chromosomal DNA.** Pulsed-field gel electrophoresis (PFGE) is often considered the "gold standard" of molecular typing methods. For PFGE, bacterial isolates grown either in broth or on solid media are combined with molten agarose and poured into small molds. The results are agarose plugs containing the whole bacteria. The embedded bacteria are then subjected to in situ detergent-enzyme lysis and digestion with an infrequently cutting restriction enzyme. The digested bacterial plugs are then inserted into an agarose gel and subjected to electrophoresis in an apparatus in which the polarity of the current is changed at predetermined intervals. The pulsed field allows clear separation of very-large-molecular-length DNA fragments ranging from 10 to 800 kb (65). The electrophoretic patterns are visualized following staining of the gels with a fluorescent dye such as ethidium bromide. Gel results can be photographed, and the data can be stored by using one of the commercially available digital systems, such as those manufactured by Alpha-Innotech, Bio-Rad, Hitachi, or Molecular Dynamics. Data analysis can be accomplished by using any of a number of commercially available software

\* Corresponding author. Mailing address: Millennium Strategies, 3404 Country Grove Dr., Madison, WI 53719. Phone: (608) 848-1509. Fax: (608) 848-1757. E-mail: molive@madison.tds.net.

packages available from Applied Math, Bio-Rad, BioSystematics, Media Cybernetics, or Scanalytics.

Tenover et al. (75) have proposed a system for standardizing the interpretation of PFGE patterns in relation to determining strain relatedness. In their scheme, bacterial isolates yielding the same PFGE pattern are considered the same strain. Bacterial isolates differing by a single genetic event, reflected as a difference of one to three bands, are closely related. Isolates differing by four to six bands, representing two independent genetic changes, are possibly related. Bacterial isolates containing six or more band differences, representative of three or more genetic changes, are considered unrelated. These criteria are applicable to small, local studies in which genetic variability is presumed to be limited.

Numerous studies which define the banding patterns of a variety of bacteria digested with a number of restriction enzymes have been conducted. PFGE has proven to be superior to most other methods for biochemical and molecular typing. It is highly discriminatory and superior to most methods for analysis of *Escherichia coli*, vancomycin-resistant enterococci, *S. aureus*, *Acinetobacter* species, *Pseudomonas aeruginosa*, and *M. avium* (4, 5, 7, 30, 46, 52, 57, 63, 64, 73). PFGE was more efficient at differentiating strains of methicillin-resistant *S. aureus* than other methods tested (58, 63). PFGE was also more discriminatory than repetitive element sequence-based PCR (Rep-PCR) for differentiating strains of the *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex, vancomycin-resistant enterococci, *Neisseria gonorrhoeae*, and *P. aeruginosa* (7, 30, 46, 56).

With the aid of the computerized gel scanning and analysis software, it is possible to create data banks of PFGE patterns for all organisms, enabling the creation of reference databases to which any new strain could be compared for identifying its phylogenetic relationship to other similar strains. This capability is reflected in a recent report to the President on food safety initiatives (22). Two of the recommendations were (i) to enhance surveillance of food-borne outbreaks of disease by equipping FoodNet sites with the equipment to carry out DNA fingerprinting by PFGE in order to identify the source of infectious agents and (ii) to establish an electronic network for rapid fingerprint comparison which would allow sentinel sites to rapidly share data (22).

One of the factors that has limited the use of PFGE is the time involved in completing the analysis. While the procedural steps are straightforward, the time needed to complete the procedure can be 2 to 3 days. This can reduce the laboratory's ability to analyze large numbers of samples.

**Southern blotting and RFLP methods.** Southern blotting has been used for years to detect and locate genomic sequences from a variety of prokaryotic and eukaryotic organisms. For gene detection, whole chromosomal DNA is digested with a restriction enzyme, and the fragments are separated by electrophoresis through an agarose gel. The separated DNA fragments are transferred from the agarose gel to either a nitrocellulose or nylon membrane by Southern blotting (69). The membrane-bound nucleic acid is then hybridized to one or more labeled probes homologous to the gene to be examined. Probes can be labeled with a number of detectable moieties, including enzyme-colorimetric substrates or enzyme-chemiluminescent substrates. This classic method has been adapted to the differentiation of bacterial strains on the basis of the observation that the locations of various restriction enzyme recognition sites within a particular genetic locus of interest can be polymorphic from strain to strain, resulting in gel bands that differ in size between unlike strains. Thus, the name restriction fragment length polymorphism (RFLP) refers to the polymor-

phic nature of the locations of restriction enzyme sites within defined genetic regions. Only the genomic DNA fragments that hybridize to the probes are visible in RFLP analysis, which simplifies the analysis greatly.

Gene-specific probes have been used to subtype *Brucella* species (29), *Legionella pneumophila* (76), and *P. aeruginosa* (47). Furthermore, ribotyping, a variation of RFLP-Southern blotting in which the probes are derived from the 16S and 23S rRNA genes, has been applied successfully in many studies for differentiating bacterial strains (12, 20, 25, 34, 40, 72). Ribotyping results in only a small number of bands, which simplifies the interpretation. However, this also limits the ability of the technique to distinguish between closely related strains.

Multiple genetic sites can also be targeted in Southern blotting studies to subtype bacteria. Grundmann et al. (30) used a combination of a *tox4* gene probe with probes for the 16S and 23S rRNA genes to subtype *P. aeruginosa* isolates. Both RFLP-Southern blotting probe sets yielded information about the evolutionary relationships of the strains examined, yet neither technique was as efficient at discriminating strain differences as PFGE. While RFLP-Southern blotting has been a useful technique for limited studies of strain subtypes, the cumbersome blotting techniques have been largely replaced by PCR-based locus-specific RFLP.

**PCR-based locus-specific RFLP.** The PCR has enabled specific genetic loci to be routinely amplified and examined for differences indicative of strain variation and antimicrobial resistance. The specific locus to be examined is amplified with gene-specific primers and subjected to RFLP analysis. The DNA fragments are separated on an agarose or small polyacrylamide gel, and the digestion patterns are visualized following ethidium bromide staining.

Locus-specific RFLP has been applied in a number of situations. Shortridge et al. (67) have used RFLP of the *ureC* gene to demonstrate the genetic diversity of *Helicobacter pylori* strains in the United States. The 16S, 23S, and 16S-23S spacer regions have also been used as targets for locus-specific RFLP (30, 83, 93). In this variation of ribotyping, the ribosomal DNA is amplified and subjected to digestion with a restriction enzyme, and the DNA fragments are visualized following separation by gel electrophoresis, alleviating the need for Southern blotting. Ribotyping can be successfully applied for differentiation of bacterial strains that display a high degree of heterogeneity within the rRNA operons.

In a novel application of locus-specific RFLP, Cockerill et al. (16) have identified point mutations in the *katG* gene of *Mycobacterium tuberculosis* that correspond to various levels of resistance to isoniazid by examination of RFLP patterns of the amplified gene. However, the discriminatory power of locus-specific PCR is generally not as good as that of other methods, due primarily to the limited region of the genome which can be examined. For example, in a study by Kühn et al. (42), PCR-ribotyping showed poor discriminatory power in comparison to PFGE and biochemical typing methods.

Locus-specific RFLP has found a significant application in epidemiological studies of hepatitis C virus (HCV). HCV can be divided into six major genotypes, numbered 1 through 6. Davidson et al. (17) have developed a simple method for determining the genotype of an HCV isolate based on PCR amplification and RFLP analysis of the 5' untranslated region of the HCV genome. Digestion with either the combination or single enzymes *HaeIII-RsaI*, *MvaI-HinFI*, *BstI*, and *ScrFI* allows differentiation of viral isolates into groups 1a, 1b, 2a, 2b, 3a, 3b, 4, 5, and 6. RFLP of the 5' untranslated region has facilitated studies of the geographical distribution of viral genotypes, the

natural history of the disease, and correlation of genotypes with response to interferon.

**RAPD assays.** The random amplified polymorphic DNA (RAPD) assay, also referred to as arbitrary primed PCR, was first described by Williams et al. (87) and Welsh and McClelland (86). RAPD assays are based on the use of short random sequence primers, 9 to 10 bases in length, which hybridize with sufficient affinity to chromosomal DNA sequences at low annealing temperatures such that they can be used to initiate amplification of regions of the bacterial genome. If two RAPD primers anneal within a few kilobases of each other in the proper orientation, a PCR product with a molecular length corresponding to the distance between the two primers results. The number and location of these random primer sites vary for different strains of a bacterial species. Thus, following separation of the amplification products by agarose gel electrophoresis, a pattern of bands which, in theory, is characteristic of the particular bacterial strain results (14, 51, 86, 87).

In most cases the sequences of the RAPD primers which generate the best DNA pattern for differentiation must be determined empirically. More recently, a consensus M13 DNA sequencing primer has been used in RAPD fingerprinting assays, which allows for some standardization of the procedure (83, 85).

A number of studies have reported success in using RAPD assays to distinguish bacterial strains among diverse species, including *A. baumannii* (60), *Candida albicans* (45, 61), *H. pylori* (1), *Proteus mirabilis* (9), *Histoplasma capsulatum* (37), *Haemophilus somnus* (53), *Leptospira* species (59), *S. aureus* (63), and *L. pneumophila* (76, 79).

In comparison to other typing techniques, Vila et al. (83) found that the RAPD assay was more discriminating than RFLP analysis of either the 16S rRNA genes or the 16S-23S rRNA spacer region but less discriminating than Rep-PCR. Furthermore, a number of problems have been reported for RAPD assays that contribute to a lack of reproducibility and standardization. Since the primers are not directed against any particular genetic locus, many of the priming events are the result of imperfect hybridization between the primer and the target site. Thus, the amplification process is extremely sensitive to slight changes in the annealing temperature which can lead to variability in the banding patterns. The use of empirically designed primers, each with its own optimal reaction conditions and reagents, also makes standardization of the technique difficult (6, 51, 86).

**Rep-PCR.** Versalovic et al. (80) described a method for fingerprinting bacterial genomes by examining strain-specific patterns obtained from PCR amplification of repetitive DNA elements present within bacterial genomes. Two main sets of repetitive elements are used for typing purposes. The repetitive extragenic palindromic (REP) elements are 38-bp sequences consisting of six degenerate positions and a 5-bp variable loop between each side of a conserved palindromic stem (71). REP sequences have been described for numerous enteric bacteria (28, 33, 66, 71, 92). The palindromic nature of the REP elements and their ability to form stem-loop structures have led to multiple proposed functions for these highly conserved, dispersed elements (27, 54, 92).

The enterobacterial repetitive intergenic consensus (ERIC) sequences are a second set of DNA sequences which have been successfully used for DNA typing. ERIC sequences are 126-bp elements which contain a highly conserved central inverted repeat and are located in extragenic regions of the bacterial genome (33, 66). They have been defined primarily based on sequence data obtained from *E. coli* and *Salmonella typhimurium* (33, 66).

While the REP and ERIC sequences are the most commonly used targets for DNA typing, another repetitive element, the BOX sequence, has been used to differentiate *Streptococcus pneumoniae* (39, 50). BOX elements are located within intergenic regions and can also form stem-loop structures due to their dyad symmetry. They are mosaic repetitive elements composed of various combinations of three subunit sequences referred to as boxA, boxB, and boxC (50). The three subunit sequences have molecular lengths of 59, 45, and 50 nucleotides, respectively (50). The BOX elements have no sequence relationship to either REP or ERIC sequences (50). While initially thought to be unique to *S. pneumoniae*, BOX elements have now been found in a number of bacterial species (43).

Rep-PCR can be performed with DNA extracted from bacterial colonies or by a modified method using unprocessed whole cells (90). REP or ERIC amplification can be performed with a single primer, a single set of primers, or multiple sets of primers. ERIC patterns are generally less complex than REP patterns, but both give good discrimination at the strain level. Application of both REP and ERIC PCR to samples to be typed increases the discriminatory power over that of either technique used alone.

Rep-PCR is fast becoming the most widely used method of DNA typing. Rep-PCR with primers based on REP and ERIC sequences has been successfully used to differentiate strains of *Bartonella* (62), *Bacillus subtilis* (80), *Citrobacter diversus* (32, 89), *Enterobacter aerogenes* (24), *Rhizobium meliloti* (18), methicillin-resistant *S. aureus* (19), *S. pneumoniae* (81, 93), *A. baumannii* (21, 60), *Burkholderia cepacia* (31), *L. pneumophila* (23), *H. pylori* (43), *N. gonorrhoeae* (56), and *Neisseria meningitidis* (91). The technique is easy to perform and can be applied to large or small numbers of isolates. Rep-PCR shows broader species applicability and better discriminatory power than either plasmid profiling or genomic fingerprinting (24). Rep-PCR has considerably better discriminatory power than restriction analysis of the 16S rRNA gene or the 16S-23S spacer region (2, 83). Furthermore, studies which have compared Rep-PCR to other typing methods such as multilocus enzyme electrophoresis (89), biochemical characterizations (15), or ribotyping (2, 68, 83) have shown Rep-PCR to be superior to these methods. Finally, several studies have shown Rep-PCR to have good correlation with PFGE results but, in general, with slightly less discriminatory power (7, 46).

Rep-PCR has been adapted to an automated format in which fluorescently labeled primers are used to create either the REP or ERIC profile and the amplified sequences are separated via a fluorescence-based DNA sequencer (19, 82). This method allows consistent pattern formation and storage of the data in a database as a digitized image. Unknown strains can be compared against the stored database for identification purposes. Thus, like PFGE, Rep-PCR lends itself to standardization, which will enable comparisons of strains isolated in different laboratories.

**CFLP.** Cleavase I is a thermostable, engineered endonuclease which consists of the 5'-nuclease domain (amino acids 1 to 281) of *Thermus aquaticus* DNA polymerase (13). Whereas restriction endonucleases recognize specific DNA sequences as their recognition sites, cleavase I recognizes structures formed by base-paired DNA at elevated temperatures (13). When denatured by heating at 95°C followed by cooling at temperatures ranging between 50 and 60°C, the single strands of DNA go through a self-base-pairing phase during which hairpin structures unique to each strand are formed prior to reannealing with the complementary DNA strand. This process creates a series of stem-loop structures connected by non-base-paired



single-stranded regions. Cleavase I recognizes the stem-loop structures and cuts the DNA 5' to the first paired base at the junction of the single- and double-stranded regions. Theoretically, in any given DNA preparation the denatured single strands represent a mixture of all possible configurations which the strands can assume. For the cleavase reaction, a PCR product labeled at one of its 5' termini is subjected to partial digestion with the enzyme such that all possible digestion products are represented in the reaction mixture. The resulting mixture is analogous to a dideoxy sequencing reaction in that it contains a collection of randomly cleaved, labeled single-stranded DNA molecules extending from the 5' labeled end to a terminal cleavage site.

Differences in the structural fingerprints between dissimilar PCR products may appear as a gain or loss of a band, a mobility shift in a band, or a change in the intensity of a particular band. While gains, losses, and shifts of bands are easily recognized, differences in band intensity, which occur frequently, are extremely difficult to detect and can often lead to false results.

The discrimination power of cleavase fragment length polymorphism (CFLP) is dependent on the DNA sequence of the locus examined. For example, while a pattern abundant in cleavage products is obtained from digestion of 16S rDNA, very few cleavage products are generated by digestion of the *rpoB* gene of *M. tuberculosis*, precluding identification of many sequence differences in the region critical for drug resistance (13). Furthermore, each DNA strand may digest with different abilities to discriminate sequence changes. For example, whereas differences in the nucleotide sequences of the intergenic regions of *Shigella dysenteriae* serotypes 1, 2, and 8 are represented in the structural fingerprints of the antisense strand, the fingerprints of the sense strand are indistinguishable. Maddox et al. (48) compared the abilities of CFLP and single-strand conformation polymorphism to detect single point mutations. In that study, single-strand conformation polymorphism detected six of six mutations analyzed, while CFLP could detect only three of the six mutations. Thus, the applicability of CFLP analysis to any given DNA cannot be predicted but must rather be determined empirically. Some DNA molecules give excellent results in terms of differentiation, while others give no results at all due to unknown reasons. CFLP is also size limited; that is, differences in the structural fingerprints of DNA fragments larger than 1 kbp cannot be reliably resolved by electrophoresis.

CFLP has been applied to genotypic differentiation of HCV isolates (49, 70). In both of these studies, genotype analysis by CFLP was successful in differentiating the strains examined. However, while the intralaboratory reproducibilities of the banding patterns for each viral type were consistent, the interlaboratory reproducibilities of the same patterns for the viral subtypes were poor. In the two studies that used CFLP for genotyping HCV, the structural fingerprints of HCV type 1a and 1b isolates in the study by Marshall et al. (49) showed significant differences from those reported by Sreevatsan et al. (70). Both studies presented CFLP fingerprints on a 244-bp fragment derived from the 5' noncoding region by PCR amplification with the Amplicor HCV kit (Roche Molecular Diagnostics, Branchburg, N.J.). The patterns produced for a given genotype in the study by Marshall et al. (49) were generally consistent, and the patterns were easily interpreted. In contrast, while a consistent, unique fingerprint was generated for each genotype tested in the study by Sreevatsan et al. (70), the fingerprints were considerably more complex and difficult to interpret. Furthermore, the fingerprints for HCV types 1a and 1b differed significantly from the patterns for types 1a and

1b presented by Marshall et al. (49). The implication of this is that while CFLP can achieve acceptable intralaboratory reproducibility, the interlaboratory reproducibility limits its utility. Patterns generated by one laboratory cannot be shared or used as a reference for other laboratories.

The need for empirical determination of the applicability of CFLP to any given genetic system coupled with the difficulty in pattern interpretation and the interlaboratory variation in the CFLP patterns of the same organisms compromise the use of this technology for clinical purposes. CFLP is also considerably more complex to perform than either locus-specific PCR-RFLP, Rep-PCR, or RAPD analysis. With the exception of DNA sequencing, the discriminatory ability of CFLP has not been compared to that of any of the other molecular typing techniques, so its capabilities remain unknown.

**AFLP assays.** Amplified fragment length polymorphism (AFLP) is a genome fingerprinting technique based on the selective amplification of a subset of DNA fragments generated by restriction enzyme digestion (84). Originally applied to the characterization of plant genomes, AFLP has been more recently applied to bacterial typing (21, 26, 35, 36, 38). Two variations of AFLP have been described, one with two different restriction enzymes and two primers for amplification and the second with a single primer and restriction enzyme (21, 26, 35, 36, 38). In the most common configuration, bacterial DNA is extracted, purified, and subjected to digestion with two different enzymes, such as *EcoRI* and *MseI*. The restriction fragments are then ligated to linkers containing each restriction site and a sequence homologous to a PCR primer binding site. The PCR primers used for amplification contain DNA sequences homologous to the linker and contain one to two selective bases at their 3' ends. For example, a selective primer directed against an *EcoRI* site might have the sequence 5'-GAA TTCAA-3', where the first six bases are complementary to the *EcoRI* site while the two A residues at the 3' end are selective and allow amplification of only those *EcoRI* sites with the sequence 3'-CTTAATT-5'. They would not amplify an *EcoRI* site with the sequence 3'-CTTAATC-5'. Thus, the selective nucleotides allow amplification of only a subset of the genomic restriction fragments. In order to visualize the patterns, one of the PCR primers contains either a radioactive (21, 26, 35, 36) or a fluorescent (38) label, or, alternatively, the gels may be stained with ethidium bromide and the patterns may be examined under UV light (26).

Studies to date have demonstrated that AFLP is reproducible (21, 35, 38) and has good ability to differentiate clonally derived strains (21, 26, 35, 36, 38). The differentiation power of AFLP appears to be greater than that of PCR-based ribotyping; however, it has not yet been compared to Rep-PCR or PFGE (38). The AFLP procedure is more labor-intensive than Rep-PCR, but results are obtained more rapidly than with PFGE. Radioactive formats are impractical for clinical use, and while significant advantages exist with the fluorescent format, the cost of a DNA sequencer may be prohibitive for most laboratories. A commercially available kit can be used in combination with enzymatic detection of the bands; however, this requires that the DNA fragments be transferred to a filter by blotting and adds considerably to the labor costs. Yet, the reproducibility of the banding patterns for a given strain facilitates storing patterns in databases for use in identifying new bacterial strains. This combined with strong discriminatory power may make AFLP attractive to those laboratories performing frequent epidemiological studies such that a DNA sequencer becomes cost-effective.

**DNA sequencing.** Ultimately, all molecular genetic methods for distinguishing organism subtypes are based on differences

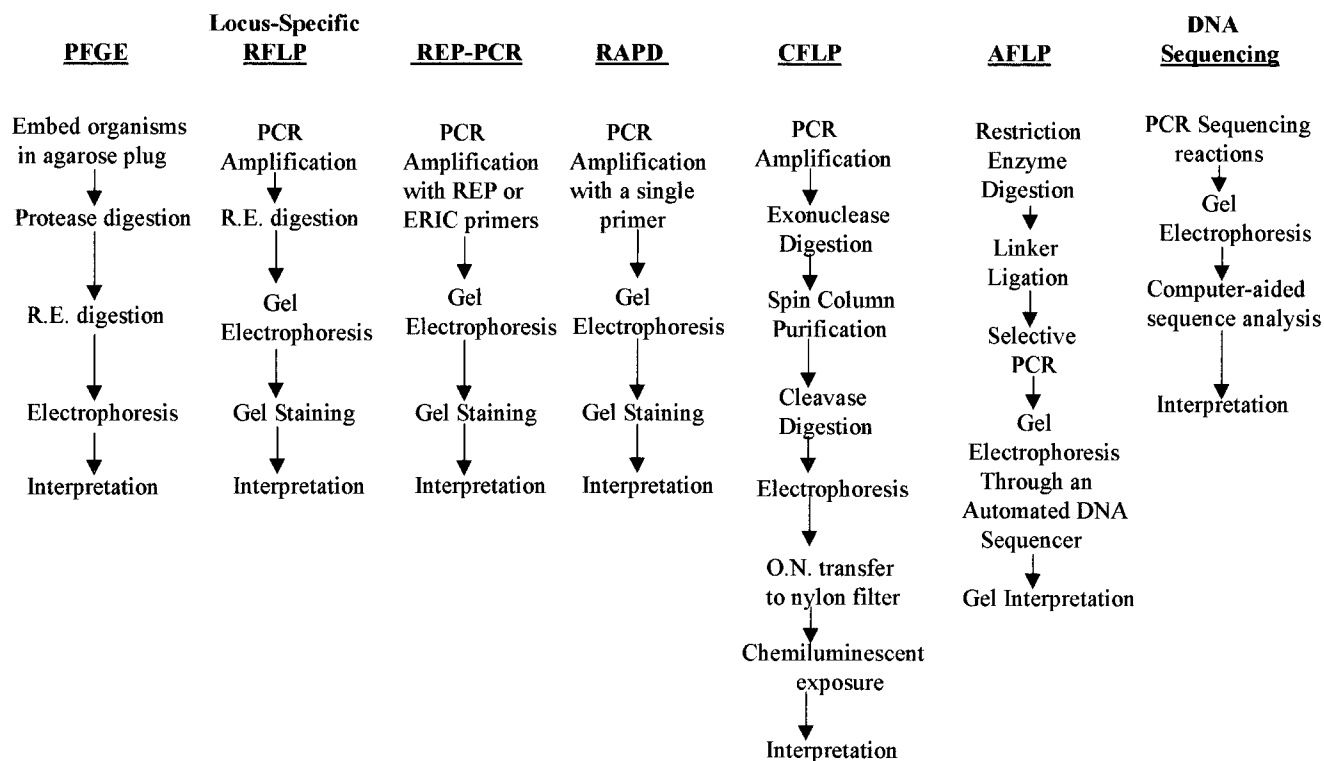


FIG. 1. Comparison of the procedural steps involved in various molecular typing methods. R.E., restriction endonuclease; O.N., overnight.

in the DNA sequence. Logically, then, DNA sequencing would appear to be the best approach to differentiating subtypes. However, as discussed below, this is not the case. DNA sequencing was originally performed by using radioactive labels for detection of the reaction products, an approach that is unsuitable for clinical use. Current DNA sequencing protocols employ fluorescent nucleotides to label the DNA. The sequence is then read with an automated instrument.

DNA sequencing generally begins with PCR amplification of a sample DNA directed at genetic regions of interest, followed by sequencing reactions with the PCR products. Modifications of the procedure can be made such that RNA may also be used as the starting material. Instruments that perform automated analysis of DNA sequencing gels are based on real-time detection of fluorescently labeled sequencing reaction products. DNA sequencing instruments most commonly utilize a modification of dideoxynucleotide chain termination chemistry in which the sequencing primer is labeled at the 5' end with one of four fluorescent dyes. Each of the four fluorescent dyes represents one of the four nucleotides which make up the DNA molecule, and hence four separate annealing and extension reactions must be performed for each DNA sample to be analyzed. Once the DNA sequencing reactions have been completed, the four sets of reaction products are combined, concentrated, and loaded in a single well on a polyacrylamide gel. During electrophoresis, these fluorescently labeled products are excited by an argon laser and automatically detected. The resulting data is stored in digital form for subsequent processing into the final sequence with the aid of specialized software.

There are several considerations that must be evaluated before undertaking the use of DNA sequencing for subtyping. First, for practical purposes in the clinical laboratory, DNA sequencing must be directed at only a very small region of the

chromosome of an organism. It is impractical to sequence multiple or large regions of the chromosome of an organism. Thus, in contrast to techniques like PFGE, Rep-PCR, or RAPD analysis, which examine the entire chromosome, DNA sequencing examines only a very small portion of the sites which can potentially vary between bacterial or fungal strains.

Furthermore, the short region of DNA used must meet several criteria before it can be used for strain differentiation. The structure of the region of DNA selected must consist of a variable sequence flanked by highly conserved regions. This enables PCR amplification and typing of all members of a species. The variability within the selected sequence must be sufficient to differentiate different strains of a particular species. Finally, the selected sequence should not be horizontally transmissible to other strains of the species.

Unfortunately, for bacteria and fungi, few sequences meet these criteria. While the 16S rRNA genes have been used to identify new species of organisms, they show limited variability between strains of a bacterial species (11, 44, 55, 88). The intergenic region between the 16S and 23S rRNA genes has also been used with variable results (8). Essentially species-specific genetic loci must be identified and tested empirically for the ability of a locus to provide strain differentiation. The major limitation of DNA sequencing approaches to bacterial and fungal subtyping is the size constraints of the DNA that can be practically sequenced.

In contrast, DNA sequencing is considered the gold standard for viral typing. Nucleic acid sequencing has much greater discriminatory power for determination of HCV genotypes and is the only molecular method by which mutations associated with human immunodeficiency virus (HIV) drug resistance can be reliably determined. The regions used for viral genotyping and detection of drug resistance mutations are in short, well-

<u>PFGE</u>	<u>RFLP</u>	<u>RAPD or REP-PCR</u>	<u>CFLP</u>	<u>AFLP</u>	<u>DNA Sequencing</u>
*PFGE Power Supply	*PCR Amplification Reagents	*PCR Amplification Reagents	*PCR Amplification Reagents	*DNA Linkers	*PCR sequencing kits
*Gel boxes	*Thermocycler	*Thermocycler	*Thermocycler	*PCR Amplification Reagents	*Thermocycler
*Proteases	*Oligonucleotide Primers	*Oligonucleotide Primers	*Oligonucleotide Primers	*Oligonucleotide Primers	*Automated DNA Sequencer
*Restriction Enzyme(s)	*Restriction Enzyme(s)	*Gel Box & Low Voltage Power Supply	*20 cm x 20 cm Gel Box	*Oligonucleotide Primers	*Analysis Software
*Photography System	*Gel Box & Low Voltage Power Supply	*Photography System	*Sequencing Power Supply	*Thermocycler	*Computer
	*Photograph System		*Exonuclease I	*Automated DNA Sequencer	
			*Spin columns	*Analysis Software	
			*CFLP kit	*Computer	
			*Nylon Membranes		
			*Chemiluminescent Detection Kits		
			*Autoradiography Supplies		

<u>ESTIMATED COSTS</u>	<u>PFGE</u>	<u>Locus-Specific RFLP</u>	<u>RAPD or Rep-PCR</u>	<u>CFLP</u>	<u>AFLP</u>	<u>DNA Sequencing</u>
Set-up	\$10K-20K	\$8K-10K	\$8-10K	\$8K-12K	\$45K-130K <sup>1</sup>	\$45-130K
Material/rx <sup>2</sup> :	\$12.00	\$5.00	\$3.00	\$20.00 <sup>3</sup>	\$10.00	\$20.00 <sup>3</sup>
Labor/rx <sup>4</sup> :	\$10.00	\$9.00	\$8.00	\$20.00 <sup>5</sup>	\$10.00	\$20.00 <sup>5</sup>
Total/rx:	\$22.00	\$14.00	\$11.00	\$40.00	\$20.00	\$40.00

FIG. 2. Equipment, reagents, and costs associated with molecular typing methods. Footnotes: 1, the large range in setup cost corresponds to the various types of automated DNA sequencing available; 2, material costs include enzymes, nucleotides, primers, disposable tubes and pipette tips, chemicals, purification kits, and electrophoresis reagents; 3, the material cost was calculated based on analysis of both DNA strands; 4, labor costs were calculated based on the total hands-on time of a technologist paid at a rate of \$20.00/h and analyzing batches of 12 samples per run; 5, labor costs were calculated based on analysis of both DNA strands. rx, reaction.

defined sequences that fulfill the criteria needed for the application of nucleic acid sequencing methodologies.

In applying DNA sequencing to genotypic studies, the laboratory should also be aware of the constraints of the sequencing methodology. One of the most problematic areas involved in DNA sequencing is the preparation of the sequencing gel. Some sequencing instruments either use precast gels or have apparatuses which greatly simplify gel preparation. Sequencing instruments that employ capillary gels, which automate gel preparation and loading, are even more convenient. However, traditional sequencing instruments in which the gel consists of a 40- by 40-cm slab gel are cumbersome to handle and pour.

Care must also be used when interpreting DNA sequencing data. The laboratory should be aware of artifacts that can complicate the interpretation. For example, repeats of a single nucleotide in a sequence can result in band compression artifacts that lead to misreading of the sequence. Both DNA strands should be sequenced and compared in order to minimize incorrect base identification.

In general, DNA sequencing is expensive and requires a high degree of technical competency to perform. Furthermore, automated DNA sequencers can be prohibitively expensive, with some costing in excess of \$100,000.

In the future, non-gel-based sequencing methods based on hybridization of an unknown sequence to a DNA array may simplify the task of DNA sequencing for typing purposes as gene-specific chips are made. Kozal et al. (41) have reported the use of high-density DNA arrays to determine the DNA

sequences of HIV clade B proteases of 167 HIV isolates taken from 102 patients in order to identify mutations associated with resistance to protease inhibitors. More recently, Troesch et al. (77) have described the use of a high-density DNA array to simultaneously identify mycobacteria at the species level and detect the presence of mutations in the *rpoB* gene associated with resistance to rifampin. Seventy isolates comprising 27 different species of mycobacteria and 15 rifampin-resistant strains of *M. tuberculosis* were analyzed. For species determination, the 16S rRNA hypervariable region was amplified by PCR, and the amplicons were hybridized to the array. Twenty-six of 27 species tested were correctly identified. Similarly, mutations associated with rifampin resistance were detected in all 15 rifampin-resistant isolates by hybridization of amplicons obtained by amplification of the *rpoB* gene to the array. This demonstrates the ability of DNA arrays to simultaneously interrogate more than one genetic locus. The multiplex capabilities of DNA arrays may significantly increase the utility of DNA sequencing-based methods for strain identification, since multiple genes could be rapidly screened for sequence-specific strain differences. The multiplexing capabilities of a DNA array would also enable simultaneous sequencing of both DNA strands, decreasing the amount of material and labor required to complete an analysis and increasing the cost-effectiveness of a chip for typing. While this technology is still some years away from commercialization, DNA arrays show intriguing prospects for the future.

TABLE 1. Summary of the characteristics of the various molecular typing methods

Methodology	Ease of use	Ease of interpretation	Discrimination power <sup>a</sup>	Time to result (days)	Intralaboratory reproducibility	Interlaboratory reproducibility	Setup cost	Cost per test
PFGE	Moderate	Easy	High	3	Good	Good	Moderate	Moderate
Locus-specific PCR-RFLP	Easy	Easy	Moderate	1	Good	Good	Moderate	Low
Rep-PCR	Easy	Easy	High	1	Good	Moderate	Moderate	Low
RAPD	Easy	Easy	High	1	Moderate	Poor	Moderate	Low
CFLP	Moderate	Moderate	Moderate	2	Good	Poor	Moderate	High
AFLP	Moderate	Easy	High	2	Good	Good	High	Moderate
Sequencing	Difficult	Moderate	High	2	Good	Good	High	High

<sup>a</sup> This table is intended to give a relative estimate of discrimination power. For more exacting comparisons, see the references cited in the text.

### PRACTICAL ASPECTS OF MOLECULAR TYPING METHODS

Tenover et al. have identified criteria to aid in the selection and interpretation of molecular typing methods for epidemiological studies (74). Although a particular typing method may have high discriminatory power and good reproducibility, the complexity of the method and interpretation of results as well as the costs involved in setting up and using the method may be beyond the capabilities of the laboratory. The choice of a molecular typing method, therefore, will depend upon the needs, skill level, and resources of the laboratory. As summarized in Fig. 1, methods such as locus-specific PCR, RAPD analysis, and Rep-PCR are similar in their procedures and are generally the easiest to implement. While PFGE appears to be time-consuming, it too is not difficult to implement. In contrast, procedures such as DNA sequencing, which require considerable knowledge about the instrumentation, methods, and potential pitfalls that can lead to uninterpretable results, can be difficult to implement reliably and require a higher skill level. Likewise, CFLP, with its need for DNA purification, careful optimization, and complex interpretation and its questionable reproducibility, can also be difficult to implement in a clinical laboratory.

These methods also vary with regard to their equipment needs, setup costs, and cost per test (Fig. 2). For example, while DNA sequencing may ensure that the subtlest of changes can be observed, the cost of an automated sequencer may be prohibitively high. Furthermore, for subtyping bacteria and fungi, the laboratory must weigh the fact that the discriminatory power of techniques such as PFGE and Rep-PCR is sufficiently high to yield excellent subtyping results for a fraction of the setup costs as well as lower costs per test.

While the costs per test may be decreased for both DNA sequencing and CFLP if only one DNA strand is examined, this must be weighed against the reliability of the result for strain differentiation. As mentioned above, CFLP can show marked DNA strand-specific differences in discriminatory power. Likewise, DNA sequencing is most accurate when both DNA strands are sequenced.

While the time to obtain a result can be as much as 4 days for PFGE, the labor costs are lower than those for a 2-day result from a technique such as CFLP (Fig. 2). Most of the time required for performance of PFGE is the reaction incubation time, whereas CFLP requires considerably more hands-on manipulation by the technician.

The pertinent characteristics that are important for introduction of a molecular typing method into a clinical laboratory situation are summarized in Table 1. One other criterion, not considered here, is the ability of a method to allow analysis of large numbers of samples. This criterion may not be of importance to some clinical laboratories, while for those routinely

involved in large epidemiological studies, high-throughput capability may be crucial. For high throughput, simple techniques with high discriminatory power and low cost, such as Rep-PCR, may be most suitable. However, if storage of genetic profiles for the creation of reference databases of organisms is desirable, then methods such as PFGE or AFLP may be equally attractive for bacterial and fungal subtyping. For viral subtyping, DNA sequencing may be the only method suitable for monitoring genetic change and detecting mutations associated with drug resistance.

Modern epidemiologists are fortunate to have a variety of tools which provide good molecular differentiation and which can be tailored to fit the needs of both the laboratory and the clinical study. Several of these molecular methods could enable the creation of large reference libraries of typed organisms to which new outbreak strains can be compared across laboratories in order to monitor changes in microbial populations.

### REFERENCES

- Akopyanz, N., N. O. Bukanov, T. U. Westblom, S. Kresovich, and D. E. Berg. 1992. DNA diversity among clinical isolates of *Helicobacter pylori* detected by PCR-based RAPD fingerprinting. *Nucleic Acids Res.* **20**:5137-5142.
- Appunamy, S., R. Parton, J. G. Coote, and H. A. Gibb. 1997. Genomic fingerprinting of *Haemophilus somnus* by a combination of PCR methods. *J. Clin. Microbiol.* **35**:288-291.
- Arbeit, R. D. 1995. Laboratory procedures for the epidemiologic analysis of microorganisms, p. 190-208. In P. R. Murray, E. J. Baron, M. A. Tenover, F. C. Tenover, and R. H. Tenover (ed.), *Manual of clinical microbiology*, 6th ed. ASM Press, Washington, D.C.
- Arbeit, R. D., M. Arthur, R. D. Dunn, C. Kim, R. K. Selander, and R. Goldstein. 1990. Resolution of recent evolutionary divergence among *Escherichia coli* from related lineages: the application of pulsed field electrophoresis to molecular epidemiology. *J. Infect. Dis.* **161**:230-235.
- Arbeit, R. D., A. Slutsky, T. W. Barber, J. N. Maslow, S. Niemczyk, J. O. Falkinham, G. T. O'Connor, and C. F. von Reyn. 1993. Genetic diversity among strains of *Mycobacterium avium* causing monoclonal and polyclonal bacteremia in patients with AIDS. *J. Infect. Dis.* **167**:1384-1390.
- Arbeit, R. D., J. N. Maslow, and M. E. Mulligan. 1994. Polymerase chain reaction-mediated genotyping in microbial epidemiology. *Clin. Infect. Dis.* **18**:1018-1019. (Letter.)
- Barbier, N., P. Saulnier, E. Chachaty, S. Dumontier, and A. Andremont. 1996. Random amplified polymorphic DNA typing versus pulsed-field gel electrophoresis for epidemiological typing of vancomycin-resistant enterococci. *J. Clin. Microbiol.* **34**:106-109.
- Barry, T., G. Collieran, M. Glennon, L. K. Dunican, and F. Gannon. 1991. The 16S/23S ribosomal spacer region as a target for DNA probes to identify eubacteria. *PCR Methods Appl.* **1**:51-56.
- Bingen, E., C. Boissinot, P. Desjardins, H. Cave, N. Lambert-Zechovsk, E. Denamur, P. Blot, and J. Elion. 1993. Arbitrary primed polymerase chain reaction provides rapid differentiation of *Proteus mirabilis* isolates from a pediatric hospital. *J. Clin. Microbiol.* **31**:1055-1059.
- Blumberg, H. M., D. Rimland, J. A. Kiehlbauch, P. M. Terry, and I. K. Wachsmuth. 1992. Epidemiologic typing of *Staphylococcus aureus* by DNA restriction fragment length polymorphisms of rRNA genes: elucidation of the clonal nature of a group of bacteriophage-nontypeable ciprofloxacin-resistant methicillin-sensitive *S. aureus* isolates. *J. Clin. Microbiol.* **30**:362-369.
- Botter, E. C. 1989. Rapid determination of bacterial ribosomal RNA sequences by direct sequencing of enzymatically amplified DNA. *FEMS Microbiol. Lett.* **53**:171-176.



12. Brenner, D. J., L. W. Mayer, G. M. Carlone, L. H. Harrison, W. F. Bibb, M. C. Brandileone, F. O. Sottnek, K. Irino, M. W. Reeves, and J. M. Swenson. 1988. Biochemical, genetic, and epidemiologic characterization of *Haemophilus influenzae* biogroup *aegyptius* (*Haemophilus aegyptius*) strains associated with Brazilian purpuric fever. *J. Clin. Microbiol.* **26**:1524–1534.
13. Brow, M. A. D., M. C. Oldenburg, V. Lyamichev, L. M. Heisler, N. Lyamicheva, J. G. Hall, N. J. Eagan, D. M. Olive, L. M. Smith, L. Fors, and J. E. Dahlberg. 1996. Differentiation of bacterial 16S rRNA genes and intergenic regions and *Mycobacterium tuberculosis* *katG* genes by structure-specific endonuclease cleavage. *J. Clin. Microbiol.* **34**:3129–3137.
14. Caetano-Anolles, G., B. J. Bassam, and P. M. Gresshoff. 1991. DNA amplification fingerprinting using very short arbitrary oligonucleotide primers. *Bio/Technology* **9**:553–557.
15. Clarridge, J. E., III, T. J. Raich, D. Pirwani, B. Simon, L. Tai, M. C. Rodriguez-Barradas, R. Regnery, A. Zollo, D. C. Jones, and C. Rambo. 1995. Strategy to detect and identify *Bartonella* species in routine clinical laboratory yields *Bartonella henselae* from human immunodeficiency virus-positive patient and unique *Bartonella* strain from his cat. *J. Clin. Microbiol.* **33**:2107–2113.
16. Cockerill, F. R., III, J. R. Uhl, Z. Temesgen, Y. Zhang, L. Strockman, G. D. Roberts, D. L. Williams, and B. C. Kline. 1995. Rapid identification of a point mutation of the *Mycobacterium tuberculosis* catalase-peroxidase (*katG*) gene associated with isoniazid resistance. *J. Infect. Dis.* **171**:240–245.
17. Davidson, F., P. Simmonds, J. C. Ferguson, L. M. Jarvis, B. C. Dow, E. A. C. Follett, C. R. G. Seed, T. Krusius, C. Lin, G. A. Medgyesi, H. Kiyokawa, G. Olim, G. Duraisamy, T. Cuyper, A. A. Saeed, D. Teo, J. Conradie, M. C. Kew, M. Lin, C. Nuchaprayoon, O. K. Ndimbie, and P. L. Yap. 1995. Survey of major genotypes and subgenotypes of hepatitis C virus using RFLP of sequences amplified from the 5' non-coding region. *J. Gen. Virol.* **76**:1197–1204.
18. De Bruijn, F. J. 1992. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergenic consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. *Appl. Environ. Microbiol.* **58**:2180–2187.
19. DelVecchio, V. G., J. M. Petroziello, M. J. Gress, F. K. McClesky, G. P. Melcher, H. K. Crouch, and J. R. Lupski. 1995. Molecular genotyping of methicillin-resistant *Staphylococcus aureus* via fluorophore-enhanced repetitive-sequence PCR. *J. Clin. Microbiol.* **33**:2141–2144.
20. Dijkshoorn, L., H. M. Aucken, P. Gerner-Smidt, M. E. Kaufmann, J. Ursing, and T. L. Pitt. 1993. Correlation of typing methods for *Acinetobacter* isolates from hospital outbreaks. *J. Clin. Microbiol.* **31**:702–705.
21. Dijkshoorn, L., H. M. Aucken, P. Gerner-Smidt, P. Janssen, M. E. Kaufman, J. Garaizar, J. Ursing, and T. L. Pitt. 1996. Comparison of outbreak and nonoutbreak *Acinetobacter baumannii* strains by genotypic and phenotypic methods. *J. Clin. Microbiol.* **34**:1519–1525.
22. Food and Drug Administration, U.S. Department of Agriculture, U.S. Environmental Protection Agency, and Centers for Disease Control and Prevention. 1997. Food safety from farm to table: a national food safety initiative report to the President, p. 3. Food and Drug Administration, Washington, D.C.
23. Georgiou, P. R., A. M. Doggett, M. A. Kielhofner, J. E. Stout, D. A. Watson, J. R. Lupski, and R. J. Hamill. 1994. Molecular fingerprinting of *Legionella* species by repetitive element PCR. *J. Clin. Microbiol.* **32**:2989–2994.
24. Georgiou, P. R., R. J. Hamill, C. E. Wright, J. Versalovic, T. Koeuth, D. A. Watson, and J. R. Lupski. 1995. Molecular epidemiology of infections due to *Enterobacter aerogenes*: identification of hospital-associated strains by molecular techniques. *Clin. Infect. Dis.* **20**:84–94.
25. Gerner-Smidt, P. 1992. Ribotyping of *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex. *J. Clin. Microbiol.* **30**:2680–2685.
26. Gibson, J. R., E. Slater, J. Xery, D. S. Tompkins, and R. J. Owen. 1998. Use of an amplified-fragment length polymorphism technique to fingerprint and differentiate isolates of *Helicobacter pylori*. *J. Clin. Microbiol.* **36**:2580–2585.
27. Gilson, E., D. Perrin, and M. Hofnung. 1990. DNA polymerase I and a protein complex bind specifically to *E. coli* palindromic unit highly repetitive DNA: implications for bacterial chromosome organization. *Nucleic Acids Res.* **18**:3941–3952.
28. Gilson, E., J. M. Clement, D. Brutlag, and D. Hofnung. 1984. A family of interspersed extragenic palindromic DNA sequences in *E. coli*. *EMBO J.* **3**:1417–1421.
29. Grimont, F., J. M. Verger, P. Cornelis, J. Limet, M. Lafevre, M. Grayon, B. Regnault, J. Van Brock, and P. A. D. Grimont. 1992. Molecular typing of *Brucella* with cloned DNA probes. *Res. Microbiol.* **143**:55–65.
30. Grundmann, H. C. Schneider, D. Hartung, F. D. Daschner, and T. L. Pitt. 1995. Discriminatory power of three DNA-based typing techniques for *Pseudomonas aeruginosa*. *J. Clin. Microbiol.* **33**:528–534.
31. Hamill, R. J., E. D. Houston, P. R. Georgiou, C. E. Wright, M. A. Koza, R. M. Cadle, P. A. Goepfert, D. A. Lewis, G. J. Zenon, and J. E. Clarridge. 1995. An outbreak of *Burkholderia* (formerly *Pseudomonas*) *cepacia* respiratory tract colonization and infection associated with nebulized albuterol therapy. *Ann. Intern. Med.* **122**:762–766.
32. Harvey, B. S., T. Koeuth, J. Versalovic, C. R. Woods, and J. R. Lupski. 1995. Vertical transmission of *Citrobacter diversus* documented by DNA fingerprinting. *Infect. Control Hosp. Epidemiol.* **16**:564–569.
33. Hulton, C. S. J., C. F. Higgins, and P. M. Sharp. 1991. ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium*, and other enterobacteria. *Mol. Microbiol.* **5**:825–834.
34. Irino, K., F. Grimont, I. Casin, and P. A. Grimont. 1988. rRNA gene restriction patterns of *Haemophilus influenzae* biogroup *aegyptius* strains associated with Brazilian purpuric fever. *J. Clin. Microbiol.* **26**:1535–1538.
35. Janssen, P., R. Coopman, G. Huys, J. Swings, M. Bleeker, P. Vos, M. Zabeau, and K. Kersters. 1996. Evaluation of the DNA fingerprinting method AFLP as a new tool in bacterial taxonomy. *Microbiology* **142**:1881–1893.
36. Janssen, P., K. Maquelin, R. Coopman, I. Tjerberg, P. Bouvet, K. Kerstens, and L. Dijkshoorn. 1997. Discrimination of *Acinetobacter* genomic species by AFLP fingerprinting. *Int. J. Syst. Bacteriol.* **47**:1179–1187.
37. Kersulyte, D., J. P. Woods, E. J. Kreath, W. E. Goldman, and D. E. Berg. 1992. Diversity among clinical isolates of *Histoplasma capsulatum* detected by polymerase chain reaction with arbitrary primers. *J. Bacteriol.* **174**:7075–7079.
38. Koelman, J. G. M., J. Stoof, D. J. Biesmans, P. H. M. Savelkoul, and C. M. J. E. Vandembroucke-Grauls. 1998. Comparison of amplified ribosomal DNA restriction analysis, random amplified polymorphic DNA analysis, and amplified fragment length polymorphism fingerprinting for identification of *Acinetobacter* genomic species and typing of *Acinetobacter baumannii*. *J. Clin. Microbiol.* **36**:2522–2529.
39. Koeuth, T., J. Versalovic, and J. R. Lupski. 1995. Differential subsequence conservation of interspersed repetitive *Streptococcus pneumoniae* BOX elements in diverse bacteria. *Genome Res.* **5**:408–418.
40. Koopman, M. B. H., A. Kasboher, G. Beckmann, B. A. M. van der Zeijst, and J. G. Kusters. 1993. Genetic similarity of intestinal spirochetes from humans and various animal species. *J. Clin. Microbiol.* **31**:711–716.
41. Kozal, M. J., N. Shah, N. Shen, R. Yang, R. Fucini, T. C. Merigan, D. D. Richman, D. Morris, E. Hubbell, M. Chee, and T. R. Gingeras. 1996. Extensive polymorphisms observed in HIV-1 clade B protease gene using high-density oligonucleotide arrays. *Nat. Med.* **2**:753–759.
42. Kühn, L., L. G. Burman, S. Hæggen, K. Tullus, and B. E. Murray. 1995. Biochemical fingerprinting compared with ribotyping and pulsed-field gel electrophoresis of DNA for epidemiological typing of enterococci. *J. Clin. Microbiol.* **33**:2812–2817.
43. Kwon, D.-H., F. A. K. El-Zaatari, J.-S. Woo, C.-L. Perng, D. Y. Graham, and M. F. Goo. 1998. Rep-PCR fragments as biomarkers for differentiating gastroenteric disease-specific *Helicobacter pylori*. *Dig. Dis. Sci.* **43**:980–987.
44. Lane, D. J., B. Pace, G. J. Olsen, D. A. Stahl, M. L. Sogin, and N. R. Pace. 1985. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analysis. *Proc. Natl. Acad. Sci. USA* **82**:6955–6959.
45. Lehmann, P. F., D. Lin, and B. A. Lasker. 1992. Genotypic identification and characterization of species and strains within the genus *Candida* using random amplified polymorphic DNA. *J. Clin. Microbiol.* **30**:3249–3254.
46. Liu, P. Y.-F., and W.-L. Wu. 1997. Use of different PCR-based DNA fingerprinting techniques and pulsed-field gel electrophoresis to investigate the epidemiology of *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex. *Diagn. Microbiol. Infect. Dis.* **28**:19–28.
47. Loutit, J. S., and L. S. Tompkins. 1991. Restriction enzyme and Southern hybridization analyses of *Pseudomonas aeruginosa* strains from patients with cystic fibrosis. *J. Clin. Microbiol.* **29**:2897–2900.
48. Maddox, L. O., P. Li, A. Bennett, M. Descartes, and J. N. Thompson. 1997. Comparison of SSCP analysis and CFLP analysis for mutation detection in the human iduronate 2-sulfatase gene. *Biochem. Mol. Biol. Int.* **43**:1163–1171.
49. Marshall, D. J., L. M. Heisler, V. Lyamichev, C. Murvine, D. M. Olive, G. D. Ehrlich, B. P. Neri, and M. De Arruda. 1997. Determination of hepatitis C virus genotypes in the United States by cleavage fragment length polymorphism analysis. *J. Clin. Microbiol.* **35**:3156–3162.
50. Martin, B., O. Humbert, M. Camara, E. Guenzi, J. Walker, T. Mitchell, P. Andrew, M. Prudhomme, G. Alloing, and R. Hakenbeck. 1992. A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. *Nucleic Acids Res.* **20**:3479–3483.
51. Meunier, J. R., and P. A. D. Grimont. 1993. Factors affecting reproducibility of random amplified polymorphic DNA fingerprinting. *Res. Microbiol.* **144**:373–379.
52. Murray, B. E., K. V. Singh, J. D. Heath, B. R. Sharma, and G. M. Weinstock. 1990. Comparison of genomic DNAs of different enterococcus isolates using restriction endonucleases with infrequent recognition sites. *J. Clin. Microbiol.* **28**:2059–2063.
53. Myers, L. E., S. V. P. S. Silva, J. D. Procunir, and P. B. Little. 1993. Genomic fingerprinting of *Haemophilus somnus* isolates using a random-amplified polymorphic DNA assay. *J. Clin. Microbiol.* **31**:512–517.
54. Newbury, S. F., N. H. Smith, E. C. Robinson, I. D. Hiles, and C. F. Higgins. 1987. Stabilization of translationally active mRNA by prokaryotic REP sequences. *Cell* **48**:297–310.
55. Olsen, G. J., and C. R. Woese. 1993. Ribosomal RNA: a key to phylogeny. *FASEB J.* **7**:113–123.
56. Poh, C. L., V. Ramachandran, and J. W. Tasall. 1996. Genetic diversity of



- Neisseria gonorrhoeae* IB-2 and IB-6 isolates revealed through whole-cell repetitive element sequence-based PCR. *J. Clin. Microbiol.* **34**:292–295.
57. Poh, C. L., C. C. Yeo, and L. Tay. 1992. Genome fingerprinting by pulsed-field gel electrophoresis and ribotyping to differentiate *Pseudomonas aeruginosa* serotype 011 strains. *Eur. J. Clin. Microbiol. Infect. Dis.* **11**:817–822.
  58. Prevost, G., B. Jaulhac, and Y. Piedmont. 1992. DNA fingerprinting by pulsed-field electrophoresis is more effective than ribotyping in distinguishing among methicillin-resistant *Staphylococcus aureus* isolates. *J. Clin. Microbiol.* **30**:967–973.
  59. Ralph, D., M. McClelland, J. Welsh, G. Baranton, and P. Perolat. 1993. *Leptospira* species categorized by arbitrarily primed polymerase chain reaction (PCR) and by mapped restriction polymorphisms in PCR-amplified rRNA genes. *J. Bacteriol.* **175**:973–981.
  60. Reboli, A. C., E. D. Houston, J. S. Monteforte, C. A. Wood, and R. J. Hamill. 1994. Discrimination of epidemic and sporadic isolates of *Acinetobacter baumannii* by repetitive element PCR-mediated DNA fingerprinting. *J. Clin. Microbiol.* **32**:2635–2640.
  61. Robert, F., F. Lebreton, M. E. Bounoux, A. Paugam, D. Wassermann, M. Schlotterer, C. Tourte-Schaefer, and J. Dupouy-Camet. 1995. Use of random amplified polymorphic DNA as a typing method for *Candida albicans* in epidemiological surveillance of a burn unit. *J. Clin. Microbiol.* **33**:2366–2371.
  62. Rodriguez-Barradas, M. C., R. J. Hamill, E. D. Houston, P. R. Georghiou, J. E. Clarridge, R. L. Regnery, and J. E. Koehler. 1995. Genomic fingerprinting of *Bartonella* species by repetitive element PCR for distinguishing species and isolates. *J. Clin. Microbiol.* **33**:1089–1093.
  63. Saulnier, P., C. Bourneix, G. Prevost, and A. Andreumont. 1993. Random amplified polymorphic DNA assay is less discriminant than pulsed-field gel electrophoresis for typing strains of methicillin-resistant *Staphylococcus aureus*. *J. Clin. Microbiol.* **31**:982–985.
  64. Schlichting, C., C. Branger, and J.-M. Fournier. 1993. Typing of *Staphylococcus aureus* by pulsed-field gel electrophoresis, zymotyping, capsular typing, and phage typing: resolution of clonal relationships. *J. Clin. Microbiol.* **31**:227–232.
  65. Schwartz, D. C., and C. R. Cantor. 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37**:67–75.
  66. Sharples, G. J., and R. G. Lloyd. 1990. A novel repeated DNA sequence located in the intergenic regions of bacterial chromosomes. *Nucleic Acids Res.* **18**:6503–6508.
  67. Shortridge, V. D., G. G. Stone, R. K. Flamm, J. Beyer, J. Versalovic, D. W. Graham, and S. K. Tanaka. 1997. Molecular typing of *Helicobacter pylori* isolates from a multicenter U.S. clinical trial by *ureC* restriction fragment length polymorphism. *J. Clin. Microbiol.* **35**:471–473.
  68. Snelling, A. M., P. Gerner-Smidt, P. M. Hawkey, J. Heritage, P. Parnell, C. Porter, A. R. Bodenham, and T. Inglis. 1996. Validation of use of whole-cell repetitive extragenic palindromic sequence-based PCR (rep-PCR) for typing strains belonging to the *Acinetobacter calcoaceticus*-*Acinetobacter baumannii* complex and application of the method to the investigation of a hospital outbreak. *J. Clin. Microbiol.* **34**:1193–1202.
  69. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503–517.
  70. Sreevatsan, S., J. B. Bookout, F. M. Ringpis, M. R. Pottathil, D. J. Marshall, M. De Arruda, C. Murvine, L. Fors, R. Pottathil, and R. R. Barathur. 1998. Algorithmic approach to high-throughput molecular screening for alpha interferon-resistant genotypes in hepatitis C patients. *J. Clin. Microbiol.* **36**:1895–1901.
  71. Stern, M. J., G. F. L. Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins. 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. *Cell* **37**:1015–1026.
  72. Stull, T. L., J. J. LiPuma, and T. D. Edlind. 1988. A broad-spectrum probe for molecular epidemiology of bacteria: ribosomal RNA. *J. Infect. Dis.* **157**:280–286.
  73. Tenover, F. C., R. D. Arbeit, G. Archer, J. Biddle, S. Byrne, R. Goering, W. Hancock, G. A. Hébert, B. Hill, R. Hollis, W. R. Jarvis, B. Kreiswirth, G. Eisner, J. Maslow, L. K. McDougal, J. M. Miller, M. Mulligan, and M. A. Pfaller. 1994. Comparison of traditional and molecular methods of typing isolates of *Staphylococcus aureus*. *J. Clin. Microbiol.* **32**:407–415.
  74. Tenover, F. C., R. D. Arbeit, and R. V. Goering. 1997. How to select and interpret molecular strain typing methods for epidemiological studies of bacterial infections: a review for healthcare epidemiologists. *Molecular Typing Working Group of the Society for Healthcare Epidemiology of America. Infect. Control Hosp. Epidemiol.* **18**:426–39.
  75. Tenover, F. C., R. D. Arbeit, R. V. Goering, P. A. Mickelsen, B. E. Murray, D. H. Persing, and B. Swaminathan. 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J. Clin. Microbiol.* **33**:2233–2239.
  76. Tram, C., M. Simonet, M.-H. Nicolas, C. Offredo, F. Grimont, M. Lefevre, E. Ageron, A. Debure, and P. A. D. Grimont. 1990. Molecular typing of nosocomial isolates of *Legionella pneumophila* serogroup 3. *J. Clin. Microbiol.* **28**:242–245.
  77. Troesch, A., H. Nguyen, C. G. Miyada, S. Desvarrenne, T. R. Gingeras, P. M. Kaplan, P. Cros, and C. Mabilat. 1999. *Mycobacterium* species identification and rifampin resistance testing with high-density DNA probe arrays. *J. Clin. Microbiol.* **37**:49–55.
  78. Tsang, A. Y., J. C. Denner, P. J. Brennen, and J. K. McClatchey. 1992. Clinical and epidemiological importance of typing *Mycobacterium avium* complex isolates. *J. Clin. Microbiol.* **30**:479–484.
  79. Van Belkum, A., J. De Jonckheere, and W. G. V. Quint. 1993. Typing *Legionella pneumophila* strains by polymerase chain reaction-mediated fingerprinting. *J. Clin. Microbiol.* **31**:2198–2200.
  80. Versalovic, J., T. Koeuth, and J. R. Lupski. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. *Nucleic Acids Res.* **19**:6823–6831.
  81. Versalovic, J., V. Kapur, E. O. Mason, Jr., U. Shah, T. Koeuth, J. R. Lupski, and J. Musser. 1993. Penicillin-resistant *Streptococcus pneumoniae* strains recovered in Houston: identification and molecular characterization of multiple clones. *J. Infect. Dis.* **167**:850–856.
  82. Versalovic, J., V. Kapur, T. Koeuth, G. H. Mazurek, T. Whittam, J. Musser, and J. R. Lupski. 1995. DNA fingerprinting of pathogenic bacteria by fluorophore-enhanced repetitive sequence-based polymerase chain reaction. *Arch. Pathol. Lab. Med.* **119**:23–29.
  83. Vila, J., M. A. Marcos, and M. T. Jimenez de Anta. 1996. A comparative study of different PCR-based DNA fingerprinting techniques for typing of the *Acinetobacter calcoaceticus*-*A. baumannii* complex. *J. Med. Microbiol.* **44**:482–489.
  84. Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Muiper, and M. Zabeau. 1995. AFLP: a new concept for DNA fingerprinting. *Nucleic Acids Res.* **21**:4407–4414.
  85. Webster, C. A., K. J. Towner, H. Humphreys, B. Ehrenstein, D. Hartung, and H. Grundmann. 1996. Comparison of rapid automated laser fluorescence analysis of DNA fingerprints with four other computer-assisted approaches for studying relationships between *Acinetobacter baumannii* isolates. *J. Med. Microbiol.* **44**:185–194.
  86. Welsh, J., and M. McClelland. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* **18**:7213–7218.
  87. Williams, J. G., A. R. Kubelik, K. J. Livak, J. A. Rafalsky, and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful genetic markers. *Nucleic Acids Res.* **18**:6531–6535.
  88. Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.
  89. Woods, C. R., J. Versalovic, T. Koeuth, and J. R. Lupski. 1992. Analysis of relationships among isolates of *Citrobacter diversus* by using DNA fingerprints generated by repetitive sequence-based primers in polymerase chain reaction. *J. Clin. Microbiol.* **30**:2921–2929.
  90. Woods, C. R., J. Versalovic, T. Koeuth, and J. R. Lupski. 1993. Whole-cell repetitive element sequence-based polymerase chain reaction allows rapid assessment of clonal relationships of bacterial isolates. *J. Clin. Microbiol.* **31**:1927–1931.
  91. Woods, C. R., T. Koeuth, M. M. Estabrook, and J. R. Lupski. 1996. Rapid determination of outbreak-related strains of *Neisseria meningitidis* by repetitive element-based polymerase chain reaction genotyping. *J. Infect. Dis.* **174**:760–767.
  92. Yang, Y., and G. F. L. Ames. 1988. DNA gyrase binds to the family of prokaryotic repetitive extragenic palindromic sequences. *Proc. Natl. Acad. Sci. USA* **85**:8850–8854.
  93. Zee-Jordans, J., J. Paul, J. Bates, C. Beaumont, J. Kimari, and C. Gilks. 1995. Characterization of *Streptococcus pneumoniae* from human immunodeficiency virus-seropositive patients with acute recurrent pneumonia. *J. Infect. Dis.* **172**:983–987.