# 13

# Genetic techniques: molecular subtyping methods

**C. Fitzgerald and B. Swaminathan, Centers for Disease Control and Prevention, USA and A. Sails, Newcastle General Hospital, UK**

## 13.1 Introduction

The ability to discriminate or subtype foodborne pathogens below the level of species has been applied successfully to aid the epidemiological investigation of outbreaks of gastrointestinal disease. Reliable, sensitive and informative subtyping methods are required to recognize outbreaks of infection, match case isolates with those from potential vehicles of infection and discriminate these from unrelated strains. Sensitive and discriminatory subtyping methods also are required for surveillance programs to identify new or emergent strains or clones that may present a new risk to public health. These methods also are fundamental to epidemiological research projects to identify potential reservoirs of strains that cause disease in humans, identify routes of transmission and improve our understanding of the epidemiology of foodborne disease (Swaminathan *et al.*, 2001).

The criteria used to evaluate subtyping methods include:

- *typability*: the ability of the method to produce an unambiguous result for all strains of a particular species
- *reproducibility*: the ability of the method to produce the same result when a particular strain is tested repeatedly (this includes both intra- and inter-laboratory reproducibility)
- *discriminatory power*: how well the method discriminates between unrelated strains of the same species
- *ease of use*: technical difficulty and expertise necessary to perform the test
- *ease of interpretation*: how easily results can be interpreted and compared between laboratories
- *time and cost factors*: time taken to obtain a result and the cost per isolate.

In the absence of a definitive 'gold standard', new subtyping methods must be vigorously evaluated using sets of well-characterized isolates whose relatedness has been unequivocally determined previously.

Phenotypic subtyping methods include biotyping, which discriminates between strains on the basis of specific biochemical reactions, and their ability to grow in the presence of certain chemicals, pH gradients, temperatures or gaseous atmospheres. Biotyping is often used to identify isolates to species level, but the method has poor discriminatory power and a limited ability to differentiate between isolates below the level of species. Serotyping, which is based on the reaction of antigenic determinants on the cell surface with specific antibodies or antisera, is relatively rapid and technically easy to perform, and is considered to be one of the classic tools for the epidemiological discrimination of organisms such as *Salmonella, Escherichia coli* and *Listeria monocytogenes*. Certain serotypes have well-established associations with particular clinical syndromes, such as the association between haemolytic-uremic syndrome and *E. coli* O157:H7.

Phage typing is based on the presence or absence of a phage receptor on the exposed bacterial surface and has been used to subtype a number of foodborne pathogens, including *Salmonella*, *L. monocytogenes*, *Campylobacter jejuni* and *E. coli* O157:H7. For some organisms phage typing has poor reproducibility, and typability can range from good to moderate with significant numbers of strains lacking phage receptors and being untypable. Despite this, phage typing notably has been of value for the phenotypic characterization of *S*. Enteritidis, *S*. Typimurium and *E. coli* O157:H7. Phage typing also requires constant quality control of the phage suspensions and significant expertise to perform and interpret the results. These factors often limit its use to reference laboratories.

Phenotypic methods such as serotyping and phage typing have been used effectively for the subtyping of many foodborne pathogens, and for some pathogens many years of historical data have been generated. However, these methods do have limitations. Many phenotypic methods are not universally applicable and are useful only for the species for which they were developed. Variability in gene expression and the acquisition or loss of DNA-carrying genes, which encode phenotypic traits, can lead to changes in the phenotype displayed by the strain. This can lead to closely-related strains exhibiting different phenotypic traits and unrelated strains exhibiting indistinguishable subtypes. In addition, serotyping and phage typing require the production and maintenance of large panels of reagents, which is time-consuming and laborious. For example, approximately 350 antisera are required to detect the 2,523 recognized serotypes of *Salmonella*, therefore limiting identification of all serotypes to reference laboratories. In addition, specific reagents may not be available for some serotypes or phage types, leading to significant numbers of strains being untypable using these methods.

The limitations of phenotypic subtyping methods and the rapid growth of molecular biological techniques have led to the development of a range of molecular subtyping methods. Molecular subtyping methods target genotypic variation within the DNA sequence of the organism, which may reduce or

eliminate problems encountered with untypability, theoretically providing 100 per cent typability. Molecular approaches are often both more universally applicable and useful for a wide range of organisms. Plasmid profiling was the first molecular typing to be described. However, many strains may contain no plasmids, or may gain or lose plasmids, making this method an unreliable marker for routine surveillance. One of the most useful molecular subtyping techniques to emerge in the last ten years is pulsed-field gel electrophoresis of macrorestricted chromosomal DNA (PFGE). Rare cutting restriction enzymes cleave the DNA into 10–800 kb fragments, which are then separated by gel electrophoresis to produce a PFGE 'fingerprint' containing between 5 and 30 bands. This technique is highly discriminatory and is a proven tool for epidemiological investigations (Swaminathan *et al.*, 2001).

Some molecular methods are based on the polymerase chain reaction (PCR), which provides results in less than a day. Many of the currently available molecular subtyping methods rely on the electrophoretic separation of DNA fragments of different sizes using gel electrophoresis. The resulting pattern of bands can be complex and difficult to interpret, and often provide results that are difficult to compare between laboratories or individual gels. The rapid progress of DNA sequencing technology has made this method increasingly more available and cost-effective. This in turn has led to the sequencing of the complete genome of a number of foodborne pathogens and is facilitating the discovery of new targets for molecular subtyping methods, which may provide more precise information on strain relatedness. A comparison of the utility of phenotypic and genotypic subtyping methods for *C. jejuni* and *L. monocytogenes* is presented in Table 13.1.

In this chapter we will review the most commonly used molecular typing methods for foodborne pathogens and describe their potential advantages and disadvantages for epidemiological typing.

## 13.2 Approaches to molecular subtyping

Since many laboratories have limited resources and time yet handle large numbers of strains, a single typing method to determine strain relatedness would be optimal in the absence of a currently available definitive 'gold standard' typing technique applicable to all foodborne pathogens. In the next section we describe some of the methods currently used for bacterial subtyping and discuss some of the advantages and limitations of each method. The subtyping method of choice for a particular foodborne pathogen is ultimately determined by the nature of the microbiological question being asked.

### 13.2.1 Restriction endonuclease analysis

One of the first techniques to be described for molecular subtyping of bacteria species was the restriction endonuclease analysis (REA) of chromosomal

**Table 13.1** Comparison of the utility of phenotypic and genotypic subtyping methods for *C. jejuni* and *L. monocytogenes*

| Method and organism | Typability (%) | Reproducibility | Discriminatory power (DI)[c] | Ease of use | Ease of interpretation | Time and costs |
|---|---|---|---|---|---|---|
| *C. jejuni*[a] | | | | | | |
| Serotyping | ~80 | Good | Average | Fairly simple | Fairly simple | <1 day, low cost |
| Phage typing | ~80 | Fair | Fair | Fairly simple | Requires expertise | 2 days, lack of availability of phages |
| RAPD | ~80 | Low | Average | Fairly simple | Difficult | <1 day, low cost |
| AFLP | 100 | Good | Good | Complex method | Complex | 2–3 days, moderate cost |
| Ribotyping | 100 | Good | Poor | Complex method | Difficult | 3-4 days, moderate cost |
| PFGE | 100 | Good | Good | Moderately complex | Requires normalization of data | 2 days, moderate cost |
| *L. monocytogenes*[b] | | | | | | |
| Serotyping | 100 | 82–100 | 0.68 | Fairly simple | Simple | 1 day, low cost |
| Phage typing | 49–80 | 79 | ND[d] | Fairly simple | Requires expertise | 2 days, problems with availability of phages |
| RAPD | 100 | Range 0–100 (median 86.5) | 0.75-0.95 (for 3 primers) | Fairly simple | Fairly simple | 1 day, low cost |
| Ribotyping | 100 | 80-100 | 0.83-0.88 | Difficult | Fairly simple | 2 days, moderate cost |
| PFGE | 100 | 84 | 0.95-0.96 | Fairly simple | Requires normalization of data | 2 days, moderate cost |

[a] Modified from Wassenaar and Newell (2000). [b]Modified from Graves *et al*. (1999). [c] DI: Simpson's Index of Diversity; [d] ND: Not determined.

DNA. This technique is universally applicable, rapid, inexpensive and relatively easy to perform. However, the multi-banding patterns produced are complex and can be difficult to analyze, making REA unsuitable for the comparison of large numbers of isolates. In addition, the presence of plasmid DNA in some strains can affect the profile produced and further complicate interpretation of the data.

### 13.2.2 Plasmid profile analysis

Plasmid profiling emerged in the 1980s as one of the first DNA-based typing methods to be applied to epidemiological studies. The method is based on the isolation and then separation of intact plasmids by agarose gel electrophoresis. REA can be used to provide further evidence of the similarities and differences between strains. Plasmid profile analysis is a simple yet powerful tool and has proven to be useful when used in conjunction with other subtyping data in epidemiological investigations of a number of foodborne pathogens, especially when the strains being studied are indistinguishable by other methods. A review by Mayer (1988) provides a summary of the utility of plasmid profiles. For some foodborne pathogens, plasmid profile analysis has proven to be an effective typing method, particularly for certain serotypes of *Salmonella*. However, such analysis has been of limited value for other foodborne pathogens, such as *C. jejuni*, in which only about 20 to 30 per cent of isolates carry plasmid DNA.

Although plasmid profile analysis provides a sensitive and specific epidemic strain marker when present, not all strains carry plasmids. In addition, strains may readily acquire or lose plasmids, and thus the results of plasmid profiling should always be interpreted with caution. Despite these limitations, the continuing emergence of antimicrobial resistance has led to renewed interest in the molecular characterization of plasmids. Advantages of plasmid profile analysis include the use of a single set of reagents and equipment, which are applicable to many bacterial pathogens. The technique is relatively rapid, technically simple and relatively inexpensive to perform.

### 13.2.3 Hybridization techniques

Ribosomal RNA restriction analysis (rRNA) is a technique that examines restriction fragment length polymorphisms (RFLP) associated with the ribosomal operon(s). The presence of multiple copies of the rRNA genes and their highly conserved nature among bacteria make them appropriate targets for typing purposes. Chromosomal DNA is digested with the appropriate restriction endonuclease and the fragments are separated using conventional agarose gel electrophoresis as in REA. The resulting complex DNA patterns are made easier to interpret by Southern blot hybridization with a probe specific for rRNA genes to generate a ribopattern. Because the genes encoding for rRNA are highly conserved, ribotyping of many foodborne pathogens is performed using an appropriately labeled 16S and 23S ribosomal RNA (rRNA) or rDNA general-

purpose probe from *E. coli*. In addition, the use of probes specifically designed based on type strains of the organism being examined have also been described. For example, an intragenic probe, generated by PCR from *C. jejuni* (NCTC 11168) was developed to examine polymorphisms around the 16S rRNA gene of this organism. This generates simplified ribotypes related to gene copy number.

Ribotyping has been used extensively in molecular epidemiological investigations of many foodborne pathogens including *C. jejuni*, *L. monocytogenes*, *Salmonella* sp., and *Vibrio* sp. (Hunter and Swaminathan, 1998). The main disadvantage of ribotyping methods is the multiple steps involved that make it labor intensive and time-consuming. However, reproducibility and 100 per cent typability, plus automation of the technique (RiboPrinter® system-Dupont Qualicon, see Section 13.8), have made it more accessible. The discriminatory power of the method is limited and different organisms have different ribosomal gene copies, so the suitability of the method for routine subtyping is somewhat organism-specific. For example, there are only three copies of the rRNA genes within the *C. jejuni* genome, which reduces the discriminatory power of the method for this organism. The use of a combination of restriction enzymes has been reported to increase the discriminatory power of the method (Hunter and Swaminathan, 1998). Ribotyping had proven to be a valuable epidemiological tool for *L. monocytogenes*, but its ability to discriminate within serogroup 4b strains may not be adequate for epidemiological investigations (Graves *et al.*, 1999). Figure 13.1 shows an example of RiboPrint® patterns of *Listeria monocytogenes* isolates generated using the DuPont Qualicon RiboPrinter® system. For many foodborne pathogens, ribotyping should be used in conjunction with other methods, such as PFGE. The development of the RiboPrinter® system has renewed interest in the use of ribotyping as a subtyping tool, and provides a standardized system that facilitates the interlaboratory exchange of data. However, the cost of equipment and consumables is substantial and sample throughput is small. In addition, ribotyping has been largely superseded by higher resolution subtyping techniques such as PFGE and amplified fragment length polymorphism (AFLP) typing.

### 13.2.4 Other probes

Southern blot analyses in which insertion sequences are used as probes, such as IS200 profiling for *Salmonella*, have proven to be a reproducible and moderately discriminatory method for strain identification. RFLP analysis of toxin- or virulence-associated genes has been a valuable tool for subtyping several bacterial species, including the shiga-like toxin (SLT) I and II structural genes and bacteriophage lambda ($\lambda$-RFLP) genes of *E. coli* O157:H7 and the cholerae toxin genes (*ctx*) of *V. cholerae*. The use of these labor intensive and time-consuming RFLP-Southern blotting techniques has been largely replaced by PCR-based RFLP analysis, which will be discussed later in this chapter.

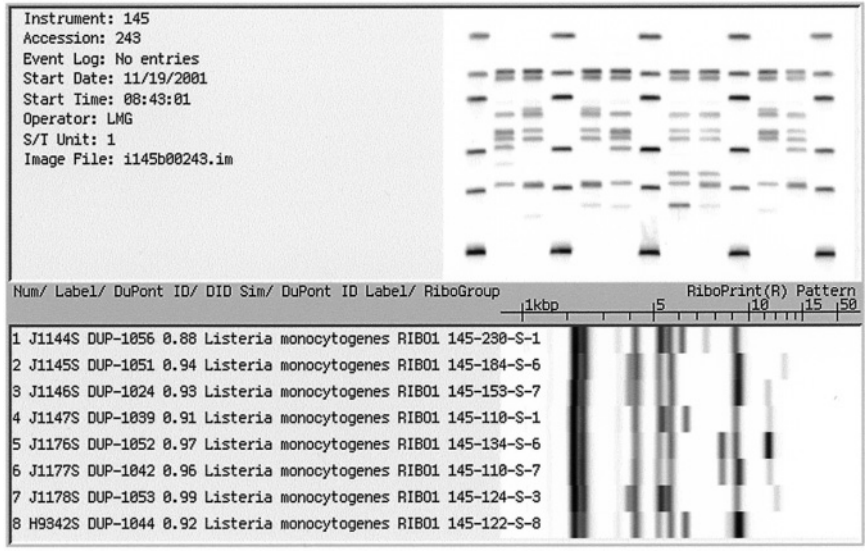


**Fig. 13.1** RiboPrint® patterns of *Listeria monocytogenes* isolates generated using DuPont Qualicon RiboPrinter® system. The unprocessed TIFF image (top right) shows ribotype patterns of 8 *L. monocytogenes* isolates (lanes 2,3,5,6,8,9,11,12). A molecular size standard is in lanes 1,4,7 and 10. The RiboGroup patterns and RiboGroup IDs of the patterns are displayed in the bottom panel of the RiboPrinter® output.

### 13.2.5 PFGE of macrorestricted chromosomal DNA

A major limitation of REA is the complexity of RFLP patterns that are generated, making it difficult to analyze the large number of overlapping, poorly resolved restriction fragments produced. DNA macrorestriction analysis by PFGE is an RFLP-based method that utilizes rare cutting restriction enzymes to cut chromosomal DNA into fewer, larger restriction fragments. Special electrophoretic conditions are used to separate these large DNA fragments. To prevent DNA shearing, bacterial cells are embedded in agarose 'plugs' and lysis of the cell wall, and digestion of cellular proteins is performed *in situ*. After washing to remove cellular debris, thin slices of the plugs containing DNA are cut and then incubated in the presence of the restriction enzyme of choice. Following restriction digestion, the DNA fragments are separated according to size by a variation of agarose gel electrophoresis in which the orientation of the electric field is changed in a pulsed manner. The resulting macrorestriction PFGE profiles are typically composed of between 5 and 30 well-resolved fragments ranging in size from approximately 10 to 800 kb, depending on the organism and the restriction enzyme used. Examples of some of the commonly used restriction enzymes for the subtyping of foodborne bacteria are presented in Table 13.2.

Since the development of PFGE as an epidemiological tool in the mid-1980s, the technique has been applied to a wide range of organisms and has become one of the most useful subtyping tools currently available for foodborne molecular epidemiology. The technique has been demonstrated to be highly reproducible

**Table 13.2** Commonly used restriction endonuclease enzymes for the subtyping of foodborne and enteric pathogens by PFGE

| Organism | Restriction endonuclease |
|---|---|
| *Campylobacter jejuni* | *Sma*I[a] *Kpn*I[b] *Sac*II *Sal*I |
| *Escherichia coli* O157:H7 | *Xba*I[a] *Bln*I[b] *Sfi*I *Swa*I *Spe*I *Not*I |
| *Listeria monocytogenes* | *Asc*I[a] ApaI[b] *Sma*I *Not*I |
| *Vibrio* species | *Not*I[a] *Sfi*I[b] *Cpo*I *Bgl*I |
| *Shigella sonnei* | *Xba*I[a] *Sfi*I *Not*I |
| *Salmonella* spp. | *Xba*I[a] *Bln*I[b] *Spe*I *Not*I |
| *Clostridium perfringens* | *Sma*I[a] *Apa*I[b] |

[a] Primary enzyme of choice.
[b] Secondary enzyme of choice.

and discriminatory. The relative simplicity of the macrorestriction profiles produced greatly facilitates the analysis and comparison of multiple isolates. In most cases PFGE has proven to be more sensitive than other subtyping methods, and is currently the subtyping method of choice for a number of foodborne pathogens, including *E. coli* O157:H7, *L. monocytogenes*, *Salmonella* and *C. jejuni*. Although the discriminatory power of PFGE is excellent, it does have a number of limitations. The equipment used for the electrophoresis is specialized and relatively expensive, the method is more laborious and time-consuming than others and it is not amenable to automation. In addition, interpretation of the results can be difficult since genetic instability can lead to changes in PFGE profiles. Such difficulties are not unique to PFGE and will be discussed later. However, the development of rapid one-day standardized PFGE protocols, which are used by participants of the PulseNet national surveillance network for foodborne pathogens (www.cdc.gov/pulsenet), has facilitated the widespread use of this subtyping method for a growing number of foodborne pathogens. Figure 13.2 shows an example of the PFGE separation of *Xba*I macrorestricted fragments of *E. coli* O157:H7 genomic DNA using the PulseNet *E. coli*. O157:H7 standardized PFGE protocol.

## 13.3 PCR-based techniques

Advances in PCR technology led the way for the emergence of PCR-based subtyping techniques in the 1990s. In general, these are relatively simple, rapid and broadly applicable typing methods that are available to any laboratory with PCR capabilities.

### 13.3.1 Random amplification of polymorphic DNA

Random amplified polymorphic DNA (RAPD) and arbitrarily primed polymerase chain reaction (AP-PCR) use single primers of arbitrary nucleotide

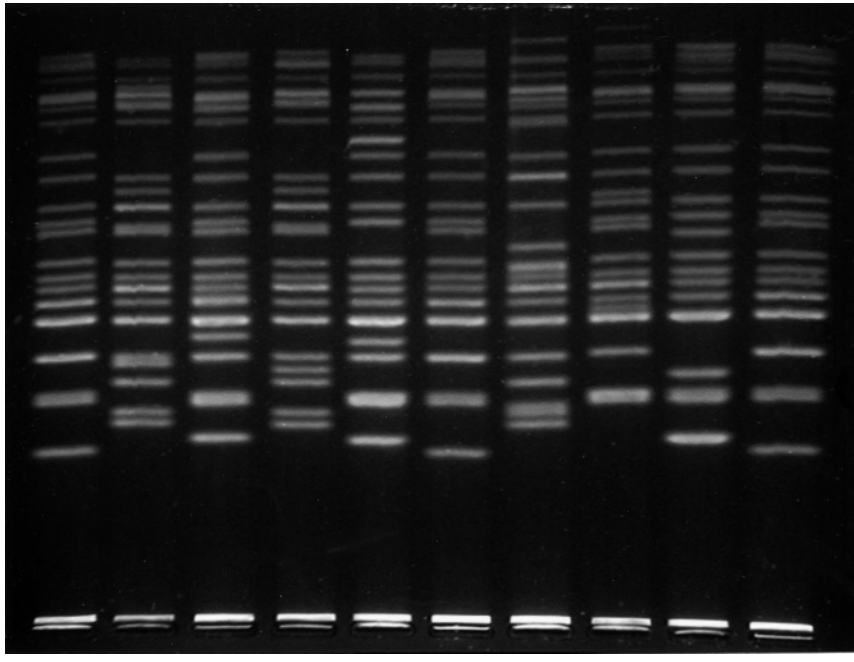

**Fig. 13.2** PFGE separation of *Xba*I macrorestricted fragments of *E. coli* O157:H7 genomic DNA using the PulseNet *E. coli* O157:H7 standardized PFGE protocol. Lanes 1, 5 and 10 contain PulseNet standard G5244.

sequence under low stringency PCR conditions. The method targets the whole genome of the organism to generate randomly amplified DNA products, with no prior knowledge of the target DNA sequence required. Typically, 10-mer primers (some of which are commercially available, see Section 13.8) are used for RAPD, whereas AP-PCR was developed with longer primers. RAPD has been used for subtyping and strain comparison for many common foodborne pathogens. The technique is technically simple, rapid, highly discriminatory and has been demonstrated to provide discrimination close if not equal to that of PFGE. However, the method suffers from significant problems with reproducibility that have restricted its widespread use. A number of parameters have been shown to have an effect on the appearance and reproducibility of the fingerprints produced; therefore optimization of PCR reaction conditions is essential (Holmberg and Feroze, 1996). The development of commercially-available ready-to-go RAPD beads (see Section 13.8) is one approach available to help reduce problems associated with optimization if cost is not a factor. However, the lack of standardization of the method, complexity of the patterns produced and the inherent difficulties in the interpretation of weak bands currently limit the widespread use of RAPD as a subtyping method for epidemiology.

### 13.3.2 Amplification of repetitive DNA sequences

Another approach to amplifying genomic DNA fragments involves using primers directed at repetitive DNA elements present within bacterial genomes (rep-PCR). Two main sets of repetitive elements are used for typing purposes: the 38-bp repetitive extragenic palindromic (REP) sequences, and the 126-bp enterobacterial repetitive intergenic consensus (ERIC) sequences (Hulton *et al.*, 1991). These motifs are genetically stable and differ only in chromosomal location and copy number between species, making them good targets for strain differentiation. REP and ERIC amplifications can be performed either with a single primer or multiple sets of primers. ERIC patterns are generally less complex than REP patterns, but both give good discrimination at the strain level. Another repetitive element, the 154-bp BOX sequence, which was initially thought to be unique to *S. pneumoniae*, has now been found in a number of other bacterial species and used for subtyping an increasing number of pathogens, including *Salmonella* and *E. coli*. PCR amplification of insertion sequences (IS), such as IS200 for some serotypes of *Salmonella* and IS3 for *E. coli* O157, has also been reported. These PCR-based methods are easy to perform and rapid. Theoretically, as REP-PCR is performed under higher stringency reaction conditions than RAPD, the patterns generated should be more reproducible. However, in practice inter-laboratory reproducibility still remains a problem. Attempts to overcome this include the use of elevated annealing temperatures and touchdown PCR. For many pathogens these typing methods produce results that correlate with PFGE, but generally they have slightly less discriminatory power. Unfortunately, lack of standardization of these PCR approaches has limited their widespread use.

### 13.3.3 Polymorphic genes used for RFLP analysis

For PCR-RFLP analysis to be a useful subtyping approach, a target containing adequate polymorphism to allow subspecies discrimination must be identified. The locus of interest is amplified by PCR with gene-specific primers; therefore, prior knowledge of the DNA sequence is necessary. The amplified PCR product is then subjected to digestion with an appropriate restriction endonuclease followed by separation of the DNA fragments by gel electrophoresis to produce the RFLP profile. The profiles produced are highly reproducible, but the discriminatory power of the method is dependent on the selection of the restriction endonuclease. A number of PCR-RFLP target sequences have been reported, many of which are organism-specific. Table 13.3 summarizes some of the commonly used loci for PCR-RFLP analysis; some examples of targets include the coagulase (*coa*) gene in *Staphylococcus aureus* (Shopsin *et al.*, 1999) and the flagellin gene (*fli*C) in *E. coli* (Machado *et al.*, 2000). A more universal approach, using the 16S, 23S and 16S-23S spacer region as PCR-RFLP targets, has also been shown to be useful. Although PCR-RFLP is a reliable and relatively simple subtyping method, it indexes variation within a limited segment of the genome, which may not be representative of the entire genome. This can reduce its discriminatory potential and complicate

**Table 13.3** Commonly used gene targets for PCR-RFLP subtyping of foodborne pathogens

| Organism | Target gene | Restriction enzyme(s) |
|---|---|---|
| *Campylobacter jejuni* | flagellin A (*fla*A) | *Dde*I *Hinf*I *Alu*I |
| *Escherichia coli* O157:H7 | flagellin gene (*fli*C) | *Hha*I |
| *Listeria monocytogenes* | *inl*A and *inl*B genes | *Alu*I |
| *Salmonella enteritidis* | flagellin (*fli*C, *flj*B) | *Hha*I and *Hph*I |
| *Salmonella enteritidis* | *rec*A gene | *Hha*I and *Sau*3AI |
| *Staphylococcus aureus* | coagulase (*coa*) | *Alu*I or *Hae*III |

interpretation of the data produced. Combining multiple polymorphic genes in a multiplex PCR has been reported to increase discriminatory power. In addition, an appreciation of the genetic stability of the marker is essential when interpreting data. For example, PCR-RFLP analysis of the flagellin (*fla*A) has been in widespread use over the last decade for the typing of several species of *Campylobacter* and has proven to be a useful epidemiological tool, but the flagellin locus has been demonstrated to undergo recombinational events. Consequently, although this method is a useful tool when the aim is to determine a precise epidemiological link in a well-defined setting (same time and place), it is unsuitable for global or long-term longitudinal epidemiological studies.

### 13.3.4 PCR-single strand conformation polymorphism typing (SSCP)

This technique is based on the single strand conformation polymorphism (SSCP) electrophoresis of PCR-amplified fragments. After PCR amplification of a specific sequence, the product is denatured and subjected to non-denaturing polyacrylamide gel electrophoresis. The structures formed by single-stranded (ss) DNA under these conditions display mobility shifts due to conformational changes as a result of nucleotide substitutions. Like PCR-RFLP analysis, PCR-SSCP requires amplification of a specific target that contains sufficient polymorphisms. Therefore, the same targets and primers used for PCR-RFLP analysis can also be used for PCR-SSCP analysis. Although this approach is not currently widely used, the main advantage of PCR-SSCP is that it can detect DNA polymorphisms and point mutations at a variety of positions in the ssDNA fragment as opposed to using restriction enzymes that detect one specific mutation. Modifications of the method include the use of fluorescent-labeled primers and an automated sequencer, which add a higher level of resolution and reproducibility.

## 13.4 AFLP analysis and emerging methods

AFLP analysis was originally developed for the genetic analysis of plants, and since then has been adapted for subtyping a number of bacteria, including

foodborne pathogens (Savelkoul *et al.*, 1999) This technique combines the reliability of RFLP with the advantages of PCR to sample a random portion of the whole genome. The technique is based on the selective amplification of a subset of DNA fragments generated by digestion of chromosomal DNA with two restriction enzymes. Specific oligonucleotide adapters are then ligated to these restriction sites. The adapters are designed so the initial restriction site is not restored after ligation, such that they create a template sequence for subsequent high stringency PCR amplifications. Adapter specific primers used for amplification were originally radioactively labeled when the method was first described, but the switch to fluorescently labeled primers (FAFLP) has since become the standard approach for detection of the PCR products in an automated sequencer. Typically, fragments that are 50 to 500 nucleotides long are separated, and a pattern of 40 to 200 bands is obtained. Incorporating into the PCR primers one or more specific nucleotides adjacent to the restriction site can reduce the number of bands generated. The use of internal lane standards corrects for variation between gels. Modifications of AFLP include the use of a single enzyme with a single adapter and analysis by agarose gel electrophoresis, and the use of a single PCR amplification with one and two selective nucleotides, respectively, on both primers.

AFLP is not dependent on prior sequence information and can be adapted to any foodborne pathogen. However, optimization of the restriction enzymes and adjacent specific nucleotides used is necessary for each species. AFLP has become established as a broadly applicable subtyping tool with other applications in taxonomy and diagnostics. Its discriminatory power is high and has been demonstrated to at least equal that of PFGE for a number of foodborne pathogens, including salmonellae. The discriminatory power of the technique can be varied systematically by using primers of specified selectivity; therefore the technique can be tailored to provide the level of discrimination required. AFLP methods can be especially useful for providing further discrimination when isolates appear apparently indistinguishable or clonal by other less discriminatory methods. AFLP has high resolution, is rapid and has a higher throughput when compared with other molecular methods for bacterial strain typing, although how differences in plasmid content influence AFLP profiles remains to be determined. The recent introduction of new multi-capillary instrumentation means that high throughput subtyping of isolates is now achievable. FAFLP profiles are suitable for rapid electronic exchange for inter-laboratory comparisons. However, recent AFLP data from the Campynet project (see Section 13.8) suggest that AFLP data comparisons between laboratories equipped with different separation and detection apparatus is problematic. Standardization of the method is also essential for inter-laboratory comparisons. However, such standardization has not been widely implemented for specific foodborne pathogens. Disadvantages include the need for high-quality DNA samples and a major capital investment (for an automated DNA sequencer and appropriate software).

### 13.4.1 Comparative DNA sequencing-based subtyping

DNA sequencing-based subtyping has emerged recently as a new subtyping method. This approach is becoming a viable alternative for genotyping bacterial isolates since the introduction of automated sequencers has made sequencing more rapid and the costs have continued to decrease in recent years. DNA sequence analysis is a highly reproducible method that does not rely on the interpretation of gel patterns. DNA sequencing also provides more precise information on strain relatedness, which is only suggested by PFGE. Comparative DNA sequencing of variable regions of a number of gene targets has been described for many foodborne pathogens. Some examples include the *rec*A, *ald*A and toxin genes (*ctx*) of *V. cholerae*, the flagellin gene (*fla*A) of *C. jejuni* and virulence-associated genes (*iap*, *inl*A and *hly*A) of *L. monocytogenes*.

### 13.4.2 Multilocus sequence typing (MLST)

One new typing method developed to take advantage of the advances in automated DNA sequencing is multilocus sequence typing (MLST) (for a review see Enright and Spratt, 1999). This typing method, first applied to *N. meningitidis*, is based on the sequencing of short 400–500 nucleotide sequences within seven housekeeping gene loci. It has emerged as a powerful tool for subtyping a number of bacterial species, including the foodborne pathogen *C. jejuni*. MLST is analogous to multilocus enzyme electrophoresis (MLEE), except that MLST indexes the variation within the housekeeping genes directly by DNA sequencing, whereas MLEE assigns alleles indirectly based upon the electrophoretic mobilities of their gene products on starch gels. The sequence data are 'portable' and can be readily compared between laboratories electronically, which facilitates the establishment of universal nomenclature systems and global databases for each pathogen, such as those at www.mlst.net. The combination of high discriminatory power and indexing variation at multiple neutral loci makes it suitable for longer term or global epidemiological studies, although this remains unproven. In addition, the DNA sequence data generated by MLST is suitable for analysis of the population genetics of the organism. However, MLST methods have been described only for *C. jejuni* and *Salmonella*, and the utility of MLST as an epidemiological tool for the investigation of outbreaks of foodborne disease has yet to be established.

### 13.4.3 Multilocus variable number tandem repeat analysis (MLVA)

Another recently described approach to subtyping of foodborne pathogens is multilocus variable number tandem repeat (VNTR) analysis (MLVA). MLVA has proven particularly useful for discriminating between isolates of highly clonal species, such as *Bacillus anthracis* (Keim *et al.*, 2000). MLVA takes advantage of the fact that many bacterial genomes are interspersed with short nucleotide sequences that are repeated multiple times. These short nucleotide

repeats often vary in copy number in different strains of a species, thus providing a means of discriminating between strains with exquisite sensitivity.

It is important to have prior knowledge of specific sequences that contain VNTRs on the bacterial genome of interest in order to develop an MLVA typing scheme. Keim *et al.* (2000) identified eight genetic loci that provide high levels of discrimination among *B. anthracis* isolates by a variety of approaches, including:

- sequencing of AFLP marker fragments
- examination of virulence plasmid sequences
- utilizing a previously described VNTR locus.

The increasing availability of whole genome sequences of bacteria will greatly facilitate the identification of VNTR loci in these bacteria and the design of specific PCR primers that amplify the regions containing the VNTRs. Appropriate fluorescent dye labels are incorporated into the PCR primers to facilitate automated genotype analysis on an automated fluorescent DNA sequencer.

Once useful VNTR loci are identified and PCR primers are designed, single or duplex PCR reactions are carried out to amplify the targets from bacterial cell preparations that have been subjected to a simple DNA preparation protocol (heat lysis followed by centrifugation to remove cell debris). Equal amounts of the PCR amplicons are combined and electrophoretically analyzed on an automated DNA sequencer. The sizes of the PCR amplicons typically range between 50 and 1000 bp. The MLVA technique has been applied to the subtyping of *Yersinia pestis* (Klevytska *et al.*, 2001) and *Francisella tularensis* (Farlow *et al.*, 2001). Preliminary results of MLVA subtyping of *E. coli* O157 indicate that MLVA may have a discriminatory power similar to that of PFGE (Keim, P., personal communication).

Although DNA-based subtyping is broadly applicable to any bacterial species, the choice of gene target and design of oligonucleotide primers is organism-specific and requires careful consideration. Loci with adequate sequence variability to permit epidemiologically useful strain differentiation must be identified. In addition, this DNA sequence must be present in all isolates and must have sufficient variability within a region that does not exceed the size constraints of the DNA that can be practically sequenced. Determination of the optimal panel of gene targets that provide the desired level of strain discrimination has yet to be identified for foodborne pathogens, though the combined use of conserved and variable genes, such as those associated with virulence, may maximize strain discrimination. In addition, DNA sequencing is still relatively expensive compared to other subtyping methods, and manipulation of the complex data generated is highly reliant on computer comparison software.

## 13.5 Standardized molecular subtyping of pathogens

Molecular subtyping moved the capacity for isolate characterization below the lowest officially recognized taxonomic levels (species and subspecies) from a few national and international public health laboratories to practically any microbiology laboratory with minimal equipment for molecular biology procedures. Though this was a welcome change, it had certain undesirable consequences as well. Before the advent of the application of molecular techniques for strain typing, available methods such as serotyping and bacteriophage typing were performed only in specialized laboratories. Often these methods were developed at one or more national public health laboratories and were standardized for use by a consortium of such laboratories. A common nomenclature was developed for use in designating strain types so that reference to a particular subtype (e.g., *Salmonella* sp. serotype Marina, *E. coli* O157:H7, *L. monocytogenes* 4b, *S.* Enteritidis phage type 4) would be universally understood. This standardization of methods and nomenclature facilitated international exchange of data and strains, allowed public health personnel to monitor changes in pathogen types and the emergence of new pathogen types, and allowed the tracking of the specific pathogen types across a region or the entire world. For example, *S.* Enteritidis phage type 4 was first seen in Europe, where it rapidly displaced the then-predominant phage types 8 and 13a. The same phenomenon began to occur in North America after a ten-year lag. Similarly, integron-mediated multi-drug resistance in *Salmonella* Typhimurium was first encountered in the United Kingdom, and the problem later surfaced in the United States. Thus, routine subtyping of pathogenic bacteria by the public health laboratories is extremely important for early recognition of the emergence of new, more virulent or resistant subtypes.

During the initial applications of molecular typing to foodborne pathogenic bacteria, no attempts were made to standardize protocols, data acquisition, or analysis. Every laboratory used its own custom protocol, making it impossible to compare results obtained in different laboratories. Even results obtained within the same laboratory could not be compared if the experiments were performed on different days or by different persons. Also, new methods of performing molecular typing were being developed at an extremely rapid pace, which further exacerbated the data comparison problems. Each laboratory had its own custom designations for subtypes; these designations were meaningless outside that laboratory. Thus, many of the advantages offered by these highly sensitive and discriminating molecular typing methods were negated by the lack of standardization and the inability to compare or exchange data. Three immediate needs were apparent:

(1) comparative evaluation of different subtyping methods for each pathogen and a quantitative assessment of each method
(2) standardization of the most useful methods and validation that results using standardized protocols enable comparison of data within and between laboratories
(3) development of a universal nomenclature for the subtypes.

During the 1980s and 1990s, several attempts were made to compare different typing methods available for specific pathogens. For foodborne pathogens, the best available example of this approach is the WHO-sponsored international collaborative study of subtyping methods for *L. monocytogenes* initiated in 1991 by Jocelyn Rocourt and Jacques Bille. The first phase of this study involved the evaluation of all phenotypic and genotypic subtyping methods that had been used by multiple investigators to subtype *L. monocytogenes*. An international set of 69 carefully selected strains was used for the evaluation. The strain set included 22 groups of epidemiologically related strains and two groups containing unrelated strains; all major outbreaks of listeriosis from 1981 to 1991 were represented. The strain set also included 11 duplicate strains (blinded to the investigators) to assess intra-laboratory reproducibility of typing results. Seven typing methods were evaluated, with two to seven investigators participating in each evaluation. The results of the study were published in a special issue of the *International Journal of Food Microbiology*. The second phase of this study was on the standardization of the most promising methods for subtyping *L. monocytogenes*, but the results of this phase have not yet been published.

Similar attempts have been made to standardize the subtyping of *C. jejuni* and *C. coli*. Once again, a multitude of subtyping methods have been applied to the subtyping of these *Campylobacter* species. Although routine subtyping of human clinical *Campylobacter* isolates may not be useful for outbreak detection, subtyping has proven to be valuable for understanding the molecular epidemiology of campylobacters in poultry operations, and is enabling the development of strategies to control poultry infection by targeted biosecurity measures (Newell *et al.*, 2000). Campynet, a three-year network project funded by the European Union, was established in 1998 with the aim to standardize and harmonize subtyping methods for *C. jejuni* and *C. coli*. Three subtyping methods were targeted for standardization in the Campynet network: *fla*-PCR RFLP, PFGE, and AFLP (www.svs.dk/campynet).

PulseNet, the national molecular subtyping network for foodborne disease surveillance established in the United States in 1996, is at the forefront of routine application of standardized subtyping for foodborne pathogenic bacteria. At this time, PulseNet laboratories use PFGE as the molecular subtyping method of choice. Public health laboratories in all 50 states participate in PulseNet, as do the laboratories of the US Department of Agriculture's Food Safety and Inspection Service and the laboratories of the US Food and Drug Administration. PulseNet addresses the standardization needs to facilitate molecular subtype data comparisons between laboratories, provides a platform for rapid exchange of DNA 'fingerprints' of bacteria between laboratories and a central repository at the Centers for Disease Control and Prevention in Atlanta, and provides a uniform nomenclature for the unique patterns of each pathogen. Currently, PulseNet has participant-accessible databases of PFGE patterns for *E. coli* O157:H7, *Salmonella* serotypes, *L. monocytogenes*, and *Shigella* spp. Databases for *C. jejuni*, *C. botulinum*, *C. perfringens*, *V. parahaemolyticus* and *V. cholerae* are under development.

Routine PFGE typing of *E. coli* O157:H7, *Salmonella* and *L. monocytogenes* has facilitated early detection of disease clusters (classified as outbreaks after epidemiologic links are found between the cases), and has allowed the linking of outbreaks that were previously investigated as separate outbreaks. This ability to link presumably different outbreaks has, in some instances, made it possible to identify the point source of the combined outbreak. Further, the availability of real-time PFGE data helps epidemiologists separate outbreak-associated cases from sporadic cases occurring in the same geographic locations during the same time period. Because listeriosis is a relatively rare disease, the availability of real-time PFGE patterns for clinical isolates enables epidemiologists to select cases with PFGE patterns different from the outbreak pattern as controls for case-control studies.

## 13.6 Interpreting molecular subtyping data

The interpretation of subtyping data is a critical factor, along with choice of typing method and an understanding of the basic microbiology of the organism, in answering the seemingly simple question: are two isolates the same or different? As the resolution of subtyping techniques has increased over the last 20 years, so has their ability to detect smaller and less frequent variations. The methods currently available are capable of detecting the small genetic changes that can potentially occur during the course of an outbreak or even during *in vitro* culture. Thus, an important component of any typing method is the ability to detect significant epidemiological differences.

In the context of public health, molecular subtyping is usually performed in support of epidemiological studies (either outbreak investigations or surveillance projects). During outbreak investigations, the number of isolates tested is often small, and usually occurs over a limited time and geographical area. The expected range of genetic differences among epidemiologically related isolates is thus relatively small. However, regardless of the method employed, the analysis of the isolates is likely to become more complex as the outbreak becomes more extensive in duration or scope. Under these circumstances, there are more opportunities for the outbreak isolates to exhibit differences such as changes in plasmid content, point mutations affecting restriction sites and insertion and/or deletion events. This is particularly true of pathogens that are disseminated by person-to-person transmission. Similarly, with surveillance studies that include geographically diverse isolates and last a year or more, one would also expect considerably more genetic variability. There is also a chance some epidemiologically unrelated isolates may have similar or indistinguishable genotypes, particularly if there is limited diversity within a species or subtype.

Many of the molecular typing techniques currently available use gel-based electrophoretic separation of DNA fragments of different molecular sizes and therefore allow only indirect chromosomal comparisons. Despite the availability

of a number of commercial software packages for data analysis, a certain level of subjectivity is still involved in interpretation of the data. For most of these subtyping methods there are no standardized criteria for analyzing the fragment patterns. Consequently, different investigators viewing the same subtyping results may come to quite different conclusions as to strain relatedness.

In the absence of epidemiologic evidence, molecular subtyping alone can neither prove nor disprove a connection between isolates. For example, there may be multiple genotypes in a contaminated food, such that the isolation of a single genotype from an incriminated food that differs from that of the patient isolates could lead to the erroneous conclusion that the food was not the source of the patient's infection. Similarly mutational events could result in changes in isolates that could lead to the erroneous conclusion that isolates that actually have the same origin are not related. In addition, co-infection may occur in the host with multiple types being present, which may also complicate interpretation of subtyping data based on single colony picks from cultures (Richardson *et al.*, 2001).

In the context of outbreak investigations and short-term studies, ideally PFGE patterns representing an outbreak strain would be indistinguishable from each other and different from epidemiologically unrelated isolates. When this occurs, identifying outbreak-related isolates on the basis of their PFGE pattern is simple. However, random genetic events such as insertion or deletion of DNA and point mutations frequently alter PFGE patterns in the course of an outbreak. From our experiences in PulseNet, single band differences are often observed, and on occasion two or three band differences in isolates from persons who were almost certainly part of the same outbreak, based on epidemiologic information, is seen. Tenover *et al.* (1995) proposed a set of guidelines for interpreting DNA macrorestriction patterns generated by PFGE. These criteria were intended for use by clinical microbiologists to examine discrete small sets of isolates in the context of short-term outbreaks in which genetic variability is presumed to be limited. In addition, a number of enteric pathogens, including *Vibrio cholerae* (Nandi *et al.*, 1997), *Shigella dysenteriae* and *Shigella flexneri* (Shu *et al.*, 2000), have been demonstrated to undergo genomic re-arrangements that can cause changes within PFGE profiles.

An inherent problem, unrelated to the methodology used, is that of genetic diversity. In general, two isolates that are subtyped as distinctly different can reasonably be assumed to represent different strains. The conclusion that two isolates are indistinguishable depends on both the discriminatory power of the method and the genetic diversity of the particular isolates being examined. For example, *E. coli* O157:H7 is a highly clonal organism, and unpublished data from the PulseNet database suggests that isolates differing by as little as one band may not be part of the same outbreak. Consequently, the general criteria used by Tenover *et al*. (1995), where isolates that differ by a single genetic event (one to three bands) are considered 'probably part of the outbreak', may be somewhat misleading when applied to highly clonal organisms. Similarly, pathogens such as *S.* Enteritidis and *C. jejuni* serotypes HS:19 and HS:11 are

highly homogeneous and therefore most isolates are indistinguishable by most subtyping methods. In these instances, differences are far more informative than similarities, and indistinguishable patterns should not in themselves be considered evidence of a common source. Interpreting subtyping data is particularly difficult when two isolates are typed as 'similar', meaning they differ in only one or two of the characteristics being considered. Any differences should be considered potentially significant, but caution should be used against over-interpretation of differences; PFGE patterns may change due to loss of genes in the course of maintenance or subcultivation of strains (Murase *et al.*, 1999; Wassenaar and Newell, 2000). Likewise, indistinguishable PFGE profiles alone do not prove that the isolates are identical and part of an outbreak.

In general, while the Tenover guidelines provide an excellent theoretical framework, our experiences suggest that criteria should be determined for each organism based on the genetic heterogeneity of that organism and the prevalence of particular subtypes in each community. Unfortunately such data are not yet available for most foodborne pathogens, and interpretation of PFGE data, as with other molecular subtyping data, is most effective when used to supplement, not replace, good epidemiologic investigations. When it is difficult to decide about differences probably caused by genetic changes, the use of more than one subtyping method may be helpful (Barrett, 1997).

There are no standard guidelines for interpreting subtyping data generated by PCR-based typing methods, and the general principles that are used for PFGE cannot readily be applied. In PCR-based methods, not only is variation in banding patterns coupled with specific genetic events, but the resulting profiles can sometimes be a combination of artifactual variation mixed with true polymorphism (Tyler *et al.*, 1997). In particular for RAPD, as reproducibility is a problem, it is especially difficult to establish the criteria for interpreting a change in the intensity of several bands or the size of a single band.

While DNA sequence data provides a direct and unambiguous chromosomal comparison of isolates, care must also be exercised when interpreting the data. Sequencing of both strands of the DNA can help to minimize incorrect base identification. The challenge when using DNA sequence-based subtyping is to consider carefully the region(s) of the genome to be sequenced, so that epidemiologically relevant information can be obtained.

For large-scale longitudinal surveillance studies, interpretation of subtyping data is not such a simple task; however, a set of interpretation guidelines should enable data comparison and integration. The speed at which the genome alters ('molecular clock') influences the data that are generated on the basis of a certain molecular marker (Van Belkum *et al.*, 2001). Therefore careful consideration should be given to the selection of the molecular marker in accordance with the scope of the study. For example, highly variable markers, such as repetitive DNA regions that have a high molecular clock speed, may be highly discriminatory in outbreak investigations yet not suitable for long-term surveillance studies. Although not yet described, these criteria will certainly be different from those used for short-term and outbreak studies. A major limitation

in comparing and interpreting molecular subtyping information from different laboratories to date has been the lack of a standardized approach to both procedures used and adoption of universal nomenclature schemes for the resulting subtyping data generated. For some commonly used methods such as PFGE, this has already been addressed with the PulseNet network.

## 13.7 The future of molecular subtyping

The emergence of automated DNA sequencers and their rapid uptake within microbiology has led to the sequencing of the complete genome of many bacterial pathogens. In the past five years, there have been tremendous advances in both sequencing technology and bioinformatics, and these will have a significant effect on the way epidemiological typing will be carried out in the future. The discovery of epidemiologically relevant sequences for subtyping by direct sequencing and technological developments in sequence detection and data analysis will have a profound effect on the development of future subtyping methods (Goering, 2002).

Determining the complete genome sequence of a number of foodborne pathogens has already been achieved, and many more genomes will be added to the list in the near future. The sequencing of additional pathogens will facilitate the identification of loci with sequence diversity suitable for subtyping. Current DNA sequencing-based subtyping methods such as MLST are considered too unwieldy and expensive for routine public health laboratories, and this has limited their use. MLST in its current form requires the direct sequencing of more than 2000 nucleotides of sequence per isolate, which is labor intensive, expensive and time-consuming. Widespread adoption of sequence-based subtyping methods will require a reduction in the number of steps for data acquisition and analysis, potentially through automation of some, if not all, parts of the process. The generation and analysis of sequence data itself will become the limiting factor in epidemiological analysis, stimulating the identification of alternative methods for the detection and analysis of the differences between strains. In MLST, multiple single nucleotide polymorphisms (SNPs) are detected between allele sequences via direct sequencing. When a suitably large dataset for a particular pathogen has been generated and the SNPs identified have been epidemiologically validated, alternative methods for their detection can be developed. There is a rapidly growing number of novel methods available for the detection of SNPs many of which come from the growing field of pharmacogenomics (Shi, 2001). Pyrosequencing and oligonucleotide microarrays are both promising methods for the rapid detection of epidemiologically relevant SNPs. Pyrosequencing detects the incorporation of specific nucleotides via the release of pyrophosphate, which leads to light production by luciferase (Nordstrom *et al.*, 2000). This method currently has the potential to detect 500 SNPs in one hour following PCR amplification. Assays could be designed to detect the epidemiologically validated SNPs within the MLST allele sequences following PCR.

DNA microarrays facilitate the simultaneous detection of many SNPs based on hybridization genotyping. High-density DNA microarrays can be constructed by attaching thousands of single oligonucleotides to a solid silicon surface in an ordered array. Fluorescently labeled nucleotides are generated by PCR from DNA from the organism being investigated and then hybridized to the array. The hybridization signals are quantified using high-resolution fluorescent scanning and are analyzed by computer software. The method is efficient for the analysis of large numbers of SNPs and it is likely that organism-specific 'chips' will be designed containing arrays specific for the typing of particular organisms. The incorporation of species identification sequences and virulence-specific sequences can also be envisaged combining typing, identification and virulence characterization on a single array. The development of these new subtyping approaches will depend on their validation using panels of previously characterized isolates whose epidemiological relationships have been unambiguously elucidated. This will ensure that the results from these new methods reflect epidemiological relationships previously established, and that they provide subtyping data relevant to future epidemiological investigations. In the future, subtyping of foodborne pathogens by DNA sequence-based methodologies will provide more rapid, sensitive and informative results. This will lead to improvements in the surveillance and recognition of foodborne diseases, and will have a significant role in the improvement of food and water safety.

## 13.8 Sources of further information and advice

*Collaborative typing networks*
Enter-Net. www.phls.co.uk/International/Enter-Net/enter-net.htm
Campynet. www.svs.dk/campynet
PulseNet. www.cdc.gov/pulsenet
Multilocus sequence typing (MLST). www.mlst.net/new/index.htm
Genetic epidemiology network for Europe (GENE). www.ewi.med.uu.nl/gene

*Commercial kits and applications.*
RiboPrinter®: www.qualicon.com/rp.html
Bacterial Barcodes: www.bacbarcodes.com
Bio-rad Laboratories (PFGE reagent kits): www.biorad.com
Ready-To-Go RAPD Analysis Beads: www.amershambiosciences.com

*Software for analysis of subtyping data*
GelCompar and BioNumerics: (www.applied-maths.com), E-mail: info@applied-maths.com. Gelcompar and BioNumerics software are modular packages for advanced fingerprint analysis.

Gene Profiler: (www.scanalytics.com), E-mail: sales@scanalytics.com. Gene

Profiler is a Windows software package that is used primarily for genotyping and DNA fingerprint analysis.

Dendron: (www.geocities.com/solltech), DENDRON corrects, processes and analyzes all forms of gel images in order to compare their banding patterns.

## 13.9 References


BARRETT T J (1997). Molecular fingerprinting of foodborne bacteria: an introduction to methods, uses, and problems. In: Tortorello, M L and Gendel, S M (eds), Food Microbiological Analysis New Technologies, New York, Marcel Dekker, pp. 249–64.

ENRIGHT M and SPRATT B (1999). 'Multilocus sequence typing', *Trend Microbiol,* **7**(12) 482–7.

FARLOW J, SMITH K L, WONG J, ABRAMS M, LYTLE M and KEIM P (2001). *Francisella tularensis* strain typing using multiple-locus, variable-number tandem repeat analysis. *J Clin Microbiol*, **39**(9) 3186–92.

GOERING R (2002), The Influence of Genomics on the Molecular Epidemiology of Nosocomial Pathogens. In: Shaw K (ed.) *Pathogen Genomics: Impact on Human Health*. Totowa, Humana Press.

GRAVES L, SWAMINATHAN B and HUNTER S (1999). Subtyping *Listeria monocytogenes*. In Elliot T, Ryser E T and Marth E H (eds), *Listeria, Listeriosis and Food Safety*. New York, Marcel Dekker.

HOLMBERG K and FEROZE F (1996). 'Evaluation of an optimized system for random amplified polymorphic DNA (RAPD)-analysis for genotypic mapping of *Candida albicans* strains', *J Clin Lab Anal*, **10**(2) 59–69.

HULTON C S, HIGGINS C F and SHARP P M (1991). 'ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria', *Mol Microbiol,* **5**(4) 825–34.

HUNTER S and SWAMINATHAN B (1998). Ribotyping as a tool for molecular epidemiology. In: Spector EA (ed.) *Rapid detection of Infectious agents*. New York, Plenum Press.

KEIM P, PRICE L B, KLEVYTSKA A M, SMITH K L, SCHUPP J M, OKINAKA R, JACKSON P J and HUGH-JONES M E (2000). 'Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis', J Bacteriol,* **182**(10) 2928–36.

KLEVYTSKA A M, PRICE L B, SCHUPP J M, WORSHAM P L, WONG J and KEIM P (2001). 'Identification and characterization of variable-number tandem repeats in the *Yersinia pestis* genome', *J Clin Microbiol,* **39**(9) 3179–85.

MACHADO J, GRIMONT F and GRIMONT P A (2000) 'Identification of *Escherichia coli* flagellar types by restriction of the amplified fliC gene', *Res Microbiol,* **151**(7) 535–46.

MAYER L (1988). 'Use of plasmid profiles in epidemiologic surveillance of disease outbreaks and in tracing the transmission of antibiotic resistance', *Clin Microbiol Rev*, **1**(2) 228–43.

MURASE T, YAMAI S and WATANABE H (1999). 'Changes in pulsed-field gel electrophoresis patterns in clinical isolates of enterohemorrhagic *Escherichia coli* O157:H7 associated with loss of Shiga toxin genes', *Cur Microbiol* **38**(1) 48–50.

NANDI S, KHETAWAT G, SENGUPTA S, MAJUMDER R, KAR S, BHADRA RK, ROYCHOUDHURY S and DAS J (1997). 'Rearrangements in the genomes of *Vibrio cholerae* strains belonging to different serovars and biovars', *Int J Syst Bacteriol*, **47**(3) 858–62.

NEWELL D, FROST J, DUIM B, WAGENAAR J, MADDEN R, VAN DER PLAS J and ON S (2000). On S. New developments in the subtyping of *Campylobacter* species. In: Nachamkin I and Blaser M, (eds), *Campylobacter*, Washington DC, American Society of Microbiology).

NORDSTROM T, RONAGHI M, FORSBERG L, DE FAIRE U, MORGENSTERN R and NYREN P (2000). 'Direct analysis of single-nucleotide polymorphism on double-stranded DNA by pyrosequencing', *Biotechnol Appl Biochem,* **31**(Pt 2) 107–12.

RICHARDSON JF, FROST JA, KRAMER JM, THWAITES RT, BOLTON FJ, WAREING DR and GORDON J A (2001). 'Coinfection with *Campylobacter* species: an epidemiological problem?', *J Appl Microbiol*, **91**(2) 206–11.

SAVELKOUL P, AARTS H, DE HAAS J, DIJKSHOORN L, DUIM B, OTSEN M, RADEMAKER J, SCHOULS L and LENSTRA J (1999). 'Amplified-fragment length polymorphism analysis: the state of an art', *J Clin Microbiol*, **37**(10) 3083–91.

SHI MM (2001). 'Enabling large-scale pharmacogenetic studies by high-throughput mutation detection and genotyping technologies', *Clin Chem*, **47**(2) 164–72.

SHOPSIN B, GOMEZ M, MONTGOMERY SO, SMITH DH, WADDINGTON M, DODGE DE, BOST DA, RIEHMAN M, NAIDICH S and KREISWIRTH BN (1999). Evaluation of protein A gene polymorphic region DNA sequencing for typing of *Staphylococcus aureus* strains, *J Clin Microbiol* **37**(11): 3556–63.

SHU S, SETIANINGRUM E, ZHAO L, LI Z, XU H, KAWAMURA Y and EZAKI T (2000). 'I-CeuI fragment analysis of the *Shigella* species: evidence for large-scale chromosome rearrangement in *S. dysenteriae* and *S. flexneri'*, *FEMS Microbiol Lett*, **182**(1) 93–8.

SWAMINATHAN B, BARRETT T, HUNTER S, TAUXE R and THE CDC PULSENET TASKFORCE (2001). 'PulseNet: The molecular subtyping network for foodborne bacterial disease surveillance, United States', *Emerg Infect Dis*, **7**(3) 382–9.

TENOVER FC, ARBEIT RD, GOERING RV, MICKELSEN PA, MURRAY BE, PERSING DH and SWAMINATHAN B (1995). 'Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing', *J Clin Microbiol* **33**(9) 2233–9.

TYLER W, WANG G, TYLER S and JOHNSON W (1997). 'Factors affecting reliability and reproducibility of amplification-based DNA fingerprinting of representative bacterial pathogens', *J Clin Microbiol*, **35**(2) 339–46.

VAN BELKUM A, STRUELENS M, DE VISSER A, VERBRUGH H and TEBAYRENC M (2001). 'Role of genome typing in taxonomy, evolutionary genetics and microbial epidemiology', *Clin Microbiol Rev*, **14**(3) 547–60.

WASSENAAR T and NEWELL D (2000). 'Genotyping of *Campylobacter* spp', *Appl Environ Microbiol*, **66**(1) 1–9.